Scalar meson fields in a conformally flat space

Dipankar Ray

Department of Physics, New York University, New York, New York 10003
(Received 12 November 1976)

Scalar meson fields are studied for the conformally flat metric $e^{+\psi}(dx^{1^2}+dx^{2^2}+dx^{3^2}-dx^{4^2})$. Complete sets of solutions of the field equations have been obtained for massless mesons, and the field equations have been reduced to a single ordinary differential equation in ψ for massive mesons. In either case ψ is found to be a function of

$$x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2} \text{ or } c_1x^1+c_2x^2+c_3x^3+c_4x^4$$

(c_1, c_2, c_3, c_4 being constants), and ϕ the meson field is found to be constant or a function of ψ.

I. INTRODUCTION

Massive and massless scalar meson fields in general relativity have been studied by several authors. In particular Penney[1] has given an exact solution for massless scalar meson fields with a conformally flat metric. The present note is an attempt to obtain a complete set of exact solutions for both massive and massless scalar mesons in a conformally flat metric.

II. FIELD EQUATIONS

The field equations for a scalar meson field are

$$R_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}R=-[\phi_{,\mu}\phi_{,\nu}-\tfrac{1}{2}g_{\mu\nu}(\phi_{,\alpha}\phi^{,\alpha}-m^2\phi^2)], \quad (1)$$

where we have set $K=1$, and where ϕ is the meson field and m the meson mass. (1) can be rewritten as

$$R_{\mu\nu}=-\phi_{,\mu}\phi_{,\nu}+\tfrac{1}{2}g_{\mu\nu}m^2\phi^2. \quad (2)$$

The metric is

$$ds^2=\exp(-\psi)(dx^{1^2}+dx^{2^2}+dx^{3^2}-dx^{4^2}),$$

i.e.,

$$g_{\mu\nu}=\exp(+\psi)\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&-1\end{pmatrix}\equiv\exp(+\psi)\eta_{\mu\nu} \text{ (say).} \quad (3)$$

We shall first prove that for solutions of (2) of the form (3), ψ must be a function of $(x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2})$ or $(c_1x^1+c_2x^2+c_3x^3+c_4x^4)$, ($c_1, c_2, c_3, c_4$ being constants) and ϕ is either a constant or a function of ψ. For (3), $R_{\mu\nu}$ is given by

$$R_{\mu\nu}=\psi_{,\mu,\nu}-\tfrac{1}{2}\psi_{,\mu}\psi_{,\nu}+\tfrac{1}{2}\eta_{\mu\nu}\chi, \quad (4)$$

where

$$\chi=\psi_{,1,1}+\psi_{,2,2}+\psi_{,3,3}-\psi_{,4,4}+\psi_{,1}^2+\psi_{,2}^2+\psi_{,3}^2-\psi_{,4}^2.$$

A. Case I. ϕ is constant

In this case we see from Eqs. (2) that the space is an Einstein space (i.e., a space for which $R_{\mu\nu}=\text{constant}\times g_{\mu\nu}$), and we know[2] that only a conformally flat Einstein space is a space of constant curvature and the metric takes the form

$$g_{\mu\nu}=\{1+(K_0/4)(x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2})\}^{-1}\times(dx^{1^2}+dx^{2^2}+dx^{3^2}-dx^{4^2}).$$

$R_{\mu\nu}$ becomes

$$R_{\mu\nu}=-3K_0g_{\mu\nu},$$

where K_0 is the constant curvature. Thus (2) and hence (1) is satisfied if

$$K_0<0 \text{ and } \phi^2=-6K_0/m^2$$

Thus ψ is function of $x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2}$ and ϕ is constant.

B. Case II. ϕ is not constant

Comparing (2) and (4) and noting (3)

$$-2\exp(\psi/2)(\exp(-\psi/2))_{,\mu,\nu}=-\phi_{,\mu}\phi_{,\nu} \text{ for } \mu\neq\nu.$$

In particular

$$\{(\exp(-\psi/2))_{,4}\}_{,i}=(\phi_{,4}\exp(-\psi/2)/2)\phi_{,i}, \quad i=1,2,3. \quad (5)$$

From (5) we note that the derivatives of $(\exp(-\psi/2))_{,4}$ with respect to x^1, x^2, and x^3 are proportional to the derivatives ϕ with respect to x^1, x^2, and x^3. Thus, if x^4 is treated as a constant, $(\exp(-\psi/2))_{,4}$ and ϕ are functionally dependent[3] and since ϕ is not constant for the case under consideration we can say that $(\exp(-\psi/2))_{,4}$ is a function of ϕ if x^4 is treated as a constant; in other words

$$(\exp(-\psi/2))_{,4}=\delta(x^4,\phi), \text{ where } \delta \text{ is some function.}$$

Similarly,

$$(\exp(-\psi/2))_{,1}=\alpha(x^1,\phi),$$
$$(\exp(-\psi/2))_{,2}=\beta(\phi,x^2),$$
$$(\exp(-\psi/2))_{,3}=\gamma(\phi,x^3), \quad (6)$$

where α, β, γ are some functions. From (5) and (6)

$$\alpha_\phi=\tfrac{1}{2}\exp(-\psi/2)\phi_{,1}, \quad \beta_\phi=\tfrac{1}{2}\exp(-\psi/2)\phi_{,2},$$
$$\gamma_\phi=\tfrac{1}{2}\exp(-\psi/2)\phi_{,3}, \quad \delta_\phi=\tfrac{1}{2}\exp(-\psi/2)\phi_{,4}, \quad (7)$$

where $\alpha_\phi=\partial\alpha/\partial\phi|_{x^1 \text{ as constant}}$ and so on, and $\alpha_{x^1}=\partial\alpha/\partial x^1|_{\phi \text{ as constant}}$, etc. From (2)–(4)

$$-2\exp(\psi/2)(\exp(-\psi/2))_{,i,i}+\tfrac{1}{2}\chi=-\phi_{,i}^2+(\exp(+\psi)/2)m^2\phi^2$$
$$-2\exp(\psi/2)(\exp(-\psi/2))_{,4,4}-\tfrac{1}{2}\chi=-\phi_{,4}^2-(\exp(+\psi)/2)m^2\phi^2. \quad (8)$$

Using (6), (7), and (8)

$$\alpha_{x^1}=\beta_{x^2}=\gamma_{x^3}=-\delta_{x^4}=-(\exp(-\psi/2)/4)(m^2\phi^2\exp(-\psi)-\chi). \quad (9)$$

However, $\alpha_{,1}$ is a function of ϕ and x^1, $\beta_{,2}$ is a function of ϕ and x^2, and so on. Thus (9) is possible only if $\exp(-\psi/2)(m^2\phi^2\exp(-\psi)-\chi)$ is a function of ϕ alone,[4] or

$$\alpha = x^1 v + p, \quad \beta = x^2 v + q, \quad \gamma = x^3 v + r, \quad \delta = -x^4 v + s, \tag{10}$$

where $v = -\frac{1}{4}\exp(-\psi/2)(m^2\phi^2\exp(-\psi)-\chi)$; v, p, q, r, s are functions of ϕ only. From (6), (7), (10) and using relations as $\phi_{,1,2} = \phi_{,2,1}$,

$$\frac{2(x^1 v_{\phi\phi}+\phi_{\phi\phi})+(x^1 v+p)}{x^1 v_\phi + p_\phi} = \frac{2(x^2 v_{\phi\phi}+q_{\phi\phi})+(x^2 v+q)}{x^2 v_\phi + q_\phi}$$
$$= \frac{2(x^3 v_{\phi\phi}+r_{\phi\phi})+(x^3 v+r)}{x^3 v_\phi + r_\phi} = \frac{2(-x^4 v_{\phi\phi}+s_{\phi\phi})+(-x^4 v+s)}{-x^4 v_\phi + s_\phi},$$
$$v_\phi \equiv dv/d\phi \text{ and so on,} \tag{11}$$

where in view of (7) and (10) and the fact that ϕ is not constant, all the denominators of (11) cannot be zero. If any of the denominators of (11) is zero, (11) is to be interpreted to mean that the corresponding numerator is zero.

In (11), we have four different expressions that are equal to each other, but one of them is a function of ϕ and x^1, another, a function of ϕ and x^2, and so on. This is possible only if each of them is equal to a function of ϕ only, say $\xi(\phi)$,[5]

$$\frac{2v_{\phi\phi}+v}{v_\phi} = \frac{2p_{\phi\phi}+p}{p_\phi} = \frac{2q_{\phi\phi}+q}{q_\phi} = \frac{r_{\phi\phi}+r}{r_\phi} = \frac{s_{\phi\phi}+s}{s_\phi} = \xi(\phi), \tag{12}$$

where as before, if any of the denominators vanishes, the corresponding numerator vanishes.

Thus v, p, q, r are solutions of an ordinary linear homogeneous second order differential equation and hence at most two of them can be independent. Thus, let

$$v = A\eta + B\xi, \quad p = A_1\eta + B_1\xi, \quad q = A_2\eta + B_2\xi,$$
$$r = A_3\eta + B_3\xi, \quad s = A_4\eta + B_4\xi, \tag{13}$$

where η and ξ are solutions of the same differential equation and A, B, A_1, B_1, A_2, B_2, A_3, A_4, B_4 are constants. By (10) and (13), (6) and (7) reduce to

$$(\exp(-\psi/2))_{,\mu} = \eta P_{,\mu} + \xi Q_{,\mu},$$
$$\tfrac{1}{2}\exp(-\psi/2)\,\phi_{,\mu} = \eta_\phi P_{,\mu} + \xi Q_{,\mu}, \tag{14}$$

where

$$P = (A/2)(x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2}) + (A_1x^1 + A_2x^2 + A_3x^3 + A_4x^4),$$
$$Q = (B/2)(x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2}) + (B_1x^1 + B_2x^2 + B_3x^3 + B_4x^2),$$

which readily gives

$$\tfrac{1}{2}\exp(-\psi/2)(\epsilon+\psi)_{,\mu} = (\xi_\phi\eta/\eta_\phi - \xi)\,Q_{,\mu}$$
where $\epsilon = \int(\eta\,d\phi/\eta\phi)$. (15)

From (15) we have either of the following:

$$\xi_\phi\eta/\eta_\phi - \xi = 0, \tag{16a}$$

or

$$\epsilon + \psi = f(Q), \quad \text{where } f \text{ is some function.} \tag{16b}$$

If (16a) is true, $\xi = K\eta$, K is a constant, and we get, from (14),

$$(\exp(-\psi/2))_{,\mu} = \eta(P+KQ)_{,\mu},$$
$$\tfrac{1}{2}\exp(-\psi/2)\,\phi_{,\mu} = \eta_\phi(P+KQ)_{,\mu},$$

i.e., ψ and ϕ are functions of $P+KQ$. If, on the other hand, (16b) is true, we get from (15) and (16b),

$$2\exp(-\epsilon/2)\left(\frac{\xi_\phi\eta}{\eta_\phi} - \xi\right) = \frac{f'(Q)}{\exp(f(Q)/2)}. \tag{17}$$

The left-hand side of (17) is a function of ϕ and the right-hand side of (17) is a function of Q. Thus ϕ is a function of Q. This implies that ϵ is a function of ϕ, and we get from (16b) that ψ is a function of Q. Thus in case II, for either (16a) or (16b) we have

$$\psi = \psi(u), \quad \phi = \phi(u), \tag{18}$$

where

$$u = C(x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2}) + (C_1x^1 + C_2x^2 + C_3x^3 + C_4x^4), \tag{19}$$

C, C_1, C_2, C_3, C_4 being constants.

We note that case I of ϕ is a constant and is also a special case of (18) and (19). Also, for $C \neq 0$, it is obvious that, without loss of generality, one can set

$$u = x^{1^2}+x^{2^2}+x^{3^2}-x^{4^2}. \tag{20}$$

For $C = 0$

$$u = C_1x^1 + C_2x^2 + C_3x^3 + C_4x^4. \tag{21}$$

III. CONCLUSION

Thus the solutions of (2) for the conformally flat metric (3) are of the form (18), where u is given by (20) or (21). The field equations then reduce to

Case A—u is given by (20):

$$\left.\begin{aligned}&\psi_{uu} - \tfrac{1}{2}\psi_u^2 = -\phi_u^2 && \text{(a)}\\ &u(\psi_{uu}+\psi_u^2) + 3\psi_u = (m^2\exp(-\psi)/4)\,\phi^2 && \text{(b)}\end{aligned}\right\} \quad \psi_u \equiv \frac{d\psi}{du} \text{ etc.}$$

Case B—u is given by (21), $C_1^2 + C_2^2 + C_3^2 - C_4^2 \neq 0$:

$$\psi_{uu} - \tfrac{1}{2}\psi_u^2 = -\phi_u^2, \qquad \text{(a)}$$
$$(C_1^2 + C_2^2 + C_3^2 - C_4^2)(\psi_{uu} + \psi_u^2) = \exp(+\psi)\,m^2\phi^2. \qquad \text{(b)}$$

We note also that in either case A or case B, one can get from (a) and (b) an ordinary differential equation for ψ and equation (a) is integrable for ϕ. Complete integration is easily possible for $m = 0$.

Case C—u is given by (21), $C_1^2 + C_2^2 + C_3^2 - C_4^2 = 0$:

$$\psi_{uu} - \tfrac{1}{2}\psi_u^2 = -\phi_u^2, \quad \text{(a)}$$
$$m = 0, \quad \text{(b)}$$

i.e., this case is only possible for massless mesons. (a) is readily integrable for ϕ. ψ is an arbitrary function of u except for $\psi_{uu} - \frac{1}{2}\psi_u^2 < 0$.

[1]R.V. Penney, Phys. Rev. D 14, 910 (1975).
[2]L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1949), p. 85.
[3]It is a well-known result of the theory of functions that if two continuously differentiable functions have their partial derivatives proportional to each other then the two functions are functionally dependent.
[4]This can be seen as follows. Since ϕ is not a constant, at least one of $\phi_{,1}$, $\phi_{,2}$, $\phi_{,3}$, $\phi_{,4}$ is not zero. Let $\phi_{,1} \neq 0$; we have

$$\beta_{x^2}(x^2, \phi) = \gamma_{x^3}(x^3, \phi).$$

Consider derivatives with respect to x^1 and x^3,

$$[\beta_{x^2}]_{,\phi}\phi_{,1} = [\gamma_{x^3}]_{,\phi}\phi_{,1},$$

$$[\beta_{x^2}]_{,\phi}\phi_{,3} = [\gamma_{x^3}]_{,\phi}\phi_{,3} + [\gamma_{x^3}]_{,x^3},$$

which readily gives $[\gamma_{x^3}]_{,x^3} = 0$, i.e., γ_{x^3} is a function of ϕ. Hence so are $\alpha_{x^1}, \beta_{x^2}, \delta_{x^4}$.
[5]This argument may seem to break down if several of the denominators of equations (11) vanish. However, it is easy to see that if one or more of the denominators of (11) vanish, then we must have $v = 0$, as can be seen as follows. Let

$$x^1 v_\phi + p_\phi = 0. \qquad \text{(i)}$$

By (7) and (10),

$$\phi_{,1} = 0. \qquad \text{(ii)}$$

Also from (11) and (i)

$$2(x^1 v_{\phi\phi} + p_{\phi\phi}) + (x^1 v + p) = 0. \qquad \text{(iii)}$$

Noting (ii) and the fact that v and p are functions of ϕ, we see that (i) and (iii) can hold only if $v = 0$. For $v = 0$, all the expressions of (11) are functions of ϕ only; hence (12) is automatically true.

Invariants and wavefunctions for some time-dependent harmonic oscillator-type Hamiltonians

P. G. L. Leach

Department of Mathematics, La Trobe University, Bundoora, 3083, Australia
(Received 1 November 1976)

Recently the author has shown that the Hamiltonian, $H=(1/2)\boldsymbol{\omega}^T A(t)\boldsymbol{\omega}+B(t)^T\boldsymbol{\omega}+C(t)$, in which $A(t)$ is a positive definite symmetric matrix and $\omega^\mu=q_i$, $\mu=1,n$, $i=1,n$, $\omega^\mu=p_i$, $\mu=n+1,2n$, $i=1,n$, may be transformed to the time-independent Hamiltonian, $\bar{H}=(1/2)\bar{\boldsymbol{\omega}}^T\bar{\boldsymbol{\omega}}$, by a time-dependent linear canonical transformation, $\bar{\boldsymbol{\omega}}=S\boldsymbol{\omega}+\mathbf{r}$. $\bar{H}$ is an exact invariant of the motion described by H. A matrix invariant may also be constructed which provides a basis for the generators of the dynamical symmetry group SU(n) which may always be associated with H, usually as a noninvariance group. In this paper we examine, by way of example, an oscillator with source undergoing translation, the two-dimensional anisotropic oscillator, general one- and two-dimensional oscillators with Hamiltonians of homogeneous quadratic form and obtain explicit invariants and Schrödinger wavefunctions with the aid of the linear canonical transformations.

1. INTRODUCTION

The use of time-dependent linear canonical transformations for the construction of invariants for oscillator-type Hamiltonians has been developed recently by Leach. He showed[1] that the Hamiltonian

$$H=\tfrac{1}{2}p^2+\tfrac{1}{2}\omega^2(t)q^2, \tag{1.1}$$

in which (as hereinafter) p and q are canonically conjugate variables, could be transformed by means of the time-dependent linear canonical transformation

$$\begin{bmatrix}Q\\P\end{bmatrix}=\left[\begin{array}{c|c}C_1\cos W_1+C_2\cos W_2 & -C_1\sin W_1-C_1\sin W_2\\ \hline C_1\sin W_1-C_2\sin W_2 & C_1\cos W_1-C_2\cos W_2\end{array}\right]\times\begin{bmatrix}\rho^{-1} & 0\\ -\dot{\rho} & \rho\end{bmatrix}\begin{bmatrix}q\\p\end{bmatrix}, \tag{1.2}$$

in which $\rho(t)$ is solution of

$$\ddot{\rho}+\omega^2(t)\rho=\rho^{-3} \tag{1.3}$$

and

$$C_1^2-C_2^2=1,\qquad W_1=\int_{t_0}^{t}(\rho^{-2}-1)\,dt',\quad W_2=\int_{t_0}^{t}(\rho^{-2}+1)\,dt', \tag{1.4}$$

to the time-independent Hamiltonian

$$\bar{H}=\tfrac{1}{2}P^2+\tfrac{1}{2}Q^2. \tag{1.5}$$

$\bar{H}$ provides an invariant for the motion described by H. In terms of the original coordinates it is

$$I=\tfrac{1}{2}[\rho^{-2}q^2(C_1^2+C_2^2+2C_1C_2\cos 2W)+(\rho p-\dot{\rho}q)^2\times(C_1^2+C_2^2-2C_1C_2\cos 2W)-4\rho^{-1}q(\rho p-\dot{\rho}q)C_1C_2\sin 2W], \tag{1.6}$$

where

$$W=\int_{t_0}^{t}\rho^{-2}\,dt'. \tag{1.7}$$

In quantum mechanics any products pq are symmetrized to $\frac{1}{2}(pq+qp)$. The invariant (1.6) is a generalization of the Lewis invariant[2,3] (whose work was partly anticipated by Semour[4] with others[5]) for the same Hamiltonian. The invariant which Lewis obtained using a different method is

$$I=\tfrac{1}{2}\rho^{-2}q^2+(\rho p-\dot{\rho}q)^2. \tag{1.8}$$

Subsequently Leach[6] showed that the Hamiltonian

$$H=\tfrac{1}{2}\boldsymbol{\omega}^T A(t)\boldsymbol{\omega}+B(t)^T\boldsymbol{\omega}+C(t) \tag{1.9}$$

in which $A(t)$ is a positive definite symmetric matrix of order $2n$ and the canonical variables (q_i, p_i) are written as

$$q_i=\omega^\mu,\quad i=1,n,\quad \mu=1,n,\qquad p_i=\omega^\mu,\quad i=1,n,\quad \mu=n+1,2n, \tag{1.10}$$

may always be transformed to

$$\bar{H}=\tfrac{1}{2}\bar{\boldsymbol{\omega}}^T\bar{\boldsymbol{\omega}} \tag{1.11}$$

by means of a linear canonical transformation. The transformation is

$$\bar{\boldsymbol{\omega}}=S(t)\,\boldsymbol{\omega}+\mathbf{r}(t) \tag{1.12}$$

in which $S(t)$ is a member of the group of real symplectic matrices and $\mathbf{r}(t)$ is a real vector such that

$$\dot{S}=\epsilon S-S\epsilon A,\quad \dot{\mathbf{r}}=\epsilon\mathbf{r}-S\epsilon\mathbf{B}. \tag{1.13}$$

The matrix ϵ is the symplectic matrix

$$\epsilon=[\epsilon^{\mu\nu}]=\begin{bmatrix}0_n & I_n\\ -I_n & 0_n\end{bmatrix}. \tag{1.14}$$

The time-dependence in (1.9) is arbitrary provided that $A(t)$, $\mathbf{B}(t)$, and $C(t)$ are continuous functions of time over the interval of interest and $A(t)$ does not vary in rank (which implies that it remains positive definite).

The Hamiltonian (1.9) has an invariant

$$I=\tfrac{1}{2}(S\boldsymbol{\omega}+\mathbf{r})^T(S\boldsymbol{\omega}+\mathbf{r}). \tag{1.15}$$

Various interpretations of the nature of the invariant have been offered. Eliezer and Gray[7] suggested that (1.8) was the angular momentum associated with a two-dimensional auxiliary motion for which (1.3) was the radial equation of motion. Günther and Leach,[8] in their discussion of the three-dimensional motion with Hamiltonian

$$H=\tfrac{1}{2}\sum_{i=1}^{3}[p_i^2+\omega^2(t)q_i^2] \tag{1.16}$$

also provided an angular momentum interpretation for the corresponding invariant

$$I=\tfrac{1}{2}\sum_{i=1}^{3}[\rho^{-2}q_i^2+(\rho p_i-\dot{\rho}q_i)^2], \tag{1.17}$$

but, as an alternative, proposed that the invariant was Hamiltonian in nature. In view of Leach's more recent work on canonical transformations relating H (1.9) and $\bar{H}$ (1.11) [from which I (1.15) is obtained], the invariant appears to be the Hamiltonian of the archtypal oscillator represented by $\bar{H}$.

The invariant I (1.15) has an associated symmetric invariant matrix

$$[I^{\mu\nu}]=\tfrac{1}{2}\bar{\omega}\,\bar{\omega}^{T}+\tfrac{1}{2}(\epsilon\,\bar{\omega})(\epsilon\,\bar{\omega})^{T}$$
$$=\tfrac{1}{2}(S\omega+\mathbf{r})(S\omega+\mathbf{r})^{T}+\tfrac{1}{2}\epsilon(S\omega+\mathbf{r})(S\omega+\mathbf{r})^{T}\epsilon^{T}. \tag{1.18}$$

The Poisson bracket of each element $I^{\mu\nu}$ with I is zero. In the transformed coordinates, $\bar{\omega}^{T}=(\mathbf{Q}^{T},\mathbf{P}^{T})$, the matrix $[I^{\mu\nu}]$ has elements of the form $\frac{1}{2}(Q_iQ_j+P_iP_j)$ and $\frac{1}{2}(Q_iP_j-P_iQ_j)$. The former correspond to the elements of Fradkin's tensor[9,10] and the latter are essentially elements of the angular momentum tensor. In all there are n^2-1 elements of $[I^{\mu\nu}]$ linearly independent of I. They provide a sufficient basis for the generators of the symmetry group SU(n) which is accordingly the symmetry group of I. Since the Poisson bracket of H (1.9) and I (1.15) is nonzero in most cases, SU(n) is a noninvariance dynamical symmetry group of that Hamiltonian. Hamiltonians of the form of (1.9) are said to be characterized by the symmetry group SU(n).

Lewis and Riesenfeld[11] showed how the existence of an invariant made possible (in principle) the solution of the Schrödinger equation for that motion. Since all oscillator systems with Hamiltonians of the type (1.9) are characterized by the same symmetry group, we expect that their wavefunctions should be similar. Wolf[12] has shown that linear canonical transformations may be used to solve Schrödinger equations by introducing integral transforms which reduce to geometric transforms in special cases. It is not clear as to the nature, if any, of the time-dependence in his work.

In this paper we shall examine some simple time-dependent oscillator systems. For them we obtain the appropriate invariants and solutions of the Schrödinger equation. Thereby we illustrate a wider use for linear canonical transformations than was previously indicated.[13,14]

2. OSCILLATOR SOURCE UNDERGOING TRANSLATION[15]

The source of a time-independent one-dimensional oscillator potential is moving with respect to an inertial frame such that at time t the source is displaced $s(t)$. A particle of mass m is moving under the influence of this potential and, relative to the inertial frame, is located at q. The Hamiltonian of the motion is

$$H=\frac{1}{2m}p^2+\tfrac{1}{2}m\omega^2(q-s)^2, \tag{2.1}$$

where p is the momentum conjugate to q.

As ω is time-independent, a suitable form for the transformed Hamiltonian is

$$\bar{H}=\frac{1}{2m}p^2+\tfrac{1}{2}m\omega^2Q^2. \tag{2.2}$$

Clearly the transformation from (q,p) to (Q,P) will be translational only and we write

$$\begin{bmatrix}Q\\P\end{bmatrix}=\begin{bmatrix}q\\p\end{bmatrix}+\begin{bmatrix}r_1\\r_2\end{bmatrix}, \tag{2.3}$$

where, adapting (1.13) (cf. Leach[6]) to the different form of $\bar{H}$,

$$\begin{bmatrix}\dot{r}_1\\\dot{r}_2\end{bmatrix}=\begin{bmatrix}r_2/m & 0\\-m\omega^2r_1 & -m\omega^2s\end{bmatrix} \tag{2.4}$$

$$\Longleftrightarrow\begin{cases}\ddot{r}_1+\omega_2r_1=-\omega^2s\\ r_2=m\dot{r}_1\end{cases} \tag{2.5}$$

which have the solutions

$$r_1=A\sin\omega t+B\cos\omega t+\omega\int_{t_0}^{t}s\sin\omega(t-t')\,dt'$$
$$r_2=m\omega[A\cos\omega t-B\sin\omega t+\omega\int_{t_0}^{t}s\cos\omega(t-t')\,dt']. \tag{2.6}$$

The invariant of the motion is given by $\bar{H}$ and is

$$I=\frac{1}{2m}(p+r_2)^2+\tfrac{1}{2}m\omega^2(q+r_1)^2. \tag{2.7}$$

For a specified problem it is convenient to use the particular integrals contained in (2.6) for r_1 and r_2. Note that (2.6) remains valid if s is of period $2\pi/\omega$.

If the translation is due to the Galilean transformation

$$s=Ut, \tag{2.8}$$

the particular integrals of (2.6) are

$$r_1=-Ut,\quad r_2=-mU, \tag{2.9}$$

giving the invariant

$$I=\frac{1}{2m}(p-mU)^2+\tfrac{1}{2}m\omega^2(q-Ut)^2. \tag{2.10}$$

The invariant is simply the Hamiltonian of the motion obtained by taking the source as origin and reflects the invariance of form of Newton's equations of motion if the frame of reference is moving with constant velocity relative to an inertial frame (cf. Eliezer and Leach[16]).

If the source is accelerating uniformly so that

$$s=\tfrac{1}{2}at^2, \tag{2.11}$$

$$r_1=-\tfrac{1}{2}at^2+a/\omega^2,\quad r_2=-mat, \tag{2.12}$$

the invariant is

$$I=\frac{1}{2m}(p-mat)^2+\tfrac{1}{2}m\omega^2(q-\tfrac{1}{2}at^2+a/\omega^2)^2. \tag{2.13}$$

This is a reflection of the well-known principle of equivalence between constant acceleration and a uniform gravitational field in that, relative to the accelerating frame, the equation of motion contains a term of the form which would arise in an inertial frame from a uniform gravitational field opposite in direction to the acceleration.

3. WAVEFUNCTION WHEN THE SOURCE UNDERGOES TRANSLATION

The Schrödinger equation for the Hamiltonian $\bar{H}$ (2.2) is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\bar{\psi}}{\partial Q^2}+\tfrac{1}{2}m\omega^2Q^2\bar{\psi}=i\hbar\frac{\partial\bar{\psi}}{\partial t'} \tag{3.1}$$

with solution

$$\bar{\psi}_n(Q,t')=[2^n n!(\pi\hbar/m)^{1/2}]^{-1/2}\exp[-i(n+\tfrac{1}{2})\omega t'] \times\exp[-mQ^2/2\hbar]H_n[(m/\hbar)^{1/2}Q]. \tag{3.2}$$

The Hamiltonian H (2.1) has Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2}+\tfrac{1}{2}m\omega^2(q-s)^2\psi=i\hbar\frac{\partial\psi}{\partial t}. \tag{3.3}$$

Equation (3.3) may be transformed to Eq. (3.1) by writing

$$Q=q+r_1,\quad t'=t, \tag{3.4}$$

$$\psi(Q,t')=\exp\left(\frac{i}{\hbar}f(Q,t')\right)\bar{\psi}(Q,t'), \tag{3.5}$$

where $f(Q,t')$ is found by simple calculation to be

$$f(Q,t')=-Qr_2+\int_{t_0'}^{t'}[r_2^2/2m-\tfrac{1}{2}m\omega^2(r_1+s)^2]\,dt'' \tag{3.6}$$

with r_1 and r_2 being the same as in (2.3). The solution of (3.3) is therefore

$$\psi_n(q,t)=\exp\left(\frac{i}{\hbar}f(q+r_1,t)\right)\bar{\psi}_n(q+r_1,t). \tag{3.7}$$

Under a Galilean transformation we use (2.8) and (2.9) to obtain

$$\psi_n(q,t)=\exp\left(\frac{i}{\hbar}(mUq-\tfrac{1}{2}mU^2t)\right)\bar{\psi}_n(q-Ut,t) \tag{3.8}$$

in which the phase introduced by t_0 is made zero by making Q coincide with q when $t=0$. For a uniformly accelerating source, (2.11) and (2.12) give

$$\psi_n(q,t)=\exp\left(\frac{i}{\hbar}(matq-\tfrac{1}{3}ma^2t^3+\tfrac{1}{2}\omega^2mat)\right) \times\bar{\psi}_n(q-\tfrac{1}{2}at^2+a/\omega^2,t). \tag{3.9}$$

As Wolf[12] has noted, a translation does not produce an integral transform relationship between ψ and $\bar{\psi}$, but a geometric transform which connects ψ and $\bar{\psi}$ by a phase factor as in (3.7). The relation which he gives [Ref. 12, Eq. (2.9c)] seems to be different, probably because he is considering a function of q only and not of q and t.

4. HAMILTONIAN OF HOMOGENEOUS QUADRATIC FORM IN TWO VARIABLES

As indicated in Sec. 1, Leach[6] has shown that a time-dependent Hamiltonian of nonhomogeneous quadratic form may be transformed to a time-independent homogeneous quadratic form by means of a time-dependent linear canonical transformation. Wolf[12] has shown that the wavefunctions of the corresponding Schrödinger equations are related by an integral transform except when the transformation matrix S has the form

$$S=\begin{bmatrix}a & 0\\ c & a^{-1}\end{bmatrix}. \tag{4.1}$$

In this case, the integral transform collapses to a geometric transform, i.e., the wavefunctions are related by a phase factor as in Sec. 3 together with a possible scaling term.

As the example given in (1.2) shows, S is not usually as simple as in (4.1). However, the Hamiltonian

$$\bar{H}=\tfrac{1}{2}\rho^{-2}(P^2+Q^2) \tag{4.2}$$

has a readily solvable Schrödinger equation. Clearly the wavefunction is

$$\bar{\psi}_n(Q,t)=[2^n n!(\pi\hbar)^{1/2}]^{-1/2}\exp[-i(n+\tfrac{1}{2})\int^t\rho^{-2}dt'] \times\exp[-Q^2/2\hbar]H_n[Q/\hbar^{1/2}]. \tag{4.3}$$

We shall see that it is always possible to transform the Hamiltonian

$$H=\tfrac{1}{2}[a^2p^2+b(pq+qp)+c^2q^2],\quad a^2c^2-b^2>0, \tag{4.4}$$

in which a, b, and c may be functions of time to that of (4.2) using an S of type (4.1). Thus the wavefunction corresponding to H (4.4) will be readily accessible from that of $\bar{H}$ (4.2).

Following Leach,[1] the elements of the coefficient matrix S of the transformation

$$\begin{bmatrix}Q\\ P\end{bmatrix}=\begin{bmatrix}y_1 & y_2\\ y_3 & y_4\end{bmatrix}\begin{bmatrix}q\\ p\end{bmatrix} \tag{4.5}$$

satisfy the set of first-order differential equations

$$\begin{aligned}&\dot{y}_1+by_1-c^2y_2=y_3\rho^{-2}\\ &\dot{y}_2+a^2y_1-by_2=y_4\rho^{-2}\\ &\dot{y}_3+by_3-c^2y_4=-y_1\rho^{-2}\\ &\dot{y}_4+a^2y_3-by_4=-y_2\rho^{-2}.\end{aligned} \tag{4.6}$$

Setting y_2 at zero, a solution of (4.6) gives

$$S=\begin{bmatrix}(a\rho)^{-1} & , & 0\\ [a\rho b-(\dot{a\rho})]/a^2, & & a\rho\end{bmatrix} \tag{4.7}$$

which is of the type (4.1). The function $\rho(t)$ is a solution of

$$(\ddot{a\rho})-2\dot{a}\dot{\rho}-2\dot{a}^2\rho/a+2\dot{a}\rho b+a\rho(a^2c^2-b^2)=a\rho^{-3}. \tag{4.8}$$

If we put $a^2=1$, $b=0$, and $c^2=\omega^2(t)$ in (4.8), we obtain Eq. (1.3) which relates to the Hamiltonian (1.1). The invariant for the Hamiltonian (4.4) [corresponding to the Lewis invariant (1.8) for (1.1)] is

$$I=\tfrac{1}{2}(a\rho)^{-2}q^2+\tfrac{1}{2}[a\rho p-[(\dot{a\rho})-a\rho b]q/a^2]^2. \tag{4.9}$$

Using the relationship given by Wolfe [Ref. 22, (2.8b)], the Schrödinger wavefunction for H (4.4) is

$$\psi_n(q,t)=(a\rho)^{-1/2}\exp\left[\frac{i}{2\hbar}\left\{\frac{(\dot{a\rho})}{a\rho}-b\right\}\frac{q^2}{a^2}\right] \times\bar{\psi}_n(q/a\rho,t). \tag{4.10}$$

When a, b, and c take the particular values given above, the wavefunction (4.10) is that which has been obtained using different approaches by Camiz *et al.*[17] and Khandekar and Lawande.[18]

5. ANISOTROPIC OSCILLATOR IN TWO DIMENSIONS

The Hamiltonian of the time-dependent anisotropic oscillator (taken as two-dimensional here, but the method may be used for the n-dimensional problem) is

$$H=\tfrac{1}{2}[p_1^2+\omega_1^2(t)\,q_1^2+p_2^2+\omega_2^2(t)\,q_2^2]. \tag{5.1}$$

Under a linear canonical transformation (5.1) may be transformed to

$$\bar{H}=\tfrac{1}{2}[P_1^2+Q_1^2+P_2^2+Q_2^2]. \tag{5.2}$$

The transformation is accomplished in two stages by

$$\begin{bmatrix} q_1' \\ q_2' \\ \hline p_1' \\ p_2' \end{bmatrix} = \left[\begin{array}{cc:cc} \rho_1^{-1} & 0 & 0 & 0 \\ 0 & \rho_2^{-1} & 0 & 0 \\ \hdashline -\dot{\rho}_1 & 0 & \rho_1 & \rho_0 \\ 0 & -\dot{\rho}_2 & 0 & \rho_2 \end{array}\right] \begin{bmatrix} q_1 \\ q_2 \\ \hline p_1 \\ p_2 \end{bmatrix}, \tag{5.3}$$

$$\begin{bmatrix} Q_1 \\ Q_2 \\ \hline P_1 \\ P_2 \end{bmatrix} = \left[\begin{array}{cc:cc} C_{11}\cos W_{11}+C_{12}\cos W_{12}, & 0 & -C_{11}\sin W_{11}-C_{12}\sin W_{12}, & 0 \\ 0, & C_{11}\cos W_{21}+C_{22}\cos W_{22} & 0, & -C_{21}\sin W_{21}-C_{22}\sin W_{22} \\ \hdashline C_{11}\sin W_{11}-C_{12}\sin W_{12}, & 0 & C_{11}\cos W_{11}-C_{12}\cos W_{12}, & 0 \\ 0, & C_{21}\sin W_{21}-C_{22}\sin W_{22} & 0, & C_{21}\cos W_{21}-C_{22}\cos W_{22} \end{array}\right] \begin{bmatrix} q_1' \\ q_2' \\ \\ p_1' \\ p_2' \end{bmatrix} \tag{5.4}$$

which are found using the general method given by Leach.[6] In (5.3), $\rho_1(t)$ and $\rho_2(t)$ are solutions of

$$\ddot{\rho}_1+\omega_1^2(t)\,\rho_1=\rho_1^{-3}, \quad \ddot{\rho}_2+\omega_2^2(t)\,\rho_2=\rho_2^{-3}, \tag{5.5}$$

respectively and in (5.4)

$$\begin{aligned} W_{11}&=\int_{t_0}^{t}(\rho_1^{-2}-1)\,dt', \quad W_{12}=\int_{t_0}^{t}(\rho_1^{-2}+1)\,dt', \\ W_{21}&=\int_{t_0}^{t}(\rho_2^{-2}-1)\,dt', \quad W_{22}=\int_{t_0}^{t}(\rho_2^{-2}+1)\,dt', \end{aligned} \tag{5.6}$$

$$C_{11}^2-C_{12}^2=1, \quad C_{21}^2-C_{22}^2=1. \tag{5.7}$$

As the Hamiltonian $\bar{H}$ (5.2) possesses the dynamical symmetry group SU(2) and the Poisson bracket of H and $\bar{H}$ is nonzero, the group SU(2) is a noninvariance symmetry group of H and the two-dimensional anisotropic oscillator is said to be characterized by this group.

In Sec. 4, we saw that the Schrödinger wavefunctions ψ and $\bar{\psi}$ were simply related (4.10) when S was of the type (4.1). The corresponding matrix in four dimensions is

$$S=\begin{bmatrix} S_1 & 0 \\ S_3 & S_4 \end{bmatrix} \tag{5.8}$$

in which the submatrices satisfy the conditions[14]

$$S_1S_4^T=I, \quad S_3S_4^T=S_4S_3^T. \tag{5.9}$$

Since the transformation matrix from (q_1,q_2,p_1,p_2) to (Q_1,Q_2,P_1,P_2) is not of the form (5.8) we would not expect the method of Sec. 4 to be applicable. However, the coefficient matrix in (5.3) is of that form. The intermediate Hamiltonian produced by (5.3) is

$$H'=\tfrac{1}{2}\rho_1^{-2}(p_1^{-2}+q_1^{-2})+\tfrac{1}{2}\rho_2^{-2}(p_2^{-2}+q_2^{-2}). \tag{5.10}$$

The Schrödinger equation for (5.10) is clearly separable and the wavefunction is given by

$$\psi'_{mn}(q_1',q_2',t)=\psi'_m(q_1',t)\,\psi'_n(q_2',t), \tag{5.11}$$

where both ψ'_m and ψ'_n are of the form of $\bar{\psi}_n$ in (4.3). The problem is reduced to the product of two one-dimensional problems and Wolf's relation may be applied to each part yielding

$$\psi_{mn}(q_1,q_2,t)=(\rho_1\rho_2)^{-1/2}\exp\left[\frac{i}{2\hbar}\left(\frac{\dot{\rho}_1}{\rho_1}\,q_1^2+\frac{\dot{\rho}_2}{\rho_2}\,q_2^2\right)\right] \times\psi'_m\left(\frac{q_1}{\rho_1},t\right)\psi'_n\left(\frac{q_2}{\rho_2},t\right). \tag{5.12}$$

An invariant for H (5.1) may be obtained in the form given by (1.15) with the associated invariant matrix being given by (1.18). The matrix S used is the product of the coefficient in (5.4) by that in (5.3). The invariant is

$$\begin{aligned} I=\tfrac{1}{2}\sum_{i=1}^{2}\,&[\rho_i^{-2}q_i^2(C_{i1}^2+C_{i2}^2+2C_{i1}C_{i2}\cos 2W_i) \\ &+(\rho_ip_i-\dot{\rho}_iq_i)^2(C_{i1}^2+C_{i2}^2-2C_iC_i\cos 2W_i) \\ &-4\rho_i^{-1}q_i(\rho_ip_i-\dot{\rho}_iq_i)\,C_{i1}C_{i2}\sin 2W_i] \end{aligned} \tag{5.13}$$

in which

$$W_1=\int_{t_0}^{t}\rho_1^{-2}\,dt', \quad W_2=\int_{t_0}^{t}\rho_2^{-2}\,dt'. \tag{5.14}$$

A simpler version of the invariant may be obtained from H' (5.10). Writing

$$H'=\rho_1^{-2}I_1+\rho_2^{-2}I_2, \tag{5.15}$$

$$I=I_1+I_2=\tfrac{1}{2}\sum_{i=1}^{2}\,[\rho_i^{-2}q_i^2+(\rho_ip_i-\dot{\rho}_iq_i)^2]. \tag{5.16}$$

Formally (5.16) is (5.13) with $C_{11}=1=C_{21}$, $C_{12}=0=C_{22}$. However, (5.16) does not lend itself to the construction of a matrix invariant since this involves W_{i1} and W_{i2}. To solve the Schrödinger equation we need only transform H to H', but to obtain the symmetry group we must go to $\bar{H}$.

6. HAMILTONIAN OF HOMOGENEOUS QUADRATIC FORM IN FOUR VARIABLES

The Hamiltonian is

$$H=\tfrac{1}{2}(\mathbf{q}^T,\mathbf{p}^T)\,A\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} \tag{6.1}$$

in which A is a positive definite symmetric 4×4 matrix with time-dependent elements and $\mathbf{q}$ and $\mathbf{p}$ are two vectors. It will be amenable to a similar method of solution for the wavefunction as the anisotropic oscillator in Sec. 5 provided there exists a matrix S of the form (5.8) such that the canonical transformation

$$\begin{bmatrix} \mathbf{Q} \\ \mathbf{P} \end{bmatrix} = S \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} \tag{6.2}$$

produces the Hamiltonian

$$\bar{H} = \tfrac{1}{2}[\mathbf{Q}^T, \mathbf{P}^T] D \begin{bmatrix} \mathbf{Q} \\ \mathbf{P} \end{bmatrix}, \tag{6.3}$$

where D is the diagonal matrix

$$D = \left[\begin{array}{cc:cc} \rho_1^{-2} & 0 & 0 & 0 \\ 0 & \rho_2^{-2} & 0 & 0 \\ \hdashline 0 & 0 & \rho_1^{-2} & 0 \\ 0 & 0 & 0 & \rho_2^{-2} \end{array}\right]. \tag{6.4}$$

The matrix S satisfies the differential equation [Ref. 6, Eq. (3.5)]

$$\dot{S} = \epsilon D S - S \epsilon A. \tag{6.5}$$

In general, provided the elements of D and A are continuous functions of the interval of time of interest, (6.5) has a solution set (Ince, Ref. 19, pp. 71, 72). We now show that under the additional constraint that S be of the form (5.8), (6.5) has a solution set provided D and A have continuous second derivatives.

In block matrix form (6.5) is

$$\begin{bmatrix} \dot{S}_1 & \dot{S}_2 \\ \dot{S}_3 & \dot{S}_4 \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} R^{-2} & 0 \\ 0 & R^{-2} \end{bmatrix} \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix} - \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} U & W \\ W^T & V \end{bmatrix}. \tag{6.6}$$

Setting S_2 at zero, the other submatrices are required to satisfy (5.9) and

$$\begin{aligned} &\dot{S}_1 + S_1 W^T = R^{-2} S_3, \quad S_4 = R^2 S_1 V \\ &\dot{S}_3 + S_3 W^T = -R^{-2} S_1 + S_4 U, \quad \dot{S}_4 - S_4 W = -S_3 V. \end{aligned} \tag{6.7}$$

If T is a solution of

$$2\dot{T} = T\{VW - W^T V - \dot{V}\} V^{-1}, \tag{6.8}$$

$$\begin{aligned} &S_1 = R^{-1} T, \quad S_4 = RTV, \\ &S_3 = \tfrac{1}{2} RT(VW + W^T V - \dot{V}) V^{-1} - \dot{R} T. \end{aligned} \tag{6.9}$$

That (6.8) has a solution follows directly from Ince's theorem (the inverse of V exists since A is positive definite). Substitution for S_1, S_2, and S_3 in the third of Eqs. (6.7) yields a second order equation which R must satisfy and this involves $\ddot{V}$. Hence the stronger conditions are on D and A than in the general case. The requirements (5.9) are satisfied to within an arbitrary constant matrix multiplier which may, without loss of validity, be taken as the unit matrix.

Thus a Hamiltonian of the form (6.1) may always be transformed to the diagonal form (6.3) using a coefficient matrix of the type (5.8). Using the technique of (5.15) and (5.16), the invariant is

$$I = \tfrac{1}{2}[\mathbf{q}^T(S_1^T S_1 + S_3^T S_3)\mathbf{q} + \mathbf{q}^T S_3^T S_4 \mathbf{p} + \mathbf{p}^T S_4^T S_3 \mathbf{q} + \mathbf{p}^T S_4^T S_4 \mathbf{p}]. \tag{6.10}$$

The same remark as for the anisotropic oscillator about the invariant matrix applies in this case. It is a matter of simple algebra to show that the wavefunction of the Schrödinger equation for (6.1) is

$$\psi_{mn}(q_1, q_2, t) = \exp\left[\int_{t_0}^{t} [\tfrac{1}{2} T_r(S_1 V S_3^T) - T_r(W)]\, dt'\right] \times \exp\left[-\frac{i}{2\hbar} \mathbf{q}^T S_3^T S_1 \mathbf{q}\right] \bar{\psi}_m(Q_1, t)\, \bar{\psi}_n(Q_2, t) \tag{6.11}$$

in which $\bar{\psi}_m$ and $\bar{\psi}_n$ have the form of (4.3) and

$$\mathbf{Q} = S_1 \mathbf{q}. \tag{6.12}$$

In the particular case when

$$U = \begin{bmatrix} \omega_1^2(t) & 0 \\ 0 & \omega_2^2(t) \end{bmatrix}, \quad V = I, \quad W = 0, \tag{6.13}$$

$$S_1 = \begin{bmatrix} \rho_1^{-1} & 0 \\ 0 & \rho_2^{-2} \end{bmatrix}, \quad S_3 = \begin{bmatrix} -\dot{\rho}_1 & 0 \\ 0 & -\dot{\rho}_2 \end{bmatrix}, \tag{6.14}$$

and (6.11) is the result (5.12) for the anisotropic oscillator. We note that the result (6.11) applies equally well to higher-dimensional problems (with more $\bar{\psi}$'s) and provides a suitable generalization of Wolf's result.

7. CONCLUSION

In this paper we have examined some simple time-dependent Hamiltonians of the harmonic oscillator type. They provide illustrations of the application of linear canonical transformations to quantum mechanical systems. In particular they indicate that the wavefunction of the Schrödinger equation for the Hamiltonian

$$H = \tfrac{1}{2}(\mathbf{q}^T, \mathbf{p}^T) A(t) \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} + \mathbf{B}(t)^T \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} + C(t) \tag{7.1}$$

may always be found. A translation will always transform (7.1) to a homogeneous quadratic form which may then be treated as was (6.1). Since the linear term in (7.1) may be regarded as a forcing term, such treatment provides an alternative approach to the examination of problems such as the dynamics of coherent states (cf. Mehta *et al.*[20,21]).

The Hamiltonians considered in this paper have all been positive definite. There has been some interest in recent years in nonpositive definite time-independent Hamiltonians (cf. Moshinsky *et al.*[22]). The examination of such Hamiltonians when they are time-dependent and quadratic would be of interest, especially with regard to the existence of any characterizing symmetry groups.

[1] P.G.L. Leach, "On a Generalization of the Lewis Invariant for the Time-Dependent Harmonic Oscillator," La Trobe University Departments of Mathematics preprint.
[2] H.R. Lewis, Jr., Phys. Rev. Lett. 18, 510, 636 (1967).
[3] H.R. Lewis, Jr., J. Math. Phys. 9, 1976 (1968).
[4] P.W. Seymour, Int. J. Eng. Sci. 1, 423 (1963).
[5] P.W. Seymour, R.B. Leipnik, and A.F. Nicholson, Aust. J. Phys. 18, 553 (1965).
[6] P.G.L. Leach, "On the Theory of Time-Dependent Linear Canonical Transformations as applied to Hamiltonians of the

Harmonic Oscillator Type," La Trobe University Departments of Mathematics preprint.
[7]C.J. Eliezer and A. Gray, SIAM J. Appl. Math. **30**, 433 (1976).
[8]N.J. Günther and P.G.L. Leach, J. Math. Phys. **18**, 572 (1977).
[9]D.M. Fradkin, Am. J. Phys. **33**, 207 (1965).
[10]D.M. Fradkin, Prog. Theor. Phys. **37**, 793 (1967).
[11]H.R. Lewis, Jr. and W.R. Riesenfeld, J. Math. Phys. **10**, 1458 (1969).
[12]K.B. Wolf, J. Math. Phys. **17**, 601 (1976).
[13]M. Moshinsky, SIAM J. Appl. Math. **25**, 197 (1973).
[14]M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772 (1971).
[15]A similar problem to this has been treated in a rather *ad hoc* fashion in K. Husimi, Prog. Theor. Phys. **99**, 381 (1953).
[16]C.J. Eliezer and P.G.L. Leach, "The Equivalence Principle and Quantum Mechanics," La Trobe University Departments of Mathematics preprint (to appear in Am. J. Phys.).
[17]P. Camiz, A. Geradi, C. Marchioro, E. Presutti, and E. Scacciatelli, J. Math. Phys. **12**, 2040 (1971).
[18]D.C. Khandekar and S.V. Lawande, J. Math. Phys. **16**, 384 (1975).
[19]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[20]C.L. Mehta and E.C.G. Sudarshan, Phys. Lett. **28**, 574 (1964).
[21]C.L. Mehta, P. Chand, E.C.G. Sudarshan, and R. Vedam, Phys. Rev. **157**, 1198 (1967).
[22]M. Moshinsky, T.H. Segilman, and K.B. Wolf, J. Math. Phys. **13**, 901 (1972).

Ground state properties and lower bounds for energy levels of a particle in a uniform magnetic field and external potential

Richard Lavine* and Michael O'Carroll

Departamento de Matemática, Pontifícia Universidade Católica, do Rio de Janeiro, Brazil
(Received 9 November 1976)

The Hamiltonian $H(B)$, for a particle of mass μ and charge e in a uniform magnetic field of strength B in the z direction and an external axially symmetric potential V, is a direct sum of operators $H(m,B)$ acting in the subspace of eigenvalue m of the z component of angular momentum L_z. Let $\lambda(B)$ $[\lambda(m,B)]$ denote the smallest eigenvalue of $H(B)$ $[H(m,B)]$. For $V=-e^2/r$ $(r=|x|)$, the attractive Coulomb potential, we obtain lower bounds $l(m,B)$, for the spectrum of $H(m,B)$ such that $l(0,B)>l(0,0)=\lambda(0,0)$ for $B>0$, $l(m,B)>l(0,B)$ for $m\neq 0$, and $l(|m|,B)-l(-|m|,B)=eB|m|/\mu$, $l(m',B)>l(m,B)$ if $m'<m\leq 0$. We show at least for an interval $[0,B'>0]$ of B that the ground state of $H(B)$ is the lowest eigenvalue, $\lambda(0,B)$, of $H(0,B)$ and is an almost everywhere positive function. If $V=A/r^2+r^2$ for $B=0$, $\lambda(0)=\lambda(0,0)$ and the ground state wavefunction is an almost everywhere positive function with L_z eigenvalue zero. However, for large A, we prove that for an interval of B away from $B=0$, the lowest eigenvalue, $\lambda(-1,B)$, of $H(-1,B)$ is below the lowest eigenvalue, $\lambda(0,B)$, of $H(0,B)$ and that the ground state of $H(B)$ is not an almost everywhere positive function.

INTRODUCTION

The Hamiltonian operator for a particle of mass μ and charge e in a constant magnetic field of strength B in the z direction, and an external potential V is

$$H(B)=-\Delta/2\mu+eBL_z/2\mu+e^2B^2(x^2+y^2)/8\mu+V(x,y,z) \tag{0.1}$$

acting in the Hilbert space $L^2(R^3)$. For a wide class of potentials, V, if $H(0)$ has a lowest eigenvalue at the bottom of its spectrum, it is known that for $B=0$ the corresponding eigenvector can be chosen as an almost everywhere positive function and the eigenvalue has multiplicity one.[1] An important step in the proof of this result is that $\exp(-H(0))$ preserves positivity, since each factor involved in the Trotter product formula for $\exp(-H(0))$ preserves positivity. For $B\neq 0$ this fails because of the term $eBL_z/2\mu$. Thus questions arise about the nature of the ground state wavefunction and its eigenvalue $\lambda(B)$.

We shall mainly consider potentials V which are functions of $\rho=(x^2+y^2)^{1/2}$ and z. In this case $H(B)$ is invariant under rotations about the z axis and $H(B)$ can be written as the direct sum of operators $\hat{H}(m,B)$ $(m=0,\pm 1,\pm 2,\cdots)$, where $\hat{H}(m,B)$ acts in the subspace $\mathcal{H}(m)$ of eigenvalue m of the z component of the angular momentum, L_z. In cylindrical coordinates

$$\mathcal{H}(m)=\{\exp(im\theta)\rho^{1/2}f(\rho,z): f\in L^2((0,\infty)\times R;\, d\rho\, dz)\}$$

and

$$\hat{H}(m,B)\exp(im\theta)\rho^{1/2}f=\exp(im\theta)\rho^{1/2}H(m,B)f,$$

where

$$H(m,B)=\frac{1}{2\mu}\left[\frac{\partial^2}{\partial\rho^2}-\frac{\partial^2}{\partial z^2}-\frac{1}{4\rho^2}+\left(\frac{m}{\rho}+\frac{eB\rho}{2}\right)^2\right]+V(\rho,z). \tag{0.2}$$

If we consider $H(m,B)$ as a quadratic form and ignore domain considerations, (0.2) indicates that $\lambda(m,B)\geq\lambda(0,0)$ and $\lambda(m,B)>\lambda(-m,B)$ for $m>0$. However, the question of whether $\lambda(m,B)$, $m\neq 0$, is larger or smaller than $\lambda(0,B)$ is more difficult, since the terms $(m/\rho+eB\rho/2)^2$ for different values of m are, in general, not comparable for all ρ. Thus it is not clear if the Hamiltonian $H(B)$ has $\lambda(0,B)$ as its ground state and a ground state wavefunction which is an element of $\mathcal{H}$ $(m=0)$ and invariant under rotations about the z axis.

In Sec. 1 we note that if V is a harmonic oscillator potential, then the ground state of $H(m,B)$ is $\lambda(0,B)$ and the ground state wavefunction is an almost everywhere positive function. We also give spectral results for $V=0$.

For the attractive Coulomb potential $V(r)=-e^2/r$, Jörgens[2] has shown that $H(B)$ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$, and that $H(B)$ has essential spectrum $[eB/2\mu,\infty)$ and infinite discrete spectrum, with $\lambda(m,B)\geq\lambda(0,0)=-e^2\mu/2$. In addition numerous variational calculations have been performed for $\lambda(m,B)$ as well as for higher energy levels of $H(m,B)$ (for example, see dos Santos and Brandi[3] and the references cited there). These calculations give upper bounds for $\lambda(m,B)$ which presumably approximate $\lambda(m,B)$. They show $\lambda(0,B)$ strictly increasing with B, and $\lambda(-|m|,B)$ first decreasing and later increasing, and $\lambda(-|m|,B)>\lambda(0,B)$. Lower bounds are necessary for an estimate of the error in these calculations. We are not aware of any lower bound in the literature other than that of Jörgens, which is not good enough to verify the behavior of $\lambda(m,B)$ indicated by the variational results.

In Sec. 2 we obtain lower bounds $l(m,B)$ for $H(m,B)$ by two methods. These lower bounds have the same sort of behavior as the variational calculations, and they indicate that the ground state is a function in $\mathcal{H}(0)$ which is positive almost everywhere, at least for $\gamma=B/e^3\mu^2<8$. It seems to us that Temple's inequality[4] (see also Sec. 3 of the present paper) could be used in variational calculations without much extra labor to

obtain better lower bounds for $\lambda(0,B)$ for small B. Recall that Temple's inequality requires a rigorous lower bound τ for the first excited level, such that τ is greater than the upper bound for $\lambda(0,B)$, of the form $\langle\psi,H\psi\rangle$, which is provided by the variational calculation. One could take τ equal to the first excited level of $H(0,0)$ as long as this is larger than $\langle\psi,H\psi\rangle$. However, our lower bounds for $\lambda(0,B)$ show that this is impossible for γ greater than about 2.2.

In Sec. 3 we give an example of a cylindrically symmetric potential for which $\lambda(0,B)>\lambda(-1,B)$ for values of B in an interval. Existence of a spherically symmetric potential where this happens is a more delicate question. However, we show using Temple's inequality for a lower bound for $\lambda(0,B)$ and a variational calculation as an upper bound for $\lambda(-1,B)$ that $\lambda(-1,B)<\lambda(0,B)$ for the potential $V=A/r^2+r^2$ for large A and an interval of B. The ground state is not an a. e. positive function. In these examples the potentials are not bounded but we show by a limiting argument that unbounded potentials are not necessary for this phenomenon to occur.

We conclude in Sec. 4 with some remarks about the increase of the ground state energy for an arbitrary potential when an arbitrary magnetic field is introduced.

1. HARMONIC OSCILLATOR

In the case of a cylindrically symmetric harmonic oscillator

$$V=e^2A^2(x^2+y^2)/8\mu+c^2z^2$$

the eigenfunctions and eigenvalues of $H(B)$ can be found explicitly. First note that the eigenfunctions of the operator H' obtained from $H(B)$ by dropping the term linear in L_z are, in Cartesian coordinates

$$\Phi_{ijk}(x,y,z)=ch_i(\mu wx)\,h_j(\mu wy)\,h_k(\mu wz),$$

where h_n is the nth Hermite function, and the associated eigenvalues are

$$\mu_{ijk}=w(i+j+1)+w'(k+\tfrac{1}{2}),$$

where $w=(B^2+A^2)^{1/2}e/2\mu$, $w'=ec/2\mu$, and $i,j,k=0,1,2,3,\cdots$.

Since H' is invariant under rotations about the z axis, this operator is reduced by $\mathcal{H}(m)$. Thus for each pair of nonnegative integers, n and k, there are $n+1$ orthogonal eigenfunctions $\psi_{nmk}\in\mathcal{H}(m)$, each of which is, for fixed z, the product of an exponential in $\rho^2=x^2+y^2$ and a polynomial of degree n in $\zeta=x+iy$ and $\bar{\zeta}$. Now there are $n+1$ monomials p of degree n such that $p(\exp(i\varphi)\zeta,\overline{\exp(i\varphi)\zeta})=\exp(im\varphi)p(\zeta,\bar{\zeta})$, namely $\zeta^{(n+m)/2}\bar{\zeta}^{(n-m)/2}$ for $m=-n,-n+2,\ldots,n-2,n$. It follows by a counting argument that functions ψ_{nmk} arise for just these values of m.

The ψ_{nmk} form a complete orthonormal system, and each n an eigenfunction of both H' and L_z, and therefore of $H(B)$. The corresponding eigenvalues are

$$\lambda_{nmk}=(e/2\mu)(B^2+A^2)^{1/2}(n+1)+mB+c(k+\tfrac{1}{2}).$$

Thus the lowest eigenvalue of $\mathcal{H}(m,B)$ is

$$\lambda(m,B)=(e/2\mu)\{(B^2+A^2)^{1/2}(|m|+1)+mB+c/2\}.$$

Note that, for large B, $\lambda(m,B)\sim(e/2\mu)[(|m|+m+1)B+c/2]$. Regarding m as a continuous variable, we have $\partial\lambda/\partial m=\{B+(B^2+A^2)^{1/2}m/|m|\}e/2\mu$, so that λ increases with $|m|$. Thus the ground state eigenfunction is ψ_{000} which is positive.

If $V=0$, $H(B)$ and $H(m,B)$ have the purely continuous spectrum $[eB/2\mu,\infty)$ and $[eB/2\mu(|m|+m+1),\infty)$, respectively. In this case the Hermite function in the variable z goes over to the function $\exp(ikz)(k\in\mathbb{R})$ and the sum over k goes over to the integral over $k\in(-\infty,\infty)$ and we have the Fourier integral expansion in z.

2. ATTRACTIVE COULOMB POTENTIAL

In this section we consider $H(B)$ with $V=-e^2/r$, the attractive Coulomb potential. In subsection A we obtain lower bounds $l(m,B)$ for $\lambda(m,B)$ by partitioning the operator $H(m,B)$. In subsection B we obtain the bounds $l(m,B)$ by using differential inequalities. In subsection C we discuss the properties of the lower bounds $l(m,B)$, and by using a crude trial function we obtain an upper bound for $\lambda(0,B)$ which is less than $\lambda(-1,B)$ for an interval $[0,B']$ of B. Thus the ground state of $H(B)$ is $\lambda(0,B)$ at least for $B\in[0,B']$, and a separate argument shows that the ground state eigenfunction is an almost everywhere positive function. We point out that crude variational calculations lead to the same qualitative behavior for $\lambda(0,B)$ and $\lambda(-1,B)$.

A. Partitioning of the Hamiltonian

For notational convenience let $\gamma=B/e^3\mu^2$, $\alpha=e^2$. For $B=0$, taking into account the explicit solution for the Coulomb potential for fixed m, $l\geq|m|$, and the results of Sec. 1 for $V=0$ we have, as quadratic form inequalities on $\mathcal{D}(H(|m|)\times\mathcal{D}(H(|m|))\subset H(m)\times H(m)$,

$$-\frac{\Delta}{2\mu}+\frac{\sqrt{\alpha}}{2\mu}Bm+\frac{\alpha}{8\mu}B^2(x^2+y^2)\geq\frac{\sqrt{\alpha}B}{2\mu}[|m|+1+m],$$

$$-\frac{\Delta}{2\mu}-\frac{\alpha}{r}\geq\frac{\alpha^2\mu}{2}(|m|+1)^2, \tag{2.1}$$

so that for any $a\in[0,1]$

$$\begin{aligned}H(m,B)&=\left(-\frac{a\Delta}{2\mu}-\frac{\alpha}{r}\right)+\left(-\frac{(1-a)}{2\mu}\Delta\right.\\&\quad\left.+\frac{\sqrt{\alpha}Bm}{8\mu}+\frac{\alpha B^2}{8\mu}(x^2+y^2)\right)\\&\geq-\frac{\alpha^2\mu}{2}\left\{\frac{a^{-1}}{(|m|+1)^2}-\gamma[\sqrt{1-a}(|m|+1)+m]\right\}\\&=L(a,m,\gamma).\end{aligned} \tag{2.2}$$

Maximizing $L(a,m,\gamma)$ with respect to a, we obtain the lower bound $l(m,B)=L(a,m,\gamma)$ with $\gamma^2a^4=[2(|m|+1)^{-3}]^2\times(1-a)$. By the well-known formula for the roots of a quartic equation we can obtain $l(B)$ as an explicit function of γ, or for each $a\in[0,1]$ we have $\gamma=[4(1-a)]^{1/2}/a^2$. As a increases from 0 to 1, γ decreases from ∞ to 0. Both γ and l can be calculated in terms of a.

B. Differential inequalities

The same results can be obtained using integration by parts. Suppose that $\varphi\in C_0^\infty([0,\infty)\times\mathbb{R})$, and h is a con-

tinuously differentiable real-valued function on $[0,\infty)\times\mathbb{R}$. Then integration by parts in

$$\int_0^\infty \int_{-\infty}^\infty \left|\left[\frac{\partial}{\partial z} - h(\rho,z)\right]\varphi(\rho,z)\right|^2 dz\,d\rho \geq 0$$

gives

$$\int_0^\infty \int_{-\infty}^\infty \bar{\varphi}\left(-\frac{\partial^2}{\partial z^2} + h^2 + \frac{\partial h}{\partial z}\right)\varphi\,dz\,d\rho \geq 0. \tag{2.3}$$

Similarly, we have

$$\int_{-\infty}^\infty \int_1^\infty \bar{\varphi}\left(-\frac{\partial^2}{\partial \rho^2} + f^2 + \frac{\partial f}{\partial \rho}\right)\varphi\,d\rho\,dz \geq 0 \tag{2.4}$$

if f is a continuously differentiable real-valued function on $(0,\infty)\times\mathbb{R}$ such that $\lim_{\rho\to 0} f(\rho,z)|\varphi(\rho,z)|^2 = 0$. To apply this to $H(m,B)$, note that $C_0^\infty(\mathbb{R}^3)$ is a core for $H(B)$ and if $\Phi = \rho^{1/2}\exp(im\theta)\varphi(\rho,z)$ belongs to $C_0^\infty(\mathbb{R}^3)$, then $r^{-1}\partial\Phi/\partial\theta = n\Phi/\rho$ must be bounded, which implies $\varphi \leq C\rho^{1/2}$ if $m=0$, and $\varphi \leq C\rho^{3/2}$ if $m\neq 0$. Thus (2.6) can be applied with f if $\lim\rho^\alpha f(\rho,z) = 0$ for $\alpha = \frac{1}{2}$ $(m=0)$ or $\alpha = \frac{3}{2}$ $(m\neq 0)$. In particular, let

$$f = \frac{-\beta\rho}{2(|m|+1)(\rho^2+z^2)^{1/2}} + \frac{|m|+\frac{1}{2}}{\rho} - \frac{\rho b}{2}\sqrt{1-a}$$

$$h = \frac{-\beta z}{2(|m|+1)(\rho^2+z^2)^{1/2}}.$$

Then adding (2.3) and (2.4) gives

$$-\frac{\partial^2}{\partial\rho^2} - \frac{\partial^2}{\partial z^2} + \frac{m^2-\frac{1}{4}}{\rho^2} - \frac{\beta}{(\rho^2+z^2)^{1/2}} + \frac{b^2(1-a)\rho^2}{4}$$
$$\geq b\sqrt{1-a}\,(|m|+1) - \frac{\beta^2}{4(|m|+1)^2} - \frac{\beta b\sqrt{1-a}\,\rho^2}{2(\rho^2+z^2)^{1/2}(|m|+1)}. \tag{2.5}$$

Now

$$\frac{\beta b\sqrt{1-a}\,\rho^2}{2(\rho^2+z^2)^{1/2}(|m|+1)} \leq \frac{\beta b\sqrt{1-a}}{2(|m|+1)}\rho$$
$$\leq \frac{ab^2}{4}\rho^2 + \frac{1}{4a}\frac{(1-a)\beta^2}{(|m|+1)^2}.$$

Combining this with (2.5) gives

$$2\mu H(m,b) \geq b(\sqrt{1-a})(|m|+1) - \frac{\beta^2}{4(|m|+1)^2} + mb$$
$$-\frac{1}{4}\frac{1-a}{a}\frac{\beta^2}{(|m|+1)^2}$$
$$= b[\sqrt{1-a}\,(|m|+1)+m] - \frac{\beta^2}{4(|m|+1)^2 a},$$

which is the same as (2.2), since $\alpha = \beta/2\mu$, $b = \sqrt{\alpha}B$, and $\gamma = b/\mu^{3/2}$.

C. Properties of $l(B)$ and $l(m,B)$

Here we give some properties of the lower bounds $l(B)$ and $l(m,B)$ and show that the ground state energy is $\lambda(0,B)$ and the ground state wavefunction is positive a. e. We note that $l(B) = l(0,B)$ so we only need consider $l(m,B)$. It is useful to let $\delta = \sqrt{1-a}$; then $\gamma = 2\times(|m|+1)^{-3}(1-\delta^2)^{-2}\delta$ is an increasing function of δ. We first examine the dependence on γ for fixed m. Since $da/d\gamma = -\gamma a^4(|m|+1)^6/2[\gamma^2 a^3(|m|+1)^6+1]$, we find that

$$dl(m,B)/d\gamma = \frac{\mu\alpha^2}{2}\left[\frac{(|m|+1)^4 a^2}{2}\gamma + m\right]$$
$$= \frac{\mu\alpha^2}{2}[(|m|+1)\delta + m].$$

Thus, for $m\geq 0$, $l(m,B)$ strictly increases with B. For $B=0$, $l(m,0) = -(\mu\alpha^2/2)(|m|+1)^{-2}$, the exact Coulomb energy levels. For $m<0$, $dl(m,0)/d\gamma = \mu\alpha^2 m/2 < 0$ and, for large γ, $dl(m,B)/d\gamma$ behaves like $(\mu\alpha^2/2)(|m|+1) > 0$, $dl(m,B)/d\gamma = 0$ for $\delta = -m/(|m|+1)$. Thus for negative m, $l(m,B)$ decreases to a minimum and then increases. For fixed B, since δ only depends on $|m|$, $l(|m|,B) - l(-|m|,B) = \mu\alpha^2\gamma|m| = eB|m|/\mu$. Writing $l(m,B)$ in terms of δ, we have

$$l(\mp|m|,B) = (\mu\alpha^2/2)[(1-\delta^2)^{-1}(|m|+1)^{-2}$$
$$-2\delta(|m|+1)^{-3}(1-\delta^2)^{-2}(\delta|m|+1)\mp|m|)].$$

Treating $|m|$ as a continuous variable, taking the derivative $dl/d|m|$ with γ fixed, and noting that $d\delta/d|m| = 3\delta(1-\delta^2)(1+3\delta^2)^{-1}(|m|+1)^{-1}$, we find

$$\frac{dl}{d|m|}(\mp|m|,B)$$
$$= -\frac{\alpha\mu^2}{2}(1-\delta^2)^{-2}2(|m|+1)^{-3}(1+3\delta^2)^{-1}P_\mp(\delta),$$

where $P_\mp(\delta) = \pm 3\delta^3 - 3\delta^2 \pm -1$ and $P_\mp(\delta) < 0$ for $\delta\in[0,1)$. Thus $(dl/d|m|)(\mp|m|,B) > 0$ and the lower bounds for fixed field γ are strictly increasing functions of $|m|$. Thus we conclude that $l(m',B) > l(m,B)$ for $m' > m \geq 0$ and $l(m',B) > l(m,B)$ if $m' < m \leq 0$. In particular, we have $l(m,B) > l(-1,B) > l(0,B)$ for $m \neq 0, 1, \cdots$.

In order to show that the ground state is $\lambda(0,B)$, it is sufficient to exhibit $\psi\in\mathcal{D}(H(0))$ with $|\psi| = 1$ such that $u(0,B) \equiv (\psi, H(0,B)\psi) < l(-1,B)$. Previous variational calculations[3] for $\lambda(0,B)$ have been performed and the graph of $l(-1,B)$ indicates that $u(0,B) < l(-1,B)$ at least for $\gamma\in[0,5]$. For values of $\gamma = 25$ and $\gamma = 100$ these variational values are still above our lower bound $l(-1,B)$. Even for a crude variational wavefunction (an optimized Gaussian) we have $u(0,B) < l(-1,B)$ at least for γ in the interval $[0,8]$.

Take $\psi = C\exp(-\alpha r^2)$, $C = (2\alpha/\pi)^{3/4}$. Minimizing $u(0,B)$ with respect to α, we find $u(0,B) = (\mu\alpha^2/2)(6\alpha - 6\sqrt{2\alpha/\pi})$, where $\gamma^2 = 16\alpha^2(3/2 - \sqrt{2/\pi\alpha})$. For $\gamma = 0$, $u(0,B) = -(\mu\alpha^2/2)(8/3\pi) < -\mu\alpha^2/8 = l(-1,0)$; plotting $u(0,B)$ numerically we find that $u(0,B) < l(-1,B)$ at least for γ in the interval $[0,8]$. If we had used the ground state wavefunction for $H(0,B)$ as ψ, then $u(0,B) = -(\alpha\mu^2/2)(1-\gamma^2/2)$ and $u(0,B) > l(-1,B)$ for $\gamma \approx 1.5$. Since the ground state wavefunction $\varphi\in H(0)$, it follows by applying theorem of Hoegh-Krohn and Simon[1] to the operator $H(0,B)$ that φ is an almost everywhere positive function.

3. POTENTIALS WHERE $\lambda(-1,B) > \lambda(0,B)$

In this section we give examples of a cylindrically symmetric and spherically symmetric potential such

that $\lambda(-1,B)>\lambda(0,B)$ for an interval of B and the ground state eigenfunction is not a. e. positive, even though, for $B=0$, $\lambda(0)=\lambda(0,B)$ and the ground state is an a. e. positive element of $\mathcal{H}(m=0)$. Referring to the form of $H(m,B)$ in cylindrical coordinates, we recall that for two values of m the terms $f(m,\rho)=(m/\rho+B\rho/2)^2$ are not comparable for all ρ. However, if ρ is restricted to the interval $[b,c]$ where $b=(-2m/B_0)^{1/2}$ and $c>b$, then $f(\rho,m)<f(\rho,0)$ for $\rho\in[b,c]$, $m<0$, and $B>B_0$. Taking as our Hilbert space $L^2(Q)$, where $Q=\{\rho,z\,|\,b<\rho<c, -d<z<d\}$ and defining $H(m,B)$ with Dirichlet boundary conditions, we have $H(0,B)>H(-1,B)$ for $B>B_0$ and $V(\rho,z)$ bounded so that by the minimax principle $\lambda(0,B)>\lambda(-1,B)$. Thus the ground state ψ is a linear combination of vectors in $\mathcal{H}(m\neq 0)$. Let $\varphi\in\mathcal{H}(m=0)$ such that $\varphi>0$ a. e., then $(\varphi,\psi)=0$ and so ψ is not an a. e. positive function.

To find a spherically symmetric potential such that $\lambda(-1,B)<\lambda(0,B)$ is more difficult since for a spherically symmetric region S, $f(\rho,0)$ does not dominate $f(\rho,-1)$ for all $\rho\in S$. However, if the potential is $V=A/r^2+Cr^2$, $C=1$, then we can prove that $\lambda(-1,B)<\lambda(0,B)$ for suitable values of A and B. The values of A and B which work are indicated by an intuitive variational approximation for $\lambda(0,B)$ and $\lambda(-1,B)$. First let us properly define $H(B)$ and $H(m,B)$ and determine the spectrum for $B=0$. We define $H(m,0)$ as the direct sum of operators $H_l(m,0)$, $l=|m|,\ |m|+1,\cdots$, where $H_l(m,0)$ is the self-adjoint closure of the differential operator $-d^2/dr^2+l(l+1)/r^2+A/r^2+Cr^2$, A, $C>0$ defined on $C_0^\infty((0,\infty))$. $H_l(m,0)$ is self-adjoint by Theorem X. 11 of Reed and Simon.[5] We note that for $B^2/4<C$ that the term $mB+B^2\rho^2/4$ is a form-bounded perturbation with bound <1 of the form associated with $H(m,0)$. We define $H(m,B)$ as the unique self-adjoint operator associated with this form sum. $H(B)$ is then defined as the direct sum over m of the operators $H(m,B)$. $H(m,B)$ has a purely discrete spectrum with eigenvalues

$$\lambda_{nl}=\sqrt{C}\,\{4n+2+[(2l+1)^2+4A]^{1/2}\},\tag{3.1}$$

where $n=0,1,2,\cdots$ and $l=|m|,\ |m|+1,\cdots$. As $H(m,B)$ is invariant under three-dimensional reflections (parity), $H(m,B)$ is reduced by the subspaces associated with $H_e(m,0)=\sum_{l\ \text{even}}H_l(m,0)$ and $H_0(m,B)=\sum_{l\ \text{odd}}H_l(m,0)$. These subspaces are also reducing subspaces for $H(m,B)$. We note that for large A, $\lambda_{nl}-\lambda_{nl'}\approx\sqrt{C/A}\,(l-l')(l+l'+1)$; in particular, $\lambda_{n0}-\lambda_{n1}\approx-2\sqrt{C/A}$. The ground state eigenvalue is $\lambda_{00}=\sqrt{C}[2+(1+4A)^{1/2}]$ and the corresponding normalized eigenfunction is

$$\psi=Nr^\alpha\exp[-(\sqrt{C}/2)r^2],\tag{3.2}$$

where $A=\alpha^2+\alpha$ and $N^2=2^{\alpha+2}C^{(\alpha+1)/2+1/4}/4\pi\cdot1\cdot3\cdot5\cdots(2\alpha+1)$.

We now give an intuitive argument based on upper bounds obtained from variational calculations for $\lambda(0,B)$ and $\lambda(-1,B)$ which indicates that $\lambda(0,B)>\lambda(-1,B)$ for A large and an interval of B small but away from $B=0$. For small B we expect the lowest eigenfunction of $H(0,B)$ to be nearly radial. Thus we choose a radial function χ to get a rigorous upper bound $(\chi,H(0,B)\chi)$ for $\lambda(0,B)$. Since $(\chi,x^2,\chi)=(\chi,y^2\chi)=(\chi,z^2\chi)$ we have $(\chi,H(0,B)\chi)=(\chi,[H(0,0)+\frac{2}{3}(B^2r^2/4)]\chi)$. For $C=1$ we find, taking χ to be the ground state of $H(0,0)$ with $C=1+B^2/6$ and using (3.1),

$$\lambda(0,B)<(\chi,H(0,B)\chi)=(1+B^2/6)^{1/2}(2+\sqrt{1+4A})\equiv u(0,B).$$

Similarly a rigorous upper bound for $\lambda(-1,B)$ is given by $(\Phi,H(0,B)\Phi)$, where Φ has the form $\Phi=f(r)\,Y_{1-1}(\theta,\phi)$ [$Y_{lm}(\theta,\phi)$ is the spherical harmonic] and $f(r)$ is chosen to minimize

$$(\Phi,H(-1,B)\Phi)=(\Phi,H(-1,B)\Phi)=(\Phi,[H(-1,0)-B+\tfrac{4}{5}(B^2r^2/4)]\Phi).$$

Choosing the minimizing f and taking $C=1$ in $H(m,B)$ we find from (3.1) that

$$\lambda(-1,B)<(\varphi,H(-1,B)\varphi)=(1+B^2/5)^{1/2}(2+\sqrt{9+4A})-B\equiv u(-1,B).$$

Note that $u(0,B)$ increases with increasing B; $u(-1,B)$ decreases to a minimum at $B=B'\approx5/2\sqrt{A}$ for large A then increases for $B>B'$. For large A, $u(0,B')-u(-1,B')\approx7/24\sqrt{A}$, $u(-1,B')-u(0,0)\approx750\sqrt{A}$, and $u(0,B')-u(0,0)\approx25/24\sqrt{A}$. We find that the first crossing, $u(0,B)=u(-1,B)$, occurs for $B=b=15(1-(11/15)^{1/2})/\sqrt{A}\approx2.155/\sqrt{A}$ and $u(0,b)-u(0,0)\approx0.774/\sqrt{A}$. Thus $u(0,B)>u(-1,B)$ at least for $B\in(b,B']$.

For a rigorous argument we need a lower bound, $l(0,B)$, for $\lambda(0,B)$ to replace $u(0,B)$ in the above intuitive argument. We obtain $l(0,B)$ from Temple's inequality which states that

$$\lambda(0,B)\geq(\psi,H\psi)-[(\psi,H^2\psi)-(\psi,H\psi)^2]/[\mu-(\psi,H\psi)]=l(0,B),$$

where $H=H(0,B)$, ψ is an approximate eigenfunction, $|\psi|=1$ and $\mu=\mu(B)$ is a lower bound for the first excited level, $\mu'(B)$, of $H(0,B)$ which obeys $\mu>(\psi,H\psi)$. We note that if ψ is chosen as the eigenfunction for $\lambda(0,B)$, then $l(0,B)=\lambda(0,B)$. By the minimax principle $\mu'(B)>\mu'(0)$, so recalling that in the even parity subspace only even l's appear we may take ($C=1$ in what follows)

$$\mu(B)=\lambda_{02}=2+\sqrt{25/4A}.$$

Taking ψ from (3.2) as the approximate eigenfunction, we find that

$$\tfrac{3}{2}(\psi,\rho^2\psi)=(\psi,r^2\psi)=(2\alpha+3)/2,$$

$$\tfrac{3}{2}(\psi,\rho^2r^2\psi)=(\psi,r^4\psi)=15(\psi,\rho^4\psi)/8=(2\alpha+3)(2\alpha+5)/4,$$

$$(\psi,H\psi)=(\psi,(\lambda_{00}+\tfrac{1}{4}B^2\rho^2)\psi)=\lambda_{00}+\tfrac{1}{6}B^2(2\alpha+3)/2,$$

$$(\psi,H^2\psi)=|H\psi|^2=|(\lambda_{00}+\tfrac{1}{4}B^2\rho^2)\psi|^2=\lambda_{00}^2+\frac{\lambda_{00}B^2}{3}\,\frac{(2\alpha+3)}{2}+\frac{B^4}{2\cdot15}\,\frac{(2\alpha+3)(2\alpha+5)}{4}.$$

Substituting in Temple's inequality, we obtain

$$l(0,B)=\lambda_{00}+\frac{B^2}{6}\,\frac{(2\alpha+3)}{2}-\left[B^4\,\frac{(2\alpha+3)}{4}\left(\frac{2\alpha+5}{30}-\frac{2\alpha+3}{36}\right)\Big/\left((2+25+4A)-(2+1+4A)-\frac{B^2}{6}\,\frac{(2\alpha+3)}{2}\right)\right].$$

For large A, $\alpha\approx\sqrt{A}$ and

$$l(0,B)\approx(2+\sqrt{1+4A})+\frac{B^2\sqrt{A}}{6}\left[\frac{B^4\sqrt{A}}{2}\times\left(\frac{\sqrt{A}}{15}-\frac{\sqrt{A}}{18}\right)\Big/\left(\frac{6}{\sqrt{A}}-\frac{B^2\sqrt{A}}{6}\right)\right].$$

Now, recalling that from the variational calculation the minimum of $u(-1,B)$ occurs for $B=B'=5/2\sqrt{A}$ and $\lambda(-1,B')<u(-1,B')=2+\sqrt{1+4A}+0.750/\sqrt{A}$, we find that $l(0,B')\approx 0.998/\sqrt{A}+2+\sqrt{1+4A}$ and $\lambda(0,B')-\lambda(-1,B')\geq l(0,B')-u(-1,B')\approx 0.998/\sqrt{A}-0.750/\sqrt{A}=0.248/\sqrt{A}$. This also holds for an interval about B' by continuity. Thus we have proved that $\lambda(-1,B)<\lambda(0,B)$ for this interval. The ground state of $H(B)$ cannot be an a. e. positive function, because it is orthogonal to $\mathcal{H}(0)$ which contains functions positive a. e.

The fact that the potential $V+A/r^2+r^2$ is infinite at $r=0$ and $r=\infty$ is not crucial for showing that $\lambda(-1,B)<\lambda(0,B)$. We can take a sequence of finite cutoff potentials, V_n, which converge monotonically to V. By Theorems 3.6 and 3.15, of Kato,[6] Chap. VIII, any finite number of lowest eigenvalues of the Hamiltonian with V_n converge to the corresponding lowest eigenvalues of the Hamiltonian with V.

We remark that it is the specific relation between the coupling constants of the linear term and quadratic term in B that make the question of ground state inversion for $H(B)$ a difficult one. If, for example, the linear term in B has a large enough coupling constant then the lowest eigenvalue in $\mathcal{H}(-1)$ will be below that in $\mathcal{H}(0,)$. We conjecture that if the potential is radial and attractive $(dV/dr>0)$ and if $H(B)$ has eigenvalues, then the lowest eigenvalue $\lambda(0,B)$ in the $m=0$ subspace $\mathcal{H}(m=0)$ and thus the ground state wavefunction is positive a. e.

4. REMARKS ON GENERAL POTENTIALS

It is clear from (0.2) that in the case of cylindrical symmetry the ground state energy with no magnetic field is strictly less than the energy with the (constant) field present.[2] Simon has remarked that for a general potential V the ground state energy does not decrease with the addition of a general magnetic potential a.[7] Another argument for this can be given using the ideas of Sec. 2B.

Let f and a be real locally square integrable vector fields on $\mathbb{R}^n$. Then for $\varphi\in C_0^\infty(\mathbb{R}^n)$,

$$\|(\nabla-ia-f)\varphi\|^2=\|(\nabla-ia)\varphi\|^2+\int(\nabla\cdot f+|f|^2)(\varphi)^2\,dx, \tag{4.1}$$

where $\nabla\cdot f$ is taken in the sense of the theory of distributions, because

$$\begin{aligned}\|(\nabla-ia-f)\varphi\|^2&=\|(\nabla-ia)\varphi\|^2+\|f\varphi\|^2\\&\quad-(\nabla\cdot\varphi,f\varphi)-(f\varphi,\nabla\varphi)\\&=\|(\nabla-ia)\varphi\|^2-\int f\cdot\nabla|\varphi|^2.\end{aligned}$$

It follows that if V is locally L^1 and bounded below and if $V\geq|f|^2-\nabla\cdot f+\lambda$, then the quadratic form for $H(a)=-(\nabla-ia)^2+V$ is bounded below by λ, since the left-hand side of (4.1) is nonnegative. If u is the ground state of $H(0)$ with $H(0)u=\lambda u$, then $f=\nabla u/u$ is a natural choice, for

$$\nabla\cdot\frac{\nabla u}{u}=\frac{\Delta u}{u}-\left|\frac{\nabla u}{u}\right|^2=V-\lambda-\left|\frac{\nabla u}{u}\right|^2. \tag{4.2}$$

Thus, (4.1) and (4.2) would give $H(a)\geq\lambda$ if it was known that $\nabla u/u$ is locally L^2 and that the formal calculation (4.2) is correct. This can be done, but it is more convenient to consider $F=\nabla u/(u+\epsilon)$, which is automatically square integrable, since $u\in\mathcal{D}(H(0))\subset\mathcal{D}(\nabla)$. We have $\Delta u=(V-\lambda)u\in L^1_{\rm loc}$,[8] and

$$\nabla\cdot\frac{\nabla u}{u+\epsilon}=\frac{1}{u+\epsilon}(V-\lambda)u-\left|\frac{\nabla u}{u+\epsilon}\right|^2. \tag{4.3}$$

This can be justified by taking a sequence u_n converging suitably to u. Then (4.1) and (4.3) give

$$(\nabla-ia)^2+V-\frac{\epsilon Vu}{u+\epsilon}\geq\lambda$$

but as $\epsilon\to 0$, $\int\epsilon Vu/(u+\epsilon)|\varphi|^2\,dx\to 0$ for any $\varphi\in C_0^\infty(\mathbb{R}^n)$, so $H(a)\geq\lambda$.

Of course, one really wants a stronger result, that the lower bound increases when the field is added (and information about how much it increases.) The identity (4.1) suggests the following argument for the increase of the ground state energy unless there is a gauge transformation which transforms $H(a)$ into $H(0)$. It can probably be made rigorous, but since it does not give numerical information we do not attempt this.

Let φ be the ground state of $H(a)$ and suppose the energy is λ; then $(\nabla-ia-\nabla u/u)\varphi=0$, or $\nabla(\varphi/u)=ia(\varphi/u)$, so φ/u provides the required gauge transformation.

This indicates that even when the magnetic field is zero, the ground state energy can rise in the presence of a vector potential a such that $\oint a\cdot dr\neq 2\pi n$ around some closed path. This points to another physical effect of the vector potential.[9]

*On leave from the Department of Mathematics, University of Rochester, New York.

[1]R. Hoegh-Krohn and B. Simon, J. Func. Anal. **9**, 121 (1972).
[2]K. Jorgens, Math. Z. **96**, 355 (1967).
[3]R. R. dos Santos and H. S. Brandi, "Addendum to Hydrogen atoms in strong Magnetic Fields," Nota Cientifica 17-75, Departmento de Fisica, PUC/RJ Brazil.
[4]G. Temple, Proc. Roy. Soc. A **119**, 276 (1928) (see also Ref. 5, Chap. XIII).
[5]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II* (Academic, New York, 1975).
[6]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[7]B. Simon, private communication, to be published.
[8]See Ref. 5, X.32.
[9]Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).

On Schwinger functionals—positive extensions, moment problems, and representations

John L. Challifour*

ZiF, Universität Bielefeld, 48 Bielefeld, Germany

Steven P. Slinker

Department of Mathematics, Texas A & M University, College Station, Texas 77843
(Received 29 November 1976)

We construct extensions for a class of Schwinger functions at noncoincident arguments to symmetric states on the Borchers algebra. Conditions are given for these states to be strongly positive. For strongly positive states the relation between uniqueness of the Euclidean measure, polynomial density, and self-adjointness for the Euclidean field is examined.

1. INTRODUCTION

Success in recent years of Euclidean techniques in constructive quantum field theory poses questions as to the general mathematical relation between relativistic and Euclidean quantum fields. When sharp time fields exist this connection has been given by Nelson[1] and Simon[2] encompassing the usual models in two and three space—time dimensions. For more singular theories expected in four dimensions, in Ref. 3, we showed the connection to be given by extending the Schwinger functions of Osterwalder and Schrader[4] to a symmetric, strongly positive state on the Borchers algebra over the underlying test function space. From strong positivity, the existence of a probability measure on real tempered distributions was derived and the Euclidean field appeared as the related generalized random process. This was also shown independently in Ref. 5 and in a somewhat different formulation in Ref. 6. The central mathematical problem is then extension of a linear functional from a closed subspace of a nuclear *-algebra to the whole algebra, subject to requirements of symmetry and strong positivity.

If one is only concerned with the existence of a measure, possibly complex, then necessary and sufficient conditions for the extension have been found by Borchers and Yngvason[5] as continuity ($\hat{\tau}$-continuity) with respect to a topology weaker than Mackey topology (τ-continuity) for the field algebra. In Sec. 2 we settle an intermediate question by showing existence of symmetric, positive extensions for all Osterwalder—Schrader theories which are order bounded. This is also a purely topological constraint requiring the singularity of the Schwinger functions at coincident arguments not to become more severe as their order increases. Such is the case for all presently known models. Our method extends a technique first employed in this context by Yngvason[7] and leads to a positive state dominated by a strongly positive state. We show in Sec. 3 that if the Euclidean field defined by the strongly positive state is essentially self-adjoint, then the dominated state is also strongly positive. Here it is necessary for us to return to the moment problem for the Euclidean measure to correct and extend results in Ref. 3 due to a gap in the proof of Lemma 3.4 of that paper. A maximal measure will be one for which polynomials in the Euclidean field are dense in all L_p, $1 \leq p < \infty$, and an extremal measure one with density only for $1 \leq p \leq 2$. We show that maximal measures are always unique and define self-adjoint standard representations of the symmetric Borchers algebra in the sense of Powers.[8] The Carleman criterion leads to a maximal measure and unique extremal measures arise when the Euclidean field is essentially self-adjoint. The converse question is posed in the last section as an open problem.

2. POSITIVE EXTENSIONS

Throughout our notation follows Ref. 3, which should also be consulted for motivation. Let $\underline{S} = \{1, S_1, S_2, \ldots, S_n, \cdots\}$ denote a symmetric linear functional on $\underline{\mathcal{S}}_0 = \oplus_n \mathcal{S}(\mathcal{E}_n)$, where $\mathcal{S}(\mathcal{E}_n)$ consists of functions in the Schwartz space $\mathcal{S}_n := \mathcal{S}(\mathbb{R}^{4n})$ vanishing with all derivatives whenever two arguments coincide. $\underline{\mathcal{S}}_0$ is a closed subspace of the Borchers algebra $\underline{\mathcal{S}} := \oplus_n \mathcal{S}_n$ carrying the locally convex direct sum topology (τ-topology) inherited from $\mathcal{S} := \mathcal{S}_1$. In terms of finite sequences $\underline{f} = \{f_0, f_1, f_2, \ldots, f_n, 0, \cdots\}$ in $\underline{\mathcal{S}}$, $f_0 \in \mathbb{C} := \mathcal{S}_0$; the product

$$(\underline{f} \times \underline{g})_n(x_1, \ldots, x_n) := \sum_{k=0}^{n} f_k(x_1, \ldots, x_k) g_{n-k}(x_{k+1}, \ldots, x_n),$$

and involution $(\underline{f}^*)_n(x_1, \ldots, x_n) = \bar{f}_n(x_n, \ldots, x_1)$ make $\underline{\mathcal{S}}$ into a complete nuclear *-algebra with unit $\underline{1} = \{1, 0, \cdots\}$. A state $\underline{T} \in \underline{\mathcal{S}}'$ is normalized, $\underline{T}(\underline{1}) = 1$, and positive, $\underline{T}(\underline{f}^* \times \underline{f}) \geq 0$ for all $\underline{f} \in \underline{\mathcal{S}}$. For $\underline{S} \in \underline{\mathcal{S}}_0'$, the extension problem is to find a symmetric, positive extension for $\underline{S}$. In this paper all states will be symmetric without further comment.

To define the order-bounded topology for $\underline{\mathcal{S}}$ introduce Hilbert seminorms by

$$h_n^{(N)}(f_n)^2 := \sum_{\substack{|\alpha_i| \leq N_2 \\ 1 \leq i \leq n}} \int dx_1 \cdots dx_n \times \prod_{j=1}^{n} (1 + \|x_j\|^2)^{N_1} |D_1^{\alpha_1} \cdots D_n^{\alpha_n} f_n(x_1, \ldots, x_n)|^2,$$

wherein $\|x\|^2 = x_0^2 + \mathbf{x}^2$, $D_k^\alpha = \partial^{|\alpha|} / \partial x_{k,0}^{\alpha_0} \partial x_{k,1}^{\alpha_1} \cdots \partial x_{k,3}^{\alpha_3}$, $|\alpha| = \alpha_0 + \cdots + \alpha_3$; and according on $\underline{\mathcal{S}}$ with

$$q_n^{(N)}(f_n)^2 := C_n^2 h_n^{(N)}(f_n)^2 \text{ by } q^{(N)}(\underline{f})^2 := \sum_{n=0}^{\infty} q_n^{(N)}(f_n)^2,$$

with $\{C_n\}$ a sequence of positive numbers and $N=(N_1,N_2)$ an ordered pair of nonnegative integers.

Definition 2.1: $\underline{T}\in\underline{\mathcal{S}}'$ is order-bounded if $|\underline{T}(\underline{f})| \le q^{(N)}(\underline{f})$, $\underline{f}\in\underline{\mathcal{S}}$, for some seminorm $q^{(N)}$.

The result of this section is then

Theorem 2.2: Let $\underline{S}\in\underline{\mathcal{S}}'_0$ be symmetric, normalized, and order-bounded. Then there exists a (nonunique) symmetric, positive extension to $\underline{\mathcal{S}}$ with the same order bound.

Remark: If $N_1=0$, a Euclidean invariant extension is obtained by averaging over the Euclidean group provided $\underline{S}$ is Euclidean invariant. In this way ext$\underline{S}$ satisfies the extended Eucliean axioms (Ref. 3 Definition 2.1).

The proof generalizes a construction due to Yngvason.[7] Denote real functions and functionals by the subscript R and consider the Gaussian process ϕ_0 on $\mathcal{S}'_R$ with mean zero and covariance

$$\langle\phi_0(f_1)\phi_0(f_2)\rangle := \delta S_2^0(f_1\times f_2) = \sum_{|\alpha|\le N_2}\int dx(1+\|x\|^2)^{N_1}D^\alpha f_1(x)\,D^\alpha f_2(x).$$

The nonvanishing moments are even and define $\delta S_{2n}^0 \in \mathcal{S}'_{2n}$, $n=1,2,\cdots$. Finally define $\delta\underline{S}\in\underline{\mathcal{S}}'$ by the sequence

$$\delta\underline{S} := \{\lambda_0,0,\lambda_2C_1^2\,\delta S_2^0,0,\ldots,0,\lambda_{2n}C_n^2\,\delta S_{2n}^0,0,\cdots\},$$

in which $\{\lambda_{2n}\}$ is a sequence of positive numbers.

Lemma 2.3: For every seminorm of the form $q^{(N)}$ there exists $\delta\underline{S}\in\underline{\mathcal{S}}'$ such that

(a) each δS_{2n}^0 is symmetric and $\mathrm{supp}\,\delta S_{2n}^0\subset\mathbb{R}^{8n}/\mathcal{E}_{2n}$,

(b) $\delta S_{2n+1}^0=0$, $n=0,1,2,\cdots$,

(c) $\delta\underline{S}(\underline{f}^*\times\underline{f})\ge q^{(N)}(\underline{f})^2\ \forall\,\underline{f}\in\underline{\mathcal{S}}$.

Proof of Theorem 2.2: We use Lemma 2.3 to prove the extension theorem. As $S_0=1$ the choice $C_0=1$ will be made in $q^{(N)}$. This means that

$$\begin{aligned} q^{(N)}(\underline{f}^*\times\underline{g}) &\le \sum_{i,j=0}^{\infty} q_{i+j}^{(N)}(f_i^*\times g_j)\\ &\le\Big(\sum_{i=0}^{\infty} d_i\,h_i^{(N)}(f_i)\Big)\Big(\sum_{j=0}^{\infty} d_j\,h_j^{(N)}(g_j)\Big)\\ &\le\Big(\sum_{i=0}^{\infty}\tilde{c}_i^2\,h_i^{(N)}(f_i)^2\Big)^{1/2}\Big(\sum_{j=0}^{\infty}\tilde{c}_j^2\,h_j^{(N)}(g_j)^2\Big)^{1/2}\\ &:=\tilde{q}^{(N)}(\underline{f})\,\tilde{q}^{(N)}(\underline{g}), \end{aligned}$$

for two suitably chosen sequences $\{d_i\,|\,c_n\le d_i\,d_j i+j\le n\}$ and $\{\tilde{c}_n\}$. It is easy to see that one may take $d_0=1$, $\tilde{c}_0^2=1+\delta>1$. Now use Lemma 2.3 to find $\delta\underline{S}$ appropriate to the $\tilde{q}^{(N)}$ seminorm where as the proof shows $\lambda_0=(1+\epsilon_0)(1+\delta)$, $\epsilon_0>0$. Consider the functional $\underline{S}'=\underline{S}-2$ and let ext$\underline{S}'$ be any symmetric $\tilde{q}^{(N)}$-continuous extension to $\underline{\mathcal{S}}$. Setting ext$\underline{S}=$ ext$\underline{S}'+\delta\underline{S}$ and $\epsilon_0=(1-\delta)/(1+\delta)$ to normalize ext$\underline{S}$ gives the desired symmetric positive extension.

Proof of Lemma 2.3: Suppose $q_0(f_0)=c_0|f_0|$; then put $\lambda_0=(1+\epsilon_0)c_0^2$ with $\epsilon_0>0$. The construction proceeds by induction with $\delta S_{2n+1}^0=0$ at each step and suitable choice of $\{\lambda_{2n}\}$ to guarantee positivity. The even order distributions are

$$\begin{aligned}\delta S_{2n} &= \lambda_{2n}C_n^2\prod_{j=1}^{2n}(1+\|X_j\|^2)^{N_1/2}\\ &\times\sum_{\substack{|\alpha_i|\le N_2\\ 1\le i\le n}}(\text{Symmetrization of }\delta(X_n-X_{n+1})\delta(X_{n-1}-X_{n+2})\\ &\times\cdots\times(X_1-X_{2n})D_1^{\alpha_n}\otimes D_2^{\alpha_{n-1}}\otimes\cdots\\ &\otimes D_n^{\alpha_1}\otimes D_{n+1}^{\alpha_1}\otimes D_{n+2}^{\alpha_2}\otimes\cdots\otimes D_{2n}^{\alpha_n})\end{aligned}$$

consisting of $2n!/2^n n!$ distinct terms. One proceeds now to catalogue the contributions to $\delta S_{2n}^0(f_n^*\times f_n)$ which is to be the dominant positive term at the nth step. For this purpose it is convenient to use some new seminorms:

$$\begin{aligned}&|\beta_{n;k}^{(N)}(f_n)|^2\\ &=C_n^2\sum_{\substack{|\alpha_i|\le N_2\\ 1\le i\le n-2k}}\sum_{\substack{|\gamma_j|\le N2\\ 1\le j\le K}}\sum_{\substack{|\delta_i|\le N_2\\ 1\le i\le N}}\int dy_1\cdots dy_k\\ &\times dz_1\cdots dz_k\,dx_1\cdots dx_{n-2k}\prod_{i=1}^{n-2k}(1+\|x_i\|^2)^{N_1}\\ &\times\prod_{j=1}^{k}(1+\|y_j\|^2)^{N_1}(1+\|z_j\|^2)^{N_1}D_1^{\alpha_1}\otimes\cdots\\ &\otimes D_{n-2k}^{\alpha_{n-2k}}\otimes D_{n-2k+1}^{\gamma_1}\otimes D_{n-2k+2}^{\gamma_1}\otimes\cdots\otimes D_{n-1}^{\gamma_k}\\ &\otimes D_n^{\gamma_k}\bar{f}_n(X_1,\ldots,X_{n-2k},y_1,\ldots,y_k)D_1^{\alpha_1}\otimes\cdots\\ &\otimes D_{n-2k}^{\alpha_{n-2k}}\otimes D_{n-2k+1}^{\delta_1}\otimes D_{n-2k+2}^{\delta_1}\otimes\cdots\otimes D_{n-1}^{\delta_k}\\ &\otimes D_n^{\delta_k}f_n(X_1,\ldots,X_{n-2k},z_1,\ldots,z_k).\end{aligned}$$

Without loss of generality it will be assumed in the calculation that follows that f is symmetric in all arguments and the notation $\lceil n/2\rceil=n/2$ for n even and $(n-1)/2$ for n odd will be used. After some computation in which k stands for the number of δ-distribution contractions setting variables equal in the set $(X_1,\ldots,X_n)$, the positive even terms are a sum of positive quantities

$$\delta S_{2n}(f_n^*\times f_n)=\lambda_{2n}\sum_{k=0}^{\lfloor n/2\rfloor}\left(\frac{n!}{[k!\,2^k(n-2k)!]}\right)^2\times(n-2k)!\,|\beta_{n;k}^{(N)}(f_n)|^2, \tag{2.1}$$

where it should be pointed out $q_n^{(N)}(f_n)^2=|\beta_{n;0}^{(N)}(f_n)|^2$ would be the only terms appearing without the symmetrization.

To carry out the induction step $n-1$ to n, isolate the highest nonvanishing component in $\underline{f}=\underline{\hat{f}}+f_n$ and assume an induction hypothesis

$$\delta\underline{S}(\underline{\hat{f}}^*\times\underline{\hat{f}})\ge(1+\epsilon_{n-1})q^{(N)}(\underline{\hat{f}})^2+\epsilon_{n-1}\left[\sum_{l=2}^{n-1}\sum_{k=1}^{\lfloor l/2\rfloor}|\beta_{l;k}^{(N)}(f_l)|^2\right] \tag{2.2}$$

and $0<\epsilon_{n-1}<\epsilon_{n-2}<\cdots<\epsilon_1<\epsilon_0$. It is necessary to estimate a positive lower bound for

$$\delta\underline{S}(\underline{f}^*\times\underline{f})=\delta\underline{S}(\underline{\hat{f}}^*\times\underline{\hat{f}})+\delta\underline{S}_{2n}(f_n^*\times f_n)+2\mathrm{Re}\,\delta\underline{S}(\underline{\hat{f}}^*\times f_n), \tag{2.3}$$

which is accomplished by means of the expression

$$\delta\underline{S}(\underline{\hat{f}}^* \times f_n) = \sum_{s=\{n/2\}}^{n-1} \delta S_{2s}(f^*_{2s-n} \times f_n)$$

wherein $\{n/2\} = n/2$ for n even and $(n+1)/2$ for n odd. Repeated use of the Cauchy–Schwarz inequality produces the estimate

$$\begin{aligned}
&|\delta\underline{S}(\underline{\hat{f}}^* \times f_n)| \\
&\quad \leq \sum_{s=\{n/2\}}^{n-1} \frac{\lambda_{2s} C_s^2}{(C_{2s-n} C_n)} \\
&\quad \times \sum_{k=0}^{[(2s-n)/2]} \left(\frac{n!\,(2s-n)!}{(k!\, 2^{n-s+2k}(n-s+k)!\,(2s-n-2k)!} \right) \\
&\quad \times |\beta^{(N)}_{2s-n;k}(f_{2s-n})| \; |\beta^{(N)}_{n;n-s+k}(f_n)| . \qquad (2.4)
\end{aligned}$$

Substituting (2.1), (2.2), and (2.4) into (2.3) leads to the induction lower bound (2.3) in which n replaces $n-1$ provided the λ_{2n} are chosen to increase rather rapidly with n. By distributing the crossterms evenly among the positive terms, sufficient conditions on the λ_{2n} are given by $0 < \epsilon_n < \epsilon_{n-1}$ and λ_{2n} larger than any of the following:

$$(1+\epsilon_n)/n!, \quad (k!)^2 2^{2k}(n-2k)!\,\epsilon_n/(n!)^2 \text{ for } 1 \leq k \leq [n/2],$$

and for n even

$$\begin{aligned}
&[(n/2+k-r)!\,]^2 2^{n+2k-2r}(2r-2k)!\,\epsilon_n/(n!)^2 \\
&\quad + \delta_{n/2+k-r,0} \lambda^2_{2r+n} c^4_{r+n/2} n (2r!)^2/[c^2_{2r} c^2_n (k!)^2 2^{2k+1} \\
&\quad \times (2r-2k)!\,(\epsilon_{n-1} - \epsilon_n)]
\end{aligned}$$

for each $0 \leq r \leq n/2 - 1$ and $0 \leq k \leq r$, while for n odd there is a similar expression with $n-1$ replacing n and $2r+1$ for $2r$.

Remark: The combinatorial estimates above illustrate that λ_{2n} increases roughly as $\exp(\lambda_0^{2n})$ and this rapid growth dominates any increase from c_n. This is also the reason why this method yields easily positive rather than strongly positive extensions.

In order to cast Theorem 2.2 in a slightly different form, we recall the notion of strong positivity. If $P: \mathbb{R}^n \to \mathbb{R}_+$ is a polynomial and $\{f_1, f_2, \ldots, f_n\} \subset \mathcal{S}_R$, a positive polynomial on $\mathcal{S}'_R$ is of the form $P(\omega) := P(\langle\omega, f_1\rangle, \ldots, \langle\omega, f_n\rangle) \geq 0$ for all $\omega \in \mathcal{S}'_R$. A state $\underline{T}$ is called strongly positive if $T(P) \geq 0$ for every positive polynomial P. By $\delta\underline{S}^{(\alpha)}$ denote the linear functional on $\underline{\mathcal{S}}$ whose components are $\delta S^{(\alpha)}_{2n} = \alpha^{2n} \delta S^0_{2n}$, α real, with odd components zero; then $\{\lambda_{2n}\}$ may be chosen so that $\delta\underline{S}$ in Lemma 2.3 is strongly positive. This allows

$$\delta\underline{S} = \int_{-\infty}^{\infty} d\rho(\alpha)\, \delta\underline{S}^{(\alpha)},$$

where ρ is a measure solving a Hamburger moment problem for the moments $\rho_{2n+1} = 0$, $\rho_{2n} = \lambda_{2n} c_n^2$, $n = 0, 1, 2, \cdots$, and leads to bounds for the extensions in Theorem 2.2

$$0 \leq \text{ext}\underline{S}(\underline{f}^* \times f) \leq 2\delta\underline{S}(\underline{f}^* \times \underline{f}).$$

In the next section we give conditions on $\delta\underline{S}$ which imply that ext$\underline{S}$ is also strongly positive.

3. STRONG POSITIVITY AND SELF-ADJOINT REPRESENTATIONS

In the ensuing discussion the polynomial algebra with complex coefficients over $\mathcal{S}_R$ will be written as $\mathcal{P}(\mathcal{S}_R)$. Given $\omega \in \mathcal{S}'_R$, $\otimes^n\omega$ stands for an element of $\mathcal{S}'_{n,R}$ obtained by continuous extension from evaluation on elements in $\otimes^n_\pi \mathcal{S}$, where $\underline{f}(\omega) := \sum_{n=0}^{\infty} \langle \otimes^n \omega, f_n \rangle$ with $\otimes^0\omega := 1$. It is shown in Ref. 3 and independently in Ref. 5 that for each strongly positive state $\underline{T}$ on $\mathcal{P}(\mathcal{S}_R)$ there exists at least one probability measure μ on the σ-algebra $\mathcal{B}$ generated by the Borel cylinder sets in $\mathcal{S}'_R$ such that

$$\underline{T}(\underline{f}) = \int_{\mathcal{S}'_R} \underline{f}(\omega) d\mu(\omega), \quad \underline{f} \in \underline{\mathcal{S}}.$$

In general μ is not uniquely determined by $\underline{T}$. Among such measures two particular types seems to play a central role.

Definition 3.1: A measure μ is called maximal if $\mathcal{P}(\mathcal{S}_R)$ is dense in $\mathcal{L}_p(\mathcal{S}'_R, \mathcal{B}, \mu)$ for all $1 \leq p < \infty$. When this is true only for $1 \leq p \leq 2$, μ is called extremal.

It should be pointed out that this differs from definition 4.2 in Ref. 3. The alteration is needed as Lemma 3.4 in Ref. 3 contains a gap. After showing polynomial density for the one-dimensional cylinder set measures ν_f, $f \in \mathcal{S}_R$, in $\mathcal{L}_p(\mathbb{R}, d\nu_f)$ for $1 \leq p \leq 2$, we then inadvertantly assumed they in fact formed a basis and then used duality to deal with the remaining $2 < p < \infty$. Whether or not N-extremal solutions for ν_f have polynomials with the basis property for $p \neq 2$ is not part of the Riesz–Nevanlinna theory and appears to be an open question. For this the condition

$$\sup_{\|F\|_p = 1} \Big\| \sum_{n=0}^{N} \omega_n \langle F, \omega_n \rangle \Big\|_p \leq M < \infty, \quad n = 0, 1, 2, \cdots,$$

is necessary and sufficient while the stronger $\sum_{n=0}^{\infty} \|\omega_n\|_p \|\omega_n\|_{p/p-1} < \infty$ is sufficient. $\{\omega_n\}$ are orthonormalized polynomials for ν_f. After these remarks let us recall Theorems 3.5, 4.3 of Ref. 3:

Theorem 3.2: Suppose each one dimensional cylinder set measure ν_f, $f \in \mathcal{S}_R$ is maximal. Then μ is unique and maximal.

We shall shortly show the uniqueness of maximal measures. In fact the class of maximal measures is large.

Proposition 3.3: Suppose $\underline{T}$ is a symmetric, strongly positive state on $\underline{\mathcal{S}}$ for which $\sum_{p=1}^{\infty} \underline{T}(f^{2p})^{-1/2p} = \infty$ for each $f \in \mathcal{S}_R$. Then the measure for $\underline{T}$ is maximal.

Proof: Notice that if μ is a measure for $\underline{T}$ each $\underline{f} \in \underline{\mathcal{S}}$ defines a multiplication operator on $\mathcal{L}_2(\mu)$ and $\underline{T}(f^{2p})^{1/2p} = \|f\|_{2p}$. By Hölder's inequality this is a nondecreasing function of p so the Carleman quasianalyticity condition satisfies

$$\sum_{p=1}^{\infty} (\|f\|_{2p})^{-1} = \infty \text{ if and only if } \sum_{p=1}^{\infty} (\|f\|_{2^K p})^{-1} = \infty$$

for $K = 0, 1, 2, \cdots$. Suppose ν_f is the unique solution to the one-dimensional moment problem associated with f; then polynomials in f are dense in $\mathcal{L}_p(\mathbb{R}, d\nu_f)$ for $1 \leq p \leq 2$ (Ref. 3, Lemma 3.4). Pick some $2 < q < \infty$ and

$F \neq 0$ in $\mathcal{L}_{q/q-1}$ for which $\int F(t)t^n d\nu_f(t) = 0$ when $n = 0, 1, 2, \cdots$. Now $\tilde{F}(s) = \int d\nu_f(t) e^{ist} F(t)$ is C^∞ and $D^n\tilde{F}(0) = 0$. The derivatives also obey estimates

$$\| D^n\tilde{F}(0) \|_\infty \leq \| f^n \|_q \| F \|_{q/q-1}, \quad n = 0, 1, 2, \cdots,$$

and again by Hölder's inequality $\| f^n \|_q^2 \leq \| f^{n-1} \|_q \| f^{n+1} \|_q$. This means $\tilde{F}$ is quasianalytic and identically zero by the Denjoy–Carleman theorem.[9] By choosing K so that $2^K > q$, F is identically zero and ν_f is maximal. The result now follows from Theorem 3.2.

Remark: Maximality for Euclidean measures associated with Schwinger states removes possible pathologies in the Euclidean field theory. For the weakly coupled $P(\phi)_2$ Euclidean quantum field theory and the even $P(\phi)_2$ models with half-Dirichlet boundary conditions and a lower mass gap, Fröhlich[10] has shown the stronger analytic vector condition $\sum_{p=0}^\infty \lambda^p \underline{T}(f^p)/p! < \infty$ in the case of both the nonsharp time and time zero Euclidean field. Hence these theories have maximal measures. For an application of maximality see Ref. 11.

Associated with $\underline{T}$ is a commutative, cyclic $*$-representation of $\underline{S}$ by an algebra of unbounded operators on a separable Hilbert space H_E. If μ is any measure for T, this representation is unitarily equivalent to $\{H_E, \phi, 1\}$, where $H_E \cong \overline{P(S_R)} \subset \mathcal{L}_2(\mu)$ and

$$\phi(\underline{f})\psi(\omega) = \underline{f}(\omega)\psi(\omega) \quad \forall \underline{f}, \psi \in \underline{S}.$$

We follow Powers[8] and denote the closure of this representation by $\tilde{\phi}$ and the adjoint by ϕ^*. For notation on the Euclidean field, see (Ref. 3, Sec. 4). A linear manifold $D \subset H_E$ is a core for the representation if $\widetilde{\phi \upharpoonright_D} = \tilde{\phi}$ and ϕ is essentially self-adjoint if for the domains $D(\tilde{\phi}) = D(\phi^*)$. For each μ there is an essentially self-adjoint representation ϕ_μ which extends ϕ in the sense of an extension out of H_E to $\mathcal{L}_2(\mu)$. $\phi_\mu(\underline{f})$ is multiplication by $\underline{f}$ on the domain

$$D(\phi_\mu) = \{\psi \in \mathcal{L}_2(\mu) \mid \underline{f}\psi \in \mathcal{L}_2(\mu) \quad \forall \underline{f} \in \underline{S}\}$$

and the closure $\phi_\mu(\underline{f})$ is the corresponding maximal multiplication operator on $\mathcal{L}_2(\mu)$. Generally, different measures μ lead to different extensions ϕ_μ. Moreover, ϕ_μ is standard in the sense of Powers[8] for when $\underline{f} \subset S_R$, $\overline{\phi_\mu(\underline{f})}$ is self-adjoint and the spectral projections commute. These features of the representation are related to properties of μ in the following way.

Proposition 3.4: Suppose $\underline{T}$ is a strongly positive state on $\underline{S}$ with cyclic representation $\{H_T, \pi_T, \Omega_T\}$. If $\tilde{\pi}_T$ is self-adjoint, then the measure associated with $\underline{T}$ is unique and extremal.

Proof: Powers (Ref. 8, Theorem 7.3) has shown that for $\underline{f} \in S_R$, each $\overline{\pi_T(\underline{f})}$ is self-adjoint and their spectral projections, say $E_{\underline{f}}$, commute. For linearly independent $\{f_1, \ldots, f_n\} \subset S_R$ define Borel measures on $\mathbb{R}^n$ by

$$d\nu_{f_1,\ldots,f_n}(t_1, \ldots, t_n) := d(E_{f_1}(t_1) \cdots E_{f_n}(t_n)\Omega_T, \Omega_T).$$

One readily verifies these define a consistent set of cylinder set measures satisfying the continuity property (see Sec. 3 of Ref. 3) and so determine a unique probability measure μ on β. If $A = A_1 \times A_2 \times \cdots \times A_n$ is a product of Borel sets on $\mathbb{R}$, then

$$\mu(A) = (E_{f_1}(A_1) \cdots E_{f_n}(A_n)\Omega_T, \Omega_T),$$

where by the spectral theorem $E_{f_1}(A_1) \cdots E_{f_n}(A_n)\Omega_T \in H_T$. This means there exists a sequence $\{P_K\} \subset P(S_R)$ for which

$$\| \chi_A - P_K \|_{\mathcal{L}_2(\mu)} = \| E_{f_1}(A_1) \cdots E_{f_n}(A_n)\Omega_T - P_K\Omega_T \| \to 0$$

as $K \to \infty$. So μ is extremal. Uniqueness for μ follows from that for the $\nu_{f_1 \ldots f_n}$, which is a well-known proposition for the n-dimensional moment problem.[12]

Corollary: For each $f \in S_R$ suppose polynomials are dense in $\mathcal{L}_{2+\epsilon}(\mathbb{R}, d\nu_f)$ for some $\epsilon > 0$. Then $\tilde{\pi}_T$ is self-adjoint and standard. The resulting measure μ for $\underline{T}$ is unique and $P(S_R)$ is dense in $\mathcal{L}_2(\mu)$, $1 \leq p < 2 + \epsilon$.

Proof: Again take a Borel set $A = A_1 \times A_2 \times \cdots \times A_n$ and choose a sequence of polynomials $P_{\alpha_K}(f_K)$, $K = 1, 2, \ldots, n$, such that

$$\| \chi_{A_K} - P_{\alpha_K} \|_{2+\epsilon} < \epsilon / (n \| P_{\alpha_1} \cdots P_{\alpha_{K-1}} \|_{(2+\delta)(2+\epsilon)/\epsilon}), \quad 0 < \delta < \epsilon.$$

Then

$$\| \chi_A - P_{\alpha_1}(f_1) \cdots P_{\alpha_n}(f_n) \|_{2+\delta} \leq \sum_{K=1}^n \| P_{\alpha_1}(f_1) \cdots P_{\alpha_{K-1}}(f_{K-1})(\chi_{A_K} - P_{\alpha_K}(f_K)) \|_{2+\delta} < \epsilon,$$

which is the claimed density for polynomials. Finally, observe each $\pi_T(\underline{f}) \in \mathcal{L}_p(\mu)$ for all $1 \leq p < \infty$ and hence for $p = 2(2+\delta)/\delta$. This means $P(S_R)$ is a core for the maximal multiplication operator by $\underline{f}$ on $\mathcal{L}_2(\mu)$ so $\pi_T(\underline{f})$ is essentially self-adjoint.

Remark: We see in particular that maximal measures lead to self-adjoint standard representations of $\underline{S}$ with the uniqueness of the representing measure for $\underline{T}$. A converse to Proposition 3.4 is well known for the one-dimensional case but is open in more dimensions. The following corrects the Corollary to Theorem 4.3 in Ref. 3 and may be proved as above.

Proposition 3.5: Suppose a strongly positive, symmetric state on $\underline{S}$ determines a unique measure μ. Then $P(S_R)$ is dense in $\mathcal{L}_p(\mu)$ for $1 \leq p < 2$.

Uniqueness for μ implies the same for ν_f and hence polynomial density in $\mathcal{L}_2(\mathbb{R}, d\nu_f)$.

Finally, let us consider the case in which one state is dominated by another. The extension Theorem 2.2 suggests a situation for which this might arise.

Theorem 3.6: Let $\underline{S}$ be a symmetric, positive state and $\underline{T}$ a symmetric, strongly positive state such that $\underline{S}(\underline{f}^* \times \underline{f}) \leq \underline{T}(\underline{f}^* \times \underline{f})$, $\underline{f} \in \underline{S}$. Then, if $\{H_T, \pi_T, \Omega_T\}$ is essentially self-adjoint, $\underline{S}$ is strongly positive and $\{H_s, \pi_s, \Omega_s\}$ essentially self-adjoint.

Proof: Consider the sesquilinear form $\underline{S}(\underline{f}^* \times \underline{g}) = (\psi_s(\underline{f}), \psi_s(\underline{g}))$ for which $|\underline{S}(\underline{f}^* \times \underline{g})| \leq \|\psi_T(\underline{f})\| \, \|\psi_T(\underline{g})\|$. We follow the notation in Ref. 3, p. 44, wherein $\psi_{s,T}$

are respectively cosets in $\mathcal{S}/N_{\underline{S},\underline{T}}$ corresponding to the states $\underline{S}, \underline{T}$. There is a unique self-adjoint operator $\underline{B}$ which is positive, of norm, one, and such that $(\psi_s(\underline{f}), \psi_s(\underline{g})) = (\psi_T(\underline{f}),\ B\psi_T(\underline{g}))$. From $(\psi_T(\underline{f}), B\pi_T(\underline{g})\psi_T(\underline{h})) = (\pi_T(\underline{g})^*\psi_T(\underline{f}), B\psi_T(\underline{g}))$, B commutes weakly with the representation π_T hence when $\underline{f} \in \mathcal{S}_R$ both B and $\sqrt{B}$ commute strongly with $\overline{\pi_T(\underline{f})}$. This means when $\underline{f}(\omega) \geq 0$, $\omega \in \mathcal{S}_R'$,

$$\underline{S}(\underline{f}) = (\Omega_T, B\pi_T(\underline{f})\Omega_T) = (\sqrt{B}\ \Omega_T, \overline{\pi_T(\underline{f})}\ \sqrt{B}\ \Omega_T) \geq 0.$$

Next suppose $\psi_s \in \mathrm{Ker}(\pi_s(\underline{f})^* + i)$. Then, for all $\underline{g} \in \underline{\mathcal{S}}$,

$$\begin{aligned} 0 &= (\psi_s, [\pi_s(\underline{f}) - i]\psi_s(\underline{g})) \\ &= \lim_n (\psi_T(\underline{f}_n),\ B[\pi_T(\underline{f}) - i]\psi_T(\underline{g})) \end{aligned}$$

for some sequence $\psi_s(\underline{f}_n) \xrightarrow{s} \psi_s$. From $\| \sqrt{B}\ \psi_T(\underline{f}_n - \underline{f}_m) \|^2 = \| \psi_s(\underline{f}_n - \underline{f}_m) \|^2$ one has $\psi_T \in \mathcal{H}_T$ with $B\psi_T(\underline{f}_n) \xrightarrow{T} \sqrt{B}\psi_T$. This implies $\sqrt{B}\psi_T = 0$; hence, $\psi_s \in \mathrm{Ker}(\pi_s(\underline{f})^* \pm i)$ and $\pi_s(\underline{f})$ is essentially self-adjoint. Proposition 3.4 now requires uniqueness of the measure for $\underline{S}$, one concluding easily that $\tilde{\pi}_s$ is self-adjoint and standard.

Remark: Returning to the construction at the end of Sec. 2, if $\mathrm{ext}\underline{S}$ is to be strongly positive by Theorem 3.6 it is necessary to show $\delta\underline{S}$ determines a self-adjoint representation. However any lack of uniqueness in the choice of ρ shows up in nonuniqueness for the measure representing $\delta\underline{S}$. Certainly by Hamburger's criteria the one-dimensional measures for such $\delta\underline{S}$ are never unique.

4. TWO OPEN QUESTIONS

Our discussion in Sec. 3 suggests two problems of interest for the Schwinger functional moment problem.

(1) For the one-dimensional moment problem on $\mathbb{R}$ is there a characterization of solutions which are maximal?

The class of maximal solutions is certainly large as indicated by Proposition 3.3 and most likely a description of these must use more than the moments alone.

(2) Let μ be uniquely determined by a symmetric, strongly positive state $\underline{T}$ on $\mathcal{S}$. Then is $\{\mathcal{H}_T, \tilde{\pi}_T, \Omega_T\}$ self-adjoint? The heart of the matter here is whether or not $\tilde{\pi}_T$ has self-adjoint extensions analogously to the case of a single symmetric operator. If such extensions exist, then uniqueness of μ implies that there is at most one, μ is extremal, and the extension is $\{\mathcal{H}_E, \phi, \mathbf{1}\}$. Now one needs $\tilde{\pi}_T \cong \phi$ for which a theoretical criterion exists which is not easy to implement in practice.

Proposition (*Singer, Powers*[8]): Let $P_T : \mathcal{L}_2(\mu) \to \mathcal{H}_T$ be orthogonal projection. Then π_T is essentially self-adjoint if and only if $P_T(\overline{\phi(\underline{f})} \pm i)^{-1}\Omega_T \in \mathcal{D}(\tilde{\pi}_T)$ for all $\underline{f} \in \mathcal{S}_R$.

The relation between the weak commutant of π_T and extensions beyond $\mathcal{H}_T$ have been examined in the second paper of Ref. 5.

*On leave from Indiana University, Bloomington, Indiana 47401.

[1]E. Nelson, J. Func. Anal. 12, 97—112 (1973).
[2]B. Simon, Helv. Phys. Acta 46, 686—96 (1973).
[3]J. L. Challifour and S. P. Slinker, Commun. Math. Phys. 43, 41—58 (1975), Paper I of a series (see Ref. 11).
[4]K. Osterwalder and R. Schrader, Commun. Math. Phys. 31, 83—112 (1973); 42, 281—305 (1975).
[5]H. J. Borchers and J. Yngvason, Commun. Math. Phys. 43, 255—71 (1975); 47, 197—213 (1976).
[6]M. Dubois-Violette, "A Generalization of the Classical Problem on *-Algebras with Applications to Relativistic Quantum Theory": I. Commun. Math. Phys. 43, 225—54 (1975); II, preprint 1975.
[7]J. Yngvason, Commun. Math. Phys. 34, 315—33 (1973).
[8]R. T. Powers, Commun. Math. Phys. 21, 85—124 (1971).
[9]W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 1966), 1st ed.
[10]J. Fröhlich, Helv. Phys. Acta 47, 265—306 (1974).
[11]J. L. Challifour, J. Math. Phys. 17, 1889—92 (1976), Paper II of a series (see Ref. 3).
[12]A. Devinatz, Trans. Am. Math. Soc. 74, 56—77 (1953), Errata, p. 536.

On the evolution equations for Killing fields

Bartolomé Coll

Laboratoire de Physique Mathématique, Collège de France, Paris, France
(Received 29 November 1976)

The problem of finding necessary and sufficient conditions on the Cauchy data for Einstein equations which insure the existence of Killing fields in a neighborhood of an initial hypersurface has been considered recently by Berezdivin, Coll, and Moncrief. Nevertheless, it can be shown that the evolution equations obtained in all these cases are of nonstrictly hyperbolic type, and, thus, the Cauchy data must belong to a special class of functions. We prove here that, for the vacuum and Einstein–Maxwell space–times and in a coordinate independent way, one can always choose, as evolution equations for the Killing fields, a strictly hyperbolic system: The above theorems can be thus extended to *all* Cauchy data for which the Einstein evolution problem has been proved to be well set.

1. INTRODUCTION

In connection with several topics in general relativity (i.e., evolution, matching problems, propagation of discontinuities, stability, etc.) it is interesting to solve the following problem: Σ being a nonnull hypersurface in space–time and C_Σ being a set of Cauchy data on Σ for the Einstein equations, how do we obtain the conditions on C_Σ which insure the existence of Killing fields in the neighborhood of Σ?

This problem has been considered recently by Berezdivin, Moncrief, and Coll. Berezdivin[1] (vacuum case, no lightlike isometries) starts from the analyticity of the Cauchy data and of the coordinate system, and gives no explicit evolution equations for the Killing fields. The methods used independently by Moncrief[2] (vacuum case) and Coll[3–5] (vacuum, perfect fluid, Einstein–Maxwell cases) are essentially the same: We work in the Gauss coordinate gauge and we obtain the wanted Killing constraints on C_Σ from an evolution system and under suitable differentiability conditions. In an earlier paper[6] I have given a coordinate independent formulation corresponding to a slightly different evolution system.

The study of all the above evolution systems for the Killing fields reveals that their associated differential operators are of nonstrictly hyperbolic type. This situation is unsatisfactory for two reasons. On one hand, this fact means that the results obtained from these evolution systems are only valid when applied to Cauchy data C_Σ belonging to a particular class of functions (Gevrey class; see Ref. 7). On the other hand, it is generally assumed that, when we are dealing with a differential system corresponding to a certain physical situation, their characteristic manifolds represent the different waves that can propagate in the medium; nevertheless, in the present case, the characteristic manifolds associated with the above evolution systems can not, in any way, be interpreted as physical waves. Thus, it seems interesting to analyze when and why the above "anomalies" appear and, if possible, how to eliminate them.

In the present article, after such an analysis, we prove, for vacuum and Einstein–Maxwell space–times and in a coordinate independent way, that it is always possible to choose, as evolution equations for Killing fields, a strictly hyperbolic system. This new evolution system is exempt from the above two anomalous features.

In Sec. 2 we shall select the basic equations and construct the evolution system. Some common arguments on the initial data for the two cases considered below are explained in Sec. 3. Sections 4 and 5 are devoted to proving the strictly hyperbolic character of the evolution system for the vacuum and the Einstein–Maxwell space–times, respectively. Both sections conclude with the more general version (Theorems 1 and 2) that can be given for the results obtained in the previous papers (Ref. 1–5) from the point of view of the differentiability class of the Cauchy data and of the choice of local charts.

2. EVOLUTION SYSTEM

(a) Let $(V_{n+1},\hat{g})$, $n>1$, be a Lorentzian manifold. In what follows we use a caret $\hat{}$ for tensors over V_{n+1} to avoid confusion with the tensors over the submanifold Σ introduced below and, when a distinction between covariant and contravariant tensors will seem convenient, we shall use the superscript $*$ for the latter. On the other hand, all the expressions are, in general, given in their covariant form, and it is to be understood that all the operators used below act over the covariant form of the tensor fields.

Let us consider in $(V_{n+1},\hat{g})$ an arbitrary vector field $\hat{s}^*$ and denote by $\hat{L}$ the Lie derivative of the metric tensor $\hat{g}$ with respect to $\hat{s}^*$: $\hat{L}\equiv\pounds(\hat{s}^*)\hat{g}$. The classical expression for the Lie derivative of the Ricci tensor $\hat{R}$ in terms of $\hat{L}$ may be written

$$2\pounds(\hat{s}^*)\hat{R}=\Delta\hat{L}-\pounds(\delta\hat{L})\hat{g}-\nabla d\,\mathrm{tr}\hat{L}, \tag{1}$$

where Δ is the Lichnerowicz laplacian for arbitrary tensors,[8] d, ∇, and δ are respectively the exterior differentiation, the covariant derivative, and the divergence operator (up to a sign), and "tr" means "trace of."

The energy tensor $\hat{T}$ of V_{n+1} is related to the Ricci tensor $\hat{R}$ by the Einstein equations

$$\hat{R}=\hat{T}-\tfrac{1}{2}\mathrm{tr}\hat{T}\cdot\hat{g}. \tag{2}$$

Taking the Lie derivative of (2), we obtain

$$\pounds(\hat{s}^*)\hat{R}=\pounds(\hat{s}^*)\hat{T}+\tfrac{1}{2}\mathrm{tr}\{\hat{L}\times\hat{T}-\pounds(\hat{s}^*)\hat{T}\}\hat{g}-\tfrac{1}{2}\mathrm{tr}\hat{T}\cdot\hat{L} \tag{3}$$

because, for any second rank tensor $\hat{Q}$, we have the

relation

$$[\mathcal{L}(\hat{s}^*),\mathrm{tr}]\hat{Q} = -\mathrm{tr}(\hat{L}\times\hat{Q}), \tag{4}$$

where $\times$ is the *cross product* (contraction over the even indices of the tensorial product).

From (1) and (3) we can eliminate the term $\mathcal{L}(\hat{s}^*)\hat{R}$. It follows that

$$\Delta\hat{L} - \mathcal{L}(\delta\hat{L})\hat{g} - \nabla d\,\mathrm{tr}\hat{L} = 2\hat{M} - \mathrm{tr}\hat{T}\cdot\hat{L} + \mathrm{tr}(\hat{L}\times\hat{T})\hat{g}, \tag{5}$$

where we have written

$$\hat{M} \equiv \mathcal{L}(\hat{s}^*)\hat{T} - \tfrac{1}{2}\mathrm{tr}\mathcal{L}(\hat{s}^*)\hat{T}\cdot\hat{g}. \tag{6}$$

Equation (5) is the starting equation for the study of the propagation of Killing fields. Its complexity comes essentially from the form of the tensor $\hat{M}$ which, according to the definition (6), depends on the choice of $\hat{T}$. In what follows, we shall assume that $\hat{M}$ depends only on $\hat{L}$ and not on its partial derivatives. Such an assumption is sufficient for the two cases which we examine here.

Let us consider now Eq. (5) as a differential system of of $\binom{n+2}{2}$ equations in the $\binom{n+2}{2}$ unknowns $\hat{L}$, for a certain $\hat{T}$. This system is *degenerate* (incomplete) in the sense that its characteristic polynomial vanishes identically; to see it, it is sufficient, for example, to remark that the differential operator defined by its principal part is nothing else but the operator defined in the same way by the Einstein equations which, as it is well known, is degenerate.

We are thus led to add, to the system (5), $n+1$ supplementary conditions: The analogs, in the present case, of the $n+1$ coordinate conditions added to the Einstein equations in the study of the Cauchy problem. By combination with (5) of the $n+1$ supplementary conditions, we shall obtain a new system which, in general, will be nondegenerate. Nevertheless, from the point of view of the applicability of the general existence and unicity theorems for partial differential equations, the appropiate properties of this new system are not yet insured. For example, in the Cauchy problem for the Einstein equations, the choice of the Gauss coordinate conditions gives an evolution system which is not strictly hyperbolic due to the presence of multiple (double) characteristic manifolds, those generated by the timelike geodesics canonically defined by the coordinate system, whereas the choice of the harmonic coordinate conditions gives, as it is well known, a strictly hyperbolic system.

(b) We must therefore look for $n+1$ conditions for the system (5) playing a role similar to that played by the harmonic coordinate conditions in the case of the Einstein system. To do this, let us consider the laplacian $\Delta\hat{s}$ of the 1-form $\hat{s}$ associated by the metric $\hat{g}$ to the vector field $\hat{s}^*$. If we denote by $\hat{Q}^t$ the tensor obtained by matrix transposition of the components of the second rank tensor $\hat{Q}$, we may write $\hat{L} = \nabla\hat{s} + (\nabla\hat{s})^t$, $d\hat{s} = \nabla\hat{s} - (\nabla\hat{s})^t$ so that $\mathrm{tr}\hat{L} = -2\delta\hat{s}$ and $\hat{L} = 2\nabla\hat{s} + d\hat{s}$ and it follows that

$$\Delta\hat{s} \equiv (d\delta + \delta d)\hat{s} = d(-\tfrac{1}{2}\mathrm{tr}\hat{L}) + \delta(2\nabla\hat{s} - \hat{L}). \tag{7}$$

On the other hand, $\Delta\hat{s}$ may be also written as $\Delta\hat{s} = \delta\nabla\hat{s} + i(\hat{s}^*)\hat{R}$ and so the term $2\delta\nabla\hat{s}$ in (7) may be eliminated. One obtains

$$\Delta\hat{s} = \delta\hat{L} + \tfrac{1}{2}d\,\mathrm{tr}\hat{L} + 2i(\hat{s}^*)\hat{R}. \tag{8}$$

We now eliminate $\delta\hat{L}$ between (5) and (8); taking into account that $\mathcal{L}(d\,\mathrm{tr}\hat{L})\hat{g} = \nabla d\,\mathrm{tr}\hat{L} + (\nabla d\,\mathrm{tr}\hat{L})^t = 2\nabla d\,\mathrm{tr}\hat{L}$, it follows that

$$\Delta\hat{L} - \mathcal{L}(\Delta\hat{s} - 2i(\hat{s}^*)\hat{R})\hat{g} - 2\hat{M} + \mathrm{tr}\hat{T}\cdot\hat{L} - \mathrm{tr}(\hat{L}\times\hat{T})\hat{g} = 0. \tag{9}$$

We are thus led to take, as $n+1$ supplementary conditions, the following ones:

$$\Delta\hat{s} - 2i(\hat{s}^*)\hat{R} = 0. \tag{10}$$

Let us denote by H the class of all vector fields verifying (10). As Eq. (10) is a strictly hyperbolic system for $\hat{s}^*$, it follows that H is never empty. In addition, it is clear from (8) that all Killing fields belong to H. For the elements of H, Eq. (9) takes the form

$$\Delta\hat{L} - 2\hat{M} + \mathrm{tr}\hat{T}\cdot\hat{L} - \mathrm{tr}(\hat{L}\times\hat{T})\hat{g} = 0 \tag{11}$$

which is a strictly hyperbolic system in $\hat{L}$ under the assumption made on $\hat{M}$.

3. φ CHARACTERIZATIONS

(a) Let us consider, in a domain of $(V_{n+1},\hat{g})$, a one-parameter family of spacelike hypersurfaces with local equations $\varphi(x) = \mathrm{const}$. On each hypersurface of the family, every (covariant) tensor of $(V_{n+1},\hat{g})$ induces a unique (covariant) tensor of the same rank; in particular, the tensor g induced by $\hat{g}$ endows the hypersurface $\varphi = \mathrm{const}$ with a Riemannian structure (φ,g).

Let $\hat{n}$ be the unit normal 1-form to each hypersurface: $i(\hat{n}^*)\hat{n} \equiv \hat{g}(\hat{n}^*,\hat{n}^*) = 1$, where $i(\ )$ stands for the interior product. It is easy to see that, on each hypersurface $\varphi = \mathrm{const}$, every p-tensor $\hat{Q}$ (tensor of rank p) of $(V_{n+1},\hat{g})$ is biunivocally characterized by the following set of 2^p tensors of (φ,g): the $\binom{p}{s}$ s-tensors $(s = 0,1,\ldots,p)$ induced by the s-tensors of $(V_{n+1},\hat{g})$ obtained taking all the possible $p-s$ interior products of $\hat{n}^*$ with $\hat{Q}$. Such a set will be called the φ *characterization* of $\hat{Q}$.

Thus, the φ characterization of a vector $\hat{s}^*$ is the set $\{\sigma,s\}$ where the scalar σ is given by $\sigma \equiv i(\hat{n}^*)\hat{s}$ and s is the covector induced by $\hat{s}$. The strict elements of the φ characterization of a second rank symmetric tensor $\hat{A}$ are $\{\alpha,a,A\}$ where $\alpha \equiv i^2(\hat{n}^*)\hat{A}$ and a and A are respectively the vector and the (second rank symmetric) tensor induced by $i(\hat{n}^*)\hat{A}$ and $\hat{A}$. In particular, the φ characterizations of $\hat{n}$ and $\hat{g}$ are respectively $\{1,0\}$ and $\{1,0,g\}$.

(b) Let us consider now a timelike vector field $\hat{m}^*$ and let γ be its canonically parametrized integral curves. Every pair $[\varphi;\hat{m}^*]$ defines, in a natural manner, a class $C[\varphi,\hat{m}^*]$ of local charts of V_{n+1} in the following way: a chart (u,Ψ) belongs to $C[\varphi;\hat{m}^*]$ if it is adapted simultaneously to φ and to $\hat{m}^*$, that is to say if in a neighborhood of every hypersurface $\varphi = \mathrm{const}$ and of every curve γ one has

$$\Psi([\varphi = \mathrm{const}]\cap U) = \mathbb{R}^n\times\{\kappa\},\quad \Psi(\gamma\cap U) = \{\theta\}\times\mathbb{R},$$

where κ and θ are respectively two fixed points of $\mathbb{R}$ and $\mathbb{R}^n$. Conversely, every "physically admissible" local chart (U,Ψ) defines univocally, in the domain U, a pair $[\varphi;\hat{m}^*]$: If x^0 is the timelike coordinate function of (U,Ψ), then $\varphi \equiv x^0 = \mathrm{const}$ and $\hat{m}^* = \partial/\partial x^0$.

(c) Let $[\varphi, \hat{m}^*]$ be the pair defined by a physically admissible local chart (U,Ψ) and let $\{\mu, m\}$ be the φ characterization of $\hat{m}^*$. The three strict components of the φ characterization of the Einstein equations (2) are equivalent to

$$C_1\colon \operatorname{tr}(K\times K) - (\operatorname{tr}K)^2 + \operatorname{tr}R = -2\tau,$$

$$C_2\colon \delta[K - (\operatorname{tr}K)g] = -\iota,$$

$$E_2\colon \partial_\gamma K = \mu[R + 2S - P] - \nabla d\mu + \pounds(m^*)K,$$

where K, the extrinsic curvature of $\varphi = \text{const}$, is related to $\partial_\gamma g$ by

$$E_1\colon \partial_\gamma g = 2\mu K + \pounds(m^*)g,$$

$\{\tau, \iota, T\}$ is the (strict) φ characterization of the energy tensor $\hat{T}$, R is the Ricci tensor of g, and ∂_γ is the derivative along the integral curves γ of $\hat{m}^*$. It is understood that the operators tr, $\times$, δ, and ∇ refer now to the induced metric g. We have defined

$$S \equiv K\times K - \tfrac{1}{2}(\operatorname{tr}K)K, \quad P \equiv T - \tfrac{1}{2}(\operatorname{tr}T + \tau)g. \tag{12}$$

(d) A local chart (U,Ψ) defines a pair $[\varphi;\hat{m}^*]$, but φ characterizations are, by definition, independent of $\hat{m}^*$. Let us denote by $\{\sigma, s\}$ and $\{\lambda, l, L\}$ the φ characterization of $\hat{s}$ and $\hat{L}$, respectively, and let us consider, for a moment, a local chart such that $\hat{m}^* = \hat{s}^*$; in this chart we have obviously $\hat{L} \equiv \pounds(\hat{s}^*)\hat{g} = \partial_\gamma \hat{g} \Rightarrow L = \partial_\gamma g$ and thus, taking into account equation E_1

$$L = 2\sigma K + \pounds(\hat{s}^*)g. \tag{13}$$

Since both members of (13) relate only quantities and operations defined on $\varphi = \text{const}$, it follows that expression (13) of L is also valid when $\hat{m}^* \neq \hat{s}^*$.

On the other hand, as φ characterizations are linear mappings and $\hat{L} = \nabla\hat{s} + (\nabla\hat{s})^t$, λ and l must be linear combinations of σ, s, and their partial derivatives. Denoting by $\Phi(\cdots)$ a linear homogeneous polynomial in its arguments and by $\bar{x}$ the set of all partial derivatives tangent to $\varphi = \text{const}$ of a variable x, we may write, for arbitrary local charts,

$$\begin{aligned} \lambda &= \Phi^1(\sigma, s;\ \bar{\sigma}, \bar{s};\ \partial_\gamma\sigma,\ \partial_\gamma s), \\ l &= \Phi^2(\sigma, s;\ \bar{\sigma}, \bar{s};\ \partial_\gamma\sigma,\ \partial_\gamma s). \end{aligned} \tag{14}$$

(e) Let us consider the tensor $\hat{A} \equiv \pounds(\hat{m}^*)\hat{L}$ and let $\{\alpha, a, A\}$ be its φ characterization. In every local chart associated to a pair $\{\varphi, \hat{m}^*\}$ we have $\pounds(\hat{m}^*) \equiv \partial_\gamma$, and thus it follows immediately that

$$\begin{aligned} \alpha &= \Phi^3(\lambda, l, L;\ \partial_\gamma\lambda), \\ a &= \Phi^4(\lambda, l, L;\ \partial_\gamma l), \\ A &= \partial_\gamma L. \end{aligned} \tag{15}$$

Let us suppose now that $\hat{s}^*$ is a Killing vector; then in a local chart such that $\hat{m}^* = \hat{s}^*$ we have, from E_2, $p = 0$ where the quantity p is defined by

$$p \equiv \sigma[R + 2S - P] - \nabla d\sigma + \pounds(s^*)K. \tag{16}$$

It is interesting to relate the quantities p and A. From the expression of A given by (15), and taking into account (12) and the commutation rule

$$[\partial_\gamma,\ \pounds(s^*)] = \pounds(\partial_\gamma s^*)$$

we have

$$A = \partial_\gamma L = 2\sigma\partial_\gamma K + 2\partial_\gamma\sigma \circ K + \pounds(s^*)\partial_\gamma g + \pounds(\partial_\gamma s^*)g. \tag{17}$$

The expressions of $\partial_\gamma\sigma$ and $\partial_\gamma s$ in terms of λ and l may be obtained from (14): they take the form

$$\partial_\gamma\sigma = \Phi^5(\lambda, l, \sigma, s;\ \bar{\sigma}, \bar{s}), \quad \partial_\gamma s = \Phi^6(\lambda, l, \sigma, s;\ \bar{\sigma}, \bar{s}). \tag{18}$$

On the other hand, equation E_1 gives $\partial_\gamma g$ as a linear function of σ, s and $\bar{s}$ and thus Eq. (17) may be written

$$A = 2\sigma\partial_\gamma K + \Phi^7(\lambda, l, \sigma, s;\ \bar{\lambda}, \bar{l}, \bar{\sigma}, \bar{s};\ \bar{\bar{\sigma}}, \bar{\bar{s}}). \tag{19}$$

The terms $\bar{\bar{\sigma}}, \bar{\bar{s}}$ come in because in (17) there appear the spatial derivatives of the vector $\partial_\gamma s^*$. Now, eliminating between equation E_2 and (16) the φ invariant quantity $R + 2S - P$, it follows that

$$\begin{aligned} 2\sigma\partial_\gamma K &= 2\mu[p + \nabla d\sigma - \pounds(s^*)K] - 2\sigma[\nabla d\mu + \pounds(m^*)K] \\ &= 2\mu p + \Phi^8(\sigma, s;\ \bar{\sigma}, \bar{s};\ \bar{\bar{\sigma}}) \end{aligned}$$

and so Eq. (19) takes the form

$$A = 2\mu p + \Phi^9(\lambda, l, \sigma, s;\ \bar{\lambda}, \bar{l}, \bar{\sigma}, \bar{s};\ \bar{\bar{\sigma}}, \bar{\bar{s}}). \tag{20}$$

But we know that for every Killing vector field $\hat{s}^* \Leftrightarrow \{\sigma, s\}$ one has $\hat{A} = 0$, $\hat{L} = 0 \Rightarrow \lambda = l = L = A = p = 0$ and thus the polynomial term of (19) must be of the form

$$\Phi^9(\lambda, l, L;\ \bar{\lambda}, \bar{l}, \bar{L}).$$

The terms in $\hat{L}$ appear because of the terms in $\bar{\bar{\sigma}}$ and $\bar{\bar{s}}$ in (20) and relation (13). Hence, the relation between A and p may be written

$$A = \partial_\gamma L = 2\mu p + \Phi^9(\lambda, l, L;\ \bar{\lambda}, \bar{l}, \bar{L}). \tag{21}$$

(f) Finally, let us consider the φ characterization $\{\beta, b\}$ of the vector field $\hat{b} \equiv \Delta\hat{s} - 2i(\hat{s})R$; from (8), it is evident that it is of the form

$$\begin{aligned} \beta &= \Phi^{10}(\lambda, l, L;\ \bar{\lambda}, \bar{l}, \bar{L};\ \partial_\gamma\lambda, \partial_\gamma l, \partial_\gamma L), \\ b &= \Phi^{11}(\lambda, l, L;\ \bar{\lambda}, \bar{l}, \bar{L};\ \partial_\gamma\lambda, \partial_\gamma l, \partial_\gamma L). \end{aligned} \tag{22}$$

The class H of vector fields considered above is just defined by $\beta = b = 0$.

4. THE VACUUM CASE

In the vacuum case, $\hat{T} = 0 \Leftrightarrow \hat{R} = 0$, the evolution system (11) reduces to

$$\Delta\hat{L} = 0. \tag{23}$$

Let Σ be an initial hypersurface which is covered by a set of adapted local charts and let $(g,K)|_\Sigma$ be a set of initial data on Σ for the Einstein equations. Suppose that the equations $L|_\Sigma = 0$ and $p|_\Sigma = 0$, where L is given by (13) and p by (16), have a solution $\{\sigma, s\}|_\Sigma$; then, according to (14), we can associate with the pair $\{\sigma, s\}|_\Sigma$ the transversal derivatives $\partial_\gamma\sigma|_\Sigma$ and $\partial_\gamma s|_\Sigma$ such that $\lambda|_\Sigma = l|_\Sigma = 0$. Since we now have $\lambda|_\Sigma = l|_\Sigma = L|_\Sigma = 0$, $p|_\Sigma = 0$, Eq. (21) gives $A|_\Sigma = \partial_\gamma L|_\Sigma = 0$. For the class H of vector fields defined above, which reduce here to

$$\Delta\hat{s} = 0, \tag{24}$$

(22) gives $\partial_\gamma\lambda|_\Sigma = 0$ and $\partial_\gamma L|_\Sigma = 0$, and thus from (15) we have $\alpha|_\Sigma = a|_\Sigma = 0$. Since φ characterizations are isomorphisms, it follows that $L|_\Sigma = 0$ and $\partial_\gamma L|_\Sigma = 0$ and, for these initial values, Eq. (23) has the unique solution $\hat{L} = 0$ in the neighborhood of Σ: The pair $\{\sigma, s\}|_\Sigma$ determines a Killing field.

Since the systems (23) and (24) are obviously strictly hyperbolic, it follows that the above argument is valid for all Cauchy data $(g,K)|_\Sigma$ for wihch the Cauchy problem for the Einstein equations is proved to be well set. On the other hand, it is easy to see, from the same argument, that different solutions of the equations $L=p=0$ on Σ give rise to different Killing fields in its neighborhood. This result may be summarized in the following theorem.

Theorem 1: A necessary and sufficient condition which insures that the vacuum space–time determined by an initial set $\{\Sigma,(g,K)\}$ admits r linearly independent Killing fields is that the dimension of the vector space of solutions of the system

$$2\sigma K+\mathcal{L}(s^*)g=0,$$

$$\sigma[R+2S]-\nabla d\sigma+\mathcal{L}(s^*)K=0,$$

in $\{\sigma,s\}$ on Σ, be r.

5. THE EINSTEIN-MAXWELL CASE

(a) The energy tensor of a Einstein–Maxwell space–time is given by

$$\hat{T}=\tfrac{1}{4}\mathrm{tr}(\hat{F}\times\hat{F})\hat{g}-\hat{F}\times\hat{F}, \tag{25}$$

where $\hat{F}$ is the electromagnetic 2-form.

On the other hand, it is easy to see that, for an arbitrary second rank tensor $\hat{Q}$, the commutation operator $[\mathcal{L}(\hat{s}^*),\mathrm{tr}]$ is

$$[\mathcal{L}(\hat{s}^*),\mathrm{tr}]\hat{Q}=-\mathrm{tr}[\hat{L}\times\hat{Q}]$$

and that, for two second rank tensors $\hat{P}$ and $\hat{Q}$, the Lie derivative of the cross product may be written

$$\mathcal{L}(\hat{s}^*)(\hat{P}\times\hat{Q})=\mathcal{L}(\hat{s}^*)\hat{P}\times\hat{Q}+\hat{P}\times\mathcal{L}(\hat{s}^*)\hat{Q}-\hat{P}\times\hat{L}\times\hat{Q}.$$

From these expressions it follows that the Lie derivative of the energy tensor (25) takes the form

$$\mathcal{L}(\hat{s}^*)\hat{T}=\hat{\Phi}^1(\hat{L})+\tfrac{1}{2}\mathrm{tr}[\hat{F}\times\mathcal{L}(\hat{s}^*)\hat{F}]\hat{g}$$
$$-\hat{F}\times\mathcal{L}(\hat{s}^*)\hat{F}-\mathcal{L}(\hat{s}^*)\hat{F}\times\hat{F}, \tag{26}$$

where, as usual, $\hat{\Phi}(\ )$ denotes a linear homogeneous polynomial in its arguments.

Now, let us introduce in (26) the 2-form $\hat{G}$ given by

$$\hat{G}\equiv\mathcal{L}(\hat{s}^*)\hat{F}-\varkappa*\hat{F} \tag{27}$$

where $*$ is the duality operator; the value of the scalar $\varkappa$ is not important here and will be defined below. Taking into account the well-known identity

$$\hat{F}\times*\hat{F}+*\hat{F}\times\hat{F}=\tfrac{1}{2}\mathrm{tr}(\hat{F}\times*\hat{F})\hat{g}, \tag{28}$$

one finds

$$\mathcal{L}(\hat{s}^*)\hat{T}=\hat{\Phi}^1(\hat{L})-\hat{\Phi}^2(\hat{G}) \tag{29}$$

with

$$\hat{\Phi}^2(\hat{G})\equiv\hat{F}\times\hat{G}+\hat{G}\times\hat{F}-\tfrac{1}{2}\mathrm{tr}(\hat{F}\times\hat{G})\hat{g} \tag{30}$$

and thus, from the definition (6) of $\hat{M}$, it follows that

$$\hat{M}=\hat{\Phi}^3(\hat{L})-\hat{\Phi}^2(\hat{G}).$$

Hence, for vector fields $\hat{s}^*$ belonging to the class $\mathcal{H}$ defined above, the system (11) may be written

$$\Delta\hat{L}+\hat{\Phi}^4(\hat{L})-\hat{\Phi}^2(\hat{G})=0. \tag{31}$$

(b) In terms of the basic variables $\hat{g}$, $\hat{F}$ of the Einstein–Maxwell space–time, Eq. (31) contains implicitly the scalar $\varkappa$. Its definition will be taken from the following lemma.

Lemma: If an Einstein–Maxwell space–time admits a Killing field $\hat{s}^*$, then the electromagnetic field $\hat{F}$ verifies

$$\mathcal{L}(\hat{s}^*)\hat{F}=\varkappa*\hat{F} \tag{32}$$

where $\varkappa$ is such that $d\varkappa\,/\!/\,\hat{l}$ if $\hat{F}$ is null with fundamental null vector $\hat{l}$, $\hat{l}\wedge d\hat{l}=0$, and $\varkappa=\mathrm{const}$ otherwise.

For a nonull $\hat{F}$, this result was partially obtained by Wooley[9]; Ray and Thompson[10] and Michalski and Wainwright[11] have given a proof based on the Rainich theory and Yaremowicz[12] has obtained it as a particular case of homothetic motions. McIntosh[13] applies this result to obtain some general properties of certain Einstein–Maxwell space–times. The extension to the null case has been given independently by Yaremowicz[12] and Coll.[14] Nevertheless, the shortest proof, which is valid for arbitrary $\hat{F}$ is the following one: For a Killing field, Eq. (31), which is nothing else but $\mathcal{L}(\hat{s}^*)\hat{R}=0$, gives $\hat{\Phi}^2(\hat{G})=0$ and from (28) and (30) it is clear that $\hat{G}=\lambda*\hat{F}$ is a solution. That this is the only solution may be easily seen by writing (30) as a function of the electromagnetic components of $\hat{F}$ and $\hat{G}$ with respect to an arbitrary timelike unit vector. Then, taking into account the definition (27) of $\hat{G}$, one obtains (32). The function $\varkappa$ is obtained immediately by taking the exterior derivative of (32) and its dual under the assumption that $\hat{F}$ is a vacuum Maxwellian field, that is

$$d\hat{F}=0,\quad \delta\hat{F}=0. \tag{33}$$

(c) For an arbitrary 2-form $\hat{Q}$, the commutator $[\mathcal{L}(\hat{s}^*),*]$ is given by

$$[\mathcal{L}(\hat{s}^*),*]\hat{Q}=*(\hat{L}\times\hat{Q}-\hat{Q}\times\hat{L}-\tfrac{1}{2}\mathrm{tr}\hat{L}\cdot\hat{Q})$$

and thus, by differentiation of (27) and its dual, taking into account (33), one finds

$$d\hat{G}=-d\varkappa\wedge*\hat{F},\quad \delta\hat{G}=*(d\varkappa\wedge\hat{F})+\hat{\Phi}^5(\hat{L}), \tag{34}$$

where, as usual, $\hat{\Phi}^5(\hat{L})$ stands for terms which are linear in $\hat{L}$. Using the definition of $\varkappa$ given by the lemma, (34) reduces to

$$d\hat{G}=0, \tag{35}$$

$$\delta\hat{G}=\hat{\Phi}^5(\hat{L}). \tag{36}$$

From (35) it follows that there always exists a local 1-form $\hat{\alpha}$ such that

$$d\hat{\alpha}=\hat{G},\quad \delta\hat{\alpha}=0 \tag{37}$$

and, in terms of $\hat{\alpha}$, (36) gives

$$\Delta\hat{\alpha}=\hat{\Phi}^5(\hat{L}). \tag{38}$$

The coupled system $\{(31),(38)\}$, that we write as

$$\Delta\hat{L}+\hat{\Phi}^4(\hat{L})-\hat{\Phi}^2(d\hat{\alpha})=0,\quad \Delta\hat{\alpha}-\hat{\Phi}^5(\hat{L})=0 \tag{39}$$

is now the evolution system for the study of Killing fields on the Einstein–Maxwell space–times. It is easy to see that it is a second order *strictly hyperbolic* system with indices $s(\hat{L})=s(\hat{\alpha})=2$ for the unknowns and

indices $t(\Delta\hat{L}) = t(\Delta\hat{\alpha}) = 1$ for the equations (for a simple definition of strictly hyperbolic systems, see Ref. 15).

Taking into account the local equivalence of the systems $\{(35), (36)\}$ and $\{(37)\}$, it is clear that in order to insure the existence of Killing fields we must take $\hat{L}|_\Sigma = 0$, $\partial_\gamma \hat{L}|_\Sigma = 0$, $\hat{G}|_\Sigma = 0$ as initial data for the system $\{(39)\}$. By an argument similar to that made in the vacuum case, it is possible to show that the conditions $\hat{L}|_\Sigma = 0$, $\partial_\gamma \hat{L}|_\Sigma = 0$ are equivalent, in terms of the Cauchy data for the Einstein equations and in terms of the Σ characterization $\{\sigma, s\}$ of the Killing fields, to

$$2\sigma K + \pounds(s^*)g = 0,$$

$$\sigma[R + 2S - P] - \nabla d\sigma + \pounds(s^*)K = 0.$$

The explicit expression for the condition $\hat{G}|_\Sigma = 0$ is somewhat longer to obtain: Let us express the first element a of the φ characterization $\{a, b\}$ of $\hat{G}$, obtained by induction of the vector fields $\hat{a} \equiv i(\hat{n})\hat{G}$, $\hat{b} \equiv i(\hat{n}) * \hat{G}$, in terms of σ, s, g, K, and of the φ characterization $\{e, h\}$ of $\hat{F}$, obtained in a similar way to that of $\hat{G}$; and then, let us eliminate the terms in $\partial_\gamma e$ using the φ characterization of the Maxwell equations (33). The result is

$$a(e,h) \equiv \pounds(s^*)e - \varkappa h - \sigma[*dh + \mathrm{tr}K \cdot e - 2i(e)K] + i(d\sigma) * h.$$

By duality, we obtain $b(e,h) \equiv a(h,-e)$. Since $\hat{G}|_\Sigma = 0 \iff \{a|_\Sigma = 0,\ b|_\Sigma = 0\}$, we have the following theorem.

Theorem 2: A necessary and sufficient condition which insures that the Einstein–Maxwell space–time determined by an initial set $\{\Sigma, (g,K), (e,h)\}$ admits r linearly independent Killing fields is that the dimension of the vector space of solutions of the system

$$2\sigma K + \pounds(s)g = 0,$$

$$\sigma[R + 2S - P] - \nabla d\sigma + \pounds(s)K = 0,$$

$$\pounds(s)e = \varkappa h + \sigma[* dh + \mathrm{tr}K \cdot e - 2i(e)K] - i(d\sigma) * h,$$

$$\pounds(s)h = \varkappa e + \sigma[-* de + \mathrm{tr}K \cdot h - 2i(e)K] + i(d\sigma) * e,$$

in $\{\sigma, s\}$ on Σ, be r.

[1] R. Berezdivin, J. Math. Phys. **15**, 1963–6 (1974).
[2] V. Moncrief, J. Math. Phys. **16**, 493–8 (1975).
[3] B. Coll, C.R. Acad. Sci. Paris **280**, A393–6 (1975).
[4] B. Coll, Ann. Inst. Henri Poincaré **25**, 393–410 (1976).
[5] B. Coll, C.R. Acad. Sci. Paris **281**, A1109–12 (1975).
[6] B. Coll, C.R. Acad. Sci. Paris **282**, A247–50 (1976).
[7] J. Leray and Y. Ohya, *Colloque Analyse Fonctionnelle* 2, *Liège* (Gauthiers-Villars, Paris, 1964).
[8] A. Lichnerowicz, "Propagateurs et commutateurs en relativité générale," I.H.E.S., Pub. Math. **10** (1961).
[9] M.L. Wooley, Commun. Math. Phys. **31**, 75–81 (1973).
[10] J.R. Ray and E.L. Thompson, J. Math. Phys. **16**, 345–6 (1975).
[11] H. Michalski and J. Wainwright, J. Gen. Rel. Grav. **6**, 289–317 (1975).
[12] P.E.A. Yaremowicz, thesis, University of Waterloo, Canada (1975).
[13] C.B.G. McIntosh, preprint, Mathematics department, Monash University, Clayton, Victoria, 3168, Australia.
[14] B. Coll, C.R. Acad. Sci. Paris **280**, A1773–6 (1975).
[15] A. Licknerowicz, *Relativistic Hydrodynamics and Magnetohydrodynamics* (Benjamin, New York, 1967).

Geometric aspects of supergauge theory

D. J. R. Lloyd-Evans

45c Corsica Street, London, N5 1JT, England
(Received 29 November 1976)

It is shown that the Wess–Zumino supergauge algebra can be obtained from the theory of automorphisms of a complex structure on spacetime, and that this is related to a formal geometric scheme of quantization. The geometric interpretation requires $\hbar$ to be proportional to the square of the basic unit of length.

I. INTRODUCTION

Recently considerable interest has been shown in supergauge theories, both from the point of view of the classification of graded Lie algebras[1,2] and of their application to geometry and gravity[3–10]; however despite the fact that GLA's first appeared in differential geometry the precise status of supergeometries and supergravity remains uncertain, and it is with this question that this paper is concerned, through the theory of automorphisms of G-structures.

To date two initially distinct approaches to supergravity have been adopted; the first is a Kaluza–Klein type formalism[3–6] in which the methods of Riemannian geometry are applied to a superspace possessing a subspace of Majorana spinors, while in the second local supergauge transformations are shown to imply general covariance and hence gravity.[7–10] The first method has proved successful in the theory of unified fields, with a Weinberg–Salam type unification of gravity and electromagnetism[3] and a possible description of the Yang–Mills field through the supermetric components $g_{\mu\alpha}$, where $0 \leq \mu \leq 3$ is a space–time index and α a Majorana spinor index, while the local supergauge method gives a consistently coupled Einstein–Maxwell theory[10] and the hope of cancellations of divergences in the full quantum theory. This difference of approach mirrors a difference in approach to the ordinary theory of gravity, namely general relativity versus local Poincare gauge invariance,[11] and this paper shows how supergauge transformations arise in a third such approach and how they are related to a geometric theory of quantization.

The method used here is one of broken gauge invariance using $\mathrm{GL}(4,\mathbb{R})/\mathrm{O}(1,3)$ which was shown by the author to arise in the frame bundle description of the metric and used in a geometric theory of formal quantization.[12,13] There are two $\mathrm{GL}(4,\mathbb{R})$ groups involved in the theory of gravity, one is that of general linear basis frame transformations and the other that of nonsingular infinitesimal coordinate transformations and it is the breaking of the first of these which describes the metric field. The local Lorentz transformations constitute a subbundle of that of general linear frames, and the effect of requiring invariance of the Lagrangian under local Lorentz transformations is to relax the condition that the connection in the frame bundle be the unique metric connection and thereby to introduce torsion. Local transformations correspond to elements of the fibers isomorphous to $\mathbb{R}^4$ associated with the bundle of general affine frames with structure group $\mathbb{R}^4 \circledS \mathrm{GL}(4,\mathbb{R})$, but because of the presence of the canonical 1-form of the frame bundle they have no dynamical significance in the purely geometric theory and serve merely to introduce a choice of local coordinates whereas for an arbitrary principal fiber bundle with isomorphous structure group they would lead to an independent Yang–Mills vierbein field.

The geometric scheme of formal quantization is based on the method initiated by Segal[14] in which the basic structures required are complex and Hermitian structures on the phase space of the system to be quantized, and on Kostant's closely related theory of quantization of sections of a complex line bundle.[15,16] Such structures arise naturally in the frame bundle description of geometry and were used to outline a formal theory of quantization of space–time. A complex structure J on $\mathbb{R}^4$ is a linear endomorphism of $\mathbb{R}^4$ such that $J^2 = -1$, where 1 denotes the identity transformation, and defines a reduction of $\mathrm{GL}(4,\mathbb{R})$ to $\mathrm{GL}(2,\mathbb{C})$ where an element $(A+iB) \in \mathrm{GL}(2,\mathbb{C})$ corresponds to

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathrm{GL}(4,\mathbb{R}),$$

where A and B are 2×2 matrices and J is in canonical form. The reduction of $\mathrm{GL}(4,\mathbb{R})$ to $\mathrm{O}(1,3)$ defines a pseudo-Riemannian metric structure $g(X,Y)$, $X, Y \in \mathbb{R}^4$, and the special feature of this geometry is that this is Hermitian with respect to the complex structure J implicit in the definition of spinors belonging to the group $\mathrm{SL}(2,\mathbb{C})$. Cohomological aspects of this quantization scheme can be examined by application of the Kostant technique[16] to sections of the line bundle with structure group $\mathrm{GL}(2,\mathbb{C})/\mathrm{SL}(2,\mathbb{C})$. On its own this does not give a full quantum theory and must be supplemented by further measures, but it does lead to supergauge transformations as its automorphisms.

In Sec. II the Bose sector of the Wess–Zumino supergauge algebra[17] is shown to come from the standard theory of automorphisms of a complex structure,[18] while the subalgebra preserving the Hermitian structure consists of the Poincaré algebra. In Sec. III the use of the Segal technique for Fermi–Dirac quantization is shown to be related to a four-component Majorana spinor formulation of geometry related to local twistor theory, thereby clarifying the relationship between the Wess–Zumino supergauge algebra and twistors noted earlier by Salam and Strathdee.[19] The significance of the Wess–Zumino transformations is that they preserve the complex structure and hence the distinction between the

creation and annihilation operators of the formal geometric quantization theory, while the Volkov–Akulov subalgebra preserves also the magnitude of the commutation relations.

II. BOSE SECTOR

A G-structure is defined[18] to be a differentiable subbundle with structure group G of the bundle of general linear frames. As mentioned in the introduction the main G-structures to be considered are the pseudo-Riemannian metric structure and the almost complex structure with $G=\mathrm{SO}(1,3)$ and $\mathrm{GL}(2,\mathbb{C})$ respectively. The light-cone structure, which is often regarded as fundamental to quantization, is a $\mathrm{CO}(1,3)$-structure in this formalism, and differs from the almost complex structure by an additional phase factor corresponding to the one-parameter group of transformations whose real generator is the almost complex structure J, where $J^2=-\mathbf{1}$. A spinor structure is defined up to a sign as an $\mathrm{SL}(2,\mathbb{C})$-structure, but if this sign ambiguity is to be removed so enabling distinction between positive and negative energy spinors, the spinor structure must be defined as a subbundle of the bundle of metalinear frames[20] with group $\mathrm{ML}(2,\mathbb{C})$, where $\mathrm{ML}(2,\mathbb{C})$ is the double covering of $\mathrm{GL}(2,\mathbb{C})$ corresponding to the sign chosen in taking the square root of $\det(g)$, $g\in\mathrm{GL}(2,\mathbb{C})$ in defining spinor transformations. The close relationship between metric, spinor, conformal (or causal), and complex structures is unique to the signature of space–time, and in a Riemannian space–time there would be no possibility of a covariant Segal-type quantization since $\mathrm{O}(4)\cap\mathrm{GL}(2,\mathbb{C})=\mathrm{U}(2)$.

Automorphisms are defined for integrable G-structures, where a G-structure over the base space M is said to be integrable if every point of M has a coordinate neighborhood U with local coordinate system $(x^0,\ldots,x^3)$ such that the cross section $(\partial/\partial x^0,\ldots,\partial/\partial x^3)$ over U is a cross section of the G-structure over U. Every $\mathrm{GL}(4,\mathbb{R})$-structure is integrable, but in general integrability requires the vanishing of some tensor field; the Riemannian curvature for an $\mathrm{SO}(1,3)$-structure, the conformal curvature for a $\mathrm{CO}(1,3)$ structure, and the Nijenhuis tensor $N(X,Y)$ for a $\mathrm{GL}(2,\mathbb{C})$-structure, where

$$N(X,Y)=2\{[JX,JY]-J[JX,Y]-J[X,JY]-[X,Y]\}, \tag{2.1}$$

where $X,Y\in\mathbb{R}^4$. Since the Nijenhuis tensor is related to the torsion tensor of an almost complex connection (2.2) below, a sufficient (but not necessary) condition for integrability of a $\mathrm{GL}(2,\mathbb{C})$ structure is that there should be no torsion

$$\begin{aligned}N(X,Y)=&\,2T(X,Y)+2J(T(X,Y))\\&+2J(T(JX,Y))-2T(JX,JY),\end{aligned} \tag{2.2}$$

where $T(X,Y)$, $X,Y\in\mathbb{R}^4$, denotes the torsion tensor.

A vector field X on M is an infinitesimal automorphism of a G-structure P if it generates a one-parameter group of automorphisms of P, and hence if the Lie derivative with respect to X of the corresponding tensor field on M vanishes. Thus for the complex structure tensor J

$$\mathcal{L}_X J=0 \tag{2.3}$$

or equivalently

$$[X,JY]=J([X,Y]),\quad \forall\, Y\in D(M), \tag{2.4}$$

where $D(M)$ is the space of vector fields on M.

For the present approach automorphisms are considered in terms of their local behavior, for which there is no loss in generality in taking $M=\mathbb{R}^4$ so that any G-structure can be expressed as $\mathbb{R}^4\times G$. Any vector field X can be expressed locally as a power series

$$X=\sum_{\mu=0}^{3}\lambda^\mu\,\partial/\partial x^\mu, \tag{2.5}$$

where

$$\lambda^\mu\approx\sum_k\frac{1}{k!}\sum_{\mu_1,\ldots,\mu_k=0}^{\mu}a^\mu_{\mu_1\cdots\mu_k}x^{\mu_1}\cdots x^{\mu_k} \tag{2.6}$$

and the coefficients $a^\mu_{\mu_1\cdots\mu_k}$ are symmetric in the subscripts. X is an infinitesimal automorphism of a G-structure iff the matrix $(\partial\lambda^\mu/\partial x^\nu)$, $0\le\mu,\nu\le3$, belongs to the Lie algebra of G, and hence iff for each fixed $\mu_2\cdots\mu_k$ the matrix $a^\mu_{\mu_1\cdots\mu_k}$ belongs to the Lie algebra of G. For some groups G this is trivial in the sense that the coefficients are necessarily zero for $k>1$, but this is not always so, and a prolongation of the Lie algebra G may be defined having a graded structure $\sum_j\mathbf{G}_j$, $j=0,1,2,\cdots$, whose order is said to be the first positive integer j such that $\mathbf{G}_j=0$. The jth prolongation $\mathbf{G}_j$ is defined as the space of symmetric multilinear mappings

$$\underbrace{\mathbb{R}^4\times\cdots\times\mathbb{R}^4}_{(j+1)\text{ times}}\to\mathbb{R}^4$$

such that for each fixed $v_1,\ldots,v_j\in\mathbb{R}^4$ the linear transformation $v\in\mathbb{R}^4\mapsto t(v,v_1,\ldots,v_j)\in\mathbb{R}^4$ belongs to $\mathbf{G}$.

In the cases of Riemannian or pseudo-Riemannian structures this prolongation is trivial, with $\mathbf{G}_1=0$, and the theory of infinitesimal automorphisms of the G-structure is that of Killing vector fields (of flat space). The first nontrivial case is that of a conformal structure, i.e., $G\equiv\mathrm{CO}(1,3)$. Only the dilatation subalgebra of $\mathbf{G}$ prolongs, and $\mathbf{G}_1$ consists of the four derivations $\frac{1}{4}\sum_{i=0}^{3}t^i_{ij}$ of the scale factor $\frac{1}{4}\sum_i t^i_i$. These four elements commute with each other and generate the Lie algebra of special conformal transformations (or "accelerations"), while $\mathbf{G}_2$, can be shown to vanish.[18]

The Lie algebra of $\mathrm{GL}(2,\mathbb{C})$ regarded as a real subgroup of $\mathrm{GL}(4,\mathbb{R})$ via the mapping

$$(A+iB)\mapsto\begin{pmatrix}A & B\\ -B & A\end{pmatrix},$$

(where A,B are 2×2 real matrices) is equivalent to that of $\mathrm{CO}(1,3)$ with the addition of an element corresponding to J.

$$J\equiv\begin{pmatrix}0 & 1_2\\ -1_2 & 0\end{pmatrix},$$

where 1_2 denotes the 2×2 unit matrix. This element is

skew-symmetric, and so does not have any prolongation; hence the prolonged Lie algebra of $gl(2,\mathbb{C})$ is isomorphous to $\mathbb{R}^4 \oplus U(1) \oplus CO(1,3) \oplus \mathbb{R}^{4*} \approx U(2,2)$. The Bose sector of the Wess–Zumino GLA appears therefore in the geometric quantization scheme as the infinitesimal automorphisms of the complex structure J.

Infinitesimal coordinate transformations appear in this formalism as the elements of the prolonged Lie algebra of $GL(4,\mathbb{R})$ which is of infinite type, thus they generate automorphisms of the bundle of general linear frames.

III. FERMI SECTOR

In addition to the usual automorphisms of the complex structure considered in the previous section, it is also preserved by certain spinor transformations. The ability of the abstract quantization scheme to treat bosons and fermions in a unified manner results from the standard decomposition of the Hermitian metric $g(X, Y)$ into real and imaginary parts.

$$g(X, Y) = S(X, Y) + iA(X, Y), \tag{3.1}$$

where $X, Y \in \mathbb{R}^4$, and

$$S(X, Y) = S(Y, X) \tag{3.2}$$

$$A(X, Y) = -A(Y, X) = -S(JX, Y). \tag{3.3}$$

In principle this means that the metric can be used to define two formal quantum theories with the commutation relations

$$\exp[i\varphi(x)]\exp[i\varphi(y)] = \exp[iA(x, y)]\exp[i\varphi(y)]\exp[i\varphi(x)], \tag{3.4}$$

$$\psi(x)\psi(y) + \psi(y)\psi(x) = S(x, y)\,1, \tag{3.5}$$

where (3.4) is the Weyl form of the commutation relations avoiding the basic unboundedness problem of the Bose operators. These two quantizations were considered elsewhere[12] and their significance here lies in the relation of (3.5) to Majorana spinors.

It was shown[12] that as a result of the Cartan isotropic vector ⟷ spinor correspondence in three dimensions[21]

$$(\xi^A) \longleftrightarrow \{x^i \mid \sum_{i=1}^{3} (x^i)^2 = 0\} \tag{3.6}$$

a formal quantum theory (3.5) could be written in which the Dirac gamma matrices acted as formal quantum operators for points of 3-D hypersurfaces. The space of spinors on which the Dirac matrices act is the direct sum of left- and right-handed $\mathbb{C}^2$, where the space of two component spinors corresponds to the mapping $GL(4,\mathbb{R}) \to GL(2,\mathbb{C})$ defined by the complex structure J. Since J defines a 4-orientation, the two spaces $\mathbb{C}^2_R$ and $\mathbb{C}^2_L$ must have opposite time orientation in addition to opposite chirality. This formalism was compared with the real Dirac algebra approach to space–time of Hestenes[22] in which the physically uninterpreted quantity (i) in the Dirac equation was represented by γ_5 in the Dirac algebra, and in the present case $\mathbb{C}^4$ is the complexification of the $\mathbb{R}^4$ space of Majorana spinors with respect to γ_5.

The Majorana spinors so defined determine a fiber isomorphous to $\mathbb{R}^4$ over each point of space–time, and together these fibers constitute a spinor bundle of metalinear frames with structure group $ML(2,\mathbb{C})$, where $ML(2,\mathbb{C})$ is the two-fold covering of $GL(2,\mathbb{C})$ corresponding to the choice of sign of the square root of the determinant of elements $GL(2,\mathbb{C})$ taken for the spinor transformations. Globally the existence of the metalinear frame bundle requires the cohomology condition $H^2(M, Z_2) = 0$, where M denotes space–time.[20] The Majorana spinor bundle is the spinor analog of the tangent bundle which is the associated vector bundle with standard fiber $\mathbb{R}^4$ of the general linear frame bundle $L(M)$ with structure group $GL(4,\mathbb{R})$. Local spinor transformations

$$\psi_\alpha(x) \to \psi_\alpha(x) + \delta\psi_\alpha(x) \quad 1 \le \alpha \le 4$$

are defined vertically on the fiber $\pi^{-1}(x)$, and so preserve the complex structure J.

Supergauge theory is very closely linked to the spinor formulation of geometry, as is most evident from the two-component spinor formations of Ferrara *et al.*[23] This includes the basic anticommutator

$$\{S_\alpha, \bar{S}_{\dot\beta}\} = 2\sigma^\mu_{\alpha\dot\beta} P_\mu, \tag{3.7}$$

where P_μ is the momentum operator and $\sigma^\mu_{\alpha\dot\beta}$ the Van der Waerden symbol where the bar denotes the Hermitian conjugate, and the relation of dotted to undotted spinors is

$$\bar{S}_{\dot\alpha} = (S_\alpha)^*, \quad (\dot{S}_\alpha)^* = \bar{S}_\alpha. \tag{3.8}$$

Formula (2.7) may be inverted to give

$$P_\mu = \sigma_\mu{}^{\alpha\dot\beta} S_\alpha \bar{S}_{\dot\beta} - \tfrac{1}{2}\sigma_\mu{}^{\alpha\dot\beta}[S_\alpha, \bar{S}_{\dot\beta}]. \tag{3.9}$$

This is similar to the usual Penrose two-component spinor description (3.10) of a vector x_μ except for an extra term

$$x_\mu = \sigma_\mu{}^{\alpha\dot\beta} S_\alpha \dot{S}_{\dot\beta}. \tag{3.10}$$

Since the spinor components $S_\alpha, S_{\dot\beta}$ define a basis the commutator $[S_\alpha, \bar{S}_{\dot\beta}]$ satisfies

$$[S_\alpha, \bar{S}_{\dot\beta}] = \epsilon_{\alpha\dot\beta}, \tag{3.11}$$

where $\epsilon_{\alpha\dot\beta}$ is the Levi-Civita symbol, so that P_μ differs from x_μ by an extra component in the spin plane of the spinor. The significance of this is that whereas the commutator $[x^\mu, x_\mu]$ vanishes, $[x^\mu, P_\mu]$ does not, in accordance with quantum mechanics.

A similar formalism exists for the Majorana spinor formulation, thus inverting the anticommutator

$$\{\psi_\alpha, \psi_\beta\} = (C^{-1}\gamma^\mu)_{\alpha\beta} P_\mu, \tag{3.12}$$

where C denotes the charge conjugation matrix for the representation chosen, gives

$$P_\mu = (\gamma_\mu C)^{\alpha\beta}\{\psi_\alpha, \psi_\beta\}. \tag{3.13}$$

The Majorana spinor formulation of geometry does not seem to have been fully investigated, and it may be noted that instead of using the conversions of vector to spinor formalisms employing a summation over spinor indices, a homogeneous function approach more analogous to that of twistor theory may be adopted,

which is essentially the defining relation for the Majorana representation.

Using the single complex variable form of the homogeneous functions the relationship between the vector and Majorana representations is given by the theory of covariant vector operators for $SL(2,\mathbb{C})$.[24] An operator $T(x|z)$, where x and z are complex variables, is a covariant vector operator if for any pair of functions $\theta \in \mathcal{D}_\chi$ and $\psi \in \mathcal{D}_{-\chi'}$

$$f(x) = \int \theta(z)\, T(x|z)\, \psi(z)\, dz \tag{3.14}$$

is an element of the space $\mathcal{D}_{\chi^*}$ where $\chi^* = (2,2)$ and if the operator equation (3.15) is satisfied

$$T_{a^{-1}}^{\chi'}\, T(x)\, T_a^{\chi} = (a_{12}x + a_{22})\overline{(a_{12}x + a_{22})}\, T(x_a), \tag{3.15}$$

where $(a_{ij}) \in SL(2,\mathbb{C})$, $1 \le i,j \le 2$, and $\mathcal{D}_\chi$ with $x = (n_1, n_2)$ is the space of homogeneous functions such that for

$\theta(z) \in \mathcal{D}_x$,

$$T_a^{\chi}\theta(z) = (a_{12}z + a_{22})^{n_1 - 1}\overline{(a_{12}z + a_{22})}^{n_2 - 1}\, \theta\left(\frac{a_{11}z + a_{21}}{a_{12}z + a_{22}}\right). \tag{3.16}$$

The Majorana representation used here has $\chi, \chi' = \pm(\frac{1}{2}, -\frac{1}{2})$ and belongs to the principal series of representations, while the other Majorana representation series with $\chi = (-\frac{1}{2}, -\frac{1}{2})$ belonging to the supplementary series is infinite dimensional and does not fit into the geometric quantization scheme.[12] For the case $\chi = \pm(\frac{1}{2}, -\frac{1}{2})$ the covariant vector operator $T(x|z)$ has the respective forms

$$T(x|z) = (z - x)^{1/2}(\bar{z} - \bar{x})\, \frac{\partial}{\partial z}\, (z - x)^{1/2}, \tag{3.17}$$

$$T(x|z) = (\bar{z} - \bar{x})^{1/2}(z - x)\, \frac{\partial}{\partial \bar{z}}\, (\bar{z} - \bar{x})^{1/2}, \tag{3.18}$$

where bar denotes complex conjugation.

The formalism presented so far here has not depended on the essential feature of supergauge theory, namely that the supergauge transformations should generate a graded Lie algebra, thereby enabling the possibility of a unified non-Abelian gauge theory of both bosons and fermions. $L = \sum_{k \in N} L_k$ is a graded Lie algebra with index set N if given a bilinear map $[\ ,\ ]$ of $L \times L \to L$ the following conditions hold:

(a) $[L_k, L_l] \subset L_{k+l}$,

(b) $[x, y] = (-1)^{kl}[y, x]$,

(c) $[x, [y, z]] = [[x, y], z] + (-1)^{kl}[y, [x, z]]$, where

$x \in L_k$, $y \in L_l$, $z \in L_m$, $k, l, m \in N$. The GLA of Wess–Zumino has the form $L = L_2 \oplus L_{-1} \oplus L_0 \oplus L_1 \oplus L_2$, where the even elements generate a 16-D algebra isomorphous to u(2,2), corresponding to the infinitesimal automorphisms of the complex structure, and the odd elements are Majorana spinors defined via the complex structure on the tangent space $\mathbb{R}^4$ and also on the dual space $\mathbb{R}^{4*}$ of special conformal transformations (or "accelerations"). The nontrivial feature of this is the satisfaction of condition (c) which leads to use of the representation[1]

$$\begin{pmatrix} A & 0 & B \\ 0 & \lambda_A & 0 \\ C & 0 & -A^* \end{pmatrix}$$

of the general element

$$\begin{pmatrix} A & B \\ C & -A^* \end{pmatrix} \quad \text{of u(2,2)}$$

on the Fermi sector, where $A \in gl(2,\mathbb{C})$, $B = B^*$, $C = C^*$, and $\lambda_A = -2i\,\mathrm{Im}(\mathrm{Tr}A)$ and $*$ denotes Hermitian conjugation, instead of the obvious action of u(2,2) on basic column and row vectors. Since the element u(1) in the center of u(2,2) generates automorphisms of the commutators and anti-commutators in the Segal-type geometric quantization its eigenstates correspond to generalized charge eigenstates, so the significance of the GLA requirement (c) is that the charge of the Majorana spinors must be three times the value otherwise expected, in accordance with the commutator

$$[E, \psi_\alpha] = -(3i/2)(\gamma_5)_\alpha^\beta\, \psi_\beta, \quad \text{where } E \in u(1). \tag{3.19}$$

The other consequence of the GLA requirement is that the pair of Majorana spinors must define a twistor representation under su(2,2).

A number of different types of twistors have been developed by Penrose of which that involved in supergauge theory is usually regarded as the most trivial; it results from applying the Penrose construction to the tangent space at each point of spacetime separately so giving a field of flat space twistors. Curvature in this model is described by the position dependence of the complex structure tensor J, i.e., by the nonlinear realization $GL(4,\mathbb{R})/GL(2,\mathbb{C})$, or alternatively by the complex connection coefficients corresponding to the nonlinear realization $G^2(4)/U(2,2)$, where $G^2(4)$ is the structure group of the bundle of second order frames[18] and consists of elements of the form (a_j^i, a_{jk}^i), where $a_j^i \in GL(4,\mathbb{R})$ and $a_{jk}^i = a_{kj}^i$. One difference from the Penrose approach to twistors is that according to the preceding analysis quantum mechanics is already built into the theory through solving the supergauge commutation relations for the momentum operator, so giving the modified spinor formalism (3.9), (3.13) for $\partial/\partial x^\mu$.

The 14-dimensional subalgebra which generates the Volkov–Akulov supergauge transformations consists of the generators of the Poincaré group, together with the components of the Majorana spinor ψ_α, $1 \le \alpha \le 4$, and defines the algebra of infinitesimal automorphisms of the Hermitian metric g, (X, Y) of space–time. The geometrical approach to supergauge theory followed here shows the transition from the Wess–Zumino supergauge algebra to the 14-dimensional subalgebra as due explicitly to the breaking of conformal symmetry, and has one nontrivial consequence; namely, it requires $\rho^2 \propto \hbar$, where ρ is the basic unit of length used in the geometric quantization. To obtain the explicit appearance of the gravitional coupling constant in this relationship, it is necessary to consider the symmetry breaking through the Ricci curvature form of the line bundle with structure group $GL(2,\mathbb{C})/SL(2,\mathbb{C})$ and its relation to the field equations of general relativity.

IV. DISCUSSION

The main result of this paper is to show via the theory of geometric quantization how Wess–Zumino supergauge theory is related to the spinor and twistor[25] formulations of geometry, subject to the proportionality $\rho^2 \propto \hbar$ where it suggests two extensions of the spinor formalism:

(1) the introduction of quantum mechanics through inversion of the supergauge commutation relations, and

(2) the use of Majorana four-component spinors instead of the usual two-component spinors. It also shows how the theory of quantization based on the line bundle over spacetime with structure group $GL(2,\mathbb{C})/SL(2,\mathbb{C})$ is related to twistor theory.

Regarding supergravity and the curved space formulation of supergauge theory this theory amounts to one of local supergauge transformations depending only linearly on position, but with the definition of the Majorana spinor depending arbitrarily on position through the local reduction $GL(4,\mathbb{R})/SL(2,\mathbb{C})$. The infinitesimal automorphisms of the set of all possible such reductions $GL(4,\mathbb{R})/SL(2,\mathbb{C})$ are simply the infinitesimal automorphisms of the $GL(4,\mathbb{R})$ structure itself which are the nonsingular infinitesimal coordinate transformations. This theory is much more restricted than the Kaluza–Klein type supergeometry considered by Arnowitt and Nath because the space–time metric $g_{\mu\nu}$ determines the definition of the Majorana spinors and hence also the spinor metric components $g_{\alpha\beta}$ of the supermetric g_{AB}. Furthermore fields defined on Majorana spinor space must be related to those on spacetime through the covariant vector operator formalism (3.14) for $SL(2,\mathbb{C})$, or more generally by the relation of quantum mechanical twistor functions to fields on spacetime. This geometrical interpretation also limits the scope for unified field theories of the Kaluza–Klein type, but retains the possibility that gauge symmetries can be described as due to the invariance of functions of several twistors under transformations of the twistors[26] and hence under the corresponding supergauge transformations.

The present geometric approach to supergauge theory corresponds to a very specific attitude to quantum gravity; namely that much of quantum theory is the theory of nonlinear realizations of $\mathbb{R}^4 \circledS GL(4,\mathbb{R})$. According to this view the reduction $GL(4,\mathbb{R})/GL(2,\mathbb{C})$ defines the complex structure underlying all second quantization, while remaining unquantized itself, and the reduction $GL(2,\mathbb{C})/SL(2,\mathbb{C})$ defines a section of the line bundle which has to be quantized, and is the part of the gravitational field which couples to $T_{\mu\nu}$. In terms of the curvature tensor the conformal part which is determined by J together with a phase factor remains unquantized while the Ricci curvature has to be quantized: some considerations of this scheme have been given elsewhere,[12] but it remains to be shown that a consistent quantum theory can be so obtained.

[1] L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. 47, 573 (1975).
[2] M. Scheunert, W. Nahm, and V. Rittenberg, J. Math. Phys. 17, 1626, 1640 (1976).
[3] R. Arnowitt and P. Nath, J. Gen. Rel. Grav. 7, 89 (1976).
[4] R. Arnowitt and P. Nath, Phys. Rev. Lett. 36, 1527 (1976).
[5] Y. M. Cho and P. G. O. Freund, Phys. Rev. D 12, 1711 (1975).
[6] P. G. O. Freund and I. Kaplansky, J. Math. Phys. 17, 228 (1976).
[7] D. Z. Freedman, P. van Nieuwenhuizen, and S. Ferrara, Phys. Rev. D 13, 3214 (1976).
[8] D. Z. Freedman and P. van Nieuwenhuizen, Phys. Rev. D 14, 912 (1976).
[9] S. Deser and B. Zumino, Phys. Lett. B 62, 335 (1976).
[10] S. Ferrara, J. Scherk, and P. van Nieuwenhuizen, Phys. Rev. Lett. 37, 1035 (1976).
[11] T. W. B. Kibble, J. Math. Phys. 2, 212 (1961).
[12] D. J. R. Lloyd-Evans, Int. J. Theor. Phys. (to be published).
[13] D. J. R. Lloyd-Evans, Nuovo Cimento Lett. 16, 321 (1976).
[14] I. Segal, J. Math. Phys. 1, 468 (1960).
[15] B. Kostant, Lect. Notes Phys. 6, 237 (1970).
[16] B. Kostant, Lect. Notes Math. 170, 87 (1970).
[17] J. Wess and B. Zumino, Nucl. Phys. B 70, 39 (1974).
[18] S. Kobayashi, *Transformation Groups* (Springer-Verlag, Berlin, 1972).
[19] A. Salam and J. Strathdee, Nucl. Phys. B 76, 477 (1974).
[20] D. J. Simms and N. M. J. Woodhouse, Lect. Notes Phys. 53, 1 (1976).
[21] E. Cartan, *Spinors*, M.I.T. transl. (M.I.T. Press, Cambridge, Mass., 1966).
[22] D. Hestenes, J. Math. Phys. 8, 798, 809, 1046 (1967).
[23] S. Ferrara, J. Wess, and B. Zumino, Phys. Lett. B 51, 239 (1974).
[24] W. Ruhl, *The Lorentz Group and Harmonic Analysis* (Benjamin, New York, 1970), Chap. III.
[25] R. Penrose and M. A. H. MacCallum, Phys. Rep. C 6, 241 (1972).
[26] R. Penrose, *Quantum Gravity*, edited by C. J. Isham, R. Penrose, and D. Sciama (Clarendon, Oxford, 1975), Chap. VI.

Correlations near phase transition in a simple fluid

G. V. Ramanathan*

Department of Mathematics, University of Illinois at Chicago Circle, Box 4348, Chicago, Illinois 60680
(Received 22 September 1976)

Correlations in a system with a weak, long-ranged attractive potential are studied. Using a natural small parameter, the asymptotic orders of all correlations are established. Explicit leading order solutions are obtained for all correlations when the system is away from phase change. As it approaches transition, it is shown that the many-body correlations, even though still small, develop very long ranges and that they are extremely slowly varying functions over space. As a result all the correlations contribute equally significantly to the computation of the pair correlation. In an asymptotic region near transition, a hierarchy is derived in which each correlation is expressed in terms of its immediate predecessor, its immediate successor and the pair correlation. The pair potential does not appear in these equations. Instead the correlations are expressed in terms of a Yukawa potential whose strength and range are related to the density and other parameters but not to the detailed form of the potential. In the leading approximation, the pair correlation is proportional to this Yukawa function.

INTRODUCTION

The purpose of this paper is to investigate the behavior of correlations in a simple fluid in a region near change of phase. We consider a model in which the molecular interaction consists of a small hard core and a weak, long-range attractive force. The precise nature of the attractive force is unimportant for the analysis here and is never used. The only requirement made is that the integral over all space of the attractive potential be of order one. The smallness of the parameter describing the weak nature of the potential (as well as its long range) is used to obtain a consistent asymptotic ordering of all correlations. The asymptotic analysis of this problem away from the region of condensation turns out to be the same as that in the so-called "plasma limit."[1] This model resembles that of Kac, Uhlenbeck, and Hemmer,[2,3] but the actual inverse-range parametrization is similar to that suggested by Stell and Theumann[4] and by Høye and Stell.[5]

The usual approach to problems of this kind is to perform a virial expansion and then resum the significant terms.[2,6] Graph-theoretical techniques are useful and powerful in doing this and as Stell[7] has shown, they can be used to solve the formal closure problem of expressing the many-body correlations in terms of the two-body function.

We present here a somewhat different approach. Starting from the BBGKY hierarchy, we make use of a small parameter to order all the correlations and to obtain explicit asymptotic solutions for them, which would be uniformly valid throughout space. We show that in this asymptotic limit, the three-body and higher correlations are uniformly small. Actually, the higher correlations are succesively smaller. So long as the system is not close to condensation, the contributions of higher correlations to the computation of the pair correlation are succesively more and more insignificant. As a result one obtains, in an asymptotic limit, the well known hypernetted chain equation for the pair correlation function.[7] But as the system approaches phase change,[8] even though the higher correlations remain small, their range becomes large. Therefore, their contributions to certain integrals become significant. More precisely, *every* correlation contributes to the leading order of the pair correlation. For this asymptotic region, an infinite chain of equations is derived and it cannot be legitimately truncated. This chain is simpler than the BBGKY hierarchy in two senses. One is that it contains far fewer terms. The other is that whereas the BBGKY hierarchy expresses each correlation in terms itself, its successor and all its predecessors, our hierarchy expresses each correlation in terms of its immediate predecessor, its immediate successor, and the pair correlation. In other words, whereas each equation in the BBGKY hierarchy is an integral equation, ours is obtained by asymptotically solving for each correlation in terms of its predecessor, its sucessor, and the pair correlation. The pair potential does not appear in any of these equations. Thus these results are independent of the detailed form of the potential. Instead, what appears is a Yukawa potential whose strength and range are related to the strength, range, and the second moment of the pair potential. In the leading approximation, truncation of the hierarchy at the first level gives an expression of the pair correlation of this Yukawa form.

In Sec. 1, we define correlation functions through a cluster expansion suitable for our purposes. We present the equation for the general s-body correlation. In the next section the potential and the origin of the small parameter are described. Using that, all the correlations are ordered and leading order solutions are obtained for the three-body and higher correlations. In Sec. 3, these solutions are used to derive the hypernetted chain equation, away from the region of condensation. Then in Sec. 4, the entire asymptotics is redone as the system is assumed to be close to phase change.

1. BASIC EQUATIONS

The starting point for our analysis is a hierarchy of equations for correlation functions, derived from the BBGKY hierarchy through a cluster expansion. No simplifying assumptions are made in this process. The details and proofs for the derivation are given in Ref. 1. Here we shall only symmarize the results. It should

be pointed out that the cluster expansion used here is different from the well-known one due to Mayer and Ursell. Also the functions we call s-body correlations are different from the Mayer–Ursell correlations except for $s=2$. Our functions are related to potentials of average force.[1,9] The advantages of using such a cluster expansion are many. Not only can one write down the sth equation of the hierarchy explicitly, but all the correlations can be consistently ordered in terms of a small parameter. It can be shown that for $s>2$, all the s-body correlations are uniformly small throughout space. Under certain conditions, this leads to a legitimate truncation of the hierarchy and one can obtain explicit solutions for the three-body and higher correlations. These solutions are functionals of the pair correlation and do not depend upon the detailed nature of the potential.[1,7,10]

Consider a spatially homogeneous system of identical particles, each of mass m, interacting through a two-body potential ϕ. We shall assume that the bulk limit exists and that the system is being described in that limit. Let n be the number density of particles. We shall denote by $\{p\}_r^s$ $(1\leq p\leq s-r+1)$, a set of p particles chosen from the set $(r, r+1, \ldots, s)$ of $s-r+1$ particles. $f_{\{p\}_r^s}$ will stand for the p-particle "reduced" distribution function[11] for the set $\{p\}_r^s$. Sometimes we shall find it convenient to name some or all the particles. Then we shall use the notation $f_{a_1,a_2,\ldots,a_q,\{p\}_r^s}$ to stand for the $(q+p)$-particle reduced distribution of the q particles, $a_1, a_2, \ldots, a_q$ and the p particles belonging to the set $\{p\}_r^s$. For example, f_{34} will stand for the two-body reduced distribution of the particles 3 and 4. We shall denote by $\{p/j'\}_r^s$ a set of p particles from the set obtained by removing the jth particle from $\{r, r+1, \ldots, s\}$.

Then we have, for $s=2,3,\cdots$ the well-known BBGKY hierarchy,

$$\left[\frac{\partial}{\partial t}+\sum_{i=1}^{s}\mathbf{v}_i\cdot\frac{\partial}{\partial \mathbf{x}_i}-\frac{1}{m}\sum_{i\neq j=1}^{s}\sum\frac{\partial\phi_{ij}}{\partial\mathbf{x}_i}\cdot\frac{\partial}{\partial\mathbf{v}_i}\right]f_{\{s\}_1^s}=\frac{n}{m}\int\sum_{i=1}^{s}\frac{\partial\phi_{i,s+1}}{\partial\mathbf{x}_i}\cdot\frac{\partial}{\partial\mathbf{v}_i}f_{\{s+1\}_1^{s+1}}\,d\mathbf{x}_{s+1}\,d\mathbf{v}_{s+1}. \quad (1.1)$$

Here $\mathbf{x}_i$ and $\mathbf{v}_i$ denote the position and velocity, respectively of particle i, ϕ_{ij} denotes the potential between particles i and j and $\partial/\partial\mathbf{x}_i$ and $\partial/\partial\mathbf{v}_i$ stand for the gradients with respect to the position vector and velocity vector, respectively, of particle i.

We define the s-particle correlation $\alpha_{\{s\}_1^s}$, for $s\geq 2$, through the following equation:

$$f_{\{s\}_1^s}\equiv\frac{\Pi f_{\{s-1\}_1^s}\ \Pi f_{\{s-3\}_1^s}}{\Pi f_{\{s-2\}_1^s}\ \Pi f_{\{s-4\}_1^s}}\cdots\left(\Pi f_{\{1\}_1^s}\right)^{(-1)^s}\left[1+\alpha_{\{s\}_1^s}\right]. \quad (1.2)$$

Here $\Pi f_{\{p\}_1^s}$ stands for the product of all distinct p-particle distribution functions that can be formed from the s particles $1, 2, \ldots, s$. This product will contain $\binom{s}{p}$ factors. If one neglects $\alpha_{\{s\}_1^s}$, (1.2) can be thought of as a generalization of Kirkwood's superposition approximation. For $s=2$, 3, and 4, for example, the correlations are defined by

$$f_{12}=f_1f_2(1+\alpha_{12}),$$

$$f_{123}=\frac{f_{12}f_{13}f_{23}}{f_1f_2f_3}(1+\alpha_{123}),$$

$$f_{1234}=\frac{f_{123}f_{124}f_{134}f_{234}}{f_{12}f_{13}f_{14}f_{23}f_{24}f_{34}}f_1f_2f_3f_4(1+\alpha_{1234}).$$

α_{12} is what is usually known as the pair correlation function. As for the distribution functions, $\alpha_{a_1,a_2,\ldots,a_q,\{p\}_r^s}$ will stand for the $(q+p)$-particle correlation function of the set of particles denoted by the subscripts.

In thermal equilibrium, all the correlations are independent of velocities and are functions only of positions. Also,

$$f_1=\left(\frac{m}{2\pi kT}\right)^{3/2}\exp\left(-\frac{mv_1^2}{kT}\right), \quad (1.3)$$

where k is the Boltzmann's constant and T is the temperature. In that case, we can substitute for the distributions the expressions (1.2) and (1.3) and obtain from the BBGKY hierarchy, the following hierarchy for correlations:

$$\frac{\partial}{\partial\mathbf{x}_1}\log(1+\alpha_{12})=-\frac{1}{kT}\frac{\partial\phi_{12}}{\partial\mathbf{x}_1}-\frac{n}{kT}\int\frac{\partial\phi_{13}}{\partial\mathbf{x}_1}(1+\alpha_{13})\times(\alpha_{23}+\alpha_{123}+\alpha_{23}\alpha_{123})\,d\mathbf{x}_3. \quad (1.4)$$

For $s\geq 3$,

$$\frac{\partial}{\partial\mathbf{x}_1}\log(1+\alpha_{\{s\}_1^s})=-\frac{n}{kT}\int\frac{\partial\phi_{1,s+1}}{\partial\mathbf{x}_1}\times(1+\alpha_{1,s+1})\sum\prod_{i=1}^{m}\alpha_{\{p_i\}_1^s,s+1}, \quad (1.5)$$

where the summation is over all products satisfying the following conditions:

(i) The sets $\{p_i\}_1^s$ occurring in each product are distinct.

(ii) If $p_i=1$, the particle 1 is not a member of $\{p_i\}_1^s$.

(iii) Let $\cup_{i=1}^m\{p_i\}_1^s=\{p\}_1^s$. Then $p=s-1$ or $p=s$.

(iv) If $p=s-1$, the particle 1 is not a member of $\{p\}_1^s$.

This means that, for example, for $s=4$, terms such as $\alpha_{25}\alpha_{235}$ and $\alpha_{125}\alpha_{235}$ cannot occur in the sum, the former because of (iii) and the latter because of (iv). These statements are proved in Ref. 1.

2. THE POTENTIAL AND ITS EFFECTS ON ORDERING

In order to obtain useful solutions to the hierarchy [(1.4), (1.5)], we look for methods of simplifying it. If there is a small parameter ϵ natural to the system, it can be used to estimate the orders of the correlation functions and one can attempt an asymptotic solution of the hierarchy. For this purpose, let us suppose that the potential consists of a small hard core and a weak long-ranged attraction. The small parameter will be related to the ratio between the range of the hard core and that of the attractive potential. There is a natural length scale in the problem, namely, $n^{-1/3}$. We shall

use this as a standard of reference. For the sake of facility of expression, sometimes we shall refer to a distance as being of a certain order instead of saying that ratio of that distance to $n^{-1/3}$ is of that order. The symbols O and o will be used in their usual sense. We shall say that two quantities $y(\epsilon)$ and $z(\epsilon)$, are precisely of the same order and denote it by $y \sim z$ if $y = O(z)$ and $z = O(y)$.

The radius of the hard core will be assumed to be of the order of $\epsilon^{2/3}$ where ϵ is $o(1)$. The exponent $\frac{2}{3}$ is chosen so that the effect of the hard core would be significant in an asymptotic region near phase change. The attractive part of the potential will be assumed to have a range of the order of ϵ^{-1} and its strength for distances of its range to be of order ϵ^3. Outside the hard core, $\phi(r)$ will be assumed to be a smooth, negative, increasing function of r such that

$$\frac{1}{kT}\int_{\delta}^{\infty} \phi(r) r^2\, dr \sim 1,$$

where δ is the radius of the hard core. The analysis that follows is independent of the detailed form of the potential. For the purpose of visualizing it, however, one could think of the attractive part as something like $-\epsilon^2 A \exp(-\epsilon\beta r)/r$, where $A, \beta \sim 1$.

Under these assumptions, away from the region of phase change, the asymptotic analysis turns out to be similar to that in the so-called "plasma limit." The technique employed is one of successive estimates and consistency arguments. It can be summarized as follows. Suppose that the asymptotic orders of $\alpha_{\{j\}_1^j}$ for $j = 1, 2, \ldots, s-1$ are known. In the equations for $\alpha_{\{s\}_1^s}$, all the terms involving $\alpha_{\{s+1\}_1^{s+1}}$ are neglected and the order of $\alpha_{\{s\}_1^s}$ is estimated. Then in the equation for $\alpha_{\{s+1\}_1^{s+1}}$, all the terms containing the $(s+2)$-particle correlation are neglected and this estimate of $\alpha_{\{s\}_1^s}$ and those known for lower correlations are used to estimate the order of $\alpha_{\{s\}_1^{s+1}}$. This is then put back in the full equation for $\alpha_{\{s\}_1^s}$ to check if it affects the original estimate. Finally, after having found the orders of all the correlations, as a result of this ordering procedure, one can obtain explicit solutions to the leading order for *all* correlations, away from the region of phase change. The consistency argument is completed and verified by observing that these solutions concur in order with the earlier asymptotic estimates.

To illustrate this, let us start with (1.4). Neglecting α_{123}, suppose one seeks a solution to the resulting differentio-integral equation,

$$\frac{\partial \log(1+\alpha_{12})}{\partial \mathbf{x}_1} = \frac{1}{kT}\frac{\partial \phi_{12}}{\partial \mathbf{x}_1} - \frac{n}{kT}\int \frac{\partial \phi_{13}}{\partial \mathbf{x}_1}(1+\alpha_{13})\alpha_{23}\, d\mathbf{x}_3 \tag{2.1}$$

by iteration. One would start by setting

$$\frac{\partial \log(1+\alpha_{12}^0)}{\partial \mathbf{x}_1} = -\frac{1}{kT}\frac{\partial \phi_{12}}{\partial \mathbf{x}_1}, \tag{2.2}$$

solve this, substitute it for α_{13} and α_{23} inside the integral in (2.1), and solve it again and so on. If the iteration converges or if it is asymptotic (i.e., yielding successively smaller correction terms as $\epsilon \to 0$), the solution to (2.2) will be a good pointwise estimate of the solution to (2.1). Integrating (2.2) with the condition that $\alpha_{12} \to 0$ as $|\mathbf{x}_1 - \mathbf{x}_2| \to \infty$ we have

$$\alpha_{12}^0 = \Psi_{12} \equiv -1 + \exp(-\phi_{12}/kT). \tag{2.3}$$

For $|\mathbf{x}_1 - \mathbf{x}_2| = O(\epsilon^{2/3})$, $\Psi_{12} \sim 1$ and if $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1$ or larger, $\Psi_{12} \sim \phi_{12}/kT = o(1)$. Substituting (2.3) in (2.1), we find that the integral term becomes

$$\frac{\partial}{\partial \mathbf{x}_1} n \int \Psi_{13}\Psi_{23}\, d\mathbf{x}_3.$$

Let $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}$ and $\mathbf{x}_1 - \mathbf{x}_3 = \boldsymbol{\xi}$. Then,

$$n\int \Psi_{13}\Psi_{23}\, d\mathbf{x}_3 \sim n|\boldsymbol{\xi}|^3 \Psi(|\boldsymbol{\xi}|)\Psi(|\mathbf{r}-\boldsymbol{\xi}|). \tag{2.4}$$

If $|\mathbf{r}| \sim \epsilon^{2/3}$, the major contribution comes from $|\boldsymbol{\xi}| \sim \epsilon^{2/3}$ and this term is of the order ϵ^2. Hence the iteration is asymptotic yielding the virial expansion, with successive terms of higher order in ϵ. Therefore, for $|\mathbf{r}| \sim \epsilon^{2/3}$, $\alpha_{12} \sim 1$. For $|\mathbf{r}| \sim 1$ or bigger, $\Psi_{12} \sim \phi_{12}/kT$. When $|\mathbf{r}| \sim 1/\epsilon$, i.e., the range of the potential, with $|\boldsymbol{\xi}| \sim 1/\epsilon$, the right-hand side of (2.4) becomes precisely of the order of $\Psi(\mathbf{r})$. In other words, the integral term is significant to the leading order when $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1/\epsilon$. Differentiating the logarithm in (2.1),

$$\frac{\partial \alpha_{12}}{\partial \mathbf{x}_1} = -\frac{1}{kT}\frac{\partial \phi_{12}}{\partial \mathbf{x}_1}(1+\alpha_{12}) - \frac{n}{kT}(1+\alpha_{12})\int \frac{\partial \phi_{13}}{\partial \mathbf{x}_1}(1+\alpha_{13})\alpha_{23}\, d\mathbf{x}_3.$$

Since α_{12} in the first estimate is Ψ_{12} which is $o(1)$ for $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1$ or larger, and since the significant contribution to the integral when $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1/\epsilon$, comes from $|\mathbf{x}_1 - \mathbf{x}_3| \sim 1/\epsilon$, we can write the above equation in the form

$$\frac{\partial \alpha_{12}}{\partial \mathbf{x}_1} - n\int \frac{\partial \Psi_{13}}{\partial \mathbf{x}_1}\alpha_{23}\, d\mathbf{x}_3 = \frac{\partial \Psi_{12}}{\partial x_1} + o\left[\frac{\partial \alpha_{12}}{\partial \mathbf{x}_1}\right]. \tag{2.4'}$$

Observe that this equation is valid for all distances. When $|\mathbf{x}_1 - \mathbf{x}_2| \sim \epsilon^{2/3}$, the integral term is small, but when $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1/\epsilon$, it is significant. The operator on the left side of (2.4'), acting on the pair-correlation function, is linear. The term $o(\partial\alpha_{12}/\partial\mathbf{x}_1)$ will generate an asymptotic iteration. The rest can be integrated to obtain

$$\alpha_{12} - n\int \Psi_{13}\alpha_{23}\, d\mathbf{x}_3 = \Psi_{12} + o(\alpha_{12}). \tag{2.5}$$

Since $\Psi_{12} = -1 + \exp(-\phi_{12}/kT)$ and ϕ_{12} outside the hard core is assumed to be a negative, monotonically increasing, integrable function, Ψ_{12} is square integrable. Therefore, so long as the quantity

$$\mu \equiv 1 - n\int \Psi_{12}\, d\mathbf{x}_2 \tag{2.6}$$

is a fixed positive number (~ 1), by classical Fredholm theory[12] the iteration converges everywhere. Thus the order of α_{12} as given by (2.3) remains unaltered, and we have

$$\alpha_{12} \sim \Psi_{12} \quad \text{everywhere.} \tag{2.7}$$

Now we go to Eq. (1.5) with $s = 3$ and set α_{1234} to zero. The leading iterant is

$$\frac{\partial \log(1+\alpha_{123}^0)}{\partial \mathbf{x}_1} = -\frac{n}{kT}\int \frac{\partial \phi_{14}}{\partial \mathbf{x}_1}(1+\alpha_{14})\alpha_{24}\alpha_{34}\, d\mathbf{x}_4. \tag{2.8}$$

Using (2.7) and the definition of Ψ_{12} as given by (2.3) and integrating with the condition that $\alpha_{123} \to 0$ as the particle separations tend to infinity, we get the estimate,

$$\log(1+\alpha^0_{123}) \sim n \int \Psi_{14}\Psi_{24}\Psi_{34}\, d\mathbf{x}_4. \tag{2.9}$$

Let $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}$, $\mathbf{x}_1 - \mathbf{x}_3 = \boldsymbol{\xi}$, and $\mathbf{x}_1 - \mathbf{x}_4 = \boldsymbol{\eta}$. Then the right side of (2.9) is of the order

$$n|\boldsymbol{\eta}|^3 \Psi(|\boldsymbol{\eta}|)\Psi(|\mathbf{r}-\boldsymbol{\eta}|)\Psi(|\boldsymbol{\xi}-\boldsymbol{\eta}|).$$

It is easily verified that for *all* values of $\mathbf{r}$ and $\boldsymbol{\xi}$, this quantity is $O(\epsilon^6)$. Hence $\alpha_{123} = o(1)$ and we can write (2.9) as

$$\alpha^0_{123} \sim n \int \Psi_{14}\Psi_{24}\Psi_{34}\, d\mathbf{x}_4 \sim \epsilon^6. \tag{2.10}$$

Since $\alpha_{12} \sim \Psi_{12}$ everywhere, in the equation for α_{123} all terms which are integrals over products of pair correlation and three-particle correlation are $o(\partial\alpha_{123}/\partial\mathbf{x}_1)$ and we can write (1.5) for $s=3$, after setting α_{1234} to zero, as

$$\frac{\partial\alpha_{123}}{\partial\mathbf{x}_1} - n \int \frac{\partial\Psi_{14}}{\partial\mathbf{x}_1}\,\alpha_{234}\, d\mathbf{x}_4 = n \int \frac{\partial\Psi_{14}}{\partial\mathbf{x}_1}\,\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 + o\left[\frac{\partial\alpha_{123}}{\partial\mathbf{x}_1}\right]. \tag{2.11}$$

As before, so long as $\mu \equiv 1 - n\int \Psi_{12}\, d\mathbf{x}_2$ is a fixed positive number, the iteration converges and the estimate (2.10) remains unaltered.

Now we have to put this estimate in (1.4) and verify that (2.7) still remains valid. This is immediate, since from (2.10) $\alpha_{123} = o(\Psi_{12})$ everywhere and from (2.11) the integral terms in (1.4) containing α_{123} are $o(\partial\Psi_{12}/\partial\mathbf{x}_1)$. But $\partial\Psi_{12}/\partial\mathbf{x}_1$ is the leading iterant and hence α_{123} will make asymptotically small contributions to the iteration. This completes the argument for the pair-correlation and partly for the three-body correlation. Next one proceeds to the equation for α_{1234}, estimates it and verifies that it contributes asymptotically small terms to α_{123}. In general one can show that for $s \geq 3$,

$$\alpha_{\{s\}_1^s} \sim n \int \prod_{j=1}^{s} \Psi_{j,s+1}\, d\mathbf{x}_{s+1}. \tag{2.12}$$

This result is demonstrated by a different argument in Ref. 2. It is easily seen from (2.12) that for $s > 2$, $\alpha_{\{s\}_1^s} = o(1)$ for *all* values of its argument. In particular, if every pair of particles is separated by a distance of the order $\epsilon^{2/3}$, $\alpha_{\{s\}_1^s} \sim \epsilon^2$ and if every pair is separated by a distance $1/\epsilon$, viz., the range of the potential, $\alpha_{\{s\}_1^s} \sim \epsilon^{3s-3}$. It should be pointed out that the range of every correlation is of the order of the potential. In other words, if any pair of particles is separated by a distance much larger than the potential [i.e., $1/\epsilon = o(\text{distance})$], $\alpha_{\{s\}_1^s}$, where $\{s\}_1^s$ contains this pair, essentially vanishes as fast as the potential. This is obvious from the estimate (2.12). This is due to the fact that we have assumed that $\mu \sim 1$. If μ is sufficiently small, however, as we shall see in Sec. 4, the range of each correlation is much longer than that of the potential.

It should be emphasized that in all these ordering arguments, ϵ is considered to be an infinitesimal in the strict asymptotic sense. In other words, the sum of *any* finite number of infinitesimally small terms is still an infinitesimal.

Now for the final step of obtaining leading order solutions let us first illustrate this with α_{12} and α_{123}. The equation (2.5) can be solved by taking Fourier transforms and it is easily verified that α_{12} is of the same order as Ψ_{12}. The leading order equation (2.11) for α_{123} can be integrated and written as

$$\alpha_{123} - n\int \Psi_{14}\,\alpha_{234}\, d\mathbf{x}_4 = n\int \Psi_{14}\,\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 + o(\alpha_{123}). \tag{2.13}$$

Substituting for Ψ_{14} on the right side from (2.5) and rearranging terms, we have

$$\alpha_{123} - n\int \alpha_{14}\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 - n\int \Psi_{14}[\alpha_{234} - n\int \alpha_{25}\alpha_{35}\alpha_{45}\, d\mathbf{x}_5]\, d\mathbf{x}_4 = o(\alpha_{123}).$$

From a well-known uniqueness theorem for Fredholm equations we conclude that

$$\alpha_{123} = n\int \alpha_{14}\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 + o(\alpha_{123}). \tag{2.14}$$

Similarly, one can show that

$$\alpha_{1234} = n\int \alpha_{15}(\alpha_{25}\alpha_{35}\alpha_{45} + \alpha_{25}\alpha_{345} + \alpha_{35}\alpha_{245} + \alpha_{45}\alpha_{235})\, d\mathbf{x}_5 + o(\alpha_{1234}). \tag{2.15}$$

In general for $s \geq 3$,

$$\alpha_{\{s\}_1^s} = n\int \alpha_{1,s+1} \sum \prod_{j=1}^{m} \alpha_{\{p_j\}_2^s,\, s+1}\, d\mathbf{x}_{s+1} + o(\alpha_{\{s\}_1^s}), \tag{2.16}$$

where the summation is over all products such that

(i) In any one product all the $\{p_j\}$'s are distinct,

(ii) $p_j < s-1$,

(iii) $\cup_{j=1}^{m} \{p_j\} = \{s-1\}_2^s$,

(iv) all the $\{p_j\}_2^s$'s are pairwise disjoint.

A proof of this is given in Ref. 1. The symmetry of the expression (2.16) under permutations of subscripts can be verified. The term consisting of $p_j = 1$ for all j along with the fact that $\alpha_{12} \sim \Psi_{12}$, corroborates the estimate (2.12). It may be noted that the potential does not appear explicitly in (2.16). Thus all the correlations are expressed as functionals of the pair correlation.

3. THE HYPERNETTED CHAIN EQUATION

This is obtained by retaining terms up to and including the leading order in α_{123} in Eq. (1.4). Since $\alpha_{23} \sim 1$ only for distances of order $\epsilon^{2/3}$ and $o(1)$ for longer distances, the term $\int (\partial\alpha_{13}/\partial\mathbf{x}_1)\,\alpha_{23}\alpha_{123}\, d\mathbf{x}_3$ is $o[\int (\partial\phi_{13}/\partial\mathbf{x}_1)\,\alpha_{123}\, d\mathbf{x}_3]$ and hence can be neglected to this order. Substituting for α_{123} its leading order solution (2.14), (1.4) now becomes

$$\frac{\partial \log(1+\alpha_{12})}{\partial\mathbf{x}_1} = -\frac{1}{kT}\frac{\partial\phi_{12}}{\partial\mathbf{x}_1} - \frac{n}{kT}\int \frac{\partial\phi_{13}}{\partial\mathbf{x}_1}(1+\alpha_{13})\,\alpha_{23}\, d\mathbf{x}_3 - \frac{n^2}{kT}\iint \frac{\partial\phi_{13}}{\partial\mathbf{x}_1}(1+\alpha_{13})\,\alpha_{14}\alpha_{24}\alpha_{34}\, d\mathbf{x}_4\, d\mathbf{x}_3.$$

Note that the last term on the right side is of a higher order than $\partial\log(1+\alpha_{12})/\partial\mathbf{x}_1$. Therefore, we can sub-

stitute to this order,

$$-\frac{n}{kT}\int\frac{\partial\phi_{13}}{\partial\mathbf{x}_1}(1+\alpha_{13})\alpha_{34}\,d\mathbf{x}_3$$
$$=\frac{\partial}{\partial\mathbf{x}_1}\log(1+\alpha_{14})+\frac{1}{kT}\frac{\partial\phi_{14}}{\partial\mathbf{x}_1}+o\left[\frac{\partial\alpha_{14}}{\partial\mathbf{x}_1}\right]$$

to obtain

$$\frac{\partial\log(1+\alpha_{12})}{\partial\mathbf{x}_1}=-\frac{1}{kT}\frac{\partial\phi_{12}}{\partial\mathbf{x}_1}-\frac{n}{kT}\int\frac{\partial\phi_{13}}{\partial\mathbf{x}_1}(1+\alpha_{13})\alpha_{23}\,d\mathbf{x}_3$$
$$+\int\frac{\partial\log(1+\alpha_{14})}{\partial\mathbf{x}_1}\alpha_{14}\alpha_{24}\,d\mathbf{x}_4$$
$$+\frac{n}{kT}\int\frac{\partial\phi_{14}}{\partial\mathbf{x}_1}\alpha_{14}\alpha_{24}\,d\mathbf{x}_4+o\left[\epsilon^3\frac{\partial\alpha_{12}}{\partial\mathbf{x}_1}\right].$$

This, upon simplification and integration, yields the hypernetted chain equation,

$$\log(1+\alpha_{12})=-\frac{\phi_{12}}{kT}-n\int\left(\frac{\phi_{13}}{kT}+\log(1+\alpha_{13})-\alpha_{13}\right)\alpha_{23}\,d\mathbf{x}_3. \tag{3.1}$$

The last two terms on the right side are $o(\alpha_{12})$ if $\mu\sim1$. As μ becomes small, however, they become significant. Specifically we shall show that when $\mu\sim\epsilon^2$ they are of the same order as the terms retained. It is easily verified that the range of α_{12} is of the order of $\epsilon^{-1}\mu^{-1/2}$ and its magnitude in its range is of the order $\epsilon^3\mu^{1/2}$. Suppose that $\mu\ll1$, and let $|\mathbf{x}_1-\mathbf{x}_2|\sim\epsilon^{-1}\mu^{-1/2}$ so that the particles are separated by distances much larger than the range of the potential. The contribution to the integral from the hard core will be $O(\epsilon^2\alpha_{12})$ and hence is insignificant to the leading order in α_{12}. So we can replace ϕ_{13}/kT by $-\Psi_{13}\equiv1-\exp(-\phi_{13}/kT)$ and expand the logarithms to obtain,

$$\alpha_{12}-n\int\Psi_{13}\alpha_{23}\,d\mathbf{x}_3=-\phi_{12}/kT+\tfrac{1}{2}n\int\alpha_{13}^2\alpha_{23}\,d\mathbf{x}_3. \tag{3.2}$$

Let $\mathbf{x}_1-\mathbf{x}_2=\mathbf{r}$ and $\mathbf{x}_1-\mathbf{x}_3=\boldsymbol{\xi}$. If $|\mathbf{r}|\gg1/\epsilon$, the range of the potential, $|\mathbf{r}|\gg|\boldsymbol{\xi}|$ in the integral term on the left side. We can expand $\alpha_{23}=\alpha(\mathbf{r}-\boldsymbol{\xi})$ in a Taylor series around $\mathbf{r}$.

$$\alpha_{23}=\alpha_{12}-\xi_i\frac{\partial}{\partial r_i}\alpha_{12}+\tfrac{1}{2}\xi_i\xi_j\frac{\partial^2\alpha_{12}}{\partial r_i\,\partial r_j}+\cdots, \tag{3.3}$$

where the subscripts i,j denote components and the summation convention is used. Because of the spherical symmetry of Ψ_{13}, all the odd derivative terms integrate to zero. So if one keeps the second derivative terms, the remainder is of the order $|\boldsymbol{\xi}|^4(d^4/d|\mathbf{r}|^4)\alpha(|\mathbf{r}|)$. We anticipate that $(d^4/d|\mathbf{r}|^4)\alpha(|\mathbf{r}|)\sim\alpha(|\mathbf{r}|)/|\mathbf{r}|^4$. Since Ψ has a range of $1/\epsilon$, and is of magnitude ϵ^3 in its range,

$$\int\Psi_{13}|\mathbf{x}_1-\mathbf{x}_3|^4\,d\mathbf{x}_3\sim1/\epsilon^4.$$

Therefore, the remainder term upon integration is of the order $\alpha(|\mathbf{r}|)/|\mathbf{r}|^4\epsilon^4$, which is much smaller than the second derivative term which is of the order $\alpha(|\mathbf{r}|)/|\mathbf{r}|^2\epsilon^2$. Now (3.2) can be written as

$$\mu\alpha_{12}-\frac{n}{2}\nabla^2\alpha_{12}\int|\mathbf{x}_1-\mathbf{x}_3|^2\Psi_{13}\,d\mathbf{x}_3=\frac{n}{2}\int\alpha_{13}^2\alpha_{23}\,d\mathbf{x}_3. \tag{3.4}$$

The ϕ_{12} term does not appear because $|\mathbf{x}_1-\mathbf{x}_2|\gg1/\epsilon$. If one neglects the right side, which arises from α_{123}, the solution is

$$\alpha_{12}=\frac{\epsilon^2A}{kT}\frac{\exp(-\epsilon\nu|\mathbf{r}|)}{|\mathbf{r}|}, \tag{3.5}$$

where

$$\epsilon^2\nu^2=\mu/\int|\mathbf{x}_1-\mathbf{x}_3|^2\Psi_{13}\,d\mathbf{x}_3 \tag{3.6}$$

and

$$A=-\lim_{\epsilon\to0}\phi(1)\epsilon^{-2}. \tag{3.7}$$

The constant A is determined from the matching condition that when $|\mathbf{r}|\sim1$, $\alpha(|\mathbf{r}|)\sim-(1/kT)\phi(|\mathbf{r}|)$. It is easily verified that both terms on the left side of (3.4) are of the same order.

Also from (3.6), one has

$$\nu\sim\mu^{1/2}. \tag{3.8}$$

Now we can estimate the right side of (3.4) to be of the order $\epsilon^2\alpha_{12}$. Therefore, when $\mu\sim\epsilon^2$, the right side of (3.4) becomes significant. Finally it can be readily checked that if α_{12} is given by (3.5), $\nabla^2\alpha_{12}\sim\alpha_{12}/|\mathbf{x}_1-\mathbf{x}_2|^2$ as anticipated.

4. REGION OF PHASE CHANGE

In the last section we showed that α_{123}, even though small, makes a significant contribution to the leading order to α_{12} if $\mu\sim\epsilon^2$. This implies that the asymptotic analysis presented in Sec. 2 is not valid as $\mu\to0$ with ϵ. In this limit, the range of α_{12} becomes infinite and so does the compressibility. Hence we refer to the region $\mu\sim\epsilon^2$ as the "region of phase change." In this section we redo the asymptotics on the BBGKY hierarchy in this limit.

In particular, we study the behavior of the correlation functions as $\mu\to0$, when the particles are separated by distances much larger than the range of the potential.

The calculations presented in the last section indicate that in the equation for the s-particle correlation, one must retain the $(s+1)$-particle correlation. The correlation functions themselves are small in magnitude but because of their long range, integrals over them are significant. Therefore, if we seek a solution to the leading order for $\alpha_{\{j\}_1^j}$, we can still neglect all those terms containing products of $\alpha_{\{j\}_1^j}$ with other correlations. We can replace $\log(1+\alpha_{\{j\}_1^j})$ by $\alpha_{\{j\}_1^j}$. We should, however, retain the term containing $\alpha_{\{j+1\}_1^{j+1}}$ singly. Thus (1.4) and (1.5) for $s=3$ become in this limit,

$$\frac{\partial\alpha_{12}}{\partial\mathbf{x}_1}-n\int\frac{\partial\Psi_{13}}{\partial\mathbf{x}_1}\alpha_{23}\,d\mathbf{x}_3$$
$$=\frac{\partial\Psi_{12}}{\partial\mathbf{x}_1}+n\int\frac{\partial\Psi_{13}}{\partial\mathbf{x}_1}\alpha_{123}\,d\mathbf{x}_3+o\left(\frac{\partial\Psi_{12}}{\partial\mathbf{x}_1}\right), \tag{4.1}$$

$$\frac{\partial\alpha_{123}}{\partial\mathbf{x}_1}-n\int\frac{\partial\Psi_{14}}{\partial\mathbf{x}_1}\alpha_{234}\,d\mathbf{x}_4$$
$$=n\int\frac{\partial\Psi_{14}}{\partial\mathbf{x}_1}\alpha_{24}\alpha_{34}\,d\mathbf{x}_4+n\int\frac{\partial\Psi_{14}}{\partial\mathbf{x}_1}\alpha_{1234}\,d\mathbf{x}_4+o\left(\frac{\partial\alpha_{123}}{\partial\mathbf{x}_1}\right). \tag{4.2}$$

In general, for $s \geq 3$,

$$\frac{\partial \alpha_{\{s\}_1^s}}{\partial \mathbf{x}_1} - n \int \frac{\partial \Psi_{1,s+1}}{\partial \mathbf{x}_1} \alpha_{\{s\}_2^{s+1}} \, d\mathbf{x}_{s+1}$$
$$= n \int \frac{\partial \Psi_{1,s+1}}{\partial \mathbf{x}_1} \sum \prod \alpha_{\{p_j\}_2^s, s+1}$$
$$+ \frac{n}{2} \int \frac{\partial \Psi_{1,s+1}}{\partial \mathbf{x}_1} \alpha_{\{s+1\}_1^{s+1}} \, d\mathbf{x}_{s+1} + o\ (\alpha_{\{s+1\}_1^{s+1}}), \qquad (4.3)$$

where the summation is over the same products as given by (2.16). Notice that we have replaced $\partial \phi_{ij}/\partial \mathbf{x}_i$ $(1 + \alpha_{ij})$ by $\partial \Psi_{ij}/\partial \mathbf{x}_i$, because these two quantities are equal to the leading order, inside the hard core, and outside it, $\Psi_{ij} \sim \phi_{ij}$ and α_{ij} is small. We shall assume that $\mu \sim \epsilon^2$ and $|\mathbf{x}_1 - \mathbf{x}_2| \sim 1/\epsilon^2$. Let $L\alpha_{123} \equiv \alpha_{123} - n \int \Psi_{14} \alpha_{234} \, d\mathbf{x}_4$. Expanding α_{234} around α_{123} in a Taylor series it is readily seen that

$$L\alpha_{123} \sim \mu \alpha_{123} = o(\alpha_{123}). \qquad (4.4)$$

Hence to this order,

$$\int \frac{\partial \Psi_{13}}{\partial \mathbf{x}_1} \alpha_{123} \, d\mathbf{x}_3 = n^2 \iint \frac{\partial \Psi_{13}}{\partial \mathbf{x}_1} \Psi_{14} \alpha_{234} \, d\mathbf{x}_4 \, d\mathbf{x}_3$$
$$= n^2 \iint \frac{\partial \Psi_{14}}{\partial \mathbf{x}_1} \Psi_{13} \alpha_{234} \, d\mathbf{x}_4 \, d\mathbf{x}_3$$

(by interchanging 3 and 4)

$$= \frac{n^2}{2} \frac{\partial}{\partial \mathbf{x}_1} \iint \Psi_{13} \Psi_{14} \alpha_{234} \, d\mathbf{x}_4 \, d\mathbf{x}_3$$
$$= \frac{n}{2} \frac{\partial}{\partial \mathbf{x}_1} \int \Psi_{13} \alpha_{123} \, d\mathbf{x}_3 \ \text{[from (4.4)]}.$$

Substituting this in (4.1), and integrating it we have

$$\alpha_{12} - n \int \Psi_{13} \alpha_{23} \, d\mathbf{x}_3$$
$$= \Psi_{12} + \frac{n}{2} \int \Psi_{13} \alpha_{123} \, d\mathbf{x}_3 + o(\alpha_{12}). \qquad (4.5)$$

Similarly in (4.2) we replace α_{1234} by $n \int \Psi_{15} \alpha_{2345} \, d\mathbf{x}_5$, interchange 4 and 5, and integrate to obtain

$$\alpha_{123} - n \int \Psi_{14} \alpha_{234} \, d\mathbf{x}_4$$
$$= n \int \Psi_{14} \alpha_{24} \alpha_{34} \, d\mathbf{x}_4 + \frac{n}{2} \int \Psi_{14} \alpha_{1234} \, d\mathbf{x}_4 + o(\alpha_{123}). \qquad (4.6)$$

Just as we did in the last section, we expand α_{23} in (4.5) and α_{234} in (4.6) in Taylor series and retain the second derivative terms. The proof that in each case the remainder is small is the same as presented earlier. Then we have

$$\nabla_1^2 \alpha_{12} - \epsilon^4 \frac{\theta}{\lambda} \alpha_{12}$$
$$= -\frac{\epsilon^2}{\lambda} \Psi_{12} - \frac{\epsilon^2 n}{2\lambda} \int \Psi_{13} \alpha_{123} \, d\mathbf{x}_3 + o(\epsilon^2 \alpha_{12}), \qquad (4.6')$$

$$\nabla_1^2 \alpha_{123} - \epsilon^4 \frac{\theta}{\lambda} \alpha_{123}$$
$$= -\frac{\epsilon^2}{\lambda} n \int \Psi_{14} \alpha_{24} \alpha_{34} \, d\mathbf{x}_4 - \frac{\epsilon^2 n}{2\lambda} \int \Psi_{14} \alpha_{1234} \, d\mathbf{x}_4$$
$$+ o(\epsilon^2 \alpha_{123}). \qquad (4.7')$$

Here we have set

$$\mu = \epsilon^2 \theta, \quad \lambda = \frac{\epsilon^2}{2} n \int |\mathbf{x}_1 - \mathbf{x}_2|^2 \Psi_{12} \, d\mathbf{x}_2, \quad \theta \sim 1, \quad \lambda \sim 1. \qquad (4.7)$$

Define

$$G_{12}(\mathbf{x}_1 - \mathbf{x}_2) \equiv \frac{1}{4\pi |\mathbf{x}_1 - \mathbf{x}_2|} \exp[-\epsilon^2 (\theta/\lambda)^{1/2} |\mathbf{x}_1 - \mathbf{x}_2|]. \qquad (4.8)$$

Then (4.6′) and (4.7′) are solved by

$$\alpha_{12} = \frac{\epsilon^2}{\lambda} \int G_{13} \Psi_{23} \, d\mathbf{x}_3 + \frac{\epsilon^2 n}{2\lambda} \iint G_{14} \Psi_{34} \alpha_{234} \, d\mathbf{x}_3 \, d\mathbf{x}_4 + o(\alpha_{12}),$$

$$\alpha_{123} = \frac{\epsilon^2 n}{\lambda} \iint G_{15} \Psi_{45} \alpha_{24} \alpha_{34} \, d\mathbf{x}_4 \, d\mathbf{x}_5$$
$$+ \frac{\epsilon^2 n}{2\lambda} \iint G_{15} \Psi_{45} \alpha_{2345} \, d\mathbf{x}_4 \, d\mathbf{x}_5 + o(\alpha_{123}).$$

Note that G has a range of the order of $1/\epsilon^2$. This, together with the fact that $n \int \Psi_{12} \, d\mathbf{x}_2 = 1 + o(1)$, can be used to simplify the above equations as

$$\alpha_{12} = \frac{\epsilon^2}{\lambda n} G_{12} + \frac{\epsilon^2 n}{2\lambda} \iint G_{14} \Psi_{34} \alpha_{234} \, d\mathbf{x}_4 \, d\mathbf{x}_3 + o(\alpha_{12}), \qquad (4.9)$$

$$\alpha_{123} = \frac{\epsilon^2}{\lambda} \int G_{14} \alpha_{24} \alpha_{34} \, d\mathbf{x}_4$$
$$+ \frac{\epsilon^2 n}{2\lambda} \iint G_{15} \Psi_{45} \alpha_{2345} \, d\mathbf{x}_4 \, d\mathbf{x}_5 + o(\alpha_{123}). \qquad (4.10)$$

In general for $s \geq 3$,

$$\alpha_{\{s\}_1^s} = -\frac{\epsilon^2}{\lambda} \int G_{1,s+1} \sum \prod \alpha_{\{p_j\}_2^s, s+1} \, d\mathbf{x}_{s+1}$$
$$- \frac{\epsilon^2 n}{2\lambda} \iint G_{1,s+2} \Psi_{s+1,s+2} \alpha_{\{s+1\}_2^{s+2}} \, d\mathbf{x}_{s+1} \, d\mathbf{x}_{s+2}$$
$$+ o(\alpha_{\{s\}_1^s}), \qquad (4.11)$$

where the summation is over all products given by (2.16).

It can be easily verified that if α_{123} is of the same order as the first term on the right side of (4.10), then the integral term in (4.9) is of the same order as α_{12} and hence cannot be ignored. Similarly, the equation for α_{1234} can be written as

$$\alpha_{1234} = \frac{\epsilon^2}{\lambda} \int G_{15} (\alpha_{25} \alpha_{35} \alpha_{45} + \alpha_{25} \alpha_{345} + \alpha_{35} \alpha_{245} + \alpha_{45} \alpha_{235}) \, d\mathbf{x}_5$$
$$+ \frac{\epsilon^2 n}{2\lambda} \iint G_{16} \Psi_{56} \alpha_{23456} \, d\mathbf{x}_5 \, d\mathbf{x}_6 + o(\alpha_{1234}). \qquad (4.12)$$

By following an argument similar to the one in Sec. 2, we presume that

$$\alpha_{12} \sim \epsilon^2 G_{12} \sim \epsilon^4 \ \text{when} \ |\mathbf{x}_1 - \mathbf{x}_2| \sim \epsilon^{-2}, \qquad (4.13)$$

$$\alpha_{123} \sim \epsilon^2 \int G_{14}\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 \sim \epsilon^6. \tag{4.14}$$

Then in the equation for α_{1234}, $\alpha_{25}\alpha_{35}\alpha_{45} \sim \epsilon^{12}$, whereas $\alpha_{25}\alpha_{345} \sim \epsilon^{10}$ and hence the former can be neglected and we have $\alpha_{1234} \sim \epsilon^8$. This is the same order as when all the particles are separated by distances of order 1. Now we can show inductively that $\alpha_{\{s\}_1^s} \sim \epsilon^{2s}$. Consider a product $\prod \alpha_{\{p_j\}_{2,s+1}^s}$ of Eq. (4.12). Suppose that this product has n terms with indices $p_1, \ldots, p_n$ respectively, $2 \leq p_k \leq s-2$ and $\sum_{k=1}^n p_k = s-1$, so that it satisfies all the conditions given for (2.16). If we assume that $\alpha_{\{j\}_1^j} \sim \epsilon^{2j}$ for $j = 2, \ldots, s-1$, the order of this product is $\epsilon^{2(s-1+n)}$. Since $p_k \leq s-2$ for each k, this quantity is largest when $n=2$, i.e., when the product has only two terms. Therefore, (4.11) is simplified to

$$\alpha_{\{s\}_1^s} = \frac{\epsilon^2}{\lambda} \int G_{1,s+1} \sum_{j=2}^{s} \alpha_{j,s+1}\, \alpha_{\{s-2/j'\}_2^s,\, s+1} + \frac{\epsilon^2 n}{2\lambda} \iint G_{1,s+2}\Psi_{s+1,s+2}\, \alpha_{\{s+1\}_2^{s+2}}\, d\mathbf{x}_{s+1}\, d\mathbf{x}_{s+2} + o(\alpha_{\{s\}_1^s}). \tag{4.15}$$

Here $\{s-2/j'\}_2^s$ stands for the set of $s-2$ particles obtained by removing the jth particle from the set $2, 3, \ldots, s$.

Using the fact that G has a range of the order of $1/\epsilon^2$ it is immediately seen that $\alpha_{\{s\}_1^s} \sim \epsilon^{2s}$. This completes the induction argument. It can also be easily verified that the integral over the $(s+1)$-particle term is of the order ϵ^{2s}. For let $|\mathbf{x}_1 - \mathbf{x}_{s+2}| \sim 1/\epsilon^2$, $|\mathbf{x}_{s+1} - \mathbf{x}_{s+2}| \sim 1/\epsilon$. Then since the $(s+1)$-particle correlation is of the order $\epsilon^{(2s+2)}$, the integral term containing it in (4.15) is of the order

$$\epsilon^2 \circ \frac{1}{\epsilon^6} \circ \epsilon^2 \circ \frac{1}{\epsilon^3} \circ \epsilon^3 \circ \epsilon^{(2s+2)} = \epsilon^{2s}.$$

As an example, suppose that the attractive potential is of the form $-\epsilon^2 A \exp(-\epsilon\beta r)/r$. From the estimate (2.12) we see that, when the system is away from the region of phase change, if every pair of particles is separated by a distance of order 1, $\alpha_{\{s\}_1^s} \sim \epsilon^{2s}$ for $s > 2$. From this value it decays to ϵ^{3s-3} as the particle separations increase to the order of ϵ^{-1}. Near phase transition, however, it stays to be of the same order even when every pair of masses is separated by a very large distance, viz., of the order ϵ^{-2}. This suggests that for $s \geq 3$, $\alpha_{\{s\}_1^s}$ is a very slowly varying function of its arguments. Note that this is not necessarily true for $s=2$, because the pair correlation is qualitatively like a Debye potential and its order varies with the separation. This property of extremely slow variation of the three-body and higher correlations can be exploited to simplify the hierarchy (4.15). We expand these higher correlations around configurations obtained by replacing $\mathbf{x}_{s+1}$ by $\mathbf{x}_1$. Further, using the fact that $\mu \ll 1$, we obtain the following:

$$\alpha_{12} = \frac{\epsilon^2}{\lambda n} G_{12} + \frac{\epsilon^2}{2\lambda} \int G_{13}\alpha_{123}\, d\mathbf{x}_3 + o(\alpha_{12}), \tag{4.16}$$

$$\alpha_{123} = \frac{\epsilon^2}{\lambda} \int G_{14}\alpha_{24}\alpha_{34}\, d\mathbf{x}_4 + \frac{\epsilon^2}{2\lambda} \int G_{14}\alpha_{1234}\, d\mathbf{x}_4 + o(\alpha_{123}), \tag{4.17}$$

and for $s > 3$,

$$\alpha_{\{s\}} = \frac{\epsilon^2}{\lambda} \sum_{j=2}^{s} \alpha_{\{s-1/j'\}_1^s} \int G_{1,s+1}\alpha_{j,s+1}\, d\mathbf{x}_{s+1} + \frac{\epsilon^2}{2\lambda} \int G_{1,s+1}\alpha_{\{s+1\}_1^{s+1}}\, d\mathbf{x}_{s+1} + o(\alpha_{\{s\}_1^s}). \tag{4.18}$$

By truncating this hierarchy at various levels by neglecting $\alpha_{\{s+1\}_1^{s+1}}$ it is possible to solve all the equations in terms of the pair correlation and obtain single nonlinear equations. The convergence and/or asymptoticity of this procedure as $s \to \infty$ is being investigated. Convergence may not be as relevant here as asymptoticity. If $\epsilon^2 \ll \mu \ll 1$, it might be asymptotic.

As an example, suppose one neglects α_{1234}. We then get, on substituting for α_{123} in (4.16),

$$\alpha_{12} = \frac{\epsilon^2}{\lambda n} G_{12} + \int C_{13}\alpha_{23}\, d\mathbf{x}_3, \tag{4.19}$$

where

$$C_{12} = \frac{\epsilon^4}{2\lambda^2} G_{12} \int G_{13}\alpha_{23}\, d\mathbf{x}_3. \tag{4.20}$$

CONCLUSION

We have derived a hierarchy for correlations in a region near change of phase in a simple fluid. This is an infinite chain of equations given by (4.16)–(4.18). Each correlation function is expressed in terms of its immediate predecessor, immediate successor, and the pair correlation function. We have shown that the magnitude of the correlation functions becomes successively smaller as one includes more and more particles but their range is very large. As a result *all* correlations contribute to the leading order in the computation of the pair correlation. It is also shown that the correlation functions vary extremely slowly over all space. It may be possible to utilize this fact to obtain approximate solutions to the hierarchy. One could conceivably obtain formal solutions of it by truncating it at the sth level and letting s go to infinity. But in order to do this one perhaps should approach the region from large values of θ [defined in Eq. (4.7)]—i.e., approach from the gaseous state—and obtain an asymptotic expansion in terms of θ. Work along these lines is in progress.

* Supported by a grant from the National Science Foundation, NSF75 07391.

[1] G. V. Ramanathan, J. Math. Phys. 7, 1507 (1966).
[2] M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. 4, 216 (1963).
[3] P. C. Hemmer, J. Math. Phys. 5, 75 (1964).
[4] G. Stell and W. K. Theumann, Phys. Rev. 186, 581 (1969).
[5] J. S. Høye and G. Stell, Mol. Phys. 32, 195 (1976).
[6] E. Hiis Hauge and P. C. Hemmer, J. Chem. Phys. 45, 323 (1966).
[7] G. Stell, *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz (Benjamin, New York, 1964).
[8] M. Fisher, J. Math. Phys. 5, 944 (1966).
[9] G. V. Ramanathan, J. M. Dawson, and M. D. Kruskal, J. Math. Phys. 11, 339 (1970).
[10] G. Stell, Physica 29, 517 (1963).
[11] N. N. Bogoliubov, *Studies in Statistical Mechanics*, edited by J. de Boer and G. Uhlenbeck (North-Holland, Amsterdam, 1962).
[12] S. G. Mikhlin, *Integral Equations* (McMillan, New York, 1964).

Structure of the Azzarelli–Collas representation for the scattering amplitude and generalization to the Rice representation and the Euler–Pochhammer representation

Alfred C. T. Wu

Department of Physics, University of Michigan, Ann Arbor, Michigan 48109
(Received 11 April 1977)

The Azzarelli–Collas integral representation for the scattering amplitude is extended to the Rice representation. The latter representation in turn is viewed as an example of the Euler–Pochhammer representation.

I. INTRODUCTION

Recently, Azzarelli and Collas have derived a new integral representation for the scattering amplitude.[1] The representation originated from their study of Coulomb scattering in the parabolic coordinate system where the Schrödinger equation is separable and the wavefunctions are expressible as a superposition of products of Whittaker functions. A general derivation is obtained[1] by recasting, in the partial wave expansion, the Legendre function in terms of the Bateman function.[2] It appears that the partial wave amplitude smeared over by the Bateman function is a much smoother function in the new variable. Usual Regge poles in the angular momentum plane may now appear however as zeroes in this new formalism.[1]

We study here the structure of the Azzarelli–Collas integral representation and the generalization thereof.

The Azzarelli–Collas representation may be viewed as a special Mellin transform. It so happens that the same relationship exists between the Bateman function and the Legendre function.[3] This last fact permits the kernel in the integral representation (or the expansion coefficients in the power series expansion via a Sommerfeld–Watson transformation) to be simply related to the partial wave amplitude.

In their derivation of the integral representation, Azzarelli–Collas used an identity between the Legendre function and the Bateman function, and cited the reference to Rice.[3] Rice has actually studied a generalization of the Bateman function, which we shall refer to as the Rice function. The use of the Rice function in place of the Bateman function gives an immediate generalization of the Azzarelli–Collas representation to a new representation which we shall refer to as the Rice representation.

The Rice representation may be viewed further as an example of the Euler–Pochhammer representation.[4] Previously, other examples of the generalized Euler–Pochhammer representations have been found in the study of Feynman amplitudes,[5] Veneziano amplitudes,[6] angular momentum recoupling coefficients,[7] and SU(N) basis functions.[8]

For the sake of readability, we give briefly in Sec. II a derivation of the Azzarelli–Collas representation from first principles. In Sec. III, we emphasize the Euler–Pochhammer aspect of the representation. In Sec. IV, we generalize the Azzarelli–Collas representation to the Rice representation. The Euler–Pochhammer structure is manifest.

Although the Azzarelli–Collas representation owes its origin to the analysis of the Coulomb scattering in the parabolic coordinates, the heuristic argument presented in Ref. 1 regarding the asymptotic behavior of the Whittaker functions does not seem to provide a deeper insight to the *raison d'être* of the integral representation. One possible interpretation from the group-theoretic point of view would be as follows. The spherical functions pertaining to the rank-one O(3) group and labelled by the angular momentum l (beside the third component label m) belong to the family of the Gauss hypergeometric function [such as $P_l(z) = {}_2F_1(-l, l+1; 1; (1-z)/2]$. The case of the Coulomb scattering admits, as is well known, the group O(4), which is of rank two. Therefore, it is quite conceivable to arrange, at least, for some of the representation functions to correspond to the higher-hierarchy generalized hypergeometric functions [such as ${}_3F_2(-l, l+1, \nu; 1, p; z)$ whence the Bateman and the Rice functions] to accommodate the extra label. In this spirit, it would be interesting to view the generalized Euler–Pochhammer integral representation as natural for the cases of underlying symmetry groups of higher ranks.

II. AZZARELLI-COLLAS INTEGRAL REPRESENTATION AS A MELLIN TRANSFORM

The Azzarelli–Collas integral representation for the scattering amplitude can be derived from first principles as follows.

Step 1: Expand the scattering amplitude $A(s, z)$ into partial wave amplitudes

$$A(s, z) = \sum_{l=0}^{\infty} (2l+1)\, a_l(s)\, P_l(z). \tag{1}$$

Step 2: Recognize that the Legendre function P_l is a special case of the Gauss hypergeometric function ${}_2F_1$,

$$P_l(z) = {}_2F_1\left(-l,\ l+1;\ 1;\ \tfrac{1}{2}(1-z)\right) \tag{2}$$

and that the Bateman function F_l is a special case of the ${}_3F_2$ function,

$$F_l(-2\nu-1) = {}_3F_2(-l,\ l+1,\ \nu;\ 1, 1; 1). \tag{3}$$

Step 3: Recall that there is a Euler–Pochhammer integral relation between ${}_{A+1}F_{B+1}$ and ${}_AF_B$.[4]

$$
{}_{A+1}F_{B+1}((a),\ c;\ (b),\ d;w)
$$

$$
=\frac{\Gamma(d)}{\Gamma(c)\Gamma(d-c)}\int_0^1 dt\, t^{c-1}(1-t)^{d-c-1}\quad {}_AF_B((a);(b);wt), \tag{4}
$$

where (a) and (b) are shorthand notations for the two sets of parameters $a_1, \ldots, a_A$ and $b_1, \ldots, b_B$ respectively. Applying (4) to (3) and (2) for the case $A=2$ and $B=1$ gives the Euler–Pochhammer representation for the Bateman function, namely

$$
F_l(-2\nu-1)
$$

$$
=-\frac{\sin\pi\nu}{\pi}\int_0^1 dt\, t^{-\nu-1}(1-t)^{\nu}\ {}_2F_1(-l,\ l+1;1;t)
$$

$$
=-\frac{\sin\pi\nu}{\pi}\int_0^1 dt\, t^{-\nu-1}(1-t)^{\nu}P_l(1-2t) \tag{5}
$$

$$
=-\frac{\sin\pi\nu}{\pi}\int_0^\infty d\tau\,\tau^{-\nu-1}(1+\tau)^{-1}P\left(\frac{1-\tau}{1+\tau}\right), \tag{6}
$$

where $t=\tau/(1+\tau)$ in going from (5) to (6). Equation (5) is complementary to the contour representation Eq. (B1) quoted in Ref. 1.

Step 4: Equation (6) may be viewed as the Mellin transform of $(1+\tau)^{-1}P((1-\tau)/(1+\tau))$. The inverse Mellin transform gives

$$
(1+\tau)^{-1}P\left(\frac{1-\tau}{1+\tau}\right)
$$

$$
=\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}d\nu\,\tau^{\nu}\left(\frac{-\pi F_l(-2\nu-1)}{\sin\pi\nu}\right), \tag{7}
$$

where $(1-\tau)/(1+\tau)=z=\cos\theta$. This is the Eq. (3) quoted in Ref. 1 and attributed to Rice.[3]

Step 5: Substituting (7) into (1) gives the Azzarelli–Collas representation

$$
(1+\tau)^{-1}A\left(s,\frac{1-\tau}{1+\tau}\right)
$$

$$
=\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}d\nu\,\tau^{\nu}\left(\frac{-\pi b(s,\nu)}{\sin\pi\nu}\right) \tag{8a}
$$

$$
=\sum_{\nu=0}^{\infty}b(s,\nu)(-\tau)^{\nu}, \tag{8b}
$$

$$
\frac{-\pi b(s,\nu)}{\sin\pi\nu}=\int_0^\infty d\tau\,\tau^{-\nu-1}(1+\tau)^{-1}A\left(s,\frac{1-\tau}{1+\tau}\right). \tag{8c}
$$

Equations (8a) and (8c) may be viewed simply as the Mellin transforms between $(1+\tau)^{-1}A(s,(1-\tau)/(1+\tau))$ and $-\pi b(s,\nu)/\sin\pi\nu$. This $b(s,\nu)$ is simply related to the partial wave amplitudes $a_l(s)$ in the following sense.[1] As the folded sum of $a_l(s)$ with the Legendre functions $P_l(z)$ gives $A(s,z)$, the folded sum of $a_l(s)$ with the Bateman function $F_l(-2\nu-1)$ gives $b(s,\nu)$,

$$
b(s,\nu)=\sum_{l=0}^{\infty}(2l+1)a_l(s)F_l(-2\nu-1). \tag{9}
$$

Equations (8b) and (8a) are related by a Sommerfeld–Watson transformation. A convenient way to remember the structure of the Azzarelli–Collas representation (8) is that setting $b(s,\nu)\equiv 1$ formally in (8b) gives $A=1$, and Eqs. (8a) and (8c) then merely confirm the fact that $(1+\tau)^{-1}$ and $\pi/\sin\pi\nu$ are a Mellin-transform pair.

III. EXTENSION TO THE EULER-POCHHAMMER REPRESENTATION

In terms of the variable $t=\tau/(1+\tau)$, Eq. (8c) reads

$$
b(s,\nu)=\frac{\Gamma(1)}{\Gamma(-\nu)\,\Gamma(1+\nu)}
$$

$$
\times\int_0^1 dt\, t^{-\nu-1}(1-t)^{\nu}A(s,1-2t). \tag{10}
$$

This is a special case of the Euler–Pochhammer representation

$$
g(z;c,d)=\frac{\Gamma(d)}{\Gamma(c)\,\Gamma(c-d)}
$$

$$
\times\int_0^1 dt\, t^{c-1}(1-t)^{d-c-1}f(t;z) \tag{11}
$$

evaluated at $c=-\nu$, $d=1$.

Examples of the representation (11) and the generalization thereof to multiple integrals have been cited by the statements in the Introduction. It is expected that such a generalized Euler–Pochhammer integral representation would occur, for example, in the multiple scattering.

IV. THE RICE REPRESENTATION

In place of the Bateman function (3), the Rice function[3] is given by

$$
R_l(\nu,p,z)\equiv {}_3F_2(-l,l+1,\ \nu;1,p;z). \tag{12}
$$

The Bateman function is recovered in the limit $p=1$ and $z=1$,

$$
R_l(\nu;1;1)=F_l(-2\nu-1). \tag{13}
$$

The remarkable feature of Rice's analysis is that the integral relation (4) between ${}_3F_2$ and ${}_2F_1$ is actually invertible, namely

$$
{}_2F_1(a_1,a_2;b_1;z)
$$

$$
=\frac{\Gamma(a_3)}{\Gamma(b_2)\,\Gamma(a_3-b_2)}\int_0^1 du\, u^{b_2-1}(1-u)^{a_3-b_2-1}
$$

$$
\times {}_3F_2(a_1,a_2,a_3;b_1,b_2;zu). \tag{14}
$$

This implies that from the relations (1), (2), and (14), we have

$$
A(s,1-2w)
$$

$$
=\sum_l(2l+1)a_l(s)P_l(1-2w)
$$

$$
=\frac{\Gamma(\nu)}{\Gamma(p)\Gamma(\nu-p)}\int_0^1 du\, u^{p-1}(1-u)^{\nu-p-1}r(s,\nu,p;wu), \tag{15}
$$

$$
r(s,\nu,p;y)
$$

$$
=\sum_l(2l+1)a_l(s)\,R_l(\nu,p,y) \tag{16}
$$

$$= \frac{\Gamma(p)}{\Gamma(\nu)\Gamma(\nu-p)} \int_0^1 dt\, t^{\nu-1}(1-t)^{p-\nu-1} A(s, 1-2yt). \tag{17}$$

Equation (16) is the generalization of (9) while Eq. (17) is that of (10). On the other hand, unlike the Azzarelli–Collas reprentation (8a) which is merely a Mellin transform, Eq. (15) is also in the form of the Euler–Pochhammer representation.

[1]T. Azzarelli and P. Collas, Phys. Rev. D **12**, 3237 (1975).

[2]*Higher Transcendental Functions* (Bateman Manuscript Project), edited by A. Erdelyi, *et al.* (McGraw-Hill, New York, 1953), Vol. 1, esp. p. 192.

[3]S.O. Rice, Duke Math. J. **6**, 108 (1939).

[4]See, e.g., J.J. Slater, *Generalized Hypergeometric Functions* (Cambridge U.P., Cambridge, 1966).

[5]A.C.T. Wu, Phys. Rev. D **9**, 370 (1974).

[6]A.C.T. Wu, J. Math. Phys. **12**, 2035 (1971).

[7]A.C.T. Wu, J. Math. Phys. **14**, 1222 (1973).

[8]A.C.T. Wu, J. Math. Phys. **12**, 437 (1971); C.S. Huang and A.C.T. Wu, J. Math. Phys. **14**, 263 (1973).

The quantum mechanical representations of the anisotropic harmonic oscillator group

M. E. Major

Mathematics Department, Queen Elizabeth College, London W8 7AH, England
(Received 22 September 1975)

A group G associated with the n-dimensional anisotropic harmonic oscillator is constructed : G is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system. All the quantum mechanical irreducible representations of G are evaluated, using Mackey's theory of induced representations.

1. INTRODUCTION

In an attempt to gain a greater understanding of accidental degeneracy in quantum theory, a study has been made of the n-dimensional anisotropic harmonic oscillator.

The basic object studied is a group, G, associated with the n-dimensional anisotropic harmonic oscillator: G is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system.

In the present paper, all the quantum mechanical irreducible representations of the group G are evaluated (up to unitary equivalence). This is done by using Mackey's theory of induced representations. The representations of G calculated here are used in a later paper where the accidental degeneracy of the n-dimensional anisotropic harmonic oscillator is discussed.

In Sec. 2, the Lie algebra $\mathcal{L}$ of G is defined. Next, it is shown how a heuristic construction leads from $\mathcal{L}$ to the definition of G, and its simply connected covering group $\tilde{G}$. Some properties of the groups G and $\tilde{G}$ are listed in Sec. 4. Those parts of Mackey's theory of induced representations needed to construct the representations of G and $\tilde{G}$ are summarized in Sec. 5. The theory is then applied to finding the representations of $\tilde{G}$, and hence those of G.

2. DEFINITION OF THE LIE ALGEBRA $\mathcal{L}$ OF G

The quantum mechanical n-dimensional anisotropic harmonic oscillator is a particle of mass m with Hamiltonian

$$H_{\rm osc} = \sum_{j=1}^{n} \frac{1}{2m}(P_j^2 + m^2\omega_j^2 Q_j^2),$$

where $\omega_1, \omega_2, \ldots, \omega_n$ are positive constants (the frequencies), and $\{Q_j : j = 1, 2, \ldots, n\}$ and $\{P_j : j = 1, 2, \ldots, n\}$ are sets of self-adjoint operators corresponding to position and momentum observables respectively, and satisfying the commutation relations

$$[Q_j, P_k] = i\hbar\,\delta_{jk} I \quad (1 \leq j, k \leq n).$$

To avoid unimportant constants, and to construct a Lie algebra associated with these operators, define the skew-adjoint operators:

$$E = iI, \quad B_j = i\left(\frac{m\omega_j}{\hbar}\right)^{1/2} Q_j,$$

$$A_j = i\left(\frac{1}{m\hbar\omega_j}\right)^{1/2} P_j, \quad T = i\frac{H_{\rm osc}}{\hbar}.$$

Then

$$\begin{aligned} &[A_j, B_k] = \delta_{jk} E, \quad [T, B_k] = \omega_k A_k, \\ &[T, A_k] = -\omega_k B_k, \quad [B_j, E] = 0, \\ &[A_j, E] = 0, \quad [T, E] = 0. \end{aligned} \tag{1}$$

Therefore, $E, B_1, B_2, \ldots, B_n, A_1, A_2, \ldots, A_n$, and T form a basis of the Lie algebra defined by the above commutation relations. This Lie algebra will be denoted by $\mathcal{L}$.

3. HEURISTIC CONSTRUCTION OF THE GROUPS G AND $\tilde{G}$

Every real Lie algebra is isomorphic to a subalgebra of some $\mathrm{gl}(m, \mathrm{R})$ [where $\mathrm{gl}(m, \mathrm{R})$ denotes the Lie algebra of all real $m \times m$ matrices]; therefore, there is a matrix representation of the Lie algebra $\mathcal{L}$ by elements of some $\mathrm{gl}(m, \mathrm{R})$.[1]

One straightforward representation to use is the adjoint representation. However, since E commutes with $E, B_1, B_2, \ldots, B_n, A_1, A_2, \ldots, A_n$, and T, it follows that $\mathrm{ad}E$ is the $(2n+2)\times(2n+2)$ zero matrix. A significant property of the Lie algebra $\mathcal{L}$ is that $[A_j, B_k] = \delta_{jk}E$, where E is possibly not the zero operator. Hence many important features of the Lie algebra $\mathcal{L}$ would be lost if the adjoint representation were used. It is therefore necessary to look for some other matrix representation of $\mathcal{L}$.

Let E_{jk} denote the $(2n+2)\times(2n+2)$ matrix with 1 at the intersection of the jth row and kth column, and zeros elsewhere. Define a mapping $\tilde{}$ of the chosen basis elements of $\mathcal{L}$ into $\mathrm{gl}(2n+2, \mathrm{R})$ by

$$\begin{aligned} &E \to \tilde{E} = 2E_{2n+1,2n+2}, \\ &B_j \to \tilde{B}_j = -E_{2n+1,j} + E_{n+j,2n+2} \quad (1 \leq j \leq n), \\ &A_j \to \tilde{A}_j = E_{2n+1,n+j} + E_{j,2n+2} \quad (1 \leq j \leq n), \\ &T \to \tilde{T} = \sum_{j=1}^{n} \omega_j (E_{j,n+j} - E_{n+j,j}). \end{aligned}$$

The mapping $\tilde{}$ may be extended by linearity to the whole of $\mathcal{L}$. It is easily verified that the mapping $\tilde{}$ gives a matrix representation of the Lie algebra $\mathcal{L}$.

The definition of the mapping ~ and the following heuristic construction of G are generalizations of ideas used by Streater[2] for the one-dimensional case.

Now consider the set of matrices formed by taking products of exponentials of the representative basis elements. A typical element of this set is

$$(\sigma,\beta,\alpha,\tau)=\exp(\sigma\tilde{E})\exp(\beta_1\tilde{B}_1)\exp(\beta_2\tilde{B}_2)\cdots\exp(\beta_n\tilde{B}_n)\times\exp(\alpha_1\tilde{A}_1)\exp(\alpha_2\tilde{A}_2)\cdots\exp(\alpha_n\tilde{A}_n)\exp(\tau\tilde{T}),$$

where $\sigma\in\mathrm{R}$, $\beta=(\beta_1,\beta_2,\ldots,\beta_n)\in\mathrm{R}^n$, $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathrm{R}^n$, and $\tau\in\mathrm{R}$. (The notation and the order of the exponentials have been chosen to fit in with those required in a later paper.)

A straightforward calculation, using the explicit expressions for $\tilde{E},\tilde{B}_1,\tilde{B}_2,\ldots,\tilde{B}_n,\tilde{A}_1,\tilde{A}_2,\ldots,\tilde{A}_n,\tilde{T}$, gives that

$$(\sigma,\beta,\alpha,\tau)=\left[\begin{array}{c|c|c|c}
\begin{matrix}\cos\omega_1\tau & & \\ & \ddots & \\ & & \cos\omega_n\tau\end{matrix} & \begin{matrix}\sin\omega_1\tau & & \\ & \ddots & \\ & & \sin\omega_n\tau\end{matrix} & \begin{matrix}0\\ \vdots\\ 0\end{matrix} & \begin{matrix}\alpha_1\\ \vdots\\ \alpha_n\end{matrix}\\
\hline
\begin{matrix}-\sin\omega_1\tau & & \\ & \ddots & \\ & & -\sin\omega_n\tau\end{matrix} & \begin{matrix}\cos\omega_1\tau & & \\ & \ddots & \\ & & \cos\omega_n\tau\end{matrix} & \begin{matrix}0\\ \vdots\\ 0\end{matrix} & \begin{matrix}\beta_1\\ \vdots\\ \beta_n\end{matrix}\\
\hline
\cdots,-\beta_j\cos\omega_j\tau-\alpha_j\sin\omega_j\tau,\cdots & \cdots,-\beta_j\sin\omega_j\tau+\alpha_j\cos\omega_j\tau,\cdots & 1 & 2\sigma-\alpha.\beta\\
\hline
0\ \ 0\ \cdots\ 0 & 0\ \ 0\ \cdots\ 0 & 0 & 1
\end{array}\right] \tag{2}$$

The composition law on this set of matrices is taken to be matrix multiplication. It may be shown that

$$(\sigma',\beta',\alpha',\tau')(\sigma,\beta,\alpha,\tau)=(\sigma'',\beta'',\alpha'',\tau''),$$

where

$$\sigma''=\sigma'+\sigma+\sum_{j=1}^{n}\left[\frac{\beta_j^2-\alpha_j^2}{2}\cos\omega_j\tau'\sin\omega_j\tau'-\alpha_j\beta_j\sin^2\omega_j\tau'+\alpha_j'(\beta_j\cos\omega_j\tau'-\alpha_j\sin\omega_j\tau')\right],$$

$$\beta_j''=\beta_j\cos\omega_j\tau'-\alpha_j\sin\omega_j\tau'+\beta_j',$$

$$\alpha_j''=\beta_j\sin\omega_j\tau'+\alpha_j\cos\omega_j\tau'+\alpha_j',$$

and τ'' satisfies

$$\cos\omega_j\tau''=\cos\omega_j(\tau'+\tau)$$
$$\sin\omega_j\tau''=\sin\omega_j(\tau'+\tau)\quad(1\leq j\leq n). \tag{3}$$

The ambiguity in the definition of τ'' arises from the fact that a Lie algebra determines a Lie group only up to local isomorphism. From (3), τ'' must be expressible in the form

$$\tau''=\tau'+\tau+2\pi r_j/\omega_j,$$

for each j, where $r_1,r_2,\ldots,r_n$ are appropriate integers. Hence the r_j's must satisfy

$$\frac{r_k}{\omega_k}=\frac{r_l}{\omega_l},\quad r_k,\ r_l\in\mathrm{Z}\quad(1\leq k,l\leq n).$$

When each pair of ω_j's is rationally related, it follows that condition (3) is satisfied if and only if the common value of r_k/ω_k $(1\leq k\leq n)$ lies in the set $\{\lambda s: s\in\mathbf{Z}\}$, where λ is the smallest positive number which is integral on multiplication by any ω_j. Hence τ'' is determined modulo $2\pi\lambda$ by τ,τ' and condition (3). Also, in the matrix representation of $(\sigma,\beta,\alpha,\tau)$, values of τ which differ by integral multiples of $2\pi\lambda$ give the same matrix.

When each pair of ω_j's is rationally related, this suggests regarding τ and τ' as elements of $(-\pi\lambda,\pi\lambda]$, and defining $\tau''\in(-\pi\lambda,\pi\lambda]$ by $\tau''\equiv\tau'+\tau(\mathrm{mod}2\pi\lambda)$. $(\sigma,\beta,\alpha,\tau)$ is then regarded as an element of the set $\mathrm{R}\times\mathrm{R}^n\times\mathrm{R}^n\times(-\pi\lambda,\pi\lambda]$, rather than as a matrix.

An alternative procedure for removing the ambiguity in τ'' is to define the set $\tilde{G}$ to be

$$\{(\sigma,\beta,\alpha,\tau):\sigma\in\mathrm{R},\quad\beta,\alpha\in\mathrm{R}^n,\quad\tau\in\mathrm{R}\}$$

and to take $\tau''=\tau'+\tau$ in the composition law.

If any two of the ω_j's are not rationally related, each r_j can take only the value zero, and τ'' is then uniquely determined as $\tau'+\tau$. In this case, each $\tau\in\mathrm{R}$ defines a different matrix. This suggests defining G (and $\tilde{G}$) to be the set

$$\{(\sigma,\beta,\alpha,\tau):\sigma\in\mathrm{R},\quad\beta,\alpha\in\mathrm{R}^n,\quad\tau\in\mathrm{R}\}.$$

4. DEFINITION AND PROPERTIES OF G AND $\tilde{G}$

If each pair of ω_j's is rationally related, the set G is defined to be

$$G=\{(\sigma,\beta,\alpha,\tau):\sigma\in\mathbf{R},\ \beta,\alpha\in\mathbf{R}^n,\ \tau\in(-\pi\lambda,\pi\lambda]\},\qquad(4)$$

with λ the smallest positive number such that each $\lambda\omega_j$ is integral.

If any two of the ω_j's are not rationally related, the set G is taken to be the set $\tilde{G}$ defined below.

Whatever the values of the ω_j's, the set $\tilde{G}$ is defined to be

$$\tilde{G}=\{(\sigma,\beta,\alpha,\tau):\sigma\in\mathbf{R},\ \beta,\alpha\in\mathbf{R}^n,\ \tau\in\mathbf{R}\}\ .\qquad(5)$$

Define the composition laws for G and $\tilde{G}$ by

$$(\sigma',\beta',\alpha',\tau')(\sigma,\beta,\alpha,\tau)=(\sigma'',\beta'',\alpha'',\tau'')\ ,\qquad(6)$$

where, for G and $\tilde{G}$,

$$\sigma''=\sigma'+\sigma+\sum_{j=1}^{n}\left[\frac{\beta_j^2-\alpha_j^2}{2}\cos\omega_j\tau'\sin\omega_j\tau'\right.$$

$$\left.-\alpha_j\beta_j\sin^2\omega_j\tau'+\alpha_j'(\beta_j\cos\omega_j\tau'-\alpha_j\sin\omega_j\tau')\right]\ ,$$

$$\beta_j''=\beta_j\cos\omega_j\tau'-\alpha_j\sin\omega_j\tau'+\beta_j'\ ,$$

$$\alpha_j''=\beta_j\sin\omega_j\tau'+\alpha_j\cos\omega_j\tau'+\alpha_j'\ ,$$

and, for G,

$$\tau''\equiv\tau'+\tau(\mathrm{mod}2\pi\lambda)\ ,$$

whereas, for $\tilde{G}$,

$$\tau''=\tau'+\tau\ .$$

It is readily verified that, with these composition laws, the sets G and $\tilde{G}$ are groups.

Furthermore, when the ω_j's are rationally related, G is a homomorphic image of $\tilde{G}$ under the map

$$\nu:(\sigma,\beta,\alpha,\tau)\ (\in\tilde{G})\to(\sigma,\beta,\alpha,\hat{\tau})\ (\in G)\ ,$$

where $\hat{\tau}\in(-\pi\lambda,\pi\lambda]$, and is given by $\hat{\tau}\equiv\tau\mathrm{mod}2\pi\lambda$. Hence

$$G\approx\tilde{G}/Z\ ,\qquad(7)$$

where $Z=\{(0,0,0,2\pi\lambda s):s\in\mathbf{Z}\}\approx\mathbf{Z}$, with $\approx$ denoting isomorphism.

$\tilde{G}$ is given the usual topology of $\mathbf{R}^{2n+2}$, and, when the ω_j's are rationally related, the topology on G is then chosen to be that inherited from $\tilde{G}$ under the above map $\nu:\tilde{G}\to G$. It follows that $\tilde{G}$ and G are connected Lie groups.

It remains to check that the Lie algebras of $\tilde{G}$ and G are isomorphic to the Lie algebra $\mathcal{L}$.

The left-invariant vector field X_μ, say, corresponding to the one-dimensional Lie subgroup $\mu(t)$ of a Lie group L is given by

$$(X_\mu\phi)(l)=\frac{\partial}{\partial t}\,\phi(l_\mu(t))\Big|_{t=0}\ ,$$

where $l\in L$, $\phi\in C^\infty(L)$. Now every element of the Lie groups $\tilde{G}$ and G can be specified by assigning values to the $(2n+2)$ parameters, $\sigma,\beta_1,\beta_2,\ldots,\beta_n,\alpha_1,\alpha_2,\ldots,\alpha_n,\tau$. For each such parameter μ, let $\mu(t)$ be the one-dimensional Lie subgroup formed by those group elements for which the value of every parameter except μ is zero. Then, for both $\tilde{G}$ and G, using (6),

$$X_\sigma=\frac{\partial}{\partial\sigma}\ ,$$

$$X_{\beta_j}=\alpha_j\cos\omega_j\tau\frac{\partial}{\partial\sigma}+\cos\omega_j\tau\frac{\partial}{\partial\beta_j}+\sin\omega_j\tau\frac{\partial}{\partial\alpha_j}\ ,$$

$$X_{\alpha_j}=-\alpha_j\sin\omega_j\tau\frac{\partial}{\partial\sigma}-\sin\omega_j\tau\frac{\partial}{\partial\beta_j}+\cos\omega_j\tau\frac{\partial}{\partial\alpha_j}\ ,$$

$$X_\tau=\frac{\partial}{\partial\tau}\ .$$

Direct calculation of the commutation relations for these X_μ's then shows that the Lie algebras of $\tilde{G}$ and G are each isomorphic to $\mathcal{L}$.

Moreover, since $\tilde{G}$ is simply connected, and $\tilde{G}$ and G are connected Lie groups with isomorphic Lie algebras, it follows that $\tilde{G}$ is the simply connected covering group of G.

Now, from (6), an arbitrary element of $\tilde{G}$ (or G) may be expressed as

$$(\sigma,\beta,\alpha,\tau)=(\sigma,\beta,0,0)(0,0,\alpha,0)(0,0,0,\tau)\ .\qquad(8)$$

It then follows, using (6), that

$$\tilde{G}=N\circledS\tilde{M}=(Y\circledS X)\circledS\tilde{M},\quad G=N\circledS M=(Y\circledS X)\circledS M\ ,\qquad(9)$$

where

$$N=Y\circledS X=\{(\sigma,\beta,\alpha,0):\sigma\in\mathbf{R},\ \beta,\alpha\in\mathbf{R}^n\}$$

is a normal subgroup of $\tilde{G}$ and G;

$$Y=\{(\sigma,\beta,0,0):\sigma\in\mathbf{R},\ \beta\in\mathbf{R}^n\}$$

is a normal subgroup of N, and is isomorphic to $\mathbf{R}^{n+1}$;

$$X=\{(0,0,\alpha,0):\alpha\in\mathbf{R}^n\}$$

is a subgroup of $\tilde{G}$ and G, and is isomorphic to $\mathbf{R}^n$;

$$\tilde{M}=\{(0,0,0,\tau):\tau\in\mathbf{R}\}$$

is a subgroup of $\tilde{G}$ and is isomorphic to $\mathbf{R}$;

$$M=\{(0,0,0,\tau):\tau\in(-\pi\lambda,\pi\lambda]\}$$

is a subgroup of G and is isomorphic to $\mathbf{R}/\mathbf{Z}$.

5. MACKEY'S THEORY OF GROUP REPRESENTATIONS

From here onwards, a representation will mean a unitary representation on a separable Hilbert space.

A. Induced representation (Ref. 3)

Let G be a separable locally compact group, and let H be a closed subgroup of G. Let ρ be a unitary representation of H on a separable Hilbert space $\mathfrak{H}_\rho$. Suppose there exists an invariant measure μ on G/H, and let $L^2(G/H,\mathfrak{H}_\rho,\mu)$ denote the space of functions from G/H to $\mathfrak{H}_\rho$ which are square-integrable with respect to the measure μ. From each coset $x\in G/H$, choose a coset representative $\Lambda(x)\in G$ so that $\Lambda:G/H\to G$ is a Borel function. Then the representation U of G induced by the representation ρ of H may be expressed in the form

$$(U(g)\psi)(x)=\rho(\Lambda(x)^{-1}g\Lambda(g^{-1}x))\,\psi(g^{-1}x),\qquad(10)$$

where $g\in G$, $x\in G/H$, $\psi\in L^2(G/H,\mathfrak{H}_\rho,\mu)$.

The representation U will be denoted by $\rho(H)\uparrow G$, or, when there is no risk of ambiguity, by $\rho\uparrow G$. The notation $\chi(G)\downarrow H$, or $\chi\downarrow H$, will be used to denote the restriction of a representation χ of G to a subgroup H.

In Mackey's theory of group representations, induced representations are used to find all irreducible representations of a certain type of group from those of proper subgroups. The basic result which is needed in this paper is the following.

B. Theorem (Ref. 4)

Let $G = N \circledS K$ be a separable locally compact semidirect product group. Define an action of K on the representations of N by

$$(k\chi)(n') = \chi(k^{-1}n'k), \tag{11}$$

where $k \in K, n' \in N$ and χ is a representation of N. Let $\hat{N}$ denote the set of equivalence classes of irreducible representations of N. Suppose there exists a Borel subset of $\hat{N}$ which meets each orbit of $\hat{N}$ (under the action of K) in exactly one point.

(a) General case

Let χ be an irreducible ordinary representation of N, $K_\chi = \{k \in K : k\chi \simeq \chi\}$ and $G_\chi = N \circledS K_\chi$. ($K_\chi$ is known as the little group of χ, and G_χ as the isotropy group of χ.) Then there exists a projective representation W of K_χ satisfying $(k\chi)(n') = W(k)^{-1}\chi(n')W(k)$, for each $k \in K_\chi$, $n' \in N$; W and the multiplier γ of W are unique up to multiplication by trivial multipliers. For each irreducible $1/\gamma$ representation η of K_χ, the representation of G induced from the ordinary representation $\chi W \otimes \eta : nk \to \chi(n)W(k) \otimes \eta(k)$ of G_χ is an irreducible representation of G.

Every irreducible ordinary representation of G is unitarily equivalent to one induced from a representation of the form $\chi W \otimes \eta : nk \to \chi(n)W(k) \otimes \eta(k)$, where the orbit of the irreducible representation χ of N is uniquely determined, the projective representation W of K_χ is determined up to trivial multipliers by χ, and the irreducible projective representation η of K_χ is determined up to unitary equivalence.

(b) Case when N is Abelian

Let χ be an irreducible ordinary representation of N, $K_\chi = \{k \in K : k\chi = \chi\}$ and $G_\chi = N \circledS K_\chi$. Then, for each irreducible ordinary representation η of K, the representation of G induced from the ordinary representation $\chi_\eta : nk \to \chi(n)\eta(k)$ $(n \in N, k \in K_\chi)$ of G_χ is an irreducible representation of G.

Every irreducible ordinary representation of G is unitarily equivalent to one induced from a representation of the form $\chi_\eta : nk \to \chi(n)\eta(k)$, where the orbit of the irreducible representation χ of N is uniquely determined, and the irreducible representation η of K_χ is determined up to unitary equivalence.

6. THE IRREDUCIBLE REPRESENTATIONS OF THE GROUP N

All the irreducible representations of $\tilde{G} = N \circledS \tilde{M}$ may be found by using Mackey's theory for the general case. In order to do this, all the irreducible representations of N must first be evaluated; since $N = Y \circledS X$ (where Y is Abelian), this also may be done by applying Mackey's theory.

The group law of N is, from (6),

$$(\sigma', \beta', \alpha', 0)(\sigma, \beta, \alpha, 0) = (\sigma' + \sigma + \alpha' . \beta, \beta' + \beta, \alpha' + \alpha, 0). \tag{12}$$

Since Y is isomorphic to R^{n+1}, each of its irreducible representations is of the form

$$\chi_{v,u}(\sigma, \beta, 0, 0) = \exp(iv\sigma + iu.\beta), \tag{13}$$

where $(\sigma, \beta, 0, 0) \in Y$, $v \in \mathrm{R}$, and $u \in \mathrm{R}^n$.

The action of $(0, 0, \alpha, 0) \in X$ on $\chi_{v,u}$ is, by definition (11),

$$[(0, 0, \alpha, 0)\chi_{v,u}](\sigma, \beta, 0, 0) = \chi_{v,u}[(0, 0, \alpha, 0)^{-1}(\sigma, \beta, 0, 0)(0, 0, \alpha, 0)] = \chi_{v,-v\alpha+u},$$

from (12) and (13).

The orbits of $\hat{Y}$ under the action of X are therefore

$$\begin{aligned} &O_v = \{\chi_{v,u} : u \in \mathrm{R}^n\} \quad (v \neq 0, v \in \mathrm{R}), \\ &O_u = \{\chi_{0,u}\} \quad (u \in \mathrm{R}^n). \end{aligned} \tag{14}$$

Firstly, consider the representations of N arising from orbits of the form O_v.

As a representative point of the orbit O_v, choose $\chi_{v,0}$. The little group of $\chi_{v,0}$ is just the identity $(0, 0, 0, 0)$, whose only irreducible ordinary representation is, of course, the one-dimensional identity representation. Hence the corresponding representation of N is given by inducing the representation $\chi_{v,0}$ of Y $[= Y \circledS (0, 0, 0, 0)]$ to N.

An explicit expression can be given for $\chi_{v,0}(Y) \uparrow N$ once coset representatives of N/Y in N have been chosen. Now $(\sigma, \beta, \alpha, 0)$ and $(\sigma', \beta', \alpha', 0)$ belong to the same coset of Y in N if and only if $\alpha = \alpha'$; denote the coset to which $(\sigma, \beta, \alpha, 0)$ belongs by α $(\in \mathrm{R}^n)$. Then the natural action of N on the coset $\alpha \in N/Y$ is given by

$$(\sigma', \beta', \alpha', 0)\alpha = \alpha' + \alpha.$$

Since the coset representatives $\Lambda : N/Y \to N$ must satisfy $\Lambda(\alpha)0 = \alpha$, a possible choice for Λ is $\Lambda(\alpha) = (0, 0, \alpha, 0)$ $(\in N)$.

It follows, from (10), that the representation of N induced by $\chi_{v,0}$ is

$$[U^v_N(\sigma, \beta, \alpha, 0)\psi](x) = \exp iv(\sigma - x.\beta)\psi(x - \alpha), \tag{15}$$

where v $(\neq 0) \in \mathrm{R}$, and $\psi \in L^2(\mathrm{R}^n, \mathrm{C})$.

Secondly, consider the representations of N arising from orbits of $\hat{Y}$ of the form O_u [(14)].

O_u is an orbit consisting of the single point $\chi_{0,u}$; the little group of $\chi_{0,u}$ is the whole of the subgroup X. As $X \approx \mathrm{R}^n$, each of its irreducible representations is of the form

$$\eta_t : (0, 0, \alpha, 0) \to \exp it.\alpha,$$

where $t \in \mathrm{R}^n$.

Since the isotropy group of $\chi_{0,u}$ is the whole of N, the representation of N obtained from the representations $\chi_{0,u}$ of Y and η_t of X is just their product,

$$U^{u,t}_N(\sigma, \beta, \alpha, 0) = \exp i(u.\beta + t.\alpha), \tag{16}$$

where $u,\ t \in \mathbf{R}^n$.

Hence, up to unitary equivalence, each irreducible representation of N is of the form (15) or (16).

7. THE IRREDUCIBLE REPRESENTATIONS OF THE GROUP $\tilde{G}$

As the next step towards obtaining all the irreducible representations of $\tilde{G} = N \circledS \tilde{M}$, consider the action of $(0,0,0,\tau) \in \tilde{M}$ on an irreducible representation U_N^v of N. From (6), (11), and (15),

$$\begin{aligned}&([(0,0,0,\tau)U_N^v](\sigma,\beta,\alpha,0)\psi)(x)\\ &= \exp iv\Big(\sigma + \sum_{j=1}^{n}\Big[\frac{\alpha_j^2-\beta_j^2}{2}\cos\omega_j\tau\sin\omega_j\tau\\ &\quad -\beta_j\alpha_j\sin^2\omega_j\tau - x_j(\beta_j\cos\omega_j\tau+\alpha_j\sin\omega_j\tau)\Big]\Big)\\ &\quad \times\psi(\cdots, x_j+\beta_j\sin\omega_j\tau-\alpha_j\cos\omega_j\tau,\cdots). \qquad (17)\end{aligned}$$

Since U_N^v is an irreducible representation, so is $(0,0,0,\tau)U_N^v$. Restricted to the subgroup $\{(\sigma,0,0,0)\}$, $(0,0,0,\tau)U_N^v$ is just the phase $\exp iv\sigma$. It follows that $(0,0,0,\tau)U_N^v$ must be unitarily equivalent to U_N^v, since U_N^v is the only irreducible representation of N which has the required form on restriction to the subgroup $\{(\sigma,0,0,0)\}$.

Hence there exists a unitary operator $W(\tau)$, dependent on τ, such that

$$(0,0,0,\tau)U_N^v = W(\tau)^{-1}U_N^vW(\tau), \qquad (18)$$

for each $(0,0,0,\tau) \in \tilde{M}$.

Each orbit of $\{U_N^v : v(\neq 0) \in \mathbf{R}\}$ $(\subset \hat{N})$ under the action of $\tilde{M}$ thus consists of a single point U_N^v; the little group of U_N^v is the whole of $\tilde{M}$, and so the isotropy group of U_N^v is $N \circledS \tilde{M} = \tilde{G}$.

In order to obtain the corresponding irreducible representations of $\tilde{G}$ explicitly, the operator $W(\tau)$ must be found. From Mackey's theory, W is a projective representation of $\tilde{M}$, and is unique up to trivial multipliers. Since $\tilde{M} \approx \mathbf{R}$, and R has no nontrivial multipliers,[5] W may be taken to be an ordinary unitary representation of the one-parameter subgroup $\tilde{M}$. Hence, by Stone's theorem,[6] there exists a unique skew-adjoint operator J such that

$$W(\tau) = \exp\tau J. \qquad (19)$$

It remains to find this operator J, which is determined by

$$[(0,0,0,\tau)U_N^v](n) = \exp(-\tau J)U_N^v(n)\exp(\tau J). \qquad (20)$$

Now each element $(\sigma,\beta,\alpha,0) \in N$ can be expressed as a product of elements of one-parameter subgroups, using (12), by

$$\begin{aligned}&(\sigma,\beta,\alpha,0)\\ &= (\sigma,0,0,0)\prod_{j=1}^{n}(0,\beta_je_j,0,0)\prod_{j=1}^{n}(0,0,\alpha_je_j,0), \qquad (21)\end{aligned}$$

where $e_j\ (\in \mathbf{R}^n)$ is a unit vector with 1 in the jth position and zeros elsewhere.

For each one-parameter subgroup $\mu(t)$ of N, the map $\mu(t) \to [(0,0,0,\tau)U_N^v](\mu(t))$ is an ordinary unitary representation. Hence, by Stone's theorem, there exists a unique skew-adjoint operator $Z_\mu(\tau)$ such that

$$[(0,0,0,\tau)U_N^v](\mu(t)) = \exp tZ_\mu(\tau). \qquad (22)$$

So, for each μ,

$$\exp tZ_\mu(\tau) = \exp(-\tau J)\exp[tZ_\mu(0)]\exp(\tau J). \qquad (23)$$

Differentiating with respect to t and putting $t = 0$, and then differentiating the result with respect to τ and putting $\tau = 0$ gives

$$\left.\frac{dZ_\mu(\tau)}{d\tau}\right|_{\tau=0} = [Z_\mu(0), J] \qquad (24)$$

for each $\mu \in \{\sigma,\beta_1,\ldots,\beta_n,\alpha_1,\ldots,\alpha_n\}$.

The existence of a skew-adjoint operator J satisfying (24) follows from the existence of $W(\tau) = \exp\tau J$.

Conversely, suppose $\tilde{J}$ is a skew-adjoint operator satisfying (24). To investigate the uniqueness of $\tilde{J}$, let $J' = J - \tilde{J}$. Then $[Z_\mu(0), J'] = 0$ for each μ. The operators $Z_\mu(0)$ give the representation of the Lie algebra of N corresponding to the representation U_N^v of the Lie group N. Since U_N^v is an irreducible representation, it follows, by Schur's lemma, that J' is a scalar multiple of the identity operator; since J' is a skew-adjoint operator, J' must equal ξI, where ξ is imaginary.

Hence every skew-adjoint operator $\tilde{J}$ satisfying (24) also satisfies (23), and (24) determines the skew-adjoint operator J up to an imaginary constant.

Now, from (17) and (22),

$$\begin{aligned}Z_{\beta_j}(\tau)\psi(x) &= \frac{\partial}{\partial t}([(0,0,0,\tau)U_N^v](0,te_j,0,0)\psi)(x)\Big|_{t=0}\\ &= \left(-ivx_j\cos\omega_j\tau + \sin\omega_j\tau\frac{\partial}{\partial x_j}\right)\psi(x).\end{aligned}$$

Similarly,

$$Z_{\alpha_j}(\tau) = \left(-ivx_j\sin\omega_j\tau - \cos\omega_j\tau\frac{\partial}{\partial x_j}\right),$$

$$Z_\sigma(\tau) = iv.$$

Hence, from (24), J must satisfy

$$0 = [iv, J], \quad \omega_j\frac{\partial}{\partial x_j} = [-ivx_j, J],$$

$$-iv\omega_jx_j = \left[-\frac{\partial}{\partial x_j}, J\right], \quad \text{for each } j = 1,2,\ldots,n.$$

A solution of these is

$$\tilde{J} = \frac{i}{2}\sum_{k=1}^{n}\omega_k\left(-\frac{1}{v}\frac{\partial^2}{\partial x_k^2} + vx_k^2\right).$$

With this $\tilde{J}$, and any imaginary constant ξ, $\tilde{W}(\tau) = \exp\tau(\tilde{J} + \xi I)$ is an ordinary representation of $\tilde{M}$ satisfying

$$(0,0,0,\tau)U_N^v = \tilde{W}(\tau)^{-1}U_N^v\tilde{W}(\tau). \qquad (18)$$

It follows from Mackey's theory that $U_N^v\tilde{W}: (\sigma,\beta,\alpha,\tau) \to U_N^v(\sigma,\beta,\alpha,0)\tilde{W}(\tau)$ is an irreducible representation of $N \circledS \tilde{M} = \tilde{G}$.

Now every irreducible representation of $\tilde{M}(\approx \mathbf{R})$ is of the form

$(0,0,0,\tau) \to \exp il\tau$, where $l \in \mathbf{R}$.

Hence, from Mackey's theory, every irreducible representation of $\tilde{G}$ obtained from U_N^v is unitarily equivalent to one of the form

$$(U_{\tilde{G}}^{v,h}(\sigma,\beta,\alpha,\tau)\psi)(x)$$

$$= U_N^v(\sigma,\beta,\alpha,0)\tilde{W}(\tau)\exp il\tau\psi(x)$$

$$= \exp iv\sigma \prod_{j=1}^{n} \exp(-iv\beta_j x_j) \prod_{j=1}^{n} \exp\left(-\alpha_j \frac{\partial}{\partial x_j}\right)$$

$$\times \exp i\tau\left[h + \frac{1}{2}\sum_{j=1}^{n} \omega_j\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2} + vx_j^2\right)\right]\psi(x), \tag{25}$$

where $h = l - i\xi$ can take any value in R.

In the irreducible representation $U_N^{u,t}$ [given by (16)], the generator of the subgroup $\{(\sigma,0,0,0):\sigma \in \mathbf{R}\}$ is represented by the zero operator, and it is easily seen that this is also true for the irreducible representations of $\tilde{G}$ obtained from $U_N^{u,t}$ by Mackey's method. Among the commutation relations of $\mathcal{L}$ are $[A_j, B_k] = \delta_{jk}E$ $(1 \leq j, k \leq n)$. Hence, when E is represented by the zero operator, the representatives of A_j and B_k must commute, for all $j,k = 1,\ldots,n$; this implies that the representatives of all the momentum and position observables commute. This situation corresponds to the classical case; since it is the quantum mechanical case which is of interest here, the representations arising from $U_N^{u,t}$ will not be considered further.

8. THE IRREDUCIBLE REPRESENTATIONS OF THE GROUP G

When the ω_j's are not rationally related, the group G is just the group $\tilde{G}$, so the quantum mechanical irreducible representations of G are given by (25). When the ω_j's are rationally related, the quantum mechanical irreducible representations of G may be obtained from those of its simply connected covering group $\tilde{G}$ in the following way.

From (7),

$$G \approx \frac{\tilde{G}}{Z}, \quad \text{where } Z = \{(0,0,0,2\pi\lambda s): s \in \mathbf{Z}\}.$$

Hence $U_{\tilde{G}}^{v,h}$ defines a single-valued representation of G if $U_{\tilde{G}}^{v,h}$ has the same value on each element of the subgroup Z. Now, for $s \in \mathbf{Z}$,

$$[U_{\tilde{G}}^{v,h}(0,0,0,2\pi\lambda s)\psi](x)$$

$$= \exp i2\pi\lambda s\left[h + \tfrac{1}{2}\sum_{j=1}^{n} \omega_j\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2} + vx_j^2\right)\right]\psi(x).$$

As a basis for the representation space $L^2(\mathbf{R}^n, \mathbf{C})$ of $U_{\tilde{G}}^{v,h}$, take the set $\{\psi_m : m = (m_1, m_2, \ldots, m_n) \in \mathbf{Z}_+^n\}$, where

$$\psi_m(x) = \prod_{j=1}^{n} u_{m_j}(|v|^{1/2}x_j),$$

with u_{m_j} a Hermite function of order m_j.

Then, with $\operatorname{sgn} v = \begin{cases} +1 \text{ if } v \text{ is positive,} \\ -1 \text{ if } v \text{ is negative,} \end{cases}$

$$[U_{\tilde{G}}^{v,h}(0,0,0,2\pi\lambda s)\psi_m](x) = \exp i\,2\pi\lambda s\left[h + \frac{\operatorname{sgn} v}{2}\sum_{j=1}^{n} \omega_j\right]\psi_m(x)$$

since $\lambda\omega_j$ is integral for each j. Hence $U_{\tilde{G}}^{v,h}$ defines a single-valued representation of G only if

$$h = \frac{r}{\lambda} - \frac{\operatorname{sgn} v}{2}\sum_{j=1}^{n} \omega_j, \quad \text{for some } r \in \mathbf{Z}.$$

Therefore, every quantum mechanical irreducible representation of G is unitarily equivalent to one of the form

$$[U_G^{v,r}(\sigma,\beta,\alpha,\tau)\psi](x)$$

$$= \exp iv\sigma \prod_{j=1}^{n} \exp(-iv\beta_j x_j) \prod_{j=1}^{n} \exp\left(-\alpha_j \frac{\partial}{\partial x_j}\right)$$

$$\times \exp i\tau\left[\frac{r}{\lambda} + \tfrac{1}{2}\sum_{j=1}^{n} \omega_j\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2} + vx_j^2 - \operatorname{sgn} v\right)\right]\psi(x), \tag{26}$$

where $v(\neq 0) \in \mathbf{R}$ and $r \in \mathbf{Z}$.

ACKNOWLEDGMENTS

I should like to thank Professor J. T. Lewis for his advice and encouragement. I am also grateful to the Science Research Council for financial support while most of the work for this paper was done.

[1] N. Jacobson, *Lie Algebras* (Interscience, New York, 1962).
[2] R. F. Streater, Commun. Math. Phys. **4**, 217 (1967).
[3] G. W. Mackey, Ann. Math. **55**, 101 (1952).
[4] G. W. Mackey, Ann. Math. **99**, 265 (1958).
[5] K. R. Parthasarathy, *Multipliers on Locally Compact Groups*, Lecture Notes in Mathematics, Vol. 93 (Springer, Berlin, 1969).
[6] K. Yosida, *Functional Analysis* (Springer, Berlin, 1965).

On the accidental degeneracy of the anisotropic harmonic oscillator. I

M. E. Major

Mathematics Department, Queen Elizabeth College, London W8 7AH, England
(Received 22 September 1975)

A group G associated with the n-dimensional anisotropic harmonic oscillator is shown to be embedded in a semidirect product L of the Weyl group N and the symplectic group $\mathrm{Sp}(2n,\mathbb{R})$. A particular induced representation of the group L, when restricted to G, is proved to be unitarily equivalent to $\oplus_s d_{\omega,s} U_G^{\nu,-(\mathrm{sgn}\nu)s}$, where $d_{\omega,s}$ is the degeneracy of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with frequencies $(\omega_1, \omega_2,\ldots,\omega_n) = \omega$, $U_G^{\nu,-(\mathrm{sgn}\nu)s}$ is an irreducible representation of G, and s may be regarded as indexing all distinct energy levels of the system.

1. INTRODUCTION

In an attempt to gain a greater understanding of accidental degeneracy in quantum theory, a study has been made of the n-dimensional anisotropic harmonic oscillator.

The quantum mechanical n-dimensional anisotropic harmonic oscillator is a particle of mass m with Hamiltonian

$$H_{\mathrm{osc}} = \sum_{j=1}^{n} \frac{1}{2m}(P_j^2 + m^2\omega_j^2 Q_j^2),$$

where $\omega_1, \omega_2, \ldots, \omega_n$ are positive constants (the frequencies), and $\{Q_j\colon\ j = 1,2,\ldots,n\}$ and $\{P_j\colon\ j = 1,2,\ldots,n\}$ are sets of self-adjoint operators corresponding to position and momentum observables respectively, and satisfying the commutation relations

$$[Q_j, P_k] = i\hbar\delta_{jk} I \quad (1 \leq j,k \leq n).$$

To avoid unimportant constants, define new self-adjoint operators:

$$\hat{Q}_j = \left(\frac{m\omega_j}{\hbar}\right)^{1/2} Q_j, \quad \hat{P}_j = \left(\frac{1}{m\hbar\omega_j}\right)^{1/2} P_j, \quad \hat{H} = \frac{H_{\mathrm{osc}}}{\hbar},$$

giving

$$\hat{H} = \sum_{j=1}^{n} \hat{H}_j,$$

where

$$\hat{H}_j = \tfrac{1}{2}\omega_j(\hat{P}_j^2 + \hat{Q}_j^2) \quad \text{and} \quad [\hat{Q}_j, \hat{P}_k] = i\delta_{jk} I.$$

The eigenvalues of $\hat{H}_j$ are $\omega_j(m_j + \frac{1}{2})$, where $m_j \in \mathbb{Z}_+$. Hence each eigenvalue, or energy level, E of H_{osc} is of the form

$$E = \hbar\sum_{j=1}^{n} \omega_j(m_j + \tfrac{1}{2}).$$

It is helpful to distinguish two cases:

(a) Suppose there exists a finite positive number μ such that each $\mu\omega_j$ is integral; let λ be the smallest such number μ. (In this case, the ω_j's are said to be rationally related.) The possible energy levels of the n-dimensional anisotropic harmonic oscillator with rationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$ are therefore

$$E_{\omega,s} = \hbar\left(\frac{s}{\lambda} + \tfrac{1}{2}\sum_{j=1}^{n} \omega_j\right),$$

where s $(\in \mathbb{Z}_+)$ is expressible in the form $\lambda\sum_{j=1}^{n} \omega_j m_j$ for some $m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}_+^n$.

The degeneracy $d_{\omega,s}$ of the energy level $E_{\omega,s}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}_+^n$ so that the relation $s = \lambda\sum_{j=1}^{n} \omega_j m_j$ is satisfied.

(b) Suppose no finite positive number μ exists such that each $\mu\omega_j$ is integral. The possible energy levels of the n-dimensional anisotropic harmonic oscillator with nonrationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$ are

$$E_{\omega,e} = \hbar\sum_{j=1}^{n} \omega_j(m_j + \tfrac{1}{2}),$$

where e is expressible in the form $\sum_{j=1}^{n} \omega_j(m_j + \frac{1}{2})$ for some $m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}_+^n$.

The degeneracy $d_{\omega,e}$ of the energy level $E_{\omega,e}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}_+^n$ so that the relation $e = \sum_{j=1}^{n} \omega_j(m_j + \frac{1}{2})$ is satisfied.

There have been a number of studies of the accidental degeneracy of the anisotropic harmonic oscillator. As is well known, geometric symmetry of a Hamiltonian leads to degeneracy of its eigenvalues, and to a representation on each eigenspace of a group of operators which commute with the Hamiltonian. When such a representation is reducible, sets of eigenfunctions belonging to separate irreducible representations coincide in energy; it was suggested that this extra (or accidental) degeneracy occurs because there is more symmetry present than just the obvious geometric symmetry. The group formed by the obvious geometric symmetries and these additional "hidden" symmetries would have a representation on each eigenspace of the Hamiltonian; if each such representation were irreducible, then the accidental degeneracy would be regarded as being explained. Most studies of accidental degeneracy have therefore concentrated on symmetry properties.

In 1940, Jauch and Hill[1] found classical constants of the motion for the two-dimensional anisotropic harmonic oscillator with rational frequency ratio, but when these constants were quantized in a reasonable way, they did not form a Lie algebra. In 1965, Dulock and McIntosh[2] extended these classical constants to the n-dimensional case with arbitrary frequency ratios, but because of the transcendental nature and multivaluedness of the constants, it was not possible to form their quantum me-

chanical analogs. They showed that their classical constants generated a group isomorphic to SU(n).

Demkov,[3] in 1963, investigated the particular two-dimensional case with frequency ratio 2 : 1. He showed that the eigenfunctions of the Hamiltonian may be divided into two sets, of even and odd energy values respectively; to each even value there corresponds an irreducible representation of SU(2), and each irreducible representation of SU(2) occurs in this way; the same applies to the odd values of the energy. Hence the eigenfunctions are grouped into two systems of irreducible representations of SU(2). Il'kaeva,[4] a student of Demkov, dealt with a rather more general situation. Cisneros and McIntosh[5] extended these results to the two-dimensional case with rational frequency ratio $k_2 : k_1$ (where k_1 and k_2 are mutually prime integers). They showed that the eigenfunctions may be grouped into $k_1 k_2$ systems of irreducible representations of SU(2). Recently, a somewhat different approach has been taken by Louck, Moshinsky, and Wolf.[6] They determined the canonical transformation that maps the classical two-dimensional anisotropic harmonic oscillator with rational frequency ratio $k_2 : k_1$ onto the isotropic one. Since the latter has a symmetry group of linear canonical transformations that are a representation of SU(2), these can be combined with the canonical transformation giving the mapping, to obtain the symmetry group of the classical anisotropic harmonic oscillator. Using the classical situation as a guide, ladder operators can be constructed in the quantum picture; these have different forms for the $k_1 k_2$ different systems of eigenfunctions. From the ladder operators, the generators of the SU(2) symmetry group responsible for the accidental degeneracy can then be determined. These results were later placed in a more general context of groups of canonical transformations responsible for accidental degeneracy in two-dimensional quantum mechanical problems.[7]

Less attention has been devoted to the n-dimensional anisotropic harmonic oscillator with rational frequency ratios (the n-dimensional rational oscillator) for $n > 2$. In 1968, Vendramin[8] showed that, in general, the degeneracies of the energy levels are not equal to the dimensions of any irreducible representations of SU(n). The following year, Maiella and Vilasi[9] constructed operators which commute with the Hamiltonian of the three-dimensional rational oscillator, and which generate SU(3); they showed that to any energy level there corresponds a reducible representation of SU(3). Cisneros and McIntosh,[5] generalizing their analysis of the two-dimensional rational oscillator to the n-dimensional case, found that each eigenfunction belongs to one and only one of several systems, each of which forms a basis for the symmetric tensor irreducible representations of SU(n). However, for $n > 2$, all eigenfunctions of a given energy level may not belong to the same system; hence, the representation of SU(n) corresponding to a given energy level may be reducible.

The approach adopted to the problem here is rather different. Symmetry groups are not considered; nor does the classical situation enter the picture directly. Instead, the basic object studied is a group, G, associated with the n-dimensional anisotropic harmonic oscillator: G is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system. (Such a group G can be constructed whether the frequencies are rationally related or not.) The next step is to show that G can be embedded in a semidirect product, L, of the Weyl group, N, and the symplectic group, $\mathrm{Sp}(2n,\mathrm{R})$. The anisotropic harmonic oscillator is then studied by examining representations of the groups G and L. All the quantum mechanical irreducible representations $U_G^{v,r}$ of G were evaluated in Ref. 10.

In the present paper it is proved that the degeneracies of the energy levels of the n-dimensional anisotropic harmonic oscillator occur in the following way (whether the frequencies are rationally related or not). A particular induced representation R^v of the group L, when restricted to G, is unitarily equivalent to

$$\bigoplus_s d_{\omega,s} U_G^{v,-(\mathrm{sgn}\, v)s},$$

when $v \neq 0$, where $d_{\omega,s}$ is precisely the degeneracy of the energy level $E_{\omega,s}$ of the system, $U_G^{v,-(\mathrm{sgn}v)s}$ is an irreducible representation of G, and the summation may be regarded as over all distinct energy levels of the system.

A feature of the approach taken here is that, from the outset, the emphasis is on the use of Lie groups (and their representations), rather than on Lie algebras. One advantage of this is that Mackey's theory of induced group representations can be used to find all the irreducible representations of G (as shown in Ref. 10), and to help in the analysis of reducible representations of G; the corresponding results for the Lie algebra of G would then follow from those for the group G.

The definition of L and of the representation R^v are from a paper by van Hove[11] (which is mainly concerned with a possible method of quantization). The parts of this needed for present purposes are outlined in Secs. 2 and 3. In Sec. 4, the representation R^v of L is described explicitly as an induced representation. This enables $R^v \downarrow G$ in turn to be expressed as an induced representation, using Mackey's subgroup theorem. If G were a compact group, the usual form of the Frobenius reciprocity theorem could then be applied to find the decomposition of this representation. However, G is only locally compact. Nevertheless, a formal application of the theorem is still of some help since it suggests what the result may be. Other methods are then used to prove that this suggested result is, in fact, correct.

2. DEFINITION OF VAN HOVE'S GROUP Γ AND SOME OF ITS SUBGROUPS

Define a differential form $\bar{\theta}$ on the vector space of real $(2n+1)$-tuples: $\{(s, q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n) \in \mathbf{R}^{2n+1}\}$ by

$$\bar{\theta} = ds - \sum_{j=1}^{n} p_j dq_j = ds - p.dq,$$

where $q = (q_1, q_2, \ldots, q_n)$, $p = (p_1, p_2, \ldots, p_n)$.

Let Γ be the family of invertible C^∞ transformations of this space, $\{(s, q, p)\}$, into itself which leave the differ-

ential form $\bar{\theta}$ invariant. It can be shown that Γ is a group, and that every element of Γ is a transformation of the type

$$q'=q'(q,p),\quad p'=p'(q,p),\quad s'=s+\pi(q,p), \tag{1}$$

where $p'.dq'-p.dq=d\pi$.

The following finitely-generated subgroups of Γ are of special importance. (In each case, the topology on a subgroup with m generators is taken to be that inherited from the usual topology on R^m.)

The center C: This is the subgroup given by transformations of the form

$$q'=q,\quad p'=p,\quad s'=s+\sigma, \tag{2}$$

where $\sigma\in\mathrm{R}$. As van Hove shows, the group Γ/C is essentially the group of invertible classical canonical transformations.

The subgroup N: A simple transformation of the space $\{(q,p)\}$ is that of translation,

$$q'=q+\beta,\quad p'=p+\alpha, \tag{3a}$$

where $\beta,\alpha\in\mathrm{R}^n$ (retaining van Hove's choice of notation, both here and for the other subgroups he mentions). This implies

$$s'=s+\alpha.q+\sigma, \tag{3b}$$

where $\sigma\in\mathrm{R}$. This set of transformations is a subgroup of Γ of which a typical element will be written (σ,β,α,I) (for reasons which will soon become apparent). The group law is

$$(\sigma',\beta',\alpha',I)(\sigma,\beta,\alpha,I)=(\sigma'',\beta'',\alpha'',I), \tag{4}$$

where

$$\sigma''=\sigma'+\sigma+\alpha'.\beta,$$

$$\beta''=\beta'+\beta,$$

$$\alpha''=\alpha'+\alpha.$$

It follows, from Ref. 10, Eq. (12), that this subgroup is isomorphic (as a topological group) to the subgroup N of the oscillator group G. N can therefore be identified with a subgroup of Γ.

The symplectic group $Sp(2n,\mathrm{R})$: Another simple transformation of the space $\{(q,p)\}$ is that of linear transformation,

$$\left.\begin{aligned}q'_j&=\sum_{k=1}^{n}(d_{jk}q_k+c_{jk}p_k)\\ p'_j&=\sum_{k=1}^{n}(b_{jk}q_k+a_{jk}p_k)\end{aligned}\right\}\quad\text{or}\quad\left.\begin{aligned}q'&=Dq+Cp\\ p'&=Bq+Ap\end{aligned}\right\}, \tag{5}$$

with

$$S=\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\begin{pmatrix}a_{jk} & b_{jk}\\ c_{jk} & d_{jk}\end{pmatrix}$$

$\in \mathrm{GL}(2n,\mathrm{R})$ (the group of all $2n\times 2n$ real nonsingular matrices). This leads to a transformation of $\{(s,q,p)\}$ belonging to Γ if and only if $\sum_{j=1}^{n}(p'_j dq'_j-p_j dq_j)$ is a perfect differential. This condition is equivalent to the condition that S belongs to the group $\mathrm{Sp}(2n,\mathrm{R})$ of $2n\times 2n$ real symplectic matrices. The resulting subgroup of Γ is isomorphic to the direct product of C and $\mathrm{Sp}(2n,\mathrm{R})$; in particular, $\mathrm{Sp}(2n,\mathrm{R})$ can be identified with the subgroup of Γ to which it is isomorphic.

The subgroup L: The subgroups N and $\mathrm{Sp}(2n,\mathrm{R})$ may be combined by considering transformations of the space $\{(q,p)\}$ of the form

$$q'=Dq+Cp+\beta,\quad p'=Bq+Ap+\alpha, \tag{6a}$$

where $\beta,\alpha\in\mathrm{R}^n$, $\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)\in\mathrm{Sp}(2n,\mathrm{R})$. (As before, only the subgroup $\mathrm{Sp}(2n,\mathrm{R})$ of $\mathrm{GL}(2n,\mathrm{R})$ leads to transformations belonging to Γ.) It follows that

$$s'=s+\tfrac{1}{2}Ap.Cp+\tfrac{1}{2}Bq.Dq+Bq.Cp+\alpha.(Dq+Cp)+\sigma, \tag{6b}$$

where $\sigma\in\mathrm{R}$. This set of transformations, L, is a subgroup of Γ of which a typical element may be written (σ,β,α,S). The group law is

$$(\sigma',\beta',\alpha',S')(\sigma,\beta,\alpha,S)=(\sigma'',\beta'',\alpha'',S''), \tag{7}$$

where

$$\sigma''=\sigma'+\sigma+\tfrac{1}{2}A'\alpha.C'\alpha+\tfrac{1}{2}B'\beta.D'\beta+B'\beta.C'\alpha+\alpha'.(D'\beta+C'\alpha),$$

$$\beta''=D'\beta+C'\alpha+\beta',$$

$$\alpha''=B'\beta+A'\alpha+\alpha',$$

$$S''=S'S.$$

An arbitrary element of L may be expressed in the form

$$(\sigma,\beta,\alpha,S)=(\sigma,\beta,\alpha,I)(0,0,0,S). \tag{8}$$

It is then easily verified, using (7), that

$$L=N\circledS\,\mathrm{Sp}(2n,\mathrm{R}), \tag{9}$$

where

$$N=\{(\sigma,\beta,\alpha,I):\sigma\in\mathrm{R},\ \beta,\alpha\in\mathrm{R}^n\}$$

is a normal subgroup of L, and

$$\mathrm{Sp}(2n,\mathrm{R})=\{(0,0,0,S)\in L\}.$$

[The previously mentioned identifications of N ($\subset G$) and of $\mathrm{Sp}(2n,\mathrm{R})$ with the corresponding isomorphic subgroups of L have been used here.]

The subgroup G: The isomorphism of the subgroup N ($\subset G$) with a subgroup of L raises the possibility that G (or $\tilde{G}$) is also isomorphic to a subgroup of L. The heuristic construction of G[10] suggests consideration of the subset of L given by $\{(\sigma,\beta,\alpha,S(\tau))\}$, where $S(\tau)$ is a symplectic matrix with

$$\left.\begin{aligned}a_{jk}&=\delta_{jk}\cos\omega_j\tau,\quad b_{jk}=\delta_{jk}\sin\omega_j\tau,\\ c_{jk}&=-\delta_{jk}\sin\omega_j\tau,\quad d_{jk}=\delta_{jk}\cos\omega_j\tau.\end{aligned}\right. \tag{10}$$

If the ω_j's are rationally related, then the subgroup M (of G) $=\{(0,0,0,\tau):\tau\in(-\pi\lambda,\pi\lambda]\}$ is isomorphic to the subgroup (of L) $\{(0,0,0,S(\tau)):\ \tau\in(-\pi\lambda,\pi\lambda]\}$ under the map $\tau\to S(\tau)$. It then follows that $\{(\sigma,\beta,\alpha,S(\tau)):\ \sigma\in\mathrm{R},\beta,\alpha\in\mathrm{R}^n,\tau\in(-\pi\lambda,\pi\lambda]\}$ is a subgroup of L and is isomorphic to G.

If any two of the ω_j's are not rationally related, then $\{(\sigma,\beta,\alpha,S(\tau))\colon\ \sigma\in\mathbb{R},\ \beta,\alpha\in\mathbb{R}^n,\ \tau\in\mathbb{R}\}$ is a subgroup of L isomorphic to the group G ($=\tilde{G}$, in this case).

3. THE REPRESENTATIONS R AND R^v OF Γ

Van Hove introduces two types of representation of Γ:

(i) To simplify the notation, let z denote a point of the space $\{(s,q,p)\}$ and let dz denote a volume element $dsdq_1\cdots dq_n dp_1\cdots dp_n$. The volume element dz is invariant under transformations belonging to Γ. Let $\mathfrak{H}_{2n+1}$ be the separable Hilbert space of complex-valued, measurable, square-integrable functions on $\{(s,q,p)\}$, with scalar product defined by

$$(\Phi,Z)=\int\overline{\Phi(z)}Z(z)dz\,,$$

where the integral is taken over the whole space $\{(s,q,p)\}$, and $\overline{\ }$ denotes complex conjugation. Define

$$(R(\gamma)\Phi)(z)=\Phi(\gamma^{-1}z)\,,$$

where $\gamma\in\Gamma$, $\Phi\in\mathfrak{H}_{2n+1}$, and $\gamma^{-1}z$ denotes the image of z under γ^{-1}. Since $d(\gamma z)=dz$, R is a faithful, unitary representation of Γ on the space $\mathfrak{H}_{2n+1}$.

(ii) Let w denote a point of the space $\{(q,p)\}$, and let dw denote a volume element $dq_1\cdots dq_n dp_1\cdots dp_n$. For convenience, write the transformation $\gamma\in\Gamma$ in the form

$$s'=s+\pi_\gamma(w)\,,\quad w'=\gamma\omega\,. \tag{11}$$

The volume element dw is invariant under transformations belonging to Γ. Let $\mathfrak{H}_{2n}$ be the separable Hilbert space of complex-valued, measurable, square-integrable functions on $\{(q,p)\}$, with scalar product defined by

$$(\phi,\xi)=\int\overline{\phi(w)}\xi(w)dw\,,$$

where the integral is taken over the whole space $\{(q,p)\}$. Define

$$(R^v(\gamma)\phi)(w)=\exp[iv\pi_\gamma(\gamma^{-1}w)]\phi(\gamma^{-1}w)\,, \tag{12}$$

where $\gamma\in\Gamma$, $v\in\mathbb{R}$, and $\phi\in\mathfrak{H}_{2n}$. Using $d(\gamma w)=dw$, and

$$\pi_{\gamma_1\gamma_2}(w)=\pi_{\gamma_2}(w)+\pi_{\gamma_1}(\gamma_2 w)\,, \tag{13}$$

it may be verified that R^v is a faithful, unitary representation of Γ, when $v\neq0$, and a faithful, unitary representation of Γ/C when $v=0$.

The reason for introducing the representations R^v is that the representation R is unitarily equivalent to a direct integral of them.[11]

4. EXPRESSION OF $R^v\downarrow L$ AS AN INDUCED REPRESENTATION

The representation R^v of the group Γ can be expressed as

$$(R^v(\gamma)\phi)(w)=m(w,\gamma)\phi(\gamma^{-1}w)\,,$$

where

$$m(w,\gamma)=\exp[iv\pi_\gamma(\gamma^{-1}w)]\,. \tag{14}$$

This is similar in form to the explicit expression for an induced representation $\rho(H)\uparrow G$ of a separable, locally compact group G,

$$(U(g)\psi)(x)=\rho(\Lambda(x)^{-1}g\Lambda(g^{-1}x))\psi(g^{-1}x)\,,$$

where $g\in G$, $x\in G/H$, $\psi\in L^2(G/H,\mathfrak{H}_\rho,\mu)$ [Ref. 10, Eq. (10)]. However, induced representations are usually defined only for separable locally compact groups; the group Γ is not locally compact, but its subgroup L is locally compact, and separable. From now on, attention is therefore restricted to the subgroup L of Γ. It will be shown how $R^v\downarrow L$ can be expressed as an induced representation.

The space $\{(q,p)\}$ is transitive under the action of the subgroup N (of L), and hence it is transitive under the action of L.

Let w_0 be an arbitrary point of the space $\{(q,p)\}$, and let H be the closed subgroup of L which leaves w_0 fixed. Then the space $\{(q,p)\}$ is L-isomorphic to the space L/H. Since R^v is a representation of L, m satisfies the so-called cocycle condition

$$m(w,\gamma_1\gamma_2)=m(w,\gamma_1)m(\gamma_1^{-1}w,\gamma_2)\,, \tag{15}$$

for each $w\in\{(q,p)\}$, $\gamma_1,\gamma_2\in L$.

In particular, for $w=w_0$ and $h_1,h_2\in H$,

$$m(w_0,h_1h_2)=m(w_0,h_1)m(w_0,h_2)\,.$$

Hence, since $m(w_0,h)$ is a complex number of unit modulus, $h\to m(w_0,h)$ is a unitary representation of H.

The representation of L induced from this representation of H is

$$(U(\gamma)\phi)(w)=m(w_0,\Lambda(w)^{-1}\gamma\Lambda(\gamma^{-1}w))\phi(\gamma^{-1}w)\,,$$

where $\gamma\in L$, $\phi\in L^2(L/H,\mathbb{C})$, $w\in\{(q,p)\}$, and $\Lambda\colon\ L/H\to L$ satisfies $\Lambda(w)w_0=w$.

Now, from (15),

$$m(w_0,\Lambda(w)^{-1}\gamma\Lambda(\gamma^{-1}w))$$
$$=m(w_0,\Lambda(w)^{-1})m(w,\gamma)m(\gamma^{-1}w,\Lambda(\gamma^{-1}w))\,.$$

Since $m(w,\gamma)$ is a scalar, $R^v\downarrow L$ is the induced representation U if w_0 and Λ can be chosen such that

$$m(w_0,\Lambda(w)^{-1})m(\gamma^{-1}w,\Lambda(\gamma^{-1}w))=1\,,$$

for all $\gamma\in L$, $w\in\{(q,p)\}$.

If $v=0$, this is the case for any choice of w_0 and Λ. If $v\neq0$, this holds [from (13) and (14)] if

$$\pi_{\Lambda(w)}(w_0)=\text{const},\quad\text{for all } w\in\{(q,p)\}\,.$$

A simple choice of w_0 and Λ which satisfies the above condition and $\Lambda(w)w_0=w$ is

$$w_0=(0,0)\,,\quad\Lambda(w)\hat{w}=\hat{w}+w\,,\quad\pi_{\Lambda(w)}\hat{w}=p\cdot\hat{q}\,,$$

where $w=(q,p)$ and $\hat{w}=(\hat{q},\hat{p})$.

This proves the following result.

Theorem: The representation $R^v\downarrow L$ is a representation of L induced from the representation

$$D\colon\ (\sigma,0,0,S)\to\exp iv\sigma \tag{16}$$

of the subgroup $H=C\times\mathrm{Sp}(2n,\mathbb{R})$, with choice of coset representatives

$$\Lambda(w)\hat{w}=\hat{w}+w\,,\quad\pi_{\Lambda(w)}\hat{w}=p\cdot\hat{q}\,.$$

5. EXPRESSION OF $R^v \downarrow G$ AS AN INDUCED REPRESENTATION

$$R^v \uparrow G = (R^v \uparrow L) \uparrow G = (D(H) \uparrow L) \uparrow G .$$

Hence $R^v \uparrow G$ is the restriction of an induced representation of L. Now G and H are closed subgroups of the separable locally compact group L; also, L can be expressed as a single double coset $G(0,0,0,I)H$, and, trivially, $\{(0,0,0,I)\}$ is a Borel set in L which intersects this double coset exactly once. Mackey's subgroup theorem[12] may therefore be applied, giving

$$(D(H) \uparrow L) \uparrow G = (D(H) \uparrow (G \cap H)) \uparrow G . \tag{17}$$

So

$$R^v \uparrow G = D(G \cap H) \uparrow G ,$$

where $G \cap H = \{(\sigma, 0, 0, S(\tau))\}$, and D: $(\sigma, 0, 0, S(\tau)) \to \exp iv\sigma$ is an irreducible representation of $G \cap H$.

6. FORMAL DECOMPOSITION OF $R^v \downarrow G$

When $v = 0$, the generator of the subgroup $\{(\sigma, 0, 0, I): \sigma \in \mathrm{R}\}$ of G is represented by the zero operator in the representation R^v [from (12)]. This corresponds to the classical case, and so will be ignored.

Suppose ρ is an arbitrary irreducible representation of G. A formal application of the Frobenius reciprocity theorem[12] yields

$$\langle R^v \uparrow G, \rho \rangle_G = \langle D(G \cap H) \uparrow G, \rho \rangle_G = \langle \rho \downarrow (G \cap H), D(G \cap H) \rangle_{G \cap H} . \tag{18}$$

This suggests that ρ occurs in $R^v \uparrow G$ the same number of times that $D(G \cap H)$ occurs in $\rho \downarrow (G \cap H)$.

As explained in Ref. 10, Sec. 7, for a classical irreducible representation ρ_{cl} of G, $\rho_{\mathrm{cl}}(\sigma, \beta, \alpha, S(\tau))$ is independent of the parameter σ, and so $\rho_{\mathrm{cl}} \downarrow (G \cap H)$ cannot contain the representation D: $(\sigma, 0, 0, S(\tau)) \to \exp iv\sigma$ of $G \cap H$, when $v \neq 0$. This suggests that no classical irreducible representation ρ_{cl} can occur in $R^v \uparrow G$.

(a) *Suppose the ω_j's are rationally related*

Then, from Ref. 10, Eq. (26), every quantum mechanical irreducible representation of G is unitarily equivalent to one of the form

$$\begin{aligned}[U_G^{u,r}(\sigma, \beta, \alpha, S(\tau))\psi](x) &= \exp iu\sigma \prod_{j=1}^{n} \exp(-iu\beta_j x_j) \prod_{j=1}^{n} \exp\left(-\alpha_j \frac{\partial}{\partial x_j}\right) \\ &\times \exp i\tau\left[\frac{r}{\lambda} + \tfrac{1}{2}\sum_{j=1}^{n} \omega_j \left(-\frac{1}{u}\frac{\partial^2}{\partial x_j^2} + u x_j^2 - \mathrm{sgn}\, u\right)\right] \psi(x) ,\end{aligned} \tag{19}$$

where $u (\neq 0) \in \mathrm{R}$, $r \in Z$, $\psi \in L^2(\mathrm{R}^n, \mathrm{C})$.

It now remains to find how often the representation $D(G \cap H)$ occurs in $U_G^{u,r} \downarrow (G \cap H)$.

As a basis for the representation space $L^2(\mathrm{R}^n, \mathrm{C})$ of $U_G^{u,r}$, take the set $\{\psi_m: m = (m_1, m_2, \ldots, m_n) \in Z_+^n\}$, where

$$\psi_m(x) = \prod_{j=1}^{n} u_{m_j}(|u|^{1/2} x_j) , \tag{20}$$

with u_{m_j} a Hermite function of order m_j.

Then, from (19),

$$\begin{aligned}[U_G^{u,r}(\sigma, 0, 0, S(\tau))\psi_m](x) &= \exp iu\sigma \exp i\tau\left[\frac{r}{\lambda} + (\mathrm{sgn}\, u)\sum_{j=1}^{n} \omega_j m_j\right] \psi_m(x) .\end{aligned}$$

Therefore,

$$U_G^{u,r} \downarrow (G \cap H) = \bigoplus_m \chi^{u, r + \lambda(\mathrm{sgn}\, u)\sum_{j=1}^{n} \omega_j m_j} ,$$

where $\chi^{\tilde{u},\tilde{r}}$ is the irreducible representation of $G \cap H$ given by

$$\chi^{\tilde{u},\tilde{r}}: \ (\sigma, 0, 0, S(\tau)) \to \exp i\tilde{u}\sigma \exp\left(\frac{i\tilde{r}\tau}{\lambda}\right) \quad (\tilde{u} \in \mathrm{R}, \ \tilde{r} \in Z) .$$

Hence the representation D: $(\sigma, 0, 0, S(\tau)) \to \exp iv\sigma$ of $G \cap H$ occurs in $U_G^{u,r} \downarrow (G \cap H)$ if and only if

$$u = v ,$$

and

$$r + \lambda(\mathrm{sgn}\, u)\sum_{j=1}^{n} \omega_j m_j = 0 \quad \text{for some} \quad m \in Z_+^n . \tag{21}$$

Suppose that, for fixed $r \in Z$, there are $f_{\omega,r}$ distinct ways of choosing m so that (21) is satisfied. Then $D(G \cap H)$ occurs $f_{\omega,r}$ times in $U_G^{u,r} \downarrow (G \cap H)$ provided $u = v$. Hence (18) suggests that $U_G^{u,r}$ occurs in $R^v \uparrow G$ if and only if $u = v$, and that, in this case, it occurs $f_{\omega,r}$ times. Therefore,

$$R^v \uparrow G \simeq \bigoplus_r f_{\omega,r} U_G^{v,r} \quad \text{(formally)} , \tag{22a}$$

where the summation is over all distinct values of r which are expressible in the form

$$r = -\lambda(\mathrm{sgn}\, v)\sum_{j=1}^{n} \omega_j m_j \quad (m_j \in Z_+) , \tag{23a}$$

and, for fixed r, $f_{\omega,r}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in Z_+^n$ so that (23a) is satisfied.

(b) *Suppose the ω_j's are not rationally related*

A similar argument, using the representations $U_G^{u,h}$ [Ref. 10, Eq. (25)] instead of the representations $U_G^{u,r}$ [Ref. 10, Eq. (26)] shows that

$$R^v \uparrow G \simeq \bigoplus_h f_{\omega,h} U_G^{v,h} \quad \text{(formally)} , \tag{22b}$$

where the summation is over all distinct values of h which are expressible in the form

$$h = -(\mathrm{sgn}\, v)\sum_{j=1}^{n} \omega_j (m_j + \tfrac{1}{2}) \quad (m_j \in Z_+) , \tag{23b}$$

and, for fixed h, $f_{\omega,h}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in Z_+^n$ so that (23b) is satisfied.

7. PROOF OF THE DECOMPOSITION OF $R^v \downarrow G$

Alternative expression for $R^v \uparrow G$

Since every element of G can be written in the form

$$\begin{aligned}(\sigma, \beta, \alpha, S(\tau)) = (\sigma, 0, 0, I)\prod_{j=1}^{n} (0, \beta_j e_j, 0, I) \\ \times \prod_{j=1}^{n} (0, 0, \alpha_j e_j, I)(0, 0, 0, S(\tau))\end{aligned} \tag{24}$$

[from (4) and (8)], the representation $R^v \uparrow G$ may be expressed as a product of representations of one-dimen-

sional Lie subgroups of G. These representations may be calculated explicitly by using (12), together with (6), (10), and (11), and Stone's theorem,[13]

$$[R^v(\sigma,0,0,I)\phi](q,p)=\exp iv\sigma\phi(q,p),$$

$$[R^v(0,\beta_j e_j,0,I)\phi](q,p)=\phi(q-\beta_j e_j,p)$$
$$=\exp\left(-\beta_j\frac{\partial}{\partial q_j}\right)\phi(q,p),$$

$$[R^v(0,0,\alpha_j e_j,I)\phi](q,p)=\exp(iv\alpha_j q_j)\phi(q,p-\alpha_j e_j)$$
$$=\exp\alpha_j\left(ivq_j-\frac{\partial}{\partial p_j}\right)\phi(q,p),$$

$$[R^v(0,0,0,S(\tau))\phi](q,p)$$
$$=\exp[-iv\pi_{S(\tau)^{-1}}(q,p)]\phi(S(\tau)^{-1}(q,p))$$
$$=\exp\left[iv\sum_{k=1}^{n}\left(\frac{q_k^2-p_k^2}{2}\sin\omega_k\tau\cos\omega_k\tau+p_kq_k\sin^2\omega_k\tau\right)\right]$$
$$\times\phi(\cdots,q_j\cos\omega_j\tau+p_j\sin\omega_j\tau,\ \ldots,-q_j\sin\omega_j\tau$$
$$+p_j\cos\omega_j\tau,\cdots)$$
$$=\exp\tau\left(\sum_{k=1}^{n}\omega_k\left[\frac{iv}{2}(q_k^2-p_k^2)+p_k\frac{\partial}{\partial q_k}-q_k\frac{\partial}{\partial p_k}\right]\right)\phi(q,p).$$

Therefore,

$$[R^v(\sigma,\beta,\alpha,S(\tau))\phi](q,p)$$
$$=\exp iv\sigma\prod_{j=1}^{n}\exp\left(-\beta_j\frac{\partial}{\partial q_j}\right)$$
$$\times\prod_{j=1}^{n}\exp\alpha_j\left(ivq_j-\frac{\partial}{\partial p_j}\right)\exp\tau\left(\sum_{k=1}^{n}\omega_k\left[\frac{iv}{2}(q_k^2-p_k^2)\right.\right.$$
$$\left.\left.+p_k\frac{\partial}{\partial q_k}-q_k\frac{\partial}{\partial p_k}\right]\right)\phi(q,p). \tag{25}$$

(a) *Suppose the ω_j's are rationally related*

Alternative expression for $\bigoplus_r f_{\omega,r}U_G^{v,r}$

For fixed r, the $f_{\omega,r}$ copies of $U_G^{v,r}$ arise from the $f_{\omega,r}$ distinct ways of choosing $m=(m_1,m_2,\ldots,m_n)\in Z_+^n$ so that (23a) is satisfied. The summation in $\oplus_r f_{\omega,r}U_G^{v,r}$ is over all distinct values of r which are expressible in the form $-\lambda(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j m_j$ $(m_j\in Z_+)$. Therefore,

$$\bigoplus_r f_{\omega,r}U_G^{v,r}=\bigoplus_{m\in Z_+^n}U_G^{v,-\lambda(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j m_j}. \tag{26}$$

Now, from (19),

$$[U_G^{v,-\lambda(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j m_j}(\sigma,\beta,\alpha,S(\tau))\psi](x)$$
$$=\exp iv\sigma\prod_{j=1}^{n}\exp(-iv\beta_j x_j)\prod_{j=1}^{n}\exp\left(-\alpha_j\frac{\partial}{\partial x_j}\right)$$
$$\times\exp i\tau\left[-(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j(m_j+\tfrac{1}{2})\right.$$
$$\left.+\tfrac{1}{2}\sum_{j=1}^{n}\omega_j\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2}+vx_j^2\right)\right]\psi(x).$$

The summation over m of such representations can be simplified by taking $U_G^{v,-\lambda(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j m_j}$ to act on the space $L^2(\mathbb{R}^n,\mathbb{C})\psi_m$ [where ψ_m is defined by (20)]. For each $m\in Z_+^n$,

$$\exp\left[-i\tau(\mathrm{sgn}v)\sum_{j=1}^{n}\omega_j(m_j+\tfrac{1}{2})\right]\psi_m(y)$$
$$=\exp\left[-\frac{i\tau}{2}\sum_{j=1}^{n}\omega_j\left(-\frac{1}{v}\frac{\partial^2}{\partial y_j^2}+vy_j^2\right)\right]\psi_m(y).$$

Now the set $\{\psi_m\colon\ m\in Z_+^n\}$ forms a basis of the space $L^2(\mathbb{R}^n,\mathbb{C})$ of functions from $\{y\in\mathbb{R}^n\}$ to $\mathbb{C}$; so the set $\{\psi_{m_{(1)}}(x)\psi_{m_{(2)}}(y)\colon\ m_{(1)},m_{(2)}\in Z_+^n\}$ is a basis of the space $L^2(\mathbb{R}^{2n},\mathbb{C})$ of functions from $\{(x,y)\in\mathbb{R}^{2n}\}$ to $\mathbb{C}$. Hence, from linearity, it follows that, with $\zeta\in L^2(\mathbb{R}^{2n},\mathbb{C})$,

$$\left(\left[\bigoplus_r f_{\omega,r}U_G^{v,r}\right](\sigma,\beta,\alpha,S(\tau))\zeta\right)(x,y)$$
$$=\exp iv\sigma\prod_{j=1}^{n}\exp(-iv\beta_j x_j)\prod_{j=1}^{n}\exp\left(-\alpha_j\frac{\partial}{\partial x_j}\right)$$
$$\times\exp\frac{i\tau}{2}\left\{\sum_{j=1}^{n}\omega_j\left[\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2}+vx_j^2\right)\right.\right.$$
$$\left.\left.-\left(-\frac{1}{v}\frac{\partial^2}{\partial y_j^2}+vy_j^2\right)\right]\right\}\zeta(x,y). \tag{27}$$

The unitary equivalence of $R^v\downarrow G$ and $\bigoplus_r f_{\omega,r}U_G^{v,r}$

If the representations $R^v\downarrow G$ and $\oplus_r f_{\omega,r}U_G^{v,r}$ are unitarily equivalent, there exists a unitary operator V such that

$$V^{-1}(R^v\downarrow G)V=\bigoplus_r f_{\omega,r}U_G^{v,r}.$$

For each $\mu\in\{\sigma,\beta_1,\beta_2,\ldots,\beta_n,\alpha_1,\alpha_2,\ldots,\alpha_n,\tau\}$, let $\mu(t)$ be the one-dimensional Lie subgroup of G formed by those elements for which the value of every parameter except μ is zero. Let Y_μ, Y'_μ be the representatives of the generators of $\mu(t)$ in the representations $R^v\downarrow G$ and $\oplus_r f_{\omega,r}U_G^{v,r}$, respectively. Then

$$R^v\downarrow G=\prod_\mu\exp(\mu Y_\mu)\quad\text{and}\quad\bigoplus_r f_{\omega,r}U_G^{v,r}=\prod_\mu\exp(\mu Y'_\mu) \tag{28}$$

(where the parameters in the product occur from left to right in the order $\sigma,\beta_1,\beta_2,\ldots,\beta_n,\alpha_1,\alpha_2,\ldots,\alpha_n,\tau$).

Hence, if V exists, it satisfies

$$\prod_\mu\exp(\mu Y'_\mu)=\prod_\mu V^{-1}\exp(\mu Y_\mu)V=\prod_\mu\exp\mu(V^{-1}Y_\mu V),$$

by Stone's theorem. So $Y'_\mu=V^{-1}Y_\mu V$ for each μ. Conversely, if V is a unitary operator such that $Y'_\mu=V^{-1}Y_\mu V$ for each μ, then $\oplus_r f_{\omega,r}U_G^{v,r}=V^{-1}(R^v\downarrow G)V$. Therefore, to demonstrate the unitary equivalence of $R^v\downarrow G$ and $\oplus_r f_{\omega,r}U_G^{v,r}$, it is sufficient to show that there exists a unitary operator V satisfying

$$Y'_\mu=V^{-1}Y_\mu V, \tag{29}$$

for each $\mu\in\{\sigma,\beta_1,\beta_2,\ldots,\beta_n,\alpha_1,\alpha_2,\ldots,\alpha_n,\tau\}$.

Explicitly, V must satisfy, from (25) and (27)–(29),

when $\mu=\sigma$,

$$iv=V^{-1}(iv)V; \tag{30}$$

when $\mu=\beta_j$,

$$ivx_j=V^{-1}\frac{\partial}{\partial q_j}V; \tag{31}$$

when $\mu = \alpha_j$,

$$\frac{\partial}{\partial x_j} = V^{-1}\left(-ivq_j + \frac{\partial}{\partial p_j}\right)V ; \tag{32}$$

when $\mu = \tau$,

$$\sum_{j=1}^{n} \omega_j \left[\frac{i}{2}\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j^2} + vx_j^2\right) - \frac{i}{2}\left(-\frac{1}{v}\frac{\partial^2}{\partial y_j^2} + vy_j^2\right)\right]$$

$$= V^{-1}\left(\sum_{j=1}^{n} \omega_j \left[\frac{iv}{2}(q_j^2 - p_j^2) + p_j \frac{\partial}{\partial q_j} - q_j \frac{\partial}{\partial p_j}\right]\right) V . \tag{33}$$

From the form of these relations, it should be possible to express V, if it exists, as a product of operators $V_1 V_2 \cdots V_n$, where V_j transforms operators dependent on one or both of x_j and y_j into operators dependent on one or both of q_j and p_j, and the action of V_j on operators which are independent of x_j and y_j is simply to take them into the operators obtained by replacing x_k, y_k by q_k, p_k respectively $(k = 1, \ldots, j-1, j+1, \ldots, n)$. Moreover, if $\zeta \in L^2(\mathbb{R}^{2n}, \mathbb{C})$ is a function defined on the space $\{(x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{R}^{2n}\}$, then ζ is measurable and square-integrable when considered as a function of x_j and y_j only, and hence, in this way, ζ can be regarded as an element of $L^2(\mathbb{R}^2, \mathbb{C})$. It follows that V_j can be considered as an operator on a space $L^2(\mathbb{R}^2, \mathbb{C})$. Furthermore, if V_j is unitary when regarded as acting on $L^2(\mathbb{R}^2, \mathbb{C})$, then V_j is unitary as an operator on $L^2(\mathbb{R}^{2n}, \mathbb{C})$.

This reduces the problem to showing the existence of a unitary operator $\tilde{V}$, defined on $L^2(\mathbb{R}^2, \mathbb{C})$, which satisfies [since, trivially, (30) is satisfied by any V]

$$ivx = \tilde{V}^{-1}\frac{\partial}{\partial q}\tilde{V} , \tag{34}$$

$$\frac{\partial}{\partial x} = \tilde{V}^{-1}\left(-ivq + \frac{\partial}{\partial p}\right)\tilde{V} , \tag{35}$$

$$\frac{i}{2}\left[\left(-\frac{1}{v}\frac{\partial^2}{\partial x^2} + vx^2\right) - \left(-\frac{1}{v}\frac{\partial^2}{\partial y^2} + vy^2\right)\right]$$

$$= \tilde{V}^{-1}\left[\frac{iv}{2}(q^2 - p^2) + p\frac{\partial}{\partial q} - q\frac{\partial}{\partial p}\right]\tilde{V} . \tag{36}$$

The existence of the operator $\tilde{V}$ will be shown by explicit construction, in two stages.

First, let $\zeta \in L^2(\mathbb{R}^2, \mathbb{C})$ be a function from the space $\{(x, y) \in \mathbb{R}^2\}$ to C; define a Fourier transform F_1 of ζ into the space $L^2(\mathbb{R}^2, \mathbb{C})$ of functions from the space $\{(q, p) \in \mathbb{R}^2\}$ to C by

$$(F_1\zeta)(q, p) = \left(\frac{v}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} \exp(ivxq)\zeta(x, p)dx . \tag{37}$$

By Plancherel's theorem,[13] F_1 is a unitary operator.

Let $\phi \in L^2(\mathbb{R}^2, \mathbb{C})$ be a function from the space $\{(q, p) \in \mathbb{R}^2\}$ to C; then F_1^{-1} maps ϕ into the space $L^2(\mathbb{R}^2, \mathbb{C})$ of functions from the space $\{(x, y) \in \mathbb{R}^2\}$ to C, and is given by

$$(F_1^{-1}\phi)(x, y) = \left(\frac{v}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} \exp(-ivxq)\phi(q, y)dq .$$

Then

$$F_1^{-1}\frac{\partial}{\partial q}F_1 = ivx , \quad F_1^{-1}(ivq)F_1 = -\frac{\partial}{\partial x} ,$$

$$F_1^{-1}\frac{\partial}{\partial p}F_1 = \frac{\partial}{\partial y} , \quad F_1^{-1}(ivp)F_1 = ivy . \tag{38}$$

So, for the operators occurring in (34) and (35),

$$F_1^{-1}\frac{\partial}{\partial q}F_1 = ivx ,$$

and

$$F_1^{-1}\left(-ivq + \frac{\partial}{\partial p}\right)F_1 = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} .$$

Hence the introduction of the operator F_1 reduces the problem of finding a unitary operator $\tilde{V}$ satisfying (34) and (35) to that of finding a unitary operator U satisfying

$$ivx = U^{-1}(ivx)U , \tag{39}$$

$$\frac{\partial}{\partial x} = U^{-1}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)U . \tag{40}$$

Define an operator $\tilde{U}$ on $L^2(\mathbb{R}^2, \mathbb{C})$ by

$$(\tilde{U}\zeta)(x, y) = \zeta(u_{11}x + u_{12}y, u_{21}x + u_{22}y) , \tag{41a}$$

where $u_{11}, u_{12}, u_{21}, u_{22} \in \mathbb{R}$.

$\tilde{U}$ is unitary if $|\Delta| = 1$, where $\Delta = u_{11}u_{22} - u_{12}u_{21}$.

$\tilde{U}^{-1}$ is given by

$$(\tilde{U}^{-1}\zeta)(x, y) = \zeta\left(\frac{u_{22}}{\Delta}x - \frac{u_{12}}{\Delta}y, -\frac{u_{21}}{\Delta}x + \frac{u_{11}}{\Delta}y\right) . \tag{41b}$$

It follows that

$$\tilde{U}^{-1}(ivx)\tilde{U} = \frac{iv}{\Delta}(u_{22}x - u_{12}y) ,$$

$$\tilde{U}^{-1}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)\tilde{U} = u_{11}\frac{\partial}{\partial x} + u_{21}\frac{\partial}{\partial y} + u_{12}\frac{\partial}{\partial x} + u_{22}\frac{\partial}{\partial y} .$$

Hence, when $u_{11} = 1$, $u_{12} = 0$, $u_{21} = -\Delta$, and $u_{22} = \Delta$, $\tilde{U}$ satisfies conditions (39) and (40). To make $\tilde{U}$ unitary, Δ may be taken as 1.

Therefore, when $\tilde{V} = F_1 U$, with

$$(U\zeta)(x, y) = \zeta(x, -x + y) , \tag{42}$$

conditions (34) and (35) are satisfied. It may be verified that, with this choice of $\tilde{V}$, the remaining condition (36) is also satisfied.

A unitary operator V satisfying (30)–(33) can now be constructed. Let $\zeta \in L^2(\mathbb{R}^{2n}, \mathbb{C})$ be a function from the space $\{(x, y) \in \mathbb{R}^{2n}\}$ to C, and define the operator V mapping ζ into the space $L^2(\mathbb{R}^{2n}, \mathbb{C})$ of functions from the space $\{(q, p) \in \mathbb{R}^{2n}\}$ to C by

$$(V\zeta)(q_1, \ldots, q_n, p_1, \ldots, p_n)$$

$$= \left(\frac{v}{2\pi}\right)^{n/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp iv\sum_{j=1}^{n} x_j q_j \, \zeta(x_1, \ldots, x_n,$$

$$-x_1 + p_1, \ldots, -x_n + p_n)dx_1 \cdots dx_n . \tag{43}$$

From the properties of $\tilde{V}$, it follows that V is a unitary operator on $L^2(\mathbb{R}^{2n}, \mathbb{C})$ satisfying conditions (30)–(33).

Hence $Y'_\mu = V^{-1} Y_\mu V$, for each $\mu \in \{\sigma, \beta_1, \beta_2, \ldots, \beta_n, \alpha_1, \alpha_2, \ldots, \alpha_n, \tau\}$; this implies that

$$\prod_\mu \exp(\mu Y'_\mu) = V^{-1} \left[\prod_\mu \exp(\mu Y_\mu) \right] V .$$

Therefore,

$$V^{-1}(R^v \downarrow G)V = \bigoplus_r f_{\omega,r} U_G^{v,r}, \tag{44a}$$

where V is the unitary operator defined by (43).

(b) *Suppose the ω_j's are not rationally related*

An alternative expression for $\oplus_h f_{\omega,h} U_G^{v,h}$ may be obtained by a method similar to that above, using the representations $U_G^{v,h}$ [Ref. 10, Eq. (25)] instead of the representations $U_G^{v,r}$ [Ref. 10, Eq. (26)]. It follows that, when the ω_j's are not rationally related, $([\oplus_h f_{\omega,h} U_G^{v,h}](\sigma, \beta, \alpha, S(\tau))\xi)(x, y)$ may be expressed in the form of the right-hand side of (27), where now $\tau \in \mathbf{R}$. The same argument as above then gives that

$$V^{-1}(R^v \downarrow G)V = \bigoplus_h f_{\omega,h} U_G^{v,h}, \tag{44b}$$

where V is the unitary operator defined by (43).

8. THE SIGNIFICANCE OF THE DECOMPOSITION OF $R^v \downarrow G$

(a) *Suppose the ω_j's are rationally related*

$$R^v \downarrow G \simeq \bigoplus_r f_{\omega,r} U_G^{v,r},$$

where the summation is over all distinct values of r which are expressible in the form $-\lambda(\mathrm{sgn} v)\sum_{j=1}^n \omega_j m_j$ ($m_j \in \mathbf{Z}_+$), and, for fixed r, $f_{\omega,r}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbf{Z}_+^n$ so that $r = -\lambda(\mathrm{sgn} v)\sum_{j=1}^n \omega_j m_j$ is satisfied.

For fixed r occurring in the summation, let $s = -r/\mathrm{sgn} v$. Then $s \in \mathbf{Z}_+$, and $f_{\omega,r}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbf{Z}_+^n$ so that $s = \lambda\sum_{j=1}^n \omega_j m_j$ is satisfied. Hence $f_{\omega,r}$ is precisely the degeneracy $d_{\omega,s}$ of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with rationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$; moreover, all energy levels of the system occur in this way.

Therefore,

$$R^v \downarrow G \simeq \bigoplus_s d_{\omega,s} U_G^{v,-(\mathrm{sgn} v)s}, \tag{45a}$$

where $d_{\omega,s}$ is the degeneracy of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with rationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$, and the summation may be regarded as over all distinct energy levels $E_{\omega,s}$ of the system.

(b) *Suppose the ω_j's are not rationally related*

$$R^v \downarrow G \simeq \bigoplus_h f_{\omega,h} U_G^{v,h},$$

where the summation is over all distinct values of h which are expressible in the form $-(\mathrm{sgn} v)\sum_{j=1}^n \omega_j (m_j + \frac{1}{2})$ ($m_j \in \mathbf{Z}_+$), and, for fixed h, $f_{\omega,h}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbf{Z}_+^n$ so that $h = -(\mathrm{sgn} v)\sum_{j=1}^n \omega_j (m_j + \frac{1}{2})$ is satisfied.

For fixed h occurring in the summation, let $e = -h/\mathrm{sgn} v$. Then $f_{\omega,h}$ is the number of distinct ways of choosing $m = (m_1, m_2, \ldots, m_n) \in \mathbf{Z}_+^n$ so that $e = \sum_{j=1}^n \omega_j (m_j + \frac{1}{2})$ is satisfied. Hence $f_{\omega,h}$ is precisely the degeneracy $d_{\omega,e}$ of the energy level $E_{\omega,e}$ of the n-dimensional anisotropic harmonic oscillator with nonrationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$; moreover, all energy levels of the system occur in this way.

Therefore,

$$R^v \downarrow G \simeq \bigoplus_e d_{\omega,e} U_G^{v,-(\mathrm{sgn} v)e}, \tag{45b}$$

where $d_{\omega,e}$ is the degeneracy of the energy level $E_{\omega,e}$ of the n-dimensional anisotropic harmonic oscillator with nonrationally related frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$, and the summation may be regarded as over all distinct energy levels $E_{\omega,e}$ of the system.

In conclusion, the degeneracies of the energy levels of the anisotropic harmonic oscillator can thus be viewed as multiplicities of irreducible representations of G in the decomposition of the representation R^v of L restricted to G.

The group G is intrinsically related to the anisotropic harmonic oscillator, since it is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system. On the other hand, the group L has no direct relation to the anisotropic harmonic oscillator: L is simply a semidirect product of the Weyl group, N, and the symplectic group, $\mathrm{Sp}(2n, \mathbf{R})$. As such, it would be considered a part of quantum mechanics, rather than classical mechanics. Nevertheless, it should be pointed out that the group L/C (where C is the center of L) can also be regarded as the group of invertible inhomogeneous linear canonical transformations in classical mechanics. The representation R^v of the group L will be analyzed in more detail in a later paper.

ACKNOWLEDGMENTS

I should like to thank Professor J. T. Lewis for his advice and encouragement. I am also grateful to the Science Research Council for financial support while most of the work for this paper was done.

[1] J.M. Jauch and E.L. Hill, Phys. Rev. **57**, 641 (1940).
[2] V.A. Dulock and H.V. McIntosh, Am. J. Phys. **33**, 109 (1965).
[3] Yu.N. Demkov, Zh. Eksp. Teor. Fiz. **44**, 2007 (1963) [Sov. Phys. JETP **17**, 1349 (1963)].
[4] L.A. Il'kaeva, Vestn. Leningr. Univ. **22**, 56 (1963) (in Russian).
[5] A. Cisneros and H.V. McIntosh, J. Math. Phys. **11**, 870 (1970).
[6] J.D. Louck, M. Moshinsky, and K.B. Wolf, J. Math. Phys. **14**, 692 (1973).
[7] M. Moshinsky, J. Patera, and P. Winternitz, J. Math. Phys. **16**, 82 (1975).
[8] I. Vendramin, Nuovo Cimento A **54**, 190 (1968).
[9] G. Maiella and G. Vilasi, Lett. Nuovo Cimento **1**, 57 (1969).
[10] M.E. Major, J. Math. Phys. **18**, 1938 (1977).
[11] L. van Hove, Mem. Acad. Roy. Belg. Cl. Sci. 8°, **26**, fasc. 6, 1 (1951).
[12] G.W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, and Editore Boringhieri, Torino, 1968).
[13] K. Yosida, *Functional Analysis* (Springer, Berlin, 1965).

A noninvariance group for the n-dimensional isotropic harmonic oscillator

M. E. Major

Mathematics Department, Queen Elizabeth College, London W8 7AH, England
(Received 22 September 1975)

A semidirect product of the Weyl group N, with the special unitary group SU(n), is proved to be a possible noninvariance group for the n-dimensional isotropic harmonic oscillator. In order to obtain this result, a method is developed for finding the representation W of SU(n) which intertwines the representations U_N^v and $(A+iB)U_N^v$ of N [where $A+iB\in$ SU(n)]; it is also shown that $W=\oplus_{a=0}^{\infty}W_a$, where W_a is an irreducible representation of SU(n), of dimension equal to the degeneracy of the $(a+1)$th energy level of the n-dimensional isotropic harmonic oscillator.

1. INTRODUCTION

In the literature, the term "symmetry group" or "invariance group" is sometimes used with little explanation. However, most authors, in talking about such a group, mean a degeneracy group, defined as follows. Let D be a subgroup of the group of all unitary operators commuting with the Hamiltonian of a quantum mechanical system; then the eigenfunctions of each degenerate energy level form the basis for a representation of D.[1] D is a degeneracy group for the system if:

(i) to each energy level of the system, there corresponds an irreducible representation of D;

(ii) no irreducible representation of D corresponds to more than one energy level of the system, but some irreducible representations of D may not correspond to any energy level.

In general, once such a group has been found for a system, the accidental degeneracy is regarded as being explained. In particular, SU(n) is a degeneracy group for the n-dimensional isotropic harmonic oscillator.[2]

For some systems in which the accidental degeneracy can be explained by means of a degeneracy group, larger "approximate" symmetry groups have been studied. Those irreducible representations of a degeneracy group of a system which occur as representations on the eigenspaces of the Hamiltonian are collected into one irreducible representation of a larger group. Such a larger group is known as a noninvariance group. Precisely, a noninvariance group for a system is defined to be a Lie group $G_{\rm NI}$ with the following properties:

(i) $G_{\rm NI}$ contains a degeneracy group D of the system as a subgroup;

(ii) the Lie algebra of $G_{\rm NI}$ includes certain "noninvariant" generators which do not commute with the Hamiltonian;

(iii) $G_{\rm NI}$ has an irreducible representation whose restriction to the group D is unitarily equivalent to $\oplus_m \rho_m$, where ρ_m is an irreducible representation of D of dimension equal to the degeneracy of the mth energy level of the system, and the summation may be regarded as over all energy levels of the system.

The concept of such a group was apparently first suggested by Barut,[3] in 1964; in 1965, Mukunda, O'Raifeartaigh, and Sudarshan[4] reported that SU(n, 1) is a noninvariance group for the n-dimensional isotropic harmonic oscillator.

The main result proved in this paper is that a semidirect product of the Weyl group, N, with the special unitary group, SU(n), is a possible noninvariance group for the n-dimensional isotropic harmonic oscillator.

First, all the quantum mechanical irreducible representations of $N\textcircled{S}$ SU(n) are found explicitly, using Mackey's theory of induced representations. In order to do this, a method is developed for finding the projective representation W which intertwines the representations U_N^v and $(A+iB)U_N^v$ of N [where $A+iB\in$ SU(n)]. Some of the simplest of the quantum mechanical irreducible representations of $N\textcircled{S}$ SU(n), when restricted to SU(n), become just the representation W of SU(n). It is proved that $W=\oplus_{a=0}^{\infty}W_a$, where W_a is an irreducible representation of SU(n), of dimension equal to the degeneracy of the $(a+1)$th energy level of the n-dimensional isotropic harmonic oscillator: These results show that $N\textcircled{S}$ SU(n) is a possible noninvariance group for the n-dimensional isotropic harmonic oscillator.

2. THE LITTLE GROUP OF THE REPRESENTATION U_N^v

A typical element of the semidirect product $N\textcircled{S}$ SU(n) of the Weyl group N and the special unitary group SU(n) is denoted by (σ,β,α,U), where $\sigma\in\mathbb{R}$, $\beta\in\mathbb{R}^n$, $\alpha\in\mathbb{R}^n$, and $U\in$ SU(n); the group law is taken to be

$$(\sigma',\beta',\alpha',U')(\sigma,\beta,\alpha,U)=(\sigma'',\beta'',\alpha'',U''), \tag{1}$$

where

$$\sigma''=\sigma'+\sigma+\tfrac{1}{2}A'\beta\cdot B'\beta-\tfrac{1}{2}A'\alpha\cdot B'\alpha - B'\beta\cdot B'\alpha+\alpha'\cdot(A'\beta-B'\alpha),$$

$$\beta''=A'\beta-B'\alpha+\beta',$$

$$\alpha''=B'\beta+A'\alpha+\alpha',$$

$$U''=U'U,$$

with $U'=A'+iB'\in$ SU(n), and $U\in$ SU(n). [This definition is just the restriction to the subgroup $N\textcircled{S}$ SU(n) of the group law of $N\textcircled{S}$ Sp($2n$, R) ($=L$), introduced in Ref. 5; SU(n) has been identified with a subgroup of Sp($2n$, R) under the isomorphism

$$A+iB\ [\in \mathrm{SU}(n)] \leftrightarrow \begin{pmatrix} A & B \\ -B & A \end{pmatrix}\ [\in \mathrm{Sp}(2n,\mathbb{R})].$$

Every quantum mechanical irreducible representation of N is unitarily equivalent to one of the form [Ref. 6, Eq. (15)]

$$[U_N^v(\sigma,\beta,\alpha,I)\psi](x) = \exp iv(\sigma - x\,.\,\beta)\,\psi(x-\alpha), \tag{2}$$

where $v(\neq 0)\in \mathbb{R}$, $\psi \in L^2(\mathbb{R}^n,\mathbb{C})$.

The action of $A+iB \in \mathrm{SU}(n)$ on U_N^v is, by definition [Ref. 6, Eq. (11)],

$$\begin{aligned}&[((A+iB)\,U_N^v)(\sigma,\beta,\alpha,I)\psi](x)\\ &\quad=[U_N^v((0,0,0,(A+iB)^{-1})\\ &\qquad\times(\sigma,\beta,\alpha,I)(0,0,0,A+iB))\psi](x)\\ &\quad=\exp iv(\sigma+\tfrac{1}{2}(A^T\alpha\,.\,B^T\alpha - A^T\beta\,.\,B^T\beta) - B^T\beta\,.\,B^T\alpha\\ &\qquad - x\,.\,(A^T\beta+B^T\alpha))\,\psi(x+B^T\beta-A^T\alpha),\end{aligned} \tag{3}$$

from (1) and (2), with T denoting transpose.

$(A+iB)\,U_N^v$ is an irreducible representation of N. Restricted to the subgroup $\{(\sigma,0,0,I)\}$, it is just the phase $\exp iv\sigma$. Hence $(A+iB)\,U_N^v$ must be unitarily equivalent to U_N^v, since U_N^v is the only irreducible representation of N which has the required form on restriction to the subgroup $\{(\sigma,0,0,I)\}$.

Therefore, there exists a unitary operator $W(A+iB)$, dependent on $A+iB$, such that

$$(A+iB)\,U_N^v = W(A+iB)^{-1}\,U_N^v\,W(A+iB), \tag{4}$$

for each $A+iB \in \mathrm{SU}(n)$.

In order to obtain all the irreducible representations of $N \circledS \mathrm{SU}(n)$ explicitly, the operator W must be found (Ref. 6, Sec. 5).

3. A METHOD OF FINDING THE OPERATOR W

A. General approach

Since $\mathrm{SU}(n)$ is a separable, compact, connected and simply connected group, it has no nontrivial multipliers[7]; hence W can always be chosen to be an ordinary representation W_0 of $\mathrm{SU}(n)$. The following method of finding W_0 depends on the use of one-parameter subgroups of N and $\mathrm{SU}(n)$.

Let H be an r-dimensional connected Lie group. It may be proved that every element of H can be expressed as a finite product of the form

$$\begin{aligned}h =\ & h_1(\tau_{11})h_2(\tau_{21})\cdots h_r(\tau_{r1})h_1(\tau_{12})h_2(\tau_{22})\\ &\times\cdots h_r(\tau_{r2})\cdots h_1(\tau_{1s})h_2(\tau_{2s})\cdots h_r(\tau_{rs}),\end{aligned}$$

where $h_1(\tau),h_2(\tau),\ldots,h_r(\tau)$ are independent one-parameter subgroups of H.[8] For convenience of notation, this will be written as

$$h = \prod_j \prod_i h_i(\tau_{ij})$$

(where the order of the terms is important).

Since $\mathrm{SU}(n)$ is a connected Lie group, every element of $\mathrm{SU}(n)$ can be expressed as a finite product of the above form, with r = dimension of $\mathrm{SU}(n) = n-1$.

Since W_0 is an ordinary representation of $\mathrm{SU}(n)$,

$$W_0(h) = \prod_j \prod_i W_0(h_i(\tau_{ij})). \tag{5}$$

Hence, in order to find W_0, it is enough to determine the restriction of W_0 to a set of $(n-1)$ independent one-parameter subgroups of $\mathrm{SU}(n)$.

For each i, the map $h_i(\tau) \to W_0(h_i(\tau))$ is an ordinary unitary representation of the one-parameter subgroup $h_i(\tau)$ of $\mathrm{SU}(n)$. Hence, by Stone's theorem,[9] there exists a unique skew-adjoint operator J_i such that

$$W_0(h_i(\tau)) = \exp \tau J_i. \tag{6}$$

It is thus sufficient to find the J_i's corresponding to a set of $(n-1)$ independent one-parameter subgroups of $\mathrm{SU}(n)$.

For each one-parameter subgroup $h(\tau)$ of $\mathrm{SU}(n)$, the unique skew-adjoint operator determined by $W_0(h(\tau)) = \exp \tau J$ satisfies, from (4),

$$\begin{aligned}&(h(\tau)\,U_N^v)(n')\\ &\quad=\exp(-\tau J)\,U_N^v(n')\exp(\tau J), \quad \text{for each } n'\in N.\end{aligned} \tag{7}$$

Some of the properties of the operator J will now be examined: This will indicate how the operator J can be calculated.

Now each element $(\sigma,\beta,\alpha,I)\in N$ can be expressed as a product of elements of one-parameter subgroups, using (1), by

$$\begin{aligned}&(\sigma,\beta,\alpha,I)\\ &\quad=(\sigma,0,0,I)\prod_j(0,\beta_j e_j,0,I)\prod_j(0,0,\alpha_j e_j,I),\end{aligned} \tag{8}$$

where e_j $(\in \mathbb{R}^n)$ is a unit vector with 1 in the jth position and zeros elsewhere. For each parameter $\mu \in \{\sigma,\beta_1,\beta_2,\ldots,\beta_n,\alpha_1,\alpha_2,\ldots,\alpha_n\}$, let $\mu(t)$ be the one-dimensional Lie subgroup of N formed by those group elements for which the value of every parameter except μ is zero. For each such one-parameter subgroup $\mu(t)$, the map $\mu(t) \to (h(\tau)\,U_N^v)(\mu(t))$ is an ordinary unitary representation. Hence, by Stone's theorem, there exists a unique skew-adjoint operator $Z_\mu(\tau)$ such that

$$(h(\tau)\,U_N^v)(\mu(t)) = \exp tZ_\mu(\tau). \tag{9}$$

So, for each μ,

$$\exp tZ_\mu(\tau) = \exp(-\tau J)\exp(tZ_\mu(0))\exp(\tau J). \tag{10}$$

Differentiating with respect to t and putting $t=0$, and then differentiating the result with respect to τ and putting $\tau = 0$ gives

$$\left.\frac{dZ_\mu(\tau)}{d\tau}\right|_{\tau=0} = [Z_\mu(0), J] \tag{11}$$

for each μ.

The existence of a skew-adjoint operator J satisfying (11) follows from the existence of $W_0(\tau) = \exp \tau J$.

Conversely, suppose $\tilde{J}$ is a skew-adjoint operator satisfying (11). To investigate the uniqueness of J, let $J' = J - \tilde{J}$. Then $[Z_\mu(0), J'] = 0$ for each μ. The operators

$Z_\mu(0)$ give the representation of the Lie algebra of N corresponding to the representation U_N^v of the Lie group N. Since U_N^v is an irreducible representation, it follows, by Schur's lemma, that J' is a scalar multiple of the identity operator; since J' is a skew-adjoint operator, J' must equal ξI, where ξ is imaginary.

Hence every skew-adjoint operator $\tilde{J}$ satisfying (11) also satisfies (10), and (11) determines the skew-adjoint operator J up to an arbitrary constant.

Therefore, for each i, the ordinary representation W: $h_i(\tau) \to \exp\tau(\tilde{J}_i + \xi_i I)$ of the one-parameter subgroup $h_i(\tau)$ of SU(n) satisfies the restriction of (4) to $h_i(\tau)$ [where $\tilde{J}_i$ is a skew-adjoint operator satisfying the condition analogous to (11)]. However, because of the arbitrary constants ξ_i, the map

$$W: \prod_j \prod_i h_i(\tau_{ij}) \to \prod_j \prod_i \exp\tau_{ij}(\tilde{J}_i + \xi_i I)$$

may not be an ordinary representation of SU(n). To ensure that W is an ordinary representation of SU(n), the constants ξ_i cannot be allowed to be arbitrary, but must be appropriately chosen. If W is an ordinary representation of SU(n), it must give an ordinary representation w of the Lie algebra of SU(n); this leads to constraints on the values of the constants ξ_i.

B. Detailed calculation

In order to find a set of independent one-parameter subgroups which generate SU(n), a basis for the Lie algebra of SU(n) is first obtained. Now the Lie algebra of SU(n) is the algebra of all $n\times n$ complex skew-adjoint matrices of trace zero, denoted by su(n).[10]

A basis for su(n) is the set of (n^2-1) elements,

$$\begin{aligned} R_{jk}^{(1)} &= i(E_{jk}+E_{kj}) \quad (1\le j<k\le n), \\ R_{jk}^{(2)} &= E_{jk}-E_{kj} \quad (1\le j<k\le n), \\ H_k &= i\left(E_{kk}-\frac{1}{n}I_n\right) \quad (2\le k\le n) \end{aligned} \tag{12}$$

[where E_{lm} is the $n\times n$ matrix with 1 in the (l,m) position and zeros elsewhere, and I_n is the $n\times n$ identity matrix].

If $h(\tau)=A(\tau)+iB(\tau)=(a_{jk}(\tau)+ib_{jk}(\tau))$ is a one-parameter subgroup of SU(n), then, from (3) and (9), the representatives of the generators of N in the representation $h(\tau)\,U_N^v$ of N are

$$\begin{aligned} Z_\sigma(\tau) &= iv, \\ Z_{\beta_j}(\tau) &= -iv\sum_{k=1}^n a_{jk}(\tau)x_k + \sum_{k=1}^n b_{jk}(\tau)\frac{\partial}{\partial x_k}, \\ Z_{\alpha_j}(\tau) &= -iv\sum_{k=1}^n b_{jk}(\tau)x_k - \sum_{k=1}^n a_{jk}(\tau)\frac{\partial}{\partial x_k}. \end{aligned} \tag{13}$$

From (11), J must satisfy
when $\mu=\sigma$,

$$0=[iv,J],$$

which is satisfied by any J;
when $\mu=\beta_j$,

$$-iv\sum_{k=1}^n a'_{jk}(0)x_k + \sum_{k=1}^n b'_{jk}(0)\frac{\partial}{\partial x_k} = [-ivx_j, J]; \tag{14}$$

when $\mu=\alpha_j$,

$$-iv\sum_{k=1}^n b'_{jk}(0)x_k - \sum_{k=1}^n a'_{jk}(0)\frac{\partial}{\partial x_k} = \left[-\frac{\partial}{\partial x_j}, J\right] \tag{15}$$

(where $'$ denotes differentiation with respect to τ).

(i) *The operator J for the subgroup generated by $R_{lm}^{(1)}$*

Suppose, with the above notation, that $R_{lm}^{(1)}$ is the generator of the one-parameter subgroup $h(\tau)=A(\tau)+iB(\tau)=(a_{jk}(\tau)+ib_{jk}(\tau))$. Then

$$R_{lm}^{(1)} = A'(0)+iB'(0) = (a'_{jk}(0)+ib'_{jk}(0)).$$

In this case, the only nonzero elements of $A'(0)$ and $B'(0)$ are

$$b'_{lm}(0) = 1 = b'_{ml}(0).$$

Hence the conditions (14) and (15) become

for $j\neq l$ or m,

$$0=[-ivx_j,J] \quad \text{and} \quad 0=\left[-\frac{\partial}{\partial x_j},J\right];$$

for $j=l$ $(\neq m)$,

$$\frac{\partial}{\partial x_m}=[-ivx_l,J] \quad \text{and} \quad -ivx_m=\left[-\frac{\partial}{\partial x_l},J\right];$$

for $j=m$ $(\neq l)$,

$$\frac{\partial}{\partial x_l}=[-ivx_m,J] \quad \text{and} \quad -ivx_l=\left[-\frac{\partial}{\partial x_m},J\right].$$

These relations are all satisfied by the skew-adjoint operator

$$J = w(R_{jk}^{(1)}) = \frac{1}{iv}\frac{\partial^2}{\partial x_l \partial x_m} + ivx_l x_m + \xi_{lm}^{(1)} I, \tag{16a}$$

where $\xi_{lm}^{(1)}$ is an arbitrary imaginary constant.

(ii) *The operator J for the subgroup generated by $R_{lm}^{(2)}$*

Similarly, the operator J for the subgroup generated by $R_{lm}^{(2)}$ must satisfy

for $j\neq l$ or m,

$$0=[-ivx_j,J] \quad \text{and} \quad 0=\left[-\frac{\partial}{\partial x_j},J\right];$$

for $j=l$ $(\neq m)$,

$$-ivx_m=[-ivx_l,J] \quad \text{and} \quad -\frac{\partial}{\partial x_m}=\left[-\frac{\partial}{\partial x_l},J\right];$$

for $j=m$ $(\neq l)$,

$$ivx_l=[-ivx_m,J] \quad \text{and} \quad \frac{\partial}{\partial x_l}=\left[-\frac{\partial}{\partial x_m},J\right].$$

These relations are all satisfied by the skew-adjoint operator

$$J = w(R_{lm}^{(2)}) = x_l\frac{\partial}{\partial x_m} - x_m\frac{\partial}{\partial x_l} + \xi_{lm}^{(2)} I, \tag{16b}$$

where $\xi_{lm}^{(2)}$ is an arbitrary imaginary constant.

(iii) *The operator J for the subgroup generated by H_m*

The operator J for the subgroup generated by H_m must satisfy

for $j \neq m$,

$$-\frac{1}{n}\frac{\partial}{\partial x_j} = [-ivx_j, J] \quad \text{and} \quad \frac{iv}{n}x_j = \left[-\frac{\partial}{\partial x_j}, J\right];$$

for $j = m$,

$$\left(1-\frac{1}{n}\right)\frac{\partial}{\partial x_m} = [-ivx_m, J] \quad \text{and} \quad -iv\left(1-\frac{1}{n}\right)x_m = \left[-\frac{\partial}{\partial x_m}, J\right].$$

These relations are all satisfied by the skew-adjoint operator

$$J = w(H_m) = \frac{i}{2}\left(-\frac{1}{v}\frac{\partial^2}{\partial x_m^2} + vx_m^2\right) - \frac{i}{2n}\sum_{k=1}^{n}\left(-\frac{1}{v}\frac{\partial^2}{\partial x_k^2} + vx_k^2\right) + \xi_m I, \tag{16c}$$

where ξ_m is an arbitrary imaginary constant.

(iv) *Calculation of the values of the constants* $\xi_{jm}^{(1)}$, $\xi_{jm}^{(2)}$, *and* ξ_m

It may be shown that the commutation relations of su(n), in terms of the basis $R_{jk}^{(1)}, R_{jk}^{(2)}, H_k$ [see (12)], are

$$\begin{aligned}
&[R_{jk}^{(1)}, R_{lm}^{(1)}] = -(\delta_{km}R_{jl}^{(2)} + \delta_{kl}R_{jm}^{(2)} + \delta_{jm}R_{kl}^{(2)} + \delta_{jl}R_{km}^{(2)}),\\
&[R_{jk}^{(1)}, R_{lm}^{(2)}] = -\delta_{km}R_{jl}^{(1)} + \delta_{kl}R_{jm}^{(1)} - \delta_{jm}R_{kl}^{(1)} + \delta_{jl}R_{km}^{(1)},\\
&[R_{jk}^{(2)}, R_{lm}^{(2)}] = -\delta_{km}R_{jl}^{(2)} + \delta_{kl}R_{jm}^{(2)} + \delta_{jm}R_{kl}^{(2)} - \delta_{jl}R_{km}^{(2)},\\
&[R_{jk}^{(1)}, H_l] = (\delta_{jl} - \delta_{kl})R_{jk}^{(2)},\\
&[R_{jk}^{(2)}, H_l] = (\delta_{kl} - \delta_{jl})R_{jk}^{(1)},\\
&[H_l, H_m] = 0,
\end{aligned} \tag{17}$$

where, for convenience of notation, define

$$R_{jk}^{(1)} = R_{kj}^{(1)}, \quad R_{jk}^{(2)} = -R_{kj}^{(2)}, \quad \text{for } j > k;$$

$$H_1 = -\sum_{j=2}^{n} H_j;$$

$$R_{jj}^{(1)} = 2H_j, \quad R_{jj}^{(2)} = 0, \quad \text{for } j = 1, 2, \ldots, n.$$

Since w is to be an ordinary representation of su(n), $w(R_{jk}^{(1)})$, $w(R_{jk}^{(2)})$, and $w(H_k)$ must satisfy the commutation relations of su(n). Explicit calculation shows that this implies

$$\delta_{km}\xi_{jl}^{(2)} + \delta_{kl}\xi_{jm}^{(2)} + \delta_{jm}\xi_{kl}^{(2)} + \delta_{jl}\xi_{km}^{(2)} = 0, \tag{18a}$$

$$\delta_{km}\xi_{jl}^{(1)} - \delta_{kl}\xi_{jm}^{(1)} + \delta_{jm}\xi_{kl}^{(1)} - \delta_{jl}\xi_{km}^{(1)} = 0, \tag{18b}$$

$$\delta_{km}\xi_{jl}^{(2)} - \delta_{kl}\xi_{jm}^{(2)} - \delta_{jm}\xi_{kl}^{(2)} + \delta_{jl}\xi_{km}^{(2)} = 0, \tag{18c}$$

$$(-\delta_{jl} + \delta_{kl})\xi_{jk}^{(2)} = 0, \tag{18d}$$

$$(-\delta_{kl} + \delta_{jl})\xi_{jk}^{(1)} = 0, \tag{18e}$$

where, for convenience of notation, define

$$\xi_{jk}^{(1)} = \xi_{kj}^{(1)}, \quad \xi_{jk}^{(2)} = -\xi_{kj}^{(2)}, \quad \text{for } j > k;$$

$$\xi_1 = -\sum_{j=2}^{n}\xi_j;$$

$$\xi_{jj}^{(1)} = 2\xi_j, \quad \xi_{jj}^{(2)} = 0, \quad \text{for } j = 1, 2, \ldots, n.$$

Conditions (18d) and (18e) give that $\xi_{jk}^{(1)} = 0 = \xi_{jk}^{(2)}$ for $j < k$ (and hence for $j > k$ as well). Condition (18b) gives that, when $j = l$, $k = m$ $(j < k)$, $\xi_{jj}^{(1)} - \xi_{kk}^{(1)} = 0$, so $\xi_j = \xi_k$; but $\sum_{j=1}^{n}\xi_j = 0$, so $\xi_j = 0$ for all $j = 1, 2, \ldots, n$.

Hence the condition that w is an ordinary representation of su(n) is sufficient to determine the ξ's uniquely: They must all be zero. Therefore, the ordinary representation W of SU(n) satisfying (4) is uniquely determined. $W(A + iB)$ can be written down explicitly, once $A + iB$ has been expressed as a product of elements of the one-parameter subgroups of SU(n) generated by $R_{jk}^{(1)}$, $R_{jk}^{(2)}$, and H_k.

Note: An alternative expression for $W(A + iB)$ can be obtained as follows. Since SU(n) is a connected, simply connected, exponential Lie group,[11] the ordinary representation of su(n) determined by w [(16)] with all the ξ's zero exponentiates to an ordinary representation of SU(n);[12] moreover, this representation is precisely the representation W determined by (4).

Therefore, if $A + iB \in \mathrm{SU}(n)$ is parametrized by $\tau_{jk}^{(1)}$, $\tau_{jk}^{(2)}$, $\tau_k \in \mathrm{R}$ $(1 \leq j < k \leq n)$, so that

$$A + iB = \exp\left[\sum_{1 \leq j < k \leq n}(\tau_{jk}^{(1)}R_{jk}^{(1)} + \tau_{jk}^{(2)}R_{jk}^{(2)}) + \sum_{k=2}^{n}\tau_k H_k\right],$$

then

$$\begin{aligned}
W(A + iB) = \exp\Bigg\{&\sum_{1 \leq j < k \leq n}\Bigg[i\tau_{jk}^{(1)}\left(-\frac{1}{v}\frac{\partial^2}{\partial x_j \partial x_k} + vx_j x_k\right)\\
&+ \tau_{jk}^{(2)}\left(x_j\frac{\partial}{\partial x_k} - x_k\frac{\partial}{\partial x_j}\right)\Bigg] + \sum_{k=2}^{n}\frac{i}{2}\tau_k\Bigg[-\frac{1}{v}\frac{\partial^2}{\partial x_k^2}\\
&+ vx_k^2 - \frac{1}{n}\sum_{m=1}^{n}\left(-\frac{1}{v}\frac{\partial^2}{\partial x_m^2} + vx_m^2\right)\Bigg]\Bigg\}
\end{aligned} \tag{19}$$

is the unique ordinary representation of SU(n) satisfying (4).

4. THE IRREDUCIBLE REPRESENTATIONS OF NⓈ SU(n)

It now follows from Mackey's theory for semidirect product groups (summarized in Ref. 6) that every quantum mechanical irreducible (ordinary) representation of NⓈ SU(n) is unitarily equivalent to one of the form

$$\begin{aligned}
U_N^v W \otimes \eta\colon (\sigma, \beta, \alpha, A + iB) &\to \exp iv(\sigma - x \cdot \beta)\exp\left(-\sum_{j=1}^{n}\alpha_j\frac{\partial}{\partial x_j}\right)\\
&\times W(A + iB) \otimes \eta(A + iB),
\end{aligned} \tag{20}$$

where $(\sigma, \beta, \alpha, A + iB) \in N$Ⓢ SU($n$), W is the representation of SU(n) defined by (19), η is an irreducible representation of SU(n), and the operators in the first part of the inner Kronecker product are defined on a dense subspace of $L^2(\mathrm{R}^n, \mathrm{C})$.

5. THE DECOMPOSITION OF THE REPRESENTATION W OF SU(n)

Some of the simplest of the quantum mechanical irreducible representations of NⓈ SU(n) [(20)] are those for which η is the one-dimensional identity representation of SU(n): $\eta(A + iB) = 1$ for all $A + iB \in \mathrm{SU}(n)$. The restriction of $U_N^v W \otimes \eta$ to SU(n) is then just the representation W of SU(n).

In order to decompose W into irreducible representations, the closed subspaces of $L^2(\mathrm{R}^n, \mathrm{C})$ which are in-

variant and irreducible under W will be found. Suppose w is the representation of su(n) corresponding to the representation W of SU(n). Then a closed subspace Ω is invariant under W if and only if the dense subspace of Ω on which w is defined is invariant under w.[13] Hence it is sufficient to find the subspaces of $L^2(\mathbb{R}^n, \mathbb{C})$ which are invariant and irreducible under the representation w.

As a basis for the representation space $L^2(\mathbb{R}^n, \mathbb{C})$ of W, take the set $\{\psi_m : m = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}_+^n\}$, where

$$\psi_m(x) = \prod_{j=1}^{n} u_{m_j}(|v|^{1/2} x_j),$$

with u_{m_j} a Hermite function of order m_j.

From the properties of Hermite functions,[14] it follows that the function $\phi_m(x) = u_m(|v|^{1/2} x)$ satisfies

$$\frac{1}{2}\left(-\frac{1}{v}\frac{d^2}{dx^2} + vx^2\right)\phi_m(x) = (\mathrm{sgn}v)(m + \tfrac{1}{2})\,\phi_m(x), \tag{21a}$$

$$x\phi_m(x) = \frac{1}{|v|^{1/2}}\left[\left(\frac{m}{2}\right)^{1/2}\phi_{m-1}(x) + \left(\frac{m+1}{2}\right)^{1/2}\phi_{m+1}(x)\right], \tag{21b}$$

$$\frac{d}{dx}\phi_m(x) = |v|^{1/2}\left[\left(\frac{m}{2}\right)^{1/2}\phi_{m-1}(x) - \left(\frac{m+1}{2}\right)^{1/2}\phi_{m+1}(x)\right]. \tag{21c}$$

Therefore, from (19) and (21),

$$\begin{aligned} w(R_{jk}^{(1)})\psi_m(x) &= i(\mathrm{sgn}v)[\sqrt{m_j(m_k+1)}\,\phi_{m_j-1}(x_j)\,\phi_{m_k+1}(x_k) \\ &\quad + \sqrt{(m_j+1)m_k}\,\phi_{m_j+1}(x_j)\,\phi_{m_k-1}(x_k)] \prod_{l \neq j,k} \phi_{m_l}(x_l), \end{aligned} \tag{22a}$$

$$\begin{aligned} w(R_{jk}^{(2)})\psi_m(x) &= (\mathrm{sgn}v)[-\sqrt{m_j(m_k+1)}\,\phi_{m_j-1}(x_j)\,\phi_{m_k+1}(x_k) \\ &\quad + \sqrt{(m_j+1)m_k}\,\phi_{m_j+1}(x_j)\,\phi_{m_k-1}(x_k)] \prod_{l \neq j,k} \phi_{m_l}(x_l), \end{aligned} \tag{22b}$$

$$w(H_k)\psi_m(x) = i(\mathrm{sgn}v)\left(m_k - \frac{1}{n}\sum_{l=1}^{n} m_l\right)\psi_m(x). \tag{22c}$$

Hence the representatives of each of the basis elements of su(n) transform any $L^2(\mathbb{R}^n, \mathbb{C})$ basis vector $\phi_{m_1}(x_1)\cdots\phi_{m_n}(x_n)$ with $\sum_{j=1}^n m_j = a$ into a linear combination of $L^2(\mathbb{R}^n, \mathbb{C})$ basis vectors, each of which has the sum of the subscripts of the component functions also equal to a.

Therefore, the subspace Ω_a of $L^2(\mathbb{R}^n, \mathbb{C})$ spanned by the set

$$\left\{\psi_m : \sum_{j=1}^{n} m_j = a\right\}$$

contains a dense subspace on which the representation w of su(n) is defined, and which is invariant under w. It follows that, for each $a \in \mathbb{Z}_+$, Ω_a is invariant under the representation W of SU(n). Hence

$$W = \bigoplus_{a=0}^{\infty} W_a, \tag{23}$$

where W_a is the restriction of W to the subspace Ω_a.

The representations W_a must now be decomposed into irreducible representations. This will be done with the aid of two preliminary lemmas.

Lemma A: Any one-dimensional subspace of Ω_a which is invariant under the subalgebra generated by $w(H_k)$ $(k = 2, 3, \ldots, n)$ consists of scalar multiples of some $\psi_m(x) = \phi_{m_1}(x_1)\cdots\phi_{m_n}(x_n)$, where $\sum_{j=1}^n m_j = a$.

Proof: From (22c), each ψ_m spans a one-dimensional subspace which is invariant under the subalgebra generated by $w(H_k)$ $(k = 2, \ldots, n)$.

Suppose $\zeta_a \in \Omega_a$ also spans a one-dimensional subspace of Ω_a which is invariant under the subalgebra generated by $w(H_k)$ $(k = 2, \ldots, n)$. Then, for each $k = 2, \ldots, n$,

$$w(H_k)\zeta_a = h_k\zeta_a, \quad \text{for some } h_k \in \mathbb{C}. \tag{24}$$

Now since $\zeta_a \in \Omega_a$, ζ_a can be expressed in the form

$$\zeta_a = \sum c_m \psi_m, \quad \text{for some } c_m \in \mathbb{C}, \tag{25}$$

where the summation is over all m such that $\sum_{j=1}^n m_j = a$. Hence, using (22c),

$$\begin{aligned} w(H_k)\zeta_a &= \sum c_m w(H_k)\psi_m \\ &= i(\mathrm{sgn}v)\sum c_m(m_k - a/n)\psi_m. \end{aligned}$$

Also, from (24) and (25),

$$w(H_k)\zeta_a = h_k\sum c_m\psi_m.$$

Since the ψ_m are linearly independent, it follows that

$$i(\mathrm{sgn}v)\,c_m(m_k - a/n) = h_k c_m,$$

for all $m = (m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ such that $\sum_{j=1}^n m_j = a$. So either $c_m = 0$, or $m_k = a/n - ih_k$. Hence all nonzero terms in the sum $\sum c_m\psi_m$ must have the same value of m_k, for each $k = 2, \ldots, n$; moreover, since $\sum_{j=1}^n m_j = a$ for each term in the sum $\sum c_m\psi_m$, it follows that each term must have the same value of m_1 as well. Hence $\zeta_a = c_m\psi_m$ for some ψ_m satisfying $\sum_{j=1}^n m_j = a$.

Lemma B: Each ψ_m is contained in an irreducible invariant subspace under the representation W.

Proof: Choose $m = (m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ arbitrarily. Then $\psi_m \in \Omega_a$, where $a = \sum_{j=1}^n m_j$. Suppose Ω_a splits into $s(a)$ irreducible invariant subspaces under W,

$$\Omega_a = \bigoplus_{r=1}^{s(a)} \Omega_{ar},$$

where Ω_{ar} is an irreducible invariant subspace under W. Then

$$\psi_m = \sum_{r=1}^{s(a)} \zeta_{ar}, \quad \text{where } \zeta_{ar} \in \Omega_{ar}. \tag{26}$$

So

$$w(H_k)\psi_m = \sum_{r=1}^{s(a)} w(H_k)\zeta_{ar} = \sum_{r=1}^{s(a)} \zeta'_{ar},$$

where $\zeta'_{ar} = w(H_k)\zeta_{ar} \in \Omega_{ar}$, since Ω_{ar} is invariant under W, and therefore under $w(H_k)$. Also, from (22c) and (26),

$$w(H_k)\psi_m = i(\mathrm{sgn}v)(m_k - a/n)\sum_{r=1}^{s(a)} \zeta_{ar}.$$

Since ζ_{ar} and $\zeta'_{ar} \in \Omega_{ar}$, and the Ω_{ar} $(r = 1, \ldots, s(a))$ are disjoint subspaces, it follows that

$$\zeta'_{ar} = w(H_k)\,\zeta_{ar} = i(\mathrm{sgn}v)(m_k - a/n)\,\zeta_{ar}\,.$$

Hence ζ_{ar} spans a one-dimensional subspace of Ω_a, which is invariant under the subalgebra generated by $w(H_k)$ $(k = 2, \ldots, n)$. So, by Lemma A, $\zeta_{ar} = c_{\tilde{m}}\psi_{\tilde{m}}$, for some $c_{\tilde{m}} \in \mathrm{C}$ and $\psi_{\tilde{m}}$ satisfying $\sum_{j=1}^{n}(\tilde{m})_j = a$. Since the ψ's are linearly independent functions, it follows, from (26), that the chosen ψ_m is itself of the form ζ_{ar} (for some r), and so $\psi_m \in \Omega_{ar}$.

Proposition: For each $a \in \mathbb{Z}_+$, Ω_a is an irreducible invariant subspace under the representation W.

Proof: Choose $a \in \mathbb{Z}_+$ arbitrarily. From Lemma B, $\psi_{(a,0,\ldots,0)}(x)$ belongs to an irreducible invariant subspace, Θ say, of Ω_a. Now, from (22b),

$$w(R_{12}^{(2)})\,\psi_{(a,0,\ldots,0)}(x) = -(\mathrm{sgn}v)\sqrt{a}\,\psi_{a-1}(x_1)\,\psi_1(x_2)\,\psi_0(x_3)\cdots\psi_0(x_n)\,,$$

unless $a = 0$. If $a = 0$, then $\psi_{(a,0,\ldots,0)} = \psi_0$ spans the one-dimensional subspace, Ω_0, invariant under W. If $a \neq 0$, then $\psi_{(a-1,1,\ldots,0)} \in \Theta$, since Θ is invariant under W and therefore under $w(R_{12}^{(2)})$. Using (22b) again,

$$w(R_{12}^{(2)})\,\psi_{(a-1,1,0,\ldots,0)}(x) = (\mathrm{sgn}v)[-\sqrt{2(a-1)}\,\psi_{a-2}(x_1)\,\psi_2(x_2) + \sqrt{a}\,\psi_a(x_1)\,\psi_0(x_2)]\,\psi_0(x_3)\cdots\psi_0(x_n)\,.$$

Since $\psi_{(a,0,\ldots,0)} \in \Theta$, it follows that, unless $a = 1$, $\psi_{(a-2,2,0,\ldots,0)} \in \Theta$ as well. Then, by induction, the invariant subspace Θ containing $\psi_{(a,0,\ldots,0)}$ also contains $\psi_{(a-j,j,0,\ldots,0)}$, for $j = 1, 2, \ldots, a$.

Now choose arbitrarily $m = (m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ satisfying $\sum_{j=1}^{n} m_j = a$. Since $\psi_{(a,0,\ldots,0)} \in \Theta$, it follows, from above, that $\psi_{(a-m_2,\,m_2,0,\ldots,0)} \in \Theta$. A similar argument to that given before, using $R_{13}^{(2)}$ instead of $R_{12}^{(2)}$, gives that $\psi_{(a-m_2-m_3,\,m_2,\,m_3,0,\ldots,0)} \in \Theta$. Similarly, using $R_{1k}^{(2)}$ $(k = 4, \ldots, n)$ it follows that $\psi_{(a-\sum_{j=2}^{n} m_j,\,m_2,\,m_3,\ldots,m_n)} \in \Theta$. So, since $\sum_{j=1}^{n} m_j = a$, $\psi_{(m_1,\,m_2,\ldots,\,m_n)} \in \Theta$. Hence Θ contains all the basis functions ψ_m for which $\sum_{j=1}^{n} m_j = a$, and so $\Theta = \Omega_a$. Therefore, Ω_a is an irreducible invariant subspace under W. (This completes the proof of the proposition.)

It follows, from (23), that the decomposition of the representation W of SU(n) is

$$W = \bigoplus_{a=0}^{\infty} W_a\,,$$

where each W_a is an irreducible representation (namely the restriction of W to the subspace Ω_a).

6. NⓈ SU(n) AS A NONINVARIANCE GROUP

The dimension of Ω_a is the number of distinct ways of choosing $m = (m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ such that $\sum_{j=1}^{n} m_j = a$. This is precisely the degeneracy of the $(a+1)$th energy level of the n-dimensional isotropic harmonic oscillator. The results of Secs. 4 and 5 then lead to the following.

Theorem: For each $v(\neq 0) \in \mathrm{R}$, the ordinary representation $U_N^v W$ of NⓈ SU(n), given by

$$U_N^v W : (\sigma, \beta, \alpha, A + iB) \to \exp iv(\sigma - x\,.\,\beta)\exp\left(-\sum_{j=1}^{n}\alpha_j\frac{\partial}{\partial x_j}\right)W(A + iB) \tag{27}$$

[where W is the representation of SU(n) defined by (19) and the operators are defined on a dense subspace of $L^2(\mathrm{R}^n, \mathrm{C})$], is irreducible as a representation of NⓈ SU(n).

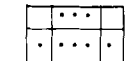
The restriction of $U_N^v W$ to SU(n), which is a degeneracy group for the n-dimensional isotropic harmonic oscillator, is

$$(U_N^v W)\downarrow \mathrm{SU}(n) = W = \bigoplus_{a=0}^{\infty} W_a\,, \tag{28}$$

where W_a is an irreducible representation of SU(n), of dimension equal to the degeneracy of the $(a+1)$th energy level of the n-dimensional isotropic harmonic oscillator.

Therefore, NⓈ SU(n) is a noninvariance group for the n-dimensional isotropic harmonic oscillator.

Note: Each irreducible representation of SU(n) may be described by means of a Young tableau. (There is a review article on this subject by Itzykson and Nauenberg.[11]) By examining the highest weight of the representation W_a, it is readily seen that

when $v > 0$, $W_a =$  $= (a, 0, \ldots, 0)$, (29a)

when $v < 0$, $W_a =$ $= (a, \ldots, a, 0)$. (29b)

[1] M. Hamermesh, *Group Theory* (Addison-Wesley, Massachusetts, 1962).
[2] G.A. Baker, Phys. Rev. 103, 1119 (1956).
[3] A.O. Barut, Phys. Rev. 135, 839B (1964).
[4] N. Mukunda, L. O'Raifeartaigh, and E.C.G. Sudarshan, Phys. Rev. Lett. 15, 1041 (1965).
[5] M.E. Major, J. Math. Phys. 18, 1944 (1977).
[6] M.E. Major, J. Math. Phys. 18, 1938 (1977).
[7] K.R. Parthasarathy, *Multipliers on Locally Compact Groups*, Lecture Notes in Mathematics, Vol. 93 (Springer, Berlin, 1969).
[8] L.S. Pontryagin, *Topological Groups* (Gordon and Breach, New York, 1966).
[9] K. Yosida, *Functional Analysis* (Springer, Berlin, 1965).
[10] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[11] C. Itzykson and M. Nauenberg, Rev. Mod. Phys. 38, 95 (1966).
[12] E. Nelson, Ann. Math. 70, 572 (1959).
[13] M.A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, Oxford, 1964).
[14] K. Gottfried, *Quantum Mechanics* (Benjamin, New York, 1966), Vol. 1.

On the accidental degeneracy of the *n*-dimensional anisotropic harmonic oscillator. II

M. E. Major

Mathematics Department, Queen Elizabeth College, London W8 7AH, England
(Received 22 September 1975)

In an earlier paper, a group, G, associated with the n-dimensional anisotropic harmonic oscillator was shown to be embedded in a semidirect product, L, of the Weyl group N and the symplectic group $\mathrm{Sp}(2n,\mathrm{R})$. A particular representation R^ν of L, when restricted to G, was proved to be unitarily equivalent to $\oplus_s d_{\omega,s}\, U_G^{\nu,-(\mathrm{sgn}\nu)s}$, where $d_{\omega,s}$ is the degeneracy of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with frequencies $(\omega_1,\omega_2,\ldots,\omega_n)=\omega$, $U_G^{\nu,-(\mathrm{sgn}\nu)s}$ is an irreducible representation of G and s may be regarded as indexing all distinct energy levels of the system. In the present paper, the representation R^ν of L is shown to be unitarily equivalent to the representation $U_N^\nu W\otimes\overline{W}$ of L, where U_N^ν is an irreducible representation of N, W is the projective representation of $\mathrm{Sp}(2n,\mathrm{R})$ which intertwines the representations U_N^ν and SU_N^ν of N [where $S\in\mathrm{Sp}(2n,\mathrm{R})$], and $\overline{W}$ is the complex conjugate of W. This alternative form for the representation R^ν of L enables it to be decomposed, into two irreducible representations.

1. INTRODUCTION

In Ref. 1, a group G associated with the n-dimensional anisotropic harmonic oscillator was constructed: G is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system.

G was shown to be embedded in a group, L, which is a semidirect product of the Weyl group, N, and the symplectic group, $\mathrm{Sp}(2n,\mathrm{R})$.[2] A particular representation R^ν of L, when restricted to G, was proved to be unitarily equivalent to $\oplus_s d_{\omega,s} U_G^{\nu,-(\mathrm{sgn}\nu)s}$, where $d_{\omega,s}$ is precisely the degeneracy of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with frequencies $(\omega_1,\omega_2,\ldots,\omega_n)$, $U_G^{\nu,-(\mathrm{sgn}\nu)s}$ is an irreducible representation of G and the summation may be regarded as over all distinct energy levels $E_{\omega,s}$ of the system.[2]

In the present paper, the representation R^ν of L is studied in greater detail. R^ν is shown to be unitarily equivalent to the representation $U_N^\nu W\otimes\overline{W}$ of L, where U_N^ν is an irreducible representation of N, W is the projective representation of $\mathrm{Sp}(2n,\mathrm{R})$ which intertwines the representations U_N^ν and SU_N^ν of N [where $S\in\mathrm{Sp}(2n,\mathrm{R})$], and $\overline{W}$ is the complex conjugate of W. This alternative form for the representation R^ν enables it to be decomposed, into two irreducible representations.

In Secs. 2–4, the quantum mechanical irreducible representations of L are obtained, using Mackey's theory of induced representations (summarized in Ref. 1). In Sec. 5, various informal arguments which suggest the unitary equivalence of R^ν and $U_N^\nu W\otimes\overline{W}$ are given. The result is proved in Sec. 6. The decomposition of the projective representation W of $\mathrm{Sp}(2n,\mathrm{R})$ is found in Sec. 7, and, from this, the decomposition of the representation R^ν of L is obtained.

2. THE LITTLE GROUP OF U_N^ν

From Ref. 2, Eq. (7), the group law of L is given by

$$(\sigma',\beta',\alpha',S')(\sigma,\beta,\alpha,S)=(\sigma'',\beta'',\alpha'',S''), \tag{1}$$

where

$$\sigma''=\sigma'+\sigma+\tfrac{1}{2}(A'\alpha.C'\alpha+B'\beta.D'\beta)+B'\beta.C'\alpha+\alpha'.(D'\beta+C'\alpha),$$

$$\beta''=D'\beta+C'\alpha+\beta',$$

$$\alpha''=B'\beta+A'\alpha+\alpha',$$

$$S''=S'S,$$

with $S'=\left(\begin{smallmatrix}A' & B'\\ C' & D'\end{smallmatrix}\right)\in\mathrm{Sp}(2n,\mathrm{R})$ and $S\in\mathrm{Sp}(2n,\mathrm{R})$.

Every quantum mechanical irreducible representation of N is unitarily equivalent to one of the form [Ref. 1, Eq. (15)]

$$[U_N^\nu(\sigma,\beta,\alpha,I)\psi](x)=\exp i\nu(\sigma-x.\beta)\psi(x-\alpha), \tag{2}$$

where $\nu(\neq 0)\in\mathrm{R}$, $\psi\in L^2(\mathrm{R}^n,\mathrm{C})$.

The action of $S=\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)\in\mathrm{Sp}(2n,\mathrm{R})$ on U_N^ν is, by definition [Ref. 1, Eq. (11)],

$$\begin{aligned}[(SU_N^\nu)(\sigma,\beta,\alpha,I)\psi](x)&=[U_N^\nu((0,0,0,S)^{-1}(\sigma,\beta,\alpha,I)(0,0,0,S))\psi](x)\\&=\exp i\nu(\sigma+\tfrac{1}{2}(-D^T\alpha.C^T\alpha-B^T\beta.A^T\beta)+B^T\beta.C^T\alpha\\&\quad-x.(A^T\beta-C^T\alpha))\psi(x+B^T\beta-D^T\alpha),\end{aligned} \tag{3}$$

with T denoting transpose, using $\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)^{-1}=\left(\begin{smallmatrix}D^T & -B^T\\ -C^T & A^T\end{smallmatrix}\right)$, together with (1) and (2).

SU_N^ν is an irreducible representation of N. Restricted to the subgroup $\{(\sigma,0,0,I)\}$, it is just the phase $\exp i\nu\sigma$. Hence SU_N^ν must be unitarily equivalent to U_N^ν, since U_N^ν is the only irreducible representation of N which has the required form on restriction to the subgroup $\{(\sigma,0,0,I)\}$.

Therefore, there exists a unitary operator $W(S)$, dependent on S, such that

$$SU_N^\nu=W(S)^{-1}U_N^\nu W(S)\quad\text{for each } S\in\mathrm{Sp}(2n,\mathrm{R}). \tag{4}$$

The orbits of the quantum mechanical part of $\hat{N}$ under the action of $\mathrm{Sp}(2n,\mathrm{R})$ thus consist of single points U_N^ν. The little group of U_N^ν under the action of $\mathrm{Sp}(2n,\mathrm{R})$ is the whole of $\mathrm{Sp}(2n,\mathrm{R})$, and so the isotropy group of U_N^ν is $N\circledS\mathrm{Sp}(2n,\mathrm{R})$ $(=L)$.

3. THE OPERATOR W FOR THE GROUP Sp($2n$, R)

A. General approach

In Ref. 3, a method was developed for finding the operator W for the group SU(n) [which can be regarded as a subgroup of Sp($2n$, R)]. This method depends partly on the fact that SU(n) has no nontrivial multipliers, and so W can be taken to be an ordinary representation of SU(n). Now, the group Sp($2n$, R) possesses nontrivial multipliers, and so it is possible that W is a projective representation; hence the method of Ref. 3 cannot be applied directly in the present case. Nevertheless, since Sp($2n$,R) is a connected semisimple Lie group, every multiplier is locally trivial;[4] it follows that the (possibly) projective representation W of Sp($2n$,R) can be chosen in such a way that the corresponding representation w of the Lie algebra of Sp($2n$,R) is ordinary.

Suppose $h_i(\tau)$ is a one-dimensional Lie subgroup of Sp($2n$,R). For each one-parameter subgroup $\mu(t)$ of N, let $Z_{\mu i}(\tau)$ be the unique skew-adjoint operator defined by

$$(h_i(\tau)U_N^v)(\mu(t)) = \exp tZ_{\mu i}(\tau) \text{ [cf. Ref. 3, Eq. (9)]}. \quad (5)$$

Then [cf. Ref. 3, Eq. (11)] the skew-adjoint operator $dW(h_i(\tau))/d\tau\,|_{\tau=0}$ is determined up to an arbitrary imaginary constant ξ_i by

$$\left.\frac{dZ_{\mu i}(\tau)}{d\tau}\right|_{\tau=0} = \left[Z_{\mu i}(0),\ \left.\frac{dW(h_i(\tau))}{d\tau}\right|_{\tau=0}\right]$$

$$\text{for each } \mu \in \{\sigma, \beta_1, \ldots, \beta_n, \alpha_1, \ldots, \alpha_n\}. \quad (6)$$

[The proof of this result is similar to the corresponding one of Ref. 3, Sec. 3 A, with $\exp\tau J$ replaced by $W(h_i(\tau))$.]

Since the operators $dW(h_i(\tau))/d\tau\,|_{\tau=0}$ can be chosen to give an ordinary representation w of the Lie algebra of Sp($2n$,R), there are constraints on the constants ξ_i; as will be shown, these are sufficient to determine the constants ξ_i uniquely.

Since w is an ordinary representation of the Lie algebra of the connected group Sp($2n$,R), w exponentiates to an ordinary representation, $\tilde{W}$ say, of the connected, simply connected, covering group $\widetilde{\mathrm{Sp}}(2n,\mathrm{R})$ of Sp($2n$,R).[5] Sp($2n$,R) is the image of $\widetilde{\mathrm{Sp}}(2n,\mathrm{R})$ under a homomorphism δ whose kernel is a discrete central subgroup of $\widetilde{\mathrm{Sp}}(2n,\mathrm{R})$: $\mathrm{Sp}(2n,\mathrm{R}) \approx \widetilde{\mathrm{Sp}}(2n,\mathrm{R})/\ker\delta$. It is possible to show that $\tilde{W}$ maps $\ker\delta$ into the unit circle, and hence $\tilde{W}$ determines the required projective representation W of Sp($2n$,R).[6] However, for what follows, it is not necessary to know W explicitly; it is sufficient that the ordinary representation w of the Lie algebra of Sp($2n$,R) is known in detail.

B. Detailed calculation

The symplectic group Sp($2n$,R) is the set of matrices $S \in \mathrm{GL}(2n,\mathrm{R})$ for which $S^TJS = J$, where $J = \sum_{j=1}^{n}(E_{j,j+n} - E_{j+n,j})$ [with E_{jk} a $2n \times 2n$ matrix having 1 in the (j,k) position and zeros elsewhere]. Let $S(\tau)$ be an analytic curve in Sp($2n$,R), and suppose $S(0) = I$. Then, differentiating $S(\tau)^TJS(\tau) = J$ with respect to τ, and putting $\tau = 0$, gives

$$S'(0)^TJ + JS'(0) = 0,$$

where $'$ denotes differentiation with respect to τ. Hence the Lie algebra of Sp($2n$,R) is the algebra of all matrices of the form $\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$, where A, B, and C are $n \times n$ real matrices, with A arbitrary, $B = B^T$, and $C = C^T$; this algebra will be denoted by sp($2n$,R).

A basis for sp($2n$,R) is

$$X_{jk} = \begin{pmatrix} E_{jk} & 0 \\ 0 & -E_{kj} \end{pmatrix} \quad (1 \le j,k \le n),$$

$$Y_{jk} = \begin{pmatrix} 0 & E_{jk}+E_{kj} \\ 0 & 0 \end{pmatrix} \quad (1 \le j \le k \le n), \quad (7)$$

$$Z_{jk} = \begin{pmatrix} 0 & 0 \\ E_{jk}+E_{kj} & 0 \end{pmatrix} \quad (1 \le j \le k \le n)$$

(where E_{lm} is an $n \times n$ matrix having 1 in the (l,m) position and zeros elsewhere).

If

$$h(\tau) = \begin{pmatrix} a_{jk}(\tau) & b_{jk}(\tau) \\ c_{jk}(\tau) & d_{jk}(\tau) \end{pmatrix}$$

is a one-dimensional Lie subgroup of Sp($2n$,R), then, from (3) and (5), the representatives of the generators of N in the ordinary representation $h(\tau)U_N^v$ of N are

$$\begin{aligned} Z_\sigma(\tau) &= iv, \\ Z_{\beta_j}(\tau) &= -iv\sum_{k=1}^{n} a_{jk}(\tau)x_k + \sum_{k=1}^{n} b_{jk}(\tau)\frac{\partial}{\partial x_k}, \\ Z_{\alpha_j}(\tau) &= iv\sum_{k=1}^{n} c_{jk}(\tau)x_k - \sum_{k=1}^{n} d_{jk}(\tau)\frac{\partial}{\partial x_k}. \end{aligned} \quad (8)$$

From (6),

$$dW(h(\tau))/d\tau\,|_{\tau=0} = w(h'(0))$$

must satisfy,
when $\mu = \sigma$,

$$0 = [iv, w(h'(0))],$$

which is satisfied by any $w(h'(0))$;
when $\mu = \beta_j$,

$$-iv\sum_{k=1}^{n} a'_{jk}(0)x_k + \sum_{k=1}^{n} b'_{jk}(0)\frac{\partial}{\partial x_k} = [-ivx_j, w(h'(0))]; \quad (9)$$

when $\mu = \alpha_j$,

$$iv\sum_{k=1}^{n} c'_{jk}(0)x_k - \sum_{k=1}^{n} d'_{jk}(0)\frac{\partial}{\partial x_k} = \left[-\frac{\partial}{\partial x_j}, w(h'(0))\right] \quad (10)$$

(where $'$ denotes differentiation with respect to τ).

The operator $w(X_{lm})$

Suppose, with the above notation, that X_{lm} is the tangent at the identity of a one-dimensional Lie subgroup

$$h(\tau) = \begin{pmatrix} a_{jk}(\tau) & b_{jk}(\tau) \\ c_{jk}(\tau) & d_{jk}(\tau) \end{pmatrix}$$

in Sp($2n$,R). Then

$$X_{lm} = h'(0) = \begin{pmatrix} a'_{jk}(0) & b'_{jk}(0) \\ c'_{jk}(0) & d'_{jk}(0) \end{pmatrix}.$$

In this case, the only nonzero elements of $h'(0)$ are

$$a'_{lm}(0) = 1, \quad d'_{ml}(0) = -1.$$

Hence the conditions (9) and (10) become,

for $j \neq l$ or m,

$$0 = [-ivx_j, w(X_{lm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_j}, w(X_{lm})\right];$$

for $j = l$,

$$-ivx_m = [-ivx_l, w(X_{lm})] \text{ and } \delta_{lm}\frac{\partial}{\partial x_l} = \left[-\frac{\partial}{\partial x_l}, w(X_{lm})\right];$$

for $j = m$,

$$-iv\delta_{lm}x_m = [-ivx_m, w(X_{lm})] \text{ and } \frac{\partial}{\partial x_l} = \left[-\frac{\partial}{\partial x_m}, w(X_{lm})\right].$$

If $l \neq m$, these relations are all satisfied by the skew-adjoint operator

$$\tilde{w}(X_{lm}) = -x_m\frac{\partial}{\partial x_l} + \alpha_{lm},$$

where α_{lm} is an arbitrary imaginary constant.

If $l = m$, these relations are all satisfied by

$$\tilde{w}(X_{mm}) = -x_m\frac{\partial}{\partial x_m} + \epsilon_m,$$

where ϵ_m is an arbitrary constant. Now the adjoint of $\tilde{w}(X_{mm})$ is $\tilde{w}(X_{mm})^* = (-\partial/\partial x_m)^* x_m^* + \bar{\epsilon}_m = (\partial/\partial x_m)x_m + \bar{\epsilon}_m$ $= x_m(\partial/\partial x_m) + 1 + \bar{\epsilon}_m$. Thus $\tilde{w}(X_{mm})$ is a skew-adjoint operator provided $\epsilon_m = -\frac{1}{2} + \alpha_{mm}$, where α_{mm} is an arbitrary imaginary constant. Therefore, for $1 \leq l, m \leq n$,

$$\tilde{w}(X_{lm}) = -x_m\frac{\partial}{\partial x_l} - \tfrac{1}{2}\delta_{lm} + \alpha_{lm}. \tag{11a}$$

The operator $w(Y_{lm})$ $(l < m)$

Similarly, the operator $w(Y_{lm})$ must satisfy the relations,

for $j \neq l$ or m,

$$0 = [-ivx_j, w(Y_{lm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_j}, w(Y_{lm})\right];$$

for $j = l$ $(\neq m)$,

$$\frac{\partial}{\partial x_m} = [-ivx_l, w(Y_{lm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_l}, w(Y_{lm})\right];$$

for $j = m$ $(\neq l)$,

$$\frac{\partial}{\partial x_l} = [-ivx_m, w(Y_{lm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_m}, w(Y_{lm})\right].$$

These relations are all satisfied by the skew-adjoint operator

$$\tilde{w}(Y_{lm}) = \frac{1}{iv}\frac{\partial^2}{\partial x_l \partial x_m} + \beta_{lm}, \tag{11b}$$

where β_{lm} is an arbitrary imaginary constant.

The operator $w(Y_{mm})$

Similarly, the operator $w(Y_{mm})$ must satisfy the relations,

for $j \neq m$,

$$0 = [-ivx_j, w(Y_{mm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_j}, w(Y_{mm})\right];$$

for $j = m$,

$$2\frac{\partial}{\partial x_m} = [-ivx_m, w(Y_{mm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_m}, w(Y_{mm})\right].$$

These relations are all satisfied by the skew-adjoint operator

$$\tilde{w}(Y_{mm}) = \frac{1}{iv}\frac{\partial^2}{\partial x_m^2} + \beta_{mm}, \tag{11c}$$

where β_{mm} is an arbitrary imaginary constant.

The operator $w(Z_{lm})$ $(l < m)$

Similarly, the operator $w(Z_{lm})$ must satisfy the relations,

for $j \neq l$ or m,

$$0 = [-ivx_j, w(Z_{lm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_j}, w(Z_{lm})\right];$$

for $j = l$ $(\neq m)$,

$$0 = [-ivx_l, w(Z_{lm})] \text{ and } ivx_m = \left[-\frac{\partial}{\partial x_l}, w(Z_{lm})\right];$$

for $j = m$ $(\neq l)$,

$$0 = [-ivx_m, w(Z_{lm})] \text{ and } ivx_l = \left[-\frac{\partial}{\partial x_m}, w(Z_{lm})\right].$$

These relations are all satisfied by the skew-adjoint operator

$$\tilde{w}(Z_{lm}) = -ivx_l x_m + \gamma_{lm}, \tag{11d}$$

where γ_{lm} is an arbitrary imaginary constant.

The operator $w(Z_{mm})$

Similarly, the operator $w(Z_{mm})$ must satisfy the relations,

for $j \neq m$,

$$0 = [-ivx_j, w(Z_{mm})] \text{ and } 0 = \left[-\frac{\partial}{\partial x_j}, w(Z_{mm})\right];$$

for $j = m$,

$$0 = [-ivx_m, w(Z_{mm})] \text{ and } 2ivx_m = \left[-\frac{\partial}{\partial x_m}, w(Z_{mm})\right].$$

These relations are all satisfied by the skew-adjoint operator

$$\tilde{w}(Z_{mm}) = -ivx_m^2 + \gamma_{mm}, \tag{11e}$$

where γ_{mm} is an arbitrary imaginary constant.

C. Calculation of the values of the constants α_{lm}, β_{lm}, γ_{lm}

It may be shown that the commutation relations of $\mathrm{sp}(2n, \mathrm{R})$, in terms of the basis X_{jk}, Y_{jk}, Z_{jk} [see (7)] are

$$[X_{jk}, X_{lm}] = \delta_{kl}X_{jm} - \delta_{jm}X_{lk}, \tag{12a}$$

$$[X_{jk}, Y_{lm}] = \delta_{kl}Y_{jm} + \delta_{km}Y_{jl}, \tag{12b}$$

$$[X_{jk}, Z_{lm}] = -\delta_{jl}Z_{km} - \delta_{jm}Z_{kl}, \tag{12c}$$

$$[Y_{jk}, Y_{lm}] = 0, \tag{12d}$$

$$[Y_{jk}, Z_{lm}] = \delta_{km}X_{jl} + \delta_{kl}X_{jm} + \delta_{jm}X_{kl} + \delta_{jl}X_{km}, \tag{12e}$$

$$[Z_{jk}, Z_{lm}] = 0, \tag{12f}$$

where, for convenience of notation, define

$$Y_{jk}=Y_{kj},\quad Z_{jk}=Z_{kj},\quad \text{for } j>k.$$

Since $\tilde{w}$ is to be an ordinary representation of $\mathrm{sp}(2n,\mathrm{R})$, $\tilde{w}(X_{jk})$, $\tilde{w}(Y_{jk})$, and $\tilde{w}(Z_{jk})$ must satisfy the commutation relations of $\mathrm{sp}(2n,\mathrm{R})$. Explicit calculation shows that this implies

$$-\delta_{kl}\alpha_{jm}+\delta_{jm}\alpha_{lk}=0, \tag{13a}$$

$$-\delta_{kl}\beta_{jm}-\delta_{km}\beta_{jl}=0, \tag{13b}$$

$$\delta_{jl}\gamma_{km}+\delta_{jm}\gamma_{kl}=0, \tag{13c}$$

$$-\delta_{km}\alpha_{jl}-\delta_{kl}\alpha_{jm}-\delta_{jm}\alpha_{kl}-\delta_{jl}\alpha_{km}=0, \tag{13d}$$

where, for convenience of notation, define

$$\beta_{jk}=\beta_{kj},\quad \gamma_{jk}=\gamma_{kj},\quad \text{for } j>k.$$

Condition (13a) gives that $\alpha_{jm}=0$ for $j\neq m$ (putting $k=l$). Condition (13d) gives that $\alpha_{jj}=0$ for $1\leq j\leq n$ (putting $j=k=l=m$). From condition (13b), it follows that $\beta_{jm}=0$ for $1\leq j,m\leq n$ (putting $k=l\neq m$, if $n>1$, and $j=k=l=m=1$, if $n=1$). Similarly, from condition (13c), $\gamma_{km}=0$ for $1\leq k,m\leq n$. Hence, $\tilde{w}$ is an ordinary representation w of $\mathrm{sp}(2n,\mathrm{R})$ if and only if the α's, β's, and γ's are all zero. Therefore, from (11), the ordinary representation w of $\mathrm{sp}(2n,\mathrm{R})$ is determined by

$$\begin{aligned}
w(X_{jk})&=-x_k\frac{\partial}{\partial x_j}-\tfrac{1}{2}\delta_{jk}\quad(1\leq j,k\leq n),\\
w(Y_{jk})&=-\frac{i}{v}\frac{\partial^2}{\partial x_j\partial x_k}\quad(1\leq j\leq k\leq n),\\
w(Z_{jk})&=-ivx_jx_k\quad(1\leq j\leq k\leq n).
\end{aligned} \tag{14}$$

4. THE IRREDUCIBLE ORDINARY REPRESENTATIONS OF L

It now follows from Mackey's theory for semidirect product groups that every quantum mechanical irreducible ordinary representation of $L=N\circledS\mathrm{Sp}(2n,\mathrm{R})$ is unitarily equivalent to one of the form

$$U_N^vW\otimes\eta:(\sigma,\beta,\alpha,S)$$

$$\to\exp iv(\sigma-x.\beta)\exp\left(-\sum_{j=1}^{n}\alpha_j\frac{\partial}{\partial x_j}\right)W(S)\otimes\eta(S), \tag{15}$$

where $(\sigma,\beta,\alpha,S)\in L$, W is the projective representation of $\mathrm{Sp}(2n,\mathrm{R})$ determined up to trivial multipliers by (14), η is an irreducible projective representation of $\mathrm{Sp}(2n,\mathrm{R})$ with multiplier inverse to that of W, and the operators in the first part of the inner Kronecker product are defined on a dense subspace of $L^2(\mathrm{R}^n,\mathrm{C})$.

5. HEURISTIC ARGUMENTS SUGGESTING THAT $(R^v\downarrow L)\simeq U_N^vW\otimes\overline{W}$

To simplify the notation denote $R^v\downarrow L$ by R_L^v. The result stated above was found through an attempt to decompose R_L^v into irreducible representations of L, of the form (15). It was hoped that the decomposition of R_L^v would be suggested by that of $R_L^v\downarrow(N\circledS\mathrm{SU}(n))$. The decomposition of $R_L^v\downarrow(N\circledS\mathrm{SU}(n))$ may be obtained as follows. From Ref. 2, Sec. 5,

$$R_L^v\downarrow(N\circledS\mathrm{SU}(n))=(D(H)\uparrow L)\downarrow(N\circledS\mathrm{SU}(n)),$$

where D is the representation of the subgroup $H=\{(\sigma,0,0,I)\}\times\mathrm{Sp}(2n,\mathrm{R})$ of L, given by

$$D:(\sigma,0,0,S)\to\exp iv\sigma.$$

Hence $R_L^v\downarrow(N\circledS\mathrm{SU}(n))$ is the restriction of an induced representation of L. Since $N\circledS\mathrm{SU}(n)$ and H are closed subgroups of the separable locally compact group L, and L can be expressed as a single double coset of $N\circledS\mathrm{SU}(n)$ and H: $L=N\circledS\mathrm{SU}(n)(0,0,0,I)H$ [from (1)], Mackey's subgroup theorem[7] may be applied,

$$(D(H)\uparrow L)\downarrow(N\circledS\mathrm{SU}(n))=D((N\circledS\mathrm{SU}(n))\cap H)\uparrow(N\circledS\mathrm{SU}(n)).$$

Therefore,

$$R_L^v\downarrow(N\circledS\mathrm{SU}(n))=D(C\times\mathrm{SU}(n))\uparrow(N\circledS\mathrm{SU}(n)), \tag{16}$$

where $C=\{(\sigma,0,0,I)\}$.

The group $N\circledS\mathrm{SU}(n)$ is locally compact, but not compact, so the Frobenius reciprocity theorem[7] does not necessarily hold for it. However, a formal application of the theorem may, nevertheless, give the correct decomposition of $R_L^v\downarrow(N\circledS\mathrm{SU}(n))$.

Suppose $U_N^tW\otimes\eta$ is an arbitrary quantum mechanical irreducible representation of $N\circledS\mathrm{SU}(n)$, where W is now regarded as an ordinary representation of $\mathrm{SU}(n)$ (Ref. 3, Sec. 4). Then, formally,

$$\begin{aligned}
&\langle R_L^v\downarrow(N\circledS\mathrm{SU}(n)),U_N^tW\otimes\eta\rangle_{N\circledS\mathrm{SU}(n)}\\
&=\langle D(C\times\mathrm{SU}(n))\uparrow(N\circledS\mathrm{SU}(n)),U_N^tW\otimes\eta\rangle_{N\circledS\mathrm{SU}(n)}\\
&=\langle(U_N^tW\otimes\eta)\downarrow(C\times\mathrm{SU}(n)),D(C\times\mathrm{SU}(n))\rangle_{C\times\mathrm{SU}(n)}.
\end{aligned} \tag{17}$$

The representation $D:(\sigma,0,0,U)\to\exp iv\sigma$ of $C\times\mathrm{SU}(n)$ is contained in the representation $U_N^tW\otimes\eta:(\sigma,0,0,U)\to\exp it\sigma W(U)\otimes\eta(U)$ of $C\times\mathrm{SU}(n)$ only when $t=v$; in this case, the number of times that D is contained in $U_N^tW\otimes\eta$ [as representations of $C\times\mathrm{SU}(n)$] equals the frequency of the one-dimensional identity representation $I_{\mathrm{SU}(n)}$ in the representation $U\to W(U)\otimes\eta(U)$ of $\mathrm{SU}(n)$.

Now, from Ref. 3, Eq. (23),

$$W\otimes\eta=\Big(\bigoplus_{a=0}^{\infty}W_a\Big)\otimes\eta=\bigoplus_{a=0}^{\infty}(W_a\otimes\eta). \tag{18}$$

It thus remains to determine, for each a, how often $I_{\mathrm{SU}(n)}$ occurs in $W_a\otimes\eta$. A general method for decomposing inner Kronecker products of irreducible representations of $\mathrm{SU}(n)$ is given in Ref. 8.

When $v>0$, from Ref. 3, Eq. (29a),

$$W_a=\underbrace{\boxed{\cdots}}_{a\text{ boxes}}=(a,0,\ldots,0).$$

It follows that, when $v>0$, $W_a\otimes\eta$ contains $I_{\mathrm{SU}(n)}$ if and only if

$$\eta=\underbrace{\begin{array}{|c|c|c|c|}\hline\cdot&\cdot&\cdot&\\\hline\cdot&\cdot&\cdot&\cdot\\\hline\cdot&\cdot&\cdot&\cdot\\\hline&\cdot&\cdot&\\\hline\end{array}}_{a\text{ columns}}\Bigg\}\,n-1\text{ rows}=(a,\ldots,a,0).$$

When $v<0$, from Ref. 3, Eq. (29b), $W_a=(a,\ldots,a,0)$. It follows that, when $v<0$, $W_a\otimes\eta$ contains $I_{\mathrm{SU}(n)}$ if and only if $\eta=(a,0,\ldots,0)$.

Now the representations $(a, 0, \ldots, 0)$ and $(a, \ldots, a, 0)$ are mutually contragredient,[8] so, since W_a is unitary, the representation contragredient to W_a is just the complex conjugate $\overline{W}_a$ of W_a.

So, for any $v(\neq 0)$, $I_{SU(n)}$ occurs in $W_a \otimes \eta$ if and only if $\eta = \overline{W}_a$, and then it occurs exactly once. Hence, using (18), D is contained in $U_N^t W \otimes \eta$ [as representations of $C \times SU(n)$] if and only if $t = v$ and $\eta = \overline{W}_a$ for some a; in this case, D occurs exactly once.

A formal application [(17)] of the Frobenius reciprocity theorem therefore suggests that the decomposition of $R_L^v \downarrow (N \circledS SU(n))$ into irreducible representations is

$$R_L^v \downarrow (N \circledS SU(n)) \simeq \bigoplus_{a=0}^{\infty} (U_N^v W \otimes \overline{W}_a).$$

Now

$$\bigoplus_{a=0}^{\infty} (U_N^v W \otimes \overline{W}_a) = U_N^v W \otimes \left(\bigoplus_{a=0}^{\infty} \overline{W}_a\right) = U_N^v W \otimes \overline{W}.$$

This indicates that, perhaps, the representation R_L^v of L is unitarily equivalent to $U_N^v W \otimes \overline{W}$, where W is now regarded as a projective representation of $Sp(2n, R)$. $U_N^v W \otimes \overline{W}$ is an ordinary representation of L, since $\overline{W}$ has multiplier inverse to that of W.

From Ref. 1, Eq. (26), and Ref. 2, Eqs. (27) and (44), it follows that the restrictions of R_L^v and $U_N^v W \otimes \overline{W}$ to the subgroup G of L are certainly unitarily equivalent,

$$V^{-1}(R_L^v \downarrow G)V = (U_N^v W \otimes \overline{W}) \downarrow G, \tag{19}$$

where V is the operator defined in Ref. 2, Eq. (43).

The operator V was chosen so that the representations of the generators of the subgroup N of G transformed in the required manner. Once V had been chosen in this way, it happened that the representations of the remaining generator of G also transformed in the required manner.

These heuristic arguments suggest that, perhaps, $V^{-1}R_L^v V = U_N^v W \otimes \overline{W}$ as representations of L.

6. STATEMENT AND PROOF OF RESULT

Theorem:

$$V^{-1}R_L^v V = U_N^v W \otimes \overline{W} \quad (v \neq 0), \tag{20}$$

where $R_L^v = R^v \downarrow L$ is van Hove's representation of L [Ref. 2, Eq. (12)], U_N^v is a quantum mechanical irreducible representation of N [(2)], V is a unitary operator defined on $L^2(R^{2n}, C)$ [Ref. 2, Eq. (43)] and W is a projective representation of $Sp(2n, R)$ [Sec. 3].

Proof:

Method: Every element $l \in L = N \circledS Sp(2n, R)$ can be expressed uniquely in the form $l = n'S$, where $n' \in N$, $S \in Sp(2n, R)$. From (19), the restriction of the theorem to N is certainly true. Hence, since $V^{-1}R_L^v V$ is an ordinary representation of L, it is sufficient to show that the theorem holds for $Sp(2n, R)$.

Since $Sp(2n, R)$ is a connected Lie group, every element may be expressed as a product of elements of one-parameter subgroups.[9] As $(V^{-1}R_L^v V) \downarrow Sp(2n, R)$ is an ordinary representation of $Sp(2n, R)$, it is thus enough to prove the theorem for a set of independent one-parameter subgroups which generate $Sp(2n, R)$.

If $h(\tau)$ is a one-parameter subgroup of $Sp(2n, R)$, then $(V^{-1}R_L^v V) \downarrow h(\tau)$ and $(W \otimes \overline{W}) \downarrow h(\tau)$ are unitary ordinary representations of the one-parameter subgroup $h(\tau)$, to which Stone's theorem may be applied. Hence, it is sufficient to show that V transforms the representation r_L^v of $sp(2n, R)$ into the representation $w \otimes \overline{w}$ of $sp(2n, R)$ [where $r_L^v, w \otimes \overline{w}$ are the representations of $sp(2n, R)$ corresponding to the representations R_L^v, $W \otimes \overline{W}$ of $Sp(2n, R)$, respectively].

The representation $w \otimes \overline{w}$

Suppose W acts on functions $[\in L^2(R^n, C)]$ dependent on $x \in R^n$, and that $\overline{W}$ acts on functions $[\in L^2(R^n, C)]$ dependent on $y \in R^n$. Let $h(\tau)$ be a one-parameter subgroup of $Sp(2n, R)$. Then

$$(w \otimes \overline{w})(h'(0)) = w(h'(0)) + \overline{w}(h'(0)) \tag{21}$$

(where $'$ denotes differentiation with respect to τ). $\overline{w}(h'(0))$ may be obtained from $w(h'(0))$ by replacing x_m by y_m $(m = 1, 2, \ldots, n)$, and taking the complex conjugate.

From (14) and (21), it follows that

$$\begin{aligned}
&(w \otimes \overline{w})(X_{jk}) = -x_k \frac{\partial}{\partial x_j} - y_k \frac{\partial}{\partial y_j} - \delta_{jk} \quad (1 \leq j, k \leq n), \\
&(w \otimes \overline{w})(Y_{jk}) = \frac{i}{v}\left(-\frac{\partial^2}{\partial x_j \partial x_k} + \frac{\partial^2}{\partial y_j \partial y_k}\right) \quad (1 \leq j \leq k \leq n), \\
&(w \otimes \overline{w})(Z_{jk}) = iv(-x_j x_k + y_j y_k) \quad (1 \leq j \leq k \leq n).
\end{aligned} \tag{22}$$

Explicit expression for $r_L^v(h'(0))$

By definition [Ref. 2, Eq. (12)], $R_L^v \downarrow Sp(2n, R)$ is given by

$$[R_L^v(\gamma)\phi](q, p) = \exp iv\pi_\gamma(\gamma^{-1}(q, p))\phi(\gamma^{-1}(q, p)),$$

where

$$\gamma = \left(0, 0, 0, \begin{pmatrix} A & B \\ C & D \end{pmatrix}\right) \in Sp(2n, R),$$

$$\pi_\gamma(q, p) = \tfrac{1}{2}Ap.Cp + \tfrac{1}{2}Bq.Dq + Bq.Cp,$$

and $\phi \in L^2(R^{2n}, C)$. Therefore,

$$\begin{aligned}
[R_L^v(\gamma)\phi](q, p) = \exp iv[&\tfrac{1}{2}(-AB^T q + AD^T p).(-CB^T q + CD^T p) \\
&+ \tfrac{1}{2}(BA^T q - BC^T p).(DA^T q - DC^T p) \\
&+ (BA^T q - BC^T p).(-CB^T q + CD^T p)] \\
&\times \phi(A^T q - C^T p, -B^T q + D^T p).
\end{aligned} \tag{23}$$

Suppose $h(\tau) = \left(\begin{smallmatrix} A(\tau) & B(\tau) \\ C(\tau) & D(\tau) \end{smallmatrix}\right)$ is a one-parameter subgroup of $Sp(2n, R)$. Then, from (23), since $A(0) = I_n = D(0)$ and $B(0) = 0 = C(0)$,

$$\begin{aligned}
[r_L^v(h'(0))\phi](q, p) &= \frac{d}{d\tau}[R_L^v(h(\tau))\phi](q, p)\Big|_{\tau=0} \\
&= \Bigg[\frac{iv}{2}(p.C'(0)p + B'(0)q.q) \\
&\quad + \sum_{m=1}^{n}\left((A'(0)^T q - C'(0)^T p)_m \frac{\partial}{\partial q_m} + (-B'(0)^T q \right. \\
&\quad \left. + D'(0)^T p)_m \frac{\partial}{\partial p_m}\right)\Bigg]\phi(q, p).
\end{aligned} \tag{24}$$

The action of V

From the definition of V [Ref. 2, Eq. (43)],

$$\begin{aligned} V^{-1}q_jV &= \frac{i}{v}\left(\frac{\partial}{\partial x_j} - \frac{\partial}{\partial y_j}\right), \\ V^{-1}p_jV &= x_j + y_j, \\ V^{-1}\frac{\partial}{\partial q_j}V &= ivx_j, \\ V^{-1}\frac{\partial}{\partial p_j}V &= \frac{\partial}{\partial y_j}. \end{aligned} \tag{25}$$

Evaluation of $V^{-1}r_L^v(X_{jk})V$

In this case, from (7),

$$A'(0)^T = E_{kj}, \quad B'(0) = 0,$$

$$C'(0) = 0, \quad D'(0)^T = -E_{jk}.$$

Hence, from (24),

$$r_L^v(X_{jk}) = q_j\frac{\partial}{\partial q_k} - p_k\frac{\partial}{\partial p_j}.$$

Therefore, using (25),

$$\begin{aligned} V^{-1}r_L^v(X_{jk})V &= (V^{-1}q_jV)\left(V^{-1}\frac{\partial}{\partial q_k}V\right) - (V^{-1}p_kV)\left(V^{-1}\frac{\partial}{\partial p_j}V\right) \\ &= -x_k\frac{\partial}{\partial x_j} - y_k\frac{\partial}{\partial y_j} - \delta_{jk} \\ &= (w\otimes\bar{w})(X_{jk}) \text{ from (22).} \end{aligned}$$

Evaluation of $V^{-1}r_L^v(Y_{jk})V$

From (7) and (24),

$$r_L^v(Y_{jk}) = ivq_jq_k - q_k\frac{\partial}{\partial p_j} - q_j\frac{\partial}{\partial p_k}.$$

Therefore, using (25),

$$\begin{aligned} V^{-1}r_L^v(Y_{jk})V &= -\frac{i}{v}\left(\frac{\partial}{\partial x_j} - \frac{\partial}{\partial y_j}\right)\left(\frac{\partial}{\partial x_k} - \frac{\partial}{\partial y_k}\right) - \frac{i}{v}\left(\frac{\partial}{\partial x_k}\right. \\ &\quad \left. - \frac{\partial}{\partial y_k}\right)\left(\frac{\partial}{\partial y_j}\right) - \frac{i}{v}\left(\frac{\partial}{\partial x_j} - \frac{\partial}{\partial y_j}\right)\left(\frac{\partial}{\partial y_k}\right) \\ &= (w\otimes\bar{w})(Y_{jk}) \text{ from (22).} \end{aligned}$$

Evaluation of $V^{-1}r_L^v(Z_{jk})V$

From (7) and (24),

$$r_L^v(Z_{jk}) = ivp_jp_k - p_k\frac{\partial}{\partial q_j} - p_j\frac{\partial}{\partial q_k}.$$

Therefore, using (25),

$$\begin{aligned} V^{-1}r_L^v(Z_{jk})V &= iv(x_j+y_j)(x_k+y_k) - (x_k+y_k)(ivx_j) - (x_j+y_j)(ivx_k) \\ &= (w\otimes\bar{w})(Z_{jk}) \text{ from (22).} \end{aligned}$$

This completes the proof of the theorem.

7. THE DECOMPOSITION OF THE PROJECTIVE REPRESENTATION W OF $\mathrm{Sp}(2n,\mathbb{R})$

From Ref. 3, Eq. (28),

$$W(\mathrm{Sp}(2n,\mathrm{R}))\downarrow \mathrm{SU}(n) = \bigoplus_{a=0}^{\infty} W_a,$$

where W_a is an irreducible ordinary representation of $\mathrm{SU}(n)$ defined on the subspace Ω_a of $L^2(\mathrm{R}^n,\mathrm{C})$ (Ref. 3, Sec. 5).

Suppose Ω is a subspace of $L^2(\mathrm{R}^n,\mathrm{C})$ which is invariant and irreducible under the projective representation W of $\mathrm{Sp}(2n,\mathrm{R})$. Then, *a priori*, Ω is invariant, although not necessarily irreducible, under $W(\mathrm{Sp}(2n,\mathrm{R}))\downarrow \mathrm{SU}(n)$. Hence Ω is a direct sum of closed subspaces of $L^2(\mathrm{R}^n,\mathrm{C})$ which are invariant and irreducible under $W(\mathrm{Sp}(2n,\mathrm{R}))\downarrow \mathrm{SU}(n)$. Since the irreducible representations W_a of $\mathrm{SU}(n)$ have different dimensions, they are inequivalent, and hence, by the uniqueness (up to unitary equivalence) of the decomposition of $W(\mathrm{Sp}(2n,\mathrm{R}))\downarrow \mathrm{SU}(n)$, it follows that Ω must be a direct sum of subspaces of the form Ω_a.

Suppose Ω contains the subspace Ω_b spanned by the set $\{\psi_m : \sum_{j=1}^n m_j = b\}$. Since Ω is invariant under the projective representation W of $\mathrm{Sp}(2n,\mathrm{R})$, the space generated by Ω_b, $w(X_{jk})\Omega_b$ $(1\le j,k\le n)$, $w(Y_{jk})\Omega_b$ and $w(Z_{jk})\Omega_b$ $(1\le j \le k\le n)$ must be contained in Ω. Now, from (14) and Ref. 3, Eq. (21), for $j\neq k$,

$$\begin{aligned} &w(X_{jk})\psi_m(x) \\ &= -\left[\left(\frac{m_k}{2}\right)^{1/2}\psi_{m_k-1}(x_k) + \left(\frac{m_k+1}{2}\right)^{1/2}\psi_{m_k+1}(x_k)\right] \\ &\quad\times\left[\left(\frac{m_j}{2}\right)^{1/2}\psi_{m_j-1}(x_j) - \left(\frac{m_j+1}{2}\right)^{1/2}\psi_{m_j+1}(x_j)\right] \\ &\quad\times\prod_{l\neq j,k}\psi_{m_l}(x_l). \end{aligned}$$

Therefore, Ω contains elements which are linear combinations of elements belonging to Ω_{b-2}, Ω_b, and Ω_{b+2}. Hence, since Ω is the direct sum of subspaces of the form Ω_a, Ω must contain the subspaces Ω_{b-2}, Ω_{b+2} as well as Ω_b. By induction, it follows that Ω contains all the subspaces Ω_a for which a has the same parity as b.

Hence, the projective representation W of $\mathrm{Sp}(2n,\mathrm{R})$ splits into at most two irreducible projective representations.

From the form of the representation w of the remaining basis elements of $\mathrm{sp}(2n,\mathrm{R})$ [(14)], it follows, again using Ref. 3, Eq. (21), that the subspaces

$$\Omega_{\text{even}}, \quad \text{spanned by } \{\psi_m : \sum_{j=1}^n m_j \text{ is even}\}, \tag{26a}$$

and

$$\Omega_{\text{odd}}, \quad \text{spanned by } \{\psi_m : \sum_{j=1}^n m_j \text{ is odd}\}, \tag{26b}$$

are each invariant under the representation w of $\mathrm{sp}(2n,\mathrm{R})$.

Therefore, the decomposition of the projective representation W of $\mathrm{Sp}(2n,\mathrm{R})$ is

$$W = W_{\text{even}} \oplus W_{\text{odd}}, \tag{27}$$

where W_{even}, W_{odd} are irreducible projective representations of $\mathrm{Sp}(2n,\mathrm{R})$ (namely the restriction of W to the subspaces Ω_{even}, Ω_{odd} respectively).

8. THE DECOMPOSITION OF THE REPRESENTATION $R^v \downarrow L$

$$R^v \downarrow L \simeq U_N^v W \otimes \overline{W}$$
$$= U_N^v W \otimes (\overline{W_{\text{even}}} \oplus \overline{W_{\text{odd}}})$$
$$= (U_N^v W \otimes \overline{W_{\text{even}}}) \oplus (U_N^v W \otimes \overline{W_{\text{odd}}}). \quad (28)$$

Now W_{even} and W_{odd} are irreducible projective representations each of which has the same multiplier as W; so $\overline{W_{\text{even}}}$ and $\overline{W_{\text{odd}}}$ are irreducible projective representations each of which has multiplier inverse to that of W. Hence, from (15), $U_N^v W \otimes \overline{W_{\text{even}}}$ and $U_N^v W \otimes \overline{W_{\text{odd}}}$ are irreducible ordinary representations of L.

It follows that the representation $R^v \downarrow L$ $(v \neq 0)$ splits into two irreducible ordinary representations of L.

9. CONCLUSION

As far as the anisotropic harmonic oscillator is concerned, almost all reference to van Hove's paper[10] can be removed. For clarity, let $W_{\text{Sp}(2n,\text{R})}$ denote the projective representation W of $\text{Sp}(2n,\text{R})$ determined by (14). Van Hove's representation $R^v \downarrow L$ [defined in Ref. 2, Eq. (12)] can now be replaced by the unitarily equivalent representation $U_N^v W_{\text{Sp}(2n,\text{R})} \otimes \overline{W_{\text{Sp}(2n,\text{R})}}$; all that is then required from van Hove's paper is the action of $\text{Sp}(2n, \text{R})$ on N in the definition of $N \circledS \text{Sp}(2n,\text{R})$. The main results obtained can then be summarized as follows.

Main results

I. A group G intrinsically related to the anisotropic harmonic oscillator has been constructed: G is essentially a group generated by the position and momentum observables, the identity operator, and the Hamiltonian of the system. G can be regarded as a subgroup of the group $L = N \circledS \text{Sp}(2n,\text{R})$ (where N is the Weyl group). Let $W_{\text{Sp}(2n,\text{R})}$ be the projective representation of $\text{Sp}(2n,\text{R})$ which intertwines the irreducible representations U_N^v and SU_N^v of N [where $S \in \text{Sp}(2n,\text{R})$]. Then the degeneracies of the energy levels of the anisotropic harmonic oscillator occur in the following way (whether the frequencies are rationally related or not).

$$(U_N^v W_{\text{Sp}(2n,\text{R})} \otimes \overline{W_{\text{Sp}(2n,\text{R})}}) \downarrow G = \bigoplus_s d_{\omega,s} U_G^{v,-(\text{sgn}v)s},$$

where $d_{\omega,s}$ is the degeneracy of the energy level $E_{\omega,s}$ of the n-dimensional anisotropic harmonic oscillator with frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$, $U_G^{v,-(\text{sgn}v)s}$ is an irreducible representation of G, and the summation may be regarded as over all distinct energy levels $E_{\omega,s}$ of the system.

II. Every quantum mechanical irreducible (ordinary) representation of L is unitarily equivalent to one of the form $U_N^v W_{\text{Sp}(2n,\text{R})} \otimes \eta_{\text{Sp}(2n,\text{R})}$, where $\eta_{\text{Sp}(2n,\text{R})}$ is an irreducible projective representation of $\text{Sp}(2n,\text{R})$, with multiplier inverse to that of $W_{\text{Sp}(2n,\text{R})}$.

III. The (ordinary) representation $U_N^v W_{\text{Sp}(2n,\text{R})} \otimes \overline{W_{\text{Sp}(2n,\text{R})}}$ of L (see I) splits into two irreducible (ordinary) representations of L.

IV. Denote $W_{\text{Sp}(2n,\text{R})} \downarrow \text{SU}(n)$ by $W_{\text{SU}(n)}$. Then $U_N^v W_{\text{SU}(n)}$ is an irreducible (ordinary) representation of $N \circledS \text{SU}(n)$; when $U_N^v W_{\text{SU}(n)}$ is restricted to $\text{SU}(n)$, which is a degeneracy group for the n-dimensional isotropic harmonic oscillator, its decomposition is

$$(U_N^v W_{\text{SU}(n)}) \downarrow \text{SU}(n) = \bigoplus_{a=0}^{\infty} W_a,$$

where W_a is an irreducible (ordinary) representation of $\text{SU}(n)$, of dimension equal to the degeneracy of the $(a+1)$th energy level of the n-dimensional isotropic harmonic oscillator. Hence, $N \circledS \text{SU}(n)$ is a noninvariance group for the n-dimensional isotropic harmonic oscillator.

This alternative expression of the results illustrates more clearly the structure of the representation of L which yields the degeneracies of the anisotropic harmonic oscillator (see I). The Weyl group N, together with its irreducible representation U_N^v, is also seen to be important. Lastly, the alternative expression emphasizes the significant role played by the projective representation $W_{\text{Sp}(2n,\text{R})}$.

Note: The original parametrization of G was chosen in such a way that it would immediately fit in with that used by van Hove.[10] Now that the connection with van Hove's representation R^v is no longer required, the group L can be parametrized in other ways. One alternative parametrization which involves $\text{Sp}(2n,\text{R})$ in a more intrinsic way is obtained as follows.

The Weyl form of the commutation relations $[\hat{Q}_j, \hat{P}_k] = i\delta_{jk} I$ is

$$U(\alpha)V(\beta) = \exp i\alpha.\beta V(\beta)U(\alpha), \quad (29)$$

where $\alpha \to U(\alpha)$, $\beta \to V(\beta)$ are unitary representations of the additive groups of momentum space $(\simeq \text{R}^n)$ and configuration space $(\simeq \text{R}^n)$ respectively.[11]

Putting $(\sigma, \beta, \alpha) = \exp i\sigma\, V(\beta)U(\alpha)$, with $\sigma \in \text{R}$, yields the group law of the Weyl group N in the form used earlier [Ref. 1, Eq. (12)]

$$(\sigma', \beta', \alpha')(\sigma, \beta, \alpha) = (\sigma' + \sigma + \alpha'.\beta, \beta' + \beta, \alpha' + \alpha).$$

Alternatively, let M = configuration space $\oplus$ momentum space $(\simeq \text{R}^{2n})$. Define the nondegenerate skew-symmetric bilinear form $[\ ,\ ]$: $M \times M \to \text{R}$ by $[\gamma_1, \gamma_2] = \alpha_1.\beta_2 - \alpha_2.\beta_1$, where $\gamma_i = (\beta_i, \alpha_i) \in M$. Now put $\langle \sigma, \gamma \rangle = \exp i\sigma Z(\gamma)$, where $\sigma \in \text{R}$, and $Z(\gamma) = \exp(i/2)\alpha.\beta V(\beta)U(\alpha)$. Using (29), the group law of N then takes the form

$$\langle \sigma', \gamma' \rangle \langle \sigma, \gamma \rangle = \langle \sigma' + \sigma + \tfrac{1}{2}[\gamma', \gamma], \gamma' + \gamma \rangle.$$

It is easily verified that the group $\text{Sp}(2n,\text{R})$ can be characterized as the set of all $S \in \text{GL}(2n,\text{R})$ which satisfy $[S\gamma_1, S\gamma_2] = [\gamma_1, \gamma_2]$, for any $\gamma_i \in \text{R}^{2n}$.

The group law of $L = N \circledS \text{Sp}(2n,\text{R})$ can then be taken as

$$\langle \sigma', \gamma', S' \rangle \langle \sigma, \gamma, S \rangle = \langle \sigma' + \sigma + \tfrac{1}{2}[\gamma', S'\gamma], \gamma' + S\gamma, S'S \rangle.$$

The connection between the two parametrizations of L is $\langle \sigma, \beta, \alpha, S \rangle = (\sigma + \tfrac{1}{2}\alpha.\beta, \beta, \alpha, S)$. This change in parametrization means that the relation $(\sigma, \beta, \alpha, I) = (\sigma, 0, 0, I)(0, \beta, 0, I)(0, 0, \alpha, I)$ is replaced by $\langle \sigma, \beta, \alpha, I \rangle = \langle \sigma + \tfrac{1}{2}\alpha.\beta, 0, 0, I \rangle \langle 0, \beta, 0, I \rangle \langle 0, 0, \alpha, I \rangle$, with corresponding minor modifications in several places. The characterization of $\text{Sp}(2n,\text{R})$ as a group leaving $[\ ,\ ]$ invariant results in easier calculations in one or two places, but does not lead to any overall simplifications.

[1]M.E. Major, J. Math. Phys. 18, 1938 (1977).
[2]M.E. Major, J. Math. Phys. 18, 1944 (1977).
[3]M.E. Major, J. Math. Phys. 18, 1952 (1977).
[4]K.R. Parthasarathy, *Multipliers on Locally Compact Groups*, Lecture Notes in Mathematics, Vol. 93 (Springer, Berlin, 1969).
[5]E. Nelson, Ann. Math. 70, 572 (1959).
[6]V.S. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, Princeton, N. J., 1970), Vol. II, Chapter 10.
[7]G.W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, and Editore Boringhieri, Torino, 1968).
[8]C. Itzykson and M. Nauenberg, Rev. Mod. Phys. 38, 95 (1966).
[9]L.S. Pontryagin, *Topological Groups* (Gordon and Breach, New York, 1966).
[10]L. van Hove, Mem. Acad. Roy. Belg. Cl. Sci. 8°, 26, fasc. 6, 1 (1951).
[11]I.E. Segal and G.W. Mackey, *Mathematical Problems of Relativistic Physics*, Lectures in Applied Mathematics, Vol. 2. (American Mathematical Society, Providence, R.I., 1967).

A finite quantum electrodynamical base for many photon collective phenomena*

M. Stelzer, J. Mühlstein, and V. Ernst

Sektion Physik der Universität Müchen, Theoretische Physik, Theresienstr. 37, 8000 München 2, Germany
(Received 29 September 1976)

We analyze the interaction of many "nonoverlapping" atoms with the common radiation field of transverse photons over a continuum of modes. Typical effects of quantum optics arising from this interaction require extensions to very high powers in the coupling constant e and diverge therefore within a systematic perturbation approach. Since renormalization to such high orders is not feasible in practice, we formulate for the purpose of quantum optics "a new peace treaty between QED and its infinities." This is a compromise between the necessity of including effects of very high powers in e and the trivial demands for finite results. We show that a unitary time evolution operator exists if, in essence, only any finite number of levels of each atom is treated as "existent." We convince ourselves that the limits of the concept of "nonoverlapping atoms" are reached long before any infinities can enter. The resulting theories meet still reasonable requirements with respect to the dualism of light, Dicke principle, and causality. They are formulated so that some practical demands of quantum optics (homogeneous and inhomogeneous line broadening, Doppler effects) can be met easily if desired.

1. INTRODUCTION

Typical processes of quantum optics, the most common one being the creation of a spike in a laser, involve the simultaneous emission of many photons by many atoms. In the "order hierarchy" of the Feynman–Dyson expansion of quantum electrodynamics (QED) such processes are of very high order m, typically $m \geq 10^6 - 10^{20}$. The Feynman rules tell us namely that a "coherent" process involving the emission of $\sim m$ photons requires at least an expansion to the order m in the coupling constant e. Actually, the needed order is much higher because the typical processes of quantum optics arise from *many* photons exchanged between the active atoms and each absorption–reemission act of a photon by any atom requires at least two additional orders in e. The above m accounts only for the "final" emission of each photon. But for $m \geq 2$ the results of a systematic perturbation approach are divergent, in general, and must be renormalized to the high orders mentioned. This is not possible *in practice*. Therefore, "the peaceful coexistence of QED with its infinities"[1] cannot be extended to quantum optics, its most challenging offspring. We therefore formulate here for the typical problems of quantum optics a "new peace treaty" which guarantees finite results in any finite step of calculation and, in our opinion, still meets all essential and reasonable demands of a quantum field theory of optical phenomena. These demands are:

1. Von Laue[2] showed in 1907 that resolvable interference effects cannot occur in light from two different macroscopical sources if one assumes that the acts of emission or absorption of wavetrains by different atoms are statistically independent processes. Weisskopf[3] pointed out in 1930 that the acts of emission or absorption of photons by different atoms in a given region of space are not independent *in principle*. Dicke[4] seems to have recognized first that the necessary coupling of all active atoms to the one common, "universal" radiation field R may lead to novel, macroscopically observable many photon collective phenomena. In the meantime, much work has been done along these lines (cf. Refs. 5–13 and the long lists of references given especially in Refs. 11 and 13). It has become clear[5–7] that Dicke's collective phenomena must be identified with the laser and superradiance effects detected later. The "Dicke principle," the necessity of coupling all active atoms to the one radiation field R of transverse photons, offers indeed the only convincing arguments against the theorem of von Laue. For lack of space we cannot treat here the attempts to ignore this[14] and alternative explanations of laser activity. Compliance with the Dicke principle, however, excludes already at this point the use of rate equations and necessitates the high order expansions mentioned.

The Dicke principle is an immediate consequence of the Bose principle the "photons from different atoms" are subject to, and of causality ("exchanged photons"). It can be visualized as follows: Let $\mathbf{A}(\mathbf{x}, t)$ be the Heisenberg operator of the observable "transverse part of the vector potential" of R, and let $\psi^j(\mathbf{x}, t)$ be the Heisenberg operator of a fermion field amplitude "associated" with an optically active atom A^j, $j = 1, \ldots, J$. We assume that A^j carries out a prescribed, nonrelativistic motion so that $\mathbf{X}_t^j$ is its position at time t, and its potential is given by a real c-number function $V_t^j(\mathbf{x} - \mathbf{X}_t^j)$ "centered" about $\mathbf{X}_t^j$. Let m_0^j denote the effective mass of the electrons of A^j and $\boldsymbol{\alpha}, \beta$ the usual Dirac matrices. Then the mutual coupling of atoms through R is reflected best by the following Heisenberg equations of motion:

$$i \frac{d}{dt} \psi^j(\mathbf{x}, t) = \left[-i\boldsymbol{\alpha}\nabla + \beta m_0^j + V_t^j(\mathbf{x} - \mathbf{X}_t^j) + e\boldsymbol{\alpha}\mathbf{A}(\mathbf{x}, t)\right] \psi^j(\mathbf{x}, t), \quad j = 1, \ldots, J, \tag{1a}$$

and

$$\left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right) \mathbf{A}(\mathbf{x}, t) = e \sum_{j=1}^{J} (\psi^{*j}(\mathbf{x}, t)\boldsymbol{\alpha}\psi^j(\mathbf{x}, t))_{\mathrm{tr}}. \tag{1b}$$

For $J = 1$ this yields a version of QED restricted to transverse photons. For $J > 1$ each atom A^j reacts, by

*Supported in part by Deutsche Forschungsgemeinschaft, Contracts No. ER 68/12-4.

(1a), on the common R, "represented" by its observable $\mathbf{A}(\mathbf{x}, t)$. $\mathbf{A}(\mathbf{x}, t)$ plays in (1a) the role of an operator vector potential for A^j. But by (1b) it has *all* atoms as "sources," namely the transverse part $(\cdots)_{tr}$ of the "current" $e\psi^{+j}(\mathbf{x}, t)\boldsymbol{\alpha}\psi^j(\mathbf{x}, t)$ associated with A^j and, in essence, "localized" about $\mathbf{X}_t^j$. So all atoms "see each other" through R, and it is no longer allowed to speak of interaction processes of single atoms: There are only collective interaction processes. Note that *all possible* collective processes are described by the *one* operator solution $(\psi^j(\mathbf{x}, t), \mathbf{A}(\mathbf{x}, t))$ of (1). These Heisenberg operators namely have the corresponding Schrödinger operators as initial values. As the latter are always *given* and unique, only this one solution of (1) can be of interest to us.

2. The "dualism of light" requires the quantization of the electromagnetic field as in (1) (or some equivalent not yet known). It has been shown beyond doubt[15, 16] that a laser beam shows dualistic features which cannot be accounted for by a "classical electromagnetic wave" $\mathbf{E}(\mathbf{x}, t), \mathbf{B}(\mathbf{x}, t)$.

3. Signals transferred by photons should *a priori* have the freedom to move and spread like classical electromagnetic waves in free space. This requires at least a *complete* set of "modes" of R and thus forbids the usual "self-consistent coupling of the active atoms to few distinguished modes" only. Because of (1) we can expect causality to hold in a theory *defined* by (1); care must be given, however, that this is not lost in later approximations (cf. Sec. 7).

4. A finite base of quantum optics should be as simple as possible, provided that this is compatible with 1–3. If the desired solution of (1) exists it contains the results on all possible states of the coupled system of one Bose field R and J Fermi fields associated with J atoms. In particular, each fermion field could still be in a state of s^j quanta, i.e., A^j can still be any "s^j-electron atom," $s^j = 0, 1, 2, \cdots$ (cf. Sec. 3, however). The desired solution of (1) describes all these possibilities. Therefore, an equivalent of (1) must be found which allows us, for example, to consider "one-electron atoms" only. In principle we can achieve this easily, and without any loss in comparison with (1), but in practice this requires some tedious work (Sec. 3).

5. In practice it is necessary to account also for the influence of the surrounding of the active atoms (heat bath, finite temperature $T, \cdots$) and for the influence of their motion (Doppler effects) on the effects arising from the Dicke principle. Details on this are formulated in Sec. 2.

Our compromise between these demands and the trivial necessity for finite results is based on the hierarchy of unitary Weisskopf–Wigner theories systematized recently.[17] Its main clauses will be discussed in Sec. 7 when more details are available. Following Ref. 6, we extend here the ideas of Ref. 17 to the interaction of *many* optically active atoms $A^1 \cdots A^j$ with R. As no new mathematical ideas are needed, we take the occasion to meet also the practical demands 4 and 5 (cf. Sec. 2). Our main concern will be the absence of infinities, however. The literature on quantum optics is marred with infinities which have nothing to do with the field theoretical problems of (1). Our results should be useful to avoid such erroneous, but nonetheless irritating, infinities. One may also forget the irritating scruples one *had* to have when infinities did not occur in processes, e.g. Ref. 6, which in perturbation theory are of order $\sim 10^{20}$. Not much is known about the *convergence* of the approximation method suggested (cf. Sec. 7), but this holds for the renormalized perturbation series as well.

In Sec. 2 we formulate physical postulates which lead to (1) and take account of the above demands. In Sec. 3 we discuss additional postulates which are *sufficient* to guarantee the existence of finite results. In Sec. 4 we "restrict" Eqs. (1) to one-electron atoms. Sections 5 and 6 contain the proofs of existence of finite results. In Sec. 7 we collect some conclusions and an outlook on further problems.

2. A FORMAL DERIVATION OF EQs. (1)

At the beginning we formulate nine postulates P_1–P_9 which allow the formal derivation of (1). This permits a review of the rich physical content of (1) and the introduction of the necessary notation.

P_1: *Each active atom A^j is determined by a given, in general, time-dependent c-number potential $V_t^j(\mathbf{x}, t)$ "centered" about the origin $\mathbf{x}^j = 0$ of the coordinates $\mathbf{x}^j$ of the rest frame of A^j.* The dependence on j accounts in a phenomenological way for effects of the "individual" surrounding of A^j on the properties of A^j, in particular on its eigenvalues and eigenstates. The atoms are therefore "different," in general. The time dependence accounts for "dynamical" surrounding effects causing fluctuations of the effective potential, as arising, e.g., from lattice vibrations. Motion of "A^j as a whole" is introduced by

P_2: *Each atom A^j moves in the laboratory system on a prescribed space curve $\mathbf{X}_t^j$ so that the atoms never overlap and the instantaneous velocity $\dot{\mathbf{X}}_t^j$ of A^j relative to the laboratory system is always small against the velocity of light.* This implies that Galilei instead of Lorentz transformations can be used for the transition from the common laboratory system to the individual rest frames. The potential of A^j in the laboratory system thus is $V_t^j(\mathbf{x} - \mathbf{X}_t^j)$.

P_3: *In the laboratory system the equation of motion for one electron in the potential $V_t^j(\mathbf{x} - \mathbf{X}_t^j)$ shall read*

$$i\frac{d}{dt}v^j(\mathbf{x} - \mathbf{X}_t^j, t) = \left[\boldsymbol{\alpha}\left(-i\frac{\partial}{\partial \mathbf{x}}\right) + \beta m_0^j + V_t^j(\mathbf{x} - \mathbf{X}_t^j)\right] v^j(\mathbf{x} - \mathbf{X}_t^j, t). \tag{2}$$

The choice of $V_t^j(\mathbf{x}^j)$ shall be restricted by the condition that there exists a set of solutions of (2) of the form

$$v^j(\mathbf{x} - \mathbf{X}_t^j, t) = u_{t,a_j}^j(\mathbf{x} - \mathbf{X}_t^j)\exp\left(-i\int_0^t dt\, E_{t,a_j}^j\right) \tag{3}$$

where $E_{t,a_j}^j > 0$ and $u_{t,a_j}^j(\mathbf{x}^j)$ depends "weakly" on t. The set of all functions $u_{t,a_j}^j(\mathbf{x} - \mathbf{X}_t^j)$, characterized and distinguished from each other by the values of an index a_j

comprising a set of necessary quantum numbers, shall at any time t be orthonormal and complete in the Hilbert space $L_2(\mathbb{R}^3) \oplus L_2(\mathbb{R}^3)$ *of "electron" solutions of the Dirac equation (2).* Since we consider "electron" states only, we have $E^j_t \sim m^j_0$ and condition $E^j_{t,a_j} > 0$ poses not an essential restriction. The completeness of the $u^j_{t,a_j}(\mathbf{x}^j)$ is needed for the "derivation" of (1). A Dirac, instead of a Schrödinger or Pauli equation, was chosen to avoid the later occurrence of a term $\mathbf{A}^2(\mathbf{x})$ which might be small in the unquantized versions[18] but becomes infinite in the quantized theory. Since $V^j_t(\mathbf{x}-\mathbf{X}^j_t)$ is "concentrated" about the point $\mathbf{x}=\mathbf{X}^j_t$, the essential solutions of (2) will in essence also be "concentrated" about the point $\mathbf{x}=\mathbf{X}^j_t$. We say that "$A^j$ is at $\mathbf{X}^j_t$."

The "weak" time dependence of the "pseudo eigenfunctions" $u^j_{t,a_j}(\mathbf{x}^j)$ need not be specified quantitatively. Its purpose is the following: If $V^j_t(\mathbf{x}^j)$ is independent of t, $v^j(\mathbf{x}-\mathbf{X}^j_t, t)$ shall assume the form $u^j_{a_j}(\mathbf{x}^j)\exp(-itE^j_{a_j})$ with "proper," time-independent eigenfunctions $u^j_{a_j}(\mathbf{x}^j)$ and eigenvalues $E^j_{a_j} > 0$. We accept so that in the transition from the laboratory system to an individual rest frame a term $\dot{\mathbf{X}}^j_t\nabla$ is neglected against the term $\alpha\nabla$ in the Dirac equation; note that $|\alpha| = 1 \gg |\dot{\mathbf{X}}^j_t|$. A time dependent potential in general does not allow for such solutions. But even under strong "dynamic perturbations" we can expect that solutions of form (3) exist, which contain the *strong* dependence on t in the form of a phase factor. Otherwise the fiction of "atomic states" and of "transitions between them" loses its meaning altogether.

We so try to account for homogeneous (and, if desired, some additional inhomogeneous) broadening[16] effects, as now no "static" eigenvalues and no "sharp transition frequencies" exist. So we avoid the field theoretical problems, namely additional infinities, which arise if the field operators $\mathbf{A}(\mathbf{x},t)$ of R and $\psi^j(\mathbf{x},t)$ of A^j are coupled by corresponding Hamiltonians[18] to the observables of other fields which account for homogeneous broadening by surrounding effects ("heat baths"). We emphasize that homogeneous broadening by the *natural width* of emission lines and its *strong modifications arising from the Dicke principle*[4] are fully contained in our theory. One of its aims, in fact, is to serve as a reliable base for the derivation of such effects. *Inhomogeneous broadening* has been introduced by P_1 (different eigenvalues) and by P_2 (Doppler broadening).

P_4: *Electrons belonging to different atoms A^j can be considered as quanta of different fermion fields.* This postulate is used silently in practically all works on quantum optics. It can be looked upon as a consequence of the "nonoverlapping of the active atoms" which permits the exclusion of tunneling and exchange effects of electrons belonging to different atoms. The "more correct" alternative had been the assumption that all involved electrons are quanta of one fermion field, with $\sum_j V^j_t(\mathbf{x}-\mathbf{X}^j_t)$ as potential. But electrons of "optically inactive atoms" then have the same right to be considered as quanta of this field. We have made the necessary "cut" at a stage where we can gain some advantages without visibly prejudicing the desired results on quantum optics. P_4 implies that a field operator $\psi^j(\mathbf{x},t)$ (and an individual mass m^j_0, if desired) is associated with each atom; in the alternative case only one $\psi(\mathbf{x},t)$ (and only one m_0) would be introduced.

P_5: *Creation and destruction of electron positron pairs play no role.* This excludes probably the worst infinities of field theory, namely vacuum polarization effects. In quantum optics it appears well justified. In connection with P_2 and P_4 it means that the *Schrödinger* operator $\psi^j_t(\mathbf{x})$ associated with A^j can be spanned up by

$$\psi^j_t(\mathbf{x}) = \sum_{a_j} u^j_{t,a_j}(\mathbf{x}-\mathbf{X}^j_t) b^j_{a_j}, \tag{4}$$

the sum going over all pseudo one-electron eigenstates in the potential of A^j. $b^j_{a_j}$ "destructs" an electron in the state $u^j_{t,a_j}(\mathbf{x}-\mathbf{X}^j_t)$. With the corresponding creation operator $b^{*j}_{a_j}$ it satisfies the usual anticommutation relations. $\psi^j_t(\mathbf{x})$ depends "weakly" on t because it refers to a "field" subject to "external," yet prescribed "perturbations." It is important, however, that the anticommutators

$$[\psi^{*j}_t(\mathbf{x}), \psi^j_t(\mathbf{x}')]_+ = \delta(\mathbf{x}-\mathbf{x}'), \quad \text{etc.}, \tag{5}$$

are independent of t.

P_6: *The state spaces of all occuring fermion fields are Fock spaces.* We are aware of the problems involved in choosing Fock spaces as state spaces for interaction fields.[19] We try to avoid them by "heuristic" restrictions on the interacting systems. By P_6 the operator $\psi^j_t(\mathbf{x})$ acts on a fermion Fock space $\mathfrak{F}^j_e$ whose elements $|f^j_e\rangle$ will be written in the form

$$|f^j_e\rangle = f^j_0|v^j_e\rangle + \sum_{\rho=1}^{\infty}\sum_{a^1_j}\cdots\sum_{a^\rho_j} f^j_\rho(a^1_j,\ldots,a^\rho_j) \times b^{*j}_{a^1_j}\cdots b^{*j}_{a^\rho_j}|v^j_e\rangle. \tag{6}$$

$|v^j_e\rangle$ is the vacuum in $\mathfrak{F}^j_e$ satisfying $b^j_{a_j}|v^j_e\rangle = 0$ for any a_j. $f^j_0 \in \mathbb{C}$. The $f^j_\rho(a^1_j,\ldots,a^\rho_j)$ are antisymmetric c-number functions of the "discrete" variables $a^1_j,\ldots,a^\rho_j$, subject to normalization conditions which restrict $|f^j_e\rangle$ to an element of $\mathfrak{F}^j_e$.

P_7: *Only transverse photons need to be considered; their state space is a Bose Fock space.* We consider here only the *formal* implications of P_7; other consequences are drawn later below. Let $\mathfrak{F}_p$ denote the Fock space of elements $|f^p\rangle$ to be written in the form[1]

$$|f^p\rangle = f^p_0|v_p\rangle + \sum_{n=0}^{\infty}\frac{1}{(n!)^{1/2}} \times \int d^3\kappa_1\cdots\int d^3\kappa_n f^p_n(\kappa_1,\ldots,\kappa_n) a^*_{\kappa_1}\cdots a^*_{\kappa_n}|v_p\rangle. \tag{7}$$

$|v_p\rangle$ is the normalized photon vacuum stated defined by $a_\kappa|v_p\rangle = 0$ for any $\kappa \in K$. $\kappa := (\mathbf{k},\lambda)$ comprises the wave vector $\mathbf{k} \in \mathbb{R}^3$ and polarization index $\lambda \in \{1,2\}$ and varies over the set $K := \mathbb{R}^3 \times \{1,2\}$. $\int d^3\kappa \cdots$ denotes the integral over $\mathbf{k} \in \mathbb{R}^3$ *and* the sum over $\lambda \in \{1,2\}$. a_κ and a^*_κ are the usual Bose destruction and creation operators satisfying $[a_\kappa, a^*_{\kappa'}] = \delta(\mathbf{k}-\mathbf{k}')\delta_{\lambda\lambda'}$. $f^p_0 \in \mathbb{C}$. The

$f_n^p(\kappa_1,\ldots,\kappa_n)$, $n\geq 1$, are Lebesque square integrable, symmetric c-number functions of n arguments $\kappa_1,\ldots,\kappa_n$ subject to normalization conditions so that $|f^p\rangle$ is element of a Hilbert space. The Schrödinger operator of the observable $\mathbf{A}(\mathbf{x})$ of the radiation field R is given by

$$\mathbf{A}(\mathbf{x})=\frac{1}{(2\pi)^{3/2}}\int d^3\kappa\,\frac{\boldsymbol{\epsilon}(\kappa)}{[2\omega(\kappa)]^{1/2}}\,(e^{i\mathbf{k}\mathbf{x}}a_\kappa+e^{-i\mathbf{k}\mathbf{x}}a_\kappa^*), \tag{8}$$

to be understood as an operator on $\mathfrak{F}_p$. $\boldsymbol{\epsilon}(\kappa)$ is the usual unit polarization vector satisfying $\mathbf{k}\boldsymbol{\epsilon}(\mathbf{k},\lambda)=0$ and $\boldsymbol{\epsilon}(\mathbf{k},\lambda_1)\boldsymbol{\epsilon}(\mathbf{k},\lambda_2)=\delta_{\lambda_1\lambda_2}$. $\omega(\kappa):=\omega(\mathbf{k}):=\omega(|\mathbf{k}|):=(\mathbf{k}^2+\mu^2)^{1/2}$; $\mu\geq 0$ is a photon mass parameter discussed below.

P_8: *The state space of the interacting system* $R+A^1+\cdots+A^J$ *shall be the Hilbert space*

$$H:=\mathfrak{F}_p\otimes\mathfrak{F}_e^1\otimes\cdots\otimes\mathfrak{F}_e^J, \tag{9}$$

where $\otimes$ *denotes the usual tensor product.* H is the *smallest* Hilbert space the basic Eqs. (1) may be defined on possibly. For the practical purposes of quantum optics it is still too large (cf. Secs. 3, 4).

P_9: *The Hamiltonian of the interacting system* $R+A^1+\cdots+A^J$ *shall be the operator* H_t *defined formally on* H *by*

$$\begin{aligned}H_t:=&\int d^3\kappa\omega(\kappa)a_\kappa^*a_\kappa\\&+\sum_{j=1}^{J}\Big\{\int d^3x\,\psi_t^{*j}(\mathbf{x})\Big[\boldsymbol{\alpha}\Big(-i\frac{\partial}{\partial\mathbf{x}}\Big)+m_0^j\beta+V_t^j(\mathbf{x}-\mathbf{X}_t^j)-i\frac{d}{dt}\Big]\\&\times\psi_t^j(\mathbf{x})+e\int d^3x\,\mathbf{A}(\mathbf{x})(\psi_t^{*j}(\mathbf{x})\boldsymbol{\alpha}\psi_t^j(\mathbf{x}))_{\mathrm{tr}}\Big\}.\end{aligned} \tag{10}$$

Due to the external forces causing the fluctuations and motion of $V_t^j(\mathbf{x}-\mathbf{X}_t^j)$ the Hamiltonian depends on t. The first term is the usual Hamiltonian of free photons. The first term in the sum over j is the Hamiltonian H_0^j of the "not interacting" atom A^j, chosen so that Eqs. (1) result from (10). The time derivative operator in H_0^j acts only on the time dependence contained in $\psi_t^j(\mathbf{x})$. Inserting (4) and (3) into H_0^j we find, using (2) and the completeness relation,

$$H_0^j=\sum_{a_j}E_{t,a_j}^j b_{a_j}^{*j}b_{a_j}^j. \tag{11}$$

This reflects once more that the "fluctuations" of $V_t^j(\mathbf{x}^j)$ cause homogeneous and inhomogeneous broadening. The second term in the sum over j in (10) is the usual coupling Hamiltonian of quantum electrodynamics, restricted to transverse photons. The sum over the atoms reflects the Dicke principle: All atoms "see" the same field R and by this become coupled to each other even in the absence of *direct* atom–atom interaction.

Postulates P_1–P_9 are sufficient for the formal derivation of Eqs. (1). Indeed, let $U(t)$ be the unitary solution of the equation

$$i\frac{d}{dt}U(t)=H_tU(t),\quad U(0)=1, \tag{12}$$

and define the Heisenberg operators corresponding to $\psi_t^j(\mathbf{x})$, $\mathbf{A}(\mathbf{x})$ by

$$\begin{aligned}&\psi^j(\mathbf{x},t):=U^*(t)\psi_t^j(\mathbf{x})U(t),\\&\mathbf{A}(\mathbf{x},t):=U^*(t)\mathbf{A}(\mathbf{x})U(t).\end{aligned} \tag{13}$$

Using the above definitions and commutation relations, we easily find that expressions (13) *solve* Eqs. (1) formally. The postulates P_1–P_9 so elucidate, indeed, the richly structured physical content of (1).

3. ADDITIONAL POSTULATES AND THE EXISTENCE THEOREM

Unfortunately, however, by all we know, the desired $U(t)$ does not exist in a naive sense. Therefore we must set up additional postulates P_{10}–P_{13} which allow the *proof of existence* of a unitary time evolution operator which guarantees finite results.

The exclusion of longitudinal and scalar photons by P_7 excludes the Coulomb interaction between the electrons of A^j, which must be accounted for on a phenomenological base. This has been done, in part, by P_1 so that P_7 is in a sense a necessary consequency of P_1. The atoms described by (1) are therefore not realistic unless we restrict them by

P_{10}: *All active atoms are one-electron atoms.* From the field theoretical point of view this is probably not decisive, but the treatment of atoms with more electrons were rather inconvenient. It is of some practical importance, however.

So far it has not yet been necessary to speak of laser mirrors: On the "atomic level" a part of the active atoms can be chosen and arranged always as "atoms of laser mirrors." But mirrors have most probably nothing to do with the infinities we want to exclude here, nor with the "microscopic creation mechanism" of the many photon collectives we want to get a base for. They are probably much better introduced at a later stage and in a much more phenomenological way, just as "electrical wires" for good reasons do not occur in the first principles of classical electrodynamics. In view of this we have already imposed conditions which in essence require the omission of mirrors. Care has been taken that the desired phenomena of quantum optics are not prejudiced and the resulting equations of motion still contain solutions describing radiation of "black body radiation consistency" as well as solutions describing laser or superradiance phenomena. It is a later problem to find the condition under which this or that occurs.

P_{11}: *Only a finite number* $q_j<\infty$ *of pseudoeigenstates* $u_{t,a_j}^j(\mathbf{x}^j)$ *of each* A^j *are essential for the interaction between* R *and* $A^1+\cdots+A^j$; *all other pseudoeigenstates can be treated as nonexisting.* This is probably the main contribution to the price we have to pay for finite results. For the typical problems of quantum optics it seems to be acceptable, however: Since the "diameters" of atoms tend to infinity as the eigenenergies approach infinity or the ionization limit,[20] the condition of "nonoverlapping" can be met only as long as all atoms remain in their lower states. Therefore the fiction of "many atoms" is realizable only with atoms with a finite, not too large number of

levels. *In fact, including too many states of each atom would introduce effects which in the form, as they are introduced, do not occur in reality*: Electrons in overlapping states are subject to the Pauli principle which does not play a role in our model. One could think also of infinitely high potential walls between the atoms. But then the eigenstates are not complete and the "second quantization" of such atoms is not obvious; for example, (4) and (5) were not compatible then. It is questionable, finally, whether this was more realistic than the fiction of finite level atoms.

P_{12}: *For any couple* $u^j_{t,a_j}(\mathbf{x}^j), u^j_{t,a'_j}(\mathbf{x}^j)$ *of "essential" pseudoeigensolutions of* A^j *the expression*

$$M^{*j}_t(a_j,a'_j;\kappa) := \frac{e\epsilon(\kappa)}{(2\pi)^{3/2}[2\omega(\kappa)]^{1/2}} \times \int d^3x\, e^{i\kappa\mathbf{x}} u^{j*}_{t,a_j}(\mathbf{x}^j)\alpha u^j_{t,a'_j}(\mathbf{x}^j) \qquad (14)$$

is an element of $L_2(K)$, *i.e., for* $\lambda = 1,2$ *it is square integrable in* $\mathbf{k}$, *and its norm in* $L_2(K)$ *is uniformly bounded with respect to* t. Formally this puts another restriction on the choice of $V^j_t(\mathbf{x}^j)$. It is satisfied[17] if $V^j_t(\mathbf{x}^j)$ is the Coulomb potential. In general it is a consequence of the *finite spatial extension of atomic* (pseudo) *eigenstates*. In practice, P_{12} "forbids" only the use of unphysical secondary approximations such as the dipole approximation.[21]

P_{13}: *Only a finite number* $n \le N < \infty$ *of photons of mass* $\mu \ge 0$ *take part in the interaction with the atoms*. Alternatively we postulate:

P'_{13}: *The photons have mass* $\mu > 0$; *all atoms have time independent potentials* $V^j(\mathbf{x}^j)$ *and move with the same velocity* v. It is unlikely that the relative motion of the atoms or fluctuations of the $V^j(\mathbf{x}^j)$ lead to problems, but we have no technical proof of this. It is clear that the restriction to any finite number of photons is *a priori* much more serious that the imposition of a photon mass. But in practice it excludes only infrared problems which typically involve infinitely many low frequency photons. Infinitely many photons of finite energy require an infinite total energy. They do not occur in realistic quantum optics problems.

The precise meaning of P_{11}–P_{13}, P'_{13} will be given in connection with the proof of the existence theorem, our main results:

Postulates P_1–P_{13} *or, alternatively, postulates* P_1–P_{12} *and* P'_{13} *are sufficient to guarantee for any* $t < \infty$ *the existence of a unitary time evolution operator* $U(t)$ *for the interacting system* $R + A^1 + \cdots + A^J$.

This implies, in particular, the absence of any infinities in observable quantities, as all transition probabilities are finite with certainty and properly normed to 1. See also Sec. 7.

4. LINEARIZATION OF THE PROBLEM IN THE SCHRÖDINGER PICTURE

In Sec. 2 the *linear* Schrödinger equation (12) provided us formally with the exact solution of the *nonlinear* Heisenberg equations of motion (1). For practical purposes the latter have still another disadvantage: their generality. Despite their formal appearance, the operators $\psi^j(\mathbf{x},t)$ in (1) namely do *not* refer to one-electron atoms A^j only; they refer simultaneously to the "states with ρ^j electrons in the potential $V^j(\mathbf{x}^j)$, $\rho^j = 1,2,\ldots$, *without* mutual Coulomb interaction," and to all states with "unsharp electron numbers" as well. But, as mentioned already, H is the "smallest" Hilbert space where a meaning can be given to (1). To shake off some unnecessary ballast we first "reduce," following P_{10}, Eqs. (1) to one-electron atoms. In the Schrödinger picture this can be achieved without any loss of physical content with respect to one-electron atoms.

We note that the Hamiltonian H_t defined in (10) commutes with the electron number operator

$$\sum_{a_j} b^{*j}_{a_j} b^j_{a_j} \qquad (15)$$

for each atom. Therefore, if a time evolution operator $U(t)$ exists at all, it must be the direct sum[22] of unitary operators $U_{\rho_1,\ldots,\rho_J}(t)$ defined and acting on the subspaces

$$S_{\rho_1\cdots\rho_J} := H^1_{e,\rho_1} \otimes \cdots \otimes H^J_{e,\rho_J} \otimes \mathfrak{F}_p \qquad (16)$$

of H, H^j_{e,ρ_j} being the ρ_j-electron subspace of $\mathfrak{F}^j_e$. Correspondingly, H_t must be a direct sum

$$H_t = \bigoplus_{\rho_1=1}^{\infty} \cdots \bigoplus_{\rho_J=1}^{\infty} H_{t,\rho_1,\ldots,\rho_J} \qquad (17)$$

of Hamiltonians acting on $S_{\rho_1,\ldots,\rho_J}$ only. Therefore, we can restrict our considerations without any loss to one-electron atoms by considering the time evolution on the sector $S := S_{1,\ldots,1}$ only. Any element $|\alpha\rangle \in S$ can be written in the form

$$|\alpha\rangle = \sum_{a_1} \cdots \sum_{a_J} \sum_{n=0}^{\infty} \frac{1}{(n!)^{1/2}} \times \int d^3\kappa_1 \cdots \int d^3\kappa_n \alpha[a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n) \times a^*_{\kappa_1}\cdots a^*_{\kappa_n} b^{*1}_{a_1}\cdots b^{*J}_{a_J}|v_e\rangle|v_p\rangle. \qquad (18)$$

For $n = 0$ the integrals are to be understood as a factor 1.

Let the restriction of H_t on S, $H_{t,1,\ldots,1}$, be denoted by H_{tS}. If the theory existed, H_{tS} would generate on S the unitary time evolution operator $U_{1,\ldots,1}(t) := U_S(t)$ so that the states of the interacting system $R + A^1 + \cdots + A^J$ at different times t, $|\alpha(t)\rangle = U_S(t)|\chi\rangle$, form in S a "curve" defined by the Schrödinger equation

$$i\frac{d}{dt}|\alpha(t)\rangle = H_{tS}|\alpha(t)\rangle, \quad |\alpha(0)\rangle = |\chi\rangle, \qquad (19)$$

and an arbitrary $|\chi\rangle \in S$ as "initial state." The best definition of the restriction H_{tS} seems to consist of writing (19) in terms of the amplitudes $\alpha[t;a_1,\ldots,a_J]_n \times (\kappa_1,\ldots,\kappa_n)$ of the vector $|\alpha(t)\rangle$ and the amplitudes $\chi[a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)$ of the vector $|\chi\rangle = |\alpha(0)\rangle$ in the sense of (18). Equation (19) then reads

$$i\frac{d}{dt}\alpha[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)=(\omega(\kappa_1)+\cdots+\omega(\kappa_n)+E^1_{a_1}+\cdots+E^J_{a_J})\,\alpha[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)$$
$$+(n+1)^{1/2}\sum_{j=1}^{J}\sum_{a'_j}\int d^3\kappa M_t^{*j}(\mathbf{X}_t^j;a_j,a'_j;\kappa)$$
$$\times\,\alpha[t;a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n+1}(\kappa,\kappa_1,\ldots,\kappa_n)+(n)^{-1/2}\sum_{\nu=1}^{n}\sum_{j=1}^{J}\sum_{a'_j}M_t^j(\mathbf{X}_t^j;a'_j,a_j;\kappa_\nu)$$
$$\times\,\alpha[t;a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n-1}(\kappa_1,\ldots,\kappa_{\nu-1},\kappa_{\nu+1},\ldots,\kappa_n),$$
$$n=0,1,2,\ldots,\quad(20)$$

with initial conditions

$$\alpha[0;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)=\chi[a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n),\quad n=0,1,2,\cdots.\quad(21)$$

The integral kernels $M_t^{*j}(\mathbf{X}_t^j;a_j,a'_j;\kappa)$ are given by

$$M_t^{*j}(\mathbf{X}_t^j;a_j,a'_j;\kappa):=\frac{e\boldsymbol{\epsilon}(\kappa)}{(2\pi)^{3/2}[2\omega(\kappa)]^{1/2}}\times\int d^3x\,e^{i\mathbf{k}\mathbf{x}}u^{*j}_{t,a_j}(\mathbf{x}-\mathbf{X}_t^j)\alpha u^j_{t,a'_j}(\mathbf{x}-\mathbf{X}_t^j)$$
$$=e^{i\mathbf{k}\mathbf{X}_t^j}M_t^{*j}(a_j,a'_j;\kappa).\quad(22)$$

The last expression, with $M_t^{*j}(a_j,a'_j;\kappa)$ defined in (14) and subject to P_{12}, has been obtained by a change of the integration variable.

Equation (20) can be derived from the general Schrödinger equation (12) on H by making for $|\alpha(t)\rangle=U(t)|\chi\rangle$ an "ansatz" (18), inserting this into (12), introducing on the right-hand side a complete set of "intermediate states," computing the necessary "elements" of H_t, and carrying out the integration over the many δ functions occurring so. We do not reproduce this tedious derivation. Some details have been given in Ref. 6 in a simpler, but still characteristic case. We have chosen here only an other notation and other normalization conventions, namely those of functional analysis on Fock spaces.[23,24] We emphasize that Eq. (20) is *exact* in the same sense as (1); unfortunately, also the solutions do not exist in this sense. Yet, though exact, Eq. (20) is much more "special" than (1); the right-hand side of (20) namely, *defines* only the one term $H_{t,1,\ldots,1}$ of the sum (17) referring to one-electron atoms. We know of no convenient way to express this *restricted*, but exact, theory in the Heisenberg picture.

The sums over the atoms in Eqs. (20) mirror the Dicke principle as clearly as (1). It is an advantage that the "correspondence principle" now is hidden more and thus cannot be misused so easily. A second gain is the possibility to consider "special solutions," for example, in dependence on various initial conditions; we recall, that (1) has only *one interesting solution* (if at all). A third gain is the *linearization* of the problem achieved so. This allows the use of various "secondary" approximation methods. We emphasize that this linearization is *not* an approximation. The interpretation[25] of a corresponding linearization in Ref. 6 as a first order perturbation approximation must be rejected.

5. EXISTENCE OF FINITE SOLUTIONS UNDER CONDITIONS P_1-P_{12} AND P'_{13}

For technical reasons we prove first the existence theorem in the more special case defined by P_1–P_{12} and P'_{13}.

Postulate P_{12} of all atoms having a finite number q_j of levels is an "important state hypothesis" in the sense of the systematic Weisskopf–Wigner approximation scheme.[17] We "mathematize" it in the following way: Let Q_j denote the finite set of quantum numbers a_j of the q_j states $u^j_{t,a_j}(\mathbf{x}^j)$ of A^j which by some physical arguments are "important" with respect to some desired purposes. In simple cases we could take only two levels for A^j, for example. Generalizing Refs. 6 and 17 we assume that all amplitudes $\alpha[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)$ vanish which for at least one A^j contain the index a_j of an "unimportant," "omitted" level of A^j, and we ignore in (20) the equations for these amplitudes. Equations (20) thus are written down only for the finite set $Q:=Q_1\times\cdots\times Q_J$ of all sets $a:=(a_1,\ldots,a_J)$ of indices of "important" levels. We so get a new, "reduced" set of equations on the Hilbert space

$$\mathcal{S}_Q:=\bigoplus_{a\in Q}\mathcal{F}^a=\bigoplus_{a\in Q}\bigoplus_{n=0}^{\infty}H^a_n,\quad(23)$$

where $\mathcal{F}^a:=\mathcal{F}^{(a_1,\ldots,a_J)}$ is the Fock space of photon states "with A^1 in the state $u^1_{t,a_1}(\mathbf{x}^1)$, A^2 in $u^2_{t,a_2}(\mathbf{x}^2),\ldots$, and A^J in $u^J_{t,a_J}(\mathbf{x}^J)$." $\mathcal{F}^a$ is the direct sum of the Hilbert spaces H^a_n of all states of precisely n photons, $n=0,1,2,\ldots$, "with A^1 in $u^1_{t,a_1}(\mathbf{x}^1),\ldots$, and A^J in $u^J_{t,a_J}(\mathbf{x}^J)$." $\mathcal{S}_Q$ is a subpsace of $\mathcal{S}$; $\mathcal{S}$ namely is the direct sum of *all* $\mathcal{F}^a$.

It is most likely, but not yet proven, that $\mu>0$ alone is sufficient to guarantee on $\mathcal{S}_Q$ the existence of a unitary time evolution operator $U_Q(t)$, defined by the "reduced" equations. To get proven results, we make the auxiliary assumption P'_{13} that all atoms are stationary $[E^j_{t,a_j}=E^j_{a_j},\ u^j_{t,a_j}(\mathbf{x}^j)=u^j_{a_j}(\mathbf{x}^j)]$ and move with the same velocity $\mathbf{v}$. This includes the important case $\mathbf{v}=0$ of all atoms at rest, the only case that has been considered[5–9,12,13,17] up to now. So we have now

$$\mathbf{X}_t^j=\mathbf{X}_0^j+\mathbf{v}t.\quad(24)$$

For the amplitudes of the reduced equations we make an ansatz:

$$\alpha[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)=\exp[-it\mathbf{v}(\mathbf{k}_1+\cdots+\mathbf{k}_n)]\ \beta[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n). \tag{25}$$

This is a unitary transformation on $\mathcal{S}_Q$. Inserting (25) into (20) we find that (25) yields a solution of (20), as reduced above, if the β's satisfy the equations

$$\begin{aligned}
i\frac{d}{dt}\beta[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n) &= [(\omega(\kappa_1)-\mathbf{v}\mathbf{k}_1)+\cdots+(\omega(\kappa_n)-\mathbf{v}\mathbf{k}_n)+E^1_{a_1}+\cdots+E^J_{a_J}]\ \beta[t;a_1,\ldots,a_J]_n(\kappa_1,\ldots,\kappa_n)\\
&+(n+1)^{1/2}\sum_{j=1}^{J}\sum_{a'_j\in Q_j}\int d^3\kappa M^{*j}(\mathbf{X}^j_0;a_j,a'_j;\kappa)\\
&\times\beta[t;a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n+1}(\kappa,\kappa_1,\ldots,\kappa_n)\\
&+(n)^{-1/2}\sum_{j=1}^{J}\sum_{a'_j\in Q_j}\sum_{\nu=1}^{n}M^j(\mathbf{X}^j_0;a'_j,a_j;\kappa_\nu)\\
&\times\beta[t;a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n-1}(\kappa_1,\ldots,\kappa_{\nu-1},\kappa_{\nu+1},\ldots,\kappa_n).
\end{aligned} \tag{26}$$

The initial state can be any state of $\mathcal{S}_Q$. The point of (25) is the elimination of the explicit dependence on t from the operator H_Q which defines the time evolution in terms of the β's. H_Q is defined by the right-hand side side of (26).

As (25) is a unitary transformation on $\mathcal{S}_Q$ our statement on the existnece of finite solutions under the present conditions can be proved by showing that H_Q is *self-adjoint* on $\mathcal{S}_Q$. For this we follow, with a slight modification, the proof idea of Sec. 4 of Ref. 17: We show that H_Q is the sum of a self-adjoint operator H^0_Q, defined by the first term of the right-hand side of (26), and a *symmetric* "interaction" operator H'_Q, defined by the other terms, which is *bounded relative* to H^0_Q. The theorem of Rellich[26] then guarantees the selfadjointness of $H_Q=H^0_Q+H'_Q$. H_Q generates, therefore, for any $t<\infty$ a *unitary* time evolution operator $U_Q(t)$ which together with (25) guarantees finite transition probability amplitudes and finite transition probabilities normed to 1.

H^0_Q is self-adjoint because it is a direct sum of self-adjoint operators of multiplication, on H^a_n, with the real function $[(\omega(\kappa_1)-\mathbf{k}_1\mathbf{v})+\cdots+(\omega(\kappa_n)-\mathbf{k}_n\mathbf{v})+E^1_{a_1}+\cdots+E^J_{a_J}]$. We note the inequality

$$\begin{aligned}
&(\omega(\kappa_1)-\mathbf{k}_1\mathbf{v})+\cdots+(\omega(\kappa_n)-\mathbf{k}_n\mathbf{v})\\
&\quad\geq(\omega(\kappa_1)-|\mathbf{k}_1|\ |\mathbf{v}|)+\cdots+(\omega(\kappa_n)-|\mathbf{k}_n|\ |\mathbf{v}|)\\
&\quad\geq n\mu(1-\mathbf{v}^2)^{1/2}
\end{aligned} \tag{27}$$

where $\mu(1-\mathbf{v}^2)^{1/2}$ is the minimum of the function $(\mathbf{k}^2+\mu^2)^{1/2}-|\mathbf{v}|\ |\mathbf{k}|$ at $|\mathbf{k}|\geq 0$. This means that the dense domain $D(H^0_Q)$ of H^0_Q is contained in the domain $D(\mathfrak{N})$ of the direct sum $\mathfrak{N}$ of the operators of multiplication, on H^a_n, with the number $n, n=0,1,2,\ldots$. Therefore, and because $E^1_{a_1}+\cdots+E^J_{a_J}>0$, we have for any $|\beta\rangle\in D(H^0_Q)$

$$\|H^0_Q|\beta\rangle\|\geq\mu(1-\mathbf{v}^2)^{1/2}\|\mathfrak{N}|\beta\rangle\|, \tag{28}$$

with $\|\cdots\|$ denoting the norm on $\mathcal{S}_Q$. As in Ref. 17 we can find for any number $\epsilon>0$ a number $b_\epsilon>0$ so that for any $|\beta\rangle\in D(\mathfrak{N})$ we get

$$\|(\mathfrak{N}+1)^{1/2}|\beta\rangle\|\leq\epsilon\|\mathfrak{N}|\beta\rangle\|+b_\epsilon\||\beta\rangle\|. \tag{29}$$

$(\mathfrak{N}+1)^{1/2}$ is the direct sum of the operators of multiplication, on H^a_n, with the number $(n+1)^{1/2}$. Of course, $D((\mathfrak{N}+1)^{1/2})\supset D(\mathfrak{N}+1)=D(\mathfrak{N})$. Combining (28) and (29) we see that for any $\epsilon>0$ we can find a $b_\epsilon>0$ so that for any $|\beta\rangle\in D(H^0_Q)$ we have

$$\|(\mathfrak{N}+1)^{1/2}|\beta\rangle\|\leq\frac{\epsilon}{\mu(1-\mathbf{v}^2)^{1/2}}\|H^0_Q|\beta\rangle\|+b_\epsilon\||\beta\rangle\|. \tag{30}$$

We show below that the interaction operator H'_Q satisfies for any $|\beta\rangle\in D((\mathfrak{N}+1)^{1/2})$ an inequality

$$\|H'_Q|\beta\rangle\|\leq 2qm\|(\mathfrak{N}+1)^{1/2}|\beta\rangle\| \tag{31}$$

with finite constants q, m. This means that H'_Q is defined at least on the dense domain $D((\mathfrak{N}+1)^{1/2})$, i.e., $D(H'_Q)\supseteq D((\mathfrak{N}+1)^{1/2})$. Equation (28) implies $D(H^0_Q)\subseteq D(\mathfrak{N})$ and because of $D(\mathfrak{N})\subset D((\mathfrak{N}+1)^{1/2})$ we get (i) $D(H^0_Q)\subset D(H'_Q)$. It poses no problem to verify (ii) that H'_Q is symmetric, i.e., that $\langle\beta|H'_Q|\beta\rangle$ is real for any $|\beta\rangle\in D(H'_Q)$. Combining (30) and (31), we get for any $|\beta\rangle\in D(H^0_Q)$

$$\|H'_Q|\beta\rangle\|\leq\frac{2qm\epsilon}{\mu(1-\mathbf{v}^2)^{1/2}}\|H^0_Q|\beta\rangle\|+b_\epsilon 2qm\||\beta\rangle\|. \tag{32}$$

Choosing (iii) ϵ smaller than $\mu(1-\mathbf{v}^2)^{1/2}(2qm)^{-1}$, we meet the conditions (i)–(iii) of the theorem of Rellich and $H_Q=H^0_Q+H'_Q$ is self-adjoint.

We must prove, finally, the inequality (31). By the definition of H'_Q we have

$$\begin{aligned}
\|H'_Q|\beta\rangle\|^2=\sum_{a\in Q}\sum_{n=0}^{\infty}\int d^3\kappa_1\cdots\int d^3\kappa_n\\
\times\Bigg|\sum_{j=1}^{J}\sum_{a'_j\in Q_j}\Big\{(n+1)^{1/2}\int d^3\kappa M^{*j}(\mathbf{X}^j_0;a_j,a'_j;\kappa)\\
\times\beta[a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n+1}\\
\times(\kappa,\kappa_1,\ldots,\kappa_n)+(n)^{-1/2}\sum_{\nu=1}^{n}M^j(\mathbf{X}^j_0;a'_j,a_j;\kappa_\nu)\\
\times\beta[a_1,\ldots,a_{j-1},a'_j,a_{j+1},\ldots,a_J]_{n-1}\\
\times(\kappa_1,\ldots,\kappa_{\nu-1},\kappa_{\nu+1},\ldots,\kappa_n)\Big\}\Bigg|^2.
\end{aligned} \tag{33}$$

For given a and n the integral $\int d^3\kappa_1\cdots\int d^3\kappa_n|\cdots|^2$ defines the square of the norm $\|\cdots\|_n$ of a vector of the

Hilbert space H_n^a, this vector being the sum over several other such vectors, $\alpha_i+\beta_i$, namely the terms of the sum over j and a_j'. Therefore, we can use the triangle relation

$$\|\sum_i(\alpha_i+\beta_i)\|_n^2 \le \left\{\sum_i (\|\alpha_i\|_n+\|\beta_i\|_n)\right\}^2 \tag{34}$$

with respect to the sum over j and a_j'. This yields

$$\begin{aligned}\|H_Q'|\beta\rangle\|^2 \le \sum_{a\in Q}\sum_{n=0}^{\infty}\Big\{\sum_{j=1}^{J}\sum_{a_j'\in Q_j} \big(\|(n+1)^{1/2}\int d^2\kappa M^{*j}(\mathbf{X}_0^j;a_j,a_j';\kappa)\\ \times\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_{n+1}\\ \times(\kappa,\bigstar,\ldots,\bigstar)\|_n + \|(n)^{1/2}M^j(\mathbf{X}_0^j;a_j',a_j;\bigstar_1)\\ \times\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_{n-1}\\ \times(\bigstar_2,\ldots,\bigstar_n)\|_n\big)\Big\}^2.\end{aligned} \tag{35}$$

The stars indicate the integration variables of the norm $\|\cdots\|_n$. In the last term we have taken into account that the sum over ν in (33), after the use of (34), yields only a factor n. Let $m_j(a_j,a_j')=m_j(a_j',a_j)$ denote the norm of $M^j(\mathbf{X}_0^j;a_j',a_j;\kappa)$ with respect to κ, which is finite by P_{12}. Applying now the Schwarz inequality to the integration over κ in the first term and using the fact that the square of the norm in the second term factors, we get

$$\begin{aligned}\|H_Q'|\beta\rangle\|^2 \le \sum_{a\in Q}\sum_{n=0}^{\infty}\Big\{\sum_{j=1}^{J}\sum_{a_j'\in Q_j}[m_j(a_j,a_j')\|(n+1)^{1/2}\\ \times\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_{n+1}\\ \times(\bigstar,\ldots,\bigstar)\|_{n+1} + m_j(a_j,a_j')\|(n)^{1/2}\\ \times\beta[(a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_{n-1}\\ \times(\bigstar,\ldots,\bigstar)\|_{n-1}]\Big\}^2.\end{aligned} \tag{36}$$

By a simple renaming $n+1\to n$ and $n-1\to n$ of the terms of the sum over n and the use of $\|(n)^{1/2}\beta\cdots\|_n = (n)^{1/2}\|\beta\cdots\|_n < (n+1)^{1/2}\|\beta_n\cdots\|_n = \|(n+1)^{1/2}\beta_n\cdots\|_n$, we obtain

$$\begin{aligned}\|H_Q'|\beta\rangle\|^2 \le 4\sum_{a\in Q}\sum_{n=0}^{\infty}\Big\{\sum_{j=1}^{J}\sum_{a_j'\in Q_j} m_j(a_j,a_j')\|(n+1)^{1/2}\\ \times\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_n\\ (\bigstar,\ldots,\bigstar)\|_n\Big\}^2.\end{aligned} \tag{37}$$

Using now $|\sum_i c_i d_i| \le (\sum_i |c_i|^2)^{1/2}(\sum_i |d_i|^2)^{1/2}$ with respect to the sum over j and a_j', we get

$$\|H_Q'|\beta\rangle\|^2 \le 4\sum_{a\in Q}\sum_{n=0}^{\infty}\Big\{\sum_{j=1}^{J}\sum_{a_j'\in Q_j} m_j(a_j,a_j')^2\Big\}\Big\{\sum_{j=1}^{J}\sum_{a_j'\in Q_j}\|(n+1)^{1/2}\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_n(\bigstar,\ldots,\bigstar)\|_n^2\Big\}. \tag{38}$$

Let m denote the finite maximum of all occuring $m_j(a_j,a_j')$, let q_j denote the number of levels of A^j, i.e., the number elements in Q_j, and put $q:=\sum_{j=1}^J q_j$. Then the first term in (38) is not greater than m^2q. The sum over $a\in Q$, of course, equals the multiple sum over all $a_j\in Q_j$, and after a suitable change of the order of summation we find

$$\begin{aligned}\|H_Q'|\beta\rangle\|^2 \le 4\,qm^2\sum_{j=1}^{J}\sum_{a_j\in Q_j}\Big\{\sum_{n=0}^{\infty}\sum_{a_1\in Q_1}\cdots\sum_{a_{j-1}\in Q_{j-1}}\sum_{a_j'\in Q_j}\sum_{a_{j+1}\in Q_{j+1}}\cdots\sum_{a_J\in Q_J}\|(n+1)^{1/2}\\ \times\beta[a_1,\ldots,a_{j-1},a_j',a_{j+1},\ldots,a_J]_n(\bigstar,\ldots,\bigstar)\|_n^2\Big\}.\end{aligned} \tag{39}$$

The sum in the bracket is equal, for any value of j, to $\|(\mathfrak{N}+1)^{1/2}|\beta\rangle\|^2$. The sum in front of the bracket therefore yields a factor q again. So we get $\|H_Q'|\beta\rangle\|^2 \le 4q^2m^2\|(\mathfrak{N}+1)^{1/2}|\beta\rangle\|^2$, which is identical to (31) and implies $\mathcal{D}(H_Q')\supseteq\mathcal{D}(\mathfrak{N}+1)^{1/2}$.

6. EXISTENCE OF UNITARY TIME EVOLUTION UNDER CONDITIONS P_1-P_{13}

We now relax condition $\mu>0$ to $\mu\ge 0$, the condition of a *common* velocity $\mathbf{v}$ of all atoms to arbitrary motion, and allow the atomic potentials $V_i^j(\mathbf{x}^j)$ to depend on t. Instead we assume that there exists a number $N<\infty$ so that never more than N photons exist.

Together with the assumption of any atom having a finite number of levels this also is an important state hypothesis.[21] It will be convenient, however, and for reasons of practicability it is even necessary, to consider also theories defined by stronger, more restrictive important state hypotheses. For example we might wish to treat all atoms as "two-level atoms" and assume, *ab initio*, that for any emitted (absorbed) photon one atom must make a transition from the upper (lower) to the lower (upper) level. This still does not yet prejudice the typical effects of quantum optics. We now prove the existence theorem for *any* such assumption.

The above important state hypotheses can be math-

ematized in the following way: Let I denote *any* subset of the finite set $Q_1 \times \cdots \times Q_J \times N^*$ with $N^* := \{0, 1, 2, \ldots, N\}$. Following Refs. 6, 17, we assume

$$\alpha[t; a_1, \ldots, a_J]_n(\kappa_1, \ldots, \kappa_n) \equiv 0 \quad \text{for } (a_1, \ldots, a_J, n) \notin I \tag{40}$$

and ignore the equations of motion for these amplitudes. Physically this implies the *ab initio* hypothesis that transitions are *possible only between the states* of $R + A_1 + \cdots + A_J$ "with n photons *and* A_1 in the state $u^1_{t,a_1}(\mathbf{x}^1)$, A_2 in $u^2_{t,a_2}(\mathbf{x}^2)$, ..., and A_J in $u^J_{t,a_J}(\mathbf{x}^J)$" with $(a_1, \ldots, a_J, n) \in I$. The "reduction" of Eqs. (20) by (40) leads to the following equations of motion:

$$\begin{aligned} i\frac{d}{dt}&\alpha[t; a_1, \ldots, a_J]_n(\kappa_1, \ldots, \kappa_n) \\ &= (\omega(\kappa_1) + \cdots + \omega(\kappa_n) + E^1_{t,a_1} + \cdots + E^J_{t,a_J}) \\ &\quad \times \alpha[t; a_1, \ldots, a_J]_n(\kappa_1, \ldots, \kappa_n) \\ &\quad + (n+1)^{1/2} \sum_{j=1}^{J} \sum_{(a'_j)}^{+} \int d^3\kappa M^{*j}_t(\mathbf{X}^j_t; a_j, a'_j; \kappa) \\ &\quad \times \alpha[t; a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J]_{n+1}(\kappa, \kappa_1, \ldots, \kappa_n) \\ &\quad + n^{-1/2} \sum_{j=1}^{J} \sum_{(a'_j)}^{-} \sum_{\nu=1}^{n} M^j_t(\mathbf{X}^j_t; a'_j, a_j; \kappa_\nu) \\ &\quad \times \alpha[t; a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J]_{n-1} \\ &\quad \times (\kappa_1, \ldots, \kappa_{\nu-1}, \kappa_{\nu+1}, \ldots, \kappa_n), \end{aligned} \tag{41}$$

$$\alpha[0, a_1, \ldots, a_J]_n(\kappa_1, \ldots, \kappa_n) = \chi[a_1, \ldots, a_J]_n(\kappa_1, \ldots, \kappa_n). \tag{42}$$

The sums $\sum^{\pm}$ over (a'_j) are to be understood to cover all terms where, given $(a_1, \ldots, a_J; n) \in I$, $(a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J; n \pm 1)$ is also an element of I. Equations (41) are defined on the Hilbert space

$$\mathcal{S}_I := \bigoplus_{(a_1, \ldots, a_J, n) \in I} \mathcal{H}_n^{(a_1, \ldots, a_J)}. \tag{43}$$

We show that they define on $\mathcal{S}_I$ a unitary time evolution operator $U_I(t)$. In particular, for $I = Q_1 \times \cdots \times Q_J \times N^*$ we consider the theory of J atoms with a finite number of levels interacting with up to N photons. Another choice of I leads to the theory considered in Ref. 6; we prove here its existence even in the case when the atoms move in an arbitrary, yet *prescribed* way, and are subject to arbitrary external perturbations.

Let $H'_{t,I}$ denote the time-dependent interaction operator defined in (41) by the sums over j. Then $\|H'_{t,I}|\alpha\rangle\|$, with $\|\cdots\|$ denoting now the norm on $\mathcal{S}_I$, is given by the right-hand side of (33), if the $\mathbf{X}^j_0$ are replaced by $\mathbf{X}^j_t$, the M's are allowed to depend on t by the $\mathbf{v}$'s, if the sum over $a \in Q$ and n is replaced by the sum over $(a_1, \ldots, a_J, n) \in I$, and if finally the sum over $a'_j \in Q_j$ is replaced by the sums occurring in (41). Since all these sums contain a finite number of finite terms only, the right-hand side of (33) now remains finite for *any* $|\alpha\rangle \in \mathcal{S}_I$. Therefore, $H'_{t,I}$ is *defined everywhere* on $\mathcal{S}_I$. It poses no problem to verify that $\langle\alpha|H'_{t,I}|\alpha\rangle$ is real for any $|\alpha\rangle \in \mathcal{S}_I$. Therefore, $H'_{t,I}$ is bounded and self-adjoint for every t, by well-known theorems of functional analysis.

It will be convenient to know in addition that a time independent bound B of $H'_{t,I}$ exists, i.e., that

$$\|H'_{t,I}|\alpha\| \leq B \||\alpha\rangle\| \tag{44}$$

for any $|\alpha\rangle \in \mathcal{S}_I$, and to learn something about the magnitude of this bound.

For this we look at (39), the right-hand side being modified as described above. We find easily that the considerations leading to (36) remain valid also under the present conditions. Since by P_{12} the norms of the time dependent M's are uniformly bounded in t, we can use $m_{t,j}(a_j, a'_j) \leq m$. Using also $(n+1)^{1/2} < (N+1)^{1/2}$, $(n)^{1/2} < (n+1)^{1/2}$, we obtain from (36)

$$\begin{aligned} \|H'_{t,I}|\alpha\rangle\|^2 \leq m^2(N+1) \sum_{(a_1, \ldots, a_J, n) \in I} \Bigg\{ &\sum_{j=1}^{J} \sum_{(a'_j)}^{+} \|\alpha[a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J]_{n+1}(\bigstar, \ldots, \bigstar)\|_{n+1} \\ &+ \sum_{j=1}^{J} \sum_{(a'_j)}^{-} \|\alpha[a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J]_{n-1}(\bigstar, \ldots, \bigstar)\|_{n-1} \Bigg\}^2. \end{aligned} \tag{45}$$

Using now for $A \geq 0, B \geq 0$ the trivial estimate $(A+B)^2 = A^2 + B^2 + 2AB \leq 3A^2 + 3B^2$, and with $A_j \geq 0$ and $A_{\max}$ denoting the maximal A_j the trivial estimate

$$\left(\sum_{j=1}^{J} A_j\right)^2 = \sum_{j=1}^{J} \sum_{j'=1}^{J} A_j A_{j'} \leq J^2 A^2_{\max} \leq J^2 \sum_{j=1}^{J} A^2_j, \tag{46}$$

we get

$$\|H'_{t,I}|\alpha\rangle\|^2 \leq 3(N+1)m^2J^2 \sum_{j=1}^{J} \Bigg\{ \sum_{(a_1, \ldots, a_J, n) \in I} \left(\sum_{(a'_j)}^{+} \|\cdots\|^2_{n+1} + \sum_{(a'_j)}^{-} \|\cdots\|^2_{n-1} \right) \Bigg\}, \tag{47}$$

with the same terms in the norms $\|\cdots\|_{n\pm1}$ as in (45). The double sums in the bracket go over all pairs $(a_1, \ldots, a_J, n) \in I$ *and* $(a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J, n \pm 1) \in I$. The set of these pairs is identical with the set "all $(a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_J, n \pm 1) \in I$ with $(a_1, \ldots, a_J, n)$ *also* $\in I$." By dropping the second condition, i.e., by adding some nonnegative terms to the right-hand side, the double sums can be completed to the square of the

norm of $|\alpha\rangle$. The sum over j then yields a factor J, and finally we get (44) with

$$B \leq m[6(N+1)J^3]^{1/2}. \tag{48}$$

Checking critically this derivation, we see that only (46) might be poor by a factor $\sim J$; in realistic cases the completion of (47) to the norm yields a factor of magnitude 1 only. Therefore, B is proportional, in essence, to the number J of active atoms, which is plausible. Note that m contains a factor e^2, but the smallness of this factor is compensated by far by the other factors which typically are of the order $J \sim 10^{20}$ (20 cm^3 pink ruby) and $N \sim 10^6$ (laser threshold) to $N \sim 10^{20}$ (giant pulse).

The first term on the right-hand side of (41) defines, as in Sec. 4, on $\mathcal{S}_I$ a self-adjoint, unbounded, yet time dependent operator $H^0_{t,I}$. Since it is a direct sum of operators of multiplication with time-dependent real functions, the solution $U^0_I(t)$ of the equation

$$i\frac{d}{dt}U^0_I(t) = H^0_{t,I}U^0_I(t), \quad U^0_I(0) = 1 \tag{49}$$

is given by [cf. P_3 and Eq. (3)]

$$U^0_I(t) := \exp(-i\int_0^t dt' H^0_{t',I}), \tag{50}$$

and is unitary on $\mathcal{S}_I$; $U^0_I(t)$ again is a direct sum of operators of multiplication with complex numbers of modulus 1.

Since $H^0_{t,I}$ is self-adjoint and $H'_{t,I}$ is self-adjoint and bounded, the total Hamiltonian $H^0_{t,I}+H'_{t,I}$ is also self-adjoint for any t.

We are ready now to prove the existence of a unitary time evolution operator defined by (41). There are theorems[27] giving sufficient conditions that the equation

$$i\frac{d}{dt}U_I(t) = (H^0_{t,I}+H'_{t,I})U_I(t), \quad U_I(0) = 1 \tag{51}$$

with a time dependent, self-adjoint Hamiltonian has a unique unitary solution. It is not convenient to check whether these conditions are met in our case. Fortunately, we can construct an existence proof by "summing up the time-dependent perturbation series." For this we put $U_I(t) = U^0_I(t)V_I(t)$ and obtain for $V_I(t)$ the integral equation

$$V_I(t) = 1 - i\int_0^t dt' H_I(t')V_I(t') \tag{52}$$

with $H_I(t) = U^{*0}_I(t)H'_{t,I}U^0_I(t)$. Since $U^0_I(t)$ is unitary on $\mathcal{S}_I$, $H_I(t)$ is also bounded with bound B. The usual Dyson-expansion

$$V_I(t) = 1 - i\int_0^t dt' H_I(t') + (-i)^2\int_0^t dt'\int_0^{t'} dt'' H_I(t')H_I(t'') + \cdots \tag{53}$$

then converges by the norm, the ρth term being bounded by $(Bt)^\rho/\rho!$. (53) therefore provides the solution of (52) which is equivalent to the solution of (51), (41). Unitarity and uniqueness follow from standard arguments.

The maximal difference between an exact solution of (41), $|\alpha(t)\rangle$, and the approximate expression $|\alpha_\rho(t)\rangle$ obtained for it in ρth order perturbation theory is bounded by

$$\| |\alpha(t)\rangle - \alpha_\rho(t)\rangle\| \leq e^{Bt} - \sum_{s=0}^{\rho}\frac{(Bt)^s}{s!}. \tag{54}$$

For given $t < \infty$ this can be made smaller than any $\epsilon > 0$, but at the price of sufficiently large orders ρ. The necessary minimal order ρ_ϵ depends on B. A glance on (48) and its discussion shows that time-dependent perturbation theory cannot be used practically in typical quantum optics problem (see Ref. 21 for problems of principle): If for given t the order ρ_ϵ guarantees an error smaller than ϵ in a one-atom problem, we must sum the perturbation series to order $\sim J \cdot \rho_\epsilon$ to get the same accuracy in a problem with J atoms. The situation is *not* better in the Heisenberg picture.

7. CONCLUSION AND OUTLOOK

The proofs in Secs. 5 and 6 are certainly and necessarily somewhat "mathematical"; the motivation behind all this is purely physical, however. We mentioned in Sec. 1, and found still more convincing arguments in Sec. 6, that the well-understood and simple order hierarchy of perturbation theory cannot be used in quantum optics. Therefore, it is *necessary* to consider alternatives, maybe the one presented. Its "order hierarchy" can be defined (cf. Ref. 17) by any sequence

$$I^0 \subset I^1 \subset \cdots \subseteq I^\infty \tag{55}$$

of the index sets I introduced in (40), which "tend" to the index set $I^\infty := Q_1^\infty \times \cdots \times Q_J^\infty \times \{0,1,2,\cdots\}$ of the "exact" theory of Sec. 4. Q_j^∞ denotes the set of indices a_j of *all* eigenstates of A_j. The sequence of unitary quantum theories corresponding to $I^0, I^1, \ldots$, all of which *exist with certainty* if all I's are finite or meet the conditions of Sec. 5, then *defines* the "exact" theory as its limit $I^\rho \to I^\infty$. However, in the present case this potential limit theory is of questionable value: We saw in the discussion of P_{11} that the inclusion of too many states of each single atom introduces effects which because of the Pauli principle cannot occur in reality. The limit theory therefore contains in any case some unrealistic idealizations which are the actual cause of the "infinities" that occur in the "exact" theory.

The familiar divergencies of QED appear in a systematic perturbation expansion already in the order $m \geq 2$ in e. But they occur always in sums over "intermediate" atomic states which because of diverging diameters cannot occur in reality. *In a unitary theory the "intermediate" states must also be possible as final or initial states.* Therefore, the main clauses of our peace treaty should be acceptable by "physical arguments": We postpone the calculation of the influence of higher atomic states (which never exist in reality if they are high enough!) upon the transitions between the lower states by packing them into the "higher orders" in the sense of (55), but in the lowest orders $I^0, I^1, I^2, \ldots$ we include already effects which in a systematic expansion in e are of high, even infinite order. The "divergence problems" are thereby postponed to

the "convergence" of the sequence of Weisskopf–Wigner theories defined by $I^0, I^1, I^2, \ldots$. For this we get a strictly unitary time evolution operator in any finite step. We leave it to the reader to decide whether the *"total" exclusion of atomic states which never can occur in reality* is a "heuristic ad hoc principle," or to see in it a first step to exclude *unrealistic idealizations* contained in (1).

At present we aim at another, much more modest goal, namely a practical working base for quantum optics. This will give us also an example of a "reasonable choice" $I^0, I^1, \ldots$. *It appears namely that one can get reliable results already by the following choice of the lowest order* I^0[6]: In quantum optics it appears "obvious" to consider "two-level atoms," and in view of validity of the Bohr relations $\omega = E_2 - E_1$ it is likewise suggestive to assume, as *the principle defining the lowest order, and thus subject to higher order corrections*, that one photon is emitted (absorbed) whenever one atom makes a transition from the upper (lower) to the lower (upper) level. Then I^0 is given by

$$I^0 := \{(a_1, \ldots, a_J, n) : n + m_{\text{exc}} = N + M_{\text{exc}}\}, \tag{56}$$

where m_{exc} is the number of times the index of the upper level occurs among $a_1, \ldots, a_J$. M_{exc} is the number of excited atoms and N the number of photons in the initial state.

"First corrections" to this theory can be obtained from the higher order theory defined by the choice $I^1 \supset I^0$

$$I^1 := \{(a_1, \ldots, a_J, n) : a_j = 1, 2 \quad \text{for } j = 1, \ldots, J; n = 0, 1, 2, \ldots\}. \tag{57}$$

Each atom has now still "two levels," but in addition to the above "ordinary Bohr transitions," each A^j is now allowed to make "virtual" transitions from *a given level to itself* and virtual transitions from the upper (lower) to the lower (upper) level *under absorption (emission)* of a photon. One can show that the "lower" theory I^0 is obtained from the higher by the use of the familiar rotating wave approximation. One can further show that the "first corrections" to the theory I^0 amount to *finite* level shifts of the order of magnitude of Lamb shifts, and that these corrections have little influence upon those mechanisms which possibly build up the interesting many photon collective phenomena. For one atom the corresponding results are already published.[28]

Our hope for reliable results in the lowest order I^0 rests also upon the experience with numerous Weisskopf–Wigner type theories practically all of which are of type I^0, and thus provides examples for the present existence theorem:

The case $J = N = 1$, atom at rest, has been reanalyzed recently.[21] It contains Källén's version[29] of the Weisskopf–Wigner theory of natural line width which agrees with experience. We shall show elsewhere that motion of the atoms, as introduced here, "shifts" the natural lines in agreement with the nonrelativistic Doppler theory. Therefore, reliable and realistic results can also be expected if a finite temperature of the active atoms is introduced later by allowing the atoms to move, either like molecules in a gas, or like host atoms in a crystal subject to various lattice vibrations. We saw (Sec. 6) that this can be achieved without any additional mathematical expenses.

The case $J > 1, N = 1$, atoms at rest, leads in a secondary approximation to photon inprisonment and superradiance effects.[30] One can show[31] that this theory, without the mentioned secondary approximation, contains *all* time-of-flight effects of the photon which must be expected in a causal theory. So we can say that even the crudest, "lowest" order I^0 of our approximations meets our demand 3 of Sec. 1 (at least with respect to the photon, cf. Ref. 17). Cases $J \gg 1, N \gg 1$, atoms at rest[6,12] lead to the formation of a directed beam of light capable of interference.[7] All this, previously computed in "formal" theories and now put on a firm base, should not be accidental.

Our existence theorem covers still many other cases of interest: few photons may impinge upon many excited atoms (all the cases $M_{\text{exc}} \lesseqgtr J$ are possible here) and thus "stimulate" their emission process, many photons may impinge upon few atoms initially excited or not, etc. Much work remains to be done along these lines and with respect to more realistic (temperature etc.), yet still "microscopic" and causal approaches to laser activity and related phenomena. We hope that our existence theorem is of some help for this.

We note finally that the transition to a description by density matrices makes no problems as the present "pure" theory is formulated on separable Hilbert spaces.

[1] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Field Theory* (McGraw-Hill, New York, 1965).
[2] M. von Laue, Ann. Phys. (Leipzig) **20**, 365 (1906); **23**, 1 (1907); **23**, 795 (1907).
[3] V. Weisskopf, Ann. Phys. (Leipzig) **9**, 23 (1931).
[4] R. H. Dicke, Phys. Rev. **93**, 99 (1955).
[5] F. Schwabl and W. Thirring, Ergeb. Exakt. Naturw. **36**, 219 (1964).
[6] V. Ernst and P. Stehle, Phys. Rev. **176**, 1456 (1968).
[7] V. Ernst, Z. Physik **229**, 432 (1969).
[8] H. Steudel, Ann. Phys. (Leipzig) **27**, 57 (1971).
[9] P. Glas, Ann. Phys. (Leipzig) **29**, 121 (1973).
[10] P. Schwendimann, Z. Physik **265**, 267 (1973); Z. Physik B **21**, 301 (1975).
[11] R. Jodoin and L. Mandel, Phys. Rev. A **9**, 873 (1974).
[12] R. J. Glauber and F. Haake, Phys. Rev. A **13**, 357 (1976).
[13] M. Gronchi and L. A. Lugiato, Phys. Rev. A **13**, 830 (1976).
[14] L. Allen, in *Coherence and Quantum Optics*, edited by L. Mandel and E. Wolf (Plenum, New York and London, 1973).
[15] R. L. Pfleegor and L. Mandel, Phys. Rev. **159**, 1084 (1967).
[16] L. Csillag, M. Jánossy, and Zs. Náray, Acta Phys. Hung. **32**, 275 (1972).
[17] E. Grimm and V. Ernst, J. Phys. A: Math. Nucl. Gen. **7**, 1664 (1974).
[18] H. Haken, in *Encyclopaedia of Physics, Vol. XXV/2C* (Springer, Berlin, Heidelberg and New York, 1970).
[19] G. G. Emch, *Algebraic Methods in Statistical Mechanics and*

Quantum Field Theory (Wiley, New York, London, Sidney, and Toronto, 1970).
[20]W. Weizel, *Lehrbuch der Theoretischen Physik, Vol. II* (Springer, Berlin, Göttingen, Heidelberg, 1958), p. 851.
[21]E. Grimm and V. Ernst, Z. Phys. A **274**, 293 (1975).
[22]N. I. Achieser and I. M. Glasmann, *Theorie der linearen Operatoren im Hilbertraum*, 5.Aufl. (Akademie-Verlag, Berlin, 1968).
[23]K. O. Friedrichs, *Mathematical Aspects of the Quantum Theory of Fields* (Interscience, New York, 1953).
[24]J. M. Cook, J. Math. Phys. **2**, 33–45 (1961).
[25]R. H. Picard and C. R. Willis, Phys. Rev. A **8**, 1536–61 (1973).
[26]T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[27]K. Yosida, J. Fac. Sci. Univ. Tokyo Sect. 1, **9**, Part 5, 397–402 (1963).
[28]V. Ernst, Z. Phys. B **23**, 103 (1976).
[29]G. Källén, in *Encyclopaedia of Physics, Vol. V/1* (Springer-Verlag, Berlin, 1958), p. 274.
[30]V. Ernst, Z. Phys. **218**, 111 (1969).
[31]J. Mühlstein and V. Ernst, Ann. Phys. (Leipzig) **34**, 105 (1977).

A multiple-scales space–time analysis of a randomly perturbed one-dimensional wave equation[a)]

Werner E. Kohler

Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
(Received 23 March 1977)

An initial value problem for one-dimensional wave propagation is considered; the medium is assumed to be randomly perturbed as a function of both space and time. The stochastic perturbation theory of Papanicolaou and Keller [SIAM J. Appl. Math. **21**, 287 (1971)] is applied directly in the space–time regime to derive transport equations for the first and second moments of the solution. These equations are solved in special cases.

I. INTRODUCTION

The problem of characterizing long-range acoustic transmission in the ocean is essentially that of understanding the dynamics of wave propagation in a medium subjected to small random spatial and temporal perturbations. While the ocean problem is further complicated by a deterministic sound-speed profile (forming the SOFAR channel) and randomly irregular boundaries, understanding the effects of the medium itself represents a necessary prerequisite.

In this paper, a *model problem* involving one-dimensional wave propagation is considered. The properties of the medium are assumed to be randomly perturbed in both space and time. The stochastic perturbation theory of Papanicolaou and Keller[1] is applied directly in space–time; transport equations for the first and second moments of the solution emerge as necessary relations for the suppression of secular growth in the two characteristic directions of the unperturbed wave operator.

Dealing with the problem directly in space–time permits us to study the evolution of wavepackets in a spatially and temporally fluctuating environment. Fourier transforms are used, but only after the formal stochastic asymptotic analysis is complete. Only the infinite spatial domain is considered; however, the formalism can be developed as well for the semi-infinite spatial domain. It is hoped that subsequent analytis of this latter problem will provide insights into the nature of the boundary conditions that must accompany the limiting transport equations in cases (like the ocean) where boundaries exist.

In Sec. II the asymptotic formalism is developed while equations for the first moment of the solution are derived in Sec. III. In Sec. IV similar equations are derived for the second moment (i.e., the mutual coherence function). Sec. V deals with a specific example.

[a)]Research supported by the Office of Naval Research under contract No. N00014-76-C-0056. The research reported in this paper was initiated at the 1976 Applied Mathematics Summer Institute; this institute was supported by the Office of Naval Research under contract No. N00014-75-C-0921.

II. DEVELOPMENT OF THE FORMALISM

The initial value problem that we shall consider is the following:

$$\frac{\partial^2}{\partial x^2}u(x,t,\omega,\epsilon)-\frac{\partial}{\partial t}\left(\bar{c}^{-2}(x,t,\omega,\epsilon)\frac{\partial}{\partial t}u(x,t,\omega,\epsilon)\right)=0,\quad -\infty<x<\infty,\quad t>0 \tag{1}$$

$$u(x,0,\omega,\epsilon)=f(x),\quad \frac{\partial u}{\partial t}(x,0,\omega,\epsilon)=g(x),\quad -\infty<x<\infty, \tag{2}$$

where ϵ is a small real parameter and ω is an element of some underlying probability space. We assume that

$$\bar{c}^{-2}(x,t,\omega,\epsilon)=c^{-2}(1+\epsilon\mu(x,t,\omega)), \tag{3}$$

i.e., the sum of a constant and a small randomly fluctuating quantity. The random field μ is assumed to be a zero mean wide-sense stationary function of both space and time; consequently, we have

$$\langle\mu(x,t,\omega)\rangle=0,$$
$$\langle\mu(x,t,\omega)\mu(x',t',\omega)\rangle=R(x-x',t-t'), \tag{4}$$
$$-\infty<x,x'<\infty,\quad 0\le t,t'<\infty,$$

where $\langle\cdot\rangle$ denotes expectation, i.e., integration with respect to the underlying probability measure. We shall further assume that the random field μ is mixing (Ref. 2) in the sense that as the space–time separation of (x_1,t_1) and (x_2,t_2) tends to infinity, the random variables $\mu(x_1,t_1,\omega)$ and $\mu(x_2,t_2,\omega)$ become asymptotically independent.

We are ultimately interested in $\langle u(x,t,\omega,\epsilon)\rangle$ and $\langle u(x,t,\omega,\epsilon)u(x',t',\omega,\epsilon)\rangle$ in the asymptotic limit where $\epsilon\to 0$ but where the space–time propagation paths (i.e., distances along the characteristics) tend to infinity. Prior work (Refs. 1, 2) has shown that since $\langle\mu\rangle=0$, interesting probabilistic effects will energe on ϵ^{-2} scales. Accordingly, we introduce the following slow spatial and temporal variables:

$$\xi\equiv\epsilon^2x,\quad \tau\equiv\epsilon^2t \tag{5}$$

and view the solution u as a function of $x,t,\xi,\tau,\epsilon,\omega$. The differential operators transform as follows:

$$\frac{\partial}{\partial x} \to \frac{\partial}{\partial x} + \epsilon^2 \frac{\partial}{\partial \xi}, \quad \frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + \epsilon^2 \frac{\partial}{\partial \tau} \tag{6}$$

and Eqs. (1) and (2) become

$$\left(\frac{\partial^2}{\partial x^2} + 2\epsilon^2 \frac{\partial^2}{\partial x \partial \xi} + \epsilon^4 \frac{\partial^2}{\partial \xi^2}\right) u(x, t, \xi, \tau, \omega, \epsilon) - \left(\frac{\partial}{\partial t} + \epsilon^2 \frac{\partial}{\partial \tau}\right) \times [c^{-2}(1 + \epsilon \mu(x, t, \omega))\left(\frac{\partial}{\partial t} + \epsilon^2 \frac{\partial}{\partial \tau}\right) u(x, t, \xi, \tau, \omega, \epsilon)] \equiv \left(\sum_{n=0}^{4} \epsilon^n \mathcal{L}_n\right) u = 0, \tag{7}$$

$$u(x, 0, \xi, 0, \omega, \epsilon) = f(x),$$
$$\frac{\partial}{\partial t} u(x, 0, \xi, 0, \omega, \epsilon) + \epsilon^2 \frac{\partial}{\partial \tau} u(x, 0, \xi, 0, \omega, \epsilon) = g(x). \tag{8}$$

In particular:

$$\mathcal{L}_0 = \frac{\partial^2}{\partial x^2} - c^{-2} \frac{\partial^2}{\partial t^2},$$
$$\mathcal{L}_1 = -c^{-2} \frac{\partial}{\partial t}\left(\mu(x, t, \omega) \frac{\partial}{\partial t}(\cdot)\right), \tag{9}$$
$$\mathcal{L}_2 = 2\left(\frac{\partial^2}{\partial x \partial \xi} - c^{-2} \frac{\partial^2}{\partial t \partial \tau}\right).$$

The solution u is expanded in a power series in ϵ,

$$u(x, t, \xi, \tau, \omega, \epsilon) \equiv \sum_{n=0}^{\infty} \epsilon^n u_n(x, t, \xi, \tau, \omega). \tag{10}$$

When (10) is substituted into (7), (8) and coefficients of the same power of ϵ are equated, we obtain the following hierarchy of problems:

$$\text{(i)}\ \mathcal{L}_0 u_0 = 0, \quad u_0(x, 0, \xi, 0) = f(x), \quad \frac{\partial}{\partial t} u_0(x, 0, \xi, 0) = g(x),$$

$$\text{(ii)}\ \mathcal{L}_0 u_1 = -\mathcal{L}_1 u_0, \quad u_1(x, 0, \xi, 0, \omega) = 0,$$
$$\frac{\partial}{\partial t} u_1(x, 0, \xi, 0, \omega) = 0,$$

$$\text{(iii)}\ \mathcal{L}_0 u_2 = -\mathcal{L}_1 u_1 - \mathcal{L}_2 u_0, \quad u_2(x, 0, \xi, 0, \omega) = 0,$$
$$\frac{\partial}{\partial t} u_2(x, 0, \xi, 0, \omega) = -\frac{\partial}{\partial t} u_0(x, 0, \xi, 0), \tag{11}$$
$$\vdots$$

Stochastic effects in the actual solution gradually build up over long space–time propagation paths. Because of the assumed mixing property of the random field, the solution becomes essentially independent of the random field contained in any given space–time correlation cell. The stochastic perturbation formalism incorporates these features in the sense that computationally u_0 is a deterministic quantity and yet its dependence upon the slow variables ξ and τ will ultimately be dictated by properties of the random field.

The u_0 problem is solved by imposing the initial conditions upon the D'Alembert general solution; we obtain

$$u_0(x, t, \xi, \tau) = v_1(\xi, \tau, x - ct) + v_2(\xi, \tau, x + ct) \tag{12}$$

with

$$v_1(\xi, 0, x - ct) = \tfrac{1}{2} f(x - ct) - \frac{1}{2c} \int_{-\infty}^{x-ct} g(\lambda)\, d\lambda,$$
$$v_2(\xi, 0, x + ct) = \tfrac{1}{2} f(x + ct) + \frac{1}{2c} \int_{-\infty}^{x+ct} g(\lambda)\, d\lambda. \tag{13}$$

Equations (13) will essentially provide the initial conditions for the resulting $\xi\tau$-transport equations. The equations themselves will emerge from the need to suppress secular growth in the expression for $\langle u_2 \rangle$.

We now solve the u_1 problem [cf. (11ii)], imposing the more stringent initial conditions $u_1(x, 0, \xi, \tau, \omega) = (\partial/\partial t) u_1(x, 0, \xi, \tau, \omega) = 0$ $(\tau \geq 0)$. Using Duhamel's Method[3] we obtain

$$u_1(x, t, \xi, \tau, \omega) = -\frac{1}{2c} \int_0^t \int_{x-c(t-s)}^{x+c(t-s)} \frac{\partial}{\partial s}\left(\mu(\lambda, s, \omega) \frac{\partial}{\partial s} u_0(\lambda, s, \xi, \tau)\right) d\lambda ds, \tag{14}$$

where u_0 is given by (12). In solving the u_2 problem (11iii) we again impose the more stringent initial conditions $u_2(x, 0, \xi, \tau) = 0$, $(\partial/\partial t) u_2(x, 0, \xi, \tau) = -(\partial/\partial \tau) u_0(x, 0, \xi, \tau)$. Then, using superposition and Duhamel's Method, we obtain

$$u_2(x, t, \xi, \tau, \omega) = -\frac{1}{2c} \int_0^{\tau} \int_{x-c(t-s)}^{x+c(t-s)} \left[\frac{\partial}{\partial s}\left(\mu(\lambda, s, \omega) \frac{\partial}{\partial s} u_1(\lambda, s, \xi, \tau, \omega)\right) - 2\left(c^2 \frac{\partial^2}{\partial \xi \partial \lambda} - \frac{\partial^2}{\partial \tau \partial s}\right) u_0(\lambda, s, \xi, \tau)\right] d\lambda ds - \frac{1}{2c} \int_{x-ct}^{x+ct} \frac{\partial}{\partial \tau} u_0(\lambda, 0, \xi, \tau) d\lambda. \tag{15}$$

III. EQUATIONS FOR THE FIRST MOMENT

In this section we shall derive equations of evolution for v_1 and v_2 (as functions of ξ and τ). Taking expected values of the terms in (10), we have

$$\langle u \rangle = u_0 + \epsilon \langle u_1 \rangle + \epsilon^2 \langle u_2 \rangle + \cdots. \tag{16}$$

Recall that u_0 is a deterministic quantity. Therefore, noting (14) and the fact that $\langle \mu \rangle = 0$, a formal exchange of operations leads to $\langle u_1 \rangle = 0$. Therefore, $\langle u \rangle = u_0 + \epsilon^2 \langle u_2 \rangle + \cdots$. An examination of the terms comprising $\langle u_2 \rangle$ will reveal that some terms grow secularly with t; suppression of this growth, which is required to make the correction $\epsilon^2 \langle u_2 \rangle$ truly small on $O(1)$ $\xi\tau$ scales, will also determine the equations of evolution for v_1 and v_2 (i.e., u_0).

We are basically interested in the evolution of wave-packets in a statistically fluctuating environment. Thus we are tacitly assuming the initial data f and g to be such that $v_i(\xi, 0, x)$, $i = 1, 2$, are suitably smooth with (essentially) compact support in x. [As (13) indicates, $v_i(\xi, 0, x)$, $i = 1, 2$, are actually independent of ξ.] In the absence of random fluctuations, v_1 and v_2 would propagate undistorted along the characteristics. In the presence of a spatially and temporally fluctuating medium,

however, the packets will become distorted and could also conceivably grow in strength (having energy "pumped in" by the medium) as they evolve. We shall assume, however, that any such accrual of energy occurs, on the average, at a sufficiently slow rate. Specifically, we shall assume that

$$\lim_{t\to\infty}\frac{1}{t}\int_{x-ct}^{x+ct}|v_i(\xi,\tau,\lambda)|d\lambda=0,\quad i=1,2 \tag{17}$$

uniformly in ξ,τ,x. We shall also assume that similar averages of various partial derivatives of v_1 and v_2 vanish in the limit as $t\to\infty$.

Consider $\langle u_2\rangle$, where u_2 is given by (15). Noting (12) the latter portion can be expressed as

$$\begin{aligned}&\frac{1}{c}\int_0^t\int_{x-c(t-s)}^{x+c(t-s)}\left(c^2\frac{\partial^2}{\partial\xi\partial\lambda}-\frac{\partial^2}{\partial\tau\partial s}\right)u_0(\lambda,s,\xi,\tau)d\lambda ds\\&-\frac{1}{2c}\int_{x-ct}^{x+ct}\frac{\partial}{\partial\tau}u_0(\lambda,0,\xi,\tau)d\lambda\\&=t\Bigg[-\left(\frac{\partial}{\partial\tau}-c\frac{\partial}{\partial\xi}\right)v_1(\xi,\tau,x-ct)\\&-\left(\frac{\partial}{\partial\tau}-c\frac{\partial}{\partial\xi}\right)v_2(\xi,\tau,x+ct)\\&+\frac{1}{2t}\int_{x-ct}^{x+ct}\left(\frac{\partial}{\partial\xi}v_1(\xi,\tau,\lambda)-\frac{\partial}{\partial\xi}v_2(\xi,\tau,\lambda)\right)d\lambda\Bigg].\end{aligned} \tag{18}$$

In the context of our assumptions, the last term on the right-hand side of (18) will not contribute to secular growth in t. The remainder of (15) is evaluated using the expression for u_1 given by (14). For brevity, let ∂_i denote partial differentiation with respect to the ith argument. Then

$$\begin{aligned}&-\frac{1}{2c}\int_0^t\int_{x-c(t-s)}^{x+c(t-s)}\left\langle\frac{\partial}{\partial s}\left(\mu(\lambda,s,\omega)\frac{\partial}{\partial s}u_1(\lambda,s,\xi,\tau,\omega)\right)\right\rangle d\lambda ds\\&=t\Bigg[\frac{1}{8c^2t}\int_0^{2ct}\int_0^{2ct-\sigma}[c^2R(-\sigma,c^{-1}\sigma)(\partial_{33}^2v_1(\xi,\tau,2\sigma+x-ct)\\&+\partial_{33}^2v_2(\xi,\tau,\sigma+\eta+x-ct)+\partial_{33}^2v_1(\xi,\tau,\sigma-\eta+x+ct)\\&+\partial_{33}^2v_2(\xi,\tau,x+ct))-c\partial_2R(-\sigma,c^{-1}\sigma)(-\partial_3v_1(\xi,\tau,2\sigma\\&+x-ct)+\partial_3v_2(\xi,\tau,\sigma+\eta+x-ct)\\&-\partial_3v_1(\xi,\tau,\sigma-\eta+x+ct)+\partial_3v_2(\xi,\tau,x+ct))\\&+c^2R(\sigma,c^{-1}\sigma)(\partial_{33}^2v_1(\xi,\tau,x-ct)\\&+\partial_{33}^2v_2(\xi,\tau,\eta-\sigma+x-ct)+\partial_{33}^2v_1(\xi,\tau,-\sigma-\eta+x+ct)\\&+\partial_{33}^2v_2(\xi,\tau,-2\sigma+x+ct))-c\partial_2R(\sigma,c^{-1}\sigma)\\&\times(-\partial_3v_1(\xi,\tau,x-ct)+\partial_3v_2(\xi,\tau,\eta-\sigma+x-ct)\\&-\partial_3v_1(\xi,\tau,-\eta-\sigma+x+ct)\\&+\partial_3v_2(\xi,\tau,-2\sigma+x+ct))]d\eta d\sigma\Bigg].\end{aligned} \tag{19}$$

Note that $R(\pm\sigma,c^{-1}\sigma)$ corresponds to correlations along the two characteristic directions. We shall assume that both R and ∂_2R decrease rapidly as a function of σ in the sense that

$$\int_0^\infty[R(\pm\sigma,c^{-1}\sigma)+|\partial_2R(\pm\sigma,c^{-1}\sigma)|]\sigma^n d\sigma<\infty \tag{20}$$

for $n=0,1$. In view of assumptions (17) and similar assumptions for the partial derivatives, contributions to the secular term will arise from those portions of the integrand of (19) which are independent of η. For brevity of notation, let

$$\alpha=x-ct,\quad \beta=x+ct. \tag{21}$$

Combining the secular terms in (19) with those of (18), we observe that

$$\begin{aligned}&-\left(\frac{\partial}{\partial\tau}+c\frac{\partial}{\partial\xi}\right)v_1(\xi,\tau,\alpha)-\left(\frac{\partial}{\partial\tau}-c\frac{\partial}{\partial\xi}\right)v_2(\xi,\tau,\beta)\\&+\frac{c}{4}\int_0^{2ct}R(-\sigma,c^{-1}\sigma)\partial_{33}^2v_1(\xi,\tau,2\sigma+\alpha)d\sigma\\&+\frac{1}{4}\int_0^{2ct}\partial_2R(-\sigma,c^{-1}\sigma)\partial_3v_1(\xi,\tau,2\sigma+\alpha)d\sigma\\&+\frac{c}{4}\partial_{33}^2v_1(\xi,\tau,\alpha)\int_0^{2ct}R(\sigma,c^{-1}\sigma)d\sigma\\&+\frac{1}{4}\partial_3v_1(\xi,\tau,\alpha)\int_0^{2ct}\partial_2R(\sigma,c^{-1}\sigma)d\sigma\\&+\frac{c}{4}\partial_{33}^2v_2(\xi,\tau,\beta)\int_0^{2ct}R(-\sigma,c^{-1}\sigma)d\sigma\\&-\frac{1}{4}\partial_3v_2(\xi,\tau,\beta)\int_0^{2ct}\partial_2R(-\sigma,c^{-1}\sigma)d\sigma\\&+\frac{c}{4}\int_0^{2ct}R(\sigma,c^{-1}\sigma)\partial_{33}^2v_2(\xi,\tau,-2\sigma+\beta)d\sigma\\&-\frac{1}{4}\int_0^{2ct}\partial_2R(\sigma,c^{-1}\sigma)\partial_3v_2(\xi,\tau,-2\sigma+\beta)d\sigma\end{aligned} \tag{22}$$

must necessarily vanish as $t\to\infty$ if $\epsilon^2\langle u_2\rangle$ is to be genuinely small on $O(1)$ $\xi\tau$ scales. Recall that we have tacitly assumed initial data corresponding to wavepacket propagation. In the absence of random fluctuations, an initial localized disturbance would split into forward and backward propagating components; these components, in turn, would propagate undistorted along the characteristics.

In the randomly perturbed case, we shall assume that the same gross qualitative features exist, i.e., that the support of v_1 is concentrated on a family of characteristics $x-ct=\alpha=\text{const}$ while the support of v_2 remains concentrated on a family of characteristics $x+ct=\beta=\text{const}$. Thus, while the packets may be distorted, smeared or otherwise affected by the random fluctuations, we assume that these fluctuations have not totally obliterated the packets.

Suppression of the secular term (22) as $t\to\infty$ reduces, therefore, to the suppression of secular growth in the two characteristic directions. Setting $\alpha\equiv\text{const}$ and letting $t\to\infty$, we obtain the equation

$$\begin{aligned}&\left(\frac{\partial}{\partial\tau}+c\frac{\partial}{\partial\xi}\right)v_1(\xi,\tau,\alpha)\\&=\frac{c}{4}\int_0^\infty R(-\sigma,c^{-1}\sigma)\frac{\partial^2}{\partial\alpha^2}v_1(\xi,\tau,2\sigma+\alpha)d\sigma\\&+\frac{1}{4}\int_0^\infty\partial_2R(-\sigma,c^{-1}\sigma)\frac{\partial}{\partial\alpha}v_1(\xi,\tau,2\sigma+\alpha)d\sigma\\&+\frac{c}{4}\frac{\partial^2}{\partial\alpha^2}v_1(\xi,\tau,\alpha)\int_0^\infty R(\sigma,c^{-1}\sigma)d\sigma\\&+\frac{1}{4}\frac{\partial}{\partial\alpha}v_1(\xi,\tau,\alpha)\int_0^\infty\partial_2R(\sigma,c^{-1}\sigma)d\sigma.\end{aligned} \tag{23a}$$

The partial derivatives with respect to the third argument have been interpreted as partial derivatives with respect to α. Since $\beta = \alpha + 2ct$, taking the limit $t \to \infty$ along a family of characteristics $\alpha = \text{const}$ will suppress the v_2 terms in (22); we are assuming that v_2 and its various partial derivatives with respect to β vanish as $\beta \to \infty$. [This added assumption is similar but more restrictive than the ones made in (17).]

Setting $\beta \equiv \text{const}$ (so that $\alpha = \beta - 2ct$) and letting $t \to \infty$, we obtain the second equation,

$$\left(\frac{\partial}{\partial\tau} - c\frac{\partial}{\partial\xi}\right) v_2(\xi,\tau,\beta) = \frac{c}{4}\frac{\partial^2}{\partial\beta^2} v_2(\xi,\tau,\beta)\int_0^\infty R(-\sigma, c^{-1}\sigma)d\sigma - \frac{1}{4}\frac{\partial}{\partial\beta} v_2(\xi,\tau,\beta)\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma)d\sigma + \frac{c}{4}\int_0^\infty R(\sigma, c^{-1}\sigma)\frac{\partial^2}{\partial\beta^2} v_2(\xi,\tau,-2\sigma+\beta)d\sigma - \frac{1}{4}\int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma)\frac{\partial}{\partial\beta} v_2(\xi,\tau,-2\sigma+\beta)d\sigma. \tag{23b}$$

Together with Eq. (23), we have initial conditions (13), which can be recast as

$$v_1(\xi,0,\alpha) = \tfrac{1}{2}f(\alpha) - \frac{1}{2c}\int_{-\infty}^{\alpha} g(s)ds, \tag{24a}$$

$$v_2(\xi,0,\beta) = \tfrac{1}{2}f(\beta) + \frac{1}{2c}\int_{-\infty}^{\beta} g(s)ds. \tag{24b}$$

Both α and β range from $-\infty$ to $+\infty$ as x and t vary over the half-plane $-\infty < x < \infty$, $t \geq 0$. Introducing Fourier transforms greatly simplifies problems (23) and (24). Define

$$\hat{v}_i(\xi,\tau,\gamma) \equiv \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{\infty} v_i(\xi,\tau,z)e^{-i\gamma z}dz, \quad i = 1,2. \tag{25}$$

Then, (23) and (24) transform into the following pair of first order Cauchy problems:

$$\left(\frac{\partial}{\partial\tau} + c\frac{\partial}{\partial\xi}\right)\hat{v}_1(\xi,\tau,\gamma) = -\Gamma_1(\gamma)\hat{v}_1(\xi,\tau,\gamma), \tag{26}$$

$$-\infty < \xi < \infty, \quad \tau \geq 0, \quad -\infty < \gamma < \infty$$

$$\Gamma_1(\gamma) \equiv \frac{c\gamma^2}{4}\int_0^\infty R(-\sigma, c^{-1}\sigma)e^{i2\gamma\sigma}d\sigma + \frac{c\gamma^2}{4}\int_0^\infty R(\sigma, c^{-1}\sigma)d\sigma - \frac{i\gamma}{4}\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma)e^{i2\gamma\sigma}d\sigma - \frac{i\gamma}{4}\int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma)d\sigma,$$

$$\left(\frac{\partial}{\partial\tau} - c\frac{\partial}{\partial\xi}\right)\hat{v}_2(\xi,\tau,\gamma) = -\Gamma_2(\gamma)\hat{v}_2(\xi,\tau,\gamma), \tag{27}$$

$$-\infty < \xi < \infty, \quad \tau \geq 0, \quad -\infty < \gamma < \infty,$$

$$\Gamma_2(\gamma) \equiv \frac{c\gamma^2}{4}\int_0^\infty R(-\sigma, c^{-1}\sigma)d\sigma + \frac{c\gamma^2}{4}\int_0^\infty R(\sigma, c^{-1}\sigma)e^{-i2\gamma\sigma}d\sigma + \frac{i\gamma}{4}\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma)d\sigma + \frac{i\gamma}{4}\int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma)e^{-i2\gamma\sigma}d\sigma,$$

where the initial data $\hat{v}_i(\xi,0,\gamma)$, $i = 1,2$, are obtained by the Fourier transformation of Eq. (24). Note that the initial data is actually independent of ξ; to make this point explicit, we shall set $\hat{v}_i(\xi,0,\gamma) \equiv \Phi_i(\gamma)$, $i = 1,2$. Then, the solutions of (26) and (27) are

$$\hat{v}_i(\xi,\tau,\gamma) = \Phi_i(\gamma)\exp[-\Gamma_i(\gamma)\tau], \quad i = 1,2. \tag{28}$$

The desired functions $v_1(\xi,\tau,\alpha)$ and $v_2(\xi,\tau,\beta)$ must then be determined by inverse Fourier transformation.

We shall conclude this section by considering an idealized special case for which the computations are particularly simple. Let

$$f(x) = \frac{e^{-x^2/2\aleph^2}}{(2\pi)^{1/2}\aleph}, \quad g(x) = -c\frac{df}{dx}(x). \tag{29}$$

This choice of initial data corresponds to

$$v_1(\xi,0,\alpha) = f(\alpha), \quad v_2(\xi,0,\beta) = 0. \tag{30}$$

Thus we consider the case of a right-propagating Gaussian pulse; in the absence of random fluctuations, this pulse would propagate undistorted. For simplicity, assume that the random field μ is independent of time and spatially delta-correlated, i.e.,

$$R(x-x', t-t') = S_0\delta(x-x'). \tag{31}$$

Then, the inverse Fourier transformation of expressions (28) for this example leads to

$$v_1(\xi,\tau,x-ct) = \frac{\exp[-(x-ct)^2/2(\aleph^2 + \frac{1}{2}cS_0\tau)]}{(2\pi)^{1/2}(\aleph^2 + \frac{1}{2}cS_0\tau)^{1/2}}, \qquad v_2(\xi,\tau,x+ct) = 0. \tag{32}$$

Therefore,

$$\langle u(x,t,\omega,\epsilon)\rangle \sim \frac{\exp[-(x-ct)^2/2(\aleph^2 + \frac{1}{2}\epsilon^2 cS_0 t)]}{(2\pi)^{1/2}(\aleph^2 + \frac{1}{2}\epsilon^2 cS_0 t)^{1/2}}. \tag{33}$$

IV. EQUATIONS FOR THE MUTUAL COHERENCE FUNCTION

In this section we shall study the asymptotic behavior of $\langle u(x_1,t_1,\omega,\epsilon)u(x_2,t_2,\omega,\epsilon)\rangle$. We again introduce slow variables ξ_i, τ_i, $i = 1,2$, and develop the solution at each space–time point in an ϵ power series [cf. (10)]. For brevity, let $u^{(i)}$ and $u_n^{(i)}$ denote $u(x_i,t_i,\xi_i,\tau_i,\omega,\epsilon)$ and $u_n(x_i,t_i,\xi_i,\tau_i,\omega)$, respectively (where $i = 1,2$ and $n = 0,1,2,\cdots$). Then

$$u^{(1)}u^{(2)} = u_0^{(1)}u_0^{(2)} + \epsilon(u_0^{(1)}u_1^{(2)} + u_1^{(1)}u_0^{(2)}) + \epsilon^2(u_0^{(1)}u_0^{(2)} + u_1^{(1)}u_1^{(2)} + u_2^{(1)}u_0^{(2)}) + \cdots. \tag{34}$$

The product $u_0^{(1)}u_0^{(2)}$ is computationally a deterministic quantity. Noting (14) we obtain

$$\langle u^{(1)}u^{(2)}\rangle = u_0^{(1)}u_0^{(2)} + \epsilon^2(u_0^{(1)}\langle u_2^{(2)}\rangle + \langle u_1^{(1)}u_1^{(2)}\rangle + \langle u_2^{(1)}\rangle u_0^{(2)}) + \cdots. \tag{35}$$

Recall that we are essentially interested in the evolution of wavepackets in a random environment. Thus, the spatial and temporal offsets of interest, i.e.,

$$\Delta x \equiv x_1 - x_2, \quad \Delta t \equiv t_1 - t_2 \tag{36}$$

will be $O(1)$, being limited basically by the support of the packet. Consequently, the case of interest will correspond to $O(\epsilon^2)$ offsets in the slow variables ξ and τ. Therefore, we shall be ultimately concerned (to leading order) with $\xi_1 = \xi_2 \equiv \xi$ and $\tau_1 = \tau_2 \equiv \tau$.

Note that the consideration of the mutual coherence function involves a "cross-coupling" term $\langle u_1^{(1)} u_1^{(2)} \rangle$ which will contribute to the ultimate equations of evolution. Let

$$\begin{aligned} &\alpha_i = x_i - ct_i, \quad \beta_i = x_i + ct_i, \quad i = 1, 2, \\ &X = \tfrac{1}{2}(x_1 + x_2), \quad T \equiv \tfrac{1}{2}(t_1 + t_2). \end{aligned} \tag{37}$$

Again, we must identify the terms in the ϵ^2 coefficient of (35) which grow secularly as $T \to \infty$. Recall that the $v_1(\xi_i, \tau_i, \alpha_i)$ terms have their support located near $X - cT = \text{const}$ while the $v_2(\xi_i, \tau_i, \beta_i)$ terms have their support in β located near $X + cT = \text{const}$. Therefore, as $T \to \infty$, two equations will emerge from the need to suppress secular growth along the two families of characteristics. Note that cross products of the form $v_1(\xi_i, \tau_i, \alpha_i) v_2(\xi_j, \tau_j, \beta_j)$, $i, j = 1, 2$, will tend to zero as T increases since the supports of the two terms forming the product become essentially disjoint.

Suppression of secular growth along the characteristic family $\alpha = \text{const}$ necessitates the vanishing of

$$\begin{aligned}
&-v_1(\xi_1, \tau_1, \alpha_1)\left(\frac{\partial}{\partial \tau_2} + c\frac{\partial}{\partial \xi_2}\right) v_1(\xi_2, \tau_2, \alpha_2) - v_1(\xi_2, \tau_2, \alpha_2) \\
&\times \left(\frac{\partial}{\partial \tau_1} + c\frac{\partial}{\partial \xi_1}\right) v_1(\xi_1, \tau_1, \alpha_1) + \frac{c}{4}\int_0^\infty R(-\sigma, c^{-1}\sigma) \\
&\times \left(v_1(\xi_1, \tau_1, \alpha_1)\frac{\partial^2}{\partial \alpha_2^2} v_1(\xi_2, \tau_2, 2\sigma + \alpha_2) + v_1(\xi_2, \tau_2, \alpha_2)\right. \\
&\left.\times \frac{\partial^2}{\partial \alpha_1^2} v_1(\xi_1, \tau_1, 2\sigma + \alpha_1)\right) d\sigma \\
&+\frac{1}{4}\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma)\left(v_1(\xi_1, \tau_1, \alpha_1)\frac{\partial}{\partial \alpha_2} v_1(\xi_2, \tau_2, 2\sigma + \alpha_2)\right. \\
&\left.+ v_1(\xi_2, \tau_2, \alpha_2)\frac{\partial}{\partial \alpha_1} v_1(\xi_1, \tau_1, 2\sigma + \alpha_1)\right) d\sigma \\
&+\frac{c}{4}\left(v_1(\xi_1, \tau_1, \alpha_1)\frac{\partial^2}{\partial \alpha_2^2} v_1(\xi_2, \tau_2, \alpha_2) + v_1(\xi_2, \tau_2, \alpha_2)\right. \\
&\left.\times \frac{\partial^2}{\partial \alpha_1^2} v_1(\xi_1, \tau_1, \alpha_1)\right)\int_0^\infty R(\sigma, c^{-1}\sigma) d\sigma + \frac{1}{4}\left(v_1(\xi_1, \tau_1, \alpha_1)\right. \\
&\left.\times \frac{\partial}{\partial \alpha_2} v_1(\xi_2, \tau_2, \alpha_2) + v_1(\xi_2, \tau_2, \alpha_2)\frac{\partial}{\partial \alpha_1} v_1(\xi_1, \tau_1, \alpha_1)\right) \\
&\times \int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma) d\sigma + \frac{c}{4}\frac{\partial}{\partial \alpha_1} v_1(\xi_1, \tau_1, \alpha_1) \\
&\times \frac{\partial}{\partial \alpha_2} v_2(\xi_2, \tau_2, \alpha_2)\int_{-\infty}^\infty R(\sigma, c^{-1}(\sigma - \alpha_1 + \alpha_2)) d\sigma .
\end{aligned} \tag{38}$$

Suppression of secular growth along the characteristic family $\beta = \text{const}$ leads to the vanishing of

$$\begin{aligned}
&-v_2(\xi_1, \tau_1, \beta_1)\left(\frac{\partial}{\partial \tau_2} - c\frac{\partial}{\partial \xi_2}\right) v_2(\xi_2, \tau_2, \beta_2) - v_2(\xi_2, \tau_2, \beta_2) \\
&\times \left(\frac{\partial}{\partial \tau_1} - c\frac{\partial}{\partial \xi_1}\right) v_2(\xi_1, \tau_1, \beta_1) + \frac{c}{4}\left(v_2(\xi_1, \tau_1, \beta_1)\frac{\partial^2}{\partial \beta_2^2}\right. \\
&\left.\times v_2(\xi_2, \tau_2, \beta_2) + v_2(\xi_2, \tau_2, \beta_2)\frac{\partial^2}{\partial \beta_1^2} v_2(\xi_1, \tau_1, \beta_1)\right) \\
&\times \int_0^\infty R(-\sigma, c^{-1}\sigma) d\sigma - \frac{1}{4}\left(v_2(\xi_1, \tau_1, \beta_1)\frac{\partial}{\partial \beta_2} v_2(\xi_2, \tau_2, \beta_2)\right. \\
&\left.+ v_2(\xi_2, \tau_2, \beta_2)\frac{\partial}{\partial \beta_1} v_2(\xi_1, \tau_1, \beta_1)\right)\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma) d\sigma \\
&+\frac{c}{4}\int_0^\infty R(\sigma, c^{-1}\sigma)\left(v_2(\xi_1, \tau_1, \beta_1)\frac{\partial^2}{\partial \beta_2^2} v_2(\xi_2, \tau_2, \beta_2 - 2\sigma)\right. \\
&\left.+ v_2(\xi_2, \tau_2, \beta_2)\frac{\partial^2}{\partial \beta_1^2} v_2(\xi_1, \tau_1, \beta_1 - 2\sigma)\right) d\sigma \\
&-\frac{1}{4}\int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma)\left(v_2(\xi_1, \tau_1, \beta_1)\frac{\partial}{\partial \beta_2} v_2(\xi_2, \tau_2, \beta_2 - 2\sigma)\right. \\
&\left.+ v_2(\xi_2, \tau_2, \beta_2)\frac{\partial}{\partial \beta_1} v_2(\xi_1, \tau_1, \beta_1 - 2\sigma)\right) d\sigma \\
&+\frac{c}{4}\frac{\partial}{\partial \beta_1} v_2(\xi_1, \tau_1, \beta_1)\frac{\partial}{\partial \beta_2} v_2(\xi_2, \tau_2, \beta_2) \\
&\times \int_{-\infty}^\infty R(\sigma, c^{-1}(\beta_1 - \beta_2 - \sigma)) d\sigma .
\end{aligned} \tag{39}$$

Recall that we are interested in the case where $\xi_1 = \xi_2 \equiv \xi$ and $\tau_1 = \tau_2 \equiv \tau$; define

$$\begin{aligned} &w_1(\xi, \tau, \alpha_1, \alpha_2) \equiv v_1(\xi, \tau, \alpha_1) v_1(\xi, \tau, \alpha_2), \\ &w_2(\xi, \tau, \beta_1, \beta_2) \equiv v_2(\xi, \tau, \beta_1) v_2(\xi, \tau, \beta_2). \end{aligned} \tag{40}$$

Then, (38) and (39) can be recast as the following equations for w_1 and w_2:

$$\begin{aligned}
&\left(\frac{\partial}{\partial \tau} + c\frac{\partial}{\partial \xi}\right) w_1(\xi, \tau, \alpha_1, \alpha_2) \\
&= \frac{c}{4}\int_0^\infty R(-\sigma, c^{-1}\sigma)\left(\frac{\partial^2}{\partial \alpha_1^2} w_1(\xi, \tau, 2\sigma + \alpha_1, \alpha_2)\right. \\
&\left.+ \frac{\partial^2}{\partial \alpha_2^2} w_1(\xi, \tau, \alpha_1, 2\sigma + \alpha_2)\right) d\sigma \\
&+\frac{1}{4}\int_0^\infty \partial_2 R(-\sigma, c^{-1}\sigma)\left(\frac{\partial}{\partial \alpha_1} w_1(\xi, \tau, 2\sigma + \alpha_1, \alpha_2)\right. \\
&\left.+ \frac{\partial}{\partial \alpha_2} w_1(\xi, \tau, \alpha_1, 2\sigma + \alpha_2)\right) d\sigma \\
&+\frac{c}{4}\left(\frac{\partial^2}{\partial \alpha_1^2} + \frac{\partial^2}{\partial \alpha_2^2}\right) w_1(\xi, \tau, \alpha_1, \alpha_2)\int_0^\infty R(\sigma, c^{-1}\sigma) d\sigma \\
&+\frac{1}{4}\left(\frac{\partial}{\partial \alpha_1} + \frac{\partial}{\partial \alpha_2}\right) w_1(\xi, \tau, \alpha_1, \alpha_2)\int_0^\infty \partial_2 R(\sigma, c^{-1}\sigma) d\sigma \\
&+\frac{c}{4}\frac{\partial^2}{\partial \alpha_1 \partial \alpha_2} w_1(\xi, \tau, \alpha_1, \alpha_2)\int_{-\infty}^\infty R(\sigma, c^{-1}(\sigma - \alpha_1 + \alpha_2)) d\sigma ,
\end{aligned} \tag{41}$$

$$\left(\frac{\partial}{\partial\tau}-\frac{c\partial}{\partial\xi}\right)w_2(\xi,\tau,\beta_1,\beta_2)$$
$$=\frac{c}{4}\left(\frac{\partial^2}{\partial\beta_1^2}+\frac{\partial^2}{\partial\beta_2^2}\right)w_2(\xi,\tau,\beta_1,\beta_2)\int_0^\infty R(-\sigma,c^{-1}\sigma)d\sigma$$
$$-\frac{1}{4}\left(\frac{\partial}{\partial\beta_1}+\frac{\partial}{\partial\beta_2}\right)w_2(\xi,\tau,\beta_1,\beta_2)\int_0^\infty \partial_2R(-\sigma,c^{-1}\sigma)d\sigma$$
$$+\frac{c}{4}\int_0^\infty R(\sigma,c^{-1}\sigma)\left(\frac{\partial^2}{\partial\beta_1^2}w_2(\xi,\tau,\beta_1-2\sigma,\beta_2)\right.$$
$$\left.+\frac{\partial^2}{\partial\beta_2^2}w_2(\xi,\tau,\beta_1,\beta_2-2\sigma)\right)d\sigma-\frac{1}{4}\int_0^\infty \partial_2R(\sigma,c^{-1}\sigma)$$
$$\times\left(\frac{\partial}{\partial\beta_1}w_2(\xi,\tau,\beta_1-2\sigma,\beta_2)+\frac{\partial}{\partial\beta_2}w_2(\xi,\tau,\beta_1,\beta_2-2\sigma)\right)d\sigma$$
$$+\frac{c}{4}\frac{\partial^2}{\partial\beta_1\beta_2}w_2(\xi,\tau,\beta_1,\beta_2)\int_{-\infty}^\infty R(\sigma,c^{-1}(\beta_1-\beta_2-\sigma))d\sigma. \tag{42}$$

The analysis of (41) and (42) is again greatly facilitated by the use of Fourier transforms; define

$$\hat{w}_i(\xi,\tau,\gamma_1,\gamma_2)\equiv\frac{1}{2\pi}\int_{-\infty}^\infty\int_{-\infty}^\infty w_i(\xi,\tau,z_1,z_2)$$
$$\times\exp[-i(\gamma_1z_1+\gamma_2z_2)]dz_1dz_2,\quad i=1,2. \tag{43}$$

Note that if the random field varies solely as a function of space or time, the equations for $\hat{w}_1$ and $\hat{w}_2$ will be first order linear constant coefficient equations. The general case, however, will not be so simple; the "cross-product" terms [i.e., the last terms in (41) and (42)] become convolution integrals in the transform domain. Let

$$\rho_1(z)\equiv\int_{-\infty}^\infty R(\sigma,c^{-1}(\sigma-z))d\sigma,$$
$$\rho_2(z)\equiv\int_{-\infty}^\infty R(\sigma,c^{-1}(z-\sigma))d\sigma \tag{44}$$

and let $\hat{\rho}_i$, $i=1,2$, denote the corresponding Fourier transforms. Then, a Fourier transformation of (41) and (42) leads to

$$\left(\frac{\partial}{\partial\tau}+c\frac{\partial}{\partial\xi}\right)\hat{w}_1(\xi,\tau,\gamma_1,\gamma_2)$$
$$=-\Lambda(\gamma_1,\gamma_2)\hat{w}_1(\xi,\tau,\gamma_1,\gamma_2)-\frac{c}{4(2\pi)^{1/2}}\int_{-\infty}^\infty\hat{\rho}_1(\gamma)$$
$$\times(\gamma_1-\gamma)(\gamma_2+\gamma)\hat{w}_1(\xi,\tau,\gamma_1-\gamma,\gamma_2+\gamma)d\gamma,$$
$$\Lambda_1(\gamma_1,\gamma_2) \tag{45}$$
$$\equiv\frac{c}{4}\left(\gamma_1^2\int_0^\infty R(-\sigma,c^{-1}\sigma)\exp(i2\gamma_1\sigma)d\sigma\right.$$
$$+\gamma_2^2\int_0^\infty R(-\sigma,c^{-1}\sigma)\exp(i2\gamma_2\sigma)d\sigma$$
$$\left.+(\gamma_1^2+\gamma_2^2)\int_0^\infty R(\sigma,c^{-1}\sigma)d\sigma\right)$$
$$-\frac{i}{4}\left(\gamma_1\int_0^\infty \partial_2R(-\sigma,c^{-1}\sigma)\exp(i2\gamma_1\sigma)d\sigma\right.$$
$$+\gamma_2\int_0^\infty \partial_2R(-\sigma,\ c^{-1}\sigma)\exp(i2\gamma_2\sigma)d\sigma$$
$$\left.+(\gamma_1+\gamma_2)\int_0^\infty \partial_2R(\sigma,c^{-1}\sigma)d\sigma\right)$$

and

$$\left(\frac{\partial}{\partial\tau}-c\frac{\partial}{\partial\xi}\right)\hat{u}_2(\xi,\tau,\gamma_1,\gamma_2)$$
$$=-\Lambda_2(\gamma_1,\gamma_2)\hat{w}_2(\xi,\tau,\gamma_1,\gamma_2)-\frac{c}{4(2\pi)^{1/2}}$$
$$\times\int_{-\infty}^\infty\hat{\rho}_2(\gamma)(\gamma_1-\gamma)(\gamma_2+\gamma)\hat{w}_2(\xi,\tau,\gamma_1-\gamma,\gamma_2+\gamma)d\gamma, \tag{46}$$

$$\Lambda_2(\gamma_1,\gamma_2)$$
$$\equiv\frac{c}{4}\left((\gamma_1^2+\gamma_2^2)\int_0^\infty R(-\sigma,c^{-1}\sigma)d\sigma\right.$$
$$+\gamma_1^2\int_0^\infty R(\sigma,c^{-1}\sigma)\exp(-i2\gamma_1\sigma)d\sigma$$
$$\left.+\gamma_2^2\int_0^\infty R(\sigma,c^{-1}\sigma)\exp(-i2\gamma_2\sigma)d\sigma\right)$$
$$+\frac{i}{4}\left((\gamma_1+\gamma_2)\int_0^\infty \partial_2R(\sigma,c^{-1}\sigma)d\sigma\right.$$
$$+\gamma_1\int_0^\infty \partial_2R(\sigma,c^{-1}\sigma)\exp(-i2\gamma_1\sigma)d\sigma$$
$$\left.+\gamma_2\int_0^\infty \partial_2R(\sigma,c^{-1}\sigma)\exp(-i2\gamma_2\sigma)d\sigma\right).$$

We conclude this section by considering again the example discussed at the end of Sec. III; we consider a right-propagating Gaussian pulse in a random field that is time independent and spatially delta-correlated [cf. (29)–(31)]. In this case, (45) and (46) simplify to

$$\left(\frac{\partial}{\partial\tau}+\frac{c\partial}{\partial\xi}\right)\hat{w}_1=-\frac{c}{4}S_0(\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2)\hat{w}_1,$$
$$\hat{w}_1(\xi,0,\gamma_1,\gamma_2)=\frac{e^{-\aleph^2(\gamma_1^2+\gamma_2^2)/2}}{2\pi},$$
$$\left(\frac{\partial}{\partial\tau}-\frac{c\partial}{\partial\xi}\right)\hat{w}_2=-\frac{c}{4}S_0(\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2)\hat{w}_2, \tag{47}$$
$$\hat{w}_2(\xi,0,\gamma_1,\gamma_2)=0.$$

Solving these equations for $\hat{w}_i$, $i=1,2$, and taking inverse Fourier transforms lead to

$$\langle u(x_1,t_1,\omega,\epsilon)u(x_2,t_2,\omega,\epsilon)\rangle\sim\exp\left(-\frac{(\Delta x-c\Delta t)^2}{4(\aleph^2+\frac{1}{4}\epsilon^2cS_0T)}-\frac{(X-cT)^2}{(\aleph^2+\frac{3}{4}\epsilon^2cS_0T)}\right)\Big/[2\pi(\aleph^2+\tfrac{1}{4}\epsilon^2cS_0T)^{1/2}(\aleph^2+\tfrac{3}{4}\epsilon^2cS_0T)^{1/2}]. \tag{48}$$

Note that if $x_1=x_2=ct_1=ct_2$, the effect of the random field upon the mutual coherence function reduces to an attenuation due to the demoninator term. In general, the fluctuations in the solution are given by

$$\langle u^2(x,t,\omega,\epsilon)\rangle - \langle u(x,t,\omega,\epsilon)\rangle^2$$

$$\sim \exp\left(\frac{-(x-ct)^2}{(\aleph^2+\frac{3}{4}\epsilon^2 cS_0 t)}\right)\Big/\left[2\pi(\aleph^2+\tfrac{1}{4}\epsilon^2 cS_0 t)^{1/2}\ (\aleph^2+\tfrac{3}{4}\epsilon^2 cS_0 t)^{1/2}\right] - \exp\left(\frac{-(x-ct)^2}{(\aleph^2+\frac{1}{2}\epsilon^2 cS_0 t)}\right)\Big/ 2\pi(\aleph^2+\tfrac{1}{2}\epsilon^2 cS_0 t)\,. \tag{49}$$

V. A SPECIAL CASE

If the random field fluctuates solely as a function of space or time, the functions $\hat{w}_i(\xi,\tau,\gamma_1,\gamma_2)$, $i=1,2$, are readily determined as solutions of first-order linear equations. The mutual coherence function is then obtained by inverse Fourier transformation. In this section, we consider the case of a spatially fluctuating Gaussian random field; let

$$R(x,t)\equiv R(x)=\exp(-x^2/2\sigma^2)/(2\pi)^{1/2}\sigma\,. \tag{50}$$

We again adopt initial data corresponding to a right-propagating Gaussian pulse [cf. (29)]. Then, (45) and (46) [with $\hat{\rho}_1(\gamma)=\hat{\rho}_2(\gamma)=(2\pi)^{1/2}\delta(\gamma)$] imply

$$\hat{w}_1(\xi,\tau,\gamma_1,\gamma_2)$$

$$=\frac{1}{2\pi}\exp\left[-\frac{\gamma_1^2}{2}\left(\aleph^2+\frac{c\tau}{4}[1+\exp(-2\sigma^2\gamma_1^2)]\right)\right.$$

$$-\frac{\gamma_2^2}{2}\left(\aleph^2+\frac{c\tau}{4}[1+\exp(-2\sigma^2\gamma_2^2)]\right)-\frac{c\tau}{4}\gamma_1\gamma_2-\frac{ic\tau}{4(2\pi)^{1/2}}$$

$$\times\left(\gamma_1^2\exp(-2\sigma^2\gamma_1^2)\int_0^{\sigma\gamma_1}\exp(\lambda^2/2)d\lambda+\gamma_2^2\right.$$

$$\left.\left.\times\int_0^{\sigma\gamma_2}\exp(\lambda^2/2)d\lambda\right)\right]\,, \tag{51}$$

$$w_2(\xi,\tau,\gamma_1,\gamma_2)=0\,.$$

For simplicity, we shall study the parametric dependence of $w_1(\xi,\tau,0,0)$ (i.e., $x_1=x_2=ct_1=ct_2$) upon σ. Noting (51) we have

$$w_1(\xi,\tau,0,0)$$

$$=\frac{1}{2\pi^2}\int_0^\infty\int_{-\infty}^\infty\exp\left[-\frac{\gamma_1^2}{2}\left(\aleph^2+\frac{c\tau}{4}[1+\exp(-2\sigma^2\gamma_1^2)]\right)\right.$$

$$\left.-\frac{\gamma_2^2}{2}\left(\aleph^2+\frac{c\tau}{4}[1+\exp(-2\sigma^2\gamma_2^2)]\right)-\frac{c\tau}{4}\gamma_1\gamma_2\right]$$

$$\times\cos\left(\frac{c\tau}{4(2\pi)^{1/2}}[\gamma_1^2\exp(-2\sigma^2\gamma_1^2)\int_0^{\sigma\gamma_1}\exp(\lambda^2/2)d\lambda\right.$$

$$\left.+\gamma_2^2\exp(-2\sigma^2\gamma_2^2)\int_0^{\sigma\gamma_2}\exp(\lambda^2/2)d\lambda]\right)d\gamma_1 d\gamma_2 \tag{52}$$

and

$$\frac{\langle u^2(x,t,\omega,\epsilon)\rangle|_{x=ct}}{\langle u^2(0,0,\omega,\epsilon)\rangle}\sim 2\pi\aleph^2 w_1(\xi,\tau,0,0)\,. \tag{53}$$

The variation of this ratio as a function of $c\tau$ ($=\epsilon^2 ct$) is shown in Fig. 1 for several values of σ. The case $\sigma=0$ corresponds to the delta-correlated case with $S_0=1$.

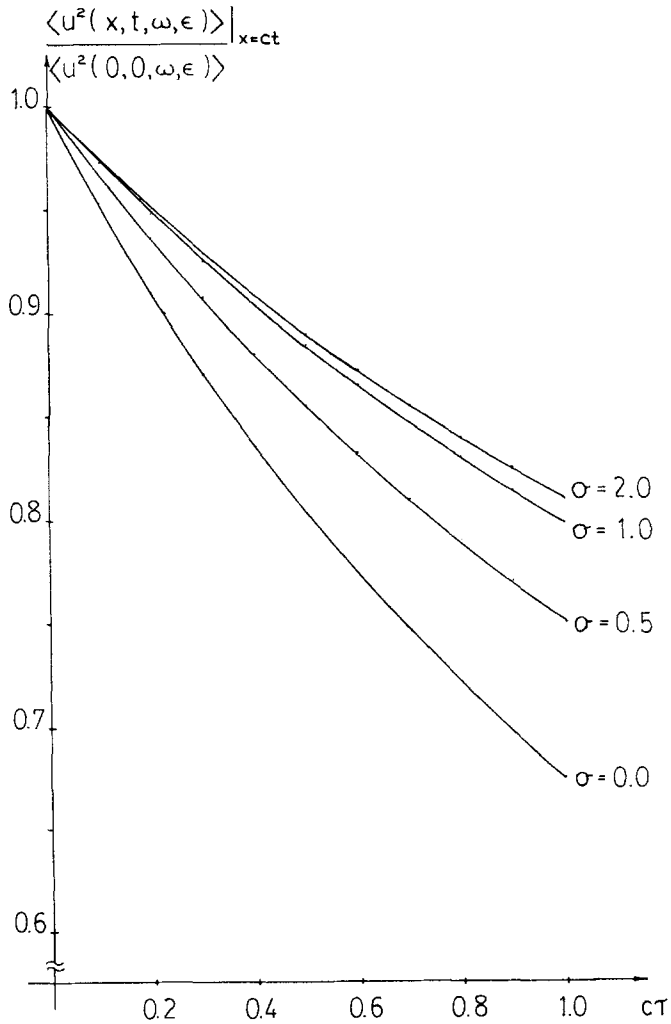


FIG. 1.

[1] G. C. Papanicolaou and J. B. Keller, SIAM J. Appl. Math. **21**, 287 (1971).

[2] G. C. Papanicolaou and S. R. S. Varadhan, Commun. Pure Appl. Math. **26**, 497 (1973).

[3] F. John, *Partial Differential Equations* (Springer, New York, 1975), 2nd ed.

The interaction function and lattice duals

William Greenberg

Department of Mathematics, Virginia Polytechnic Institute & State University, Blacksburg, Virginia 24061
(Received 29 April 1975)

An interaction function is defined for lattice models in statistical mechanics. A correlation function expansion is derived, giving a direct proof of the duality relations for correlation functions.

A general theory of duality transformations between pairs of classical spin-$\frac{1}{2}$ lattice models has been developed by Gruber and Merlini[1] and independently by Wegner.[2] The theory of Gruber and Merlini is constructive, providing explicitly a family of "dual" lattices and Hamiltonians for any given spin-$\frac{1}{2}$ system. These duals are exact, all requisite boundary terms being provided for, which is necessary in considerations of correlation functions below criticality.

We define in this article the interaction functions $u_{H^*}(A,B)$ of lattice duals G and G^*, and express them in terms of correlation functions. This gives an easy derivation of the relationship between correlation functions of a lattice and its duals. The notation in this article, while somewhat different from Ref. 1 and some current usage, has the advantage, in addition to simplifying the derivations, of generalizing to higher spin lattices.[3] The reader is referred to Ref. 1 for details on the construction of dual spin-$\frac{1}{2}$ lattices.

1. DUAL LATTICES

We suppose we are given a finite set Λ of lattice sites in a ν-dimensional space, along with a Hamiltonian H defined on the configuration of Λ. It is convenient to take as the configuration space the group $P_2(\Lambda)$ of functions from Λ to Z_2, the integers modulo 2, with group multiplication

$$fg(\lambda)=f(\lambda)+g(\lambda)\ \mathrm{mod}2\,.$$

Considering H as a function $H:P_2(\Lambda)\to\mathbf{C}$, its Fourier decomposition

$$H(g)=\sum_{\sigma\in\hat{G}} H_\sigma\sigma(g)\,,\quad g\in P_2(\Lambda)$$

in terms of the elements of the character group $\hat{G}$ of $G=P_2(\Lambda)$ is just the usual decomposition of H into a sum of products of spin matrices, since the characters of G are products of characters of Z_2. Define the set of nonzero interactions

$$B=\{\sigma\in G\,|\,H_\sigma\neq 0\}\,.$$

Dual lattices are constructed with the set B. Defining $P_2(B)$ as the group of functions from B to Z_2, let p be the group homomorphism

$$p:P_2(B)\to G\ \text{ by }\ p(f)=\prod_{\sigma\in B}f(\sigma)^\sigma$$

and denote its kernel by K_p. Suppose X is any set which generates K_p as a group. Then X defines a dual of Λ, with configuration space $G^*=P_2(X)$ and dual Hamiltonian H^* defined as follows. Let

$$q:B\to B^*\subset G^*\ \text{ by }\ q(\sigma):h\to h(\sigma)$$

for $\sigma\in B$ and $h\in X\subset P_2(B)$. $q(\sigma)$ is indeed a character on G^*, and these $q(\sigma)$ are to be the nonzero interactions of the dual. The coefficients $H_{q(\sigma)}$ are given by

$$H_{q(\sigma)}=\tfrac{1}{2}\beta\log\prod_{\substack{\sigma'\in B\\ q(\sigma')=q(\sigma)}}\tanh\beta H_{\sigma'}\,,\tag{1}$$

and

$$H^*=\sum_{q(\sigma)}H_{q(\sigma)}q(\sigma)\,.$$

In most models of physical interest, q is one–one, except perhaps near the boundary. Thus

$$H_{q(\sigma)}=\frac{1}{2\beta}\log\tanh\beta H_\sigma$$

except near the boundary, where (1) must be used.

The partition functions $Z(\beta H)=\sum_{g\subset G}\exp(-\beta H(g))$ of G and $Z(\beta H^*)=\sum_{g\subset G}\exp(-\beta H^*(g))$ of its duals G^* are related then by

$$Z(\beta H)=\frac{N(K_t^*)}{N(G)}\prod_{\sigma\in B}[\sinh(-\beta H_\sigma)\cosh(-\beta H_\sigma)]^{1/2}Z(\beta H^*)$$

where $N(S)$ is the cardinality of S, and K_t^* is defined after Eq. (2).

2. THE INTERACTION FUNCTION

The correlation functions $\rho(\sigma)$ of G are defined by

$$\rho(\sigma)=Z(\beta H)^{-1}\sum_{g\subset G}\exp(-\beta H(g))\sigma(g)\,,\quad\sigma\in\hat{G}$$

with H^* replacing H for the correlation functions $\rho(\sigma^*)$ of G^*, $\sigma^*\in\hat{G}^*$. Note that $\rho(\sigma)=0$ if σ is not a product of elements of B.[1]

Define the characteristic projection $t:G^*\to P_2(B^*)$ by

$$t(g^*):\sigma\to\tfrac{1}{2}(1-\sigma(g^*))\,,\quad\sigma\in B^*\,.\tag{2}$$

The support of $t(g^*)$ is precisely those characters $\sigma\in B^*$ whose value at g^* is -1. Now if the kernel and range of t are denoted, respectively, by K_t^* and R_t^*, then the map $Q:K_p\to P_2(B^*)$ given by

$$Q(f)(q(\sigma))=f(\sigma)\,,\quad\sigma\in B\,,$$

is a group isomorphism $K_p\to R_t^*$. In particular, $f\in K_p$,

$$f(\sigma)=\begin{cases}1, & \sigma\in S,\\ 0, & \sigma\in Y-S,\end{cases}\quad\text{if and only if } Q(f)\in R_t^*\,,$$

$$Q(f)(q(\sigma))=\begin{cases}1, & \sigma\in S,\\ 0, & \sigma\in Y-S,\end{cases}$$

and then

$$\prod_{\sigma\in f^{-1}(1)} \tanh(-\beta H_\sigma) = \prod_{\sigma\in Q(f)^{-1}(1)} \exp(2\beta H_\sigma).$$

Let the symbol $\sum_f(S,T)$ with $S,T\subset B$ indicate that the summation [over $f\in P_2(B)$, say] is to be restricted to f satisfying $f(\sigma)=0$ if $\sigma\in S$, $f(\sigma)=1$ if $\sigma\in T$. Then the interaction function $u_{H^*}(A,C)$ is given by

$$u_{H^*}(A,C) = \sum_{f\in R_t^*}^{(A,C)} \prod_{\sigma\in f^{-1}(1)} \exp(2\beta H_\sigma)$$

for $A,C\subset B^*$.

We wish to evaluate u_{H^*} in terms of the correlation functions of G^*. Thus, suppose Y and W are any disjoint subsets of B^*. Writing $\overline{Y}$ for $\pi\sigma$, $\sigma\in Y$, etc., obtain from (2):

$$\left(\prod_{\sigma\in B^*} e^{-\beta H\sigma}\right)^{-1} Z^*\rho(\overline{Y\cup W})$$

$$= \sum_{g\in G^*} \prod_{\sigma\in B^*} e^{-\beta H_\sigma(\sigma(g)-1)} \prod_{\sigma'\in Y\cup W} \sigma'(g)$$

$$= N(K_t^*) \sum_{f\in R_t^*} \prod_{\sigma\in f^{-1}(1)} e^{2\beta H\sigma} \prod_{\sigma'\in Y} (-2f(\sigma')+1)$$

$$\times \prod_{\sigma''\in W} (2-2f(\sigma'')-1).$$

Now, expanding the product

$$\prod_{\sigma'\in Y}(-2f(\sigma')+1) = \sum_{L\subset f^{-1}(1)\cap Y} (-2)^{N(L)},$$

and similarly with $\Pi_{\sigma''\in W}(2-2f(\sigma'')-1)$, this becomes

$$N(K_t^*) \sum_{\substack{L\subset Y\\ M\subset W}} u_{H^*}(M,L)(-2)^{N(L)+N(M)}(-1)^{N(W)}.$$

Therefore, with a change in summation variable,

$$\left(\prod_{\sigma\in B^*} e^{-\beta H_\sigma}\right)^{-1} Z^* \sum_{\substack{Y\subset C\\ W\subset A}} (-1)^{N(Y)}\rho(\overline{Y\cup W})$$

$$= N(K_t^*) \sum_{\substack{L\subset C\\ M\subset A}} \sum_{\substack{Z\subset C-L\\ V\subset A-M}} u_{H^*}(M,L)2^{N(L)+N(M)}(-1)^{N(Z)+N(V)}$$

$$= 2^{N(A)+N(C)} N(K_t^*) u_{H^*}(A,C), \tag{3}$$

which gives the desired expression.

3. DUAL CORRELATION FUNCTIONS

The interaction functions can be used to derive directly the duality relations for correlation functions. Let $Y\subset B$. Then, using

$$\exp(-\beta H_\sigma(g)) = \cosh(-\beta H_\sigma) + \sigma(g)\sinh(-\beta H_\sigma)$$

and the orthonormality of the characters,

$$Z\rho(\overline{Y}) = N(G) \prod_{\sigma\in B} \cosh(-\beta H_\sigma) \sum_{\substack{f\in P_2(B)\\ p(f)=Y}} \prod_{\sigma'\in f^{-1}(1)} \tanh(-\beta H_{\sigma'}).$$

From the one–one correspondence between $f\in P_2(B)$ with $p(f)=\overline{Y}$ and $f'\in K_p$ with

$$f'\colon \sigma \to \begin{cases} f(\sigma), & \text{if } \sigma\notin Y,\\ f(\sigma)+1, & \text{if } \sigma\in Y, \end{cases}$$

the expansion can be written as

$$Z\left(N(G)\prod_{\sigma\in B}\cosh(-\beta H_\sigma)\right)^{-1}\rho(\overline{Y})$$

$$= \sum_{S\subset Y} \sum_{\substack{f\in P_2(B)\\ p(f)=\overline{Y}}}^{(S,Y-S)} \prod_{\sigma'\in f^{-1}(1)} \tanh(-\beta H_{\sigma'})$$

$$= \sum_{S\subset Y} \sum_{f\in K_p}^{(Y-S,S)} \prod_{\sigma'\in f^{-1}(1)} \tanh(-\beta H_{\sigma'}) \prod_{\sigma\in S} [\tanh(-\beta H_\sigma)]^{-1}$$

$$\times \prod_{\sigma\in Y-S} \tanh(-\beta H_\sigma)$$

$$= \sum_{\substack{S\subset Y\\ S^*\cap(Y-S)^*=\phi}} u_{H^*}(Y^*-S^*,S^*) \prod_{\sigma\in S} (\tanh(-\beta H_\sigma))^{-1}$$

$$\times \prod_{\sigma\in Y-S} \tanh(-\beta H_\sigma),$$

where it has been necessary to consider in the sum over S only sets $S\subset Y$ for which $S^*=\{q(\sigma)\,|\,\sigma\in S\}$ and $(Y-S)^*$ are disjoint. Thus, the interaction function expansion (3) gives the general relation between the correlation functions of G and the correlation functions of a dual G^*,

$$\rho_G(\overline{Y}) = \sum_{T^*\subset W^*} \rho_{G^*}(\overline{T}^*)K(W,T^*), \tag{4}$$

where

$$K(W,T^*) = 2^{-N(W^*)} \sum_{\substack{S\subset W\\ S^*\cap(W-S)^*=\phi}} (-1)^{N(S^*\cap T^*)}$$

$$\times \prod_{\sigma\in S} (\tanh(-\beta H_\sigma))^{-1} \prod_{\sigma\in W-S} \tanh(-\beta H_\sigma)$$

for any $W\subset B$ such that $\overline{W}=\overline{Y}$.

In the event that the duality map q is one–one, Eq. (4) simplifies to the path formula of Kadanoff and Ceva.[4] Injectivity of q is equivalent to requiring that the elements of K_p separate the bonds σ of B, and is satisfied, for example, by a hexagonal Ising lattice with periodic boundary conditions, or with an external field at the boundary, but is not satisfied by this lattice with open boundary conditions.

[1] D. Merlini and C. Gruber, J. Math. Phys. **13**, 1814 (1972).
[2] F. Wegner, J. Math. Phys. **12**, 2259 (1971).
[3] W. Greenberg, Commun. Math. Phys. **29**, 163 (1973).
[4] L. P. Kadanoff and H. Ceva, Phys. Rev. B **3**, 3918 (1971).

Extensions of Lie-graded algebras

Hans Tilgner

Mathematische Physik, Freie Universität Berlin, 1 Berlin 33, Germany
(Received 28 February 1977)

For Lie-graded algebras which are generalizations of Lie algebras with respect to a graduation, used recently in physics for the classification of elementary particles, and extension **G** of **F** by **T**, i.e., a short exact sequence $\mathbf{T} > \rightarrow \mathbf{G} \rightarrow > \mathbf{F}$ can be described by a Lie-graded composition on $\mathbf{T} \oplus \mathbf{F}$, which is formulated in terms of a pair of mappings $\partial : \mathbf{F} \rightarrow \mathrm{der}\mathbf{T}$ and $\Delta : \mathbf{F} \times \mathbf{F} \rightarrow \mathbf{T}$. The congruence of two extensions of **F** by **T**, i. e., the equivalence of the corresponding short exact sequences, is related to an equivalence relation on the set $Z^2(\mathbf{F},\mathbf{T})$ of such 2-cocycles (∂,Δ) such that there exists a bijection between the set of congruence classes of extensions of **F** by **T** and the set $H^2(\mathbf{F},\mathbf{T})$ of classes in $Z^2(\mathbf{F},\mathbf{T})$. This generalizes Lie algebraical results which are also known for groups. Examples for two special cases are given: the semidirect sums with Δ trivial and the almost direct sums with ∂ trivial. Both generalize the concepts of tangent and cotangent algebras of Lie algebras and their central extensions with **R**, the latter being used in the Bargmann theory of ray representations of these semidirect sums.

1. INTRODUCTION

An algebra **A** (over the ground field **K**) is called *graded* (more precisely Z_2-*graded* or Z-*graded*, the only cases we are interested in) if (i) the vector space **A** is graded, i.e., there is a direct decomposition $\mathbf{A} = \oplus_i \mathbf{A}_i$ (i in Z_2 or in Z), which in addition is (ii) *compatible* with the algebra composition, i.e.,

$$\mathbf{A}_i \mathbf{A}_k \subset \mathbf{A}_{i+k} \tag{1}$$

(Ref. 1, p. 163). An endomorphism M of the graded vector space $\mathbf{V} = \oplus_i \mathbf{V}_i$ is called *graded* (*of degree* k) if for all $\mathbf{V}_i$,

$$M\mathbf{V}_i \subset \mathbf{V}_{i+k}. \tag{2}$$

Let $\mathrm{end}_k\mathbf{V}$ be the vector space of graded endomorphisms of degree k. M in $\mathrm{end}_k\mathbf{V}$ and N in $\mathrm{end}_l\mathbf{V}$ implies $M \circ N$ in $\mathrm{end}_{k+l}\mathbf{V}$. Hence $\oplus_i \mathrm{end}_i\mathbf{V}$ is a graded associative algebra.

Examples of graded Lie algebras of type $\mathbf{L}_{-2} \oplus \mathbf{L}_{-1} \oplus \mathbf{L}_0 \oplus \mathbf{L}_1 \oplus \mathbf{L}_2$ (rest $\{0\}$) are given by the Lie algebras of the infinitesimal automorphisms of the generalized Siegel domains, as was shown by Kaup, Matsushima, and Ochiai in Ref. 2. In Ref. 3 it was shown that the Lie algebras of infinitesimal conformal transformations and colineations in a pseudoorthogonal vector space, which are nonlinear polynomial transformations, admit a $\mathbf{L}_{-1} \oplus \mathbf{L}_0 \oplus \mathbf{L}_1$ (rest $\{0\}$) graduation.

Now recently in particle physics another graded generalization of Lie algebras, involving anticommutators (symmetric algebra compositions), was introduced. A general mathematical discussion and references to the physical motivation is given by Corwin, Ne'eman, and Sternberg in Ref. 4. Also in a forthcoming paper Sternber and Wolf[5] discuss a differential geometric application of this new structure in the theory of generalized Siegel domains. To avoid confusion with the above notion of graded Lie algebras in the following this structure is called a "Lie-graded algebra."

In Secs. 2 and 3 we generalize the Calabi-,[6] Hochschild-,[7,8] MacLane,[9] Mori,[10] Zassenhaus-[11] structure theory of Lie algebra extensions to Lie-graded algebras, thus getting semidirect sums and a (frequently in Bargmann's theory of ray representations used) type of extensions (sometimes called "factor sets") which we call "almost direct" sums. The examples of the last two sections again are generalizations of Lie algebra (and Lie group) constructions.

2. LIE-GRADED ALGEBRAS

A Lie-graded algebra **L** with composition $[\,,\,]_\pm$ is a graded algebra such that for all x_k in $\mathbf{L}_k$, y_l in $\mathbf{L}_l$, z in **L** L (arbitrary)

(LGA.1) $[x_k, y_l]_\pm = -(-1)^{kl}[y_l, x_k]_\pm$ (*graded antisymmetry*),

(LGA.2) $[[x_k, y_l]_\pm, z]_\pm = [x_k, [y_l, z]_\pm]_\pm - (-1)^{kl}[y_l, [x_k, z]_\pm]_\pm$ (*graded Jacobi identity*).

Introducing the left multiplication $\mathrm{ad}^\pm(x)y = [x, y]_\pm$, the graded Jacobi identity is equivalent to

(LGA.3) $\mathrm{ad}^\pm([x_k, y_l]_\pm) = \mathrm{ad}^\pm(x_k)\mathrm{ad}^\pm(y_l) - (-1)^{kl}\mathrm{ad}^\pm(y_l)\mathrm{ad}^\pm(x_k)$.

Obviously in a Lie-graded algebra $\mathbf{L}_0$ and $\mathbf{L}^{\mathrm{even}} = \oplus_i \mathbf{L}_{2i}$ are Lie subalgebras and $\mathbf{L}^{\mathrm{even}}$ is a graded Lie algebra; trivially any Lie algebra **L** is a Lie-graded algebra by $\mathbf{L} = \mathbf{L}_0$ (rest $\{0\}$).

A graded associative algebra **A** is a Lie-graded algebra with respect to the *graded commutator* $[x_k, y_l]_\pm = x_k y_l - (-1)^{kl} y_l x_k$, especially those of type $\oplus_i \mathrm{end}_i\mathbf{V}$ for some graded vector space **V**. It will be written $\mathrm{end}^\pm\mathbf{V}$ in the following. (LGA.3) shows that for a Lie-graded algebra **L** the set of left multiplications $\mathrm{ad}^\pm(\mathbf{L})$ is a Lie-graded subalgebra of $\mathrm{end}^\pm\mathbf{L}$. Another class of examples of Lie-graded algebras is given by the set of graded derivations of a graded algebra **A**: D in $\mathrm{end}_k\mathbf{A}$ is a *graded derivation* (*of degree* k) of **A** if for all y_l in $\mathbf{A}_l$, z in **A**,

$$D(y_l z) = (Dy_l) + (-1)^{kl} y_l Dz.$$

Given two graded derivations of degree k resp. l the graded commutator of them is a graded derivation of degree $k+l$. Hence the graded derivations $\mathrm{der}^\pm\mathbf{A} = \oplus_i \mathrm{der}_i\mathbf{A}$ of **A** is a Lie-graded subalgebra of $\mathrm{end}^\pm\mathbf{A}$. If in addition the graded algebra is Lie-graded the

graded Jacobi identity

$$\text{(LGA.4)}\quad \mathrm{ad}^{\pm}(x_k)[y_l, z]_{\pm} = [\mathrm{ad}^{\pm}(x_k)y_l, z]_{\pm} + (-1)^{kl}[y_l, \mathrm{ad}^{\pm}(x_k)z]_{\pm}$$

shows that $\mathrm{ad}^{\pm}(\mathbf{L})$ is a Lie-graded subalgebra (and even an ideal) of $\mathrm{der}^{\pm}\mathbf{L}$. The elements of $\mathrm{ad}^{\pm}(\mathbf{L})$ may be called *inner* derivations, other graded derivations *outer*.

A morphism of Lie-graded algebras is a morphism of the algebra structure which in addition is compatible with the graduation (the same type of graduation being assumed on both algebras). A *representation* of a Lie-graded algebra $\mathbf{L}$ is a morphism into $\mathrm{end}^{\pm}\mathbf{V}$ for some graded vector space $\mathbf{V}$. From (LGA.3) the epimorphism $\mathrm{ad}^{\pm}: \mathbf{L} \to \mathrm{ad}^{\pm}(\mathbf{L})$ is a representation, called the *adjoint* representation of $\mathbf{L}$, the kernel of which is the *center* of $\mathbf{L}$, i.e., the set of all x in $\mathbf{L}$ such that $[x, y]_{\pm} = 0$ for all y in $\mathbf{L}$. $\mathbf{L}$ is called *graded-commutative* if it equals its center and obviously the center is a graded commutative ideal in $\mathbf{L}$. (LGA.3) shows that the restrictions to $\mathbf{L}_i$ of the adjoint representation $x_0 \mapsto \mathrm{ad}^{\pm}(x_0)|_{\mathbf{L}_i}$ for all i are representations of the Lie algebra $\mathbf{L}_0$, the first one being the adjoint representation of $\mathbf{L}_0$. More examples, especially those Lie-graded algebras of the Virasoro type, are given in Ref. 4.

Given a graded vector space $\mathbf{V}$ and a bilinear form $\langle\ ,\ \rangle$ on $\mathbf{V}$, a graded endomorphism Φ of degree i of $\mathbf{V}$ is called a *graded derivation of* $(\mathbf{V}, \langle\ ,\ \rangle)$ *of degree* i if for all x_k in $\mathbf{V}_k$, z in $\mathbf{V}$,

$$\langle \Phi(x_k), z\rangle + (-1)^{ik}\langle x_k, \Phi(z)\rangle = 0.$$

In this case $\langle\ ,\ \rangle$ is called *invariant under* Φ. Given two such graded derivations of $(\mathbf{V}, \langle\ ,\ \rangle)$ of degree k and l respectively, their graded commutator is a graded derivation of $(\mathbf{V}, \langle\ ,\ \rangle)$ of degree $k+l$. Hence the graded derivations $\mathrm{der}^{\pm}(\mathbf{V}, \langle\ ,\ \rangle)$ of $(\mathbf{V}, \langle\ ,\ \rangle)$ are a Lie-graded subalgebra of $\mathrm{end}^{\pm}(\mathbf{V})$. *Graded-symmetric*, resp. *graded-skew* bilinear forms, i.e., bilinear forms with $\langle x_k, y_l\rangle = \pm(-1)^{kl}\langle y_l, x_k\rangle$, may be constructed on a finite-dimensional graded vector space $\mathbf{V}$ by taking the matrix $\mathrm{diag}(\cdots, I_i, \cdots)$ as matrix of $\langle\ ,\ \rangle$ in some basis, where I_i is a square $\dim\mathbf{V}_i$-dimensional matrix with the desired symmetry properties, the sums $\oplus$ being now $\langle\ ,\ \rangle$ orthogonal.

3. COCYCLE SUMS OF LIE-GRADED ALGEBRAS

Given two Lie-graded algebras $\mathbf{T}$ and $\mathbf{F}$ (over the same field, with the same graduation, both compositions being written $[\ ,\ ]_{\pm}$), we consider a pair of mappings

$$\partial: f \mapsto \partial_f,\quad \partial: \mathbf{F} \to \mathrm{der}^{\pm}\mathbf{T} \qquad \text{(linear)}$$

$$\Delta: (f, g) \mapsto \Delta(f, g),\quad \Delta: \mathbf{F}\times\mathbf{F} \to \mathbf{T} \qquad \text{(bilinear)}$$

which are *graded*, i.e., $\partial(\mathbf{F}_i) \subset \mathrm{der}_i\mathbf{T}$ and $\Delta(\mathbf{F}_i, \mathbf{F}_k) \subset \mathbf{T}_{i+k}$. Such a pair (∂, Δ) is called a 2-*cocycle of* $\mathbf{F}$ *with values in* $\mathbf{T}$ if for all t in $\mathbf{T}_i$, f in $\mathbf{F}_k$, g in $\mathbf{F}_l$, h in $\mathbf{F}_m$,

$$\text{(C.1)}\quad \Delta(f, g) = -(-1)^{kl}\Delta(g, f),$$

$$\text{(C.2)}\quad \partial_{[f,g]_{\pm}}(t) = \partial_f\circ\partial_g(t) - (-1)^{kl}\partial_g\circ\partial_f(t) - [\Delta(f, g), t]_{\pm},$$

$$\text{(C.3)}\quad \Delta([f, g]_{\pm}h) - (-1)^{(k+l)m}\partial_h(\Delta(f, g)) = \Delta(f, [g, h]_{\pm}) - (-1)^{kl}\Delta(g, [f, h]_{\pm}) + \partial_f(\Delta(g, h)) - (-1)^{kl}\partial_g(\Delta(f, h)).$$

Obviously ∂ is a morphism of Lie-graded algebras if and only if Δ maps into the center of $\mathbf{T}$. Let $Z^2(\mathbf{F}, \mathbf{T})$ be the set of such 2-cocycles. The motivation of these axioms is given by the following lemma.

Lemma 1: If $\mathbf{G}_i = \mathbf{T}_i \oplus \mathbf{F}_i$, for every (∂, Δ) in $Z^2(\mathbf{F}, \mathbf{T})$ the composition $[t\oplus f, u\oplus g]_{\pm} = [t, u]_{\pm} + \partial_f(u) - (-1)^{kl}\partial_g(t) + \Delta(f, g) \oplus [f, g]_{\pm}$ with t in $\mathbf{T}_k$, u in $\mathbf{T}_l$, f in $\mathbf{F}_k$, g in $\mathbf{F}_l$, defines a Lie-graded algebra on $\oplus_i\mathbf{G}_i = \mathbf{G}$.

Here the graded skew symmetry follows from (C.1) and the graded Jacobi identity from (C.2) and (C.3). Linear continuation gives the same composition for now t in $\mathbf{T}_k$, u in $\mathbf{T}_l$, f in $\mathbf{F}_m$, and g in $\mathbf{F}_r$, with the right-hand side elements in $\mathbf{T}_{k+l}, \mathbf{T}_{m+l}, \mathbf{T}_{r+k}, \mathbf{T}_{m+r}$, and $\mathbf{F}_{m+r}$ respectively. Obviously $\partial_f(u) = [f, u]_{\pm} = \mathrm{ad}^{\pm}_{\mathbf{G}}(f)u$. In the following we write $\mathbf{G} = \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F}$ and call this algebra the 2-*cocycle sum of* $\mathbf{T}$ *and* $\mathbf{F}$ *with respect to* (∂, Δ), shorter the (∂, Δ)-*sum of* $\mathbf{T}$ *and* $\mathbf{F}$. $Z^2(\mathbf{F}, \mathbf{T})$, always contains the *direct* 2-cocycle $(f \to 0, (f, g) \to 0)$, the corresponding sum being the *direct sum of* $\mathbf{T}$ *and* $\mathbf{F}$. Given a morphism $\partial: \mathbf{F} \to \mathrm{der}^{\pm}\mathbf{T}$, the 2-cocycle $(\partial, (f, g) \mapsto 0)$ will be called ∂-*semidirect*, the corresponding sum being the ∂-*semidirect sum*. A 2-cocycle $(f \mapsto 0, \Delta)$ will be called Δ-*almost-direct*, the corresponding sum the Δ-*almost-direct sum*.

Lemma 2: (a) the linear mapping $\chi: t \to t\oplus 0$, $\chi: \mathbf{T} \to \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F}$ is an isomorphism onto an ideal of $\mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F}$,

(b) *the linear mapping* $\psi: t\oplus f \mapsto f$, $\psi: \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F} \to \mathbf{F}$ *is an epimorphism whose kernel is the image of* χ.

We omit the simple proof. A (∂, Δ)-sum of Lie-graded algebras hence leads to a short exact sequence of Lie-graded algebras

$$\mathbf{T} \overset{\chi}{\rightarrowtail} \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F} \overset{\psi}{\twoheadrightarrow} \mathbf{F}.$$

Conversely, following Mac Lane's construction (Ref. 9, p. 124) it is straightforward to show that for any short exact sequence $\mathbf{T} \overset{\chi}{\rightarrowtail} \mathbf{G} \overset{\psi}{\twoheadrightarrow} \mathbf{F}$ of Lie-graded algebras and any *section* $\mu: \mathbf{F} \to \mathbf{G}$, i.e., any graded-linear mapping μ with $\psi\circ\mu$ trivial, there is a 2-cocycle (∂, Δ) of $\mathbf{F}$ with values in $\mathbf{T}$, which however does not depend on the chosen section μ. This correspondence between extensions and 2-cocycle sums makes the search of examples equivalent to that of Lie-graded algebras $\mathbf{G}$ having an ideal $\mathbf{T}$. Some explicit examples will be given below.

There is the well known *split* property of semidirect sums which states that the linear map $\omega: f \mapsto 0\oplus f$, $\omega: \mathbf{F} \to \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F}$ is a (necessarily mono-) morphism (or $\{0\}\oplus\mathbf{F}$ a subalgebra) if and only if Δ vanishes, and the *retract* property of direct sums: $\theta: t\oplus f \mapsto t$, $\theta: \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F} \to \mathbf{T}$ is a (necessarily epi-) morphism if and only if ∂ and Δ vanish.

Two 2-cocycle sums $\mathbf{G} = \mathbf{T}_{\partial} + {}^{\Delta}\mathbf{F}$ and $\tilde{\mathbf{G}} = \mathbf{T}_{\tilde{\partial}} + {}^{\tilde{\Delta}}\mathbf{F}$ of the same Lie-graded algebras $\mathbf{T}$ and $\mathbf{F}$ are said to be *congruent* if there is a (necessarily iso-) morphism

$\rho : \tilde{\mathbf{G}} \to \mathbf{G}$ such that $\rho \circ \chi = \chi$ and $\psi \circ \rho = \psi$, i.e., such that the corresponding short exact sequences are equivalent. Congruence is an equivalence relation on the set of 2-cocycle sums of $\mathbf{T}$ and $\mathbf{F}$. Let $\rho(t_k \oplus f_k) = \Lambda(t_k \oplus f_k) \oplus \Lambda'(t_k \oplus f_k)$. Both Λ's define graded-linear mappings and $f_k = \psi(t_k \oplus f_k) = \psi(\rho(t_k \oplus f_k)) = \Lambda'(t_k \oplus f_k)$; from this $\rho(t_k \oplus f_k) = \rho(\chi(t_k)) + \Lambda(0, f_k) \oplus \Lambda'(0, f_k) = t_k + \Lambda(f_k) \oplus f_k$, where $\Lambda : \mathbf{F} \to \mathbf{T}$ is a graded-linear mapping. It is now straightforward to prove

Lemma 3: (a) The isomorphism $\rho : \tilde{\mathbf{G}} \to \mathbf{G}$ has the form $\rho : t_k \oplus f_k \mapsto t_k + \Lambda(f_k) \oplus f_k$, with $\Lambda : \mathbf{F} \to \mathbf{T}$ graded-linear and

$$\tilde{\partial}_{f_k}(u_l) = \partial_{f_k}(u_l) + [\Lambda(f_k), u_l]_{\pm},$$

$$\tilde{\Delta}(f_k, g_{l\pm}) = \Delta(f_k, g_l) + \partial_{f_k}(\Lambda(g_l)) - (-1)^{kl}\partial_{g_l}(\Lambda(f_k)) + [\Lambda(f_k), \Lambda(g_l)]_{\pm} - \Lambda([f_k, g_l]_{\pm}).$$

(b) Conversely, given a graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$ such that for two 2-cocycles $(\tilde{\partial}, \tilde{\Delta})$ and (∂, Δ) of $\mathbf{F}$ with values in $\mathbf{T}$ the last two equations hold, the mapping ρ in (a) is an isomorphism and the corresponding short exact sequences are equivalent.

The set of congruence classes is denoted by $\mathrm{ext}(\mathbf{F}, \mathbf{T})$. It is not a Lie-graded algebra unless $\mathbf{T}$ is graded-commutative.

4. THE SECOND COHOMOLOGY OF A PAIR OF LIE-GRADED ALGEBRAS

The preceding lemma suggests the following, which can be proven by thorough bookkeeping (the indices of Lemma 3 are dropped).

Lemma 4: Given a graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$, (∂, Δ) in $Z^2(\mathbf{F}, \mathbf{T})$ implies $(\partial^\Lambda, \Delta^\Lambda)$ in $Z^2(\mathbf{F}, \mathbf{T})$, where

$$\partial_f^\Lambda : u \mapsto \partial_f(u) + [\Lambda(f), u]_{\pm},$$

$$\Delta^\Lambda : (f, g) \mapsto \Delta(f, g) + \partial_f(\Lambda(g)) - (-1)^{kl}\partial_g(\Lambda(f)) + [\Lambda(f), \Lambda(g)]_{\pm} - \Lambda([f, g]_{\pm}).$$

Two 2-cocycles (∂, Δ) and $(\tilde{\partial}, \tilde{\Delta})$ for which there is a graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$ such that $\tilde{\partial} = \partial^\Lambda$ and $\tilde{\Delta} = \Delta^\Lambda$ are called *cohomologous*. This defines an equivalence relation on $Z^2(\mathbf{F}, \mathbf{T})$, the set of classes of which $H^2(\mathbf{F}, \mathbf{T})$ is called the *second cohomology space of* $\mathbf{F}$ *with values in* $\mathbf{T}$. If $\mathbf{T}$ is graded-commutative, ∂ remains unchanged. From the above we get a result which was proven first by Zassenhaus for groups (Ref. 11, Sec. 15.1).

Theorem: For any two Lie-graded algebras $\mathbf{T}$ and $\mathbf{F}$ we have:

(a) $\mathbf{T}_\sigma + {}^\Delta\mathbf{F}$ congruent $\mathbf{T}_{\tilde{\partial}} + {}^{\tilde{\Delta}}\mathbf{F} \iff (\partial, \Delta)$ equivalent $(\tilde{\partial}, \tilde{\Delta})$,

(b) $\Gamma : [\mathbf{T}_\partial + {}^\Delta\mathbf{F}]_{\mathrm{conj}} \mapsto [(\partial, \Delta)]_\sim$ is a bijection of $\mathrm{ext}(\mathbf{F}, \mathbf{T})$ onto $H^2(\mathbf{F}, \mathbf{T})$.

2-cocycles which are equivalent to semidirect (resp. direct) ones are called 2-*coboundaries* (resp. *direct* 2-*coboundaries*); the set of 2-coboundaries which does not necessarily contain only one class, is denoted by $B^2(\mathbf{F}, \mathbf{T})$, the single class of direct 2-coboundaries by $B_d^2(\mathbf{F}, \mathbf{T})$. Given a graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$ we write

$$\delta_\partial \Lambda : (f, g) \mapsto \partial_f(\Lambda(g)) - (-1)^{kl}\partial_g(\Lambda(f)) + [\Lambda(f), \Lambda(g)]_{\pm} - \Lambda([f, g]_{\pm}). \tag{3}$$

From (C.2) and the preceding lemma one easily deduces that $(\partial^\Lambda, \delta_\partial \Lambda)$ is in $Z^2(\mathbf{F}, \mathbf{T})$ if and only if $\partial : \mathbf{F} \to \mathrm{der}^\pm\mathbf{T}$ is a morphism. Hence we get the following characterization of 2-coboundaries and direct 2-coboundaries:

(ϕ, Δ) in $B^2(\mathbf{F}, \mathbf{T}) \iff$ there is a morphism $\partial : \mathbf{F} \to \mathrm{der}^\pm\mathbf{T}$ and a graded-linear mapping Λ such that $\partial^\Lambda = \phi$ and $\delta_\partial \Lambda = \Delta$,

(∂, Δ) in $B_d^2(\mathbf{F}, \mathbf{T}) \iff \partial : \mathbf{F} \to \mathrm{der}^\pm\mathbf{T}$ is trivial and $\Delta = \delta_\partial \Lambda$ for a suitable graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$.

The following results are generalizations to $B^2(\mathbf{F}, \mathbf{T})$, resp. $B_d^2(\mathbf{F}, \mathbf{T})$, of well-known results for semidirect, resp. direct, 2-cocycles of Lie algebras:

Lemma 5: Given a graded-linear mapping $\Lambda : \mathbf{F} \to \mathbf{T}$, then:

(a) the graded-linear mapping $\omega_\Lambda : \mathbf{F} \mapsto \mathbf{T}_\partial + {}^\Delta\mathbf{F}$, defined by $\omega_\Lambda : f_k \mapsto \Lambda(f_k) \oplus f_k$ is a (mono-) morphism if and only if (∂, Δ) is in $B^2(\mathbf{F}, \mathbf{T})$,

(b) the graded-linear mapping $\theta_\Lambda : \mathbf{T}_\partial + {}^\Delta\mathbf{F} \to \mathbf{T}$, defined by $\theta_\Lambda : t_k \oplus f_k \mapsto t_k - \Lambda(f_k)$ is an (epi-) morphism if and only if (∂, Δ) is in $B_d^2(\mathbf{F}, \mathbf{T})$, and in this case its kernel is the image of ω_Λ,

(c) the graded-linear mapping $E_\Lambda : \mathbf{T}_\partial + {}^\Delta\mathbf{F} \to \mathbf{T}_\partial + {}^\Delta\mathbf{F}$, defined by $E_\Lambda : t_k \oplus f_k \mapsto \Lambda(f_k) \oplus f_k$ is an endomorphism if and only if (∂, Δ) is in $B^2(\mathbf{F}, \mathbf{T})$.

The simple proof is left to the reader.

Remark: (a) describes 2-*coboundary sums* as short exact split sequences, (b) describes *direct* 2-*coboundary sums* as short exact sequences admitting an inverse short exact sequence. For (c) first note that $E_\Lambda = \omega_\Lambda \circ \psi$; since ψ is onto, E_Λ is a morphism if and only if ω_Λ is a morphism. It is easy to see, that ω_Λ resp. E_Λ are exactly those graded-linear mappings having the *split property* $\psi \circ \cdot = \mathrm{trivial}$. Moreover $E_\Lambda \circ E_\Lambda = \omega_\Lambda \circ \psi \circ \omega_\Lambda \circ \psi = \omega_\Lambda \circ \psi = E_\Lambda$, i.e., E_Λ is idempotent. This generalizes the characterization of semidirect sums by *projectors*, i.e., idempotent endomorphisms, to 2-coboundary sums. The direct sums are characterized by those projectors, the *complementary projectors* of which $g \mapsto g - E(g)$, g in $\mathbf{G}$, are projectors again.

It is easy to verify that $z_m \oplus 0$ is in the center of $\mathbf{T}_\partial + {}^\Delta\mathbf{F}$ if and only if z_m is in the center of $\mathbf{T}$ and ∂ vanishes. Hence the image of χ is in the center of $\mathbf{T}_\partial + {}^\Delta\mathbf{F}$ if and only if ∂ vanishes and $\mathbf{T}$ is graded-commutative, i.e., the almost direct sums with graded-commutative $\mathbf{T}$ correspond exactly to the central extensions of $\mathbf{F}$ by $\mathbf{T}$.

5. EXAMPLES OF SEMIDIRECT SUMS

Let $\mathbf{V} = \oplus_i \mathbf{V}_i$ be a graded vector space, let $\mathbf{F} = \oplus_i \mathbf{F}_i$ be a Lie-graded algebra with the same graduation, and let $\sigma : \mathbf{F} \to \mathrm{end}^\pm\mathbf{V}$ be a morphism of Lie-graded algebras. Then the vector space $\oplus_i(\mathbf{V}_i \oplus \mathbf{F}_i)$ is a Lie-graded alge-

bra with respect to the composition

$$[x_k \oplus f_k, y_l \oplus g_l]_\pm = \sigma(f_k)y_l - (-1)^{kl}\sigma(g_l)x_k \oplus [f_k, g_l]_\pm.$$

Obviously this Lie-graded algebra is the σ-semidirect sum $\mathbf{V}_\sigma + \mathbf{F}$ of the trivial algebra on $\mathbf{V}$ with $\mathbf{F}$. It is called the the σ-*inhomogenization* of $\mathbf{F}$. If $\mathbf{V}$ has a basis and σ is a matrix representation, then $\mathbf{V}_\sigma + \mathbf{F} =: I\mathbf{F}$ has the $(1 + \dim\mathbf{F})$-dimensional representation

$$x_k \oplus f_k \vdash \begin{pmatrix} \sigma(f_k) & x_k \\ 0 & 0 \end{pmatrix}.$$

The adjoint representation $\mathrm{ad}^\pm$ of a Lie-graded algebra $\mathbf{F}$ allows the construction of the $\mathrm{ad}^\pm$-semidirect sum of $\mathbf{F}$ with itself: $T\mathbf{F} := \mathbf{F}_{\mathrm{ad}^\pm} + \mathbf{F} = \oplus_i(\mathbf{F}_i \oplus \mathbf{F}_i)$, where the first $\mathbf{F}$ means the trivial Lie-graded structure on $\mathbf{F}$ has the Lie-graded composition

$$[x_k \oplus f_k, y_l \oplus g_l]_\pm = [f_k, y_l]_\pm - (-1)^{kl}[g_l, x_k]_\pm \oplus [f_k, g_l]_\pm.$$

$T\mathbf{F}$ is called the *tangent algebra of* $\mathbf{F}$. If $\mathbf{F}$ is an algebra of $n \times n$ matrices, then

$$x_k \oplus f_k \vdash \begin{pmatrix} f_k & x_k \\ 0 & f_k \end{pmatrix}$$

is a $2\dim\mathbf{F}$-dimensional representation of $T\mathbf{F}$ in $\mathrm{gl}^\pm(2n, \mathbf{K})$.

These two representations (the *self-*, resp. adjoint, representation) can be described as specializations of a more general construction which for simplicity is given only in matrix form: Let $\mathbf{K}^{n,r}$ be the vector space of rectangular matrices (of n rows and r columns, entries in some ground field $\mathbf{K}$), and let $\sigma : f_k \to \sigma(f_k)$, resp. $\lambda : f_k \to \lambda(f_k)$, be two representations of the Lie-graded algebra $\mathbf{F}$ by $n \times n$, resp. $r \times r$, matrices. Then

$$x_k \oplus f_k \vdash \begin{pmatrix} \sigma(f_k) & x_k \\ 0 & \lambda(f_k) \end{pmatrix} \quad x_k \text{ in } \mathbf{K}_k^{n,r}, \quad f_k \text{ in } \mathbf{F}_k$$

defines a $(n + r)$-dimensional representation (with respect to the graded commutator) of a Lie-graded algebra on the $(nr + \dim\mathbf{F})$-dimensional graded vector space $\oplus_i(\mathbf{K}_i^{n,r} \oplus \mathbf{F}_i)$ the composition of which is defined by linear continuation of

$$[x_k \oplus f_k, y_l \oplus g_l]_\pm = \partial_{f_k}(y_l) - (-1)^{kl}\partial_{g_l}(x_k) \oplus [f_k, g_l]_\pm$$

[the graduation on $\mathbf{K}^{n,r}$ being given by the graduation on $\mathrm{end}^\pm(\mathbf{K}^n \oplus \mathbf{K}^r)$] where

$$\partial_{f_k} : y_l \vdash \sigma(f_k)y_l - (-1)^{kl}y_l\lambda(f_k), \quad y_l \text{ in } \mathbf{K}_l^{n,r},$$

is a graded derivation of the trivial Lie-graded structure on $\mathbf{K}^{n,r}$. Hence we can form the ∂-semidirect sum $\mathbf{K}^{n,r}{}_{\partial *}\mathbf{F}$ (with the above composition) with respect to the morphism

$$\partial : f_k \vdash \partial_{f_k}, \quad \partial : \mathbf{F} \to \mathrm{der}^\pm\mathbf{K}^{n,r} = \mathrm{end}^\pm\mathbf{K}^{n,r}.$$

Choosing $r = 1$ and λ the trivial representation, we get the inhomogenization of $\mathbf{F}$, choosing $r = n$ and $\lambda = \sigma$, we get the tangent algebra of $\sigma(\mathbf{F})$.

Now, given μ_s in $\mathbf{K}_s^{n,r*}$, i.e., a linear form μ_s of $\mathbf{K}^{n,r}$, the *contragredient* representation ∂^* of the representation $\partial : \mathbf{F} \to \mathrm{end}^\pm\mathbf{K}^{n,r}$ is defined as usual by $[\partial_f^*(\mu_s)](y_m) = \mu_s(\partial_f(y_m))$, for y_m in $\mathbf{K}_m^{n,r}$ and f in $\mathbf{F}_k$, in $\mathbf{K}^{n,r*}$. From $(\partial_f \circ \partial_g)^* = \partial_g^* \circ \partial_f^*$ the graded-linear mapping

$$\partial^* : f_k \vdash \partial_{f_k}^*, \quad \partial^* : \mathbf{F}^{\mathrm{op}} \to \mathrm{der}^\pm\mathbf{K}^{n,r*} = \mathrm{end}^\pm\mathbf{K}^{n,r*}$$

is a morphism of the *opposite* algebra $\mathbf{F}^{\mathrm{op}}$ of $\mathbf{F}$ into the trivial Lie-graded algebra $\mathbf{K}^{n,r*}$. The ∂^*-semidirect sum $\mathbf{K}_\partial^{n,r*} + \mathbf{F}^{\mathrm{op}}$ has the Lie-graded composition

$$[\mu_k \oplus f_k, \nu_l \oplus g_l]_\pm = \partial_{f_k}^*(\nu_l) - (-1)^{kl}\partial_{g_l}^*(\mu_k) \oplus [g_l, f_k]_\pm.$$

Identifying $\mathbf{K}^{n,r*}$ with $\mathbf{K}^{r,n}$, this semidirect sum has the $(n + r)$-dimensional representation

$$\mu_k \oplus f_k \vdash \begin{pmatrix} \sigma(f_k)^t & 0 \\ 0 & \lambda(f_k)^t \end{pmatrix},$$

where $\partial_{g_l}^*(\nu_k) = \lambda(g_l)^t\nu_k - (-1)^{kl}\nu_k\sigma(g_l)^t$. The *cotangent algebra*, i.e., the contragredient of the tangent algebra, of Lie algebras is used in the quantization of Kostant and Souriau[12] for the construction of symplectic orbits of Lie groups.

Clearly in the above we can substitute for $\mathbf{K}^{n,r}$, resp. $\mathbf{K}^{r,n}$, any graded subspace invariant under the action of ∂, resp. ∂^*; in the adjoint case this means the choice of an ideal in $\mathbf{F}$.

6. EXAMPLES OF ALMOST DIRECT SUMS

For the above two ∂-, resp. ∂^*-, semidirect sums we construct almost direct sums.

Note that for almost direct sums, (C.2) implies that the image of Δ is in the center of $\mathbf{T}$ and (C.3) simplifies to

$$(\mathrm{C.3'}) \quad \Delta([f_k, g_l]_\pm, h_m) = \Delta(f_k, [g_l, h_m]_\pm) - (-1)^{kl}\Delta(g_l, [f_k, h_m]_\pm),$$

and the Lie-graded composition on $\mathbf{T} + {}^\Delta\mathbf{F}$ to

$$[t_k \oplus f_k, u_l \oplus g_l]_\pm = [t_k, \mu_l]_\pm + \Delta(f_k, g_l) \oplus [f_k, g_l]_\pm.$$

Now let $\langle\ ,\ \rangle$ be a graded-skew bilinear form on $\mathbf{K}^{n,r}$ and let γ_i be in the ith copy of $\mathbf{T} = \cdots \oplus \mathbf{K} \oplus \mathbf{K} \oplus \cdots$. If ∂_{f_k} is in $\mathrm{der}^\pm(\mathbf{K}^{n,r}, \langle\ ,\ \rangle)$ for all f_k in $\mathbf{F}_k$, then the graduation preserving bilinear mapping

$$\Delta : (x_k \oplus f_k, y_l \oplus g_l) \vdash \gamma_{k+l}\langle x_k, y_l \rangle$$

is an almost direct 2-cocycle of $\mathbf{K}^{n,r}{}_{\partial *}\mathbf{F}$ with values in $\mathbf{T}$ and we get the Lie-graded composition

$$[\alpha_k \oplus x_k \oplus f_k, \beta_l \oplus y_l \oplus g_l]_\pm = \gamma_{k+l}\langle x_k, y_l \rangle \oplus \partial_{f_k}(y_l) - (-1)^{kl}\partial_{g_l}(x_k) \oplus [f_k, g_l]_\pm$$

on $\mathbf{T} + {}^\Delta(\mathbf{K}^{n,r}{}_{\partial *}\mathbf{F})$. The same type of almost direct 2-cocycle of $\mathbf{K}_{\partial *}^{r,n} + \mathbf{F}^{\mathrm{op}}$ with values in this $\mathbf{T}$ can be defined under the corresponding assumptions.

Dropping in these two kinds of examples $\mathbf{F}$ completely we get a Lie-graded generalization of the Heisenberg Lie algebra, with the γ-dependent Lie-graded composition

$$[x_k, y_l]_\pm = \gamma_{k+l}\langle x_k, y_l \rangle, \quad x_k \text{ in } \mathbf{K}_k^{n,r}, \quad y_l \text{ in } \mathbf{K}_l^{n,r},$$

[rest zero or from (LGA.1)].

Remark: Bargmann (see Ref. 13 for a Lie group theoretical treatment) has used almost direct sums of $\mathbf{R}$ and a given Lie algebra $\mathbf{F}$ for the classification of (skew-adjoint) ray representations of $\mathbf{F}$. Souriau[12] gives applications in symplectic differential geometry of the second cohomology of some physical motivated Lie algebras with values in its dual, using the co-adjoint representation.

The above examples, specialized to Lie algebras, contain some Lie algebras which were used for the solution of some Schrödinger differential equations in quantum mechanics. Specializing in the inhomogenization of $\mathbf{F}$, resp. $\mathbf{F}^{\mathrm{op}}$, the pair $(\mathbf{V}, \langle\ ,\ \rangle)$ to a $2n$-dimensional symplectic vector space and $\mathbf{F}$ to one of the real one-dimensional algebras of matrices of the form (I is a diagonal $n \times n$ matrix having only ± 1 entries)

$$\begin{pmatrix} 0 & 0 \\ \tau\,\mathrm{id}_n & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -\tau\,\mathrm{id}_n \\ \tau\,\mathrm{id}_n & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -\tau I \\ \tau\,\mathrm{id}_n & 0 \end{pmatrix}, \quad \tau \text{ in } \mathbf{R},$$

or more general any matrix in the symplectic matrix Lie algebra of $(\mathbf{V}, \langle\ ,\ \rangle)$, the $(2n+1)$-dimensional Lie algebras $\mathbf{V}_\partial + \mathbf{F}$ become solvable spectrum-generating Lie algebras for the particle (i.e., the Galilei algebra with the rotations omitted), the harmonic oscillator, the anharmonic oscillator, or more general any dynamical system with a bilinear Hamiltonian in the positions and momentas. The classification of their skew-adjoint ray representations is equivalent to the solution of the corresponding Schrödinger equations. Due to the Bargmann theory this classification is equivalent to the classification of the skew-adjoint (ordinary) representations of the $\langle\ ,\ \rangle$-almost-direct sums of $\mathbf{R}$ and $\mathbf{V}_\partial + \mathbf{F}$, which such are the $(2n+2)$-dimensional quantum-mechanical versionsof these Lie algebras. A general treatment of the second cohomology of non-Abelian Lie algebras was given by Cattaneo in Ref. 14.

[1]N. Jacobson, *Lie Algebras* (Interscience, New York, 1962).
[2]W. Kaup, Y. Matsushima, and T. Ochiai, Am. J. Math. **42**, 475—98 (1970).
[3]H. Tilgner, Int. J. Theor. Phys. **7**, 379—88 (1973).
[4]L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573—603 (1975).
[5]S. Sternberg and J. Wolf, "Graded Lie algebras and bounded homogeneous domains," to appear in Trans. Am. Math. Soc.
[6]L. Calabi, Ann. Mat. Pura Appl. **32**, 295—370 (1951).
[7]G. Hochschild, Am. J. Math. **76**, 698—716 (1954).
[8]G. Hochschild, Am. J. Math. **76**, 763—78 (1954).
[9]S. Mac Lane, *Homology* (Springer, Berlin, 1963).
[10]M. Mori, J. Math. Soc. Jpn. **5**, 171—83 (1953).
[11]M. Hall, *The Theory of Groups* (Macmillan, New York, 1959).
[12]J.M. Souriau, *Structure des systèmes dynamiques* (Dunod, Paris, 1969).
[13]V. Bargmann, Ann. Math. **59**, 1—46 (1954).
[14]U. Cattaneo, Commun. Math. Phys. **20**, 220—44 (1971).

Invariant integrals in nonmetric supersymmetry spaces

L. Kannenberg

Physics Department, University of Lowell, Lowell, Massachusetts 01854
(Received 28 February 1977)

Definitions are given for the exterior product, infinitesimal extension of a cell, Levi-Civita symbol, and generalized Kronecker delta in order to identify invariant integrals in spaces with both Fermi and Bose dimensions. Stokes' and Green's theorems for such spaces are constructed as a preliminary to defining a generalized action principle in supersymmetry spaces not necessarily equipped with a metric or a connection.

Theories of Fermi–Bose supersymmetry provide a promising framework for a unified description of particle interactions.[1,2,3] In particular, the gauge supersymmetry scheme of Arnowitt and Nath[4,5] proposes a unification of gravitational, weak and electromagnetic interactions through the introduction of arbitrary transformations on N "supercoordinates" z^A which obey the generalized commutation relations[6]

$$z^A z^B = (-1)^{ab} z^B z^A. \tag{1}$$

Arnowitt and Nath require that the space identified by these coordinates possess an invariant interval or supermetric

$$ds^2 = dz^A ({}_A g_B)\, dz^B \tag{2}$$

and so construct a field theory remarkable in its formal similarity to classical general relativity. In addition, Arnowitt, Nath, and Zumino[7] have identified the supersymmetry form for the determinant of the metric $({}_A g_B)$ and thus find an action principle based on invariant integrals in this metric supersymmetry space.

Friedman and Srivastava[8] and Wess and Zumino[9] use the concept of supercoordinates to extend the "purely affine" unified field theory of Schrödinger[10] to a natural nonmetric gauge supersymmetry formalism. The present paper addresses the problem of constructing invariant supersymmetry space integrals in the nonmetric formalism and the generalization of Stokes' theorem for spaces with both Fermi and Bose dimensions, not necessarily equipped with either a metric or a connection. This theorem then allows development of an action principle for most gauge supersymmetry theories, including the Arnowitt–Nath formalism, the Friedman–Srivastava "purely affine" theory, and supersymmetry extensions of theories admitting torsion such as that of Hehl.[11,12]

The first requisite for constructing an invariant integral is the definition for the extension of an infinitesimal cell in the space. This flows directly from the form for the exterior product of vectors, conveniently stated as follows: If u^A and v^A are two contravariant vectors, the components of their exterior product are

$$(u \wedge v)^{AB} \equiv u^A v^B - (-1)^{ab} u^B v^A. \tag{3}$$

This structure possesses the two obvious symmetries

$$(u \wedge v)^{AB} = -(-1)^{ab} (u \wedge v)^{BA} \tag{4}$$

and

$$(u \wedge v)^{AB} = -(v \wedge u)^{AB} \tag{5}$$

so that the "diagonal elements" $(u \wedge v)^{XX}$ (no sum on X) vanish identically, and the $(u \wedge v)^{AB}$ transform as components of a contravariant second rank tensor; that is, if the vector transformation law reads

$$u^{A'} = R^{A'}{}_B u^B = u^B ({}_B L^{A'}), \tag{6}$$

then the transformation law for the exterior product is

$$(u \wedge v)^{A'B'} = R^{A'}{}_C (u \wedge v)^{CD} ({}_D L^{B'}). \tag{7}$$

The rule extends immediately to the exterior product of a sequence of vectors $u_{(\alpha)}{}^A$, $\alpha = 1, 2, \ldots, M$, as a sum of terms of the form $u_{(1)}{}^{A_1} u_{(2)}{}^{A_2} \cdots u_{(M)}{}^{A_M}$, which possesses the symmetries

$$(u_{(1)} \wedge \cdots \wedge u_{(M)})^{\cdots AB \cdots} = -(-1)^{ab} (u_{(1)} \wedge \cdots \wedge u_{(M)})^{\cdots BA \cdots} \tag{8}$$

on any adjacent pair of indices, and

$$(\cdots \wedge u_{(\alpha)} \wedge u_{(\beta)} \wedge \cdots)^{A_1 \cdots A_M} = -(\cdots \wedge u_{(\beta)} \wedge u_{(\alpha)} \wedge \cdots)^{A_1 \cdots A_M} \tag{9}$$

on any adjacent pair of vectors. This structure then transforms as a tensor of rank M, as indicated.

The form for the extension of an infinitesimal cell is now clear: consider $M \leq N$ linearly independent infinitesimal excursions from a given point, identifying the elements of the αth such excursion as $d_{(\alpha)} z^A$; the extension of an M-cell at that point is then simply the exterior product of these excursions,

$$d\tau^{A_1 \cdots A_M} = (d_{(1)} z \wedge \cdots \wedge d_{(M)} z)^{A_1 \cdots A_M}. \tag{10}$$

The order in which the excursions appear in this exterior product fixes the orientation of the cell. The extension transforms as a contravariant tensor of rank M, and thus its inner product with a covariant tensor of rank M is a scalar, allowing immediately the construction of invariant integrals, e.g.,

$$I \equiv \int (-1)^{\phi_M(a,a)} T_{A_1 \cdots A_M} d\tau^{A_1 \cdots A_M}, \tag{11}$$

where the parity factor is given as

$$\phi_M(b,a) \equiv \sum_{i=1}^{m-1} \sum_{j=i+1}^{M} b_i a_j. \tag{12}$$

Moreover, if $T_{A_1 \cdots A_{M-1}}$ is a covariant tensor, then, even though its partial derivatives $T_{A_1 \cdots A_{M-1}, A_M}$ fail to transform as tensor components, their inner product with $d\tau^{A_1 \cdots A_M}$ is a scalar, and hence a second category of

invariant integral is

$$J \equiv \int (-1)^{\phi_N(a,a)} T_{A_1 \cdots A_{M-1}, A_M} d\tau^{A_1 \cdots A_M}. \tag{13}$$

Stokes' theorem relates an integral of the second type over a closed oriented region in a subspace V_M to an integral of the first type over the boundary of the region.

In constructing Stokes' theorem it is useful to have supersymmetry space definitions for the Levi-Civita symbol and the generalized Kronecker deltas. The natural choice for the Levi-Civita symbol is that structure $\epsilon_{A_1 \cdots A_N}$ with Grassmann parity $(-1)^f$, f being the number of Fermi dimensions of the space, with $\epsilon_{12\ldots N} = +1$ and possessing the symmetry

$$\epsilon_{\ldots AB\ldots} = -(-1)^{ab} \epsilon_{\ldots BA\ldots} \tag{14}$$

for any adjacent pair of indices. In addition the Levi-Civita symbol must vanish if any two of its indices are the same, a condition not forced by Eq. (14) for Fermi indices. The superscripted Levi-Civita symbol has the same parity and symmetry structure as the subscripted symbol, and it is convenient to complete the algebraic properties of both with the rule

$$\epsilon_{12\ldots N} \epsilon^{12\cdots N} = +1 = (-1)^f \epsilon^{12\cdots N} \epsilon_{12\ldots N}. \tag{15}$$

The generalized Kronecker N-delta then follows as a combination of Levi-Civita symbols,

$$(-1)^{f(f-1)/2} D_{A_1 \cdots A_n}^{\quad B_1 \cdots B_n} \equiv (-1)^{\phi_n(b,a)} \epsilon_{A_1 \cdots A_n} \epsilon^{B_1 \cdots B_N}. \tag{16}$$

This structure exhibits the symmetries

$$D_{\ldots A_1 \ldots A_2 \ldots}^{\quad \cdots B_1 \cdots B_2 \cdots} = -(-1)^{b_1 b_2 + a_2(b_1+b_2)} D_{\ldots A_1 \ldots A_2 \ldots}^{\quad \cdots B_2 \cdots B_1 \cdots} \tag{17}$$

and

$$D_{\ldots A_1 \ldots A_2 \ldots}^{\quad \cdots B_1 \cdots B_2 \cdots} = -(-1)^{a_1 a_2 + b_1(a_1+a_2)} D_{\ldots A_2 \ldots A_1 \ldots}^{\quad \cdots B_1 \cdots B_2 \cdots} \tag{18}$$

and equals unity when $A_i = i = B_i$, $i = 1, 2, \ldots, N$. From this follows the definition of the generalized M-delta,

$$(N-M)!\, D_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M} = D_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M C_1 \cdots C_{N-M}}, \tag{19}$$

which has the same symmetries as does the N-delta.

The general reduction theorem relating generalized Kronecker deltas is the familiar form

$$(N-K)!\, D_{A_1 \cdots A_K}^{\quad B_1 \cdots B_K} = (N-M)!\, D_{A_1 \cdots A_K}^{\quad B_1 \cdots B_K C_1 \cdots C_{M-K}}. \tag{20}$$

It is possible to display $D_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M}$ in a form in which its leading term is the structure $\delta_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M}$ formed of products of the ordinary Kronecker delta $\delta_A{}^B$ with the symmetries of $D_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M}$. Additional terms must of course be appended since $\delta_{\ldots XX \ldots}^{\quad \cdots B_i B_j \cdots}$ (no sum on X) fails to vanish by itself if X is a Fermi index.

Observe that if $T^{A_1 \cdots A_M}$ is a tensor with the symmetries of an exterior product, then the identity

$$M!\, T^{B_i \cdots B_M} = (-1)^{\phi_M(a+b,b)+p_T} T^{A_1 \cdots A_M} D_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M} \tag{21}$$

obtains, where $p_T \equiv \sum_{i=1}^{M} a_i$ is the Grassmann parity of $T^{A_1 \cdots A_M}$. From this identity flows the natural definition for the dual ${}^*T_{B_1 \cdots B_{N-M}}$ of the tensor $T^{A_1 \cdots A_M}$:

$$M!\, {}^*T_{B_1 \cdots B_{N-M}} \equiv (-1)^{\phi_M(a,a)+p_T} \times T^{A_1 \cdots A_M} \epsilon_{A_1 \cdots A_M B_1 \cdots B_{N-M}}. \tag{22}$$

The inversion of this relation is just

$$T^{A_1 \cdots A_M} = (N-M)!\, (-1)^{f(f-1)/2 + \phi_{N-M}(b,b) + p_T \tilde{p}_T} \times {}^*T_{B_1 \cdots B_{N-M}} \epsilon^{A_1 \cdots A_M B_1 \cdots B_{N-M}}, \tag{23}$$

$\tilde{p}_T \equiv \sum_{i=1}^{N-M} b_i$ being the Grassmann parity of the dual. The N-delta satisfies the product rule

$$(-1)^{\phi_N(b+c,a+c)+\Sigma c_i} D_{A_1 \cdots A_N}^{\quad C_1 \cdots C_N} D_{C_1 \cdots C_N}^{\quad B_1 \cdots B_N} = N!\, D_{A_1 \cdots A_N}^{\quad B_1 \cdots B_N} \tag{24}$$

as may be checked directly from its definition in terms of the Levi-Civita symbols. Precisely similar relations apply for the M-deltas. Furthermore the related structure

$$(-1)^{f(f-1)/2} D^{A_1 \cdots A_N}_{\quad B_1 \cdots B_N} \equiv (-1)^{\phi_N(b,a)} \epsilon^{A_1 \cdots A_N} \epsilon_{B_1 \cdots B_N} \tag{25}$$

satisfies

$$(-1)^{\phi_N(b+c,a+c)} D^{A_1 \cdots A_N}_{\quad C_1 \cdots C_N} D^{C_1 \cdots C_N}_{\quad B_1 \cdots B_N} = N!\, D^{A_1 \cdots A_N}_{\quad B_1 \cdots B_N} \tag{26}$$

which may be verified by observing that it may be expressed in terms of the N-delta according to

$$(-1)^{\phi_N(a,b)+f} D^{A_1 \cdots A_N}_{\quad B_1 \cdots B_N} = (-1)^{\phi_N(b,a)} D_{B_1 \cdots B_N}^{\quad A_1 \cdots A_N}. \tag{27}$$

Now consider the mixed tensor $T_A{}^B$ and define its 2-product by

$$2!\, T_{A_1 A_2}{}^{B_1 B_2} \equiv (-1)^{\phi_2(b+c,a+c)+\phi_2(b+d,c+d)+\Sigma c_i + \Sigma d_i} \times D_{A_1 A_2}^{\quad C_1 C_2} T_{C_1}{}^{D_1} T_{C_2}{}^{D_2} D_{D_1 D_2}^{\quad B_1 B_2}. \tag{28}$$

This structure is a tensor of type indicated with exactly the symmetries of the 2-delta. Furthermore, if $T_A{}^B$ is the "matrix product" of two other tensors $U_A{}^B$ and $V_A{}^B$, that is if

$$T_A{}^B = (-1)^C U_A{}^C V_C{}^B, \tag{29}$$

then it follows at once that their 2-products are related according to

$$2!\, T_{A_1 A_2}{}^{B_1 B_2} = (-1)^{\phi_2(b+c,a+c)+\Sigma c_i} U_{A_1 A_2}^{\quad C_1 C_2} V_{C_1 C_2}^{\quad B_1 B_2}. \tag{30}$$

More generally, the M-product of the tensor $T_A{}^B$ is defined by

$$M!\, T_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M} = (-1)^{\phi_M(b+c,a+c)+\phi_M(b+d,c+d)+\Sigma c_i+\Sigma d_i} \times D_{A_1 \cdots A_M}^{\quad C_1 \cdots C_M} T_{C_1}{}^{D_1} \cdots T_{C_M}{}^{D_M} D_{D_1 \cdots D_M}^{\quad B_1 \cdots B_M} \tag{31}$$

which has the symmetries of the M-delta. If $T_A{}^B$ is the matrix product of $U_A{}^B$ and $V_A{}^B$ in the sense of Eq. (29), then their M-products are related according to the rule

$$M!\, T_{A_1 \cdots A_M}^{\quad B_1 \cdots B_M} = (-1)^{\phi_M(b+c,a+c)+\Sigma c_i} U_{A_1 \cdots A_M}^{\quad C_1 \cdots C_M} V_{C_1 \cdots C_M}^{\quad B_1 \cdots B_M} \tag{32}$$

as may be checked directly. Using the symmetry properties it is clear that the N-product of $T_A{}^B$ has a single independent element, which is convenient to take as

$$T \equiv T_{12\ldots N}^{12\ldots N} \tag{33}$$

and on using the product rule for N-products of $T_A{}^B$, $U_A{}^B$, and $V_A{}^B$, one finds directly

$$T = UV. \tag{34}$$

An alternative form for the N-product which is very

useful is just

$$T_{A_1\cdots A_N}^{B_1\cdots B_N} = TD_{A_1\cdots A_N}^{B_1\cdots B_N}. \tag{35}$$

The definition of M-products of rank two tensors of other type, e.g., $U^A{}_B$ or $({}_Ag_B)$, follows the pattern indicated above. It is also convenient to define M-products of the transformation matrix $R^A{}_B$, in the same way, and to write

$$R \equiv R_{1'2'\cdots N'}^{12\cdots N}. \tag{36}$$

A consequence of these definitions is that if $({}_Ag_B)$ is a covariant tensor of type indicated, with transformation law

$$({}_{A'}g_{B'}) = ({}_{A'}L^C)({}_Cg_D)R^D{}_{B'}, \tag{37}$$

then the quantity

$$g \equiv ({}_{12\cdots N}g_{12\cdots N}) \tag{38}$$

transforms according to

$$g' = LgR. \tag{39}$$

From the relation

$$({}_{A'}L^B) = (-1)^{a+ab}R^B{}_{A'} \tag{40}$$

follows at once $L = R$, so than an alternative form for the transformation law of g is

$$(\epsilon(g')g')^{1/2} = \epsilon(R)R(\epsilon(g)g)^{1/2}, \tag{41}$$

$\epsilon(X)$ being the indicatrix of X. It is also clearly a straightforward task to use the relation

$$R^{A'}{}_CR^C{}_{B'} = \delta^{A'}{}_{B'} \tag{42}$$

to discover that

$$R^{1'2'\cdots N'}{}_{12\cdots N} = R^{-1}. \tag{43}$$

By writing the N-product of $R^A{}_{B'}$ in the form

$$RD^{A_1\cdots A_N}{}_{B_1'\cdots B_N'} = (-1)^{\phi_N(b+c,a+c)}D^{A_1\cdots A_N}{}_{C_1\cdots C_N}R^{C_1}{}_{B_1'}\cdots R^{C_N}{}_{B_N'}, \tag{44}$$

it is possible to read the transformation law of the Levi-Civita symbol off directly as

$$\epsilon_{A_1'\cdots A_N'} = (-1)^{\phi_N(a+b,b)}R^{-1}\epsilon_{B_1\cdots B_N}R^{B_1}{}_{A_1'}\cdots R^{B_N}{}_{A_N'} \tag{45}$$

which is reasonable to identify as the transformation law of a right covariant tensor density of rank N and weight -1, provided R is defined as the determinant of the transformation matrix. This definition is consistent with the general property of N-products exhibited in Eq. (34).[13]

The above groundwork prepares the way for Stokes' theorem. Let $T_{A_1\cdots A_{N-1}}$ be a tensor with the symmetries and transformation properties of the exterior product of $N-1$ right covariant vectors, defined throughout a finite region V of the space identified by the supersymmetry coordinates z^A: Then

$$\int_V(-1)^{\phi_N(a,a)}T_{A_1\cdots A_{N-1},A_N}d\tau^{A_1\cdots A_N} = \int_S(-1)^{\phi_{N-1}(a,a)}T_{A_1\cdots A_{N-1}}d\tau^{A_1\cdots A_{N-1}}. \tag{46}$$

S is the boundary surface to V, and the orientations of the N-cell and the $(N-1)$-cell are related in the usual way; that is, if their duals are respectively $*d\tau$ and $*d\tau_B$, and if n^B is an infinitesimal vector directed out of V, then the inner product $(-1)^bn^B*d\tau_B$ has the same sign as does $*d\tau$. The proof of this result now proceeds precisely in accord with its analog in spaces without Fermi dimensions,[14] the sole new feature being the additional accounting necessary from the appearance of the parity factors. Thus in the coordinates chosen, select cells in V with edges along the coordinate lines, taken in order with the Bose coordinates first, from which is established immediately the usual relation

$$d\tau^{A_1\cdots A_N} = \epsilon^{A_1\cdots A_N}dz^1dz^2\cdots dz^N. \tag{47}$$

The integral over V appears as a sum of terms of the form

$$J_B = \int_V((-1)^{\phi_{N-1}(a,a)+bT^{(1+b)}}\epsilon^{A_1\cdots A_{N-1}B} \times T_{A_1\cdots A_{N-1}}),_Bdz^1dz^2\cdots dz^N, \tag{48}$$

no sum on B, and clearly $A_i \neq B$. One has

$$dz^1\cdots dz^N = (-1)^{k(f-k)\theta(k)}dz^Bdz^{B+1}\cdots dz^Ndz^1\cdots dz^{B-1}, \tag{49}$$

where $k \equiv B+f-N-1$, and hence the integral J_B can be taken along tubes made up of parametric lines of z^B. Assuming for the moment that V is a convex region, the integral reduces to

$$J_B = (\int_{(2)} - \int_{(1)})(-1)^{\phi_{N-1}(a,a)}T_{A_1\cdots A_{N-1}} \times(-1)^{k(f-k)\theta(k)}\epsilon^{A_1\cdots A_{N-1}B}dz^{B+1}\cdots dz^{B-1}, \tag{50}$$

where the limits indicated are the points at which the parametric lines of z^B cut the boundary S of V, with $z^B(2) > z^B(1)$. Appropriate outwardly-directed normals to V at (1) and (2) are $-dz^B$ and dz^B respectively, whence one discovers directly that the orientation of the cell

$$d\tau^{A_1\cdots A_{N-1}} = (-1)^{k(f-k)\theta(k)}\epsilon^{A_1\cdots A_{N-1}B}dz^{B+1}\cdots dz^{B-1} \tag{51}$$

at (2) is correct, and similarly for the corresponding cell at (1). The sum of all integrals J_B thus corresponds precisely to the integral over S indicated. To complete this construction, observe that a nonconvex region may be constructed by conjoining appropriate convex regions in the usual way. The general case of a tensor $T_{A_1\cdots A_{M-1}}$, $M \leq N$, defined over a finite oriented subspace V_M now appears straightforwardly as

$$\int_{V_M}(-1)^{\phi_M(a,a)}T_{A_1\cdots A_{M-1},A_M}d\tau^{A_1\cdots A_M} = \int_S(-1)^{\phi_{M-1}(a,a)}T_{A_1\cdots A_{M-1}}d\tau^{A_1\cdots A_{M-1}} \tag{52}$$

with the usual relation between the orientation of cells in V_M and on its boundary S. The proof proceeds just as before, and may be facilitated by introducing M parameters y^k

$$y^k = y^k(z^1,\cdots,z^{M-1}) \tag{53}$$

as intrinsic coordinates for V_M, and expressing the cells in terms of these intrinsic coordinates.

Green's theorem may be regarded as a restatement of Stokes' theorem in terms of duals, and appears as

$$\int_V(-1)^a(*T^A),_A*d\tau = \int_S(-1)^{a*}T^{A*}d\tau_A. \tag{54}$$

In spaces equipped with a connection, defined through

$$V^A{}_{;B} \equiv V^A{}_{,B} + V^C({}_C\Gamma^A{}_B), \tag{55}$$

it is useful to observe that requiring the covariant derivative of the Levi-Civita symbol to vanish produces the following expression for the covariant derivative of a relative scalar P of weight + 1:

$$P_{;A} = P_{,A} - P(-1)^b({}_B\Gamma^B{}_A). \tag{56}$$

If in addition the connection exhibits the affine symmetry

$$({}_A\Gamma^B{}_C) = (-1)^{a+ab+ac+bc+c}({}_C\Gamma^B{}_A), \tag{57}$$

then the covariant divergence of a relative vector A^A of weight + 1 is identical to its ordinary divergence,

$$A^A{}_{;A} = A^A{}_{,A}. \tag{58}$$

Finally, in a space with a metric $({}_Ag_B)$ as well as a connection one may associate to the dual ${}^*T_{A_1\cdots A_M}$ of a tensor an oriented tensor $T^*{}_{A_1\cdots A_M}$ via

$$T^*{}_{A_1\cdots A_M} = (\epsilon(g)g)^{1/2}{}^*T_{A_1\cdots A_M}. \tag{59}$$

The mathematical machinery exhibited here now permits construction of fairly general action principle formulations for gauge supersymmetry theories. Some of the properties of such formulations will be examined in a later paper.

[1]D.V. Volkov and V.P. Akulov, Phys. Lett. 46, 109 (1973).
[2]J. Wess and B. Zumino, Nucl. Phys. B 70, 39 (1974).
[3]A. Salam and J. Strathdee, Nucl. Phys. B 76, 477 (1974).
[4]R. Arnowitt and P. Nath, Phys. Lett. B 56, 177 (1975).
[5]R. Arnowitt and P. Nath, Gen. Rel. Grav. J. 7, 89 (1976).
[6]The notation of this paper follows that of Arnowitt and Nath, with one exception. Elements of "left covariant" vectors, transforming as the left derivatives of scalars, appear here as $({}_Au)$, that is with the supersymmetry index as a prescript. With this notation the transformation law for a contravariant vector appears as in Eq. (6), the relation between left and right transformation matrix elements being as displayed in Eq. (40).
[7]R. Arnowitt, P. Nath, and B. Zumino, Phys. Lett. B 56, 81 (1975).
[8]M. Friedman and Y.N. Srivastava, Phys. Rev. D 15, 1026 (1977).
[9]J. Wess and B. Zumino, Phys. Lett. B 66, 361 (1977).
[10]E. Schrödinger, *Space—Time Structure* (Cambridge U.P., New York, 1950).
[11]F.W. Hehl, Gen. Rel. Grav. J. 4, 333 (1973).
[12]F.W. Hehl, Gen. Rel. Grav. J. 5, 491 (1974).
[13]Comparison between the definition presented here of T as a determinant and the highly compact expression of Arnowitt, Nath, and Zumino is tedious except in spaces of low dimensionality; but a direct calculation indicates that they are equivalent at least for $N \leq 3$.
[14]The presentation given here is made to follow as closely as possible the analogous proof for the usual tensor calculus appearing in J.L. Synge and A. Schild, *Tensor Calculus* (University of Toronto, Toronto, 1949), a standard textbook. It is doubtless also possible to cast the theorem into a coordinate-independent form in the manner of M. Spivak, *Calculus on Manifolds* (Benjamin, Reading, Massachusetts, 1965).

The equivalence of a one-dimensional turbulence problem and the one-dimensional Coulomb gas

S. T. Chui[a)]

Department of Physics, SUNYA, Albany, New York 12222

H. L. Frisch[a)]

Department of Chemistry, SUNYA, Albany, New York 12222

(Received 29 April 1977)

We show that Burger's equation subject to initial conditions which are governed by a canonical (Gaussian) distribution in the kinetic energy can be related to the properties of a one-dimensional Coulomb gas in a certain limit. Some consequences of this are worked out.

Burger's equation is one of the simplest equations that can be solved exactly which exhibits intermittency. For this reason it has received a lot of interest. The statistical properties of this equation have been investigated by Burger[1] as well as others. It is now generally believed that the basic physics is well understood, but as yet no quantitative scheme has been developed by which one can calculate these statistical properties systematically. In this note we wish to point out that, using a recent $m \to 0$ trick exploited by various workers in the study of polymers,[2] in critical behavior in the presence of impurities,[3] and in spin glass problems,[4] the statistical problem of Burger's equation can be related to a one-dimensional Coulomb gas. Various consequences of this analogy are also discussed.

The problem that we are addressing ourselves to can be precisely stated as follows: Burger's equation is given by

$$U_t + UU_x = \nu U_{xx}. \tag{1}$$

Suppose the "velocity" $U(x,t)$ at time $t=0$ has a canonical distribution in the total energy, i.e.,

$$P(U(x,0)) \propto \exp(-\int U^2 dx/\sigma^2). \tag{2}$$

What are the average values of quantities such as the velocity autocorrelation function $\langle U(0,0)\, U(x,t)\rangle$, especially in the limit as $t \to \infty$? Here the angular bracket denotes the statistical average.

We have been able to show that

$$\langle U(0,0)\, U(x,t)\rangle = \sigma^2 \lim_{m\to 0} Q_1(x,m), \tag{3}$$

where

$$Q_1(x,m) = \int \prod_{i=1}^{m} d\xi_i\, b(x-\xi_1) \exp[-\beta H(m,\{\xi\})]. \tag{4}$$

$-\beta H(m,\xi)$ looks just like the Hamiltonian of a one-dimensional Coulomb gas (of one species) with m particles in an external field. It is given by

$$-\beta H = -\beta H_{\text{ext}} - \beta H_{\text{int}}, \tag{5}$$

where

$$-\beta H_{\text{ext}} = -b\sum_{i=1}^{m}(x-\xi_i)^2 - 2a\sum_{i=1}^{m}|\xi_i| \tag{6}$$

and

$$-\beta H_{\text{int}} = a\sum_{i=1}^{m}|\xi_i - \xi_j| \tag{7}$$

with

$$a = \sigma^2/32\nu^2, \quad b = (4\nu t)^{-1}. \tag{8}$$

The derivation of formula (3) is given in the Appendix. Let us now discuss some of the consequences: First of all, note that even though there has been some concern over the validity of the $m \to 0$ trick in some other applications,[5] this difficulty does not apply here. In that case the worry is about whether one can interchange the $m \to 0$ limit and the thermodynamic limit as the number of particles goes to infinity. In the present case one is interested in the limit as the number of particles goes to zero! Because of this limit one would expect the mean field approximation to be a good one. It is with this approximation that we shall discuss the consequences of our problem.

To get a feeling for what is going on, we have plotted the external potential in the limit of large and small time in Fig. 1. This figure indicates that the particles are localized around $\xi \approx x$ at small time and have "diffused" to the origin as time increases. As we see formula (3), $\langle U(0,0)\, U(x,t)\rangle$ is essentially proportional to the "ensemble average" of $x-\xi$. It must therefore

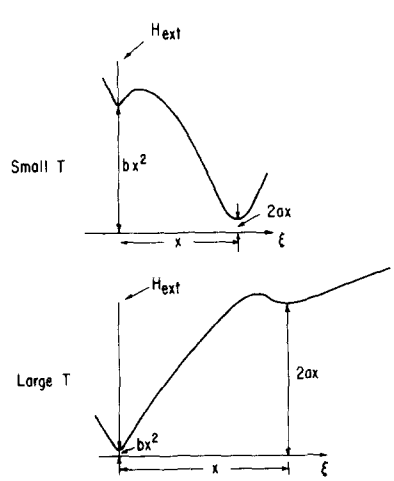


FIG. 1. A figure showing $H_{\text{ext}}(\xi)$ as a function of ξ for long and short times.

a)Supported by NSF grants.

go from zero to a finite value as $t \to \infty$ starting from zero.

For $t=0$, it is obvious that $\langle U(0,0)\, U(x,0)\rangle = 0$ for $x \neq 0$. We shall now evaluate this number in the $t \to \infty$ limit in the mean field approximation. We have, using (3),

$$\langle U(0,0)U(x,t)\rangle = \sigma^2 \lim_{m\to 0} \int \prod_{i=0}^{m} d\xi_i \; b(\xi_1 - x) \times \exp[-b\sum(x-\xi_i)^2 - 2a\sum|\xi_i| + a\sum|\xi_i - \langle\xi\rangle|]. \quad (9)$$

As we have pointed out previously, as $t \to \infty$, we expect $\langle \xi \rangle = 0$. This can also be shown to be actually a self-consistent solution. In this case, one finds straightforwardly from the above relation that

$$\langle U(0,0)U(x,t)\rangle = \sigma^2 x/4\nu t. \quad (10)$$

Note that (10) does not have the form of $f(x/\sqrt{t}\,)$, but is actually dimensionally correct. It is due to the fact the our initial distribution of U corresponds to not just one but a distribution of Reynolds number. We have no simple explanation of this simple-looking result unfortunately.

We also wish to point out that with this technique other types of initial distribution can also be treated. For example, a distribution of the form $P(u) \propto \exp(-\sum k^2 u_k^2/\sigma^2)$ (similar to one recently studied by Forster *et al.*[6]) leads to a gas of particles with δ-function interactions.

APPENDIX

By using the Cole–Hopf[7] transformation, the solution to Burger's equation is given by

$$U(x,t) = -2\nu \frac{d}{dx} \ln \int_{-\infty}^{\infty} d\xi \times \exp\left[-\frac{1}{2\nu}\left(\frac{(x-\xi)^2}{2t} + \int_0^{\xi} U(\xi_1,0)\, d\xi_1\right)\right]. \quad (A1)$$

We have written the solution of this form to bring out its similarity to a corresponding relation in the thermodynamics of amorphous systems. We are interested in the following quantity:

$$\langle U(0,0)U(x,t)\rangle = \int d\{U(x,0)\}\, P\{U(x,0)\}\, U(0,0)U(x,t). \quad (A2)$$

Making use of the formula $\ln x = \lim_{m\to 0}[(x^m - 1)/m]$, we have

$$\langle U(0,0)U(x,t)\rangle = -2\nu \lim_{m\to 0} \frac{1}{m} \int d\{U(x,0)\}\, U(0,0) \frac{d}{dx} \times \left\{\exp\left[-\frac{1}{2\nu}\left(\sum_{i=1}^{m} \frac{(x-\xi_i)^2}{2t} + \sum_i \int_0^{\xi_i} U(\xi',0)\, d\xi'\right)\right] - 1\right\} \times P\{U(x,0)\}. \quad (A3)$$

We now express $U(x,0)$ as a Fourier transform

$$U(x,0) = \sum_k U(k) e^{ikx} \frac{1}{L}. \quad (A4)$$

On substituting (A4), Eq. (A3) becomes

$$\langle U(0,0)U(x,t)\rangle = -2\nu \lim_{m\to 0} \int \prod_k dU_k \frac{\sigma}{\sqrt{L}} \left(\frac{1}{L}\sum_{k'} U_{k'}\right) \frac{d}{dx} \times \left\{\exp\left[\frac{1}{L}\sum_k U(k)\left(\sum_{i=1}^{m} \frac{\exp(i\xi_i k) - 1}{2ik\nu}\right) - \frac{1}{L}\sum_k \frac{U_{(k)}^2}{\sigma^2} + \sum_{i=1}^{m} \frac{(x-\xi_i)^2}{t\nu}\right]\right\}. \quad (A5)$$

The statistical averaging can now be easily done, and we obtain formula (3).

[1] J.M. Burger "Statistical Problems Connected with Asymptotic Solutions of the One-dimensional, Non-linear Diffusion Equation," Lecture Series #52, University of Maryland, Institute for Fluid Dynamics and Applied Mathematics, November 1973.

[2] See, for example, S.F. Edwards, Proc. Phys. Soc. (London) **88**, 265 (1966).

[3] See, for example, V.J. Emery, Phys. Rev. B **11**, 239 (1975).

[4] See, for example, S.F. Edwards and P.W. Anderson, J. Phys. F **5**, 965 (1975).

[5] See, for example, D.J. Thouless, P.W. Anderson, E. Lieb, and R.G. Palmer (to be published).

[6] D. Forster, D.R. Nelson, M.J. Stephen, Phys. Rev. Lett. **36**, 867 (1976).

[7] E. Hopf, Comm. Pure Appl. Math. **3**, 201–30 (1950); J.D. Cole, Quart. Appl. Math. **9**, 225–36 (1951).

Neutron transport in plane geometry with general anisotropic, energy-dependent scattering

Alan G. Gibbs

Battelle, Pacific Northwest Laboratory, Richland, Washington 99352

Edward W. Larsen[a)]

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

(Received 18 February 1977)

We consider the neutron transport equation in plane geometry with a general energy-dependent anisotropic scattering kernel. We construct the solution of the subcritical half-space albedo problem as a contour integral around the positive half of the spectrum of a reduced transport operator K. The integrand involves the boundary data and two operators which provide the Wiener–Hopf factorization of a third operator contained in $(\lambda I - K)^{-1}$. Bounds are obtained for the location of the spectrum of K in the complex plane. We also obtain representations of the solutions of the Milne problem and of the full-space and half-space problems with sources. Various simplifications of the general theory, which occur for particular scattering models, are discussed as an illustration of the results.

1. INTRODUCTION

We begin by determining the solution of the following half-space problem for the neutron transport equation

$$\mu\frac{\partial\psi}{\partial x}(x,\mu,E)+\sigma_t(E)\psi(x,\mu,E) - \int_0^{E_m}\int_{-1}^{1}\sigma(\mu',E'\to\mu,E)\psi(x,\mu',E')\,d\mu'\,dE' = 0, \quad 0<x<\infty, \tag{1.1}$$

$$\psi(0,\mu,E)=\psi_0(\mu,E), \quad 0<\mu\leq 1, \tag{1.2}$$

$$\lim_{x\to\infty}\psi(x,\mu,E)=0. \tag{1.3}$$

Here x is distance, μ is the cosine of the angle between the direction of propagation and the positive x axis, and E is energy; we take $0<E<E_m$. Also, ψ is the neutron flux, σ_t is the total cross section, σ is the cross section describing the production of all secondary neutrons, and $\psi_0\geq 0$ is the prescribed incident flux.[1]

A more general problem than (1.1)–(1.3) would have ψ and ψ_0 depending on the azimuthal angle ϕ. However, if scattering is rotationally invariant and ψ_0 and ψ are expanded in a Fourier series in ϕ, then the coefficients of $\exp(in\phi)$ in such a series decouple from all other components and solve a problem essentially identical to (1.1)–(1.3).[2] Therefore, we shall consider just the above problem (1.1)–(1.3) and regard it as the problem for any given Fourier mode in ϕ.

Problem (1.1)–(1.3) has received much attention over the years, and several different analytical methods have been developed to solve it.[1] All of these methods are based on the one group or multigroup approximation—or on the assumption of degenerate energy dependence—and on approximating the scattering kernel σ by a finite Legendre-polynomial expansion in the angular variables. The reason for making such approximations is that the range of the integral term in Eq. (1.1) becomes finite-dimensional in μ, E space. This simplification however does not, by any means, make problem (1.1)–(1.3) easy to solve—it only makes the problem amenable to certain types of analysis.

The first method for obtaining explicit solutions of this simplified problem was that of Wiener and Hopf, which is based on a Fourier transform of Eq. (1.1).[3,4] A more recent and now more popular method is that based on singular eigenfunction expansions;[5] this method and its development are discussed in detail in the review article by McCormick and Kuščer.[6]

It is generally acknowledged that the Wiener–Hopf and singular eigenfunction methods are equivalent, in the sense that any problem which can be solved by one method can also be solved by the other. Both methods yield truly explicit solutions only for one-group problems; for multigroup problems, solutions are expressed in terms of the solution of a matrix Riemann–Hilbert problem which, in general, has not been explicitly solved. Both methods become very cumbersome as the number of energy groups and angular moments in the scattering kernel σ increase. Also, by neither method can one construct solutions with L_1 boundary data, even though L_1 is the physically appropriate Banach space in which to solve transport problems.[7]

To overcome some of these difficulties, a new method has been developed which is based on a contour integral representation for ψ. This method was first used to derive more rigorously the singular eigenfunction formalism,[8–10] was later modified to construct L_1 solutions for one-group problems,[11] and has recently been extended to provide L_1 solutions for multigroup problems with degenerate anisotropic scattering.[12] Solutions of such multigroup problems are expressed in terms of a Wiener–Hopf factorization of a matrix $\Lambda(\lambda)$. This factorization has been explicitly obtained only for special problems,[13] but it has been shown that in general the factorization exists for both subcritical ("$C<1$") and critical ("$C=1$") media.[14,15] Some recent work on multigroup problems indicates that certain iteration schemes,

[a)] Present address: Los Alamos Scientific Laboratory, Theoretical Division, University of California, Los Alamos, New Mexico 87545.

based on the nonlinear integral equations which the factors of Λ must satisfy, converge quickly.[16]

In the present paper, we extend this contour integral method to the general problem (1.1)–(1.3), with no constraints imposed on the various quantities in these equations other than ones which are physically reasonable. We shall obtain a representation of the solution of this transport problem as a contour integral, and then perform some simple manipulations to recast the representation into a physically meaningful form. The representation is expressed, as in the previously considered multigroup problems, in terms of two operators which suitably factor a third operator; we discuss this factorization in Sec. 2.

Our representation is therefore not expressed in terms of explicitly known quantities, since closed-form expressions for the factors have not been obtained in general. In this sense, our analysis shows how to represent the solution of the full transport problem (1.1)–(1.3), without making any unphysical assumptions about the scattering terms in Eq. (1.1). In addition, our analysis provides a general viewpoint from which previous, different kinds of analysis of problem (1.1)–(1.3) can be understood. (See Secs. 7 and 8).

In Eq. (1.1), we shall require σ and $\sigma_t > 0$ to be measurable functions, satisfying:

$$\text{(a)} \quad \sup_{0<E<E_m} \frac{1}{\sigma_t(E)} \equiv \lambda_0 < \infty, \tag{1.4}$$

$$\text{(b)} \quad \sigma_t(E) > \int_0^{E_m} \int_{-1}^{1} |\sigma(\mu, E \to \mu', E')| \, d\mu' \, dE'. \tag{1.5}$$

Physically, λ_0 is the largest mean free path for neutrons with energy $E \in (0, E_m)$; condition (a) states that λ_0 is finite.

Condition (b) states that the total reaction rate of a particle with energy E before a collision is greater than the total reaction rate after a collision; hence the half-space $x > 0$ is subcritical.

If scattering is rotationally invariant (we do not actually need to assume this), then σ satisfies an additional equation, which we discuss in Sec. 4,

$$\sigma(\mu', E' \to \mu, E) = \sigma(-\mu', E' \to -\mu, E). \tag{1.6}$$

We impose the following condition on ψ_0:

$$\int_0^{E_m} \int_0^1 \mu \psi_0(\mu, E) \, d\mu \, dE < \infty. \tag{1.7}$$

This condition states that the number of particles entering the half-space through a unit cross sectional area of the boundary $x = 0$, in a unit time interval, is finite.

Finally, we define the weighted L_1 Banach space X by

$$X = \{ f(\mu, E) \mid \|f\| = \int_0^{E_m} \int_{-1}^{1} \sigma_t(E) |f(\mu, E)| \, d\mu \, dE < \infty \}. \tag{1.8}$$

Then we require ψ, the solution of problem (1.1)–(1.3), to satisfy

$$\|\psi\|(x) \equiv \|\psi(x, \mu, E)\| < \infty, \quad 0 < x < \infty. \tag{1.9}$$

Clearly, $\|\psi\|(x)$ is the total reaction rate per unit cross sectional area of the system at x. Equation (1.9) requires that this quantity be finite for each $x \in (0, \infty)$—or equivalently that $\psi \in X$ for each $x \in (0, \infty)$.

The above paragraphs provide the appropriate setting for problem (1.1)–(1.3). In Sec. 2 we shall construct a representation for ψ as an integral along a contour Γ^+ which contains the positive half of the spectrum of the "reduced" transport operator K [see Eq. (2.10)]. ψ is expressed there in terms of operators $\chi(\lambda)$ and $\mathcal{Y}(-\lambda)$ which suitably factor an operator $\Lambda(\lambda)$ contained in $(\lambda I - K)^{-1}$. The existence of this factorization for general scattering kernels is discussed in Sec. 2.

In Sec. 3 we shall cast this solution of problem (1.1)–(1.3) in a form which has a simple physical interpretation. In Sec. 4 we shall prove certain results about the location of the spectrum of K, and thereby show how the contour Γ^+ can be chosen. In Sec. 5 we discuss the Milne problem, and in Sec. 6 treat source problems. In Sec. 7 we shall indicate how, for multigroup problems with degenerate anisotropic scattering, the problem of factoring the operator $\Lambda(\lambda)$ can be reduced to factoring a matrix $\Lambda(\lambda)$. We illustrate this by explicitly rederiving the results of Ref. 11 for one-group, isotropic scattering. We conclude with a short discussion in Sec. 8.

2. CONTOUR INTEGRAL SOLUTION

To solve problem (1.1)–(1.3), we first define the operator $C : X \to X$ by

$$Cf(\mu, E) = \frac{1}{\sigma_t(E)} \int_0^{E_m} \int_{-1}^{1} \sigma(\mu', E' \to \mu, E) f(\mu', E') \, d\mu' \, dE'. \tag{2.1}$$

C is a bounded operator with

$$\|C\| \leq \sup_{\substack{-1 \leq \mu \leq 1 \\ 0 \leq E \leq E_m}} \frac{1}{\sigma_t(E)} \int_0^{E_m} \int_{-1}^{1} |\sigma(\mu, E \to \mu', E')| \, d\mu' \, dE'. \tag{2.2}$$

Thus by Eq. (1.5), $\|C\| < 1$, and so $(I - C)^{-1}$ exists.

We divide Eq. (1.1) by $\sigma_t(E)$ and operate by $(I - C)^{-1}$, obtaining

$$K \frac{\partial \psi}{\partial x} + \psi = 0. \tag{2.3}$$

Here K is the operator

$$Kf(\mu, E) = (I - C)^{-1} \frac{\mu f(\mu, E)}{\sigma_t(E)}. \tag{2.4}$$

The domain and range of K are not X, but instead the space $X_1 \supset X$ defined by

$$X_1 = \{ f \mid \mu f(\mu, E) \in X \}, \quad \|f\|_1 = \|\mu f\|, \quad f \in X_1.$$

If we define

$$c = \|C\|, \tag{2.5}$$

then by Eqs. (1.4) and (2.4), $K : X_1 \to X \subset X_1$ is a bounded operator with $\|K\|_1 \leq \lambda_0/(1 - c)$. (We introduce X_1 because the contour integral solution contains K acting on an extension of ψ_0. Since ψ_0 satisfies Eq. (1.7), this extension will be an element of X_1, but it will not in general be an element of X.)

Next we shall construct $(\lambda I - K)^{-1}$. To do this it is simplest to note that X is an invariant subspace of K, i.e., $K: X \to X$. Therefore for $f, g \in X$, we set $f = (\lambda I - K)g$ and operate by $(I - C)$ to get

$$(I - C)f(\mu, E) = \lambda(I - C)g(\mu, E) - \frac{\mu g(\mu, E)}{\sigma_t(E)}$$

$$= \left[I - \lambda C \frac{1}{\lambda - \mu/\sigma_t(E)}\right] \left(\lambda - \frac{\mu}{\sigma_t(E)}\right) g(\mu, E). \tag{2.6}$$

But

$$(I - C)f(\mu, E) = \left[I - \lambda C \frac{1}{\lambda - \mu/\sigma_t(E)}\right] f(\mu, E)$$

$$+ C \frac{\mu}{\lambda\sigma_t(E) - \mu} f(\mu, E). \tag{2.7}$$

Therefore, if we define

$$\mathbf{\Lambda}(\lambda)f(\mu, E) = \left[I - \lambda C \frac{1}{\lambda - \mu/\sigma_t(E)}\right] f(\mu, E), \tag{2.8}$$

$$N(\lambda)f(\mu, E) = \frac{\mu}{\lambda\sigma_t(E) - \mu} f(\mu, E), \tag{2.9}$$

then Eqs. (2.6) and (2.7) can be solved for $g = (\lambda I - K)^{-1}f$, obtaining

$$(\lambda I - K)^{-1}f(\mu, E) = \frac{1}{\lambda - \mu/\sigma_t(E)} \{f(\mu, E)$$

$$+ \mathbf{\Lambda}^{-1}(\lambda) CN(\lambda) f(\mu, E)\}. \tag{2.10}$$

This expression was derived just for $f \in X$, but one can directly show that it holds for all $f \in X_1$. In general, then, we interpret Eq. (2.10) as follows: For $\lambda \notin [-\lambda_0, \lambda_0]$, $N(\lambda): X_1 \to X$ is bounded, and so for $f \in X_1$, $CN(\lambda)f \in X$. By Eq. (2.8) $\mathbf{\Lambda}(\lambda): X \to X$, and if λ is such that $\mathbf{\Lambda}^{-1}(\lambda)$ is bounded on X, then $\mathbf{\Lambda}^{-1}(\lambda)CN(\lambda)f \in X$.

Thus, $\sigma(K)$, the spectrum of K, consists at most of the line segment $[-\lambda_0, \lambda_0]$ plus the points for which $\mathbf{\Lambda}^{-1}(\lambda)$ fails to exist. In Sec. 3 we show that this spectrum is confined to a "figure-eight" shaped region of the complex λ plane. (See Fig. 1 in Sec. 4.) For $\lambda \notin \sigma(K)$, $(\lambda I - K)^{-1}: X_1 \to X_1$ is an analytic, operator-valued function of λ.

Now, by analogy to Refs. (11) and (12), we shall construct the solution of problem (1.1)—(1.3) in terms of the following integral representation:

$$\psi(x, \mu, E) = \frac{1}{2\pi i} \oint_{\Gamma^+} \exp(-x/\lambda)(\lambda I - K)^{-1} g(\mu, E)\, d\lambda. \tag{2.11}$$

Here Γ^+ is a simple closed curve, contained wholly in the right-half-plane $\mathrm{Re}\lambda \geq 0$, and enclosing the right-half of $\sigma(K)$. (See Fig. 1.) The function $g \in X_1$ remains to be determined.

Since $g \in X_1$, we must interpret the integral of Eq. (2.11) in X_1. Thus, at this point, $\psi \in X_1$ for each $x \in (0, \infty)$. Later, however, we shall show that actually $\psi \in X \subset X_1$ for each $x \in (0, \infty)$.

First we shall verify that ψ, defined by Eq. (2.11), satisfies Eq. (2.3)—which is equivalent to Eq. (1.1). To do this, we introduce Eq. (2.11) into (2.3) and obtain

$$K\frac{\partial\psi}{\partial x} + \psi = \frac{1}{2\pi i} \oint_{\Gamma^+} \exp(-x/\lambda) \left(I - \frac{K}{\lambda}\right) (\lambda I - K)^{-1} g\, d\lambda$$

$$= \frac{g}{2\pi i} \oint_{\Gamma^+} \frac{\exp(-x/\lambda)}{\lambda} d\lambda. \tag{2.12}$$

This last integral is zero because for $x > 0$, $[\exp(-x/\lambda)/\lambda] \to 0$ as $\lambda \to 0$ along Γ^+. Therefore, Eq. (2.3) is satisfied identically for every $g \in X_1$.

Thus ψ, defined by Eq. (2.11), satisfies Eq. (1.1), and clearly also Eq. (1.3). It remains to determine $g \in X_1$ so that ψ satisfies the boundary condition (1.2). Introducing Eq. (2.11) into (1.2), this boundary condition becomes

$$\psi_0(\mu, E) = \frac{1}{2\pi i} \oint_{\Gamma^+} (\lambda I - K)^{-1} g(\mu, E)\, d\lambda, \quad 0 < \mu \leq 1. \tag{2.13}$$

If the contour Γ^+ extended around the entire spectrum of K, then Eq. (2.13) would be satisfied for any g such that

$$\psi_0(\mu, E) = g(\mu, E), \quad 0 < \mu \leq 1. \tag{2.14}$$

Therefore, let us require g to satisfy Eq. (2.14) and the condition

$$(\lambda I - K)^{-1} g \in X_1 \text{ is an analytic function of } \lambda \text{ for } \mathrm{Re}\lambda < 0. \tag{2.15}$$

Then without change in Eq. (2.13), Γ^+ *can* be extended to enclose the negative half of the spectrum of K since by (2.15), the contribution from such an extension is zero.

We thus require g to satisfy Eqs. (2.14) and (2.15), and then g will automatically satisfy the boundary condition (2.13). Since Eq. (2.14) defines g explicitly in terms of ψ_0 for $0 < \mu \leq 1$, then we expect—and shall verify below—that condition (2.15) determines g in terms of ψ_0 for $-1 \leq \mu < 0$. [See Eqs. (2.30) and (2.31).]

To satisfy condition (2.15), we shall utilize the following Wiener—Hopf factorization of the operator $\mathbf{\Lambda}(\lambda)$:

$$\mathbf{\Lambda}(\lambda) = \mathcal{Y}(-\lambda)\mathcal{X}(\lambda). \tag{2.16}$$

Here $\mathcal{X}(\lambda): X \to X$ and $\mathcal{Y}(\lambda): X \to X$ are operator-valued functions of λ, which are analytic and invertible in the left-half-plane $\mathrm{Re}\lambda < 0$. Since $\mathbf{\Lambda}(\infty) = I - C$, we may require $\mathcal{X}$, $\mathcal{X}^{-1}$, $\mathcal{Y}$, and $\mathcal{Y}^{-1}$ to have finite limits at $\lambda = \infty$.

The problem of obtaining the factors $\mathcal{X}(z)$ and $\mathcal{Y}(-z)$ in the case when $\Lambda(z)$ is a scalar function is classical and the solution is well known.[4,5] The existence of the factorization for a matrix $\Lambda(z)$ appropriate to multigroup neutron transport with isotropic scattering was considered by Mullikin,[14] who adapted methods appearing earlier in the Russian literature (see Ref. 14 for references). Bowden, Zweifel, and Menikoff[16] have shown that the factors for the multigroup problem can be computed by iterative solution of certain nonlinear integral equations, and that the method can be extended to the case of anisotropic scattering with a degenerate kernel. The convergence of their iterative method was proven under the assumption that the half-space is sufficiently subcritical. In a recent paper, Kelley[17] considered the case of isotropic scattering with continuous

energy dependence and, assuming a compact scattering kernel and a bounded total cross section, showed that the desired factors exist and can be constructed by iterative solution of a nonlinear integral equation. The existence of factors, and means for their computation, are being considered by the present authors under very general assumptions concerning the scattering operator. This work will be reported in a future paper.

In the present paper, we shall assume that the factors $\mathcal{X}(\lambda)$ and $\mathcal{Y}(\lambda)$ exist with the required analyticity properties. We then introduce Eq. (2.16) into (2.10), obtaining

$$(\lambda I - K)^{-1} g(\mu, E) = \frac{1}{\lambda - \mu/\sigma_t(E)} \left\{ g(\mu, E) + \mathcal{X}^{-1}(\lambda) \mathcal{Y}^{-1}(-\lambda) C N(\lambda) g(\mu, E) \right\}. \tag{2.17}$$

Then for condition (2.15) to be satisfied, we must have both

$$\mathcal{Y}^{-1}(-\lambda) C N(\lambda) g(\mu, E) \equiv \gamma(\lambda, \mu, E) \tag{2.18}$$

analytic for $\mathrm{Re}\lambda < 0$, and

$$0 = g(\mu, E) + \mathcal{X}^{-1}(\nu)\gamma(\nu, \mu, E), \quad -1 \leq \mu < 0, \tag{2.19}$$

where

$$\nu = \mu/\sigma_t(E). \tag{2.20}$$

Requiring γ to be analytic for $\mathrm{Re}\lambda < 0$ is equivalent to requiring the numerator in $(\lambda I - K)^{-1} g$ to be analytic for $\mathrm{Re}\lambda < 0$. Also, Eqs. (2.19) and (2.20) state that this numerator vanishes for those λ values where the denominator vanishes—namely $\lambda = \nu = \mu/\sigma_t(E)$. We have written this second condition as two equations to emphasize that while $\mathcal{X}^{-1}$ does act on the variables μ and E, it does *not* act on ν. Equations (2.19) and (2.20) explicitly describe g for $-1 \leq \mu < 0$ in terms of γ, which we must determine from the requirement that γ be analytic for $\mathrm{Re}\lambda < 0$.

To determine γ, we rewrite Eq. (2.18) as

$$\tau(\lambda, \mu, E) = \mathcal{Y}(-\lambda)\gamma(\lambda, \mu, E), \tag{2.21}$$

where by Eqs. (2.1) and (2.9),

$$\tau(\lambda, \mu, E) \equiv C N(\lambda) g(\mu, E) = \frac{1}{\sigma_t(E)} \int_0^{E_m} \int_{-1}^{1} \sigma(\mu', E' \to \mu, E) \times \frac{\mu' g(\mu', E')}{\lambda \sigma_t(E') - \mu'} d\mu' dE'. \tag{2.22}$$

$\tau \in X$ is an analytic function of λ for $\lambda \notin [-\lambda_0, \lambda_0]$, and $\tau(\infty, \mu, E) = 0$. Thus, by the conditions on $\mathcal{Y}(-\lambda)$ and $\gamma(\lambda, \mu, E)$ discussed above, γ is analytic for $\mathrm{Re}\lambda < 0$ if

$$\frac{1}{2\pi i}[\tau^+(\nu, \mu, E) - \tau^-(\nu, \mu, E)] = \mathcal{Y}(-\nu) \frac{1}{2\pi i}[\gamma^+(\nu, \mu, E) - \gamma^-(\nu, \mu, E)], \quad 0 < \nu < \lambda_0, \tag{2.23}$$

$$\frac{1}{2\pi i}[\tau^+(\nu, \mu, E) - \tau^-(\nu, \mu, E)] = \frac{1}{2\pi i}[\mathcal{Y}^+(-\nu) - \mathcal{Y}^-(-\nu)]\gamma(\nu, \mu, E), \quad -\lambda_0 < \nu < 0, \tag{2.24}$$

and

$$\gamma(\infty, \mu, E) = 0. \tag{2.25}$$

[We use the standard notation $f^{\pm}(\nu) = \lim_{\epsilon \to 0^+} f(\nu \pm i\epsilon)$.]

To satisfy Eqs. (2.23)–(2.25), we shall make use of the following lemma:

Lemma: Let $f(\mu, E) \in X$ be Hölder continuous in μ for each E, and define $f(\mu, E) = 0$ for $|\mu| > 1$. Also, define

$$F(\lambda) = \int_0^{E_m} \int_{-1}^{1} \frac{f(\mu', E')}{\mu' - \lambda\sigma_t(E')} d\mu' dE'.$$

Then

$$\frac{1}{2\pi i}[F^+(\nu) - F^-(\nu)] = \int_0^{E_m} f(\nu\sigma_t(E'), E')\, dE', \quad -\lambda_0 < \nu < \lambda_0.$$

Proof: Follows immediately from the Plemelj formulas.[18]

To proceed, we shall temporarily assume that $\mu g(\mu, E)$ is Hölder continuous in μ. Then we may apply the above lemma to Eq. (2.22), obtaining

$$\frac{1}{2\pi i}[\tau^+(\nu, \mu, E) - \tau^-(\nu, \mu, E)] = -\frac{1}{\sigma_t(E)} \int_0^{E_m} \sigma(\nu\sigma_t(E'), E' \to \mu, E) \times \nu\sigma_t(E') g(\nu\sigma_t(E'), E')\, dE', \quad -\lambda_0 < \nu < \lambda_0. \tag{2.26}$$

Here we have defined

$$g(\mu, E) = 0, \quad |\mu| > 1. \tag{2.27}$$

We may now solve Eqs. (2.23) and (2.25) for γ by taking

$$\gamma(\lambda, \mu, E) = \int_0^{\lambda_0} \frac{\mathcal{Y}^{-1}(-\nu)}{2\pi i} [\tau^+(\nu, \mu, E) - \tau^-(\nu, \mu, E)] \frac{d\nu}{\nu - \lambda}. \tag{2.28}$$

Introducing Eqs. (2.26) and (2.14) into (2.28), and making the change of variables $\mu' = \nu\sigma_t(E')$, we obtain

$$\gamma(\lambda, \mu, E) = \int_0^{E_m} \int_0^1 \left[\mathcal{Y}^{-1}\left(-\frac{\mu'}{\sigma_t(E')}\right) \frac{\sigma(\mu', E' \to \mu, E)}{\sigma_t(E)} \right] \times \frac{\mu' \psi_0(\mu', E')}{\lambda\sigma_t(E') - \mu'} d\mu' dE'. \tag{2.29}$$

(In this equation, $\mathcal{Y}^{-1}$ acts on the *unprimed* variables μ and E.) Equation (2.29) defines γ explicitly in terms of ψ_0, and thus Eq. (2.29) can be introduced into Eq. (2.19) to yield g, for $-1 \leq \mu < 0$, explicitly in terms of ψ_0.

It remains to solve Eq. (2.24). But since γ and g have been completely determined, then Eq. (2.24) must be automatically satisfied. To verify this directly, we use $\mathcal{Y}(-\lambda) = \Lambda(\lambda)\mathcal{X}^{-1}(\lambda)$, Eqs. (2.8), (2.19), and the lemma to get

$$\frac{1}{2\pi i}[\mathcal{Y}^+(-\nu)-\mathcal{Y}^-(-\nu)]\gamma(\nu,\mu,E)$$

$$=\frac{1}{2\pi i}[\Lambda^+(\nu)-\Lambda^-(\nu)]\chi^{-1}(\nu)\gamma(\nu,\mu,E)$$

$$=-\frac{\nu}{\sigma_t(E)}\int_0^{E_m}\sigma(\nu\sigma_t(E'),E'\to\mu,E)\sigma_t(E')$$

$$\times g(\nu\sigma_t(E'),E')\,dE',\quad -\lambda_0<\nu<0.$$

Comparing this result with Eq. (2.26), we see that Eq. (2.24) is identically satisfied by g.

We shall now summarize the results, which we have derived by requiring $\mu g(\mu,E)$ to be Hölder continuous in μ.

The solution ψ of the transport problem (1.1)–(1.3) is given by Eq. (2.11). The resolvent operator $(\lambda I-K)^{-1}$ is given explicitly in Eq. (2.10), and the function $g\in X_1$ is defined by Eqs. (2.14), (2.19), and (2.29). We may write g more compactly as

$$g(\mu,E)=E\psi_0(\mu,E)$$
$$=\begin{cases}\psi_0(\mu,E), & 0<\mu\le 1,\\ \int_0^{E_m}\int_0^1\mathcal{E}(\mu',E'\to\mu,E)\psi_0(\mu',E')\,d\mu'\,dE', & \\ & -1\le\mu<0,\end{cases}\tag{2.30}$$

where

$$\mathcal{E}(\mu',E'\to\mu,E)=\frac{\mu'}{\mu'-\nu\sigma_t(E')}\chi^{-1}(\nu)$$
$$\times\mathcal{Y}^{-1}\left(-\frac{\mu'}{\sigma_t(E')}\right)\frac{\sigma(\mu',E'\to\mu,E)}{\sigma_t(E)},\tag{2.31}$$
$$\nu=\mu/\sigma_t(E),\quad 0<\mu'\le 1,\quad -1\le\mu<0.$$

(Again, we emphasize that χ^{-1} and $\mathcal{Y}^{-1}$ act only on the unprimed variables μ and E, and *not* on ν.)

The operator E, given by Eq. (2.30), extends $\psi_0(\mu,E)$ to a function $E\psi_0(\mu,E)$ which is defined for $-1\le\mu\le 1$. For $x=0$, Eqs. (2.11) and (2.30) reduce to

$$\psi(0,\mu,E)=E\psi_0(\mu,E),\quad -1\le\mu\le 1.\tag{2.32}$$

Thus for $-1\le\mu<0$, $E\psi_0$ is the reflection of the incident beam ψ_0. (E is the "reflection" operator, defined for simpler scattering models in Refs. 8–10 and 12.)

To derive the above results, we required $\mu g(\mu,E)$ to be Hölder continuous in μ. However, a simple limiting argument shows that the results remain valid for $g\in X_1$. It is also simple to verify that for ψ_0 satisfying Eq. (1.7), $g=E\psi_0$, defined by Eq. (2.30), is in X_1.

The only remaining detail is to show that the solution ψ is in X for each x, rather than in X_1. We shall verify this in Sec. 3.

3. REFORMULATION OF THE SOLUTION

Combining Eqs. (2.11), (2.17), and (2.18), we obtain the following expression for ψ:

$$\psi(x,\mu,E)=\frac{1}{2\pi i}\oint_{\Gamma^+}\frac{\exp(-x/\lambda)}{\lambda-\mu/\sigma_t(E)}\{g(\mu,E)$$
$$+\chi^{-1}(\lambda)\gamma(\lambda,\mu,E)\}\,d\lambda.\tag{3.1}$$

From Eq. (2.30) we get

$$\frac{1}{2\pi i}\oint_{\Gamma^+}\frac{\exp(-x/\lambda)}{\lambda-\mu/\sigma_t(E)}g(\mu,E)\,d\lambda=\psi_0(\mu,E)\exp[-d(x,\mu,E)],\tag{3.2}$$

where

$$d(x,\mu,E)=\begin{cases}x\sigma_t(E)/\mu, & 0<\mu<1,\\ +\infty, & -1<\mu<0.\end{cases}\tag{3.3}$$

Here $d(x,\mu,E)$ is the optical distance an uncollided neutron must travel to reach the point x, travelling in the "direction" μ with energy E. Thus, Eq. (3.2) describes the uncollided neutron flux.

Also, from Eq. (2.29) we get

$$\frac{1}{2\pi i}\oint_{\Gamma^+}\frac{\exp(-x/\lambda)}{\lambda-\mu/\sigma_t(E)}\chi^{-1}(\lambda)\gamma(\lambda,\mu,E)\,d\lambda$$
$$=\int_0^{E_m}\int_0^1 H(\mu',E'\to x,\mu,E)\psi_0(\mu',E')\,d\mu'\,dE',\tag{3.4}$$

where

$$H(\mu',E'\to x,\mu,E)$$
$$=\mu'\sigma_t(E)\left\{\frac{1}{2\pi i}\oint_{\Gamma^+}\frac{\exp(-x/\lambda)}{[\lambda\sigma_t(E)-\mu][\lambda\sigma_t(E')-\mu']}\chi^{-1}(\lambda)\,d\lambda\right\}$$
$$\times\mathcal{Y}^{-1}\left(-\frac{\mu'}{\sigma_t(E')}\right)\frac{\sigma(\mu',E'\to\mu,E)}{\sigma_t(E)}.\tag{3.5}$$

(χ^{-1} and $\mathcal{Y}^{-1}$ act on the unprimed variables μ and E.) Equation (3.4) describes the collided neutron flux. H is a regular Green's function describing the collided flux at (x,μ,E) due to the incident flux at $(0,\mu',E')$.

Combining Eqs. (3.1), (3.2), and (3.4), we obtain

$$\psi(x,\mu,E)=\psi_0(\mu,E)\exp[-d(x,\mu,E)]$$
$$+\int_0^{E_m}\int_0^1 H(\mu',E'\to x,\mu,E)\psi_0(\mu',E')\,d\mu'\,dE'.\tag{3.6}$$

Using this form for ψ and Eq. (3.5), it is now simple to show that $\psi\in X$ for ψ_0 satisfying Eq. (1.7).

4. THE SPECTRUM OF K

We have already determined, by inspection of Eq. (2.10), that the line segment $[-\lambda_0,\lambda_0]$ is contained in the spectrum of K, and that the remaining spectrum consists at most of those values of λ for which $\mathbf{\Lambda}^{-1}(\lambda)$ fails to exist as a bounded operator mapping X into X. In the following theorems, we shall investigate the singularities of $\mathbf{\Lambda}^{-1}(\lambda)$.

Theorem 1: If $\lambda\in\sigma(K)$, then $\bar{\lambda}\in\sigma(K)$.

Proof: Let $\lambda\in\sigma(K)$ with $\lambda\notin[-\lambda_0,\lambda_0]$. Then $\mathbf{\Lambda}(\lambda)$: $X\to X$ is not invertible, and so either $\mathbf{\Lambda}(\lambda)\phi=0$ for some $\phi\in X$, or $\mathbf{\Lambda}(\lambda)X\subsetneq X$.[19] If $\mathbf{\Lambda}(\lambda)\phi=0$, then by Eq. (2.6) $\mathbf{\Lambda}(\bar{\lambda})\bar{\phi}=0$, and if $\mathbf{\Lambda}(\lambda)X\subsetneq X$, then $\mathbf{\Lambda}(\bar{\lambda})X=\mathbf{\Lambda}(\bar{\lambda})\bar{X}=\overline{\mathbf{\Lambda}(\lambda)X}$

$\not\subset X$. In either case, $\Lambda(\bar{\lambda})$ is not invertible, and so $\bar{\lambda} \in \sigma(K)$. Q.E.D.

Theorem 2: Let scattering be rotationally invariant. If $\lambda \in \sigma(K)$, then also $-\lambda \in \sigma(K)$.

Proof: Let $\lambda \in \sigma(K)$ with $\lambda \notin [-\lambda_0, \lambda_0]$. Then, as above, we have either $\mathbf{\Lambda}(\lambda)\phi = 0$ for some $\phi \in X$, or $\Lambda(\lambda)X \not\subset X$. Let us define $M: X \to X$ by $Mf(\mu, E) = f(-\mu, E)$. Then using Eqs. (2.8), (2.1), and (1.6), we may derive the identity $\mathbf{\Lambda}(-\lambda)M = M\mathbf{\Lambda}(\lambda)$.

If $\Lambda(\lambda)\phi = 0$, then $\Lambda(-\lambda)(M\phi) = M[\Lambda(\lambda)\phi] = 0$, and if $\mathbf{\Lambda}(\lambda)X \not\subset X$, then $\mathbf{\Lambda}(-\lambda)X = \mathbf{\Lambda}(-\lambda)MX = M\mathbf{\Lambda}(\lambda)X \not\subset X$. In either case, $\mathbf{\Lambda}(-\lambda)$ is not invertible on X, and so $-\lambda \in \sigma(K)$. Q.E.D.

Theorems 1 and 2 show that for rotationally invariant scattering, $\sigma(K)$ is symmetric across both the Reλ and Imλ axes. The following theorem provides a useful estimate on the location of $\sigma(K)$; this theorem is valid whether scattering is rotationally invariant or not:

Theorem 3: Let $c = \|C\|$. Then for $|\mathrm{Re}\lambda| \leq \lambda_0$, $\sigma(K)$ is restricted to the set

$$|\mathrm{Im}\lambda| \leq \frac{c}{(1-c^2)^{1/2}}|\mathrm{Re}\lambda|, \quad |\mathrm{Re}\lambda| \leq \lambda_0, \tag{4.1a}$$

and for $\lambda_0 \leq |\mathrm{Re}\lambda| \leq \lambda_0/(1-c)$, $\sigma(K)$ is restricted to the following sets:

$$\left(\mathrm{Re}\lambda \pm \frac{\lambda_0}{1-c^2}\right)^2 + (\mathrm{Im}\lambda)^2 \leq \left(\frac{\lambda_0 c}{1-c^2}\right)^2, \quad \lambda_0 \leq |\mathrm{Re}\lambda| \leq \frac{\lambda_0}{1-c}. \tag{4.1b}$$

There are no values of λ in $\sigma(K)$ satisfying $|\mathrm{Re}\lambda| > \lambda_0/(1-c)$. [Thus $\sigma(K)$ is contained on and within the curve $\Gamma(c)$, sketched in Fig. 1.]

Proof: It suffices to show that for λ outside of the sets (4.1), $\Lambda(\lambda)$ is invertible. To do this, let us consider Eq. (2.8).

$$\mathbf{\Lambda}(\lambda)f(\mu, E) = [I - CM(\lambda)]f(\mu, E),$$

where

$$M(\lambda)f(\mu, E) = \frac{\lambda f(\mu, E)}{\lambda - \mu/\sigma_t(E)}.$$

Then $\mathbf{\Lambda}(\lambda)$ is invertible if $\|CM(\lambda)\| < 1$.

But $M(\lambda)$ is simply a scalar multiplication operator, and therefore

$$\|CM(\lambda)\| < c\|M(\lambda)\| = \sup_{\substack{-1 \leq \mu \leq 1 \\ 0 \leq E \leq E_m}} \left|\frac{c\lambda}{\lambda - \mu/\sigma_t(E)}\right|.$$

Thus $\Lambda(\lambda)$ is invertible for λ satisfying

$$\sup_{\substack{-1 \leq \mu \leq 1 \\ 0 \leq E \leq E_m}} \left|\frac{c\lambda}{\lambda - \mu/\sigma_t(E)}\right| < 1. \tag{4.2}$$

The *complement* of the set (4.2) is easily shown to be the union of the sets (4.1), and therefore $\sigma(K)$ is contained in this union. Q.E.D.

In Fig. 1, $\Gamma(c)$ consists of the lines $\mathrm{Im}\lambda = \pm(\mathrm{Re}\lambda)c/(1-c^2)^{1/2}$ for $-\lambda_0 \leq \mathrm{Re}\lambda \leq \lambda_0$, and of the circles

$$\left(\mathrm{Re}\lambda \pm \frac{\lambda_0}{1-c^2}\right)^2 + (\mathrm{Im}\lambda)^2 = \left(\frac{\lambda_0 c}{1-c^2}\right)^2$$

for $\lambda_0 \leq |\mathrm{Re}\lambda| \leq \lambda_0/(1-c)$. These curves intersect for $\mathrm{Re}\lambda = \pm\lambda_0$, and at the points of intersection they are tangent.

An upper bound on $c = \|C\|$ is given in Eq. (2.2). Using this expression the curve $\Gamma(c)$ can be determined, thereby explicitly locating the portion of the complex plane containing $\sigma(K)$.

The next theorem concerns the *nature* of the spectrum of K within the set described in Theorem 3. To prove this theorem, we must assume that $C: X \to X$ is a compact (i.e., completely continuous) operator. While this assumption is not true in general, it is true for certain real physical problems, and it is always true for multigroup problems with degenerate anisotropic scattering.

Theorem 4: Let $C: X \to X$ be compact. Then in any closed subset of the complex plane which excludes the

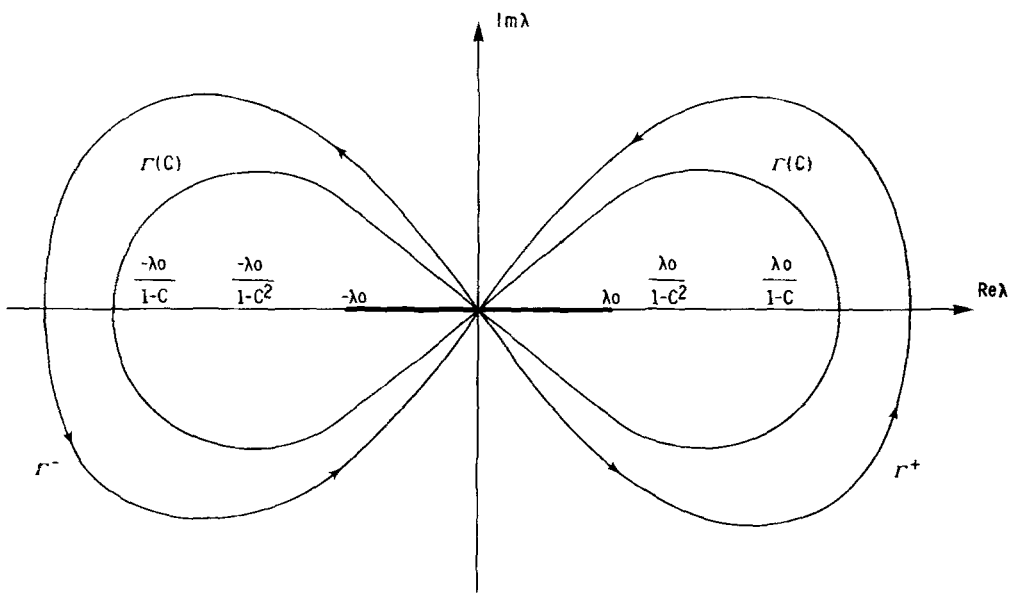


FIG. 1. The contours Γ^+, Γ^-, and the curve $\Gamma(c)$. [$\sigma(K)$ is contained in the region bounded by $\Gamma(c)$.]

interval $[-\lambda_0, \lambda_0]$, the spectrum of K consists of discrete, isolated, point eigenvalues having finite multiplicity.

Proof: Let S be any closed subset of the complex plane excluding the interval $[-\lambda_0, \lambda_0]$. Then the operator $CN(\lambda): X \to X$ is a compact operator-valued function, analytic in S. Furtheremore, $I - CN(\infty) = I - C$ is invertible. Therefore, by a theorem of Gohberg,[20] $\Lambda(\lambda) = I - CN(\lambda)$ is invertible for all $\lambda \in S$ except possibly for certain exceptional discrete, isolated λ values, λ_k. For any "exceptional" point λ_k, zero is a point eigenvalue of $\Lambda(\lambda)$ with finite multiplicity. Therefore, λ_k is a point eigenvalue of K with finite multiplicity. Q.E.D.

We note that Theorem 4 does not exclude the possibility that an infinite sequence of point eigenvalues of K exists. However, such a sequence would, by necessity, "converge" to the interval $[-\lambda_0, \lambda_0]$.

Finally, we discuss some results of Kuščer and Vidav[21] pertaining to kernels $\sigma(\mu', E' \to \mu, E)$ which can be symmetrized, e.g., which satisfy the detailed balance relation. For such scattering kernels, and in the context of an L_2 rather than an L_1 space, Kuščer and Vidav have shown that the spectrum of K lies completely on the real line. If, in addition, C is compact—in the L_2 space—then the "largest" eigenvalues $\pm\lambda_0$ of K are real and have geometrical multiplicity one with corresponding eigenfunctions $\phi_0(\pm\mu, E)$ which are positive.

These results have not been proved in L_1 or for general scattering kernels. Consideration of such problems, however, is beyond the scope of this paper.

5. THE MILNE PROBLEM

The solution of the classical Milne problem[5] can be constructed using the results of Sec. 2, provided K has a point eigenvalue $-\lambda_0$ which is real, negative, less than the real part of any other spectral point of K, and corresponding to which is an eigenfunction $\phi_{-\lambda_0}(\mu, E)$. (The existence of $-\lambda_0$ and $\phi_{-\lambda_0}$ is discussed in Sec. 4.) Then the solution of the Milne problem is

$$\psi(x, \mu, E) = a[\phi_{-\lambda_0}(\mu, E)\exp(x/\lambda_0) + \psi_1(x, \mu, E)], \quad (5.1)$$

where a is an arbitrary positive constant and ψ_1 is the solution of problem (1.1)–(1.3) with $\psi_1(0, \mu, E) = \psi_0(\mu, E) = \phi_{-\lambda_0}(\mu, E)$.

6. SOURCE PROBLEMS

In this section we shall construct a particular solution of the equation

$$\mu\frac{\partial\psi}{\partial x}(x, \mu, E) + \sigma_t(E)\psi(x, \mu, E) - \int_0^{E_m}\int_{-1}^{1}\sigma(\mu', E' \to \mu, E)\psi(x, \mu', E')\,d\mu'\,dE' = q(x, \mu, E), \quad a < x < b. \quad (6.1)$$

Here $-\infty \le a, b \le +\infty$; also $q \in X$ and $\psi \in X$ for every $x \in (a, b)$.

If we define

$$q_0(x, \mu, E) = \frac{1}{\mu}q(x, \mu, E), \quad (6.2)$$

then $q_0 \in X_1$ for each x, and so as in Sec. 2 we may rewrite Eq. (6.1) as

$$K\frac{\partial\psi}{\partial x} + \psi = Kq_0, \quad a < x < b. \quad (6.3)$$

By analogy to Eq. (2.11), we shall construct a solution of Eq. (6.3) of the form

$$\psi(x, \mu, E) = \frac{1}{2\pi i}\oint_\Gamma (\lambda I - K)^{-1}g(x, \lambda; \mu, E)\,d\lambda. \quad (6.4)$$

The above contour $\Gamma = \Gamma^+ \cup \Gamma^-$ encloses $\sigma(K)$; see Fig. 1.

Introducing Eq. (6.4) into (6.3), we obtain

$$\frac{1}{2\pi i}\oint_\Gamma (\lambda I - K)^{-1}\left(\lambda\frac{\partial g}{\partial x} + g\right)d\lambda = Kq_0.$$

This condition is met if g satisfies

$$\lambda\frac{\partial g}{\partial x} + g = \lambda q_0, \quad a < x < b, \quad (6.5a)$$

$$0 = \lim_{\substack{\lambda \to 0 \\ \lambda \in \Gamma}} g, \quad a < x < b. \quad (6.5b)$$

The general solution of Eqs. (6.5) is

$$g(x, \lambda, \mu, E) = \begin{cases} g_a(\mu, E)\exp[(a-x)/\lambda] + \displaystyle\int_a^x \frac{\exp[(x'-x)/\lambda]}{\lambda} \times q_0(x', \mu, E)\,dx', & \mathrm{Re}\,\lambda > 0, \\ g_b(\mu, E)\exp[(b-x)/\lambda] + \displaystyle\int_b^x \frac{\exp[(x'-x)/\lambda]}{\lambda} \times q_0(x', \mu, E)\,dx', & \mathrm{Re}\,\lambda < 0. \end{cases} \quad (6.6)$$

Here g_a and g_b are arbitrary functions which can be used to satisfy boundary conditions. However, to merely construct a particular solution of Eq. (6.1), it suffices to take $g_a = g_b = 0$. This particular solution is then explicitly given by Eqs. (6.4) and (6.6), and can be rewritten as

$$\psi(x, \mu, E) = \frac{1}{2\pi i}\oint_{\Gamma^+}(\lambda I - K)^{-1}\left\{\int_a^x \frac{\exp[(x'-x)/\lambda]}{\lambda} \times q_0(x', \mu, E)\,dx'\right\}d\lambda - \frac{1}{2\pi i}\oint_{\Gamma^-}(\lambda I - K)^{-1}\left\{\int_x^b \frac{\exp[(x'-x)/\lambda]}{\lambda} \times q_0(x', \mu, E)\,dx'\right\}d\lambda. \quad (6.7)$$

As we did in Sec. 2, we could reformulate ψ into a more physically meaningful form, but for brevity we shall not do this here.

We note that the above results apply to both full and half-space problems. For the full space problem one

takes $a=-\infty$ and $b=+\infty$, and for the $x>0$ half-space problem one takes $a=0$ and $b=+\infty$.

For full-space problems with sources at finite x and at $x=+\infty$ or $-\infty$, one can add to Eq. (6.7) Milne-type terms, which grow exponentially as $x\to+\infty$ or $-\infty$, to represent the effect of the infinite-range sources for finite x. [Such a term, which grows exponentially as $x\to+\infty$, occurs in Eq. (5.1).] For half-space ($0<x<\infty$) problems with sources at finite x and at $x=+\infty$, an exponentially growing Milne-type term of the type in Eq. (5.1) can be added to Eq. (6.7) to represent the source at $x=+\infty$. To satisfy the boundary condition for such problems at $x=0$, a solution of the type derived in Sec. 2 for the homogeneous half-space problem (1.1)–(1.3)—but with a modified boundary condition (1.2)—must be added to Eq. (6.7); such solutions are of course implicit in the function g_a, which we ignored in deriving Eq. (6.7) from (6.6).

In order that the representation (6.7) be defined, it is necessary that the terms in wavy brackets be elements of X_1 for each $x\in[a,b]$. This follows if $q(x,\mu,E)$ is not only an element of X for each x, but also satisfies

$$\|q\|(x)<Q<\infty, \quad x\in(a,b) \tag{6.8}$$

and

$$\|q\|(x) \text{ is locally integrable in } x. \tag{6.9}$$

Here $\|\ \|$ is just the X norm, defined in Eq. (1.8). Condition (6.9) is a technical condition, but (6.8) states physically that the total reaction rate for the source neutrons, per unit cross sectional area, is finite for each x and bounded by the positive constant Q.

7. ONE-GROUP, ISOTROPIC SCATTERING

In Secs. 2 and 3, we constructed the solution of problem (1.1)–(1.3) in terms of the contour integral representation (2.11). The major difficulty in carrying out this construction is the factorization of the operator $\Lambda(\lambda)$, which we recall occurs in Eq. (2.10).

However, by Eq. (2.10), we see that $\Lambda(\lambda)$ acts only on functions which lie in $R(C)$, the range of the operator C. Furthermore, by Eq. (2.8), $\Lambda(\lambda)$: $R(C)\to R(C)$. Hence, in the *entire* analysis of this paper, we may replace $\Lambda(\lambda)$ and $\Lambda^{-1}(\lambda)$ by their restrictions to the subspace $R(C)\subset X$. The utility of this is, as we mentioned in the Introduction, that if $R(C)$ is finite-dimensional, then with a suitable change of notation $\Lambda(\lambda)$ and $\Lambda^{-1}(\lambda)$ can be expressed as matrices whose scalar components are functions of λ. The problem of factoring $\Lambda(\lambda)$ is then reduced to the simpler problem of factoring a matrix.

As an illustration, we shall consider the transport problem (1.1)–(1.3) for the case of one-group, isotropic scattering. We may then simply ignore the energy variable E and take:

$$\sigma_t(E)=1, \quad \sigma(\mu',E'\to\mu,E)=c/2, \tag{7.1a}$$

$$(Cf)(\mu)=\frac{c}{2}\int_{-1}^{1}f(\mu')\,d\mu', \tag{7.1b}$$

$$\psi(0,\mu)=\psi_0(\mu), \quad 0<\mu\leq 1. \tag{7.1c}$$

Thus, $R(C)$ is the one-dimensional subspace of constant functions of μ. If $f(\mu)=a=\text{const}$, then by Eq. (2.8),

$$[\Lambda(\lambda)f](\mu)=a-\frac{\lambda c}{2}\int_{-1}^{1}\frac{a}{\lambda-\mu'}d\mu'$$

$$=\Lambda(\lambda)f(\mu), \quad f\in R(C). \tag{7.2}$$

Here

$$\Lambda(\lambda)=1-\frac{\lambda c}{2}\int_{-1}^{1}\frac{d\mu'}{\lambda-\mu'}$$

is the usual one-group dispersion function, which has two real zeroes at $\pm\nu_0$ for $0<c<1$.[5] Thus by Eq. (7.2), the restriction of $\Lambda(\lambda)$ to $R(C)$ is simply a scalar multiplication operator.

The operators $\mathcal{X}(\lambda)$ and $\mathcal{Y}(\lambda)$ may thus be taken to multiplication operators. In fact, let us define the scalar function $X(\lambda)$ by

$$X(\lambda)=(1-c)^{1/2}\frac{\lambda-\nu_0}{\lambda-1}\exp\left\{\frac{1}{2\pi i}\int_0^1\ln\frac{\Lambda^+(s)}{\Lambda^-(s)}\frac{ds}{s-\lambda}\right\}$$

Then $X(\lambda)$ is both analytic and invertible everywhere except for the cut $[0,1]$ and the point $\lambda=\nu_0$, and $\Lambda(\lambda)=X(\lambda)X(-\lambda)$. (See Ref. 5.) We may therefore define

$$[\mathcal{X}(\lambda)f](\mu)=X(\lambda)f(\mu), \quad f\in R(C), \tag{7.3a}$$

$$[\mathcal{Y}(\lambda)f](\mu)=X(-\lambda)f(\mu), \quad f\in R(C), \tag{7.3b}$$

and then Eq. (2.16) is satisfied.

Using the above results, we may now explicitly write the solution of the one-group transport problem (1.1)–(1.3), (7.1). This solution is, by Eqs. (3.6), (3.5), and (3.3),

$$\psi(x,\mu)=\psi_0(\mu)\exp[-d(x,\mu)]+\int_0^1 H(\mu'\to x,\mu)\psi_0(\mu')\,d\mu'. \tag{7.4}$$

Here we have introduced

$$d(x,\mu)=\begin{cases}x/\mu, & 0<\mu\leq 1,\\ +\infty, & -1\leq\mu<0,\end{cases} \tag{7.5}$$

and

$$H(\mu'\to x,\mu)=\frac{\mu'}{2\pi i}\oint_{\Gamma^+}\frac{\exp(-x/\lambda)}{(\lambda-\mu)(\lambda-\mu')}\frac{1}{X(\lambda)}\frac{1}{X(-\mu)}d\lambda. \tag{7.6}$$

Furthermore, by Eqs. (2.30)–(2.32),

$$\psi(0,\mu)=\frac{c}{2X(\mu)}\int_0^1\frac{1}{X(-\mu')}\frac{\mu'}{\mu'-\mu}\psi_0(\mu')\,d\mu', \quad -1\leq\mu<0. \tag{7.7}$$

Equations (7.4)–(7.7) agree with the results of Ref. 11. [Note however that Eq. (4.9) of Ref. 11 is in error; this equation should be identical to the present Eq. (7.7).]

If $R(C)$ is not one-dimensional but rather is finite dimensional, then as indicated above, the operator $\Lambda(\lambda)$ may be reduced to a matrix $\Lambda(\lambda)$ by a suitable change of notation. This situation is treated in detail in Ref. 12 for the general problem of multigroup, degenerate anisotropic scattering.

8. DISCUSSION

We shall now discuss a few aspects of the above analysis. The two primary features of the present work are:

(1) We use a completely general scattering operator, assuming only the subcriticality condition, Eq. (1.5).

(2) We use a physically motivated Banach space [L_1 with respect to μ and E with a $\sigma_t(E)$ weight] as a setting for our analysis.

Previous studies of the transport problem (1.1)–(1.3) have employed simplified scattering operators (degenerate anisotropic, multigroup, etc.) and have often been carried out with the assumption of Hölder continuous or L_2 integrable functions. Our analysis avoids these simplifications, and arrives at a representation of the neutron flux $\psi(x, \mu, E)$ for all of the classical plane geometry transport problems: the half-space albedo and Milne problems, the half-space with a source, and the infinite medium with a source. We have also obtained bounds on the spectrum of the reduced transport operator K (see Fig. 1) for a completely general scattering model.

Our analysis has been carried out in a physically motivated Banach space setting. We have provided complete justification for various mathematical manipulations when that could be done without excessive detail. In order to avoid losing sight of the main ideas we have sometimes used formal arguments, such as in the interchange of integration and differentiation in deriving Eq. (2.12). However, we believe that these technical details are minor and have not included proofs in our presentation.

The main result of our analysis, the representation of the neutron flux as a contour integral, has the same form as has been obtained for simpler scattering models.[11,12] The curve Γ^+ can be replaced by individual contours enclosing disjoint parts of the spectrum of K lying in the right-half-plane. The result, in general, is a contribution from the line $[0, \lambda_0]$ which occurs for all scattering models, and contributions from the remaining singularities of $\mathbf{\Lambda}^{-1}(\lambda)$. For various simplified scattering models,[11,12] these singularities are known to be isolated point eigenvalues of K. However, we cannot rule out the possibility that for general scattering models, residual spectra or additional continuous spectra might arise from certain singularities of $\mathbf{\Lambda}^{-1}(\lambda)$.

The relationship between the contour integral method employed in this paper and the singular eigenfunction method is thoroughly discussed in Ref. 11, and can quickly be described by considering the solution (7.4)–(7.6) of the one-group, isotropic scattering transport problem. To derive the singular eigenfunction form of the solution, we may rewrite H in Eq. (7.6) by collapsing the contour Γ^+ about the line segment $[0, 1]$ and the simple pole $\lambda = \nu_0$. Then we may derive the singular eigenfunction solution by interchanging certain singular integrations and making use of the Poincaré–Bertrand formula.[18] To do this, however, we must require $\mu\psi_0(\mu)$ to be Hölder continuous [or, using additional analysis, to be L_p integrable for $p > 1$[22]]. Thus, while the singular eigenfunction solution is perhaps more elegant than the contour integral solution, it is also more restrictive; the boundary data ψ_0 must satisfy unphysical smoothness conditions for the singular eigenfunction solution to exist.

The present analysis is restricted to subcritical media, but we hope to treat the critical half-space problem in a future paper. Aside from its general interest, the solution of this problem is necessary to determine asymptotic boundary layers and the boundary conditions to the asymptotic diffusion equation for large, three-dimensional, nearly-critical media.[23–25]

ACKNOWLEDGMENTS

This research was supported by the United States Energy Research and Development Administration under contract E(45-1)1830, and by the Air Force Office of Scientific Research under contract 76-3076. Part of this research was performed at Courant Institute, and part at Battelle Laboratories. The authors wish to thank these institutions for their hospitality.

[1] G. I. Bell and S. Glasstone, *Nuclear Reactor Theory* (Van Nostrand, New York, 1970).
[2] N. J. McCormick and I. Kuščer, J. Math. Phys. **7**, 2036 (1966).
[3] E. Hopf, *Mathematical Problems of Radiative Equilibrium* (Cambridge U. P., Cambridge, 1933).
[4] B. Davison, *Neutron Transport Theory* (Oxford U. P., Oxford, 1957).
[5] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Massachusetts, 1967).
[6] N. J. McCormick and I. Kuščer, Adv. Nucl. Sci. Tech. 7, 181 (1973).
[7] M. Ribaric, "Functional-Analytic Concepts and Structures of Neutron Transport Theory," Slovene Academy of Sciences and Arts, Ljubljana, Yugoslavia (1973). Transport problems can also be considered in spaces of measures. See, for example, A. Sudaholc and I. Vidav, Math. Balkanica **3**, 514 (1973).
[8] E. W. Larsen and G. J. Habetler, Commun. Pure Appl. Math. **26**, 525 (1973); E. W. Larsen, Commun. Pure Appl. Math. **27**, 523 (1974).
[9] S. Sancaktar, Trans. Theor. Stat. Phys. **4**, 71 (1975).
[10] R. L. Bowden, S. Sancaktar, and P. F. Zweifel, J. Math. Phys. **17**, 76, 82 (1976).
[11] E. W. Larsen, Commun. Pure Appl. Math. **28**, 729 (1975).
[12] E. W. Larsen and P. F. Zweifel, J. Math. Phys. **17**, 1812 (1976).
[13] E. W. Larsen and P. F. Zweifel, SIAM J. Appl. Math. **30**, 732 (1976).
[14] T. W. Mullikin, Trans. Theor. Stat. Phys. **4**, 215 (1973).
[15] D. Victory, thesis, Purdue University (1974).
[16] R. L. Bowden, P. F. Zweifel, and R. Menikoff, J. Math. Phys. **17**, 1722 (1976). See also R. L. Bowden and P. F. Zweifel, "A Banach space analysis of the Chandrasekhar *H*-equation," Astrophys. J., to appear.

[17]C. T. Kelley, J. Math. Phys. **18**, 764 (1977).
[18]N. I. Muskhelishvili, *Singular Integral Equations* (Noordhoff, Groningen, 1953).
[19]A. E. Taylor, *Introduction to Functional Analysis* (Wiley, New York, 1967).
[20]I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Non-Self-Adjoint Operators* (Am. Math. Soc., Providence, 1969), p. 21.
[21]I. Kuščer and I. Vidav, J. Math. Anal. Appl. **25**, 80 (1969).
[22]E. W. Larsen, S. Sancaktar, and P. F. Zweifel, J. Math. Phys. **16**, 1117 (1975).
[23]E. W. Larsen and J. B. Keller, J. Math. Phys. **15**, 75 (1974).
[24]E. W. Larsen and J. D'Arruda, Phys. Rev. A **13**, 1933 (1976).
[25]E. W. Larsen, "Asymptotic theory of the linear transport equation for small mean free paths. II," SIAM J. Appl. Math., to appear.

A Green's function for a linear equation associated with solitons

J. P. Keener[a] and D. W. McLaughlin[b]

Department of Mathematics, University of Arizona, Tucson, Arizona 85721
(Received 15 September 1976; revised manuscript received 24 January 1977)

A linear equation associated with nonlinear wave equations which support solitons is analyzed. A complete set of solutions of this linear equation is described through the techniques of scattering theory. This set is used to construct an explicit representation of a Green's function for perturbation theory. The cases of the nonlinear Schrödinger and sine-Gordon equations are discussed in some detail.

1. DISCUSSION OF RESULTS

Consider a nonlinear wave equation which supports solitons,

$$\partial_t u + H(u) = 0, \tag{1.1}$$

together with its associated linear equation

$$L(u)\chi \equiv \partial_t \chi + [\delta H(u)]\chi = 0. \tag{1.2}$$

Here $H(\cdot)$ is a nonlinear differential operator in one space dimension, and (1.1) is assumed to admit solution by the inverse scattering transformation; $L(u)$ denotes the linearization of $\partial_t + H$ about u. The two main purposes of this paper are the identification of a complete set of solutions of (1.2) and the use of this set to represent a particular Green's function for equation (1.2).

Associated linear equations such as (1.2) have appeared in the theoretical literature on the inverse scattering transform where it has been observed that both "squared eigenfunctions" and "conserved densities" satisfy the linear equations.[1-3] Physically, the associated linear equations can be interpreted, very semiclassically, as a meson field which surrounds a quantum soliton.[4-6] For this interpretation, consider a "breather" solution u of the "sine–Gordon" equation

$$u_{tt} - u_{xx} = -\sin u.$$

The "meson field bound to this lump" satisfies

$$\chi_{tt} - \chi_{xx} = -(\cos u)\chi.$$

A slight generalization of the material in this paper enables one to display a complete set of solutions for this "meson field" χ. Although we have not used this set for field theory, we have used it in a study of fluxon–antifluxon annihilation on the Josephson junction transmission line.[7] Of course, the linear equation (1.2) is the first step in any linearized stability analysis,[4,8] and it is central in the classical perturbation description of solitons.[9,10] In fact, as we shall see later, a Green's function which naturally arises in classical perturbation calculations is a particular solution of this associated linear equation. In each of these areas, a clear description of the complete family of solutions for the associated linear equation will be very useful.

We denote the collection of all solutions of (1.2) by $N[L(u)]$, the null space of the linear operator $L(u)$. The standard methods of describing this null space begin with the assumption that u is a traveling wave,[11] which in our case is a *single* soliton.[4,9,10] In this special instance, u depends upon space and time only through the combination $x - ct$. This simplification permits, after translation of coordinates, the reduction of Eq. (1.2) to ordinary differential equations by separation of variables.[4,9] Such methods are restricted to traveling waveforms.

In this paper, by combining classical methods with very recent results in inverse scattering theory, we present an explicit description of the entire null space $N[L(u)]$ which is valid for any solution u of (1.1). Specifically, we use scattering theory to describe a complete set of solutions of (1.2), that is, a set of solutions which spans $N[L(u)]$. The existence of an explicit characterization of this basis is another of the remarkable properties of nonlinear wave equations which are solved by inverse scattering methods.

This work can be summarized in very general terms. Consider a solution u of (1.1) which depends upon a free parameter λ, that is, a parameter which does not appear in the nonlinear wave equation itself. Clearly, $\partial_\lambda u$ belongs to the null space $N[L(u)]$. Frequently, when working with partial differential equations, this observation is not particularly useful because too few parameters are explicitly available to yield a basis for the null space. However, for the exceptional cases which are rendered completely integrable by the inverse scattering transform, the initial values of the scattering data provide enough explicit parameters to generate a basis for $N[L(u)]$. It turns out that this basis consists of (appropriately normalized) squared eigenfunctions of inverse scattering theory.[12]

Our study emphasizes several important aspects of the null space. First, $N[L(u)]$ consists of two distinct components which we call the "discrete" and "continuous" subspaces. The discrete subspace is associated with the N-soliton components of the wave u and is finite dimensional, while the continuous subspace is associated with the continuous spectrum in the inverse scattering transform and is infinite dimensional. This structure of $N[L(u)]$ is at the heart of our study of perturbations of solitons.[13] Second, the derivation emphasizes formal connections between the associated linear problem (1.2) and the theory of inverse scattering. In fact, the reason that (appropriately normalized)

[a]Supported in part by NSF Grant MPS75-07621.
[b]Supported in part by NSF Grant MPS75-07530.

squared eigenfunctions solve (1.2) and the origin of the specific normalization are clearly identified in this paper.

One important use of the basis for $N[L(u)]$ is the constructuion of a Green's function of regular perturbation theory. Consider a slight perturbation of the nonlinear wave equation (1.1):

$$\partial_t u + H(u) = \epsilon F(u), \quad 0 < \epsilon \ll 1, \tag{1.3}$$

and seek a regular expansion of the form $u = u_0 + \epsilon u_1 + \cdots$. With $u = u_0$ at $t = 0$, the first order correction u_1 satisfies the initial value problem

$$\begin{aligned} &[L(u_0)]u_1 = \partial_t u_1 + [\delta H(u_0)]u_1 = F(u_0), \\ &u_1\big|_{t=0} = 0, \end{aligned} \tag{1.4}$$

and admits a representation of the form

$$u_1 = \int_0^t G(t,t')\, F[u_0(t')]\, dt', \tag{1.5}$$

where the Green's operator (function) G satisfies

$$\begin{aligned} &L(u_0)G = \partial_t G + [\delta H(u_0(t))]G = 0, \quad 0 < t' < t, \\ &\lim_{t \to t'} G(t,t') = I. \end{aligned} \tag{1.6}$$

Since G is a particular solution of (1.2), it can be represented in terms of the basis of squared eigenfunctions. This expansion provides a new use of squared eigenfunctions. The representation of the Green's function G, together with the structure of the null space $N[L(u)]$, provides the first step, and indeed the crucial step, in a perturbation theory for solitons. It remains to use this representation of G to identify secular terms in the regular perturbation expansion and to remove these secularities by modulating the parameters in the unperturbed waveform.[13]

In this paper we restrict our attention to the nonlinear Schrödinger equation and to the sine–Gordon equation in characteristic coordinates. However, our methods easily extend to the wide class of evolution equations solved in Ref. 14.

2. THE NONLINEAR SCHRÖDINGER CASE

Consider the nonlinear Schrödinger equation

$$(-i\partial_t + \partial_{xx} + 2|r|^2)r = 0, \tag{2.1}$$

and its associated linear equation

$$[L(\mathbf{r})]\chi = [-i\sigma_3\partial_t + h(\mathbf{r})]\chi = 0, \tag{2.2}$$

where it is convenient to consider (2.1) as a system for a vector with components r and r^*; the matrix operator h is given by

$$h \equiv \begin{pmatrix} \partial_{xx} + 4|r|^2 & 2r^2 \\ 2(r^*)^2 & -\partial_{xx} - 4|r|^2 \end{pmatrix},$$

and σ_3 denotes one of the Pauli spin matrices,

$$\sigma_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 \equiv \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In this case the Green's function is a matrix $G(x,t|x',t')$ which is defined by the final value problem in (x',t'):

$$\begin{aligned} &[-i\sigma_3\partial_{t'} + h_2(\mathbf{r})]G(x,t|x',t') = 0, \\ &\forall x, x' \in (-\infty,\infty), \ \forall t' \in (0,t), \end{aligned} \tag{2.3a}$$

$$\lim_{t' \uparrow t} G(x,t|x',t') = -i\sigma_3\delta(x-x'). \tag{2.3b}$$

Here

$$h_2 G(x,t|x',t') \equiv \begin{pmatrix} \dfrac{\partial_{x'x'} + 4|r^2(x',t')}{2[r^*(x',t')]^2} & \dfrac{2r^2(x',t')}{\partial_{x'x'} + 4|r^2(x',t')|} \end{pmatrix} G(x,t|x',t').$$

The solution of the inhomogeneous equation

$$[-i\sigma_3\partial_t + h(\mathbf{r})]\mathbf{r}_1 = \mathbf{f}, \quad \mathbf{r}_1(t=0) = 0,$$

is then given by

$$\mathbf{r}_1(x,t) = \int_0^t \int_{-\infty}^{\infty} [G(x,t|x',t')]^{\dagger} \mathbf{f}(x',t')\, dx'dt',$$

where $\dagger$ denotes the Hermitian conjugate. Each column of the Green's matrix belongs to the null space $N[L(\mathbf{r})]$.

The analysis of this null space follows from the direct scattering theory of the linear eigenvalue problem[14–16]

$$\begin{aligned} &(\partial_x + i\zeta)v_1 = Qv_2, \\ &(\partial_x - i\zeta)v_2 = Rv_1, \quad -\infty < x < \infty, \end{aligned}$$

where $R = -Q^* = r$. The *scattering data* $\mathcal{S}_+$ for this linear system is defined as

$$\begin{aligned} &\mathcal{S}_+ = S_+ \cup S_+^*, \\ &S_+ \equiv \{\rho_+(\xi),\ \xi \in (-\infty,\infty);\ \zeta_j, \gamma_j^+,\ j = 1,2,\ldots,N\}, \end{aligned} \tag{2.4}$$

where ρ_+ denotes a reflection coefficient, $\{\zeta_j\}$ denotes the discrete eigenvalues, $\{\gamma_j^+\}$ the normalization of the discrete eigenfunctions, and S_+^* consists of the complex conjugates of S_+. This notation is now fairly standard; it is summarized in the Appendix.

The foundation of the inverse scattering method is the map between $\mathbf{r}$ and the scattering data $\mathcal{S}_+$. At any fixed time t, this map is one-to-one and invertible. The evolution of $\mathbf{r}$ in t induces an equivalent temporal evolution of $\mathcal{S}_+$. In particular, when $\mathbf{r}$ solves the nonlinear Schroedinger equation, the temporal behavior of the scattering data is given explicitly by[15,14,16]

$$\begin{aligned} &S_+(t) = \{\rho_+(\xi,t) = \exp(4i\xi^2 t)\rho_+(\xi,0); \\ &\zeta_j(t) = \zeta_j(0),\ \gamma_j^+(t) = \exp(4i\zeta_j^2 t)\gamma_j^+(0)\}. \end{aligned} \tag{2.5}$$

Since the map from $\mathbf{r}$ to $\mathcal{S}_+$ is invertible, knowledge of $\mathcal{S}_+(t)$ is equivalent to knowledge of $\mathbf{r}(\cdot,t)$.

The main point of the preceding paragraph is that any solution of the nonlinear Schrödinger equation is parametrized by its scattering data at $t = 0$. The variation of $\mathbf{r}$ with respect to each one of these parameters will provide a member of the null space of L. That is, the set

$$\begin{aligned} &\frac{\delta \mathbf{r}(x,t)}{\delta \mathcal{S}_+(\cdot, t=0)} \equiv \left\{ \frac{\delta \mathbf{r}(x,t)}{\delta \rho_+(\xi,0)}, \ \frac{\delta \mathbf{r}(x,t)}{\delta \overline{\rho}_+(\xi,0)}, \ \xi \in (-\infty,\infty); \ \frac{\delta \mathbf{r}(x,t)}{\delta \zeta_j(t=0)}, \right. \\ &\left. \frac{\delta \mathbf{r}(x,t)}{\delta \overline{\zeta}_j(t=0)}, \ \frac{\delta \mathbf{r}(x,t)}{\delta \gamma_j^+(t=0)}, \ \frac{\delta \mathbf{r}(x,t)}{\delta \overline{\gamma}_j^+(t=0)}, \ j = 1,2,\cdots,N \right\} \end{aligned} \tag{2.6}$$

provides an infinite number of solutions of (2.3a).

TABLE I.

$\frac{\delta \mathbf{r}(x,t)}{\delta \rho_+(\xi,t)}$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\rho}_+(\xi,t)}$
$= -\frac{i}{\pi}\sigma_2 \Psi(x,t,\xi) \quad \forall \xi \in (-\infty,\infty)$	$= \frac{i}{\pi}\sigma_2 \bar{\Psi}(x,t;\xi) \quad \forall \xi \in (-\infty,\infty),$
$\frac{\delta \mathbf{r}(x,t)}{\delta \xi_j(t)} = -2\gamma_j^+ \sigma_2 \left(\frac{d}{d\xi}\Psi(x,t;\xi)\right)_{\xi_j}$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\xi}_j(t)} = -2\bar{\gamma}_j^+ \sigma_2 \left(\frac{d}{d\xi}\bar{\Psi}(x,t;\xi)\right)_{\bar{\xi}_j}$
$\frac{\delta \mathbf{r}(x,t)}{\delta \gamma_j^+(t)} = -2\sigma_2 \Psi(x,t;\xi_j)$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\gamma}_j^+(t)} = -2\sigma_2 \bar{\Psi}(x,t;\bar{\xi}_j)$
$\frac{\delta \mathbf{r}(x,t)}{\delta \rho_-(\xi,t)}$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\rho}_-(\xi,t)}$
$= -\frac{1}{\pi}\Psi^A(x,t;\xi) \quad \xi \in (-\infty,\infty),$	$= \frac{1}{\pi}\bar{\Psi}^A(x,t,\xi) \quad \xi \in (-\infty,\infty),$
$\frac{\delta \mathbf{r}(x,t)}{\delta \xi_j(t)} = 2i\gamma_j^- \left(\frac{d}{d\xi}\bar{\Psi}^A(x,t;\xi)\right)_{\xi_i}$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\xi}_j(t)} = 2i\bar{\gamma}_j^- \left(\frac{d}{d\xi}\bar{\Psi}^A(x,t;\xi)\right)_{\bar{\xi}_j}$
$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\gamma}_j(\xi,t)} = 2i\Psi^A(x,t;\xi_j)$	$\frac{\delta \mathbf{r}(x,t)}{\delta \bar{\gamma}_j^-(t)} = 2i\bar{\Psi}^A(x,t;\bar{\xi}_j).$

Moreover, this set consists of two distinct types of members—continuous and discrete components.

It is convenient to introduce an additional set of scattering data $\mathcal{S}_-$ which is equivalent to $\mathcal{S}_+$. (The origin of the $\pm$ subscripts is that the reflection coefficient ρ_+ is defined at $x = +\infty$, while ρ_- is defined at $x = -\infty$.) This scattering data is given by

$$\mathcal{S}_- \equiv S_- \cup \bar{S}_-, \tag{2.7a}$$
$$S_- \equiv \{\rho_-(\xi),\ \xi \in (-\infty,\infty);\ \xi_j,\ \gamma_j^-,\ j = 1,2,\ldots,N\},$$

where the exact definitions are given in the Appendix. Under the nonlinear Schrödinger flow, the temporal evolution of $\mathcal{S}_-$ is given by

$$S_-(t) = \{\rho_-(\xi,t) = \exp(-4i\xi^2 t)\rho_-(\xi,0);$$
$$\xi_j(t) = \xi_j(0),\ \gamma_j^-(t) = \exp(-4i\xi_j^2 t)\gamma_j^-(0)\}. \tag{2.7b}$$

The variation of $\mathbf{r}$ with respect to $\mathcal{S}_-(t=0)$ provides another infinite set of solutions of (2.3a). The realization that either of these sets is complete follows after the variations of $\mathbf{r}$ are expressed in terms of "squared eigenfunctions,"

$$\Psi(x,t;\xi) \equiv \begin{pmatrix} \psi_1^2(x,t;\xi) \\ \\ \psi_2^2(x,t;\xi) \end{pmatrix}, \quad \bar{\Psi}(x,t;\xi) \equiv \begin{pmatrix} \bar{\psi}_1^2(x,t;\xi) \\ \\ \bar{\psi}_2^2(x,t;\xi) \end{pmatrix},$$
$$\Psi^A(x,t;\xi) \equiv -i\sigma_2 \Psi(x,t;\xi), \quad \bar{\Psi}^A(x,t;\xi) \equiv -i\sigma_2 \bar{\Psi}(x,t;\xi).$$

Here (ψ_1,ψ_2) and $(\bar{\psi}_1,\bar{\psi}_2)$ denote the components of the eigenvectors ψ and $\bar{\psi}$, which are defined in the Appendix. Table I, which expresses $\delta \mathbf{r}(x,t)/\delta \mathcal{S}_+(t)$ in terms of these squared eigenfunctions, was obtained by Newell[17]; analogous expressions for the direct scattering theory of the linear Schroedinger equation may be found in Ref. 18, where they were obtained by a somewhat different method.

The infinite collection of members of the null space of $L[\mathbf{r}]$ can now be expressed in terms of squared eigenfunctions using Table I, the chain rule, and the temporal evolution of the scattering data. For example,

$$\frac{\delta \mathbf{r}(x,t)}{\delta S_-(\xi,0)} = \frac{\delta \mathbf{r}(x,t)}{\delta \rho_-(\xi,0)} = \frac{\delta \rho_-(\xi,t)}{\delta \rho_-(\xi,0)} \cdot \frac{\delta \mathbf{r}(x,t)}{\delta \rho_-(\xi,t)}$$
$$= -\frac{1}{\pi}\exp(-4i\xi^2 t)\Psi^A(x,t;\xi). \tag{2.8}$$

Thus, the squared eigenfunction Ψ^A, when multiplied by the factor $\exp(-4i\xi^2 t)$, is a solution of the linearized equation (2.3a). Gardner, Greene, Kruskal, and Miura[19] note an analogous result for the Korteweg–de Vries equation. We believe that our derivation emphasizes the true significance of their observations about solutions of "associated linear problems." For example, the multiplicative factor $\exp(-4i\xi^2 t)$ naturally arises from the conversion of variations with respect to the parameters $\mathcal{S}_-(0)$ into variations with respect to the functions of time $\mathcal{S}_-(t)$.

In this manner the infinite collections $\delta \mathbf{r}(x,t)/\delta \mathcal{S}_\pm(\cdot, t=0)$, which belong to $N[L(\mathbf{r})]$, can be expressed in terms of squared eigenfunctions. Since Kaup[12] has established the completeness of the squared eigenfunctions, we have, for each fixed t as functions of x, two complete sets of solutions. Either of these sets will span $N[L(\mathbf{r})]$. The actual completeness relation is given by[12]

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\delta(x-x') = -\frac{1}{\pi}\int_{C_a} \frac{1}{[a(\xi)]^2}\Psi^A(x',t;\xi)[\Psi(x,t;\xi)]^T d\xi$$
$$+\frac{1}{\pi}\int_{C_b} \frac{1}{[\bar{a}(\xi)]^2}\bar{\Psi}^A(x,t;\xi)[\bar{\Psi}(x,t;\xi)]^T d\xi, \tag{2.9}$$

where the contours of integration run from $-\infty$ to ∞ along the real ξ axis, the first (C_a) indented above all zeros of $a(\xi)$, the second (C_b) below all zeros of $\bar{a}(\xi)$.

Finally, we derive a representation of the Green's function G in terms of this basis of $N[L(u)]$. First, we seek G in the form

$$G(x,t\,|\,x',t') = \int_{C_a} \frac{\delta \mathbf{r}(x',t')}{\delta S_-(\xi,t'=0)} A^T(x,t;\xi)\sigma_3 d\xi$$
$$+\int_{C_b} \frac{\delta \mathbf{r}(x',t')}{\delta S_-^*(\xi,t'=0)} \bar{A}^T(x,t;\xi)\sigma_3 d\xi, \tag{2.10}$$

where the expansion coefficients A and $\bar{A}$ are column vectors to be found. Notice that this formula is a linear combination of members of $N[L(\mathbf{r})]$; hence, it certainly solves (2.3a) as a function of (x',t'). The expansion coefficients A and $\bar{A}$ must be selected to satisfy the final data (2.3b). Using this data $\lim_{t' \to t} G = -i\sigma_3 I$ yields

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\delta(x-x') = -\frac{i}{\pi}\int_{C_a} \Psi^A(x',t;\xi)\exp(-4i\xi^2 t)A^T(x,t;\xi)d\xi$$
$$+\frac{i}{\pi}\int_{C_b} \bar{\Psi}^A(x',t;\xi)\exp(4i\xi^2 t)\bar{A}^T(x,t;\xi)d\xi, \tag{2.11}$$

where we have expressed the variations in terms of squared eigenfunctions. Comparison of (2.11) with (2.9) shows

$$A(x,t;\xi) = \{-i\exp(4i\xi^2 t)/[a(\xi)]^2\}\Psi(x,t;\xi),$$
$$\bar{A}(x,t;\xi) = \{-i\exp(-4i\xi^2 t)/[\bar{a}(\xi)]^2\}\bar{\Psi}(x,t;\xi).$$

Using these expansion coefficients, we obtain the following equivalent representations of G:

$$G(x,t|x',t') = i\pi \int_{C_a} \frac{1}{[a(\xi)]^2} \left(\frac{\delta \mathbf{r}(x',t')}{\delta S_-(\xi,0)}\right)\left(\frac{\delta \mathbf{r}(x,t)}{\delta \bar{S}_+(\xi,0)}\right)^\dagger d\xi + i\pi \int_{C_b} \frac{1}{[\bar{a}(\xi)]^2} \left(\frac{\delta \mathbf{r}(x',t')}{\delta \bar{S}_-(\xi,0)}\right)\left(\frac{\delta \mathbf{r}(x,t)}{\delta S_+(\xi,0)}\right)^\dagger d\xi; \tag{2.12a}$$

$$G(x,t|x',t') = -\frac{i}{\pi}\int_{C_a} \frac{\exp[-4i\xi^2(t'-t)]}{[a(\xi)]^2} \Psi^A(x',t';\xi)\Psi^T(x,t;\xi)\sigma_3 d\xi - \frac{i}{\pi}\int_{C_b} \frac{\exp[4i\xi^2(t'-t)]}{[\bar{a}(\xi)]^2} \bar{\Psi}^A(x',t';\xi)\bar{\Psi}^T(x,t;\xi)\sigma_3 d\xi; \tag{2.12b}$$

$$G(x,t|x',t') = G_c(x,t|x',t') + G_d(x,t|x',t'),$$

where

$$G_c(x,t|x',t') = \frac{i}{\pi}\int_{-\infty}^{\infty} \frac{\exp[-4i\xi^2(t'-t)]}{[a(\xi)]^2} \Psi^A(x',t';\xi)\Psi^T(x,t;\xi)\sigma_3 d\xi - \frac{i}{\pi}\int_{-\infty}^{\infty} \frac{\exp[+4i\xi^2(t'-t)]}{[\bar{a}(\xi)]^2} \bar{\Psi}^A(x',t';\xi)\bar{\Psi}^T(x,t;\xi)\sigma_3 d\xi$$

and

$$G_d(x,t|x',t') = 2\sum_{j=1}^{N}\left\{\frac{1}{[a_j']^2}\frac{d}{d\xi}(\exp[-4i\xi^2(t'-t)]\Psi^A(x',t';\xi)\Psi^T(x,t;\xi))_{\xi=\zeta_j}\sigma_3 - \frac{a_j''}{[a_j']^3}\exp[-4i\zeta_j^2(t'-t)]\,\Psi^A(x',t';\zeta_j)\Psi^T(x,t;\zeta_j)\sigma_3 + \frac{1}{[\bar{a}_j']^2}\frac{d}{d\xi}(\exp[+4i\xi^2(t'-t)]\,\bar{\Psi}^A(x',t';\xi)\Psi^T(x,t;\xi))_{\xi=\bar{\zeta}_j}\sigma_3 - \frac{\bar{a}_j''}{[\bar{a}_j']^3}\exp[4i\bar{\zeta}_j^2(t'-t)]\bar{\Psi}^A(x',t';\bar{\zeta}_j)\bar{\Psi}^T(x,t;\zeta_j)\sigma_3\right\}. \tag{2.12c}$$

We have listed representation (2.12a) because it is most apparent from this form that G solves the linearized equation (2.3a) in (x',t') and that $G^\dagger$ solves the same equation in (x,t). When displaying the final data as $t' \uparrow t$, representation (2.12b) is most useful as our derivation indicates. Representation (2.12c) is the most useful in the actual perturbation calculations.[13] Notice in particular the separation into discrete and continuous components. Finally, we remark that when $\mathbf{r}$ is a pure N-soliton solution, explicit representations for both the discrete and continuous squared eigenfunctions exist.[15]

3. SINE-GORDON CASE

Consider the sine–Gordon equation in characteristic coordinates,

$$\partial_{xt}\mathcal{E} - \sin\mathcal{E} = 0, \quad t \geq 0, \quad -\infty < x < +\infty, \tag{3.1a}$$

subject to the data

$$\mathcal{E}(x,t=0) = \text{given} \quad \forall x < +\infty,$$
$$\mathcal{E}(+\infty,t) = \text{given} \quad \forall t > 0.$$

The initial data is constrained to approach an integer multiple of (2π) as $x \to \infty$. In this case, the solution of the inhomogeneous equation

$$\partial_{xt}\mathcal{E}_1 - [\cos\mathcal{E}_0]\mathcal{E}_1 = f, \tag{3.2a}$$
$$\mathcal{E}_1(x,t=0) = 0 \quad \forall x < +\infty, \quad \mathcal{E}_1(x=+\infty,t) = 0 \quad \forall t > 0 \tag{3.2b}$$

is given by

$$\mathcal{E}_1(x,t) = \int_0^t \int_{-\infty}^{\infty} G(x,t|x',t') f(x',t')\,dx'\,dt', \tag{3.3}$$

where the Green's function G is defined by

$$\{\partial_{x't'} - \cos[\mathcal{E}(x',t')]\}G(x,t|x',t') = 0, \quad 0 \leq t' \leq t, \quad -\infty < x, x' < \infty, \tag{3.4a}$$
$$\lim_{x' \to -\infty} G(x,t|x',t') = 0, \quad \lim_{t' \uparrow t} \partial_{x'}G(x,t|x',t') = -\delta(x-x'). \tag{3.4b}$$

Moreover, it follows from equations (3.2) and (3.3) that, as a function of (x,t), G must satisfy

$$\{\partial_{xt} - \cos[\mathcal{E}(x,t)]\}G(x,t|x',t') = 0, \quad 0 \leq t' \leq t, \quad -\infty < x, x' < \infty, \tag{3.5a}$$
$$\lim_{x \to +\infty} G(x,t|x',t') = 0, \quad \lim_{t \downarrow t'} \partial_x G(x,t|x',t') = \delta(x-x'). \tag{3.5b}$$

All three characteristic problems (3.2), (3.4), (3.5) are depicted in Fig. 1.

In order to construct the Green's function G we must study the null space of the linear operator $L = \partial_{xt} - \cos\mathcal{E}$, where $\mathcal{E}$ is a fixed solution of the sine–Gordon equation. As in the nonlinear Schrödinger case, we use methods from scattering theory. The same eigenvalue problem applies with $Q = -R = -\frac{i}{2}\mathcal{E}_x$.

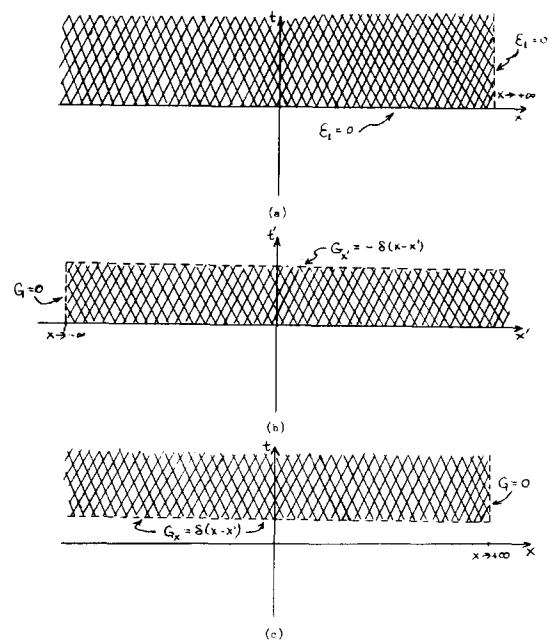


FIG. 1. Characteristic problems for $\mathcal{E}_1$, $G(\cdot,\cdot|x',t')$, $G(x,t|\cdot,\cdot)$. Solutions sought in shaded regions.

TABLE II.

$\frac{\delta \mathcal{E}_x(x,t)}{\delta \rho_+(\xi,t)} = \frac{2}{\pi}\Psi_-(x,t;\xi)$	$\frac{\delta \mathcal{E}_x(x,t)}{\delta \rho_-(\xi,t)} = -\frac{2}{\pi}\Phi_-(x,t;\xi)$
$\frac{\delta \mathcal{E}_x(x,t)}{\delta \gamma_j^+(t)} = -4i\Psi_-(x,t;\zeta_j)$	$\frac{\delta \mathcal{E}_x(x,t)}{\delta \gamma_j^-(t)} = 4i\Phi_-(x,t;\zeta_j)$
$\frac{\delta \mathcal{E}_x(x,t)}{\delta \zeta_j(x)} = -4i\gamma_j^+ \frac{d}{d\xi}\Psi_-(x,t;\xi)\Big\|_{\xi_j}$	$\frac{\delta \mathcal{E}_x(x,t)}{\delta \zeta_j(t)} = 4i\gamma_j^- \frac{d}{d\xi}\Phi_-(x,t;\xi)\Big\|_{\xi_j}$
$\frac{\delta \mathcal{E}(x,t)}{\delta \rho_+(\xi,t)} = -\frac{i}{\pi\xi}\Psi_+(x,t;\xi)$	$\frac{\delta \mathcal{E}(x,t)}{\delta \rho_-(\xi,t)} = \frac{i}{\pi\xi}\Phi_+(x,t;\xi)$
$\frac{\delta \mathcal{E}(x,t)}{\delta \gamma_j^+(t)} = \frac{2}{\zeta_j}\Psi_+(x,t;\zeta_j)$	$\frac{\delta \mathcal{E}(x,t)}{\delta \gamma_j^-(t)} = \frac{2}{\zeta_j}\Phi_+(x,t;\zeta_j)$
$\frac{\delta \mathcal{E}(x,t)}{\delta \zeta_j(t)} = -2\gamma_j^+ \frac{d}{d\xi}\left(\frac{\Psi_+(x,t;\xi)}{\xi}\right)_{\xi=\xi_j}$	$\frac{\delta \mathcal{E}(x,t)}{\delta \zeta_j(t)} = 2\gamma_j^- \frac{d}{d\xi}\left(\frac{\Phi_+(x,t;\xi)}{\xi}\right)_{\xi=\xi_j}$

In this case, as $\mathcal{E}$ evolves according to the unperturbed sine–Gordon equation, the scattering data evolves as[14]

$$S_+(t) = \{\rho_+(\xi,t) = e^{-it/2\xi}\rho_+(\xi,0);$$
$$\zeta_j(t) = \zeta_j(0),\ \gamma_j^+(t) = \exp(-it/2\xi_j)\gamma_j^+(0)\}. \tag{3.6}$$

As before, S_- is equivalent to S_+ and evolves as

$$S_-(t) = \{\rho_-(\xi,t) = e^{it/2\xi}\rho_-(\xi,0);$$
$$\xi_j(t) = \xi_j(0),\ \gamma_j^-(t) = \exp(it/2\xi_j)\gamma_j^-(0)\}. \tag{3.7}$$

Finally, in Table IIa, the variations of $\mathcal{E}_x$ with respect to the scattering data are given in terms of[16]

$$\Psi_\pm(x,t;\xi) \equiv \psi_2^2(x,t;\xi) \pm \psi_1^2(x,t;\xi),$$
$$\Phi_\pm(x,t;\xi) \equiv \phi_2^2(x,t;\xi) \pm \phi_1^2(x,t;\xi).$$

Here (ψ_1,ψ_2) and (ϕ_1,ϕ_2) are components of the eigenvectors ψ and ϕ defined in the Appendix. Moreover, by using the identities

$$\phi_-(x,t;\zeta) = (1/2i\zeta)\partial_x\Phi_+(x,t;\zeta),$$
$$\Psi_-(x,t;\zeta) = (1/2i\zeta)\partial_x\Psi_+(x,t;\zeta),$$

which follow directly from the linear eigenvalue problem, the variations of $\mathcal{E}_x$ may be integrated to yield variations of $\mathcal{E}$. These are summarized in Table IIb.

As above, the collection

$$\frac{\delta \mathcal{E}(x,t)}{\delta S_-(\cdot,0)} \equiv \left\{\frac{\delta \mathcal{E}(x,t)}{\delta \rho_-(\xi,0)},\ \xi\in(-\infty,\infty);\right.$$
$$\left.\frac{\delta \mathcal{E}(x,t)}{\delta \zeta_j(0)},\ \frac{\delta \mathcal{E}(x,t)}{\delta \gamma_j^-(0)},\ j=1,2,\cdots,N\right\}$$

constitutes an infinite set of members of the null space of L. Using Table II and the chain rule, we are lead to seek the Green's function in the form

$$G(x,t|x',t') = \int_{C_a} e^{i(t'-t)/2\zeta}A(x,t;\zeta)\Phi_+(x',t';\zeta)d\zeta.$$

Clearly G satisfies the linearized equation (3.4a). It remains to satisfy the data (3.5b). As $t'\uparrow t$, G_x approaches $-\delta(x-x')$,

$$\delta(x-x') = -2i\int_{C_a}\zeta A(x,t;\zeta)\Phi_-(x',t;\zeta)\,d\zeta. \tag{3.8}$$

Under symmetries (A2), the resolution of the identity (2.9) takes the form

$$\begin{pmatrix}1&0\\0&1\end{pmatrix}\delta(x-x') = -\frac{1}{\pi}\int_{C_a}\frac{1}{[a(\zeta)]^2}$$
$$\times[\Psi^A(x',t;\zeta)\Psi^T(x,t;\zeta) + \sigma_1\Psi^A(x',t;\zeta)\Psi^T(x,t;\zeta)\sigma_1]\,d\zeta.$$

Multiplying from the left by the vector (1.1) yields the completeness relation

$$\delta(x-x') = -\frac{1}{\pi}\int_{C_a}\frac{1}{[a(\zeta)]^2}\Psi_+(x,t;\zeta)\Phi_-(x',t;\zeta)\,d\zeta. \tag{3.9}$$

Comparing (3.8) with (3.9) yields

$$A(x,t;\zeta) = \{1/2\pi i\zeta[a(\zeta)]^2\}\Psi_+(x,t;\zeta),$$

These expansion coefficients yield the representation of G:

$$G(x,t|x',t') = -\frac{i\pi}{2}\int_{C_a}\frac{\zeta}{[a(\zeta)]^2}\left(\frac{\delta \mathcal{E}(x,t)}{\delta \rho_+(\xi,0)}\right)\left(\frac{\delta \mathcal{E}(x',t')}{\delta \rho_-(\xi,0)}\right)d\zeta, \tag{3.10a}$$

$$G(x,t|x',t') = \frac{1}{2\pi i}\int_{C_a}\frac{\exp[i(t'-t)/2\zeta]}{\zeta[a(\zeta)]^2}\Psi_+(x,t;\zeta)\Phi_+(x',t';\zeta)\,d\zeta. \tag{3.10b}$$

Finally, notice that the asymptotic conditions $\lim_{x\to+\infty} G = \lim_{x\to-\infty} G = 0$ are indeed satisfied if we define the contour (C_a) to follow $\zeta = \xi + i0^+$ as $\xi \to \pm\infty$.

ACKNOWLEDGMENT

It is our pleasure to acknowledge many helpful conversations with H. Flaschka and A.C. Scott during the development of this work. In addition, we wish to thank D. Kaup and A. Newell for keeping us informed of their work. In praticular, Refs. 12 and 17 were extremely helpful.

APPENDIX

In this appendix, we summarize our notation for the direct scattering theory.[14,16] Consider the eigenvalue problem

$$(\partial_x + i\zeta)v_1 = Qv_2,$$
$$(\partial_x - i\zeta)v_2 = Rv_1, \quad -\infty < x < +\infty.$$

For real ζ, define two pairs of solutions:

$$\phi(x,\zeta)\approx\begin{pmatrix}1\\0\end{pmatrix}e^{-i\zeta x};\ \bar\phi(x,\zeta)\approx\begin{pmatrix}0\\-1\end{pmatrix}e^{i\zeta x}\ \text{as } x\to-\infty;$$

$$\psi(x,\zeta)\approx\begin{pmatrix}0\\1\end{pmatrix}e^{i\zeta x};\ \bar\psi(x,\zeta)\approx\begin{pmatrix}1\\0\end{pmatrix}e^{-i\zeta x}\ \text{as } x\to+\infty.$$

For real ζ the pair $(\phi,\bar\phi)$ consists of linearly independent solutions as does the pair $(\psi,\bar\psi)$. These solutions are related by

$$\phi(x,\zeta) = a(\zeta)\bar\psi(x,\zeta) + b(\zeta)\psi(x,\zeta),$$
$$\bar\phi(x,\zeta) = -\bar a(\zeta)\psi(x,\zeta) + \bar b(\zeta)\bar\psi(x,\zeta),\quad \mathrm{Im}(\zeta) = 0.$$

The expansion coefficients $(a,b,\bar a,\bar b)$ can be obtained from the asymptotic behavior

$$\phi(x,\zeta)\approx\begin{pmatrix}a(\zeta)e^{-i\zeta x}\\ b(\zeta)e^{i\zeta x}\end{pmatrix},$$

$$\bar{\phi}(x,\zeta)\approx\begin{pmatrix}\bar{b}(\zeta)e^{-i\zeta x}\\ -\bar{a}(\zeta)e^{i\zeta x}\end{pmatrix},\quad \text{as } x\to+\infty,\ \mathrm{Im}(\zeta)=0.$$

The coefficient $a(\zeta)$ admits an analytic continuation into the upper half ζ plane where its only zeros occur at the bound state eigenvalues ζ_j. At these eigenvalues, ϕ and ψ are linearly dependent,

$$\phi(x,\zeta_i)=b_i\psi(x,\zeta_i),$$

a dependence which defines the normalization constant b_i. Similarly, the coefficient $\bar{a}(\zeta)$ admits a continuation into the lower half ζ plane where its only zeros occur at $\bar{\zeta}_i$, and $\bar{b}_i$ is defined by $\bar{\Phi}(x,\bar{\zeta}_i)=\bar{b}_i\bar{\psi}(x,\bar{\zeta}_i)$. Finally, the symmetry $Q=-R^*$ yields the relationships

$$\begin{aligned}
&\bar{\phi}(x,\zeta)=-i\sigma_2[\phi(x,\zeta^*)]^*, && \mathrm{Im}(\zeta)\le 0,\\
&\bar{\psi}(x,\zeta)=-i\sigma_2[\psi(x,\zeta^*)]^*, && \mathrm{Im}(\zeta)\le 0,\\
&\bar{a}(\zeta)=[a(\zeta^*)]^*, && \mathrm{Im}(\zeta)\le 0,\\
&\bar{b}(\zeta)=[b(\zeta)]^*, && \mathrm{Im}(\zeta)=0,\\
&\bar{\zeta}_j=[\zeta_j]^*,\quad \bar{b}_j=[b_j]^*,
\end{aligned}\tag{A1}$$

while the symmetry $Q=-R=-R^*$ yields

$$\begin{aligned}
&\bar{\phi}(x,\zeta)=-i\sigma_2\phi(x,-\zeta), && \mathrm{Im}(\zeta)\le 0\\
&\bar{\psi}(x,\zeta)=-i\sigma_2\psi(x,-\zeta), && \mathrm{Im}(\zeta)\le 0\\
&\bar{a}(\zeta)=a(-\zeta), && \mathrm{Im}(\zeta)\le 0\\
&\bar{b}(\zeta)=b(-\zeta), && \mathrm{Im}(\zeta)=0\\
&\bar{\zeta}_j=-\zeta_j,\quad \bar{b}_j=-b_j,
\end{aligned}\tag{A2}$$

In terms of these quantities the scattering data is defined by

$$\rho_+(\xi)\equiv b(\xi)/a(\xi),\quad \gamma_j^+\equiv b_j/a_j',$$
$$\rho_-(\xi)\equiv\bar{b}(\xi)/a(\xi),\quad \gamma_j^-\equiv[(a_j')^2\gamma_j^+]^{-1},$$

where $a_j'\equiv(d/d\zeta)a(\zeta)|_{\zeta_j}$.

[1]C.S. Gardner, J.M. Greene, M.D. Kruskal, and R. Miura, Comm. Pure Appl. Math. 27, 97 (1974). Their observations can be found in Sec. 3.3.3 of this reference and are actually for the adjoint of L.
[2]H. Flaschka, private communication.
[3]M. Wadati, "Invariances and Conservation Laws of the Korteweg—de Vries Equation," preprint, Dept. of Physics, Univ. of Alberta (1977).
[4]S. Coleman, "Classical Lumps and Their Quantum descendants," lectures delivered at the 1975 International School of Subnuclear Physics "Ettore Majorana", preprint, Dept. of Physics, Harvard University (1975). See, in particular, Eq. (2.16) and the following few paragraphs.
[5]R. Dashen, B.Hasslacher, and A. Neveu, Phys. Rev. D **11**, 3424 (1975).
[6]B. Hasslacher and A. Neveu, in Proc. N.S.F. Conf. on Solitons, Tucson, Arizona, edited by H. Flaschka and D.W. McLaughlin (1976).
[7]D.W. McLaughlin and A.C. Scott, "Fluxon Interactions," Appl. Phys. Lett. **30**, 545 (1977).
[8]A.C. Scott, F.Y. F. Chu, and D.W. McLaughlin, Proc. IEEE **61**, 1443 (1973).
[9]M.B. Fogel, S.E. Trullinger, A.R. Bishop, and J.A. Krumhansl, Phys. Rev. Lett. **36**, 1411 (1976). Also preprint (1976).
[10]R.M. Miura and M.P. Kruskal, SIAM J. Appl. Math. **26**, 376 (1974).
[11]G.B. Whitham, *Linear and Nonlinear Waves* (Wiley-Interscience, New York, 1974), Sec. 14.4.
[12]D.J. Kaup, J. Math. Anal. Appl. **54**, 849 (1976).
[13]J.P. Keener and D.W. McLaughlin, "Solitons under Perturbations," preprint, Univ. of Arizona (1976).
[14]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1974).
[15]V.E. Zakharov and A.B. Shabat, Zh. Eksp. Teor. Fiz. [Sov. Phys. JETP **34**, 62 (1972)].
[16]H. Flaschka and A.C. Newell, in *Dynamical Systems, Theory and Applications*, Battelle Seattle 1974 Rencontres, edited by J. Moser (Springer, New York, 1975), p. 441.
[17]A.C. Newell, in Proc. N.S.F. Conf. on Solitons, held in Tucson, Arizona, edited by H. Flaschka and D.W. McLaughlin (1976).
[18]V.E. Zakharov and L.D. Faddeev, Funct. Anal. Appl. **5**, 280 (1972).
[19]Ref. 1, Eq. (3.29).

Path structures on manifolds

Jürgen Ehlers and Egon Köhler

Max-Planck-Institut für Physik und Astrophysik, München, F. R. Germany
(Received 7 February 1977)

We study collections of paths—i.e., unparametrized curves—on a manifold such that through every point and every direction at that point there passes exactly one path. Among such path structures we characterize, analytically and in terms of symmetries, those which consist of geodesics of a linear connection. Examples of nongeodesic path structures are given, and some of the results are interpreted physically.

An axiomatic approach to the spacetime structure of general relativity can be based on the following concepts: a four-dimensional manifold M, a Lorentzian conformal structure C on M, and a projective structure P on M compatible with C.[1] The points of M are considered as mathematical images of events in physical spacetime, the conformal structure represents the causal structure as indicated in particular by the propagation of light, and the projective structure represents inertia which, according to Einstein, is identical with gravity, and which manifests itself particularly simply in free fall motions of test particles.

Causal structures per se have been investigated in detail by several authors.[2] Similar studies concerning inertial structures which might lead to a deeper understanding of Einstein's law of inertia seem to be lacking.

The world line of a freely falling neutral test particle is uniquely determined by one of its event and its direction at that event. One is, therefore, led to consider collections of paths—unparametrized curves (see Sec. 1)—such that given a point p and a direction ξ at p, there is exactly one path going through p in the direction ξ. The standard example of such a *path structure* consists of the paths represented by geodesics of a linear connection. The question thus arises: How can one characterize, analytically and geometrically, the geodesic path structures among the general ones? One would like particularly to have characterizations which can be interpreted physically.

In this paper we shall give one convenient analytic characterization of geodesic path structures (Theorem 1 in Sec. 3) and two geometrical ones (Theorems 2 and 3 in Sec. 4). According to Theorem 3, Einstein's law of inertia (using geodesics to represent free fall world lines) is equivalent to the statement: The set of free fall world lines is a path structure with the following property (of "local isotropy"): The collection of all path elements passing through an event p is invariant under a group of local diffeomorphisms which acts transitively on the set of all bases of the direction space (see Sec. 1) at p. (A path element is an equivalence class of paths which have a second-order contact at p; intuitively it is a second-order infinitesimal piece of a path.) So, roughly speaking, the set of all free fall world lines has the highest possible degree of local isotropy which a path structure can possess.

In Secs. 1 and 2 we develop the appropriate concepts to deal with path structures. Sections 3 and 4 contain the main results indicated above, and Sec. 5 gives some examples of nongeodesic path structures.

1. DIRECTIONS AND PATHS

Let M be a connected, paracompact Hausdorff manifold of class C^k ($k \geq 3$) and dimension $n \geq 2$. By T_pM we denote its tangent space at $p \in M$, and by π_T: $TM \to M$ its tangent bundle. Every C^k curve γ: $I \to M$ ($I \subset \mathbb{R}$) has a (canonical) lift γ^T: $I \to TM$, which we call a *special* curve in TM. The vectors $\dot{\gamma}^T(s) \in T_{\gamma^T(s)}TM$ are the tangents to γ^T. We call a vector $X \in T^2M$ a *special* vector, if it can be obtained as a tangent vector of a special curve in TM.

If π_{T*} is the differential of π_T and Π_T the bundle projection Π_T: $T^2M \to TM$, then special vectors are characterized by $\pi_{T*}X = \Pi_T X$. A special C^1 vector field X: $TM \to T^2M$ is called a differential equation of second order. Its integral curves are special curves in TM.[3]

In the sequel we shall specialize the concept of a "curve in M" by requiring: If $\gamma(s_1) = \gamma(s_2)$ and $\dot{\gamma}(s_1) = \dot{\gamma}(s_2)$, then there exist open intervals I_1, I_2 and a smooth, invertible map $\mu: I_1 \to I_2$ such that $\mu(s_1) = s_2$ and $\gamma | I_1 = (\gamma \circ \mu) | I_1$. (This excludes self-tangency.)

In many physical applications the parametrization of a curve is arbitrary or not specified *a priori*. Therefore, we need the concept of a *path* Γ as a "curve without parameter" or, more rigorously, as an equivalence class of curves with nowhere vanishing tangents which differ only by a parameter transformation.[4] Each curve γ defines a path $\underset{\sim}{\gamma}$. γ is a representative of Γ, $\gamma \in \Gamma$, if $\underset{\sim}{\gamma} = \Gamma$.

We call an equivalence class ξ of vectors $X, Y, \cdots \in T_pM$ which are proportional, $X = \lambda Y$ ($\lambda \neq 0$), a *direction* and write $\xi = \underline{X} = \underline{Y} = \cdots$. The directions generated by all vectors of $T_pM \setminus \{0\}$ form the direction space D_pM which is isomorphic to $\mathbb{P}^{n-1}$. A path $\Gamma = \underset{\sim}{\gamma}$ has a (unique) direction $\underline{\dot{\gamma}(s)}$ at each of its points $\gamma(s)$.

The collection of all D_pM over M forms another fiber bundle DM with projection π_D and compact fibers D_pM. For every curve γ in M with $\dot{\gamma} \neq 0$ there exists a canonical lift $\underline{\dot{\gamma}}$ into DM, defined by

$$\underline{\dot{\gamma}}: I \to DM: s \to \underline{\dot{\gamma}(s)}.$$

We call such curves special curves in DM.

Equivalent curves $\gamma \sim \mu$ (i.e., curves which represent the same path Γ) have equivalent lifts, $\underline{\dot{\gamma}} \sim \underline{\dot{\mu}}$. Therefore, a path $\Gamma = \underset{\sim}{\gamma}$ in M generates a unique path $\underline{\dot{\gamma}}$ in DM which

we denote by $\dot{\Gamma}$ (the *lift* of Γ to DM). As before a path in DM is said to be *special*, if it is the lift of a path in M.

DM is a C^{k-1} manifold, so we can construct the bundles TDM and $DDM = D^2M$ with the respective projections $\hat{\Pi}$ and Π_D. Besides Π_D there is another natural map from D^2M to DM which is analogous to π_{T*}: $T^2M \to TM$. It is obtained as follows. Define

$$R_1 : TM \to DM: X \to \underline{X} = \xi,$$

$$R_2 : TDM \to D^2M: \xi \to \underline{\xi} = \Xi,$$

and consider the diagram

$$\begin{array}{ccccc} TDM & \xrightarrow{\pi_{D*}} & TM & \xrightarrow{R_1} & DM \\ {\scriptstyle R_2}\downarrow & & & \nearrow_{\underline{\pi}} & \\ D^2M & & & & \end{array}$$

Since any two elements ξ_1, $\xi_2 \in R_2^{-1}\,\Xi$, $\Xi \in D^2M$, have the same image under $R_1 \circ \pi_{D*}$, there exists a map $\underline{\pi} = R_1 \circ \pi_{D*} \circ R_2^{-1}$ as indicated. In close analogy to the definition of special vectors in T^2M we define: $\Xi \in D^2M$ is a *special* direction, if

$$\Pi_D(\Xi) = \underline{\pi}(\Xi).$$

Special paths in DM have everywhere special directions directions. Moreover, every special direction field of class C^1,

$$\Xi: DM \to D^2M: \xi \to \Xi_\xi,$$

determines unique maximal special integral paths in DM and in the original manifold M.

2. PATH STRUCTURES

Definition: A *path structure* (PS) P on M is a set of paths in M such that for every point $p \in M$ and every direction $\xi_p \in D_pM$ there exists exactly one maximal path $\Gamma \in P$ which contains p and has the direction ξ_p at p.

The definition implies that through every point $\xi \in DM$ runs exactly one lifted path $\dot{\Gamma}$ of a path $\Gamma \in P$. Moreover, due to the restriction imposed on "curves in M," $\dot{\Gamma}$ has no self intersections whence we have the following lemma.

Lemma: A path structure P on M defines a special direction field (section of D^2M),

$$\Xi: DM \to D^2M: \xi \to \Xi_\xi,$$

such that for every path $\Gamma \in P$ and every point ξ on $\dot{\Gamma}$

$$\ddot{\Gamma}_\xi = \Xi_\xi.$$

This lemma enables us to define the differentiability class of a PS.

Definition: A path structure P is of class C^s if the corresponding section Ξ is of class C^s. We always demand $s \geq 1$.

Any special C^1-direction field over DM determines special maximal integral paths in DM whose projections to M satisfy the definitions given above which proves the following lemma.

Lemma: A special direction field of class C^s, $s \geq 1$, defines a C^s-path structure on M.

According to the preceding lemmas there is a one-to-one correspondence between smooth PS's and special direction fields. For the local analytical investigation of PS's special direction fields play a role analogous to that of second order differential equations for systems of curves.

For the futher analysis of PS's we describe direction fields in local coordinates of D^2M introduced as follows. Choose any local coordinate system (x^a) for $U \subset M$ and use standard coordinates (x^a, y^b) in TU such that a vector $X \in TU$ is given by

$$X = y^a \frac{\partial}{\partial x^a}.$$

In T^2U we also define standard coordinates (x^a, y^b, u^c, w^d) by

$$\mathbf{X} = u^a \frac{\partial}{\partial x^a} + w^b \frac{\partial}{\partial y^b} \in T^2U.$$

Special vector fields over TU are characterized in these coordinates by $y^b = u^b$, so that the most general special vector field is given by

$$\mathbf{X} = y^a \frac{\partial}{\partial x^a} + f^b(x^c, y^d) \frac{\partial}{\partial y^b}$$

with smooth functions f^b.

In DU we introduce coordinates (x^a, ξ^β) $\beta = 1, \ldots, n-1$, based on (x^a, y^b), by $\xi^\beta := y^\beta / y^n$ (for $y^n \neq 0$). If $y^n = 0$ we use equation $\xi^\beta_b := y^\beta / y^b$ $(\beta = 1, \ldots, b-1, b+1, \ldots, n)$ with $y^b \neq 0$. $\pi_D^{-1}(U)$ is covered by these n coordinate neighborhoods. Unless more than one of these patches is necessary we use the nth one with $(x^a, \xi^\beta_n) = (x^a, \xi^\beta)$.

This procedure can be repeated to obtain coordinates in D^2U. In TDU there are standard coordinates $(x^a, \xi^\beta, u^c, \eta^\delta)$ such that

$$\xi = u^a \frac{\partial}{\partial x^a} + \eta^\beta \frac{\partial}{\partial \xi^\beta} \in TDU.$$

To get coordinates in D^2U the $2n-1$ quantities (u^a, η^β) can be divided by one of its nonzero members. However, if $\Xi \in D^2U$ is a special direction (and only those are of interest to us), then it turns out that there exists always one nonzero component of the (u^a). For special directions only n coordinate patches are required, defined by

$$(x^a, \xi^\beta_b, \mu^\gamma_b, v^\delta_b) = \left(x^a, \frac{y^\beta}{y^b}, \frac{u^\gamma}{y^b}, \frac{\eta^\delta}{y^b}\right), \quad b = 1, \ldots, n.$$

$b = n$ will be omitted as before.

Furthermore, a special direction field is characterized by $\xi^\beta = \mu^\beta$ whence such a field is completely described by $n-1$ functions of $2n-1$ variables

$$v^\gamma = g^\gamma(x^a, \xi^\beta).$$

Note that prescribing the functions g^γ in one coordinate patch of D^2M (the nth one in this case) determines all the other $v^\gamma_b(x^a, \xi^\beta_b)$.

3. GEODESIC PATH STRUCTURES

The most important PS's are given by the geodesics of a linear connection Γ on M. (Such connections always exist since M is paracompact.) In local coordinates the

geodesics are given as the solutions of the equations

$$\ddot{x}^a = -\Gamma^a_{bc}\dot{x}^b\dot{x}^c. \tag{1}$$

These geodesics define a path structure P_Γ and a corresponding direction field Ξ_Γ. In terms of the coordinates defined above Ξ_Γ is given by

$$\Xi_\Gamma = (x^a, \xi^\beta, \xi^\gamma, v^\delta) = \left(x^a, \frac{x^\beta}{\dot{x}^n}, \frac{\dot{x}^\gamma}{\dot{x}^n}, \frac{\ddot{x}^\delta \dot{x}^n - \ddot{x}^n \dot{x}^\delta}{(\dot{x}^n)^3}\right) \tag{2}$$

or, using (1) and putting $\xi^n := 1$,

$$v^\delta = 2\Gamma^{[n}_{(cd}\delta^{\delta]}_{e)}\xi^c\xi^d\xi^e. \tag{3}$$

Thus v^δ is a polynomial in ξ^α of maximal degree 3. Clearly a projective change of the connection or a change of its torsion does not affect the functions v^δ; Ξ_Γ is determined by the projective PS implied by Γ.

Equation (3) shows further that v^δ does not contain all monomials of degree 3 (except for $n=2$) and that the coefficients of the other terms are not independent for different indices δ. Trying to construct a direction field given by v^δ's which are polynomials, but cannot be written in the form (3), one finds that there exists a discontinuity in at least one of the other coordinate patches. Thus such a direction field is not even C^0. We illustrate this by a simple

Example: $n=2$: Consider both coordinate patches,

$$(x^a, \xi^1_2, v^1_2) = (x^a, \xi, v) \text{ and } (x^a, \xi^2_1, v^2_1) = (x^a, \bar{\xi}, \bar{v}).$$

Take

$$v = \xi^p, \quad p > 3.$$

From (2) we get $\bar{\xi} = \xi^{-1}$ and

$$\bar{v} = -\bar{\xi}^{3-p}$$

which is not continuous for $\bar{\xi} = 0$.

Similarly it can be shown for $n > 2$ that if a C^1 path structure is given by polynomials $v^\delta(\xi^\alpha)$, then these polynomials can be written in the form (3). Because of the 1–1 correspondence between polynomials (3) and geodesic PS's we can reformulate this result as our first characterization of geodesic path structures:

Theorem 1: For a given special direction field Ξ of class C^1 there exists a linear connection Γ whose geodesics generate the PS of Ξ, if and only if the coordinates v^α of Ξ are polynomials in ξ^β. These polynomials are necessarily of the form (3), and Γ is determined by Ξ up to torsion and projective changes.

4. SYMMETRIES OF PATH STRUCTURES

According to a famous theorem of Helmholtz (1868) Riemannian spaces can be characterized among the more general metric spaces as being *infinitesimally isotropic*.[5] Since this is an intuitively appealing characterization (expressing free mobility of small rigid bodies) the purpose of this section is to characterize geodesic path structures in a similar way in terms of symmetries. In this section we shall show how this can be done.

We first define (finite) symmetries of PS's. Let ϕ: $M \to M$ be a diffeomorphism of class C^k. ϕ induces a mapping of the paths $\Gamma \in P$ onto a set of paths Γ', which again form a PS, P^ϕ. Let us reformulate this operation in terms of special direction fields. The differential of ϕ,

$$\phi_*: TM \to TM,$$

induces a diffeomorphism

$$\Phi_* := \underline{\phi_*}: DM \to DM \text{ via } \underline{\phi_*}(\xi) = \underline{(\phi_* X)}$$

for $X \in \xi$. In the same way the differential

$$\underline{(\phi_*)_*}: TDM \to TDM$$

induces the mapping

$$\Phi_{**} := \underline{\underline{(\phi_*)_*}}: D^2M \to D^2M,$$

for which the following is true:

Lemma: A diffeomorphism ϕ: $M \to M$ of class C^k, $k \geq 3$, induces diffeomorphisms

$$\Phi_*: DM \to DM \text{ and } \Phi_{**}: D^2M \to D^2M$$

such that for every special direction field Ξ: $DM \to D^2M$

$$\Xi^\phi := \Phi_{**} \circ \Xi \circ \Phi_*^{-1}$$

is again a special direction field.

If Ξ corresponds to P, then Ξ^ϕ corresponds to P^ϕ.

Definition: ϕ is called a *symmetry* of the path structure P if $P^\phi = P$ or, equivalently, if the corresponding direction field Ξ is invariant under ϕ,

$$\Xi^\phi = \Xi \text{ or } \Phi_{**}\Xi_\xi = \Xi_{\Phi_* \xi}. \tag{4}$$

A *local symmetry* of P is a local diffeomorphism ϕ: $U \to V$ of M which maps the restrictions of P-paths to U into restrictions of P-paths to V.

In order to consider isotropy of path structures, we denote by P_p the subset of P whose members pass through p, and formulate the following definition.

Definition: A *P-rotation* about p is a local diffeomorphism ϕ: $U \to V$ of M with fixed point p which maps the restrictions of elements of P_p to U into restrictions of elements of P_p to V.

The set of all symmetries of a path structure is a group $G(P)$, whereas the local symmetries of P and the P-rotations (for some point p) form pseudogroups. PS's determined by projectively flat linear connections have symmetry groups (or pseudogroups) acting transitively not only on D_pM but even on the set of projective bases of D_pM.

Before establishing a converse of the last assertion we introduce a weakened, infinitesimal analog of the concept of a P-rotation, guided by the analogous Helmholtz theorem. For this purpose we observe that if p is a fixed point of a local diffeomorphism ϕ of M, then Φ_* maps D_pM onto itself projectively and Φ_{**} maps $\Pi_D^{-1}(D_pM)$ into itself, so that it is meaningful to restrict the second Eq. (4) to D_pM. Accordingly, we formulate the following definition.

Definition: An *approximate P-symmetry*[6] (APS) *at p* is the restriction to $\Pi_D^{-1}(D_pM)$ of a local diffeomorphism Φ_{**} of D^2M, induced by a local diffeomorphism ϕ of M

which leaves p unchanged and satisfies the following condition:

$$\xi \in D_pM \Rightarrow \Phi_{**}\Xi_\xi = \Xi_{\Phi_*\xi}. \tag{5}$$

An $A\rho S$ at p, say Ψ, uniquely determines a projective isomorphism $\tilde{\Psi}$: $D_pM \to D_pM$ since Ψ maps fibres into fibres. The set of all $A\rho S$'s at p is a Lie group $\tilde{G}_p(\rho)$, and the set of the associated maps $\tilde{\Psi}$ is a Lie subgroup $\tilde{G}_p(\rho)$ of the full projective group $PG(D_pM)$ of D_pM. In fact, the map sending Ψ into $\tilde{\Psi}$ is a homomorphism,

$$G_p(\rho) \to \tilde{G}_p(\rho). \tag{6}$$

The differentials ϕ_{*p} of all local diffeomorphisms which induce $A\rho S$'s at p form a subgroup $L_p(\rho)$ of the full linear group $LG(T_pM)$ of T_pM. Under the standard homomorphism $h: LG(T_pM) \to PG(D_pM)$ we have

$$L_p(\rho) \to \tilde{G}_p(\rho). \tag{7}$$

We shall denote the center of $LG(T_pM)$, which is also the kernel of h, as Z_p.

By a *dilation at p* we shall mean a local diffeomorphism ϕ of M with fixed point p such that $\phi_{*p} \in Z_p$, $\phi_{*p} \neq \mathrm{id}$.

Let ϕ be a local diffeomorphism leaving p fixed. Take local coordinates in M such that $x^a(p) = 0$. ϕ is then given by functions $\phi^a(x^b)$, and the expression for Φ_{**} involves only

$$\phi^a_b = \frac{\partial \phi^a}{\partial x^b} \quad \text{and} \quad \phi^a_{bc} = \frac{\partial^2 \phi^a}{\partial x^b \partial x^c}.$$

The direction field Ξ of a path structure is given by the functions $v^\alpha(x^b, \xi^\gamma)$. If Ψ is an $A\rho S$ at p induced by ϕ, then Ψ is completely determined by the numbers

$$\Psi^a_b = \phi^a_b(0) \quad \text{and} \quad \Psi^a_{bc} = \phi^a_{bc}(0),$$

and condition (5) is expressed (in one chart) by

$$2\xi^a \Psi^{[n}_a (\Psi^{\gamma]}_\delta v^\delta(\xi^\beta) + \Psi^{\gamma]}_{bc}\xi^b\xi^c) = (\Psi^n_b \xi^b)^3 v^\gamma\left(\frac{\Psi^\beta_b \xi^b}{\Psi^n_b \xi^b}\right), \tag{8}$$

where we have written $v^\delta(\xi^\beta)$ instead of $v^\delta(0, \xi^\beta)$. ϕ is a dilatation at p iff $\Psi^a_b = f\delta^a_b$ with $f \neq 0, 1$.

Suppose $t \to \Psi(t)$ is a one-parameter subgroup of $G_p(\rho)$. With respect to local coordinates we can represent it by smooth functions $\Psi^a_b(t)$, $\Psi^a_{bc}(t)$ satisfying $\Psi^a_b(0) = \delta^a_b$, $\Psi^a_{bc}(0) = 0$. Taking derivatives at $t = 0$ we obtain parameters $\psi^a_b = \Psi^{a\prime}_b(0)$, $\psi^a_{bc} = \Psi^{a\prime}_{bc}(0)$ describing an element of the Lie algebra of $G_p(\rho)$. Applying this to Eq. (8) we obtain the infinitesimal version of the invariance condition,

$$v^\gamma_{,\beta\beta}(\xi^\alpha)\psi^{[\beta}_b \xi^{n]}\xi^b + v^\beta(\xi^\alpha)(\psi^n_b \xi^b \delta^\gamma_\beta + \psi^{[n}_\beta \xi^{\gamma]}) = \psi^{[\gamma}_{bc}\delta^{n]}_d \xi^b \xi^c \xi^d. \tag{9}$$

We now return to the discussion of path structures. A geodesic path structure ρ_Γ admits at each point p a group $G_p(\rho_\Gamma)$ of approximate symmetries whose image $G_p(\rho_\Gamma)$ [according to (6)] is the full group $PG(D_pM)$; also $L_p(\rho_\Gamma) = LG(T_pM)$. In particular ρ_Γ admits everywhere approximate symmetries induced by dilatations. We shall now prove two theorems showing that the existence of some approximate symmetrices suffices, in turn, to characterize path structures as geodesic ones.

Theorem 2: A C^1 path structure ρ is geodesic if and only if it admits at each point an approximate symmetry which is induced by a dilatation.

Proof: The necessity has been pointed out already. So, assume that ϕ is a dilatation at p which induces an approximate ρ-symmetry. We then have, in local coordinates, Eq. (8) with $\Psi^a_b = f\delta^a_b$, $f \neq 0, 1$. Consequently v^γ is a polynomial in ξ^β. Since this holds at any point p, it follows from Theorem 1 that ρ is geodesic.

This theorem can be understood intuitively: For $f > 0$, a dilatation "stretches" or "compresses" M radially away from or towards p. A path mapped into itself under this operation must be "infinitesimally straight." A similar idea applies if $f < 0$. The characterization of geodesic path structures given in Theorem 2 is closely related to Weyl's elementary method to introduce a linear connection via locally geodesic coordinates.[7]

Theorem 3: A C^ω path structure ρ is geodesic if and only if it admits, at each point $p \in M$, a group $G_p(\rho)$ of approximate symmetries which induces a transitive action in the set of projective bases of D_pM.

Proof: Again, the necessity has been established already. Let, then, $G_p(\rho)$ induce a transitive action in the set of bases of D_pM. Then $\tilde{G}_p(\rho) = PG(D_pM)$. Relation (7) then shows that the corresponding group $L_p(\rho)$ satisfies $L_p(\rho)/L_p(\rho) \cap Z_p = PG(D_pM)$ whence $\dim L_p(\rho) \geq n^2 - 1$. Introducing again local coordinates we infer that the functions v^γ describing ρ satisfy a system of equations (9) not only for one system of parameters (ψ^a_b, ψ^a_{bc}), but for a whole family of such systems containing $n^2 - 1$ linearly independent matrices ψ^a_b [which represent elements of the Lie algebra of $L_p(\rho)$]. Since ρ is analytic the functions $v^\gamma(\xi^\beta)$ can be represented by power series. If these are inserted into (9) there results an infinite system of linear homogeneous equations in the unknowns ψ^a_b, the coefficients of which contain the expansion coefficients of v^γ of degrees 4, 5, etc. (The lower degree terms appear in equations involving also the ψ^a_{bc}; these equations need not be considered.) The fact that this system admits $n^2 - 1$ linearly independent solutions (ψ^a_b) implies that all the expansion coefficients of v^γ of degree larger than 3 vanish, hence v^γ is a polynomial, so that Theorem 1 gives the desired result. We give the laborious proof of the last part of this argument only for the case $n = 2$. In this case $v = \sum_0^\infty \alpha_i \xi^i$ and (9) leads to

$$\alpha_{l+1}(l+1)\psi^2_1 + \alpha_l[(2-l)\psi^1_1 + (l-1)\psi^2_2] + \alpha_{l-1}(4-l)\psi^1_2 = 0.$$

for $l \geq 4$. The coefficient matrix A of this linear homogeneous system for $\psi = (\psi^2_1, \psi^1_1, \psi^2_2, \psi^1_2)$ reads

$$A = \begin{bmatrix} 5\alpha_5 & -2\alpha_4 & 3\alpha_4 & 0 \\ 6\alpha_6 & -3\alpha_5 & 4\alpha_5 & -\alpha_4 \\ 7\alpha_7 & -4\alpha_6 & 5\alpha_6 & -2\alpha_5 \\ 8\alpha_8 & -5\alpha_7 & 6\alpha_7 & -3\alpha_6 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Since the space of solutions is at least three-dimensional, every two-dimensional subdeterminant of A vanishes. This implies $A = 0$, i.e., v^γ is a polynomial. For $n > 2$ the argument is similar.

If $n = 4$ and M is interpreted as spacetime, we obtain

the physical interpretation of Theorem 3 given in the Introduction.

We conjecture that Theorem 3 can be generalized to PS's of lower differentiability class, even to class C^1, by exploiting (9) or (8) more effectively.

5. NONGEODESIC PATH STRUCTURES

Although nongeodesic path structures have lower symmetry than geodesic path structures, they are not altogether physically uninteresting. To show the existence of such structures, consider the example in two dimensions given by

$$v(\xi)=\frac{1-\xi^5}{1+\xi^2} \quad \text{and} \quad \bar{v}(\bar{\xi})=\frac{1-\bar{\xi}^5}{1+\bar{\xi}^2}$$

which is clearly analytic. Parametrized curves representing the paths can be obtained as solutions of the differential equations

$$\ddot{x}=\dot{y}^4/(\dot{x}^2+\dot{y}^2), \quad \ddot{y}=\dot{x}^4/(\dot{x}^2+\dot{y}^2), \quad (\dot{x},\dot{y})\neq(0,0).$$

(The corresponding spray on $T\mathbb{R}^2$ is of class C^1, but not C^2.) Another example, although only of class C^1, which is easily integrable, is

$$v(\xi)=\xi^{4/3}, \quad \bar{v}(\bar{\xi})=-\bar{\xi}^{5/3}.$$

Integral curves through $(0,0)$ are given by

$$y=\tfrac{27}{2}[(x+b)^{-2}-b^{-2}], \quad b\neq 0$$

plus the two coordinate axes.

To obtain an important physical example we slightly generalize the concept of a path structure.

Let E be an open submanifold of DM such that $\pi_D(E)=M$. An *E path structure* $\mathcal{P}_E$ with domain E is a set of paths in M such that (a) through each point $p\in M$ and each direction $\xi_p\in E_p=E\cap\pi_D^{-1}(p)$ there passes exactly one path of the set, and (b) the lift of each path of $\mathcal{P}_E$ into DM is contained in E. An E path structure will be called geodesic if there exists a connection Γ on M such that each path of $\mathcal{P}_E$ can be represented by a geodesic of Γ.

Examples of E path structures are collections of paths which are timelike with respect to a Lorentzian conformal structure of M. In this case E is the set of timelike directions in DM.

We shall now consider a "timelike", nongeodesic path structure of physical importance. Let $(\mathbb{R}^4, \eta_{ab})$ be the flat Minkowskian spacetime of special relativity, taken as time oriented, and let F_{ab} $(=-F_{ba})$ be a 2-form field on $\mathbb{R}^4$ interpreted as an electromagnetic field. Then the world lines of particles with specific charge 1 (say) are determined by

$$\dot{x}^a=F^a{}_b\dot{x}^b, \quad \eta_{ab}\dot{x}^a\dot{x}^b=-1, \quad \dot{x}^4>0, \tag{10}$$

where

$$F^a{}_b=\eta^{ac}F_{cb}.$$

Introducing direction coordinates $\xi^\alpha=\dot{x}^\alpha/\dot{x}^4$ $(\alpha=1,2,3)$ as in Sec. 1 we see that timelike directions obey

$$(\xi^1)^2+(\xi^2)^2+(\xi^2)^2<1,$$

and that the direction field corresponding to (10) is given by[8]

$$v^\alpha(\xi^\beta)=(F^\alpha{}_4-F^\alpha{}_\beta\xi^\beta-F^4{}_\beta\xi^\beta\xi^\alpha)[1-(\xi^1)^2-(\xi^2)^2-(\xi^3)^2]^{1/2}.$$

Theorem 1 implies that the spacetime paths of charged particles with a fixed specific charge form a nongeodesic, timelike path structure. This examples can easily be generalized to an arbitrary Lorentzian manifold.

[1]H. Weyl, Nachr. Ges. Wiss. Göttingen 99 (1921), *Mathematische Analyse des Raumproblems* (Springer, Berlin, 1923); J. Ehlers, F.A.E. Pirani, and A. Schild, "The Geometry of Free Fall and Light Propagation," in L. O'Raifeartaigh, Ed., *General Relativity* (Clarendon, Oxford, 1972), p. 63; M. Castagnino, J. Math. Phys. 12, 2203 (1971); F.A.E. Pirani and A. Schild, "Conformal Geometry and the Interpretation of the Weyl Tensor," in B. Hoffman, Ed., *Perspectives in Geometry and Relativity* (Indiana U.P., Bloomington, 1966), p. 291; J. Ehlers and A. Schild, Commun. Math. Phys. 32, 119 (1973); F.A.E. Pirani, "Building Spacetime from Light Rays and Free Particles," in *Symposia Mathematica Vol. XII* (Academic, London, 1973), p. 67; J. Ehlers, "Survey of General Relativity Theory," in W. Israel, Ed., *Relativity, Astrophysics and Cosmology* (Reidel, Dordrecht, 1973), p. 1; N.M.J. Woodhouse, J. Math. Phys. 14, 495 (1973).

[2]See, e.g., E.H. Kronheimer and R. Penrose, Proc. Cambridge Philos. Soc. 63, 481 (1967); B. Carter, Gen. Rel. Grav. 1, 349 (1971); S.W. Hawking and G.F.R. Ellis, *Large Scale Structure of Space Time* (Cambridge U.P., Cambridge, 1973).

[3]S. Lang, *Introduction to Differentiable Manifolds* (Interscience, New York, 1962).

[4]To avoid confusion we emphasize that the terms "curve" and "path" are not generally used in the sense described here. It is essential, however, to distinguish between these concepts, and we shall adhere consistently to the terminology proposed in the text.

[5]See, e.g., D. Laugwitz, *Differentialgeometrie* (Teubner, Stuttgart, 1960), Sec. 15.3.

[6]In contrast to an "infinitesimal" transformation which, intuitively, is a map close to the identity, an "approximate" symmetry is "finite," but preserves the figure of interest—here $\mathcal{P}_p$—only "infinitesimally."

[7]H. Weyl, *Raum, Zeit, Materie* (Springer, Berlin, 1923), 5th ed.

[8]Clearly this direction field can be continuously extended to the set of null directions on which v^α vanishes. Hence, the limiting paths are represented by null geodesics.

Uniformly valid solutions to Volterra integral equations

Rina Ling

Department of Mathematics, California State University, Los Angeles, California 90032
(Received 25 October 1976)

Volterra equations on the infinite interval with the kernel multiplied by a small parameter are studied. An approximate solution is obtained and proved to be uniformly valid for all time as the parameter tends to zero. The equation is a generalization of the model for the transport of charged particles in a random magnetic field.

1. INTRODUCTION

Integral equations of Volterra type arise in a wide variety of areas in physical and biological sciences—see for example Refs. 1–3. Usually they describe processes such as the "renewal" process, in which the unknown function at any time is expressible in terms of its values in the past.

The present work generalizes the results of Ref. 3 where an equation of the form

$$f(t;\alpha)=\phi(t)-\alpha\int_0^t K(t-\tau)f(\tau;\alpha)d\tau \tag{1.1}$$

is shown to describe a model for the transport of charged particles in a turbulent plasma, such as the cosmic rays in the interplanetary solar wind or interstellar gas. In that model, $f(t;\alpha)$ denotes the cosmic ray flux, $K(t)$ is a logarithmically growing kernel related to the two-point correlation function, and α is a small parameter—see Ref. 4. The purpose is to find, as $\alpha\to 0^+$, a solution that would be uniformly valid for all $t\in[0,\infty)$.

There are two standard methods for solving (1.1). Unfortunately, neither of them is useful for our purpose. The first is what is known as the Neumann series. For a fixed α, one expands f in a power series in α and equates coefficients of the same power.[2] It can be shown that, under mild conditions on $\phi(t)$ and $K(t)$, the series converge. However, as t becomes larger and larger, the convergence can become slower and slower. Therefore, in general, the series is not useful for studying the solution for large t.

The other method is by the use of the Laplace transform. The difficulty in the method is that the inversion process and the proof that the function so obtained is indeed a solution are nontrivial tasks–see Ref. 5. Often the inversion has to be done by numerical means.

In our problem, the exact solution depends on a parameter and the time variable t and there are two limiting processes, viz. $\alpha\to 0^+$ and $t\to\infty$. We want to find an asymptotic solution, as $\alpha\to 0^+$, that would be valid not only for every fixed t, but for all time t, i.e., to seek a function $g(t;\alpha)$, as an approximation that differs from the exact solution by an amount which tends to zero with α, uniformly in the interval $0\le t<\infty$.

First, an approximation to the solution of the integral equation is produced by a heuristic argument. Then its uniform validity for all time t is proved.

2. UNIFORMLY VALID APPROXIMATIONS

Uniformly valid solutions are found for Volterra integral equations of the form

$$f(t;\alpha)=1-\alpha\int_0^t K(t-\tau)f(\tau;\alpha)d\tau$$

where α is a small positive parameter. [For simplicity, the argument α will be dropped from the various functions from now on; for example, $f(t;\alpha)$ will be written as $f(t)$.]

In the model of cosmic ray transport (see Ref. 4), $f(t)=1-\alpha\int_0^t \ln(1+t-\tau)f(\tau)d\tau$, the kernel is positive and increasing. Consider more generally $f(t)=1-\alpha\int_0^t K(t-\tau)f(\tau)d\tau$. If the kernel is positive and increasing and if we expect the solution to be bounded, then the major contribution of the kernel to the integral would come from large values of $t-\tau$. Let $g(t)$ be the solution to the equation

$$g(t)=1-\alpha K(t)\int_0^t g(\tau)d\tau\,.$$

It would be assumed that $K\in C^2$ for $0\le t\infty$.

Theorem: If (1) $K(t)>0$, $t>0$; (2) $a/(t+b)\le K'(t)\le c/(t+d)$ where a, b, c, and d are positive constants; (3) $-a/(t+b)^2\le K''(t)<0$; and (4) $K''(t)/K'(t)$ is nondecreasing, then $f-g=o(1)$ uniformly in t as $\alpha\to 0^+$.

Proof: The proof is in the same spirit as the one given in Ref. 3.

A. Integral equation for the error

(1) Let $h(t)=f(t)-g(t)$, the error. Then h satisfies

$$h(t)=\phi(t)-\alpha\int_0^t K(t-\tau)h(\tau)d\tau\,,$$

where

$$\phi(t)=\alpha K(t)\int_0^t g(\tau)d\tau-\alpha\int_0^t K(t-\tau)g(\tau)d\tau\,.$$

(2) Since the integral equation satisfied by f has the same kernel, by the convolution theorem,[1]

$$h=\phi+\phi*f'\,.$$

B. Properties of g

(1) First note that g can be solved exactly. For if we let $G(t)=\int_0^t g(\tau)d\tau$, then

$$G'(t)+\alpha K(t)G(t)=1\,,$$

so

$$G(t)=\{1/\exp[\int_0^t \alpha K(\tau)\,d\tau]\}\int_0^t[\exp\int_0^s \alpha K(\tau)\,d\tau]\,ds\,.$$

(2) $g(t)$ has a zero: The existence comes from the fact that $G(0)=0$ and by L'Hôspital's rule $G(\infty)=0$, so there exists a t_0 such that $0<t_0<\infty$ and $G'(t_0)=0$; that is, $g(t_0)=0$.

(3) $g(t)$ has only one zero: The uniqueness comes from the equation

$$g'(t)=-\alpha K'(t)G(t)-\alpha K(t)g(t),$$

where at a zero of g, the last term is zero, the first term is negative and so $g'<0$. Since at adjacent zeroes of g, the signs of g' must be opposite, we conclude that $g(t)$ has only one zero.

C. Estimate of the zero of g

Let t_0 be the zero of $g(t)$.

(1) Lower bound: $t_0 \geq 1/\alpha K(1/\alpha)$ for α such that $K(1/\alpha)\geq 1$. Let $U(t)=\exp\int_0^t \alpha K(\tau)d\tau$, then

$$g(t)=1-\alpha K(t)\frac{1}{U(t)}\int_0^t U(t)d\tau,$$

so

$$1=\frac{\alpha K(t_0)}{U(t_0)}\int_0^{t_0} U(\tau)d\tau.$$

Since $U(t)$ is increasing, we have

$$1\leq \alpha K(t_0)t_0; \tag{2.1}$$

that is, $t_0\geq 1/\alpha K(t_0)$ for all α. Suppose $t_0\geq 1/\alpha K(1/\alpha)$ is not true for an α satisfying $K(1/\alpha)\geq 1$. Then

$$t_0<\frac{1}{\alpha K(1/\alpha)},$$

$$\alpha t_0 K(t_0)<\alpha t_0 K\left[\frac{1}{\alpha K(1/\alpha)}\right]<\frac{1}{K(1/\alpha)}K\left[\frac{1}{\alpha K(1/\alpha)}\right]$$

$$\leq 1, \quad \text{since } K(1/\alpha)\geq 1,$$

so $\alpha t_0 K(t_0)<1$, a contradiction to (2.1).

(2) Upper bound: $t_0<1/\alpha$ for α sufficiently small. It will be shown that $g(1/\alpha)<0$ for α sufficiently small. Consider again

$$g(t)=1-\frac{\alpha K(t)}{U(t)}\int_0^t U(\tau)d\tau.$$

Let t_1 be a point at which $K(t_1)\geq 1$ and write

$$\int_0^t U(\tau)d\tau=\int_0^{t_1}U(\tau)d\tau+\int_{t_1}^t U(\tau)d\tau \quad \text{for } t>t_1.$$

Evaluating the second integral by parts, we get

$$\int_{t_1}^t U(\tau)d\tau=\frac{U(t)}{\alpha K(t)}-\frac{U(t_1)}{\alpha K(t_1)}+\frac{1}{\alpha}\int_{t_1}^t U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau$$

$$\geq \frac{U(t)}{\alpha K(t)}-\frac{U(t_1)}{\alpha}+\frac{1}{\alpha}\int_{t_1}^t U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau,$$

so

$$g(t)\leq -\frac{\alpha K(t)}{U(t)}\int_0^{t_1}U(\tau)d\tau+\frac{K(t)}{U(t)}\left(U(t_1)-\int_{t_1}^t U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau\right), \tag{2.2}$$

where the first term is negative and $U(t_1)\leq \exp\int_0^{t_1}K(\tau)d\tau$, for $\alpha\leq 1$.

What remains to be shown is

$$\int_{t_1}^{1/\alpha} U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau\to\infty$$

as $\alpha\to 0^+$. We have

$$U(\tau)=\exp\left[\alpha\int_0^\tau K(S)dS\right],$$

where using integration by parts and the fact that $K'(t)\leq c/(t+d)$, we get

$$\int_0^\tau K(S)dS\geq \tau[K(\tau)-c]$$

$$=\tau[K(t)-c]+\tau[K(\tau)-K(t)],$$

where

$$K(t)-K(\tau)=\int_\tau^t K'(S)dS\leq c\,\frac{t+d}{\tau},$$

so

$$\int_0^\tau K(S)dS\geq \tau[K(t)-c]-c(t+d).$$

Therefore,

$$\int_{t_1}^t U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau$$

$$\geq \frac{K'(t)}{K^2(t)}\exp[-\alpha c(t+d)]\int_{t_1}^t \exp[\alpha\tau(K(t)-c)]d\tau$$

$$=\frac{K'(t)}{\alpha K^2(t)[K(t)-c]}\exp[-\alpha c(t+d)]$$

$$\times\{\exp[\alpha t(K(t)-c)]-\exp[\alpha t_1(K(t)-c)]\}, \tag{2.3}$$

$$\int_{t_1}^{1/\alpha} U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau$$

$$\geq \frac{a}{(1/\alpha+b)\alpha K^3(1/\alpha)}\exp(-c-\alpha cd)$$

$$\times\left\{\exp\left[K\left(\frac{1}{\alpha}\right)-c\right]-\exp\left[\alpha t_1\left(K\left(\frac{1}{\alpha}\right)-c\right)\right]\right\}$$

for α small such that $K(1/\alpha)>c$. Observing that the first term $\to\infty$ and the second term $\to 0$ as $\alpha\to 0^+$, the proof is complete.

(3) Upper bound: $t_0<ML_2/\alpha L_1$ for α sufficiently small where M is a constant, $L_1=K(1/\alpha)$, $L_2=\ln K(1/\alpha)$. Let $\bar{U}(t)=\exp\int_0^t K(\tau)d\tau$, then from (2.2)

$$\bar{U}(t_1)\geq\int_{t_1}^{t_0} U(\tau)\frac{K'(\tau)}{K^2(\tau)}d\tau,$$

so from (2.3), we obtain

$$\bar{U}(t_1)\geq\frac{K'(t_0)}{\alpha K^2(t_0)[K(t_0)-c]}\exp[-\alpha c(t_0+d)]$$

$$\times\{\exp[\alpha t_0(K(t_0)-c)]-\exp[\alpha t_1(K(t_0)-c)]\}.$$

For α sufficiently small,

$$\bar{U}(t_1)\geq\frac{K'(t_0)}{\alpha K^3(t_0)}\exp[-\alpha c(t_0+d)+\alpha t_0(K(t_0)-c)]$$

$$\times\{1-\exp[-\alpha t_0(K(t_0)-c)+\alpha t_1(K(t_0)-c)]\},$$

$$3\ln K(t_0)\geq \alpha t_0 K(t_0)-2\alpha c t_0-\alpha cd-\ln\left[\frac{\alpha\bar{U}(t_1)}{K'(t_0)}\right]$$

$$+\ln\{1-\exp[-\alpha t_0(K(t_0)-c)+\alpha t_1(K(t_0)-c)]\}. \tag{2.4}$$

Suppose for any $k>0$ and α_k no matter how small, there exists $\alpha<\alpha_k$ such that $t_0 \geq k L_2/\alpha L_1$. We will obtain a contradiction to (2.4).

1st term:

$$\frac{\alpha t_0 K(t_0)}{3\ln K(t_0)} \geq \frac{kL_2}{3L_1} K\left(\frac{kL_2}{\alpha L_1}\right) \frac{1}{\ln K(1/\alpha)},$$

for some $\alpha<\alpha_k$, α_k sufficiently small

$$= \frac{k}{3} \frac{K(kL_2/\alpha L_1)}{L_1},$$

Since $K(t)-K(\tau) \leq c\ln((t+d)/\tau)$ for $\tau<t$, we have

$$K\left(\frac{kL_2}{\alpha L_1}\right) \geq K\left(\frac{k}{\alpha}\right) - c\ln\left[\frac{k/\alpha+d}{kL_2/\alpha L_1}\right],$$

so

$$\frac{K(kL_2/\alpha L_1)}{L_1} \geq \frac{K(k/\alpha)}{L_1} - \epsilon \geq 1-\epsilon \quad \text{if } k \geq 1, \tag{2.5}$$

where

$$\epsilon = \frac{c}{L_1}[\ln(k+\alpha d)+L_2-\ln k-\ln L_2] \leq \epsilon_1 < 1$$

if α_k is sufficiently small.

Finally,

$$\frac{\alpha t_0 K(t_0)}{3\ln K(t_0)} \geq \frac{k}{3}(1-\epsilon_1) \quad \text{for } k \geq 1, \text{ some } \alpha<\alpha_k,$$

α_k arbitrarily small.

2nd term:

$$\frac{-2c\alpha t_0}{3\ln K(t_0)} > \frac{-2c}{3\ln k(t_0)} \quad \text{since } \alpha t_0 < 1 \text{ for } \alpha_k$$

sufficiently small,

$$> \frac{-2c}{3\ln K(1/\alpha L_1)} \quad \text{by estimate (1)}.$$

But $1/\alpha L_1 \to \infty$ as $\alpha \to 0^+$, therefore given any $\epsilon_2>0$, we have that

$$\frac{-2c\alpha t_0}{3\ln K(t_0)} > -\epsilon_2 \quad \text{for } \alpha_k \text{ sufficiently small.}$$

3rd term:

Clearly

$$\frac{-\alpha cd}{3\ln K(t_0)} > -\epsilon_3 \quad \text{for } \alpha_k \text{ sufficiently small.}$$

4th term:

$$\frac{-\ln[\alpha\bar{U}(t_1)/K'(t_0)]}{3\ln K(t_0)}.$$

Note that

$$\frac{\alpha\bar{U}(t_1)}{K'(t_0)} < \frac{\alpha\bar{U}(t_1)}{K'(1/\alpha)} \quad \text{for } \alpha \text{ small},$$

where

$$K'\left(\frac{1}{\alpha}\right) > \frac{a}{1/\alpha+b} = \frac{\alpha a}{1+\alpha b},$$

so

$$\frac{\alpha\bar{U}(t_1)}{K'(t_0)} < \frac{(1+\alpha b)\bar{U}(t_1)}{a} < \frac{(1+b)\bar{U}(t_1)}{a}.$$

Therefore, given any $\epsilon_4>0$,

$$-\ln[\alpha\bar{U}(t_1)/K'(t_0)]/3\ln K(t_0) > -\epsilon_4 \quad \text{for } \alpha \text{ small}.$$

5th term:

$$\frac{\ln\{1-\exp[-\alpha t_0(K(t_0)-c)+\alpha t_1(K(t_0)-c)]\}}{3\ln K(t_0)}.$$

By estimate (2),

$$\exp[-\alpha t_0(K(t_0)-c)] < \exp[c-\alpha t_0 K(t_0)]$$

for α small,

where

$$\alpha t_0 K(t_0) \geq \frac{kL_2}{L_1} K\left(\frac{kL_2}{\alpha L_1}\right)$$

$$\geq kL_2(1-\epsilon_1) \text{ by (2.5)},$$

$$\to \infty \text{ as } \alpha \to 0.$$

Therefore, the numerator of the 5th term is bounded away from $-\infty$, and given any $\epsilon_5>0$,

$$\frac{\ln\{1-\exp[-\alpha t_0(K(t_0)-c)+\alpha t_1(K(t_0)-c)]\}}{3\ln K(t_0)}$$

$$> -\epsilon_5 \quad \text{for } \alpha \text{ small}.$$

For k sufficiently large the five terms add up to a number greater than 1, a contradiction to (2.4).

(4) Lower bound: $t_0 > \ln(N\cdot K(t_0))/\alpha K(t_0)$ for α sufficiently small where $N=(1-\epsilon_0)/a$, $0<\epsilon_0<1$.

A lower and an upper bound will be put on $\int_0^\tau K(s)\,ds$. Lower bound:

$$\int_0^\tau K(s)ds = sK(s)\Big|_0^\tau - \int_0^\tau sK'(s)ds$$

$$\geq \tau K(\tau) - \int_0^\tau s\,\frac{c}{s+d}\,ds$$

$$\leq \tau[K(\tau)-c].$$

Upper bound:

$$\int_0^\tau K(s)ds = (s+b)K(s)\Big|_0^\tau - \int_0^\tau (s+b)K'(s)ds$$

$$\leq (\tau+b)K(\tau) - bK(0) - a\tau$$

$$\leq \tau[K(\tau)-a] + bK(\tau).$$

Therefore, we have

$$\exp\{\alpha t_0[K(t_0)-c]\} \leq U(t_0)$$

$$\leq \exp\{\alpha t_0[K(t_0)-a]+\alpha b K(t_0)\}$$

and since αt_0 and $\alpha K(t_0) < \alpha K(1/\alpha)$ both go to zero as $\alpha \to 0$, $U(t_0) \sim \exp[\alpha t_0 K(t_0)]$ as $\alpha \to 0$.

Now we are ready to derive the above lower bound (4) on t_0 for α small. As before,

$$1 = \frac{\alpha K(t_0)}{U(t_0)} \int_0^{t_0} U(\tau)d\tau$$

$$\leq \frac{\alpha K(t_0)}{U(t_0)} \int_0^{t_0} \exp\{\alpha\tau[K(\tau)-a]+\alpha b K(\tau)\}d\tau$$

$$\leq \frac{\alpha K(t_0)}{U(t_0)} \exp[\alpha b K(t_0)] \int_0^{t_0} \exp\{\alpha\tau[K(t_0)-a]\}\,d\tau$$

$$= [K(t_0)/U(t_0)](\exp\{\alpha t_0[K(t_0)-a]+\alpha b K(t_0)\} - \exp[\alpha b K(t_0)])/[K(t_0)-a]$$

$$< \frac{K(t_0)}{U(t_0)} \frac{U(t_0)+\epsilon_0-1}{K(t_0)-a} \quad \text{for } \alpha \text{ sufficiently small, } 0<\epsilon_0<1,$$

so

$$-aU(t_0) < (\epsilon_0 - 1)K(t_0),$$

$$U(t_0) > \left(\frac{1-\epsilon_0}{a}\right) K(t_0),$$

$$\ln U(t_0) > \ln[N \cdot K(t_0)], \quad \text{where } N = \frac{1-\epsilon_0}{a}.$$

On the other hand,

$$U(t_0) = \exp[\alpha \int_0^{t_0} K(\tau)d\tau] < \exp[\alpha t_0 K(t_0)],$$

so $\ln U(t_0) < \alpha t_0 K(t_0)$ for all α and

$$\alpha t_0 K(t_0) > \ln[N \cdot K(t_0)] \quad \text{for } \alpha \text{ small},$$

$$t_0 > \frac{\ln[N \cdot K(t_0)]}{\alpha K(t_0)}.$$

D. Bound on g

(1) Bound on g for g positive: Since

$$g'(t) = -\alpha K'(t)G(t) - \alpha K(t)g(t)$$

and $g(0)=1$, $g'(t)<0$ for $0 \leq t < t_0$ and so $0 \leq g(t) \leq 1$ on $0 \leq t \leq t_0$.

(2) Bound on g for g negative: If $t > t_0$, then

$$g(t) = 1 - \frac{\alpha K(t)}{U(t)} \int_0^{t_0} U(\tau)d\tau + \int_{t_0}^{t} U(\tau)d\tau,$$

where

$$\int_{t_0}^{t} U(\tau)d\tau = \frac{U(t)}{\alpha K(t)} - \frac{U(t_0)}{\alpha K(t_0)} + \int_{t_0}^{t} U(\tau) \frac{K'(\tau)}{\alpha K^2(\tau)}\,d\tau$$

and

$$\int_0^{t_0} U(\tau)d\tau = \frac{U(t_0)}{\alpha K(t_0)},$$

so

$$g(t) = \frac{-\alpha K(t)}{U(t)} \int_{t_0}^{t} U(\tau) \frac{K'(\tau)}{\alpha K^2(\tau)}\,d\tau. \tag{2.6}$$

A second integration by parts yields

$$\int_{t_0}^{t} U(\tau) \frac{K'(\tau)}{\alpha K^2(\tau)}\,d\tau$$

$$= \frac{U(t)K'(t)}{\alpha^2 K^3(t)} - \frac{U(t_0)K'(t_0)}{\alpha^2 K^3(t_0)} - \frac{1}{\alpha^2} \int_{t_0}^{t} \left[\frac{K''(\tau)}{K^3(\tau)} - \frac{3K^2(\tau)K'^2(\tau)}{K^6(\tau)}\right] U(\tau)d\tau.$$

Therefore,

$$g(t) = -\frac{K'(t)}{\alpha K^2(t)} + \frac{U(t_0)K'(t_0)K(t)}{\alpha U(t)K^3(t_0)} - \frac{\alpha K(t)}{U(t)} \int_{t_0}^{t} \left(\frac{3K'^2 - KK''}{\alpha^2 K^4}\right) U\,d\tau, \tag{2.7}$$

where

$$\int_{t_0}^{t} \left(\frac{3K'^2 - KK''}{\alpha^2 K^4}\right) U d\tau$$

$$= \int_{t_0}^{t} \left(\frac{3K'^2 - KK''}{\alpha K^2 K'} \cdot \frac{UK'}{\alpha K^2}\right) d\tau$$

$$\leq \frac{3K'^2(t_0) - K(t_0)K''(t_0)}{\alpha K^2(t_0)K'(t_0)} \int_{t_0}^{t} U(\tau) \frac{K'(\tau)}{\alpha K^2(\tau)}\,d\tau$$

since $-K''/K'$ is nonincreasing.

From (2.6) and (2.7) we see that

$$|g(t)| \leq \frac{K'(t)}{\alpha K^2(t)} + \frac{3K'^2(t_0) - K(t_0)K''(t_0)}{\alpha K^2(t_0)K'(t_0)} |g(t)| \tag{2.8}$$

$$\leq \frac{c}{\alpha(t+d)K^2(t)} + \left[\frac{3c}{\alpha(t_0+d)K^2(t_0)} + \frac{K(t_0)}{\alpha(t_0+b)K^2(t_0)}\right] |g(t)|$$

$$\leq \frac{c}{\alpha(t+d)K^2(t)} + \frac{3c+K(t_0)}{\alpha t_0 K^2(t_0)} |g(t)|,$$

since

$$\frac{3c+K(t_0)}{\alpha t_0 K^2(t_0)} = \frac{3c}{\alpha t_0 K^2(t_0)} + \frac{1}{\alpha t_0 K(t_0)}$$

$$< \frac{3c}{k(t_0)} + \frac{1}{\ln[N \cdot K(t_0)]} \quad \text{for } \alpha \text{ small}$$

$$\leq \delta_1 < 1 \quad \text{for } \alpha \text{ sufficiently small},$$

so

$$|g(t)| \leq \frac{c}{\alpha(t+d)K^2(t)} + \delta_1 |g(t)|,$$

$$|g(t)| \leq \frac{c}{(1-\delta_1)\alpha(t+d)K^2(t)} = \frac{m}{\alpha(t+d)K^2(t)} \quad \text{say}.$$

E. Bound on ϕ

Using the results in Secs. 2C and 2D, it will be shown that $\phi(t) = o(1)$ uniformly in $0 \leq t < \infty$ as $\alpha \to 0^+$.

(1) Bound on $\phi(t)$ for $0 \leq t \leq t_0$: We have

$$\phi(t) = \alpha K(t) \int_0^t g(\tau)d\tau - \alpha \int_0^t K(\tau)g(t-\tau)d\tau$$

$$= \alpha \int_0^t [K(t) - K(\tau)]g(t-\tau)d\tau.$$

Since

$$K(t) - K(\tau) \leq c \ln \frac{t+d}{\tau+d} \quad \text{for } \tau \leq t,$$

$$\phi(t) \leq \alpha c \int_0^t \ln \frac{t+d}{\tau+d}\,d\tau \quad \text{for } 0 \leq t \leq t_0$$

$$= -\alpha c \int_0^t \ln \frac{\tau+d}{t+d}\,d\tau$$

$$= -\alpha c(t+d) \int_{d/(t+d)}^{1} \ln s\,ds, \quad \text{where } s = \frac{\tau+d}{t+d}$$

$$\leq -\alpha c(t+d) \int_0^1 \ln s\,ds$$

$$= \alpha c(t+d)$$

$$\leq \alpha c(t_0+d),$$

where $\alpha t_0 < ML_2/L_1$ for α sufficiently small, so $\phi(t) = o(1)$ uniformly on $[0, t_0]$ as $\alpha \to 0^+$.

(2) Bound on $\phi(t)$ for $t_0 < t < \infty$: We have

$$\phi(t) = \phi_1(t_0) + \phi_2(t) \quad \text{for } t_0 < t < \infty,$$

where

$$\phi_1(t_0) = \alpha \int_0^{t_0} [K(t) - K(t-\tau)] g(\tau)\, d\tau,$$

$$\phi_2(t) = \alpha \int_{t_0}^{t} [K(t) - K(t-\tau)] g(\tau)\, d\tau.$$

We want to show that $\phi_2(t) = o(1)$ uniformly in t as $\alpha \to 0^+$. We have

$$\begin{aligned}
|\phi_2(t)| &\le \alpha \int_{t_0}^{t} [K(t) - K(t-\tau)] |g(\tau)|\, d\tau \\
&\le \int_{t_0}^{t} \left(c \ln \frac{t+d}{t-\tau+d} \right) \frac{m}{(\tau+d)K^2(\tau)}\, d\tau \\
&\le \frac{cm}{K^2(t_0)} \int_{t_0}^{t} \left| \ln\left(1 - \frac{\tau}{t+d}\right) \right| \frac{d\tau}{\tau+d} \\
&= \frac{cm}{K^2(t_0)} \int_{(t_0+d)/(t+d)}^{1} \left| \ln\left(1 - \frac{s(t+d)-d}{t+d}\right) \right| \frac{ds}{s}, \qquad \frac{\tau+d}{t+d} = s, \\
&= \frac{cm}{K^2(t_0)} \int_{(t_0+d)/(t+d)}^{1} \left| \ln\left(1 - s + \frac{d}{t+d}\right) \right| \frac{ds}{s} \\
&\le \frac{cm}{K^2(t_0)} \int_0^1 \frac{|\ln(1-s)|}{s}\, ds \\
&\le cm\, \alpha^2 t_0^2 \frac{\pi^2}{6} \\
&= o(1) \text{ as } \alpha \to 0^+.
\end{aligned}$$

F. Absolute integrability of f'

Applying the Laplace transform to the integral equation and inverting to obtain f, it can be shown that f' is absolutely integrable,

$$f(t) = 1 - \alpha \int_0^t K(t-\tau) f(\tau)\, d\tau.$$

Let $\bar{f}(s) = \int_0^\infty e^{-st} f(t)dt$, $\bar{K}(s) = \int_0^\infty e^{-st} K(t)dt$, then

$$\bar{f}(s) = \frac{1}{s(1+\alpha \bar{K}(s))} = \frac{1}{D(s)}, \text{ say}.$$

Three properties of $D(s)$ can be shown:

(1) Clearly if s is a zero of $D(s)$, then so is $\bar{s}$.

(2) $D(s) \neq 0$ for $\mathrm{Re}[s]$ sufficiently large.

(3) $D(s)$ has two zeroes for α sufficiently small.

Proof of (2): Let the Laplace transform of $K_0(t) = \ln(1+t)$ be denoted by $\bar{K}_0(s)$. We have

$$\bar{K}_0(r) = \frac{e^r}{r} E(r) \quad \text{for } r \text{ real},$$

where

$$E(r) = \int_r^\infty \frac{dt}{te^t},$$

so $\bar{K}_0(r) \le 1/r^2$.

Since $K(t) \le c \ln(1 + t/d) + K(0)$,

$$|\bar{K}(s)| \le \bar{K}(x) \le cd\, \bar{K}_0(xd) + \frac{K(0)}{x}$$

for $x > 0$, $x = \mathrm{Re}[s]$.

$$\begin{aligned}
|D(s)| &= |s|\, |1 - (-\alpha \bar{K}(s))| \\
&\ge |s|\, |1 - \alpha |\bar{K}(s)||,
\end{aligned}$$

since

$$\begin{aligned}
\alpha |\bar{K}(s)| &\le \alpha \left(cd\, \bar{K}_0(xd) + \frac{K(0)}{x} \right) \\
&\le \alpha \left(\frac{cd}{x^2 d^2} + \frac{K(0)}{x} \right) \\
&= \alpha Q(x) \\
&< 1 \quad \text{if } Q(x) < \frac{1}{\alpha},
\end{aligned}$$

so $|D(s)| > 0$ for x sufficiently large.

Proof of (3): It was shown in Ref. 3 that $D_0(s) = s + \alpha s \bar{K}_0(s)$ has only two zeroes (which are conjugates of each other) and the only singularity of $D_0(s)$ is the branch point of the logarithm at $s = 0$. [$D_0(s) = s + \alpha e^s E(s) = s + \alpha e^s(-\ln s - \gamma + e_1(s))$ where γ is the Euler constant and $e_1(s)$ is analytic.] Since

$$K(0) + a \ln\left(1 + \frac{t}{b}\right) \le K(t) \le c \ln\left(1 + \frac{t}{d}\right) + K(0),$$

$$K(t) = K(0) + a \ln\left(1 + \frac{t}{b}\right) + \theta(t),$$

where

$$0 \le \theta(t) \le c \ln\left(1 + \frac{t}{d}\right) - a \ln\left(1 + \frac{t}{b}\right).$$

Let $\bar{\theta}(s)$ be the Laplace transform of $\theta(t)$, then

$$\bar{K}(s) = \frac{K(0)}{s} + ab\, \bar{K}_0(bs) + \bar{\theta}(s),$$

so

$$\begin{aligned}
D(s) &= s(1 + \alpha \bar{K}(s)) \\
&= s[1 + \alpha ab \bar{K}_0(bs)] + \alpha K(0) + \alpha s \bar{\theta}(s),
\end{aligned} \tag{2.9}$$

where

$$|\bar{\theta}(s)| \le \bar{\theta}(x) \le cd \bar{K}_0(dx) - ab \bar{K}_0(bx),$$

$$x = \mathrm{Re}[s].$$

It can be shown, by the Rouche theorem, that $D(s)$ has only two zeroes. Consider the positive-oriented circular contour with a slit along the negative real axis, say C, $C = \cup_{i=1}^5 C_i$ where

$$C_1 = \{s = \mathrm{Re}^{i\beta}: \ 0 \le \beta \le [\pi - \arctan(\epsilon/(R^2 - \epsilon^2)^{1/2}]\},$$

$$C_2 = \{s = x + i\epsilon: \ -(R^2 - \epsilon^2)^{1/2} \le x \le 0\},$$

$$C_3 = \{s = \epsilon e^{i\beta}: \ -\pi/2 \le \beta \le \pi/2\},$$

$$C_4 = \{s = x - i\epsilon: \ -(R^2 - \epsilon^2)^{1/2} \le x \le 0\},$$

$$C_5 = \{s = Re^{i\beta}: \ -\pi + \arctan(\epsilon/(R^2 - \epsilon^2)^{1/2}) \le \beta \le 0\},$$

where R is sufficiently large and ϵ small so that for $\alpha = \alpha_0$ say, the zeroes of $D(s)$ lie inside the contour C. If $\alpha < \alpha_0$, the zeroes of $D(s)$ lie inside the circle of radius

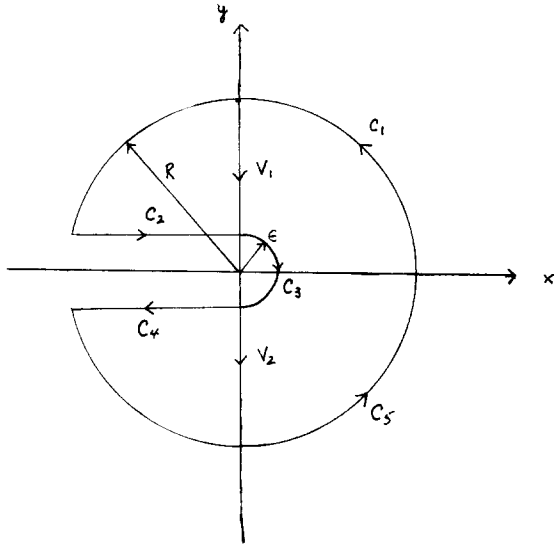


FIG. 1.

R and center at the origin, so let R be fixed and let ϵ be small so that the zeroes are inside C. See Fig. 1. On C,

$$\begin{aligned}|D_0 - D| &= \alpha|s|\,|\bar{K}_0 - \bar{K}| \\ &\le \alpha|s|(|\bar{K}_0| + |\bar{K}|) \\ &\le \alpha|s|(P_0 + P), \quad \text{where } P_0 \text{ and } P \\ &\qquad\qquad \text{are constants,} \\ |D| &= |s|\,|1 + \alpha\bar{K}| \\ &\ge |s|\,|1 - \alpha|\bar{K}||\\ &\ge |s|(1 - \alpha P) \quad \text{for } \alpha \text{ sufficiently small.}\end{aligned}$$

Therefore,

$$|D_0 - D| \le |D| \quad \text{on } C \text{ for } \alpha \text{ small}$$

and by Rouche's theorem, $D_0(s)$ and $D(s)$ have the same number of zeroes inside C.

The two zeroes of $D_0(s)$ are in the left-half-plane; if now the Rouche theorem is applied to the contour consisting of C_1, V_1, C_3, V_2, and C_4, it is clear that the two zeroes of $D(s)$ are also in the left-half-plane.

Utilizing the poles and branch cut of $[D(s)]^{-1}$, we find the inverse

$$f(t) = \frac{1}{2\pi i} \lim_{R\to\infty} \int_{r-iR}^{r+iR} \frac{e^{st}}{D(s)}\,ds,$$

where $r>0$ such that $[D(s)]^{-1}$ has no singularities for $\mathrm{Re}[s] \ge r$.

Consider the following positively oriented paths:

$$\Gamma_0 = \{s = r + iy: \ -R \le y \le R\},$$

$$\begin{aligned}\Gamma_1 = \{s = Re^{i\beta}: \ &\arctan((R^2 - r^2)^{1/2}/r) \le \beta \\ &\le \pi - \arctan(\delta/(R^2 - \delta^2)^{1/2})\},\end{aligned}$$

$$\begin{aligned}\Gamma_2 = \{s = x + i\delta: \ &-(R^2 - \delta^2)^{1/2} \le x \\ &\le (\epsilon^2 - \delta^2)^{1/2}\},\end{aligned}$$

$$\begin{aligned}\Gamma_3 = \{s = \epsilon e^{i\beta}: \ &-\pi + \arctan(\delta/(\epsilon^2 - \delta^2)^{1/2}) \le \beta \\ &\le \pi - \arctan(\delta/(\epsilon^2 - \delta^2)^{1/2})\},\end{aligned}$$

$$\Gamma_4 = \{s = x - i\delta: \ -(R^2 - \delta^2)^{1/2} \le x \le (\epsilon^2 - \delta^2)^{1/2}\},$$

$$\begin{aligned}\Gamma_5 = \{s = Re^{i\beta}: \ &-\pi + \arctan(\delta/(R^2 - \delta^2)^{1/2}) \le \beta \\ &\le -\arctan((R^2 - r^2)^{1/2}/r),\end{aligned}$$

where $0 < \delta < \epsilon < r$, ϵ small, and R sufficiently large such that the two poles of $[D(s)]^{-1}$ are inside the closed contour. See Fig. 2.

By Cauchy's residue theorem,

$$\frac{1}{2\pi i}\int_{\Gamma_0} \frac{e^{st}}{D(s)}\,ds = \sum \mathrm{Res}[e^{st}/D] - \frac{1}{2\pi i}\sum_{j=1}^{5}\int_{\Gamma_j} \frac{e^{st}}{D(s)}\,ds.$$

Let s_0 be a pole of $[D(s)]^{-1}$, then

$$\mathrm{Res}[e^{st}/D]\big|_{s=s_0} = b_1 e^{s_0 t},$$

where

$$b_1 = \lim_{s\to s_0} (s - s_0)/D(s),$$

$$\mathrm{Res}[e^{st}/D]\big|_{s=s_0^*} = b_1^* e^{s_0 t}.$$

For contours Γ_1 and Γ_5: Since $|D_0(s)| \sim R$ as $R \to \infty$, from (2.9),

$$\frac{|e^{st}|}{|D(s)|} = \frac{e^{xt}}{|D(s)|} \to 0$$

as $R \to \infty$, $x = \mathrm{Re}[s]$, so the contribution of the contours Γ_1 and Γ_5 vanish as $R \to \infty$.

For contour Γ_3: Since $|D_0(s)| \sim \alpha \ln\epsilon^{-1}$, as in above, we see that the contribution of the contour Γ_3 vanishes as $\epsilon \to 0^+$.

For contours Γ_2 and Γ_4:

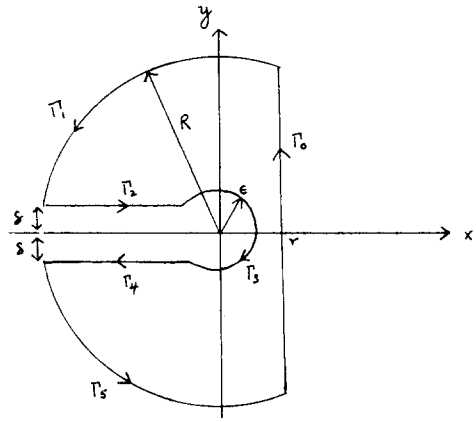


FIG. 2.

$$\lim_{\delta\to 0^+}\int_{\Gamma_2}\frac{e^{st}}{D(s)}\,ds = \int_{-R}^{-\epsilon}\frac{e^{xt}\,dx}{x+\alpha ab e^{x}[-\ln x-\gamma+e_1(x)]+\alpha K(0)+\alpha x\bar{\theta}(x)}$$
$$= -\int_{\epsilon}^{R}\frac{e^{-\xi t}d\xi}{\xi+\alpha ab e^{-\xi}[\ln\xi+i\pi+\gamma-e_1(-\xi)]-\alpha K(0)+\alpha\xi\bar{\theta}(-\xi)},$$

the result for Γ_4 is minus the complex conjugate. Let

$$e_2(\xi)=-e_1(-\xi)=\int_0^{\xi}\frac{e^{z}-1}{z}\,dz$$

and let $R\to\infty$, $\epsilon\to 0^+$; then

$$f(t)=2\mathrm{Re}[b_1 e^{s_0 t}]+I(t),$$

where

$$I(t)=-\alpha\int_0^{\infty}e^{-\xi t}B(\xi)d\xi,$$

where

$$B(\xi)=\frac{abe^{-\xi}}{\{\xi+\alpha ab e^{-\xi}[\ln\xi+\gamma+e_2(\xi)]-\alpha K(0)+\alpha\xi\bar{\theta}(-\xi)\}^2+(\alpha ab\pi e^{-\xi})^2}.$$

Since

$$B(\xi)\sim(\alpha ab\ln\xi)^{-2}\ \text{ as }\ \xi\to 0^+,\quad B(\xi)\sim[\xi+\alpha\xi\bar{\theta}(-\xi)]^{-2}e^{-\xi}\ \text{ as }\ \xi\to\infty,$$

the integral $I(t)$ converges uniformly and

$$f'(t)=2\mathrm{Re}[s_0 b_1 e^{s_0 t}]+\alpha\int_0^{\infty}\xi e^{-\xi t}B(\xi)d\xi,\quad |f'(t)|\le 2|s_0 b_1|\,e^{\mathrm{Re}[s_0]t}+\alpha\int_0^{\infty}\xi e^{-\xi t}B(\xi)d\xi.$$

Since $\mathrm{Re}[s_0]<0$,

$$\int_0^{\infty}|f'(t)|dt\le 2|s_0 b_1|\,|\mathrm{Re}[s_0]|^{-1}+\alpha\int_0^{\infty}B(\xi)d\xi,$$

but

$$1=f(0)=2\,\mathrm{Re}[b_1]-\alpha\int_0^{\infty}B(\xi)d\xi,$$

so

$$\int_0^{\infty}|f'(t)|dt\le 2|s_0 b_1|\,|\mathrm{Re}[s_0]|^{-1}+2\mathrm{Re}[b_1]-1.$$

G. Uniform validity

$$f(t)-g(t)=\phi(t)+\int_0^{t}\phi(\tau)f'(t-\tau)d\tau,$$

using the results in Secs. 2E and 2F,

$$|f-g|\le o(1)(1+\int_0^{\infty}|f'(t)|dt)=o(1)\ \text{ as }\ \alpha\to 0^+.$$

[1] R. Bellman and K. L. Cooke, *Differential-Difference Equations* (Academic, New York, 1963).
[2] F. G. Tricomi, *Integral Equations* (Interscience, New York, 1957).
[3] F. B. Hanson, A. Klimas, G. V. Ramanathan, and G. Sandri, J. Math. Phys. **14**, 1592–1600 (1973).
[4] F. B. Hanson, A. Klimas, G. V. Ramanathan, and G. Sandri, J. Math. Anal. Appl. **44**, 786–98 (1973).
[5] R. Bellman, R. Kalaba, and J. Lockett, *Numerical Inversion of the Laplace Transform* (Elsevier, New York, 1966).

On the combined Dirac–Einstein–Maxwell field equations

M. J. Hamilton[a] and A. Das[b]

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada
(Received 14 February 1977)

This paper discusses the combined Dirac–Einstein–Maxwell equations in general relativity. The combined equations are derived from a variational principle which involves the variation of tetrad fields. A class of exact, self-consistent solutions is found where the metric is static, the electromagnetic field is just electrostatic, and the spinor field is stationary in the wave mechanical sense. These solutions are analogous to Dirac's plane wave solutions which propagate along the x^3 axis and are not square integrable. It is shown that under reasonable physical conditions there do not exist solutions with finite total charge. It seems that the static electro-gravitational background is not compatible with localizable matter fields possessing intrinsic spin.

1. INTRODUCTION

In recent years[1] a lot of interest has been focused on spinor fields in curved space—time. In our opinion it is a promising sign that people are investigating the possibilities of gravitational effects on the structure of elementary particles. These pursuits might bring Einstein's gravitational theory closer to laboratory experiments.

Some years ago[2] it was realized that a matter field which generates both an attractive and a repulsive field may very well allow stable self-consistent solutions due to the possible equilibrium between two types of forces. Indeed a class of exact, self-consistent solutions of the combined Dirac—Einstein—Maxwell equations were found where the spinor affinity was assumed to have an extremely simple form[3] not allowing the general spin transformations. In special relativity a similar problem was taken up[4] for the Dirac field in the presence of electromagnetism and a Stuckelberg-type cohesive force. Again, some exact, self-consistent solutions were found. Similarly, the combined Dirac—Maxwell equations were investigated[5] in the curved geometry of Rastall's gravity. Analogous to the spin-$\frac{1}{2}$ cases, several papers[6,7] came out with exact, self-consistent solutions of the combined Klein—Gordon—Einstein—Maxwell field equations.

In this paper we investigate the combined Dirac—Einstein—Maxwell field equations using the spinor calculus in its full generality. Even the writing of the combined equations, especially the gravitational equations, is not completely obvious. The most reasonable approach is the variational derivation. In this case, however, the variation of the metric is ineffective and we have to consider the variation of the tetrad field. Fortunately, the elegant paper of Rosenfeld[8] developed the variational machinery for the combined spinor and gravitational fields. These techniques were explicitly adopted by Bergmann and Thomson[9] for the combined Dirac—Einstein equations. The complicated expression for the material energy—momentum—stress tensor of their paper is shown to reduce to the symmetrized canonical tensor in the Appendix of this paper. Brill and Wheeler[10] arrived at the same result in the the four-component formalism, though in a less rigorous way.

Following the tradition of previous papers,[2–7] we investigate here the solutions such that the metric is static, the electromagnetic field is just electrostatic, and the Dirac field is stationary in the wave-mechanical sense. From the Dirac field only "spin-up" electron-type waves are chosen. However, unlike previous investigations the solutions are not at all plentiful. Only one solution is found which corresponds in some sense to an electron plane wave propagating along the x^3 direction. Furthermore, under the assumptions given, the nonexistence of "square-integrable" solutions is concluded. Physically speaking, the "mystic" spin seems to have an impact on the symmetry of the space—time geometry and the static electro-gravitational setting is too restrictive for the existence of physically interesting solutions. There are analogous results regarding the nonexistence of neutrino solutions to the combined Weyl—Einstein equations.[11–13] In particular, Madore[11] has shown that there exist no static axially symmetric neutrino fields in general relativity, while Wainwright[12] has shown that there exist no nontrivial neutrino fields of energyclass E_2 with static metric.

2. PRELIMINARIES: SPINOR CALCULUS

The purpose of this work is to investigate the combined Dirac—Einstein—Maxwell equations. A natural formalism for these equations is the two-component spinor calculus, which has been extensively described.[14] Our definitions and notations: Let M denote the semi-Riemannian space—time manifold with signature -2. Spinor indices will be denoted by capital Latin letters, with conjugate indices dashed, and take on the values 1, 2. Small Latin letters denote space—time tensor indices ranging from 1—4, and unless otherwise noted Greek letters denote spatial tensor indices ranging from 1—3. Invariant tensor or spinor indices are denoted by parentheses: Thus $\chi^{(AB)}$ is a spinor invariant, while $\eta^{(mn)}$ is a tensor invariant. γ_{AB} denotes the antisymmetric spinor metric, and $\sigma^{mA'B}$ the spin matrices. Spin indices are raised and lowered using γ:

$$\chi^A=\gamma^{AB}\chi_B,\quad \chi_A=\chi^B\gamma_{BA},\quad \gamma_{AB}=\gamma^{1/2}\exp(i\theta)\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}.$$

[a]This paper is based on the Ph.D. thesis of this author submitted to the Mathematics Department, Simon Fraser University.

[b]Supported by NRC Grant No. -3993.

Further,

$$\gamma_A{}^B = -\gamma^B{}_A = \delta^B{}_A,$$

the Kronecker delta.

Let the solution of the combined equations be valid on the manifold M, and suppose $\tilde{M}$ is obtained from M by excising the singularities. Then the solution is valid globally over $\tilde{M}$. Geroch[15] has shown that for a C^∞ noncompact space–time manifold N the global existence of spinor fields on N is equivalent to the existence of a global field of orthonormal tetrads. If $\lambda^i{}_{(n)}$ denotes such a tetrad field, the general relativistic spin matrices may be defined by

$$\sigma^{iB'C} = \lambda^i{}_{(n)}\sigma^{(n)B'C}, \tag{2.1}$$

where $\sigma^{(n)B'C}$ are the usual special-relativistic Pauli matrices up to a real conformal factor. An equivalent approach is to require that $\sigma^{iA'B}$, g_{ij}, γ^{AB} satisfy the following axioms[16]:

Axiom I: $\sigma^{iA'B} = \sigma^{iBA'}$;

Axiom II: $\sigma_m{}^{B'}{}_A\sigma_{nB'C} = -\frac{1}{2}g_{mn}\gamma_{AC} + \frac{1}{2}i\eta_{mnsr}\sigma^{sB'}{}_A\sigma^r{}_{B'C}$,

where η_{mnsr} denotes the Levi-Civita tensor. Axiom I requires the spin matrices to be Hermitian, and Axiom II may be looked upon as the generalized commutation relations. The most important algebraic consequences of these axioms are the following:

$$\sigma_m{}^{B'}{}_A\sigma_{nB'C} + \sigma_n{}^{B'}{}_A\sigma_{mB'C} = -g_{mn}\gamma_{AC}, \tag{2.2}$$

$$\sigma_m{}^{B'}{}_A\sigma_{nB'C} - \sigma_n{}^{B'}{}_A\sigma_{mB'C} = i\eta_{mnsr}\sigma^{sB'}{}_A\sigma^r{}_{B'C}, \tag{2.3}$$

$$\sigma_m{}^{B'A}\sigma_{nB'A} = g_{mn}, \tag{2.4}$$

$$\sigma^m{}_{B'A}\sigma_{mC'D} = \gamma_{B'C'}\gamma_{AD}, \tag{2.5}$$

$$\sigma^{mB'}{}_A\sigma^n{}_{B'C}\sigma^r{}_{D'}{}^C = \tfrac{1}{2}(g^{nr}\sigma^m{}_{D'A} - g^{mr}\sigma^n{}_{D'A} + g^{mn}\sigma^r{}_{D'A} - i\eta^{mnrp}\sigma_{pD'A}). \tag{2.6}$$

Covariant differentiation of spinors is introduced through spinor affinities analogous to the usual Christoffel symbols. For a covariant 1-spinor ξ_A,

$$\xi_{A|m} = \xi_{A,m} - \Gamma^B{}_{Am}\xi_B. \tag{2.7}$$

The requirement that the usual correspondence between tensors and spinors be preserved under covariant differentiations motivates

Axiom III: $\sigma^{mA'B}{}_{|n} = 0$.

The most general expression for the spinor affinities consistent with Axiom III may be shown to be

$$\Gamma^{A'}{}_{C'k} = -\tfrac{1}{2}[\sigma_{mC'B}\sigma^{mA'B}{}_{,k} + \{{m \atop nk}\}\sigma^{nA'B}\sigma_{mC'B} + \delta^{A'}{}_{C'}(2i\epsilon\phi_k + \partial_k \ln(\gamma^{1/2}))]. \tag{2.8}$$

Here $\gamma^{1/2}$ is the real conformal factor in the spin metric γ_{AB}, ϵ is a real constant, $\{{m \atop nk}\}$ the usual Christoffel symbols, and the 4-vector ϕ_k remains to be interpreted. Under a gauge-type spin transformation $\Lambda^R{}_S = \delta^R{}_S \exp(i\epsilon\sigma)$,

$$\phi'_k = \phi_k - \phi_{,k}. \tag{2.9}$$

Together with the usual philosophy of minimal electromagnetic interaction, (2.9) provides motivation for identification of ϕ_k with the electromagnetic potential. This choice is originally due to Infeld and Van der Waerden.[17] This identification has an important mathematical consequence. Using (2.8), we have

$$\gamma^{AB}{}_{|k} = i\gamma^{AB}(2\epsilon\phi_k - \theta_{,k}). \tag{2.10}$$

Therefore, in the presence of a nonvanishing electromagnetic field $\gamma^{AB}{}_{|k}$ is not identically zero (even if the metric γ^{AB} is constant).

3. THE COMBINED EQUATIONS

The combined Dirac–Einstein–Maxwell field equations may be derived from a stationary action principle based upon the Lagrangian

$$L = \sqrt{-g}\,[\tfrac{1}{2}(R - 4\pi F^{kl}F_{kl}) - 16\sqrt{2}\pi i(\sqrt{2}m\chi^{A'}\xi_{A'} + \sigma^l{}_{B'A}\chi^A\chi^{B'}{}_{|l} + \sigma^{lB'A}\xi_{B'}\xi_{A|l})] + \text{c.c.} \tag{3.1}$$

Here $F_{kl} \equiv \phi_{k|l} - \phi_{l|k}$, units have been chosen such that $c = G = \hbar = 1$, and c.c. stands for complex conjugation. For the independent fields appearing in L, we take ϕ_k, $\lambda^{(f)}{}_j$, χ^A, ξ_A, $\chi^{A'}$, $\xi_{A'}$. The choice of the tetrad components $\lambda^{(f)}{}_j$ as the fundamental geometric variables rather than the more usual metric components g_{ij} is a characteristic feature of half-integer spin fields.[8] From (2.1), (2.8), the Lagrangian (3.1) involves the tetrad in a manner which cannot be expressed in terms of g_{ij} alone. The Euler–Lagrange equations which result from the variation of (3.1) may be written

$$D^1{}_B \equiv \sigma^l{}_{A'B}\chi^{A'}{}_{|l} - (m/\sqrt{2})\,\xi_B = 0, \tag{3.2}$$

$$D^2{}_B \equiv \sigma^{lA'B}\xi_{A'|l} + (m/\sqrt{2})\,\chi^B = 0, \tag{3.3}$$

$$M^l \equiv F^{lm}{}_{|m} - 2\sqrt{2}j^l = 0, \tag{3.4}$$

$$Q_{ij} \equiv G_{ij} + 8\pi(-F_i{}^lF_{jl} + \tfrac{1}{4}g_{ij}F^{ab}F_{ab} + M_{ij}) = 0. \tag{3.5}$$

The current j^l is given by

$$j^l \equiv \epsilon\sigma^l{}_{B'A}(\chi^A\chi^{B'} + \xi^A\xi^{B'}), \quad \epsilon \equiv \sqrt{4\pi}\,e. \tag{3.6}$$

Here e, m represent the "bare" charge, mass parameters of the wave field. M_{ij} denotes the contribution of the wave fields to the energy–momentum, and is given explicitly by

$$M_{ij} = -(i/\sqrt{2})(\sigma_{iB'A}\chi^A\chi^{B'}{}_{|j} + \sigma_{jB'A}\chi^A\chi^{B'}{}_{|i} - \sigma_i{}^{B'A}\xi_A\xi_{B'|j} - \sigma_j{}^{B'A}\xi_A\xi_{B'|i}) + \text{c.c.} \tag{3.7}$$

A summary of the derivation of (3.5), (3.7) is given in the Appendix. For convenience in later work, we also write the $Q_{ij} = 0$ equations as

$$G_{ij} = -8\pi(E_{ij} + M_{ij}) = -8\pi T_{ij}. \tag{3.8}$$

One may show that the Dirac equations[17] imply $M^l{}_{|l} = 0$, and that the Dirac and Maxwell equations together give $T^{ij}{}_{|j} = 0$.

Equations (3.2)–(3.5) contain 24 unknowns: $10(g_{ij}) + 4(\chi^A) + 4(\xi_A) + 4(\phi_k) + 1(e) + 1(m)$. Because of the freedom of imposing a gauge condition on the electromagnetic field and four coordinate conditions, there are 27 equations: $10(Q_{ij}) + 4(M^k) + 4(D^1{}_B) + 4(D^2{}_B) + 4(\text{coordinate}) + 1(\text{gauge})$. The number of independent equations is reduced by the five identities: $4(Q^{ij}{}_{|j} = 0) + 1(M^k{}_{|k} = 0)$. Thus we have 22 independent equations for the 22 unknown functions and two unknown constants.

This system may be made determinate by prescribing values for e, m. In addition to the 22 independent equations, which must hold locally everywhere, there may be global restrictions on the solutions. For example, for solutions with finite total charge, one might require that this charge be e. Such a condition appears in Das and Coffman.[7] In view of the scarcity of integral conservation laws in general relativity, this condition seems a good candidate to replace the usual square integrability of wavefunctions. Such additional global requirements could affect the arbitrariness in the choice of e, m. For our static solutions, we shall see that e and m cannot be chosen independently.

4. STATIC SOLUTIONS OF THE COMBINED EQUATIONS

We wish to find exact solutions of the combined equations such that the electric and gravitational fields are static and the Dirac field is stationary in the wave-mechanical sense. For a static metric, we use the normal form

$$\Phi = -h(dx^1)^2 - k(dx^2)^2 - l(dx^3)^2 + f(dx^4)^2. \tag{4.1}$$

The metric of any static space—time may be written locally in the form (4.1), since the metric of a 3-space may always be put in normal form.[18] With this choice, Axioms I, II will be satisfied provided the spin matrices are taken to be a conformal factor times those of special relativity:

$$\sigma^{1A'B} = \frac{h^{-1/2}}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \sigma^{2A'B} = \frac{k^{-1/2}}{\sqrt{2}}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},$$
$$\sigma^{3A'B} = \frac{l^{-1/2}}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\quad \sigma^{4A'B} = \frac{f^{-1/2}}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{4.2}$$

We set $\gamma_{AB} = \exp(i\theta)\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$. By fixing the electromagnetic gauge, one may take $\theta = 0$. Corresponding to (4.1), we assume for the electromagnetic potential

$$\phi_\alpha = 0,\quad \phi_4 = \phi. \tag{4.3}$$

Thus for mathematical simplicity our model does not take account of the magnetic field.

One form in which Dirac's equations are traditionally written in special relativity is[19]

$$\left(i\frac{\partial}{\partial t} + i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} - \beta m\right)\psi = 0. \tag{4.4}$$

Here β and the components of $\boldsymbol{\alpha}$ are certain complex constant 4×4 matrices, and ψ represents the 4-component wave field. Comparing our Dirac equations with (4.4) for the case of special relativity, we find

$$\psi_1 = -(\chi^2 + i\xi_{2'}),\quad \psi_2 = \chi^1 + i\xi_{1'},$$
$$\psi_3 = -(\chi^2 - i\xi_{2'}),\quad \psi_4 = \chi^1 - i\xi_{1'}. \tag{4.5}$$

In particular, for a pure ψ_1 wave field we have

$$\chi^1 = \xi_1 = 0,\quad \chi^2 = i\xi_{2'}. \tag{4.6}$$

For our static case, this corresponds to a "spin-up" pure electron wave.

Using (2.8), (3.2), (3.3), (4.1), (4.2), the general Dirac equations for a static metric and electrostatic Maxwell field may be written

$$-h^{-1/2}\chi^2{}_{,1} - ik^{-1/2}\chi^2{}_{,2} - l^{-1/2}\chi^1{}_{,3} + f^{-1/2}\chi^1{}_{,4}$$
$$-\left[\frac{l^{-1/2}}{4}(\ln hkf)_{,3} - i\epsilon f^{-1/2}\phi\right]\chi^1 - \tfrac{1}{4}[h^{-1/2}(\ln klf)_{,1}$$
$$+ ik^{-1/2}(\ln hlf)_{,2}]\chi^2 = m\xi_{1'}, \tag{4.7}$$

$$-h^{-1/2}\chi^1{}_{,1} + ik^{-1/2}\chi^1{}_{,2} + l^{-1/2}\chi^2{}_{,3} + f^{-1/2}\chi^2{}_{,4}$$
$$-\tfrac{1}{4}[h^{-1/2}(\ln klf)_{,1} - ik^{-1/2}(\ln hlf)_{,2}]\chi^1$$
$$+\left[\frac{l^{-1/2}}{4}(\ln hkf)_{,3} + i\epsilon f^{-1/2}\phi\right]\chi^2 = m\xi_{2'}, \tag{4.8}$$

$$h^{-1/2}\xi_{2',1} + ik^{-1/2}\xi_{2',2} + l^{-1/2}\xi_{1',3} + f^{-1/2}\xi_{1',4}$$
$$+\left[\frac{l^{-1/2}}{4}(\ln hkf)_{,3} + i\epsilon f^{-1/2}\phi\right]\xi_{1'} + \tfrac{1}{4}[h^{-1/2}(\ln klf)_{,1}$$
$$+ ik^{-1/2}(\ln hlf)_{,2}]\xi_{2'} = -m\chi^1, \tag{4.9}$$

$$h^{-1/2}\xi_{1',1} - ik^{-1/2}\xi_{1',2} - l^{-1/2}\xi_{2',3} + f^{-1/2}\xi_{2',4}$$
$$+\tfrac{1}{4}[h^{-1/2}(\ln klf)_{,1} - ik^{-1/2}(\ln hlf)_{,2}]\xi_{1'}$$
$$-\left[\frac{l^{-1/2}}{4}(\ln hkf)_{,3} - i\epsilon f^{-1/2}\phi\right]\xi_{2'} = -m\chi^2. \tag{4.10}$$

Using the pure ψ_1 conditions (4.6), these equations become

$$h^{-1/2}\xi_{2',1} + ik^{-1/2}\xi_{2',2} + \tfrac{1}{4}[h^{-1/2}(\ln klf)_{,1}$$
$$+ ik^{-1/2}(\ln hlf)_{,2}]\xi_{2'} = 0, \tag{4.11}$$

$$(\ln\xi_{2'})_{,4} = i(mf^{1/2} + \epsilon\phi), \tag{4.12}$$

$$\ln\xi_{2'} = -\tfrac{1}{4}\ln hkf + X(x^1, x^2, x^4), \tag{4.13}$$

where X is an arbitrary function. We look for solutions satisfying the separability condition

$$\psi_1 = S(\underline{x})P(t),\quad t \equiv x^4. \tag{4.14}$$

Using (4.14), (4.12) gives

$$P = \exp(-iCt), \tag{4.15}$$

$$f^{1/2} = (1/m)(C - \epsilon\phi). \tag{4.16}$$

Making the scale transformation $t' = (C/m)t$ and dropping primes, we have $C = m$. Then (4.13) may be written

$$\xi_{2'} = (i/2)S\exp(-imt),\quad S = (hkf)^{-1/4}B(x^1, x^2), \tag{4.17}$$

where B is arbitrary. Thus the general separable pure ψ_1 solution of the static Dirac equations is given by (4.15), (4.16), (4.17), and the general solution of (4.11).

For a static solution, the first three Maxwell equations are identically satisfied, and if one assumes the "balance" condition

$$e = \pm m, \tag{4.18}$$

it can be shown that $M^4 = 0$ iff $Q_{44} = 0$. One may also show that the assumption $e = Km$, $|K| \neq 1$, $K \neq 0$, leads to an inconsistency between $M^4 = 0$ and $Q_{44} = 0$, so that (4.18) does not constitute an additional assumption

For the field equations, it is convenient to use the metric form

$$\Phi = -e^{-\omega}(\bar{g}_{\alpha\beta}dx^\alpha dx^\beta) + e^\omega dt^2, \tag{4.19}$$

where $\bar{g}_{\alpha\beta}$ is in normal form. With our static pure ψ_1 assumptions, the $Q_{\alpha\beta}=0$ and $Q_{44}=0$ field equations become

$$\bar{R}_{\alpha\beta}-\tfrac{1}{2}g_{\alpha\beta}(-e^{\omega}\bar{\Delta}_2\omega+\tfrac{1}{2}e^{\omega}\bar{\Delta}_1\omega+8\pi m|S|^2)=0, \quad (4.20)$$

$$R_{44}+\tfrac{1}{4}e^{2\omega}\bar{\Delta}_1\omega+4\pi m e^{\omega}|S|^2=0. \quad (4.21)$$

Here $\bar{\Delta}_2\omega\equiv\bar{g}^{\alpha\beta}\omega_{|\alpha\beta}$, $\bar{\Delta}_1\omega=\bar{g}^{\alpha\beta}\omega_{,\alpha}\omega_{\beta}$, and $\bar{R}_{\alpha\beta}$ and other barred quantities refer to the positive definite 3-space. From (4.19) it follows[20] that $R_{44}=-\frac{1}{2}e^{2\omega}\bar{\Delta}_2\omega$. Then combining (4.20), (4.21),

$$\bar{R}_{\alpha\beta}=0, \quad (4.22)$$

which implies the associated 3-space is flat. Then we write (4.19) as

$$\Phi=-e^{-\omega}dx^{\alpha}dx^{\alpha}+e^{\omega}dt^2. \quad (4.23)$$

The $Q_{44}=0$ field equation may now be written as

$$\omega_{,\alpha\alpha}-\tfrac{1}{2}\omega_{,\alpha}\omega_{,\alpha}=8\pi m e^{-\omega}|S|^2. \quad (4.24)$$

Combining (4.17), (4.23), we have

$$S=e^{\omega/4}B(x^1,x^2). \quad (4.25)$$

The remaining Dirac equation (4.11) now becomes

$$S^{-1}S_{,1}-\tfrac{1}{4}\omega_{,1}+i(S^{-1}S_{,2}-\tfrac{1}{4}\omega_{,2})=0. \quad (4.26)$$

We have not yet discussed the $Q_{\alpha 4}=0$ field equations. For our static case, these equations reduce to $M_{\alpha 4}=0$. From (3.7) these equations become

$$|S|^2h^{-1/2}(h^{-1}h_{,2}-f^{-1}f_{,2})+2ih^{-1/2}(S\bar{S}_{,1}-\bar{S}S_{,1})=0, \quad (4.27)$$

$$|S|^2h^{-1/2}(f^{-1}f_{,1}-k^{-1}k_{,1})+2ik^{-1/2}(S\bar{S}_{,2}-\bar{S}S_{,2})=0, \quad (4.28)$$

$$S\bar{S}_{,3}-\bar{S}S_{,3}=0. \quad (4.29)$$

(4.29) is identically satisfied by (4.25) and the reality of ω. Combining (4.26)–(4.28),

$$e^{\omega/2}=|P(x^3)/B(x^1,x^2)|, \quad (4.30)$$

where P is arbitrary. (4.26) is then satisfied iff B is analytic. (4.24) implies $|B|=\text{const}$, and we set $B=1$. Then $\omega=\omega(x^3)$, and our remaining $Q_{44}=0$ equation becomes

$$\omega''-\tfrac{1}{2}(\omega')^2=8\pi m e^{-\omega/2}. \quad (4.31)$$

Setting $V=e^{-\omega/2}$, the general solution is found to be

$$V=-(3/2\pi m)\wp(x^3+C_2\,|\,g_2=0,\ g_3=C_1), \quad (4.32)$$

where $\wp$ represents the Weierstrass $\wp$-function,[21] with invariants g_2 and g_3, and C_1, C_2 are arbitrary real constants. The periodic nature of the Weierstrass $\wp$-function gives rise to singularities in the geometry which are probably true singularities. For example, using the diagonalized tetrad

$$\lambda^{i}{}_{(\alpha)}=\delta^{i}{}_{(\alpha)}e^{\omega/2}, \quad \lambda^{i}{}_{(4)}=\delta^{i}{}_{(4)}=\delta^{i}{}_{4}e^{-\omega/2}, \quad (4.33)$$

one finds $R_{(4334)}=R_{4334}=e^{\omega/2}$, which diverges as one approaches a zero of the $\wp$-function. Furthermore, the total charge for such solutions, given by

$$Q=\int_{V_3}j^4n_4\sqrt{\bar{g}}\,d^3x, \quad (4.34)$$

diverges. Here n_4 represents the unit normal to the hypersurface V_3. This work may be summarized in the propositions:

Proposition 1: A class of exact, self-consistent solutions of the combined Dirac–Einstein–Maxwell field equations (3.2)–(3.5) with static electro-gravitational fields is given by

$$\Phi=-e^{-\omega}dx^{\alpha}dx^{\alpha}+e^{\omega}dt^2,$$

$$e^{-\omega/2}=-(3/2\pi m)\wp(x^3+C_2\,|\,g_2=0,\ g_3=C_1);$$

$$\chi^2=i\xi_{2'}=-(\tfrac{1}{2})\exp[(\omega/4)-imt],$$

$$\chi^1=\xi_1=0, \quad e^{\omega/2}=1\pm\sqrt{4\pi}\phi, \quad e=\pm m.$$

(As mentioned previously, the ω–ϕ relation above is the Weyl–Majumdar relation. Unlike the Majumdar solutions[20] to the Einstein–Maxwell equations, the Dirac field here completely specifies the geometry.)

Proposition 2: There does not exist a maximally extended solution of the combined Dirac–Einstein–Maxwell equations (3.2)–(3.5) with nontrivial Dirac and Maxwell fields satisfying: (i) static space–time; (ii) algebraic spinor structure given by Axioms I–III, with minimal electromagnetic interaction; (iii) purely electrostatic Maxwell field; (iv) space–time separable pure ψ_1 Dirac field; (v) finite total charge.

APPENDIX: VARIATIONAL DERIVATION OF THE EINSTEIN EQUATIONS

The equations $Q_{ij}=0$ may be derived from the Lagrangian (3.1) either by the general method of Rosenfeld,[8] or directly by varying the tetrad $\lambda^{(f)}{}_{j}$. Here we use Rosenfeld's elegant approach. Let $\hat{\mathcal{L}}=\mathcal{L}-R\sqrt{-g}$, and define

$$\sqrt{-g}\,T^{ij}\equiv-\frac{1}{16\pi}\frac{\delta\hat{\mathcal{L}}}{\delta\lambda^{(f)}{}_{i}}\lambda^{(f)j}. \quad (A1)$$

Variation of (3.1) with respect to the tetrad then leads to the field equations

$$G^{ij}=-8\pi T^{ij}, \quad (A2)$$

where we have used

$$\frac{\partial g_{kl}}{\partial\lambda^{(f)}{}_{i}}=\lambda_{(f)l}\delta^{i}{}_{k}+\lambda_{(f)k}\delta^{i}{}_{l}. \quad (A3)$$

Let Q_{α} denote the various nongravitational field variables, and $\hat{\mathcal{L}}=L\sqrt{-g}$. Let $Q_{\alpha\|i}$ represent the usual covariant derivative for tensor fields, and for spinors define $\chi^{A}{}_{\|l}=\chi^{A}{}_{|l}-i\epsilon\phi_l\chi^{A}$, where $\chi^{A}{}_{|l}$ is the spinor derivative of Sec. 2. For the Lagrangian (3.1), $\hat{\mathcal{L}}$ depends upon the matter fields only through the Q_{α} and their first covariant derivatives $Q_{\alpha\|j}$, and is invariant under both Lorentz and general coordinate transformations. Then Rosenfeld shows

$$T^{ij}=\frac{1}{16\pi}\left[\frac{\partial L}{\partial Q_{\alpha\|i}}Q_{\alpha\|}{}^{j}-Lg^{ij}-R^{jki}{}_{\|k}\right], \quad (A4)$$

where R^{jki} depends upon the transformation properties of the fields Q_{α} and is defined in Rosenfeld. One finds that (A4) may be written

$$T^{ij}=E^{ij}+M^{ij}, \quad (A5)$$

where

$$E^{ij}=-F^{ik}F^{j}{}_{k}+\tfrac{1}{4}g^{ij}F^{ab}F_{ab}, \quad (A6)$$

$$M^{ij} = \frac{1}{16\pi}\left[\frac{\partial L}{\partial Q_{\sigma|i}}\, Q_\sigma{}^{|j} - R_\sigma{}^{jki}{}_{|k}\right], \tag{A7}$$

with Q_σ denoting the Dirac fields. After considerable calculation, one finds

$$\begin{aligned}\frac{1}{16\pi} R_\sigma{}^{jki}{}_{|k} &= \frac{i}{2\sqrt{2}}\, \sigma^j{}_{C'D}(\sigma^{iB'D}\sigma^k{}_{B'A} - \sigma^{kB'D}\sigma^i{}_{B'A})(\chi^A\chi^{C'})_{|k} \\ &\quad + \frac{i}{2\sqrt{2}}\, \sigma^{jC'D}(\sigma^{iB'A}\sigma^k{}_{B'D} - \sigma^{kB'A}\sigma^i{}_{B'D})(\xi_A\xi_{C'})_{|k} + \text{c. c.}\end{aligned} \tag{A8}$$

In another formalism, this expression appears in Bergmann and Thomson.[9] However, considerable simplification is possible by using the Dirac equations. Defining the canonical energy–momentum tensor $\tilde{M}^{ij}$ as usual by

$$\tilde{M}^{ij} = \frac{1}{16\pi}\frac{\partial L_m}{\partial Q_{\sigma|i}}\, Q_\sigma{}^{|j} - g^{ij}L_m,\quad L_m \equiv L + 4\pi F^{kl}F_{kl}, \tag{A9}$$

one finds

$$R_\sigma{}^{jki}{}_{|k} = \tfrac{1}{2}(\tilde{M}^{ij} - \tilde{M}^{ji}). \tag{A10}$$

Combining (A7), (A10), we find

$$M^{ij} = \tfrac{1}{2}(\tilde{M}^{ij} + \tilde{M}^{ji}), \tag{A11}$$

the symmetrized canonical energy–momentum tensor. (A2), (A5), (A6), (A11) then give the desired field equations.

[1] B. Kuchowicz, Gen. Rel. Grav. 5, 201 (1974).
[2] A. Das, Proc. Roy. Soc. A 267, 1 (1962).
[3] H. Tetrode, Z. Phys. 50, 336 (1928).
[4] M.N. Mukherjee, Nuovo Cimento 27, 1356 (1963).
[5] S.J. O'Hanlon, Ph.D. thesis, University of British Columbia (1970).
[6] A. Das, J. Math. Phys. 4, 45 (1963); N. De, Progr. Theoret. Phys. (Kyoto) 33, 545 (1965); Nuovo Cimento 39, 986 (1965).
[7] A. Das and C.V. Coffman, J. Math. Phys. 8, 1720 (1967).
[8] L. Rosenfeld, Mem. Roy. Acad. Belg. Cl. Sci. 18, 3 (1940).
[9] P. Bergmann and R. Thomson, Phys. Rev. 89, 400 (1953).
[10] D. Brill and J.A. Wheeler, Rev. Mod. Phys. 29, 465 (1957).
[11] J. Madore, Lett. Nuovo Cimento 5, 48 (1972).
[12] J. Wainwright, Nuovo Cimento B 22, 131 (1974).
[13] D. Trim and J. Wainwright, J. Math. Phys. 12, 2494 (1971).
[14] W. Bade and H. Jehle, Rev. Mod. Phys. 25, 714 (1953); F. Pirani, in *Brandeis Summer Institute in Theoretical Physics, Vol.* 1 (Prentice-Hall, New York, 1964); R. Penrose, Ann. Phys. 10, 171 (1960).
[15] R. Geroch, J. Math. Phys. 9, 1739 (1968).
[16] E. Schmutzer, Z. Natur. 15, 355 (1960); E. Corson, *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (Blackie, London, 1953).
[17] L. Infeld and B. Van der Waerden, Sitzber. Preuss. Akad. Wiss., Physik–Math., 380 (1933).
[18] L. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1925), p. 43.
[19] L. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), p. 472.
[20] S.D. Majumdar, Phys. Rev. 72, 390 (1947).
[21] T. MacRobert, *Functions of a Complex Variable* (MacMillan, New York, 1917), p. 183; M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (U.S. Govt., Washington, D.C., 1970), pp. 629–40.

The N-quantum approximation in the Bronzan–Lee extended-source static model

Leonard M. Scarfone[a)]

Department of Physics, University of Vermont, Burlington, Vermont 05401

James A. Burns[b)]

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts
(Received 26 April 1977)

The Bronzan–Lee model of a soluble field theory with vertex function is analyzed from the viewpoint of the N-quantum approximation. Determination of the Heisenberg field expansions in the single-source one- and two-meson sectors of the model yields three scattering amplitudes, a production amplitude, and the renormalization constants in agreement with other derivations. The solutions to one algebraic and three linear integral equations, the mathematical complexity of which is reminiscent of the Tamm–Dancoff eigenvalue approach in the Lee model with two nontrivial sources, are required for this purpose. Finite results compatible with the conventional renormalization program are obtained from the field expansions and the equations of motion by prescribing that each normal-ordered product of in-fields in an expansion have the same quantum numbers as its Heisenberg field and, so that all relevant terms are included for a sector, all possible combinations of in-fields consistent with the quantum numbers of the sector must occur. In general, this recipe is substituted in place of power counting, which evidently is inappropriate in higher sectors of the Lee model. The investigation of one- and two-meson exchange interactions of two sources in the Lee model via the N-quantum approximation has been carried out, and similar work is contemplated in charged scalar theory and the classic Chew–Low model.

I. INTRODUCTION

The N-quantum approximation (NQA) was proposed by Greenberg[1] as a nonperturbative framework of calculation which presumably upon further development would provide experimental predictions from theories of strong interactions. In the application of the NQA to a specific quantum field theory, one expresses the Heisenberg field operators as finite-degree normal-ordered expansions in an irreducible set of in-fields, and then substitutes them into the field equations where renormal ordering and comparing of different normal-ordered products of in-fields yields a set of coupled equations for the desired c-number expansion coefficients. These, at first, unknown coefficients are essentially vertex functions, or scattering and production amplitudes relevant to the various processes allowed by the theory.

Since there are relatively few published papers which use or discuss the NQA, it is convenient and perhaps worthwhile to indicate briefly the types of examples already considered, the complications encountered and the successes achieved in the evolution of the method. Greenberg's original effort gives a general introduction to the NQA and uses the Ward–Hurst–Thirring[2] (WHT) model to study, in first order of the approximation, the technical problem of making the scheme compatible with the renormalization program. A difficulty with the renormalization procedure in this application was pointed out by Halprin[3] and subsequently corrected by Greenberg,[4] but only a partial rationale for adopting this particular procedure was given. A version of the NQA for the description of bound states in relativistic quantum field theory has been given by Greenberg and Genolio,[5] who treat the deuteron in pseudoscalar-meson theory. For this purpose, they substituted Heisenberg field expansions with the smallest nontrivial number of normal-ordered terms into the field equations and obtained the simplest manifestly covariant equation for the deuteron wave function. In the weak-binding, nonrelativistic limit where renormalization effects vanish, this equation reduces to a Pauli–Schrödinger equation with the quantitative structure expected of pseudoscalar-meson theory. In another application[6] the NQA yields the formal exact solution of the vector derivative coupling model[7] in a direct way, and makes explicit that the S matrix is unity.

Some additional insight into the NQA was obtained when Pagnamenta[8] applied it to the V and $V\theta$ sectors of the Lee model and the neutral scalar field in interaction with nonrecoiling nucleons. To first order of the approximation he found some difficulties with the renormalization in the former case, such as the vanishing of the $VN\theta$ vertex function in the point source limit. Upon resorting to a renormalization scheme similar to that used by Greenberg and in view of the simplicity of the Lee model, these difficulties were overcome and correct results were established. The possibility of successfully coping with dynamical bound states and resonances from the viewpoint of the NQA was demonstrated on the examples of the $V\theta$ bound state and the unstable V particle. In the case of the neutral scalar field model, the mass renormalization and expansion coefficients are exact in the first order N-quantum solution, while the wavefunction renormalization constant Z is approximate. In each higher order the coefficients are exact, and a better approximation to Z is obtained. Halprin[9] studied the problem of assuring that the equations for the c-number coefficients are solved in a manner consistent with renormalization,

[a)]Work supported in part by Grant No. GU-3774-158 of the National Science Foundation.
[b)]Work supported in part by a Resident Study Fellowship from the IBM Corporation.

and attempted to relieve the troubles encountered by Greenberg and Pagnamenta by devising a systematic method for dropping and/or retaining terms when solving these equations. He did this by a coupling constant power-counting technique which in effect demands that, in the solution of a particular N-quantum coefficient, terms are not retained which involve higher order processes. This worked quite well when applied to the V and $V\theta$ sectors of the Lee model but encountered some difficulties in the WHT model. Thus, power counting provides a guide to the manner in which the equations for N-quantum coefficients can be solved by dropping unwanted terms, but apparently it cannot be applied straightforwardly in all cases. Thus, it remains that the principal shortcomings of the NQA lie in not knowing when to terminate the field expansions and how to solve the equations for the c-number coefficients in a manner consistent with the renormalization.

The purpose of this paper is to gain further working experience with the NQA by applying it to the lowest nontrivial sectors of the Bronzan–Lee model.[10] This model, although admittedly crude, was originally set up to provide some insight into the dynamics of strong interactions. Finite results consistent with renormalization are obtained from the field expansions and the equations of motion in a straightforward way. Each term in a field expansion is required to have the same quantum numbers as its Heisenberg operator, and all possible normal ordered products of in-fields relevant to a given sector are included. In view of the selection rules of the model, these sectors are completely solved in terms of three nonsingular integral equations of the type encountered in the Tamm–Dancoff solution of the $2V$ sector of the Lee model.[11] By contrast, in the WHT model one encounters the mathematical complications of having to solve a quadratic integral equation in lowest order and nonlinear integral equations in higher orders.

Section II describes the Bronzan-Lee theory and gives a brief survey of its literature. It then goes on to set up the field equations and identify the irreducible set of in-fields for the case of stable particles and no bound states. Section III is concerned with the field expansions and S-matrix elements in the V and U sectors. Equations for the N-quantum coefficients are established in Sec. IV, and the well-known V sector is solved. This section also includes a diagrammatic representation of the equations. In Sec. V the solutions to three linear integral equations lead to a complete description of the U sector. Some conclusions and further applications of the NQA are discussed in Sec. VI.

II. EQUATIONS OF MOTION AND ASYMPTOTIC FIELDS

The dynamical structure of the Lee model has been generalized by Bronzan, who introduced a third static and spinless fermion U and coupled it to the original V particle and the neutral scalar boson θ. This modification was implemented in terms of a second Yukawa-type interaction which describes the virtual emission and absorption of a V and a θ by the U. In the Lee model the heavy N particle has no internal states, and the only virtual elementary processes allowed by the selection rules are the emission and absorption of an N and a θ by the V. Thus, there are no radiative corrections to the vertex part, and one diagram appears as the correction to the V-particle self-energy.

In contrast, the Bronzan–Lee model has a nontrivial vertex function renormalization in addition to the renormalizations already underlying the Lee model. From the conventional diagrammatic viewpoint this function embodies all proper graphs which contain one U- and one V-particle line, plus one external meson line. The vertex function, the U- and V-particle propagators, the off-energy-shell scattering amplitude for a θ incident on a V, and all the renormalization constants were established by Bronzan via the Wigner–Brillouin perturbation series approach to transition matrix elements. The addition of a θ-particle to the U sector leads to a four-body sector involving the elastic scattering of a θ by the U particle, and the production of two and three mesons in association with a V and N particle, respectively. This sector has also been analyzed by Bronzan,[12] who found that the corresponding dynamical equations reduce to the solution of a single Fredholm integral equation in one variable. An analogous situation occurs in the $VV\theta$ sector of the Lee model.[13]

Besides being a very valuable framework for the discussion of the renormalization problem and off-energy- shell techniques, the Bronzan–Lee model has importance as a general testing ground for dynamical questions and various methods of calculation in field theory. In the former category, we mention the paper of Liossatos,[14] which demonstrates, among other things, that, in the limit where the wavefunction renormalization constant of the V particle vanishes, the Bronzan–Lee model is equivalent to a composite theory containing only Fermi-type couplings among the U, N, and θ particles. More recently, Choudhury,[15] following the procedures of Maxon[16] and Scarfone,[17] has applied the Lehmann–Symanzik–Zimmermann[18] (LSZ) formalism to the two lowest nontrivial sectors of the Bronzan–Lee model and derived all the τ functions and the relevant scattering and production amplitudes. All the renormalization constants were obtained and found to agree with the results first calculated by Bronzan, and subsequently also by Liossatos. An independent but related effort by Chen-Cheung[19] in a similarly extended Lee model deals with the renormalization problem from a Green' function approach.

Nonrelativistic versions of the Bronzan–Lee model have been studied by North,[20] and by Fonseca and Shanley.[21] The former author solves the model in terms of the old-fashioned strong-coupling theory of Wentzel and suggests a procedure for obtaining a bootstrap solution, while the latter authors use the theory as a framework for gaining insight into four-body scattering systems.

For future purposes, it is convenient to treat all heavy particles as bosons. In so far as predictions of the theory are concerned the distinction between heavy fermions and heavy bosons is irrelevant providing the

interactions are spin-independent, and we only consider states involving at most one heavy particle. In accordance with the renormalized Hamiltonian which describes the particle processes previously mentioned, we write the following set of Heisenberg field equations for the Bronzan–Lee model:

$$Z_U\left(i\frac{d}{dt} - m_U + \delta m_U\right)U(t) = \lambda Z_1 \int f(\omega)A(\mathbf{k},t)V(t)d^3k, \tag{1a}$$

$$Z_V\left(i\frac{d}{dt} - m + \delta m\right)V(t) = \lambda Z_1 \int f(\omega)A^*(\mathbf{k},t)U(t)d^3k + g\int f(\omega)A(\mathbf{k},t)N(t)d^3k, \tag{1b}$$

$$\left(i\frac{d}{dt} - m\right)N(t) = g\int f(\omega)A^*(\mathbf{k},t)V(t)d^3k, \tag{1c}$$

$$\left(i\frac{d}{dt} - \omega\right)A(\mathbf{k},t) = \lambda Z_1 f(\omega)V^*(t)U(t) + gf(\omega)N^*(t)V(t). \tag{1d}$$

The operators $U(t)$, $V(t)$, and $N(t)$ are the renormalized field operators that asymptotically annihilate U, V, and N particles, respectively. Their Hermitian conjugates are the corresponding creation operators. For simplicity, both V and N are assigned the same energy, i.e., mass m, while U has the energy m_U; Z_U and Z_V are the U- and V-field wavefunction renormalization constants; δm_U and δm are the mass renormalization counterterms, while the parameters λ and g are, respectively, the $UV\theta$ and $VN\theta$ renormalized coupling constants; Z_1 is the $UV\theta$ vertex renormalization constant. The operator $A(\mathbf{k},t)$ asymptotically annihilates a meson of 3-momentum k and relativistic energy $\omega(\mathbf{k}) \equiv \omega = (\mathbf{k}^2 + \mu^2)^{1/2}$, where μ is the rest mass. The real quantity $f(\omega)$ is an abbreviation for the ratio $\rho(\omega)/(2\pi)^{3/2}(2\omega)^{1/2}$, where $\rho(\omega)$ is an ω-dependent cutoff function that guarantees convergence of integrations over ω in the high-energy limit, and prevents the existence of any ghost states.

Equal-time commutation relations are given by

$$[U(t), U^*(t)] = Z_U^{-1}, \quad [V(t), V^*(t)] = Z_V^{-1},$$
$$[N(t), N^*(t)] = 1, \quad [A(\mathbf{k},t), A^*(\mathbf{k}',t)] = \delta(\mathbf{k}-\mathbf{k}'). \tag{2}$$

All other possible commutators vanish. The Heisenberg equations of motion are based on selection rules which allow the following operators to be constants of the motion:

$$B(t) = Z_U U^*(t)U(t) + Z_V V^*(t)V(t) + N^*(t)N(t), \tag{3a}$$

$$Q(t) = Z_U U^*(t)U(t) - N^*(t)N(t) + \int A^*(\mathbf{k},t)A(\mathbf{k},t)d^3k. \tag{3b}$$

Therefore, in the Tamm–Dancoff method of solution the state vector space is decomposed into mutually nonoverlapping subspaces which are labeled by the discrete values of B and Q. Analogously, in the NQA we may associate a pair of quantum numbers with a Heisenberg field for all values of the time t. This is to be expected since the B and Q values of a state change accordingly under the action of such field operators. The operators $U(t)$, $V(t)$, $N(t)$, and $A(\mathbf{k},t)$ have (B,Q) values given, respectively, by $(-1,-1)$, $(-1,0)$, $(-1,1)$, and $(0,-1)$. Annihilation operators have quantum numbers of the opposite sign. In contrast to the Tamm–Dancoff formalism, where a physical state is expanded in terms of all bare states having the same B and Q values, here individual contributions to the expansion of a Heisenberg field in a complete set of in-fields have the same quantum numbers obtained by adding the individual B and Q values of the corresponding in-fields.

It is advantageous to express the equations of motion with respect to the Fourier-transform variables. In this way, we obtain the following coupled field equations:

$$Z_U(H - m_U + \delta m_U)U(H) = \lambda Z_1 \int\int f(\omega)A(\mathbf{k},G)V(H-G)d^3k\,dG, \tag{4a}$$

$$Z_V(E - m + \delta m)V(E) = \lambda Z_1 \int\int f(\omega)A^*(\mathbf{k},H)U(H+E)d^3k\,dH + g\int\int f(\omega)A(\mathbf{k},G)N(E-G)d^3k\,dG, \tag{4b}$$

$$(F - m)N(F) = g\int\int f(\omega)A^*(\mathbf{k},G)V(F+G)d^3k\,dG, \tag{4c}$$

$$(G - \omega)A(\mathbf{k},G) = \lambda Z_1 f(\omega)\int V^*(H)U(H+G)dH + gf(\omega)\int N^*(E)V(E+G)dE. \tag{4d}$$

As previously stated, the central idea underlying the NQA requires that field equations of a specific theory be solved by expressing each field operator as an approximate expansion in terms of an irreducible set of asymptotic operators that create or destroy stable particles. Hence, in the first step we must identify the asymptotic fields. This is done by considering the possible energy spectrum due to the interactions in the theory. If we assume U and V are stable elementary particles so that each field has a single in-field limit defined in the LSZ sense, and if there are no bound states to account for, then the four in-fields U_{in}, V_{in}, N_{in}, and $A_{in}(\mathbf{k})$ constitute an asympotically irreducible set. The commutation relations associated with these operators have the free-field form

$$[U_{in}, U_{in}^*] = 1, \quad [V_{in}, V_{in}^*] = 1$$
$$[N_{in}, N_{in}^*] = 1, \quad [A_{in}(\mathbf{k}), A_{in}^*(\mathbf{k}')] = \delta(\mathbf{k}-\mathbf{k}'), \tag{5}$$

while all other commutators vanish. The presence of bound states would require independent in-fields to complete the present set. Since the asymptotic condition makes sense only for stable objects, unstable particles must be regarded as resonance complexes which in the remote past were separate stable particles that will in the far future again become these particles. For example, an unstable U particle will show up as a resonance in the $V\theta$ scattering cross section. In any case, there would always only be four Heisenberg fields in the theory.

III. FIELD EXPANSIONS AND S-MATRIX ELEMENTS IN THE (NONTRIVIAL) ONE-HEAVY-PARTICLE SECTORS

Having defined a complete set of asymptotic fields, we now develop the expansion for each field operator in the U- and V-particle sectors. We require that each term in an expansion have the same quantum numbers as its Heisenberg operator. In addition, so that all pertinent terms are included for a sector,

the series must contain all possible combinations of in-fields consistent with the quantum numbers of the sector. This is reminiscent of the Tamm–Dancoff procedure of superimposing all bare states having the quantum numbers of a given physical state. Thus, the in-field expansions for the Fourier-transformed fields in the V and U sectors are written in the normal-ordered forms

$$U(H)=\delta(H-m_U)U_{\rm in}+\int u_{21}(\omega)\delta(H-m-\omega)V_{\rm in}A_{\rm in}(\mathbf{k})d^3k + \iint u_{22}(\omega,\omega')\delta(H-m-\omega-\omega')N_{\rm in}A_{\rm in}(\mathbf{k})A_{\rm in}(\mathbf{k}') \times d^3k d^3k', \tag{6a}$$

$$V(E)=\delta(E-m)V_{\rm in}+\int v_{11}(\omega)\delta(E-m-\omega)N_{\rm in}A_{\rm in}(\mathbf{k})d^3k + \int v_{21}(\omega)\delta(E+\omega-m_U)A^\dagger_{\rm in}(\mathbf{k})U_{\rm in}d^3k + \iint v_{22}(\omega,\omega')\delta(E+\omega-\omega'-m)A^\dagger_{\rm in}(\mathbf{k})V_{\rm in} \times A_{\rm in}(\mathbf{k}')d^3k d^3k' + \iiint v_{23}(\omega,\omega',\omega'')\delta(E+\omega-\omega'-\omega''-m) \times A^\dagger_{\rm in}(\mathbf{k})N_{\rm in}A_{\rm in}(\mathbf{k}')A_{\rm in}(\mathbf{k}'')d^3k d^3k' d^3k'', \tag{6b}$$

$$N(F)=\delta(F-m)N_{\rm in}+\int n_{11}(\omega)\delta(F+\omega-m)A^\dagger_{\rm in}(k)V_{\rm in}d^3k + \iint n_{12}(\omega,\omega')\delta(F+\omega-\omega'-m)A^\dagger_{\rm in}(k)N_{\rm in} \times A_{\rm in}(\mathbf{k}')d^3k d^3k' + \iint n_{21}(\omega,\omega')\delta(F+\omega+\omega'-m_U) \times A^\dagger_{\rm in}(\mathbf{k})A^\dagger_{\rm in}(\mathbf{k}')U_{\rm in}d^3k d^3k' + \iiint n_{22}(\omega,\omega',\omega'')\delta(F+\omega+\omega'-\omega''-m) \times A^\dagger_{\rm in}(\mathbf{k})A^\dagger_{\rm in}(\mathbf{k}')V_{\rm in}A_{\rm in}(\mathbf{k}'')d^3k d^3k' d^3k'' + \iiint\int n_{23}(\omega,\omega',\omega''\omega''')\delta(F+\omega+\omega'-\omega''-\omega'''-m) \times A^\dagger_{\rm in}(\mathbf{k})A^\dagger_{\rm in}(\mathbf{k}')N_{\rm in}A_{\rm in}(\mathbf{k}'')A_{\rm in}(\mathbf{k}''')d^3k d^3k' d^3k'' d^3k''', \tag{6c}$$

$$A(\mathbf{k},G)=\delta(G-\omega)A_{\rm in}(\mathbf{k})+\delta(G)\alpha_{11}(\omega)N^\dagger_{\rm in}V_{\rm in} + \int \alpha_{12}(\omega,\omega'')\delta(G-\omega')N^\dagger_{\rm in}N_{\rm in}A_{\rm in}(\mathbf{k}')d^3k' + \alpha_{21}(\omega)\delta(G+m-m_U) \times V^\dagger_{\rm in}U_{\rm in}+\int \alpha_{22}(\omega,\omega')\delta(G-\omega')V^\dagger_{\rm in}V_{\rm in}A_{\rm in}(\mathbf{k}')d^3k' + \int \alpha_{23}(\omega,\omega')\delta(G+\omega'+m-m_U)A^\dagger_{\rm in}(\mathbf{k}')N^\dagger_{\rm in}U_{\rm in}d^3k' + \iint \alpha_{24}(\omega,\omega',\omega'')\delta(G-\omega'-\omega'')V^\dagger_{\rm in} \times N_{\rm in}A_{\rm in}(\mathbf{k}')A_{\rm in}(\mathbf{k}'')d^3k' d^3k'' + \iint \alpha_{25}(\omega,\omega',\omega'')\delta(G+\omega'-\omega'') \times A^\dagger_{\rm in}(\mathbf{k}')N^\dagger_{\rm in}V_{\rm in}A_{\rm in}(\mathbf{k}'')d^3k' d^3k'' + \iiint \alpha_{26}(\omega,\omega',\omega'',\omega''')\delta(G+\omega'-\omega''-\omega''') \times N^\dagger_{\rm in}A^\dagger_{\rm in}(\mathbf{k}')N_{\rm in}A_{\rm in}(\mathbf{k}'')A_{\rm in}(\mathbf{k}''')d^3k' d^3k'' d^3k'''. \tag{6d}$$

In these expansions the coefficients u, v, m, and α are labeled by two subscripts ij. The first of these indicates a V or U sector association, $i=1$ for V, and $i=2$ for U. The second index j denotes the number of the coefficient within a sector. With no bound states or unstable particles present, these expansions are complete for any analysis up through the U sector. All terms have the same quantum numbers as their corresponding Heisenberg field, and each is identified with one sector or the other. Note the omission of any terms involving more than one-heavy-particle creation or annihilation operator. In a future communication this restriction will be removed by considering the NQA technique in the VN and VV sectors of the Lee model where a two-heavy-particle bound state is already known to exist in each case. As already mentioned, Greenberg and Genolio have demonstrated how the NQA could be applied to bound state problems by treating the deuteron in pseudoscalar-meson theory. Since their effort is confined to the one-meson exchange approximation, it is worthwhile, even at the static-model level, to gain experience with the NQA in a soluble two-heavy-particle system involving the exchange of two mesons. This may be the first example of an exact NQA calculation of this type.

Diagrammatic representation of the expansion coefficients provides an illuminating means of interpreting both the expansion rules stated earlier and the structure of the equations for the c-number coefficients given in Sec. IV. We show in Fig. 1 graphs for the $UV\theta$ vertex v_{21}, the $V\theta$ scattering coefficient v_{22}, and the $V\theta \to N\theta\theta$ production coefficient v_{23} with the V leg off the mass shell in each case. The expansion coefficients denote connected parts of matrix elements of the Heisenberg fields sandwiched between in-states. Reading the diagrams from bottom to top, we can use the customary "$+i\epsilon$ trick" to obtain retarded propagators. Double lines represent the off-shell Heisenberg fields while single lines at the bottom represent in-field creation operators. Single lines at the top symbolize in-field annihilation operators, and the black box itself stands for the coefficient. A delta function expressing over-all energy conservation accompanies each graph. Double lines not terminating on a box (see Fig. 3, for example) are interpreted as renormalized propagators.

In the sectors under consideration there are four amplitudes. We determine a set of N-quantum coefficients of interest for S-matrix elements describing the elastic scatterings $N\theta$, $V\theta$, and $N\theta\theta$, and the production $V\theta \to N\theta\theta$ by contracting the N or V particle from the corresponding out-state. This leads to

$$S_{N\theta\to N\theta'}=\delta(\mathbf{k}-\mathbf{k}')+\lim_{t\to\infty} n_{12}(\omega',\omega)\exp[i(\omega'-\omega)t], \tag{7}$$

$$S_{V\theta\to V\theta'}=\delta(\mathbf{k}-\mathbf{k}')+\lim_{t\to\infty} v_{22}(\omega',\omega)\exp[i(\omega'-\omega)t], \tag{8}$$

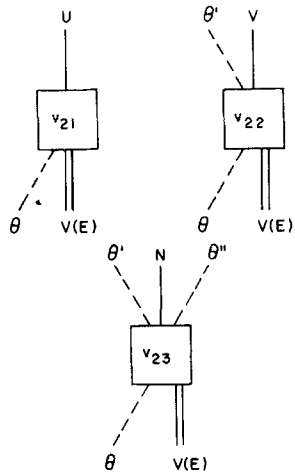


FIG. 1. U-sector coefficients appearing in the in-field expansion of the Heisenberg V field.

$$P_{V\theta\to N\theta'\theta''} = \sqrt{2}\lim_{t\to\infty} n_{22}(\omega',\omega'',\omega)\exp[i(\omega'+\omega''-\omega)t], \tag{9}$$

$$\begin{aligned} S_{N\theta\theta'\to N\theta''\theta'''} &= \tfrac{1}{2}[\delta(\mathbf{k}-\mathbf{k}'')\delta(\mathbf{k}'-\mathbf{k}''') \\ &\quad + \delta(\mathbf{k}-\mathbf{k}''')\delta(\mathbf{k}'-\mathbf{k}'')] \\ &\quad + \tfrac{1}{2}\lim_{t\to\infty}\{n_{12}(\omega'',\omega)\delta(\mathbf{k}'-\mathbf{k}''')\exp[i(\omega''-\omega)t] \\ &\quad + n_{12}(\omega'',\omega')\delta(\mathbf{k}-\mathbf{k}''')\times\exp[i(\omega''-\omega')t] \\ &\quad + n_{12}(\omega''',\omega)\delta(\mathbf{k}'-\mathbf{k}'')\exp[i(\omega'''-\omega)t] \\ &\quad + n_{12}(\omega''',\omega')\delta(\mathbf{k}-\mathbf{k}'')\exp[i(\omega'''-\omega')t]\} \\ &\quad + 2\lim_{t\to\infty} n_{23}(\omega''',\omega'',\omega',\omega) \\ &\quad \times\exp[i(\omega'''+\omega''-\omega'-\omega)t]. \end{aligned} \tag{10}$$

Before using the NQA technique to derive equations for the expansion coefficients, it should be noted that the ordering of the equal-time Heisenberg fields in the equations of motion is not unique. We choose to proceed with the ordering in Eq. (4) because the resulting equations for the coefficients are simple. This is reminiscent of dispersion theory where one's success in solving a matrix element of a current or field operator with respect to physical states depends on which particles are contracted. That there is a close correspondence between the NQA and dispersion theory is seen in the connection between such a matrix element and the expansion coefficients of one field. On the other hand, a Tamm–Dancoff expansion coefficient couples, in general, not only to one or more NQA coefficients, but to NQA coefficients of different Heisenberg fields.

By considering matrix elements of the field equations, it is possible to know beforehand which ordering will yield the simplest equations for the expansion coefficients. For example, consider writing Eq. (4a) as

$$\begin{aligned} &Z_U(H-m_U+\delta m_U)\langle 0,\ \text{in}|\,U(H)\,|\beta,\ \text{in}\rangle \\ &\quad = \lambda Z_1\sum_n\int f(\omega)\langle 0,\ \text{in}|A(\mathbf{k},G)|\,n,\ \text{in}\rangle \\ &\qquad \times\langle n,\ \text{in}|\,V(H-G)\,|\beta,\ \text{in}\rangle\, d^3k dG \end{aligned} \tag{11}$$

where the state $|\beta,\text{in}\rangle$ has the quantum numbers of the U particle, and a complete set over in-states has been inserted on the right-hand side of the equation. Because of the selection rules in the theory, it follows that the only contribution to the sum over n comes from the one-particle θ states. On the other hand, had we used the ordering $V(H-G)A(\mathbf{k},G)$ in Eq. (4a), the intermediate states would have been the physical one-particle V state and the $N\theta$ scattering states. The ordering which involves the simplest intermediate states is the desirable one. If the β state is taken to be the U-particle in-state $|U_{\text{in}}\rangle$, then the former case yields the expression

$$Z_U\delta m_U = \lambda Z_1\int f(\omega)v_{21}(\omega)d^3k, \tag{12}$$

while the latter leads to

$$\begin{aligned} Z_U\delta m_U &= \lambda Z_1\int f(\omega)\alpha_{21}(\omega)d^3k \\ &\quad \lambda Z_1\int\int f(\omega)\alpha_{23}(\omega,\omega')v_{11}(\omega')d^3kd^3k'. \end{aligned} \tag{13}$$

Compared with Eq. (12), it is obvious that Eq. (13) is a more complicated way of writing $Z_U\delta m_U$. Similar considerations of other matrix elements of the field equations show that the solutions in the sectors under discussion are found with the least mathematical resistance when the field operators are ordered as in Eqs. (4a)–(4d).

IV. EQUATIONS FOR THE N-QUANTUM COEFFICIENTS AND THEIR GRAPHICAL EQUIVALENTS: SOLUTION OF THE ONE-MESON V SECTOR

Although the procedure for implementing the NQA has been described elsewhere,[1] we briefly reiterate it here for completeness of the exposition. To solve for the unknown c-number coefficients appearing in the in-field expansions of the Fourier-transformed Heisenberg field operators, the latter expressions are substituted into the equations of motion, and products of in-field operators on the right-hand sides are renormal ordered by means of their commutation relations. Then, by virtue of the property that different products of in-field operators are independent, the resulting equations yield relations for the coefficients. Products of in-field operators associated with sectors higher than U are ignored in this process.

Using this technique in the U-field equation of motion, we come upon Eq. (12) and two other equations, namely,

$$\begin{aligned} &Z_U(\omega+m-m_U+\delta m_U)u_{21}(\omega) \\ &\quad = \lambda Z_1[f(\omega)+\int f(\omega')v_{22}(\omega',\omega)d^3k'], \end{aligned} \tag{14a}$$

$$\begin{aligned} &Z_U(\omega+\omega'+m-m_U+\delta m_U)u_{22}(\omega,\omega') \\ &\quad = \tfrac{1}{2}\lambda Z_1[f(\omega)v_{11}(\omega')+f(\omega')v_{11}(\omega)] \\ &\qquad + \lambda Z_1\int f(\omega'')v_{23}(\omega'',\omega,\omega')d^3k''. \end{aligned} \tag{14b}$$

In a similar manner, the V-field equation of motion leads to

$$Z_V\delta m = g\int f(\omega)n_{11}(\omega)d^3k, \tag{15a}$$

$$Z_V(\omega+\delta m)v_{11}(\omega) = gf(\omega)+g\int f(\omega',\omega)n_{12}(\omega',\omega)d^3k'. \tag{15b}$$

$$\begin{aligned} Z_V(m_U-m-\omega+\delta m)v_{21}(\omega) &= \lambda Z_1 f(\omega)+2g\int f(\omega')n_{21} \\ &\quad \times(\omega',\omega)d^3k'. \end{aligned} \tag{15c}$$

$$\begin{aligned} &Z_V(\omega'-\omega+\delta m)v_{22}(\omega,\omega') \\ &\quad = \lambda Z_1 f(\omega)u_{21}(\omega')+gf(\omega')n_{11}(\omega) \\ &\qquad + 2g\int f(\omega'')n_{22}(\omega'',\omega,\omega')d^3k'', \end{aligned} \tag{15d}$$

$$\begin{aligned} &Z_V(\omega'+\omega''-\omega+\delta m)v_{23}(\omega,\omega',\omega'') \\ &\quad = \lambda Z_1 f(\omega)v_{22}(\omega',\omega'')+\tfrac{1}{2}g[f(\omega')n_{12}(\omega,\omega'') \\ &\qquad + f(\omega'')n_{12}(\omega,\omega')] \\ &\qquad + 2g\int f(\omega''')n_{23}(\omega''',\omega,\omega',\omega'')d^3k'''. \end{aligned} \tag{15e}$$

The renormalization constants, and the scattering and production amplitudes in the V and U sectors can be calculated without recourse to the α coefficients in the θ-field expansion. However, to complete the goal of the NQA in the present case, we include the remaining equations for the coefficients that follow from the equations of motion for the N and θ fields. These are

$$n_{11}(\omega) = -gf(\omega)/\omega, \tag{16a}$$

$$(\omega'-\omega)n_{12}(\omega,\omega') = gf(\omega)v_{11}(\omega'), \tag{16b}$$

$$(m_U - m - \omega - \omega')n_{21}(\omega, \omega')$$
$$= \tfrac{1}{2}g[f(\omega)v_{21}(\omega') + f(\omega')v_{21}(\omega)], \tag{16c}$$
$$(\omega'' - \omega - \omega')n_{22}(\omega, \omega', \omega'')$$
$$= \tfrac{1}{2}g[f(\omega)v_{22}(\omega', \omega'') + f(\omega')v_{22}(\omega, \omega'')], \tag{16d}$$
$$(\omega''' + \omega'' - \omega' - \omega)n_{23}(\omega, \omega', \omega'', \omega''')$$
$$= \tfrac{1}{2}g[f(\omega)v_{23}(\omega', \omega'', \omega''') + f(\omega')v_{23}(\omega, \omega'', \omega''')], \tag{16e}$$
$$\alpha_{11}(\omega) = -gf(\omega)/\omega \tag{17a}$$
$$(\omega' - \omega)\alpha_{12}(\omega, \omega') = gf(\omega)v_{11}(\omega'), \tag{17b}$$
$$(m_U - m - \omega)\alpha_{21}(\omega)$$
$$= \lambda Z_1 f(\omega) + gf(\omega)\int n_{11}^*(\omega')v_{21}(\omega')d^3k', \tag{17c}$$
$$(\omega' - \omega)\alpha_{22}(\omega, \omega') = \lambda Z_1 f(\omega)u_{21}(\omega') + gf(\omega)n_{11}^*(\omega')$$
$$+ gf(\omega)\int n_{11}^*(\omega'')v_{22}(\omega'', \omega')d^3k'',$$
$$(m_U - m - \omega - \omega')\alpha_{23}(\omega, \omega') \tag{17d}$$
$$= \lambda Z_1 f(\omega)v_{11}^*(\omega') + gf(\omega)v_{21}(\omega')$$
$$+ gf(\omega)\int n_{12}^*(\omega'', \omega')v_{21}(\omega'')d^3k'', \tag{17e}$$
$$(\omega - \omega' - \omega'')\alpha_{24}(\omega, \omega', \omega'')$$
$$= \lambda Z_1 f(\omega)u_{22}(\omega', \omega'')$$
$$+ \tfrac{1}{2}g[f(\omega)n_{11}^*(\omega')v_{11}(\omega'') + f(\omega)n_{11}^*(\omega'')v_{11}(\omega')]$$
$$+ gf(\omega)\int n_{11}^*(\omega''')v_{23}(\omega''', \omega', \omega'')d^3k''', \tag{17f}$$
$$(\omega'' - \omega' - \omega)\alpha_{25}(\omega, \omega', \omega'') \tag{17g}$$
$$= \lambda Z_1 f(\omega)v_{11}^*(\omega')u_{21}(\omega'') + gf(\omega)v_{22}(\omega', \omega'')$$
$$+ gf(\omega)n_{12}^*(\omega'', \omega') + \int n_{12}^*(\omega'', \omega')v_{22}(\omega''', \omega'')d^3k''',$$
$$(\omega''' + \omega'' - \omega' - \omega)\alpha_{26}(\omega, \omega', \omega'', \omega''')$$
$$= \lambda Z_1 f(\omega)v_{11}^*(\omega')u_{22}(\omega'', \omega''')$$
$$+ gf(\omega)v_{23}(\omega', \omega'', \omega''') + \tfrac{1}{2}g[f(\omega)n_{12}^*(\omega'', \omega')v_{11}(\omega''')$$
$$+ f(\omega)n_{12}^*(\omega''', \omega')v_{11}(\omega'')]$$
$$+ gf(\omega)\int n_{12}^*(\omega'''', \omega')v_{23}(\omega'''', \omega'', \omega''')d^3k''''. \tag{17h}$$

In view of the boundary conditions imposed on the Heisenberg field operators by the in-field expansions, we want only the retarded $(+i\epsilon)$ solutions of these equations. With this in mind, we proceed with the N-quantum solution of the theory up through the U sector. The analysis of the V and $V\theta$ sectors of the Lee model by this means has already been given in Ref. 8, and the former case is included here for reference. It is obvious that the V sector is unaltered by the additional λ coupling.

The well-known expression for $Z_V\delta m$ is obtained by combining Eqs. (15a) and (16a). We find

$$Z_V\delta m = -g^2\int f^2(\omega)d\omega/\omega. \tag{18}$$

The coefficient v_{11} follows from the separable integral equation that results when Eqs. (15b) and (16b) are combined. Thus

$$h(\omega)v_{11}(\omega) = gf(\omega), \tag{19}$$

where $h(\omega)$ is a function familiar from the classic work of Källén and Pauli,[22] and defined by

$$h(\omega) = Z_V\omega + Z_V\delta m + g^2\int f^2(\omega')d^3k'/(\omega' - \omega - i\epsilon). \tag{20a}$$

In view of Eq. (18), we may also write

$$h(\omega) = \omega[Z_v + g^2\int f^2(\omega')d^3k'/\omega'(\omega' - \omega - i\epsilon)]. \tag{20b}$$

In this way, we extract the zero of $h(\omega)$ at $\omega = 0$ and obtain a function within the square brackets of Eq. (20b), which has a cut for $\mu < \omega < \infty$ and no zero or poles in the cut plane. Use of Eq. (19) in Eqs. (16b) and (17b) leads to

$$n_{12}(\omega, \omega') = \alpha_{12}(\omega, \omega') = g^2f(\omega)f(\omega)/(\omega' - \omega + i\epsilon)h(\omega'). \tag{21}$$

To finish off the NQA solution of the V sector, it remains to calculate the S matrix element for $N\theta$ elastic scattering, and the V-field renormalization constant, Z_v. The first of these follows from Eqs. (7) and (21), and is given by

$$S_{N\theta - N\theta'} = \delta(\mathbf{k} - \mathbf{k}') - 2\pi i\delta(\omega - \omega')g^2f^2(\omega)/h(\omega), \tag{22}$$

where the Riemann–Lebesgue lemma has been used in the form

$$\lim_{t\to\infty}[\exp(iWt)/(W - i\epsilon)] = 2\pi i\delta(W). \tag{23}$$

The vanishing of $h(\omega)$ at zero meson energy signifies that the exact (renormalized) $N\theta$ scattering amplitude $A_{N\theta}(\omega)$ defined via Eq. (22) by

$$A_{N\theta}(\omega) = -4\pi g^2\omega f^2(\omega)/h(\omega) \tag{24}$$

has a pole derived from the V particle. The conventional requirement that $A_{N\theta}(\omega)$ at $\omega = 0$ be equal to the Born amplitude $-4\pi g_0^2f^2(\omega)$ prefaced by g^2 instead of the square of the unrenormalized coupling constant, g_0^2, leads to the desired expression

$$Z_V = 1 - g^2\int f^2(\omega)d^3k/\omega^2. \tag{25}$$

This is equivalent to setting the residue of $h^{-1}(\omega)$ at $\omega = 0$ equal to unity or, alternatively, writing $h'(0) = 1$, where the prime denotes differentiation with respect to ω. Substitution of Eq. (25) into Eq. (20b) leads to yet another version of $h(\omega)$, namely

$$h(\omega) = \omega\left(1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h(\omega')d\omega'}{\omega'^2(\omega' - \omega - i\epsilon)}\right). \tag{20c}$$

Hereafter, the quantity in parentheses is represented by the function $\beta(\omega)$.

We next construct the diagrammatic equivalents of the equations for the coefficients associated with the V sector. Again, this has already been considered in Ref. 8, but is repeated here in our notation for convenience. Equation (16a) for n_{11} is shown in Fig. 2. The same figure with the θ particle instead of the N particle off the mass shell gives the equation for α_{11}. The V-particle mass renormalization equation is depicted in Fig. 3 and follows from Eq. (15a). Equations (15b) and (16b), showing the coupling between

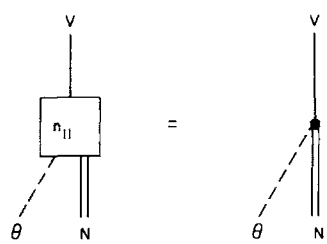


FIG. 2. Diagrammatic representation of Eq. (16a) for n_{11}.

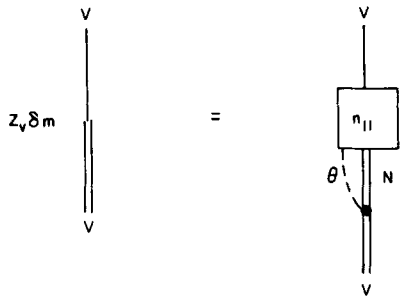


FIG. 3. V-particle mass renormalization given by Eq. (15a).

v_{11} and n_{12}, are illustrated in Fig. 4. Equation (19) is given by Fig. 5. Analogous figures representing the connection between v_{11} and α_{12} follow from Eqs. (15b) and (17b) and the equality $\alpha_{12}=n_{12}$.

It is of interest to note here in the V-sector treatment a comparison between objects of calculation in the NQA and dispersion theory. The latter approach commonly deals with matrix elements of a current operator which can be expressed in terms of the matrix element of the corresponding field operator and thereby makes connection with a c-number coefficient of the NQA. For example, the $VN\theta$ interaction factor, $gf(\omega)$, on the right-hand side of Eq. (17a) equals the matrix element $\langle N_{1n}|J|V_{1n}\rangle$, where J is the θ-particle current operator $J(t)$ at $t=0$. Here we define

$$J(t)\equiv\left(-i\frac{d}{dt}+\omega\right)A(\mathbf{k},t)=[H,A(\mathbf{k},t)]+\omega A(\mathbf{k},t), \qquad (26)$$

where H is the renormalized Hamiltonian of the theory. Using Eq. (26) to express the matrix element of J in terms of the matrix element of $A(\mathbf{k},0)$, we learn that α_{11} is equal to $-gf(\omega)/\omega$ in agreement with Eq. (17a). Likewise one defines a V-particle current operator

$$J_V(t)\equiv\left(-i\frac{d}{dt}+m\right)V(t)=[H,V(t)]+mV(t), \qquad (27)$$

and finds in dispersion theory[23] that the $VN\theta$ vertex function $\langle 0|J_v|N\theta_{1n}\rangle$ is calculated to be $-g\omega f(\omega)/h(\omega)$.

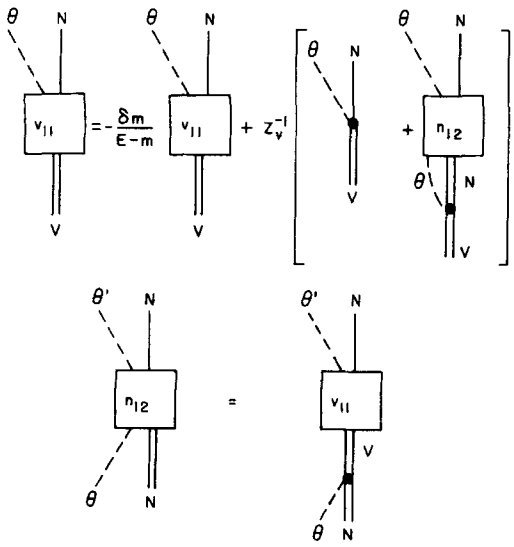


FIG. 4. Equations coupling v_{11} and n_{12}.

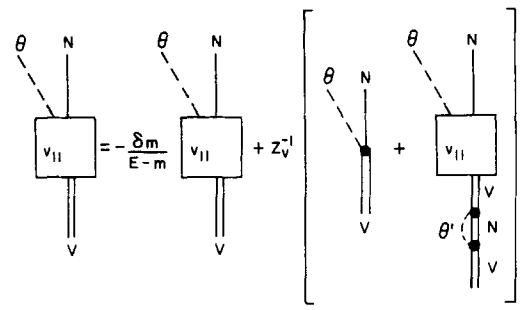


FIG. 5. Algebraic equation for v_{11} obtained from the graphs in Fig. 4.

The matrix element of the V-field operator at $t=0$ represents v_{11}, and thus Eq. (27) implies that v_{11} is given by $gf(\omega)/h(\omega)$, as expected. Again, in dispersion theory[23] the matrix element $\langle N_{1n}|J|N\theta'_{1n}\rangle$ for $N\theta$ scattering is found to be the negative of $gf(\omega)f(\omega')/h(\omega)$. Therefore, Eq. (26) implies that α_{12} is given as in Eq. (21). Of course, a dispersion relations calculation of the matrix element of a field operator itself with respect to the appropriate in-states will lead directly to the corresponding NQA coefficient. Both the NQA and dispersion theory, like the Tamm—Dancoff and LSZ formalisms, are nonperturbative methods of solution. However, in the V-sector the second named approach requires the solution of an Omnes-type and a Low-type integral equation while the other methods achieve the desired results algebraically.

The dynamics of the U-sector is carried by the three coefficients shown in Fig. 1, each of which satisfies an integral equation that can be derived from the set of equations for the c-number coefficients. For the purpose of interpreting the structure of these integral equations, it is instructive to also present them by graphical means. However, for economy of presentation we do not provide diagrams for all the equations in the set. The U-particle mass renormalization is given graphically in Fig. 6.

We first consider the appropriate graphs for a diagrammatic construction of the integral equation for v_{21}. To do this, we seek the coupling between the $UV\theta$ and $UN\theta\theta$ vertex functions v_{21} and n_{21}, each with one leg off the mass shell. Taking into account the symmetry property of identical bosons, we find the relevant diagrams are those displayed in Fig. 7, which lead to the separated integral for v_{21} shown in Fig. 8. In analytic form, this equation reads

$$h(x-\omega)v_{21}(\omega)=\lambda Z_1 f(\omega)+g^2 f(\omega)\int\frac{f(\omega')v_{21}(\omega')d^3k'}{x-\omega-\omega'+i\epsilon}, \qquad (28)$$

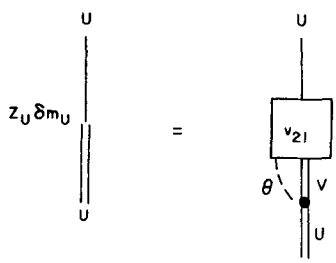


FIG. 6. U-particle mass renormalization given by Eq. (12).

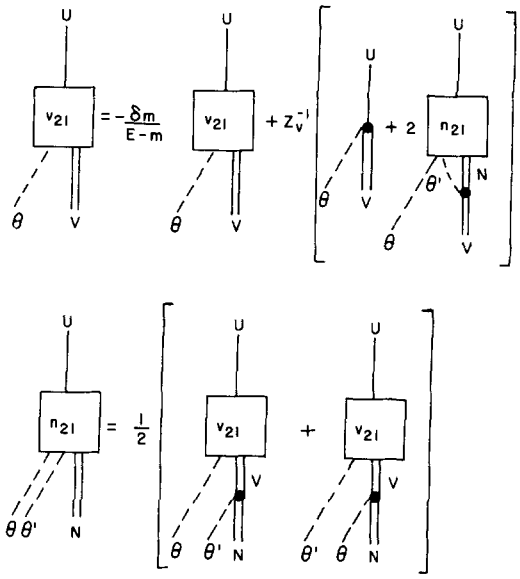


FIG. 7. Equations coupling v_{21} and n_{21}.

where $x=m_U-m$, and the it may be dropped here since for a stable U particle we have $x<\mu$. A comparison of graphs in Figs. 5 and 8 explains the factor of $h(x-\omega)$ appearing on the left-hand side of Eq. (28). We wish to rewrite Eq. (28) in terms of another function, $\phi(\omega,x)$, defined by

$$\phi(\omega,x)=\frac{h(x-\omega)\,v_{21}(\omega)}{\lambda L_1 f(\omega)}. \tag{29}$$

It is easily seen that $\phi(\omega,x)$ satisfies the integral equation

$$\phi(\omega,x)=1-\frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h(\omega')\phi(\omega',x)d\omega'}{h(x-\omega')(\omega'+\omega-x)}, \tag{30}$$

which is similar to the well-known Källén–Pauli singular integral equation. There are standard analytical techniques for solving this type of equation and generalizations thereof.[24] Let us obtain next the integral equation for the $V\theta$ scattering amplitude embodied in the coefficient v_{22} with the V leg off the mass shell. In this case we expand v_{22} in terms of those graphs having either a U field or an N field off the mass shell as shown in Fig. 9. This implicates three other coefficients, and the expansion of one of these, namely n_{22}, also given in the same figure, yields the coupling

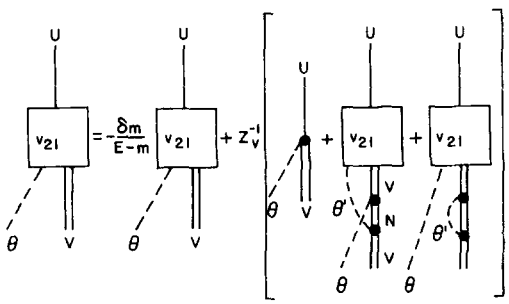


FIG. 8. Integral equation for v_{21} obtained from the graphs in Fig. 7.

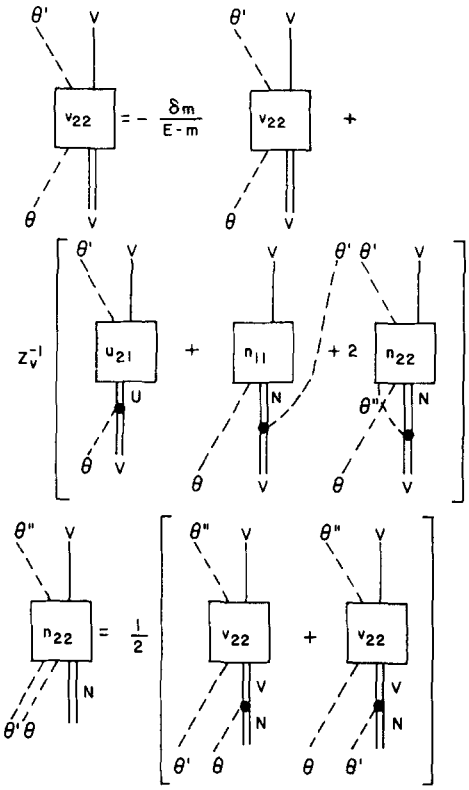


FIG. 9. Equations showing the expansion of v_{22} and its coupling with n_{22}.

of diagrammatic equations which leads to the following analytic form of the integral equation for v_{22} illustrated in Fig. 10:

$$h(\omega'-\omega)v_{22}(\omega,\omega')=-g^2f(\omega)f(\omega')/\omega+\lambda Z_1 f(\omega)u_{21}+g^2f(\omega)\int\frac{f(\omega'')v_{22}(\omega'',\omega')d^3k''}{\omega'-\omega-\omega''+i\epsilon}. \tag{31}$$

As before, the presence of the factor $h(\omega'-\omega)$ on the left-hand side of this equation is the result of combining

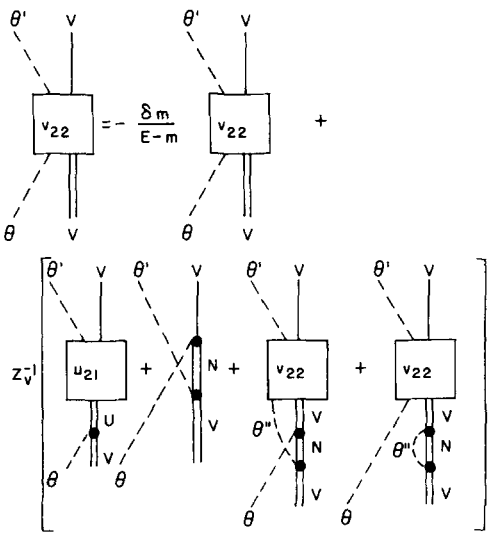


FIG. 10. Integral equation for v_{22} obtained from the graphs in Figs. 2 and 9.

the first and fifth graphs on the right side with the graph on the left side in Fig. 10. Of course, here as in Fig. 8, the closed loop involving the box signifies the occurrence of an integral equation. The absence of this feature in Fig. 5 led to an algebraic equation for v_{11} as we have seen. It is useful to introduce the function $\chi(\omega, \omega')$ defined by

$$\chi(\omega, \omega') = h(\omega' - \omega) v_{22}(\omega, \omega') / g^2 f(\omega) f(\omega') . \tag{32}$$

One can then readily establish that χ satisfies the singular integral equation

$$\chi(\omega, \omega') = -\frac{1}{\omega} + \chi_0(\omega') - \frac{1}{\pi} \int_\mu^\infty \frac{\mathrm{Im} h(\omega'') \chi(\omega'', \omega') d\omega''}{h(\omega' - \omega'')(\omega'' + \omega - \omega' - i\epsilon)} , \tag{33}$$

where $\chi_0(\omega')$ is an abbreviation for $\lambda Z_1 u_{21}(\omega') / g^2 f(\omega')$. This differs from the Källén–Pauli equation for the $V\theta$ amplitude, off the mass shell, in the usual Lee model by the presence of the additional term χ_0 in the inhomogeneous part due to the U particle.

We now proceed to set up the third integral equation in the U sector. The relevant "physical" process is the collision of a meson with a V particle and the production of an N and two mesons by these particles. The appropriate coefficient is v_{23} with the V leg off shell. As in the previous cases, we look for an expansion of v_{23} in terms of other possible coefficients permitted by the selection rules in the theory. The graphs in question are shown in Fig. 11 where the expansion of n_{23} provides the diagrammatic coupling of equations needed for the separated integral equation for v_{23} given in Fig. 12. The mathematical expression of this figure takes the form

$$\begin{aligned} & h(\omega' + \omega'' - \omega) v_{23}(\omega, \omega', \omega'') \\ &= \frac{g^3 f(\omega) f(\omega') f(\omega'')}{2} \left(\frac{1}{h(\omega'')(\omega'' - \omega + i\epsilon)} + \frac{1}{h(\omega')(\omega' - \omega + i\epsilon)} \right) \\ &\quad + \lambda Z_1 f(\omega) u_{22}(\omega', \omega'') \\ &\quad + g^2 f(\omega) \int \frac{f(\omega''') v_{23}(\omega''', \omega', \omega'') d^3 k'''}{\omega'' + \omega' - \omega - \omega''' + i\epsilon} . \end{aligned} \tag{34}$$

Again, except for the u_{22} term due to the U particle, this is the equation for the production amplitude, off the mass shell, found in the unmodified Lee model. Here, also, we introduce another function $\psi(\omega, \omega', \omega'')$ defined by

$$\psi(\omega, \omega', \omega'') = \frac{-2h(\omega' + \omega'' - \omega) h(\omega') h(\omega'') v_{23}(\omega, \omega', \omega'')}{g^3 f(\omega) f(\omega') f(\omega'')} \tag{35}$$

and obeying the integral equation

$$\begin{aligned} & \psi(\omega, \omega', \omega'') \\ &= \frac{h(\omega')}{\omega - \omega'' - i\epsilon} + \frac{h(\omega'')}{\omega - \omega' - i\epsilon} \\ &\quad + \psi_0(\omega', \omega'') \\ &\quad - \frac{1}{\pi} \int_\mu^\infty \frac{[\mathrm{Im} h(\omega''')] \psi(\omega''', \omega', \omega'') d\omega'''}{h(\omega' + \omega'' - \omega''')(\omega''' + \omega - \omega' - \omega'' - i\epsilon)} , \end{aligned} \tag{36}$$

where we let $\psi_0(\omega', \omega'')$ symbolize $-2\lambda Z_1 h(\omega') h(\omega'') u_{22}(\omega', \omega'') / g^3 f(\omega') f(\omega'')$. In solving integral equations such as (33) and (36) the unknown terms χ_0 and ψ_0 are treated as constants which are determined after the formal solutions are found. These points will be taken up in the next section which deals with the NQA solution of the full U sector.

A review of the preceeding equations shows that the V and U sectors of the Bronzan–Lee model are solved, respectively, by one algebraic and three singular integral equations in the NQA. One less integral equa-

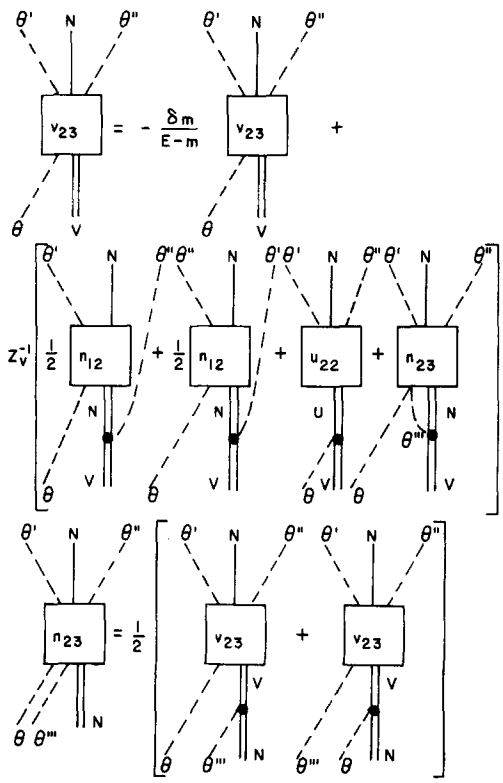


FIG. 11. Equations showing the expansion of v_{23} and its coupling with n_{23}.

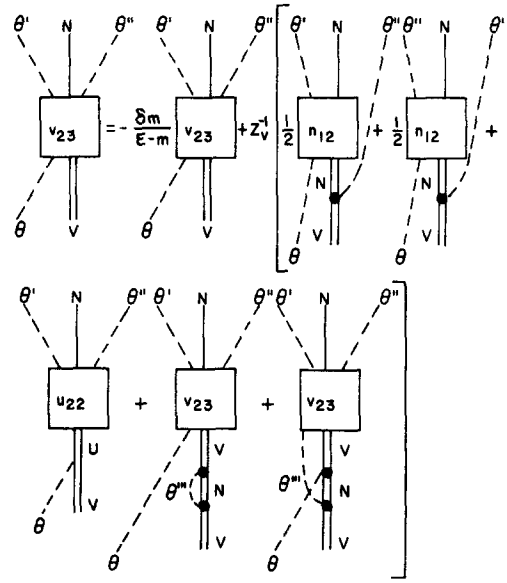


FIG. 12. Integral equation for v_{23} obtained from the graphs in Fig. 11.

tion is required in a similar treatment of the $V\theta$ sector of the Lee model. As in the Tamm—Dancoff eigenvalue method one works with off-the-mass-shell amplitudes such as those defined by Eqs. (21), (32), and (35). In the U sector this entails two separate problems. One of these involves the $V\theta$ scattering amplitude with the V leg off shell and the production process $V\theta \to N\theta\theta$ with a different off-shell extrapolation, while the other deals with the amplitude for $N\theta\theta$ scattering with the N leg off shell through its connections with the amplitude for $N\theta$ scattering and the production amplitude with the V leg off shell. In the LSZ approach the solution to two singular integral equations solves the entire U sector. One such equation performs a similar function in the $V\theta$ sector of the ordinary Lee model. There is a great deal of similarity between the LSZ analysis of the U sector and the VV sector of the Lee model. The latter problem is also solved by two singular integral equations, yielding two scattering amplitudes, a production amplitude, and an equation for the determination of the two-heavy-particle bound state energy which follows from the analytic properties of the τ function appropriate to the VV propagator. An analogous consideration of the U particle propagator leads to its mass renormalization constant.

It is known that the LSZ method of solution in models of the present type has both a formal and calculational appeal not enjoyed by the N-quantum, dispersion, and eigenvalue treatments. The reason for this is traced to the additional degree of freedom due to the time variable in the τ functions.[16] These functions are coupled through the Matthew—Salam equations which contain the dynamics of the model in question. In momentum space the τ functions in these equations carry a Fourier transform variable corresponding to the time, and the extra information in this variable reduces to a minimum the number of integral equations needed for a complete solution. From a mathematical point of view the NQA solution of the U sector is of the same order of complexity as the Tamm—Dancoff eigenvalue method. The former has the advantage of yielding the complete Heisenberg fields, while the latter uses expansions in terms of states. Of course, on the relativistic level the NQA is to be desired since it works directly with connected graphs and is manifestly covariant.[1] The inclusion of the U particle in Amado's dispersion calculation of $V\theta$ elastic scattering yields a very complicated set of dispersion relations.

V. N-QUANTUM SOLUTION OF THE TWO-MESON U SECTOR

This section is concerned with the evaluation of coefficients associated with the U sector. For this purpose, we require the solutions to the three singular integral equations developed in the previous section. After obtaining expressions for v_{21}, v_{22}, v_{23}, we make use of the c-number equations to determine the remaining U-sector coefficients. From these results we are led to the renormalization constants and the S-matrix elements.

It is not necessary to dwell on the available techniques for solving Eqs. (30), (33), and (36). As indicated earlier, standard mathematical routines exist for dealing with these equations. In spirit we adopt the procedures used in previous papers.[17,25] The mathematical structure of Eqs. (30), (33), and (36) should be compared with the integral equations that occur in the $2V$ sector of the Lee model.[25] In this comparison it will be observed that the principal difference between the two classes of equations is the use of the inverse V-particle propagator function $h(\omega)$ in the former case vis-à-vis the corresponding VN propagator function in the latter case. Both functions have a zero point, the same cut from μ to positive infinity, and the same behavior at infinity. Thus, the same techniques of analysis apply in both instances. From the solution to Eq. (30), we have via Eq. (29) that

$$v_{21}(\omega) = \frac{\lambda Z_1 f(\omega)}{Z_V[1+h(x)I_x(x)]}\left(\frac{1}{x-\omega}+\beta(x)\,[I_x(x-\omega)-I_x(x)]\right). \tag{37}$$

The integral $I_W(Z)$ was first introduced in the solution of the $V\theta$ sector of the original Lee model.[26] Its general definition in the complex Z plane is given by

$$I_W(Z)=\frac{1}{\pi}\int_\mu^\infty \mathrm{Im}\left(\frac{1}{h(\omega)}\right)\frac{d\omega}{(\omega-Z)\beta(W-\omega)}\,, \tag{38}$$

where W is a real variable. Both $h(x)$ and $I_x(x)$ are real integrals since the stability of the U particle requires that $x<\mu$.

The mass renormalization is calculated by substituting Eq. (37) into Eq. (12) and carrying out some contour integrations. In this process one establishes the expression

$$\frac{1}{\pi}\int_\mu^\infty [\mathrm{Im}\beta(\omega)]I_x(x-\omega)d\omega = 1-Z_V^{-1}-\frac{Z_V}{2}[\beta^{-1}(x)-Z_V^{-2}-xI_x(x)]. \tag{39}$$

Since the steps leading to Eq. (39) are not so obvious, the interested reader may wish to consult the Appendix of Ref. 17, where a similar integral is evaluated in detail. As expected from previous works the result for $Z_U\delta m_U$ is

$$Z_U\delta m_U=(\lambda^2Z_1^2/2g^2)[x+2\delta m-h(x)/Z_V^2\,d^+(x)], \tag{40}$$

where $d^\pm(x)\equiv 1\pm h(x)I_x(x)$. In the NQA, the mass renormalizations $Z_V\delta m$ and $Z_U\delta m_U$ are obtained from equations for c-number coefficients. Another approach makes use of the property that the V and U propagators have poles at the observed masses m and m_U, respectively. In the same context, the wavefunction renormalization constants Z_V and Z_U are derived from the fact that the corresponding residues of these propagators at these poles equal unity. To determine Z_U, we adopt here a definition of the coupling constant λ which makes the exact scattering of mesons by a V particle equal to the lowest order perturbation theory result at the unphysical energy $\omega=x$. The exact amplitude for this process will follow from the solution to Eq. (33). In the meanwhile, the vertex function renormalization constant Z_1 is determined as usual by prescribing that the $UV\theta$ vertex function be normalized to unity on the mass shell. In this way, we have

$$Z_1 = \phi^{-1}(x, x) = Z_V[1 + h(x)I_x(x)]. \tag{41}$$

Having secured v_{21} and all the V-sector coefficients, we are now in position to derive expressions for the coefficients n_{21}, α_{23}, and α_{21}. The first of these is immediately read off from Eqs. (16c) and (37). The second is obtained from Eqs. (17e), (19), (21) and the solution of Eq. (30) embodied in Eqs. (29) and (37). We write α_{23} in the convenient form

$$\alpha_{23}(\omega, \omega') = \frac{gf(\omega)f(\omega'')}{x-\omega-\omega'}\left(\frac{v_{21}(\omega')}{f(\omega')} + \frac{v_{21}(x-\omega')}{f(x-\omega')}\right). \tag{42}$$

In a like manner, we use Eqs. (16a) and (17c) to see that the third coefficient α_{21} is expressed by

$$\alpha_{21}(\omega) = \lambda f(\omega)/(x-\omega). \tag{43}$$

We now turn our attention to Eq. (33) and results derivable from its solution. Again, it is worth pointing out the similarity in the mathematical structure between this equation and the singular integral equation that describes $VN\theta$ scattering in the Lee model.[25] From the solution to Eq. (33), and in view of Eq. (32), the coefficient v_{22} for the description of $V\theta$ scattering becomes

$$\begin{aligned} v_{22}(\omega, \omega') &= \frac{g^2 f(\omega)f(\omega')}{\omega(\omega-\omega'-i\epsilon)}\left(\frac{d^*(\omega')}{\beta(\omega')d^*(\omega')} + \frac{2(\omega'-\omega)I_{\omega'}(\omega'-\omega)}{d^*(\omega')}\right) \\ &\quad - \frac{\lambda Z_1 f(\omega)u_{21}(\omega')}{(\omega-\omega'-i\epsilon)Z_V} \\ &\quad \times\left(1 + \frac{h(\omega')[(\omega'-\omega)I_{\omega'}(\omega'-\omega) - \omega' I_{\omega'}(\omega')]}{\omega d^*(\omega')}\right). \end{aligned} \tag{44}$$

At this point we find the U-sector coefficient u_{21} from Eq. (14a) by making use of the above expression for v_{22} and recalling the important relation established in Eq. (39). After a considerable amount of algebra, we obtain the desired result

$$u_{21}(\omega) = 2\lambda Z_V Z_1 f(\omega)/D(\omega, x), \tag{45}$$

where

$$\begin{aligned} D(w, x) &= Z_V^2[2Z_U - (\lambda Z_1/g)^2](\omega - x)d^*(\omega) + (\lambda Z_1/g)^2 \\ &\quad \times [h(\omega) - h(x)d^*(\omega)/d^*(x)]. \end{aligned} \tag{46}$$

Thus v_{22} is completely determined when Eqs. (44) and (45) are combined. With the solutions for v_{22} and u_{21} at hand, we can now find n_{22} and α_{22}. The former follows straightaway from Eq. (16d) while the latter is determined through Eq. (17d) and the already known coefficients appearing therein. We easily find that

$$\begin{aligned} \alpha_{22}(\omega, \omega') &= \frac{g^2 f(\omega)f(\omega')}{\omega' - \omega + i\epsilon} \\ &\quad \times\left(\frac{2(\lambda Z_1/g)^2}{D(\omega')d^*(\omega')} - \frac{d^*(\omega')}{h(\omega')d^*(\omega')}\right). \end{aligned} \tag{47}$$

Concluding this part on the solution to Eq. (33) and its use in computing coefficients, we evaluate the S-matrix elements for the elastic scattering of a meson V particle, and the production of two mesons and an N particle. At the same time, we complete the determination of the renormalization constants in the theory by obtaining Z_U. Equations (8), (23), (44), and (45) lead to

$$\begin{aligned} S_{V\theta \to V\theta'} &= \delta(\mathbf{k}-\mathbf{k}') + 2\pi i\,\delta(\omega-\omega')g^2 f^2(\omega) \\ &\quad \times\left[\frac{d^*(\omega)}{h(\omega)d^*(\omega)} - \frac{2(\lambda Z_1/g)^2}{D(\omega, x)d^*(\omega)}\right], \end{aligned} \tag{48}$$

Here the second term in the square brackets is obviously due to the λ coupling and, on the mass shell, $\omega = x$, contains the direct U-particle pole, signified by $D(x, x) = 0$. The first term alone describes $V\theta$ scattering in the Lee model, and has been derived and discussed in the literature by a variety of methods, including the one presently under examination. The possibility for a strong coupling $V\theta$ bound state in the Lee model, signaled by the vanishing of $d^*(\omega)$ for $\omega < \mu$, has been well documented.[27] Of course, it is seen from Eq. (48) that this situation can also arise in the Bronzan–Lee model. We wish further to point out that, had the meson in the out-state of $S_{V\theta}$ been contracted instead of the V particle, then Eq. (8) would contain α_{22} in place of v_{22}, as evidenced by Eq. (47).

To establish Z_U, we proceed in analogy with the $N\theta$ scattering case and define λ by comparing the exact (renormalized) expression for the $V\theta$ scattering amplitude with the result of lowest order perturbation theory at the unphysical meson energy $\omega = x$, where in the latter case the unrenormalized $UV\theta$ coupling constant λ_0 is replaced by λ. This is equivalent to setting the residue of the second term in the square bracket in Eq. (48) at $\omega = x$ equal to λ^2. This translates into the statement

$$D'(\omega, x)\big|_{\omega=x} = 2Z_V^2 d^*(x), \tag{49}$$

which leads via the definition of $D(\omega, x)$ to the form

$$\begin{aligned} Z_U &= 1 + \tfrac{1}{2}(\lambda Z_1/g)^2 - \tfrac{1}{2}(\lambda Z_1/gZ_V)^2 \\ &\quad \times [d^*(x)]^{-2}[h'(x) - h^2(x)I'_x(x)], \end{aligned} \tag{50}$$

as expected. To calculate the production amplitude, we combine Eqs. (9), (16d), and (44). The resulting expression is simplified by employing an identity[26] satisfied by $I_W(Z)$ and yields

$$\begin{aligned} P_{V\theta \to N\theta'\theta''} &= 2\pi i g^3 f(\omega)f(\omega')f(\omega'') \\ &\quad \times\sqrt{2}\left(\frac{1}{d^*(\omega)h(\omega')h(\omega'')} - \frac{(\lambda Z_1/g)^2 h(\omega)}{D(\omega, x)d^*(\omega)h(\omega')h(\omega'')}\right). \end{aligned} \tag{51}$$

The final integral equation to consider is that satisfied by $\Psi(\omega, \omega', \omega'')$, where ω' and ω'' enter only as parameters. Note the similarity between Eq. (36) and the Tamm–Dancoff equation in Ref. 25 which describes the scattering of two meson by two coincident N-particles. These equations differ from the corresponding equation for $N\theta\theta$ scattering in the Lee model only in their homogeneous terms. Equation (14b) and the formal solution to Eq. (30) provide two simultaneous relations for the evaluation of v_{23} and u_{22}. These coefficients in turn determine the remaining α coefficients in the U sector, and the S matrix element for $N\theta\theta$ scattering.

The solution to Eq. (36), together with Eq. (35), gives

$$v_{23}(\omega,\omega',\omega'')$$

$$=\frac{(\lambda Z_1/Z_V)f(\omega)u_{22}(\omega',\omega'')}{\omega'+\omega''-\omega+i\epsilon}\left(1+\frac{h(\omega'+\omega'')}{\omega d^+(\omega'+\omega'')}\right.$$

$$\times[(\omega'+\omega''-\omega)I_{\omega'+\omega''}(\omega'+\omega''-\omega)-(\omega'+\omega'')I_{\omega'+\omega''}$$

$$\left.\times(\omega'+\omega'')]\right)$$

$$-\frac{g^3f(\omega)f(\omega')f(\omega'')}{2h(\omega')h(\omega'')(\omega+\omega''-\omega+i\epsilon)}\left(\frac{\omega'}{\omega-\omega''-i\epsilon}+\frac{\omega''}{\omega-\omega'-i\epsilon}\right.$$

$$+\frac{2h(\omega'+\omega'')}{\omega d^+(\omega'+\omega'')}$$

$$\times[(\omega'+\omega''-\omega)I_{\omega'+\omega''}(\omega'+\omega''-\omega)-(\omega'+\omega'')I_{\omega'+\omega''}$$

$$\left.\times(\omega'+\omega'')]\right). \tag{52}$$

At this stage we have essentially the same mathematical situation for v_{23} and u_{22} as previously for v_{22} and u_{21}. The coupled equations for the former pair of coefficients are now given by Eqs. (14b) and (52). Again, after a tedious but straightforward computation, we arrive at

$$u_{22}(\omega,\omega')=g\lambda Z_V Z_1 f(\omega)f(\omega')h(\omega+\omega')/h(\omega)h(\omega')$$
$$\times D(\omega+\omega',x). \tag{53}$$

The substitution of this result into Eq. (52) provides the final answer for v_{22}. Then, of course, n_{23} follows at once from Eqs. (16e), (52) and (53). It remains only to determine α_{26} and α_{24} in terms of the fundamental functions in the theory. In the former case, we learn from Eq. (17h) and the known expressions for the coefficients on the right-hand side that

$$\alpha_{26}(\omega'''+\omega''-\omega'-\omega)$$

$$=[g^4f(\omega)f(\omega')f(\omega'')f(\omega''')h(\omega''+\omega''')]$$
$$\times[h(\omega')h(\omega'')h(\omega''')h(\omega'''+\omega''-\omega')$$
$$\times d^+(\omega''+\omega''')(\omega'''+\omega''-\omega'-\omega+i\epsilon)]^{-1}$$
$$\times\left(\frac{(\lambda Z_1/g)^2h(\omega''+\omega''')}{D(\omega''+\omega''',x)}-1\right). \tag{54}$$

Finally, α_{24} follows from Eq. (17f). Here also all the coefficients on the right-hand side are at hand, and we easily find that

$$\alpha_{24}(\omega,\omega',\omega'')$$

$$=h(\omega'+\omega''-\omega)v_{23}(\omega,\omega',\omega'')/(\omega-\omega'-\omega''+i\epsilon). \tag{55}$$

This completes the derivation of all the NQA coefficients in the V and U sectors of the Bronzan–Lee model. To conclude this phase which deals with the solution to Eq. (36) and its ramifications, we calculate the S-matrix element for the scattering of two mesons by an N particle. For this purpose we require the use of Eqs. (10), (16e), (21), (23), (52), and (53) to obtain

$$S_{N\theta\theta'\to N\theta''\theta'''}$$

$$=\tfrac{1}{2}[\delta(\mathbf{k}-\mathbf{k}'')\delta(\mathbf{k}'-\mathbf{k}''')+\delta(\mathbf{k}-\mathbf{k}''')\delta(\mathbf{k}'-\mathbf{k}'')]$$

$$-\pi ig^2\delta(\omega+\omega'-\omega''-\omega''')f(\omega)f(\omega')f(\omega'')f(\omega''')$$

$$\times\left\{\frac{[\delta(\mathbf{k}-\mathbf{k}'')+\delta(\mathbf{k}-\mathbf{k}''')]}{h(\omega')}f^2(\omega)+\frac{[\delta(\mathbf{k}'-\mathbf{k}'')+\delta(\mathbf{k}'-\mathbf{k}''')]}{h(\omega)}f^2(\omega')\right\}$$

$$+2\pi i[g^4\delta(\omega+\omega'-\omega''-\omega''')f(\omega)f(\omega')f(\omega'')f(\omega''')$$

$$\times h(\omega+\omega')][h(\omega)h(\omega')h(\omega'')h(\omega''')d^+(\omega+\omega')]^{-1}$$

$$\times\left(1-\frac{(\lambda Z_1/g)^2h(\omega'+\omega'')}{D(\omega'+\omega'',x)}\right). \tag{56}$$

The interpretation of the structure of $S_{N\theta\theta}$, parallels that given in the case of the scattering of two mesons by two N particles. The Dirac delta functions in the first part of Eq. (56) denote the possibility of no scattering, and the following part in curly brackets describes the elastic scattering of one meson by the N source while the other meson is unscattered. The first term in the square brackets of the connected part arises in the ordinary Lee model ($\lambda=0$). The second term is obviously due to the influence of the U particle. It is clear from the form of α_{26} in Eq. (54) that this coefficient describes the connected part of the $N\theta\theta$ elastic scattering amplitude.

VI. CONCLUSION

Over the years a number of soluble quantum field-theoretic models have been proposed as guides to understanding mass and charge renormalization, the dynamics of strong interactions, and a variety of techniques of calculation which include the Tamm–Dancoff eigenvalue approach, dispersion relations, the LSZ formalism, functional methods, and the NQA. The Bronzan–Lee model provides a very valuable context for all of these considerations, and in this paper we have used it to extend the application of the NQA and to gain insight into the methodology of this calculational tool.

We have seen that the single-source one- and two-meson sectors of this model are solved in the NQA by one algebraic and three linear integral equations, respectively, the mathematical complexity of which reminds one of the Tamm–Dancoff approach in the $2V$ sector of the Lee model. Although these two methods of solution are qualitatively different, there is the parallel aspect, at least in the present framework, of expanding a Heisenberg field operator in terms of in-fields, and a Schrödinger eigenstate of the total Hamiltonian in terms of bare states, where in each case contributions to the series have the same quantum numbers as the sector under investigation. Thus, in view of the strong selection rules in the model, there is an automatic truncation of the Heisenberg field expansions consistent with a particular sector and the conventional renormalization program. When these expansions are substituted in the field equations, normal-ordered products of in-field operators associated with sectors higher than those under consideration are neglected. In general, we substitute this approach in place of power counting,[9] introduced to achieve a properly renormalized theory, since power counting is evidently inappropriate in higher sectors of the Lee model.[28] This will be shown in a subsequent paper which deals with the $2V$ case, and amplifies the close relationship between the NQA and dispersion relations.

There are extensions of the present work which may also be worth studying. For example, one can think of solving the Bronzan–Lee model in the presence of unstable U particle, $x > \mu$, via the NQA. This would require that U_{in} be removed from the asymptotic operator ring of in-field. On the other hand the existence of a $V\theta$ dynamical bound state would necessitate introducing a new and independent in-field to complete the set of asymptotic operators. In view of the approximate relationship between the Bronzan–Lee model and the more realistic charged scalar theory[29] there is also the possibility of using the NQA to investigate the two-meson solution of the latter case which is distinguished by its two- and three-particle unitarity and a crossing symmetric scattering amplitude. Furthermore, one could approach the interesting question of dynamical bound states in meson–nucleon scattering by using the NQA in charged scalar theory with a sufficiently large coupling constant.

Work on the NQA analysis of the bound state formed by the interaction between two V particles in the Lee model has been carried out.[28] This suggests the possibility of extending the treatment to the two-source problem in charged scalar theory, and then going on to explore the NQA description of the interaction between two nucleons in the classic Chew–Low model.

[1]O.W. Greenberg, Phys. Rev. **139**, B1038 (1965).
[2]J.C. Ward, Phys. Rev. **79**, 406 (1950); C.A. Hurst, Proc. Cambridge Phil. Soc. **48**, 625 (1952); W. Thirring, Helv. Phys. Acta **26**, 331 (1953).
[3]A. Halprin, Phys. Rev. **156**, 1552 (1967).
[4]O.W. Greenberg, Phys. Rev. **156**, 1742(E) (1967).
[5]O.W. Greenberg and R.J. Genolio, Phys. Rev. **150**, 1070 (1966).
[6]O.W. Greenberg, Phys. Rev. **150**, 1076 (1966).
[7]F.J. Dyson, Phys. Rev. **73**, 929 (1948); R. Arnowitt and S. Deser, Phys. Rev. **100**, 349 (1955); J. Okubo, Nuovo Cimento **19**, 574 (1961); T. Pradhan, Proc. Natl. Inst. Sci. India A **28**, 249 (1962); Nucl. Phys. **43**, 11 (1963); B. Klaiber, Nuovo Cimento **30**, 165 (1965).
[8]A. Pagnamenta, Ann. Phys. (N.Y.) **39**, 453 (1966).
[9]A. Halprin, Phys. Rev. **172**, 1495 (1968).
[10]J.B. Bronzan, Phys. Rev. **139**, B751 (1965).
[11]L.M. Scarfone, J. Math. Phys. **9**, 977 (1968).
[12]J.B. Bronzan, Phys. Rev. **172**, 1429 (1968).
[13]L.M. Scarfone, Phys. Rev. D **1**, 584 (1970).
[14]P.S. Liossatos, Phys. Rev. **172**, 1554 (1968).
[15]A.L. Choudhury, J. Math. Phys. **14**, 90 (1973).
[16]M.S. Maxon, Phys. Rev. **149**, 1273 (1966).
[17]L.M. Scarfone, J. Math. Phys. **9**, 246 (1968).
[18]H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).
[19]Fei Shian Chen-Cheung, Phys. Rev. **152**, 1408 (1966).
[20]G.R. North, Phys. Rev. **168**, 1698 (1968).
[21]A.C. Fonseca and P.E. Shanley, Phys. Rev. D **13**, 2255 (1976).
[22]G. Källén and W. Pauli, Mat. Fys. Medd. Dan. Vid. Selsk. **30**, No. 7 (1955).
[23]L.M. Goldberger and S.B. Treiman, Phys. Rev. **113**, 1163 (1959).
[24]N.I. Muskhelishivili, *Singular Integral Equations* (Noordhoff, Groningen, The Netherlands, 1953), Chap. 5; E. Kazes, J. Math. Phys. **6**, 1772 (1965); C.M. Sommerfield, *ibid.* **6**, 1170 (1965); T.L. Trueman, Phys. Rev. **137**, B1566 (1965); R.P. Kenschaft and R.D. Amado, J. Math. Phys. **5**, 1340 (1964); A. Pagnamenta, *ibid.* **6**, 955 (1965); G.S. Litvincuk, Izv. Akad. Nauk. SSSR. Ser. Math. **25**, 871 (1961); Refs. 10 and 16.
[25]L.M. Scarfone, Ref. 11.
[26]Refer to the article by A. Pagnamenta in Ref. 24.
[27]Refer to Refs. 16 and 26, and T. Muta, Progr. Theoret. Phys. (Kyoto) **33**, 666 (1965).
[28]J.A. Burns, thesis, University of Vermont (unpublished).
[29]J.B. Bronzan, J. Math. Phys. **7**, 1351 (1966); **8**, 6 (1967); M.S. Maxon, Phys. Rev. **187**, 2129 (1969); G.R. North, Phys. Rev. **188**, 2541 (1969); L.M. Scarfone, Phys. Rev. D **2**, 3054 (1970).

Some remarks on the Wronskian technique and the inverse scattering transform[a)]

K. M. Case and S. C. Chiu

The Rockefeller University, New York, New York 10021
(Received 18 April 1977)

A connection between the Wronskian technique of Calogero and Degasperis and the Lax theorem is made by finding the time evolution of the eigenfunctions of the linear operator for the inverse scattering transform of the isospectral flow equations. The significance of the temporal half of the Bäcklund transformations is discussed via specific examples. An explicit proof of the permutability theorem of the Bäcklund transformation for the case $r = -q$ is given. New Wronskian identities with respect to bound-state constants, which hold for all potentials allowable in the inverse scattering transform, are found.

I. INTRODUCTION

In a series of remarkable papers by Calogero and Degasperis,[1-6] a generalized Wronskian technique was developed and applied to the Schrödinger and the generalized Zakharov–Shabat operator to generate classes of nonlinear evolutionary equations solvable by the inverse-scattering transform, and their Bäcklund transformations. The technique gives much insight to the connection between the linear dispersion relations and the nonlinear equations[7]; furthermore, it generalizes the known Bäcklund transformations and gives a spectral significance to them.

The purpose of the present work is to extend the theory in two respects:

(1) We make a connection between the Wronskian technique and the well-known Lax theorem[8] (in the formalism of Ablowitz *et al.*[9]) by finding the time evolution of the eigenfunctions of the inverse scattering transforms for general nonlinear equations generated by the technique.

(2) Since only the spatial half of the Bäcklund transformations were obtained in Refs. 1–4, we demonstrate how the other half may be obtained in specific examples, and point out their spectral significance with respect to the transmission coefficients.

For the purpose of explicitly proving the permutability theorem of the Bäcklund transformation, we have also shown how a quadratic Bäcklund transformation is composed of two "linear" ones.

Finally, in the Appendix, we derive new Wronskian identities for the discrete eigenstates of the Zakharov–Shabat operator, by which one can obtain the time dependence of the bound-state constants for a larger class of potentials than was assumed in Refs. 1–6.

The present work relies heavily on the results of Refs. 1–4, to which we refer the readers for much of the details.

II. SUMMARY OF RESULTS OF CALOGERO AND DEGASPERIS

In this section, we summarize some of the results of Refs. 1–4. We shall present the results corresponding to the Schrödinger operator in a slightly different form from those of Refs. 1–3 by translating their results into 2×2 matrix forms. This has the advantage of revealing the similarities and differences between the Schrödinger and the Zakharov–Shabat operator; furthermore, one obtains information about the transmission coefficients in addition to the reflection coefficients.

The linear eigenvalue problem of concern here is

$$\psi_x + ik\sigma_3\psi = V\psi, \tag{2.1}$$

where

$$\psi(x,k,t) = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{2.2}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad V(x,t) = \begin{pmatrix} 0 & q \\ r & 0 \end{pmatrix}. \tag{2.3}$$

As indicated, both ψ and V depend on x and t. The Zakharov–Shabat problem and the Schrödinger problem differ in the boundary conditions of V at $x \to \pm\infty$:

(i) For the Zakharov–Shabat problem,

$$\lim_{x\to\pm\infty} q = \lim_{x\to\pm\infty} r = 0; \tag{2.4}$$

(ii) for the Schrödinger problem,

$$r = 1, \quad \lim_{x\to\pm\infty} q = 0. \tag{2.5}$$

For every real k, there exists two independent solutions of (2.1), denoted by the columns of the matrix

$$\Phi(x,k,t) = \begin{pmatrix} \psi_1 & \bar{\psi}_1 \\ \psi_2 & \bar{\psi}_2 \end{pmatrix}. \tag{2.6}$$

Given two sets of potentials V and V' and their corresponding solutions Φ and Φ' of Eq. (2.1), the following fundamental indentity is satisfied by virtue of Eq. (2.1) for an arbitrary matrix function F:

$$\Phi'^{T}\left\{\frac{\partial F}{\partial x} + (V'^{T} - ik\sigma_3)F + F(V - ik\sigma_3)\right\}\Phi$$

$$= \frac{\partial}{\partial x}\left[\Phi'^{T} F \Phi\right], \tag{2.7}$$

[a)]This work was supported in part by the Air Force Office of Scientific Research, Grant 76-3085.

where T denotes matrix transpose. The Wronskian identities are derived from iteration schemes for F in Eq. (2.7).

A. Zakharov-Shabat problem

The eigenfunction for real k are defined by their asymptotic boundary conditions

$$\Phi(x,k,t) \underset{x\to+\infty}{\sim} \begin{pmatrix} \exp(-ikx) & \alpha^{(-)}(k)\exp(-ikx) \\ \alpha^{(+)}(k)\exp(ikx) & \exp(ikx) \end{pmatrix}. \tag{2.8}$$

$$\Phi(x,k,t) \underset{x\to-\infty}{\sim} \begin{pmatrix} \beta^{(+)}(k)\exp(-ikx) & 0 \\ 0 & \beta^{(-)}(k)\exp(ikx) \end{pmatrix}. \tag{2.9}$$

One defines the following quantities and operators:

$$A_{\pm}(k) \equiv \begin{pmatrix} \alpha^{(+)\prime}(k) \pm \alpha^{(+)}(k) & \alpha^{(+)\prime}(k)\alpha^{(-)}(k) \pm \beta^{(+)\prime}(k)\beta^{(-)}(k) \pm 1 \\ 1 \pm \alpha^{(+)}(k)\alpha^{(-)\prime}(k) - \beta^{(+)}(k)\beta^{(-)\prime}(k) & \alpha^{(-)}(k) \pm \alpha^{(-)\prime}(k) \end{pmatrix}, \tag{2.10}$$

$$B(k) \equiv -\begin{pmatrix} 0 & \beta^{(-)}(k)\beta^{(+)\prime}(k) \\ \beta^{(-)\prime}(k)\beta^{(+)}(k) & 0 \end{pmatrix}, \tag{2.11}$$

here, $\alpha^{(\pm)}$ and $\beta^{(\pm)}$ pertains to coefficients, in Eqs. (2.8) and (2.9), for solutions of Eq. (2.1) with potential V, and $\alpha^{(\pm)\prime}$, $\beta^{(\pm)\prime}$ are analogous coefficients for solutions with potential V'.

For arbitrary functions a, b, define

$$\eta\begin{pmatrix} a \\ b \end{pmatrix} \equiv \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \tag{2.12}$$

and

$$\Lambda\begin{pmatrix} a \\ b \end{pmatrix} \equiv \frac{1}{2i}\left[\sigma_3\frac{\partial}{\partial x} + \begin{pmatrix} r'Iq' + rIq & -rIr' - r'Ir \\ q'Iq + qIq' & -q'Ir' - qIr \end{pmatrix}\right]\begin{pmatrix} a \\ b \end{pmatrix}, \tag{2.13}$$

where

$$If(x,k) \equiv \int_x^{\infty} d\xi f(\xi,k). \tag{2.14}$$

For the arbitrary matrix $F(x)$ in Eq. (2.7), we define its diagonal and off-diagonal elements as

$$F(x) = \begin{pmatrix} 0 & F_1 + F_2 \\ F_1 - F_2 & 0 \end{pmatrix} + \frac{1}{2}\sigma_3 H \tag{2.15}$$

where

$$H = \sigma_3 F + F\sigma_3. \tag{2.16}$$

In Ref. 4, an iteration scheme was developed from Eq. (2.7), by which an infinite number of identities can be obtained from an initially known one. We shall use the superscript $^{(n)}$ on the quantities F, H, etc., to denote those obtained at the nth iteration. The subscripts $\pm$ will denote quantities obtained from two initial F's as defined immediately below. Starting with

$$F_{\pm}^{(0)} = \begin{pmatrix} 0 & \pm 1 \\ 1 & 0 \end{pmatrix}, \tag{2.17}$$

one obtains $H_{\pm}^{(0)} = 0$,

$$H_{\pm}^{(1)} = \eta\,(v_{\pm}^{(1)}) \equiv \begin{pmatrix} r' \pm r & 0 \\ 0 & q \pm q' \end{pmatrix}, \tag{2.18}$$

and

$$H_{\pm}^{(n)} = \eta[(i\Lambda)^{n-1}v_{\pm}^{(1)}] \equiv \eta(v_{\pm}^{(n)}), \quad n \geq 1. \tag{2.19}$$

At the nth interation, the terms inside the braces of Eq. (2.7) can be expressed in terms of H's as follows:

$$H_{\pm}^{(n+1)} - ikH_{\pm}^{(n)} = \frac{\partial}{\partial x}F_{\pm}^{(n)} + (V'^{T} - ik\sigma_3)F^{(n)} + F_{\pm}^{(n)}(V - ik\sigma_3) \tag{2.20}$$

for $n \geq 0$, with[10]

$$F_{1\pm}^{(n)} = -\tfrac{1}{4}\int_x^{\infty} d\xi(v_{+}^{(1)}(\xi), i\sigma_2 v_{\pm}^{(n)}(\xi)), \tag{2.21}$$

$$F_{2\pm}^{(n)} = -\tfrac{1}{4}\int_x^{\infty} d\xi(\sigma_3 v_{-}^{(1)}(\xi), \sigma_1 v_{\pm}^{(n)}(\xi)). \tag{2.22}$$

Given $F_{\pm}^{(0)}$ in Eqs. (2.17), Eq. (2.20) determines $H_{\pm}^{(1)}$, and Eqs. (2.19), (2.21), (2.22), (2.15) determines $F_{1\pm}^{(1)}$, $F_{2\pm}^{(1)}$ and thus $F_{\pm}^{(1)}$; and the iteration can be repeated. Integrating Eq. (2.7) for each n and using Eqs. (2.8)–(2.11) give

$$\int_{-\infty}^{\infty} dx(\Phi'^{T}H_{\pm}^{(n+1)}\Phi) = (ik)^n A_{\pm}(k) - \left[\sum_{j=0}^{n-1}(ik)^{n-j-1}(F_{1\pm}^{(j+1)}(-\infty) + F_{2\pm}^{(j+1)}(-\infty)\sigma_3)\right]B(k), \tag{2.23}$$

where $A_{\pm}$ and B are defined in Eqs. (2.10) and (2.11). Nonlinear evolutionary equations are obtained in the limit

$$r' \to r + r_t\,dt, \quad q' \to q + q_t dt. \tag{2.24}$$

Then, the operator Λ becomes

$$\Lambda \to L = \frac{1}{2i}\left[\sigma_3\frac{\partial}{\partial x} + 2\begin{pmatrix} rIq - rIr \\ qIq - qIr \end{pmatrix}\right]; \tag{2.25}$$

$H_{\pm}^{(n)}$ becomes

$$H_{+}^{(n)} - h_{+}^{(n)} = \eta\left[(iL)^{n-1}\begin{pmatrix} 2r \\ 2q \end{pmatrix}\right], \tag{2.26}$$

$$H_-^{(n)} - h_-^{(n)}\,dt = \eta\left[(iL)^{n-1}\begin{pmatrix} r_t \\ -q_t \end{pmatrix}\right]dt. \tag{2.27}$$

In this limit, a linear combination of Eq. (2.23), for different n, implies that if, for constants $C_j^{(\pm)}$,

$$\sum_{l=1}^{n} C_l^{(-)} h_-^{(l)} + \sum_{m=1}^{N} C_m^{(+)} h_+^{(m)} = 0, \tag{2.28}$$

then the evolution of $\alpha^{(\pm)}(k)$ and $\beta^{(\pm)}(k)$ are given by[11]

$$\beta_t^{(\pm)}(k) = 0, \tag{2.29}$$

$$\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}\alpha_t^{(\pm)}(k) \pm \sum_{m=1}^{N} C_m^{(+)}(ik)^{m-1}\alpha^{(\pm)}(k) = 0. \tag{2.30}$$

Equations (2.28) are the nonlinear evolution equations.

Similarly, for two different sets of potentials, (q', r') and (q, r), if, for constants $p_j^{(\pm)}$,

$$\sum_{l=1}^{n'} p_l^{(-)} H_-^{(l)} + \sum_{m=1}^{N'} p_m^{(+)} H_+^{(m)} = 0, \tag{2.31}$$

the reflection and transmission coefficients, $(\alpha^{(\pm)}(k), \beta^{(\pm)}(k))$ and $(\alpha^{(\pm)\prime}(k), \beta^{(\pm)\prime}(k))$ are related by

$$\alpha^{(\pm)\prime}(k) = \alpha^{(\pm)\prime}(k)[f(ik) \mp g(ik)]/[f(ik) \pm g(ik)] \tag{2.32}$$

and

$$\alpha^{(\pm)\prime}(k) = \beta^{(\pm)}(k)\{[f(ik) \mp g(ik)]/[f(ik) \pm g(ik)]\}\theta^{(\mp)}/\theta^{(\pm)}, \tag{2.33}$$

where

$$f(ik) = \sum_{l=1}^{n'} p_l^{(-)}(ik)^{l-1}, \tag{2.34}$$

$$g(ik) = \sum_{m=1}^{N'} p_m^{(+)}(ik)^{m-1}, \tag{2.35}$$

$$\theta^{(\pm)} = \pm\,(F_\pm(ik) + G_\pm(ik)), \tag{2.36}$$

$$F_\pm(ik) = \sum_{l=1}^{n'} p_l^{(-)}\left[\sum_{j=0}^{l-1}(ik)^{l-j-1}(F_{1_-}^{(j+1)}(-\infty) \pm F_{2_-}^{(j+1)}(-\infty))\right] \tag{2.37}$$

$$G_\pm(ik) = \sum_{m=1}^{N'} p_m^{(+)}\left[\sum_{j=0}^{m-1}(ik)^{m-j-1}(F_{1_+}^{(j+1)}(-\infty) \pm F_{2_+}^{(j+1)}(-\infty))\right]. \tag{2.38}$$

Equation (2.31) gives the generalized Bäcklund transformations. The usual Bäcklund transformation is the special case when $f(ik) \pm g(ik)$ are linear functions of k.

B. Schrödinger problem

In this case, $r = 1$. The eigenfunction of Eq. (2.1) for real k are characterized by the boundary conditions

$$\tilde{\Phi}(x,k) \underset{x\to+\infty}{\to} \begin{pmatrix} -2ik\exp(-ikx) & -2ik\bar{R}(k)\exp(-ikx) \\ \exp(-ikx) + R(k)\exp(ikx) & \exp(ikx) + \bar{R}(k)\exp(-ikx) \end{pmatrix}$$
$$\underset{x\to-\infty}{\to} \begin{pmatrix} -2ikT(k)\exp(-ikx) & 0 \\ T(k)\exp(-kx) & \bar{T}(k)\exp(ikx) \end{pmatrix}. \tag{2.39}$$

The coefficients satisfy the relations $\bar{T}(k) = T(-k)$, $\bar{R}(k) = R(-k)$.

For two potentials, q, q', we let

$$D(x) = q - q', \quad S(x) = q + q'. \tag{2.40}$$

For an arbitrary function $\phi(x)$, one defines the matrix

$$\tilde{F}(x) \equiv \begin{pmatrix} -2\phi(x) & \dfrac{\partial\phi}{\partial x} - 2ik\phi - \displaystyle\int_x^\infty dy D(y)\phi(y) \\ \dfrac{\partial\phi}{\partial x} - 2ik\phi + \displaystyle\int_x^\infty dy D(y)\,\phi(y) & -\dfrac{\partial^2\phi}{\partial x^2} + 2ik\dfrac{\partial\phi}{\partial x} + S(x)\phi \end{pmatrix} \tag{2.41}$$

It is shown in Ref. 2 that the identity (2.7) generates the iterations

$$\tilde{H}_\pm^{(n+1)} - 4k^2\tilde{H}_\pm^{(n)} = \frac{\partial\tilde{F}_\pm^{(n)}}{\partial x} + ((V'^T - i\sigma_3 k)\tilde{F}_\pm^{(n)} + \tilde{F}_\pm^{(n)}(V - i\sigma_3 k)), \tag{2.42}$$

where[12]

$$\tilde{F}_-^{(0)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{2.43}$$

$$\tilde{F}_+^{(0)} = \begin{pmatrix} -2 & -2ik - \int_x^\infty dy D(y) \\ -2ik + \int_x^\infty dy D(y) & S(x) \end{pmatrix} \tag{2.44}$$

(with $\phi_+^{(0)} = 1$),

and

$$\tilde{H}_\pm^{(n)} = \tfrac{1}{2}(1 - \sigma_3)\frac{\partial\phi_\pm^{(n)}}{\partial x} \quad (\text{with } H_\pm^{(0)} = 0), \tag{2.45}$$

$$\frac{\partial\phi_\pm^{(n+1)}}{\partial x} = -\frac{\partial^3\phi_\pm^{(n)}}{\partial x^3} + 2S(x)\frac{\partial\phi_\pm^{(n)}}{\partial x} + \frac{\partial S}{\partial x}\phi_\pm^{(n)} + D(x)\int_x^\infty D(y)\phi_\pm^{(n)}(y)\,dy. \tag{2.46}$$

Equations (2.43) and (2.44) determine $\tilde{H}_\pm^{(1)}$ in Eqs. (2.42). Equations (2.45), (2.46), and (2.41) determine

$\tilde{F}_{\pm}^{(1)}$ and $\phi_{\pm}^{(1)}$. The iteration can now be repeated. Similar to Eq. (2.23) in the Zakharov–Shabat problem, one has the Wronskian identities

$$\int_{-\infty}^{\infty} dx \Phi'^{T} \tilde{H}_{\pm}^{(n)} \Phi = \sum_{j=0}^{n-1} (4k^2)^{n-j-1} \tilde{B}_{\pm}^{(j)}, \quad n = 1, 2, \ldots . \tag{2.47}$$

where

$$\tilde{B}_{+}^{(0)} = -4k^2 \begin{pmatrix} R + R' & 1 + \bar{R}R' - \bar{T}T' \\ 1 + \bar{R}'R - \bar{T}'T & \bar{R}' + \bar{R} \end{pmatrix} + 2ik \int_{-\infty}^{\infty} D(y) dy \begin{pmatrix} 0 & -\bar{T}T' \\ \bar{T}'T & 0 \end{pmatrix}, \tag{2.48}$$

$$\tilde{B}_{-}^{(0)} = 2ik \begin{pmatrix} R - R' & 1 - \bar{R}R' - \bar{T}T' \\ -1 + \bar{R}'R + \bar{T}'T & \bar{R}' - \bar{R} \end{pmatrix}, \tag{2.49}$$

$$\tilde{B}_{\pm}^{(j)} = 4k^2 \phi_{\pm}^{(j)}(-\infty) \begin{pmatrix} 0 & \bar{T}T' \\ \bar{T}'T & 0 \end{pmatrix} - 2ik \int_{-\infty}^{\infty} D(y) \phi_{\pm}^{(j)}(y) dy \begin{pmatrix} 0 & \bar{T}T' \\ -\bar{T}'T & 0 \end{pmatrix}. \tag{2.50}$$

In the limit

$$q' \to q + q_t dt, \tag{2.51}$$

$$\tilde{H}_{+}^{(n)} \to \tilde{h}_{+}^{(n)}, \tag{2.52}$$

with

$$\tilde{h}_{+}^{(1)} = \frac{\partial q}{\partial x}(1 - \sigma_3); \tag{2.53}$$

$$\tilde{H}_{-}^{(n)} \to \tilde{h}_{-}^{(n)} dt \tag{2.54}$$

with

$$\tilde{h}_{-}^{(1)} = \tfrac{1}{2}(\sigma_3 - 1) q_t, \tag{2.55}$$

and

$$\tilde{h}_{\pm}^{(n)} = \tfrac{1}{2}(1 - \sigma_3) \frac{\partial \tilde{\phi}_{\pm}^{(n)}}{\partial x}, \tag{2.56}$$

with

$$\frac{\partial \tilde{\phi}_{\pm}^{(n+1)}}{\partial x} = -\frac{\partial^3 \tilde{\phi}_{\pm}^{(n)}}{\partial x^3} + 4q \frac{\partial \tilde{\phi}_{\pm}^{(n)}}{\partial x} + 2 \frac{\partial q}{\partial x} \tilde{\phi}_{\pm}^{(n)}. \tag{2.57}$$

In this limit, for constants $\tilde{C}_j^{(\pm)}$, the nonlinear evolution equation

$$\sum_{l=1}^{n} \tilde{C}_l^{(-)} \tilde{h}_{-}^{(l)} + \sum_{m=1}^{N} \tilde{C}_m^{(+)} \tilde{h}_{+}^{(m)} = 0 \tag{2.58}$$

and Eq. (2.47) implies that[11]

$$T_t = \bar{T}_t = 0, \tag{2.59}$$

$$R_t = \omega(k) R, \tag{2.60}$$

and

$$\bar{R}_t = -\omega(k) \bar{R}, \tag{2.61}$$

where

$$\omega(k) = i \sum_{m=1}^{N} \tilde{C}_m^{(+)} (4k^2)^m \Big/ \left[k \sum_{l=1}^{n} \tilde{C}_l^{(-)} (4k^2)^{l-1} \right]. \tag{2.62}$$

Bäcklund transformations are obtained from linear combinations of $H_{\pm}^{(l)}$:

$$\sum_{l=1}^{n'} \tilde{p}_l^{(-)} \tilde{H}_{-}^{(l)} + \sum_{m=1}^{N'} \tilde{p}_m^{(+)} \tilde{H}_{+}^{(m)} = 0, \tag{2.63}$$

which, according to Eq. (2.47), implies that

$$R' = \frac{2ik\tilde{f}(k) - \tilde{g}(k)}{2ik\tilde{f}(k) + \tilde{g}(k)} R, \tag{2.64}$$

$$\bar{R}' = \frac{2ik\tilde{f}(k) + \tilde{g}(k)}{2ik\tilde{f}(k) - \tilde{g}(k)} \bar{R}, \tag{2.65}$$

where

$$\tilde{f}(k) = \sum_{l=1}^{n'} \tilde{p}_l^{(-)} (4k^2)^{l-1}, \tag{2.66}$$

$$\tilde{g}(k) = \sum_{m=1}^{N'} \tilde{p}_m^{(+)} (4k^2)^m. \tag{2.67}$$

In addition,

$$T' = \left[\frac{2ik\tilde{f}(k) - \tilde{g}(k)}{2ik\tilde{f}(k) + \tilde{g}(k)} \right] \frac{\tilde{\theta}^{(-)}}{\tilde{\theta}^{(+)}} T, \tag{2.68}$$

$$\bar{T}' = \left[\frac{2ik\tilde{f}(k) + \tilde{g}(k)}{2ik\tilde{f}(k) - \tilde{g}(k)} \right] \frac{\tilde{\theta}^{(+)}}{\tilde{\theta}^{(-)}} \bar{T}, \tag{2.69}$$

where

$$\tilde{\theta}^{(+)}(k) = \sum_{l=1}^{n'} \tilde{p}_l^{(-)} \sum_{j=1}^{l-1} (4k^2)^{l-j-1} J_{-}^{(j)}(k) + \sum_{m=1}^{N'} \tilde{p}_m^{(+)} \left[\sum_{j=1}^{m-1} (4k^2)^{m-j-1} J_{+}^{(j)}(k) - (4k^2)^{m-1}(2ik) \times \int_{-\infty}^{\infty} D(y)\, dy \right], \tag{2.70}$$

and

$$\tilde{\theta}^{(-)}(k) = \tilde{\theta}^{(+)}(-k), \tag{2.71}$$

with

$$J_{\pm}^{(j)}(k) = 4k^2 \phi_{\pm}^{(j)}(-\infty) - 2ik \int_{-\infty}^{\infty} D(y) \phi_{\pm}^{(j)}(y)\, dy. \tag{2.72}$$

III. TIME EVOLUTION OF EIGENFUNCTIONS

We can now find the time evolution of the solution of Eq. (2.1) corresponding to the Zakharov–Shabat and the Schrödinger problem, given the time evolution of the potentials, i.e., Eqs. (2.28) and (2.58).

A. Zakharov–Shabat problem

We first note that as

$$V' \to V + V_t\, dt, \tag{3.1}$$

$F_{\pm}^{(n)}$ becomes

$$F_{+}^{(n)} \to f_{+}^{(n)} + O(dt), \quad n = 1, 2, \ldots, \tag{3.2}$$

with

$$f_{+}^{(0)} = F_{+}^{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{3.3}$$

and

$$F_-^{(n)} \to f_-^{(n)}\,dt, \quad n=1,2,\ldots, \tag{3.4}$$

But

$$f_-^{(0)} = F_-^{(0)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.5}$$

Furthermore, Eq. (2.20) becomes

$$h_\pm^{(n+1)} - ikh_\pm^{(n)} = \frac{\partial f_\pm^{(n)}}{\partial x} + (V^T - i\sigma_3 k)f_\pm^{(n)} + f_\pm^{(n)}(V - i\sigma_3 k), \tag{3.6}$$

except for $h_-^{(1)}$, which satisfies

$$h_-^{(1)} = -F_-^{(0)}V_t\ . \tag{3.7}$$

We also note that Eqs. (2.21), (2.22) imply that $f_\pm^{(n)}$, $n \geq 1$, are symmetric.

In the following, we shall prove that if the solution ψ of Eq. (2.1) evolves according to

$$\left[\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}\right]\psi_t = -F_-^{(0)}\left[\sum_{l=1}^{n} C_l^{(-)} \sum_{j=1}^{l-1}(ik)^{l-j-1}f_-^{(j)} + \sum_{m=1}^{N} C_m^{(+)} \sum_{j=0}^{m-1}(ik)^{m-j-1}f_+^{(j)}\right]\psi, \tag{3.8}$$

then the isospectral flow of the potentials in Eq. (2.1) is

$$\sum_{l=1}^{n} C_l^{(-)}h_-^{(l)} + \sum_{m=1}^{N} C_m^{(+)}h_+^{(m)} = 0. \tag{3.9}$$

The proof follows the standard procedure. First we differentiate Eq. (3.8) with respect to x and use Eq. (2.1) to obtain

$$\left[\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}\right]\frac{\partial^2\psi}{\partial t\,\partial x} = -F_-^{(0)}\left\{\sum_{l=1}^{n} C_l^{(-)} \sum_{j=1}^{l-1}(ik)^{l-j-1} \times\left[\frac{\partial f_-^{(j)}}{\partial x} + f_-^{(j)}(V - i\sigma_3 k)\right] + \sum_{m=1}^{N} C_m^{(+)} \sum_{j=0}^{m-1}(ik)^{m-j-1} \times\left[\frac{\partial f_+^{(j)}}{\partial x} + f_+^{(j)}(V - i\sigma_3 k\right]\right\}\psi. \tag{3.10}$$

By Eq. (3.6), Eq. (3.10) can be written as

$$\left[\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}\right]\frac{\partial^2\psi}{\partial t\,\partial x} = -F_-^{(0)}\left\{\sum_{l=1}^{n} C_l^{(-)} \sum_{j=1}^{l-1}(ik)^{l-j-1} \times[h_-^{(j+1)} - ikh_-^{(j)} + (i\sigma_3 k - V^T)f_-^{(j)}] + \sum_{m=1}^{N} C_m^{(+)} \sum_{j=0}^{m-1}(ik)^{m-j-1} \times[h_+^{(j+1)} - ikh_+^{(j)} + (i\sigma_3 k - V^T)f_+^{(j)}]\right\}\psi. \tag{3.11}$$

Next, we differentiate Eq. (2.1) with respect to t, multiplying the result by $\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}$ and use Eq. (3.8) to obtain

$$\left[\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}\right]\frac{\partial^2\psi}{\partial x\,\partial t} = -F_-^{(0)}\left\{\sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}F_-^{(0)}(V_t - i\sigma_3 k_t) - (V^T - i\sigma_3 k)\left[\sum_{l=1}^{n} C_l^{(-)} \sum_{j=1}^{l-1}(ik)^{l-j-1}f_-^{(j)} + \sum_{m=1}^{N} C_m^{(+)} \sum_{j=0}^{m-1}(ik)^{m-j-1}f_+^{(j)}\right]\right\}\psi. \tag{3.12}$$

Equating the right-hand sides of Eqs. (3.11) and (3.12), using Eq. (3.7), and rearranging, we obtain

$$\sum_{l=1}^{n} C_l^{(-)} \sum_{j=0}^{l-1}(ik)^{l-j-1}[h_-^{(j+1)} - ikh_-^{(j)}] + \sum_{m=1}^{N} C_m^{(+)} \sum_{j=0}^{m-1}(ik)^{m-j-1}[h_+^{(j+1)} - ikh_+^{(j)}] = i\sigma_1 k_t \sum_{l=1}^{n} C_l^{(-)}(ik)^{l-1}. \tag{3.13}$$

The isospectral flow equation is Eq. (3.13) with $k_t = 0$, which is Eq. (3.9).

As an example, let $C_i^{(-)} = 0$ for $i > 2$, $C_j^{(+)} = 0$ for $j > 4$. The nonlinear evolution equation is

$$C_2^{(-)}\begin{pmatrix} \frac{1}{2}r_{xt} + r\int_x^\infty (rq)_t\,dx' \\ \frac{1}{2}q_{xt} + q\int_x^\infty (rq)_t\,dx' \end{pmatrix} + C_1^{(-)}\begin{pmatrix} r_t \\ -q_t \end{pmatrix} + C_1^{(+)}\begin{pmatrix} 2r \\ 2q \end{pmatrix} + C_2^{(+)}\begin{pmatrix} r_x \\ -q_x \end{pmatrix} + \tfrac{1}{2}C_3^{(+)}\begin{pmatrix} r_{xx} - 2r^2q \\ q_{xx} - 2q^2r \end{pmatrix} + \tfrac{1}{4}C_4^{(+)}\begin{pmatrix} r_{xxx} - 6rqr_x \\ -q_{xxx} + 6rqq_x \end{pmatrix} = 0, \tag{3.14}$$

where subscripts x and t stands for partial derivatives. The time evolution of ψ is

$$(C_2^{(-)}ik + C_1^{(-)})\dot{\psi}(x,k,t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\left\{C_2^{(-)}\begin{pmatrix} \frac{1}{2}r_t & \frac{1}{2}\int_x^\infty (qr)_t\,dx' \\ \frac{1}{2}\int_x^\infty (qr)_t\,dx' & \frac{1}{2}q_t \end{pmatrix} + \sum_{l=0}^{3}(ik)^l C_{l+1}^{(+)}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sum_{l=0}^{2}(ik)^l C_{l+2}^{(+)}\begin{pmatrix} r & 0 \\ 0 & -q \end{pmatrix} + \tfrac{1}{2}(C_3^{(+)} + ikC_4^{(+)})\begin{pmatrix} r_x & -rq \\ -rq & q_x \end{pmatrix} + \tfrac{1}{4}C_4^{(+)}\begin{pmatrix} r_{xx} - 2r^2q & rq_x - qr_x \\ rq_x - qr_x & -q_{xx} + 2q^2r \end{pmatrix}\right\}\psi(x,k,t). \tag{3.15}$$

B. Schrödinger problem

Again, as $q' \to q + q_t\,dt$

$$\tilde{F}_+^{(n)} - \tilde{f}_+^{(n)}, \tag{3.16}$$

$$\tilde{f}_+^{(0)} = 2\begin{pmatrix} -1 & -ik \\ -ik & q \end{pmatrix}, \tag{3.17}$$

and

$$\tilde{F}_-^{(n)} - \tilde{f}_-^{(n)}\, dt \tag{3.18}$$

except for

$$\tilde{f}_-^{(0)} = \tilde{F}_-^{(0)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.19}$$

Again, Eq. (2.42) holds with $\tilde{H}_\pm^{(n)}$ replaced by $\tilde{h}_\pm^{(n)}$, and $\tilde{F}_\pm^{(n)}$ replaced by $f_\pm^{(n)}$, except for $h_-^{(1)}$, which satisfies

$$\tilde{h}_-^{(1)} = -\tilde{F}_-^{(0)} V_t. \tag{3.20}$$

We may now state that if $r = 1$ and the evolution in time of the solution $\tilde{\psi}$ of Eq. (2.1) is

$$\sum_{l=1}^{n} \tilde{C}_l^{(-)} (4k^2)^{l-1} \tilde{\psi}_t$$

$$= -\tilde{F}_-^{(0)} \left[\sum_{l=1}^{n} \tilde{C}_l^{(-)} \sum_{j=1}^{l-1} (4k^2)^{l-j-1} \tilde{f}_-^{(j)} + \sum_{m=1}^{N} \tilde{C}_m^{(+)} \sum_{j=0}^{m-1} (4k^2)^{m-j-1} \tilde{f}_+^{(j)} \right] \tilde{\psi}, \tag{3.21}$$

then the isospectral flow of q of Eq. (2.1) is Eq. (2.58). The proof goes in exactly the same way as before. The only difference is that (ik) is replaced by $(4k^2)$. As an example, let $C_i^{(+)} = C_i^{(-)} = 0$ for $i > 2$; then the equation of motion is

$$C_2^{(-)}(q_{xxt} - 4qq_t + 2q_x \int_x^\infty q_t\, dx') - C_1^{(-)} q_t + 2C_1^{(+)} q_x + 2C_2^{(+)}(6qq_x - q_{xxx}) = 0, \tag{3.22}$$

and the time evolution of ϕ is

$$(C_1^{(-)} + 4k^2 C_2^{(-)})\phi_t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \times \left\{ C_2^{(-)} \begin{pmatrix} -2\int_x^\infty q_t\, dx' & -q_t - 2ik\int_x^\infty q_t\, dx' \\ -q_t - 2ik\int_x^\infty q_t\, dx' & q_{xt} - 2ikq_t + 2q\int_x^\infty q_t\, dx' \end{pmatrix} + 2C_1^{(+)} \begin{pmatrix} ik & 2q \\ 1 & ik \end{pmatrix} + C_2^{(+)} \left[4k^2 \begin{pmatrix} ik & 2q \\ 1 & ik \end{pmatrix} + \begin{pmatrix} 2q_x - 4ikq & -2q_{xx} + 4ikq_x + 4q^2 \\ 4q & -2q_x + 4ikq \end{pmatrix} \right] \right\} \phi. \tag{3.23}$$

Remarks: A consequence of Eqs. (3.8) and (3.21) is that the time dependences of the bound-state constants are the same as those of the refelction coefficients, namely Eq. (2.30) and Eqs. (2.60)—(2.61). To derive this fact from the Wronskian identities without excessive restrictions on the potentials, one must derive the Wronskian identities for the appropriate derivatives of the Jost solutions with respect to k (see the Appendix).

IV. ON BÄCKLUND TRANSFORMATIONS

It was clear from the derivations in Refs. 3, 4 that the Bäcklund transformations (2.31) [or (2.63)] makes no reference as to what evolution equations the potentials (r', q') and (r, q) satisfy, and hence cannot guarantee that they evolve via the same equation. A remarkable result of Refs. 3, 4 is that the reflection coefficients have the same time dependence [as manifested in Eq. (2.32) or (2.64) and (2.65)] once Eq. (2.31) [or Eq. (2.63)] is satisfied with constant $p_m^{(\pm)}$ [or $\tilde{p}_m^{(\pm)}$]. However, it should also be noted that if the primed and unprimed potentials satisfy the same nonlinear equations, the transmission coefficients must also have the same time dependence, i.e., they must both be constants in the case of isospectral flows, as implied by Eq. (2.29) [or (2.59)].[13] The relations between the transmission coefficients are actually Eq. (2.33) [or Eqs. (2.68) and (2.69)], which contains $\theta^{(\pm)}$ [or $\tilde{\theta}^{(\pm)}$], whose time dependences are unknown. We must, therefore, *impose the same nonlinear equation* (2.28) [or (2.58)], on both sets of potentials, and thus guarantee that $\theta^{(\pm)}$ (or $\tilde{\theta}^{(\pm)}$) are "mutual constants of motion."[14] A look at the first few $\theta^{(\pm)}$'s shows that they are differences or products of constants of motion for the individual pair of potentials. For example, for the Zakharov—Shabat problem.

$$F_{1+}^{(1)} = F_{2-}^{(1)} = 0, \tag{4.1}$$

$$F_{2+}^{(1)}(-\infty) = -F_{1-}^{(1)}(-\infty) = \tfrac{1}{2}\int_{-\infty}^{\infty} d\xi (rq - r'q'); \tag{4.2}$$

$$F_{1+}^{(2)}(-\infty) = -F_{2-}^{(2)}(-\infty) = [F_{1-}^{(1)}(-\infty)]^2, \tag{4.3}$$

$$F_{2+}^{(2)}(-\infty) = -F_{1-}^{(2)}(-\infty) = \tfrac{1}{4}\int_{-\infty}^{\infty} d\xi \times [(r'q_t' - q'r_t') - (rq_t - qr_t)]. \tag{4.4}$$

These are easily recognized as constants of motion if the isospectral flow equations are imposed.

The nature of the temporal half of the Bäcklund transformation is clear; it must be such that together with the spatial half they imply that both sets of potentials simultaneously satisfy the same nonlinear equation,[15] thus forcing $\theta^{(\pm)}$ [or $\tilde{\theta}^{(\pm)}$] to be constants. The simplest choice, but not the only choice, is to differentiate Eq. (2.31) [or (2.63)] once with respect to time t, apply the appropriate isospectral flow equation for both sets of potentials, and integrate once with respect to x. We shall illustrate this procedure by simple examples, namely, the sine—Gordon and the Korteweg—de Vries equations and restrict ourselves to only linear and quadratic Bäcklund transformations. These should demonstrate the general idea.

First, consider the Korteweg—de Vries equation. For the linear Bäcklund transformation, $\tilde{p}_i^{(\pm)} = 0$ for $i > 1$, $\tilde{p}_1^{(+)} = 1$, and we have

$$q_x + q_x' + (q - q')\int_x^\infty dy (q - q') + \tilde{p}_1^{(-)}(q - q') = 0. \tag{4.5}$$

Differentiating with respect to time and using the fact that both q and q' satisfy

$$q_t = 3(q^2)_x - q_{xxx}, \tag{4.6}$$

and integrate with respect to x, we obtain

$$q_t + q_t' = (3q'^2 - q_{xx}' - 3q^2 + q_{xx})[\tilde{p}_1^{(-)} + \int_x^\infty (q - q')\,dy]. \tag{4.7}$$

Equations (4.5) and (4.7) are the two halves of the Bäcklund transformation. Following Ref. 3, if we let

$$q = -\omega_x, \tag{4.8}$$

$$q' = -\omega_x', \tag{4.9}$$

and choose the constant of integration appropriately, we get

$$\omega_x + \omega_x' = \tfrac{1}{2}[\tilde{p}_1^{(-)2} - (\omega - \omega')^2] \tag{4.10}$$

and

$$\omega_t + \omega_t' = \tilde{p}_1^{(-)2}(\omega_x - \omega_x') + 2(\omega - \omega')\omega_{xx} - 2\omega_x^2 - 2\omega_x\omega_x'. \tag{4.11}$$

These results are equivalent to those of Ref. 16.

The next simplest Bäcklund transformation is obtained by letting $\tilde{p}_i^{(-)} = 0$ for $i > 2$, $\tilde{p}_i^{(+)} = 0$ for $i > 1$, and $\tilde{p}_2^{(-)} = 1$. We have

$$(q_{xx} - q_{xx}') - 2(q^2 - q'^2) + (q_x + q_x')(\omega - \omega') - \tfrac{1}{2}(q - q')(\omega - \omega')^2 + \tilde{p}_1^{(-)}(q - q') + \tilde{p}_1^{(+)}[q_x - q_x' + (q - q')(\omega - \omega')] = 0, \tag{4.12}$$

where

$$\omega = \int_x^\infty q\,dy, \quad \omega' = \int_x^\infty q'\,dy. \tag{4.13}$$

Similar to Eq. (4.7), we have

$$q_{xt} - q_{xt}' - [4(q^3 - q'^3) + q_x^2 - q_x'^2 - 2(q_{xx} - q'q_{xx}')] + (q_t + q_t')(\tilde{p}_1^{(+)} + \omega - \omega') + (\omega_t - \omega_t')[(q + q') + \tfrac{1}{2}(\omega - \omega')^2 - \tilde{p}_1^{(+)}(\omega - \omega') - \tilde{p}_1^{(-)}] = 0. \tag{4.14}$$

Equation (4.14) may be considered as the temporal half of the quadratic Bäcklund transformation.

As a second example, we consider the sine–Gordon equation, which is a special case of Eq. (3.14) with $C_1^{(-)} = C_1^{(+)} = 0$, and $q = -r$, i.e.,

$$r_{xt} = 2rW_t, \tag{4.15}$$

where

$$W = \int_x^\infty r^2\,d\xi. \tag{4.16}$$

The usual form of the sine–Gordon equation is obtained by letting

$$r = -u_x; \tag{4.17}$$

and then

$$u_{xt} = \sin u \tag{4.18}$$

and

$$W_t = \tfrac{1}{2}\cos u. \tag{4.19}$$

The spatial half of the linear Bäcklund transformation is[17]

$$p_1^{(-)}(r' - r) + (r + r')_x - (r - r')(W - W') = 0, \tag{4.20}$$

or

$$u_x' + u_x = 2p_1^{(-)}\sin[(u' - u)/2)]. \tag{4.21}$$

Following the prescribed procedure, we get the temporal half of the Bäcklund transformation:

$$(r_t' - r_t)(p_1^{(-)} + W - W') + (r + r')(W_t + W_t') = 0, \tag{4.22}$$

or, equivalently

$$(u_t' - u_t) = (2/p_1^{(-)})\sin[(u + u')/2]. \tag{4.23}$$

These results are well known in the literature.[18]

For the quadratic Bäcklund transformation, we may take $p_3^{(+)} = 1$, $p_2^{(-)} = \tfrac{1}{2}i(\kappa^{(+)} + \kappa^{(-)})$, $p_1^{(+)} = -\tfrac{1}{4}\kappa^{(+)}\kappa^{(-)}$, and all other $p_i^{(\pm)} = 0$.[17] Equation (2.31) becomes

$$\tfrac{1}{4}[(r_{xx} + r_{xx}') + (r_x - r_x')(W' - W) + (r - r')^2(r + r') - i\kappa^{(+)}(r_x - r_x')] + \tfrac{1}{2}(r + r') \times \{\tfrac{1}{4}[(r + r')^2 + (W' - W)^2] - \tfrac{1}{2}i\kappa^{(+)}(W' - W)\} - \tfrac{1}{4}i\kappa^{(-)} \times [(r_x - r_x') + (r + r')(W' - W) - i\kappa^{(+)}(r + r')] = 0, \tag{4.24}$$

and the temporal half of the Bäcklund transformation is

$$(r_t + r_t')\{\tfrac{1}{2}[(r + r')^2 + (W' - W)^2] - i(\kappa^{(+)} + \kappa^{(-)})(W' - W) - \kappa^{(+)}\kappa^{(-)} + (W_t + W_t')[(r_x + r_x') + (r - r')(W' - W) + i(\kappa^{(+)} + \kappa^{(-)})(r - r')] = 0. \tag{4.25}$$

We conclude this section by showing explicitly how the quadratic Bäcklund transformation (4.24) may be considered as composed of two successive linear Bäcklund transformations of the form (4.20) or its equivalent. For this purpose, we introduce the following notations:

$$R_\pm = r \pm r', \tag{4.26}$$

$$\Omega = W' - W, \tag{4.27}$$

$$\alpha_\pm = \tfrac{1}{4}(R_\pm^2 + \Omega^2) - \tfrac{1}{2}i\kappa^{(+)}\Omega, \tag{4.28}$$

and

$$K_\pm = R_\pm^{-1}\alpha_{\pm x}. \tag{4.29}$$

Then $K_\pm = 0$, or equivalently $\alpha_\pm = 0$, are linear Bäcklund transformations as Eq. (4.20). It is clear that Eq. (4.24) can be written as

$$\tfrac{1}{2}[K_{+x} + R_+\alpha_+ - i\kappa^{(-)}K_-] = 0. \tag{4.30}$$

For some r_1, let

$$R_{1\pm} = r \pm r_1, \tag{4.31}$$

$$R_{1\pm}' = r_1 \pm r', \tag{4.32}$$

$$\Omega_1 = W_1 - W, \tag{4.33}$$

$$\Omega_1' = W' - W_1, \tag{4.34}$$

$$\alpha_{\pm 1} = \tfrac{1}{4}(R_{1\pm}^2 + \Omega_1^2) - \tfrac{1}{2} i\kappa^{(\pm)}\Omega_1, \tag{4.35}$$

$$\alpha_{1\pm}' = \tfrac{1}{4}(R_{1\pm}'^2 + \Omega_1'^2) - \tfrac{1}{2} i\kappa^{(-)}\Omega'_1, \tag{4.36}$$

$$K_{1\pm} = R_{1\pm}^{-1}\alpha_{1\pm_x}, \tag{4.37}$$

and

$$K_{1\pm}' = R_{1\pm}'^{-1}\alpha_{1\pm_x}'. \tag{4.38}$$

It can be easily shown that Eq. (4.30) can be written as

$$(K_{1+_x} - K_{1-_x}') + (R_{1-} + R_{1+}')(\alpha_{1+} + \alpha_{1-}')$$

$$- (i\kappa^{(-)}K_{1+} + i\kappa^{(+)}K_{1-}') + \Omega_1' K_{1+} + \Omega_1 K_{1-}' = 0. \tag{4.39}$$

Therefore, if $\alpha_{1+} = 0$ is a linear Bäcklund transformation from r to r_1 with constant $i\kappa^{(+)}$ and $\alpha'_{1-} = 0$ is one from r_1 to r' with constant $i\kappa^{(-)}$, then Eq. (4.39) is automatically satisfied.

Similarly, let us define

$$\beta_\pm = \tfrac{1}{4}(R_\pm^2 + \Omega^2) - \tfrac{1}{2} i\kappa^{(-)}\Omega, \tag{4.40}$$

$$J_\pm = R_\pm^{-1}\beta_{\pm x}. \tag{4.41}$$

Then Eq. (4.30) can also be written as

$$\tfrac{1}{2}[J_{+x} + R_+\beta_+ - i\kappa^{(+)}J_-] = 0. \tag{4.42}$$

Again, for some r_2, let

$$R_{2\pm} = r \pm r_2, \tag{4.43}$$

$$R_{2\pm}' = r_2 \pm r', \tag{4.44}$$

$$J_{2\pm} = R_{2\pm}^{-1}\beta_{2\pm_x} = \tfrac{1}{2}[R_{2\pm_x} + R_{2\pm}\Omega_2 - i\kappa^{(-)}R_{2\pm}], \tag{4.45}$$

$$J_{2\pm}' = R_{2\pm}'^{-1}\beta_{2\pm_x}' = \tfrac{1}{2}[R_{2\pm_x}' + R_{2\pm}'\Omega_2' - i\kappa^{(+)}R_{2\pm}']; \tag{4.46}$$

then Eq. (4.42) can be written as

$$(J_{2+_x} - J_{2-_x}') + (R_{2-} + R_{2+}')(\beta_{2+} + \beta_{2-}') - (i\kappa^{(+)}J_{2+} + i\kappa^{(-)}J_{2-}')$$

$$+ \Omega_2' J_{2+} + \Omega_2 J_{2-}' = 0, \tag{4.47}$$

so that it is also equivalent to a Bäcklund transformation, $\beta_{2+} = 0$, from r to r_2 with constant $i\kappa^{(-)}$, succeeded by another, $\beta_{2-}' = 0$, from r_2 to r' with constant $i\kappa^{(+)}$. Equations (4.39) and (4.47) demonstrate the validity of the permutability theorem. We note that when $\kappa^{(-)} = -\kappa^{(+)*}$ and r, r' are real, both r_1 and r_2 are complex, and are not solutions from the inverse-scattering theory of the sine–Gordon equation.

A consequence of the permutability theorem is that Eq. (4.24) can be replaced by either $\alpha_{1+} = 0$ and $\alpha_{1-}' = 0$ or $\beta_{2+} = 0$ and $\beta_{2-}' = 0$. If we express the latter four equations in the form of Eq. (4.21) and eliminate u_x', u_x, u_{1x}, u_{2x}, we have the well-known[18] result relating u, u_1, u_2, and u':

$$\tan\left(\frac{u' + u}{4}\right) = \frac{\kappa^{(+)} + \kappa^{(-)}}{\kappa^{(+)} - \kappa^{(-)}} \tan\left(\frac{u_1 - u_2}{4}\right). \tag{4.48}$$

Knowing three solutions u, u_1, and u_2, Eq. (4.48) enables us to calculate a fourth solution, u', algebraically. This is a remarkable simplification, considering the complexity of the original quadratic Bäcklund transformation (4.24).

V. SUMMARY

We have shown that there is a direct connection between the Wronskian technique and the Lax theorem. Given the evolutionary equation (3.9) or (2.58), the time evolution of the eigenfunction of the linear operator is given by Eq. (3.8) or (3.21), respectively.

In order to see the complete spectral meaning of the Bäcklund transformation, it is necessary to look at both the reflection and transmission coefficients. It is then seen that in order that the transmission coefficients of the two sets of potentials be constants simultaneously, the same nonlinear evolutionary equation has to be imposed on both sets of potentials. The temporal half of the Bäcklund transformation serves to identify the evolutionary equation and its effect is to force $\theta^{(\pm)}$ (or $\tilde{\theta}^{(\pm)}$), defined by (2.36) [or (2.70)], to be "mutual constants of motion."

We have also shown that the quadratic Bäcklund transformation can be regarded as two successive linear Bäcklund transformations, and given an explicit proof of the permutability theorem.

APPENDIX

We shall derive an identity for the discrete spectrum of the Zakharov–Shabat problem.

As is well known, the Jost solution defined by

$$\psi(k, x) \underset{x\to-\infty}{\rightarrow} \begin{pmatrix} \exp(-ikx) \\ 0 \end{pmatrix}$$

$$\underset{x\to+\infty}{\rightarrow} \begin{pmatrix} a(k)\exp(ikx) \\ b(k)\exp(ikx) \end{pmatrix} \tag{A1}$$

can be analytically continued into the upper half plane. Let $\psi'(k, x)$ and $\psi(k, x)$ be solutions of Eq. (2.1) corresponding to two different potentials, V' and V, respectively. Suppose, at $k = \kappa_i$, $\mathrm{Im}\kappa_i > 0$, both $\psi'(k, x)$ and $\psi(k, x)$ are simple bound states; then

$$a(\kappa_i) = a'(\kappa_i) = 0, \tag{A2}$$

but their derivatives with respect to k do not vanish. It can be easily seen that for $H_\pm^{(n)}$ defined in Eq. (2.19),

$$\int_{-\infty}^{\infty} \psi_i'^T H_\pm^{(n)} \psi_i \, dx = 0. \tag{A3}$$

Let $\dot{\psi}_i$ be the derivative of $\psi_i = \psi(\kappa_i, x)$ with respect to k at κ_i. $\dot{\psi}_i$ satisfies

$$\frac{\partial}{\partial x}\dot{\psi}_i = (V - i\sigma_3\kappa_i)\dot{\psi}_i - i\sigma_3\psi_i. \tag{A4}$$

It follows that, for the matrices $F_\pm^{(n)}$,

$$\frac{\partial}{\partial x}[\psi_i'^T F_\pm^{(n)}\dot{\psi}_i]$$

$$= \psi_i'^T[H_\pm^{(n+1)} - i\kappa_i H_\pm^{(n)}]\dot{\psi}_i - i\psi_i'^T F_\pm^{(n)}\sigma_3\psi_i. \tag{A5}$$

Similarly,

$$\frac{\partial}{\partial x}[\dot{\psi}_i^T F_\pm^{(n)}\psi_i]$$

$$= \dot{\psi}_i^T[H_\pm^{(n+1)} - i\kappa_i H_\pm^{(n)}]\psi_i - i\psi_i^T\sigma_3 F_\pm^{(n)}\psi_i. \tag{A6}$$

Summing Eqs. (A5) and (A6), integrating over x and using the definition of $H_\pm^{(n)}$ and (A3), we have

$$I_\pm^{(n+1)} = i\kappa_i I_\pm^{(n)}, \quad n \geq 1, \tag{A7}$$

where

$$I_\pm^{(n)} = \int_{-\infty}^{\infty} dx\,[\psi_i^T H_\pm^{(n)}\dot{\psi}_i + \dot{\psi}_i^T H_\pm^{(n)}\psi_i], \tag{A8}$$

and

$$I_\pm^{(1)} = C_i'\dot{a}(\kappa_i) \pm C_i\dot{a}'(\kappa_i), \tag{A9}$$

where C_i, C_i' are the bound-state constants defined by

$$\psi_i \underset{x\to+\infty}{\sim} \begin{pmatrix} 0 \\ C_i \exp(i\kappa_i x) \end{pmatrix}, \tag{A10}$$

and $\dot{a}$ is the derivative of a with respect to k.

Therefore, we have

$$I_\pm^{(n+1)} = (i\kappa_i)^n(\dot{a}(\kappa_i)C_i' \pm \dot{a}'(\kappa_i)C_i). \tag{A11}$$

Equations (A11), with $I_\pm^{(n)}$ defined by Eq. (A8) is the identity sought. In the limit $V' \to V + V_t\,dt$, linear combinations of Eqs. (A11) implies that if Eq. (2.28) is the evolutionary equation for V, the time evolution of C_i is the same as $\alpha^{(+)}$, i.e., given by (2.30) with k replaced by κ_i. A similar identity can obviously be obtained for the bound-state constants, $\bar{C}_j$, in the lower half plane and imples that $\bar{C}_j$ have the same time dependence as $\alpha^{(-)}$ with k replaced by κ_j. One may also develop similar identities for the Schrödinger operator.

[1] F. Calogero, Festschrift in honor of V. Bargmann, edited by E. Lieb, B. Simon, and A.S. Wightman.
[2] F. Calogero, Lett. Nuovo Cimento **14**, 443 (1975).
[3] F. Calogero, Lett. Nuovo Cimento **14**, 537 (1975).
[4] F. Calogero and A. Degasperis, Nuovo Cimento B **32**, 201 (1976).
[5] F. Calogero, Nuovo Cimento B **29**, 509 (1975).
[6] F. Calogero and A. Degasperis, Lett. Nuovo Cimento **15**, 65 (1976).
[7] The first to visualize such a connection was M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1974).
[8] P.D. Lax, Comm. Pure Appl. Math. **21**, 467 (1968).
[9] M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Phys. Rev. Lett. **31**, 125 (1973).
[10] It follows from Eq. (2.15) that the diagonal elements of $F_\pm^{(n)}$ are defined by $H_\pm^{(n)}$.
[11] In arriving at Eq. (2.29), it is assumed that $[\Sigma_{l=1}^n C_l^{(-)}A^{l-1}]^{-1}$ exists. A consequence is the $f_{1-}^{(j)}(-\infty)$, $j=1, \ldots, n-1$, [which are the limits of $F_{1-}^{(j)}(-\infty)$ as $V' \to V + V_t dt$] vanish and can be interpreted as the time derivatives of the first $n-1$ constants of motion. A discussion on this point was given in Ref. 8, Sec. III. The same remark holds true for the Schrödinger operator.
[12] There is no $\phi_-^{(0)}$ satisfying (2.41) with $F_-^{(0)}$ given by (2.43).
[13] Note that we cannot expect complete information about the relations between the bound-state constants because (a) the two systems have different bound states, (b) one of the two systems may be singular. However, for those values of k which are bound-state eigenvalues of both systems, a relation has been derived in the Appendix.
[14] Whether this is the only thing one can do about the time evolution is unclear, but some time evolution has to be imposed, as can be seen in the case of pure solitons when the β's vanish.
[15] We shall not consider Bäcklund transformations between two different evolutions here.
[16] H.D. Wahlquist and F.B. Estabrook, Phys. Rev. Lett. **31**, 1386 (1973).
[17] As shown in Ref. 1, the symmetry of the system requires $f(ik)g(-ik) = -f(-ik)g(ik)$. For the linear Bäcklund transformation, we may take either $p_1^{(-)} \neq 0$, $p_2^{(+)} = 1$, or $p_2^{(-)} = 1$, $p_1^{(+)} \neq 0$.
[18] See, for example, G.L. Lamb, Jr., J. Math. Phys. **15**, 2157 (1974), and Ref. 1.

Spectral properties of completely monotonic operators

L. Bonora and P. Pasti

Istituto di Fisica dell' Università—Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Italy
(Received 17 February 1977)

We introduce a special class of j-dependent analytic operator valued functions, the completely monotonic operators, and find the general properties of their spectrum and of their resolvent. Next we study an important and wide subclass of such operators, for which it is possible to reach more detailed information about the leading singularity of the resolvent in the j plane, and in some cases to completely determine it.

1. INTRODUCTION

Positive operators in ordered spaces and the perturbation theory for linear operators have been widely treated in mathematical literature. Little attention has been paid, however, to the great variety of problems that arise in the perturbation theory for positive linear operators depending on a parameter. In the study of the multiperipheral operator equation[1] we had to tackle a problem of this kind: In order to solve the multiperipheral equation one has to find the resolvent of an operator $A(j)$, depending analytically on the parameter j, and having positivity properties (4.1) (a completely monotonic operator). In the present paper we forget the particular physical features of the problem and study it from a purely mathematical point of view. We believe that the results obtained or, at least, the general lines we followed may be useful in solving other physical problems involving linear operator equations.

The subject matter is arranged as follows. In Sec. 2 we introduce the necessary mathematical formalism on Banach lattices and positive operators. In the next section we quote or derive the results needed later about the resolvent of a positive operator. In Sec. 3 analytic families of operators are considered, attention being concentrated mainly on completely monotonic operator-valued functions $A(j)$: We justify their introduction and give the most general results about their resolvent $R(\lambda, A(j))$. As a main result, we identify the position of the rightmost j-plane singularity of $R(\lambda_0, A(j))$ when λ_0 is a positive real number. In the last part of the paper we study a particular family of such operators, interesting from a physical point of view, for which we can give a detailed expression of the j-plane rightmost singularity of the resolvent.

2. MATHEMATICAL PRELIMINARIES

In this section we recall some definitions and properties about ordered vector spaces, in particular Banach lattices.

Let β be a real linear vector space. A subset $\mathbb{C}$ is a cone if it has the following two properties:

(i) If x and y belong to $\mathbb{C}$, and α and β are real nonnegative numbers, then $\alpha x + \beta y \in \mathbb{C}$.

(ii) If x and $-x$ are elements of $\mathbb{C}$, then $x = 0$.

The presence of a cone in a vector space, defines a (partial) order in it. It is enough to set $x \leq y$ when $y - x \in \mathbb{C}$. The relation so defined satisfies all properties required in order for the space to be partially ordered.[2] Later, when we mention a partially ordered vector space, we shall refer to the order induced by the cone $\mathbb{C}$. In such a space we say that a functional x^* is positive if for any $x \in \mathbb{C}$, $\langle x^*, x\rangle \geq 0$. We also say that an operator A is positive if for any $x \in \mathbb{C}$, $Ax \in \mathbb{C}$.[3] We shall write $x \geq 0$ if $x \in \mathbb{C}$, $x > 0$ if $x \in \mathbb{C}$ and $x \neq 0$. Analogous notation will be used for positive functionals and operators.

In the theory of positive operators a recurrent concept is the one of reproducing cones. A cone $\mathbb{C}$ is reproducing if every element $x \in \beta$ can be written in the form $x = u - v$, where u and v belong to $\mathbb{C}$.

A partially ordered vector space is called a lattice, if for any two elements x and y of the space, we may define the supremum $w = x \vee y$ and the infimum $z = x \wedge y$. w is such that $x \leq w$ and $y \leq w$, and, if an element u exists such that $x \leq u$ and $y \leq u$, then $w \leq u$. In an analogous way we define the infimum.

In a lattice we call positive and negative parts of an element x the vectors $x_+ = x \vee 0$ and $x_- = (-x) \vee 0$. It is obvious that $x = x_+ - x_-$, where $x_+, x_- \in \mathbb{C}$. Therefore, $\mathbb{C}$ is reproducing. The modulus of x is the vector $|x| = x_+ + x_-$. We quote here some properties[4] of a lattice that will be useful later:

$$\begin{aligned}
&\text{(a) For each } x, \quad x_+ \wedge x_- = 0,\\
&\text{(b) for any two } x, y, \quad |x - y| = |x_+ - y_+| + |x_- - y_-|,\\
&\text{(c) for any two } x, y, \quad |x + y| \leq |x| + |y|, \qquad (2.1)\\
&\text{(d) if } x, y, z \in \mathbb{C}, \quad (x + y) \wedge z \leq (x \wedge z) + (y \wedge z),\\
&\qquad\qquad (x \wedge y) + z = x \wedge z + y \wedge z.
\end{aligned}$$

The next step is to introduce a topology in a partially ordered vector space or in a lattice by means of a norm $\|\cdot\|$

A very important structure is obtained when a consistency condition between norm and order is satisfied. Explicitly, if β is both a lattice and a Banach space, and if

$$|x| \leq |y| \quad \text{implies} \quad \|x\| \leq \|y\| \qquad (2\ 2)$$

for any two elements $x, y \in \beta$, then β is called a Banach lattice. This is in fact the most relevant structure in this paper, and the remainder of this section is devoted to explaining the main properties of a Banach lattice β.

From condition (2.2) one can easily deduce that if $x, y \in \mathbb{C}$ and $x \leq y$, then $\|x\| \leq \|y\|$; for any x one has $\|x\|$

$= \| |x| \|$. Conversely these two properties imply (2.2).

The cone $\mathbb{C}$ is closed. For if the sequence $\{x_n\}$ of positive elements tends to x, from Eq. (2.1b), we have

$$\|x_n - x\| = \| |x_n - x| \| = \| |x_n - x_+| + x_- \| \geq \|x_n - x_+\|,$$

where we set $x = x_+ - x_-$. It follows that $x_n \to x_+$, that is, $x_- = 0$.

We consider now the dual β^* of β. The set of linear continuous positive functionals is a cone in β^*, which we call $\mathbb{C}^*$. $\mathbb{C}^*$ is reproducing and β^* is itself a Banach lattice.[4]

Similarly the sum and the product of two linear positive operators is a positive operator, and the bounded positive operators in β form a cone in the space $L(\beta)$ of continuous operators. It is easy to show that if $\mathbb{C}$ is closed in β, this cone is closed in $L(\beta)$. Indeed, if $A_n \to A$ and all A_n are positive, then for any $x \in \mathbb{C}$ and any n, $A_n x \in \mathbb{C}$, and therefore $\lim_{n\to\infty} A_n x = Ax \in \mathbb{C}$. We shall also take into account that the transpose of a positive operator is positive.

Proposition 2.1: If A is a positive operator, then $\|A\| = \sup_{\substack{\|x\| \leq 1 \\ x \in \mathbb{C}}} \|Ax\|$.

By definition

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| \geq \sup_{\substack{\|x\| \leq 1 \\ x \in \mathbb{C}}} \|Ax\| = M.$$

In a Banach lattice, from Eq. (2.1c), we have $|Ax| = |A(x_+ - x_-)| = |Ax_+ - Ax_-| \leq Ax_+ + Ax_- = A|x|$. Thus

$$\sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\| \leq 1} \| |Ax| \| \leq \sup_{\| |x| \| \leq 1} \|A|x|\| = \sup_{\substack{\|x\| \leq 1 \\ x \in \mathbb{C}}} \|Ax\| = M$$

and the proposition follows.

In a similar way we can show that for every $x \in \mathbb{C}$, $\|x\| = \sup_{\substack{\|x^*\| \leq 1 \\ x^* \in \mathbb{C}^*}} \langle x^*, x \rangle$.

Proposition 2.2: If $x \in \mathbb{C}$, we can find a functional $x^* \in \mathbb{C}^*$ such that $\langle x^*, x \rangle = \|x\|$ and $\|x^*\| = 1$.

A well-known result of functional analysis[5] says that for any $x \in \beta$ there exists a functional $y^* \in \beta^*$, such that $\langle y^*, x \rangle = \|x\|$ and $\|y^*\| = 1$.

In our case it is enough to choose $x^* = y_+^*$, where $y^* = y_+^* - y_-^*$. Then $\|x^*\| \leq \|y_+^* + y_-^*\| = \| |y^*| \| = \|y^*\| = 1$. On the other hand, $\langle x^*, x \rangle \geq \langle y^*, x \rangle = \|x\| \geq \langle x^*, x \rangle$, so that $\langle x^*, x \rangle = \|x\|$ and $\|x^*\| = 1$.

Proposition 2.3: If $x, y \in \beta$ and $x \neq 0$, $y \neq 0$, $x \wedge y = 0$, then a functional $x^* \in \mathbb{C}^*$ exists such that $\langle x^*, x \rangle \geq \|x\|$ and $\langle x^*, y \rangle = 0$. For the length proof see Ref. 4.

Proposition 2.4: Let $x \in \beta$ be such that for any $x^* \in \mathbb{C}^*$, $\langle x^*, x \rangle \geq 0$. Then $x \in \mathbb{C}$.

Suppose that this is not true. Then $x = x_+ - x_-$ with $x_- \neq 0$. Let $x_+ = 0$. Then we apply Proposition 2.2: There is a positive functional x_1^* such that $\langle x_1^*, x_- \rangle = \|x_-\|$. Thus $\langle x_1^*, x \rangle = -\|x_-\| < 0$, against the hypothesis. Let $x_+ \neq 0$. Then we apply Proposition 2.3: As $x_+ \wedge x_- = 0$ by Eq. (2.1a), there exists a positive functional $x_2^* \in \mathbb{C}^*$, such that $\langle x_2^*, x_- \rangle \geq \|x_-\|$ and $\langle x_2^*, x_+ \rangle = 0$. Therefore, $\langle x_2^*, x \rangle = -\langle x_2, x_- \rangle < 0$, in opposition to the hypothesis.

Corollary 2.5: Let A be an operator defined in a Banach lattice such that $\langle x^*, Ax \rangle \geq 0$ for any $x \in \mathbb{C}$ and any $x^* \in \mathbb{C}^*$. Then A is a positive operator.

In the last part of this section we introduce some definitions and properties about special topics that will be reconsidered later. An element $x \in \mathbb{C}$ is quasi-interior if $\langle x^*, x \rangle > 0$ for any $x^* > 0$. An example: Let L^p be, as usual, the space of equivalence classes of real-valued functions $f(t)$ p-integrable ($1 \leq p < \infty$) with respect to the Lebesgue measure on the real axis; L^p is a Banach lattice, the positive elements being the functions $f(t) \geq 0$ almost everywhere, whereas the quasi-interior functions satisfy $f(t) > 0$ almost everywhere.

We call the operator A quasi-interior if $x > 0$ implies Ax quasi-interior. In the above example, Let A_k be the bounded operator corresponding to the kernel $K(t, t')$ acting from L^p to L^q ($1 \leq p, q < \infty$)[6]; then, if $K(t, t') > 0$ almost everywhere, A_K is a quasi-interior operator.

So far a real Banach lattice β was considered. However we also need the concept of a complex Banach lattice.[7] The complexification of β of a vector space is obtained simply by considering the space $\beta^c = \beta + i\beta$, but in order to make it a lattice we need some hypotheses on β. We shall assume β is order complete: For any subset $\mathcal{A} \subset \beta$ such that there exists an element of β greater than any element of $\mathcal{A}$, $\mathcal{A}$ has a supremum in β. Then it is possible to define a modulus of any element $z = x + iy \in \beta^c$ by means of the formula

$$|z| = \sup_{0 \leq \theta < 2\pi} |\cos\theta x + \sin\theta y|. \tag{2.3}$$

Then, extending the norm of β by means of $\|z\| = \| |z| \|$ for any $z \in \beta^c$, one may verify that the analog of Eq. (2.2) is satisfied and therefore β^c is a Banach lattice. We say that a subset $\mathcal{A} \subset \beta^c$ is order bounded if there exists some $x \in \beta$ such that $|z| \leq x$ for any $z \in \mathcal{A}$. An example of order complete space is L^p.[8] The complexification is the obvious one.

Now let A be an operator acting in β^c, or briefly, a complex operator. We can define the linear modulus of A, only at the cost of some hypotheses on A and on β. We suppose β is order complete and A is order bounded: That is, A maps order bounded sets onto order bounded sets. Then the definition

$$|A|x = \sup_{|z| \leq x} |Az| \quad \text{for any } x \in \mathbb{C} \tag{2.4}$$

makes sense, and the operator $|A|$ is called the modulus of A.[9] It satisfies the inequality

$$|Az| \leq |A|\,|z| \quad \text{for any } z \in \beta^c. \tag{2.5}$$

Finally we give an example that we shall exploit later. Let A_K be the bounded operator corresponding to the complex kernel $K(t, t')$ from a complex L^p to a complex L^q ($1 \leq p, q < \infty$).[6] Then $|K(t, t')|$ defines a bounded positive operator $A_{|K|}$ and A_K is order bounded.

Indeed, for any complex $g = g(t)$ for which $|g(t)| \leq f(t)$, where $f = f(t)$ is nonnegative almost everywhere, we

have

$$|A_K g| = \left| \int K(t,t')g(t')\,dt' \right| \leq \int |K(t,t')|\,f(t')\,dt' = A_{|K|}f. \tag{2.6}$$

From definition (2.4) we also have

$$|A_K| \leq A_{|K|}. \tag{2.7}$$

3. SPECTRAL PROPERTIES OF POSITIVE OPERATORS

We give here some general results about the spectrum and the resolvent of a positive operator. They are both useful by themselves and necessary as an introduction to the topic dealt with in the next section. For this purpose we need some definitions and general results from spectral analysis of operators. In this section A will be a bounded operator in a Banach space β^c. The resolvent of A is the operator $R(\lambda,A)=(\lambda-A)^{-1}$, if it exists and is bounded. The set $\rho(A)$ of the complex λ plane, where this condition is satisfied, is called the resolvent set of A. The complementary set of $\rho(A)$, $\sigma(A)$, is the spectrum of A. $\sigma(A)$ is a closed and bounded set.

For $\lambda, \mu \in \rho(A)$, the equation

$$R(\lambda,A) - R(\mu,A) = (\mu-\lambda)R(\lambda,A)R(\mu,A) \tag{3.1}$$

holds. If B is a bounded operator too, we have the relation

$$R(\lambda,A) - R(\lambda,B) = R(\lambda,A)(A-B)R(\lambda,B), \qquad \lambda \in \rho(A) \cap \rho(B). \tag{3.2}$$

From the first equation one gets that, in $\rho(A)$, $R(\lambda,A)$ is an analytic operator-valued function with derivatives given by the formulas

$$\frac{d^k}{d\lambda^k}R(\lambda,A) = (-1)^k k!\,R^{k+1}(\lambda,A), \quad k=1,2,\cdots. \tag{3.3}$$

Therefore, in a neighborhood of an isolated singularity λ_0, $R(\lambda,A)$ admits a Laurent series expansion

$$R(\lambda,A) = \sum_{n=0}^{\infty} (\lambda-\lambda_0)^n C_n + \sum_{n=1}^{\infty} (\lambda-\lambda_0)^{-n} D_n, \tag{3.4}$$

where C_n, D_n are bounded operators in β^c, defined by the contour integrals:

$$C_n = \frac{1}{2\pi i}\int_\gamma (\lambda-\lambda_0)^{-n-1}R(\lambda,A)\,d\lambda, \qquad D_n = \frac{1}{2\pi i}\int_\gamma (\lambda-\lambda_0)^{n-1}R(\lambda,A)\,d\lambda. \tag{3.5}$$

γ is a counterclockwise oriented circular path with center at λ_0, contained in $\rho(A)$. It is well known that $D_1 = P$ is a projection operator, that is, $P^2 = P$. Among others the following relations hold:

$$D_n = (A-\lambda_0)^{n-1}P = (A-\lambda_0)D_{n-1}, \quad n=1,2,\cdots. \tag{3.6}$$

We observe that, if $D_k = 0$, $D_n = 0$ for $n \geq k$, so we shall say that λ_0 is a pole of order k if $D_k \neq 0$ and $D_{k+1} = 0$. In this case, λ_0 is an eigenvalue of A and the manifold spanned by all corresponding eigenvectors (the eigenmanifold) is contained in the range of the projection operator P.

A useful tool in the analysis of the spectral properties of an operator is the concept of spectral radius. We recall that the spectral radius of A is the number

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}. \tag{3.7}$$

It determines the convergence region of the series $\sum_{n=0}^{\infty} \lambda^{-n-1}A^n$. Precisely this series converges in the operator norm for $|\lambda| > r(A)$. In this region the series represents the resolvent $R(\lambda,A)$. From now on we shall suppose $r(A) > 0$.

Let us now add the condition that A is a positive operator. From now on in this section we shall suppose that β, the real Banach space subtended to β^c, has a closed reproducing cone $\mathbb{C}$.

Proposition 3.1: $R(\lambda,A)$ is a positive operator for any real $\lambda > r(A)$.

Using the above-mentioned series to represent $R(\lambda,A)$, we see that the resolvent is the limit of a series of positive operators for real $\lambda > r(A)$. The proposition follows from what has been said in the preceding section and from the fact that $\mathbb{C}$ is closed.

One shows[7] that, if β is a Banach lattice, the condition (real) $\lambda > r(A)$ is necessary (besides being sufficient) in order for $R(\lambda,A)$ to be a positive operator.

Proposition 3.2: Let A and B be two positive operators in a Banach lattice β, such that $A \leq B$. Then $r(A) \leq r(B)$.

We show first, by induction, that $A^n \leq B^n$. Suppose $A^{n-1} \leq B^{n-1}$ and observe that $B-A$, $B+A$, $B^{n-1}-A^{n-1}$, $B^{n-1}+A^{n-1}$ are positive operators. Then

$$(B-A)(B^{n-1}+A^{n-1}) = B^n - A^n + BA^{n-1} - AB^{n-1}$$

and

$$(B+A)(B^{n-1}-A^{n-1}) = B^n - A^n - BA^{n-1} + AB^{n-1}$$

are positive operators. Thus also their sum $B^n - A^n$ is a positive operator. For any $x \in \mathbb{C}$ the inequality $B^n x \geq A^n x$ holds. Therefore, from Proposition 2.1, we get

$$\|A^n\| \leq \|B^n\|.$$

The proposition follows from the definition of spectral radius.

Proposition 3.3: Let the cones $\mathbb{C}$ and $\mathbb{C}^*$ be reproducing (as is the case if β is a Banach lattice), and A be positive. Then the point of the λ plane $\lambda = r(A)$ is in the spectrum of A.

The proof[10] is based essentially on the well-known result that the intersection point of the real axis and the convergence circle of a power series with nonnegative coefficients, is a singular point of the analytic function represented by the series.

This proposition does not characterize the nature of the singularity. In order to have more detailed information about the nature of the singularity in $r(A)$, one has to consider special classes of operators. If for example, A is compact, then necessarily $\lambda = r(A)$ is a pole of finite order of the resolvent $R(\lambda,A)$. For the spectrum of a compact operator is a set of isolated points (if we exclude $\lambda = 0$), each representing a pole of the resolvent and therefore being an eigenvalue of A.[3]

As far as eigenvectors corresponding to $\lambda = r(A)$ are concerned, we have the following results.

Proposition 3.4: Let A be a positive operator and $\lambda = r(A)$ be a k-order pole of $R(\lambda, A)$. Then the corresponding D_k is a positive operator.

From Eq. (3.4) it follows that $\lim_{\lambda \to r(A)} (\lambda - r(A))^k \times R(\lambda, A) = D_k$ in the operator norm topology. If we take the limit from the right on the real axis, we can verify the assertion by applying Proposition 3.1 and remembering that the cone of positive operators is closed.

Let us call A^* the transpose of A.

Proposition 3.5: In the same hypotheses of the preceding proposition, $r(A)$ is an eigenvalue both of A and of A^*, to which there correspond respectively an eigenvector $x > 0$ and an eigenvector $x^* > 0$.[10]

The fact that $r(A)$ is an eigenvalue is immediate.[5] The rest of the proof is a corollary of the preceding proposition. Indeed, as by hypotheses $D_k \neq 0$ and $D_{k+1} = 0$, some $y \in \mathfrak{C}$ exists ($\mathfrak{C}$ is reproducing!) for which $x = D_k y \neq 0$. Obviously $x \in \mathfrak{C}$ by Proposition 3.4. Now by Eq. (3.6) we have $(A - r(A))x = D_{k+1} y = 0$. So x is the desired eigenvector. From the property $R(\lambda, A^*) = R(\bar{\lambda}, A)^*$[11] if $\bar{\lambda} \in \rho(A)$, we can see that $D_k^* > 0$ and $D_{k+1}^* = 0$. Repeating the argument used above, we conclude that there exists an eigenvector $x^* > 0$.

More detailed information about spectral properties is obtained if we restrict the hypotheses on the operator A. If A is quasi-interior (see Sec. 2), $r(A)$ is a first order pole of $R(\lambda, A)$ with one-dimensional eigenmanifold[7] (when the pole is first order, the eigenmanifold coincides with the range of the corresponding projection).

Proposition 3.6: If A is quasi-interior, $r(A)$ has a quasi-interior eigenvector. Moreover $A^* x_0^* = r(A) x_0^*$, where x_0^* has the following property: For any $x > 0$, $\langle x_0^*, x \rangle > 0$.

The first statement is obvious. From Proposition 3.5 we know that there exists an eigenvector $x_0^* > 0$ of A^* corresponding to $r(A)$. Thus for any $x > 0$ we have

$$\langle x_0^*, x \rangle = \frac{1}{r(A)} \langle A^* x_0^*, x \rangle = \frac{1}{r(A)} \langle x_0^*, Ax \rangle > 0,$$

as Ax is a quasi-interior point.

Proposition 3.7: Let T be a complex operator (i.e., acting in β^c), such that $|T|$ exists. Let λ_0, which is the point of its spectrum of largest modulus, be an eigenvalue of T, and let A and $A - |T|$ be quasi-interior. Then $|\lambda_0| < r(A)$.

For some $z \in \beta^c$ we have $Tz = \lambda_0 z$, $|\lambda|\,|z| = |Tz| \leq |T|\,|z|$; so that from Proposition 3.6

$$0 < |\lambda_0| \langle x_0^*, |z| \rangle \leq \langle x_0^*, |T|\,|z| \rangle < \langle x_0^*, A|z| \rangle = r(A) \langle x_0^*, |z| \rangle.$$

Therefore, $|\lambda_0| < r(A)$.

4. COMPLETELY MONOTONIC OPERATORS

Let us consider the function $A(j)$, defined in a domain D of the complex j plane, with values in $L(\beta^c)$, which is the space of bounded operators in the complex Banach space β^c. We shall call it an operator valued analytic function if the dependence on j is analytic, that is, if $A(j)$ is differentiable in D with respect to the uniform topology of $L(\beta^c)$. It is well known that this is equivalent to the analyticity of all numerical functions $\langle x^*, A(j)x \rangle$ for $x \in \beta^c$ and $x^* \in \beta^{c*}$.

Proposition 4.1: Let $A(j)$ be an analytic operator valued function in a domain D of the j plane. Then the resolvent $R(\lambda, A(j))$ is holomorphic in the Cartesian product $D \times \lambda$ plane, save the set of pairs (λ_0, j_0) such that λ_0 belongs to the spectrum of $A(j_0)$.[11]

We shall study a particular family of operator-valued functions, that we shall call completely monotonic operator-valued functions, or, briefly, completely monotonic operators. As in the preceding section, β has a closed reproducing cone $\mathfrak{C}$. Let j_0 be a point of the real j axis. We say that the analytic operator-valued function $A(j)$ is completely monotonic in the open real semiaxis $j > j_0$, if, there, the operators

$$(-1)^k \frac{d^k A(j)}{dj^k} \geq 0 \quad \text{for} \quad k = 0, 1, 2, \cdots. \tag{4.1}$$

It is clear that for all $x \in \mathfrak{C}$ and $x^* \in \mathfrak{C}$ the numerical functions $\langle x^*, A(j)x \rangle$ are completely monotonic (see Appendix A). Conversely if β is a Banach lattice the preceding property is a sufficient condition in order for $A(j)$ to be a completely monotonic operator, on the basis of Corollary 2.5. From the previous definition it is evident that the sum and the product of two completely monotonic operators are completely monotonic operators. Of course, for j real, $A(j)$ is understood as an operator in β but it is definite also as an operator in β^c; for j complex, it must be understood as an operator in β^c.

As an exemplification, let A be a positive operator. Then by Eq. (3.3) and Proposition 3.1, it follows that $R(\lambda, A)$ is a completely monotonic operator in λ for real $\lambda > r(A)$.

From Eq. (3.2) one gets

$$\frac{dR(\lambda, A(j))}{dj} = R(\lambda, A(j)) \frac{dA(j)}{dj} R(\lambda, A(j)). \tag{4.2}$$

Therefore, if real $\lambda_0 > r(A(j))$ and $A(j)$ is completely monotonic for $j > j_0$, $R(\lambda_0, A(j))$ is completely monotonic in the subset of the real j axis in which the two inequalities are simultaneously satisfied.

We now give two general propositions involving the spectral radius $r(j) = r(A(j))$ of an operator $A(j)$ completely monotonic for $j > j_0$ acting in a Banach lattice β.

Proposition 4.2: $r(j)$ is a nonincreasing function for real $j > j_0$.

If $j_1 > j_2$, by condition (4.1), for every $x \in \mathfrak{C}$ and every $x^* \in \mathfrak{C}^*$, we have $\langle x^*, A(j_1)x \rangle \leq \langle x^*, A(j_2)x \rangle$, so that $A(j_1) \leq A(j_2)$. Then, by Proposition 3.2, $r(j_1) \leq r(j_2)$. As a matter of fact the hypothesis of complete monotony is redundant: It is enough that the above operator inequality be satisfied.

Proposition 4.3: Let $\bar{j}$ be the rightmost point of the real j axis satisfying the equation $r(j) = \lambda_0$, and $\bar{j} \geq j_0$. Then $R(\lambda_0, A(j))$ is analytic in j on the right of the axis

$\mathrm{Re}j=\bar{j}$. Moreover the singularities on this axis are not stronger than the singularity at $j=\bar{j}$, since for any $x \in \mathbb{C}$ and $x^* \in \mathbb{C}^*$ and for $\mathrm{Re}j > \bar{j}$ one has

$$|\langle x^*, R(\lambda_0, A(\mathrm{Re}j + i\,\mathrm{Im}j))x\rangle| \leq \langle x^*, R(\lambda_0, A(\mathrm{Re}j))x\rangle. \tag{4.3}$$

By Proposition 4.2 and by assumption, it follows that $r(j) < \lambda_0$ for $j > \bar{j}$. Therefore, by Proposition 4.1, $R(\lambda_0, A(j))$ is analytic and completely monotonic in $j > \bar{j}$. If $x \in \mathbb{C}$ and $x^* \in \mathbb{C}^*$, $\langle x^*, R(\lambda_0, A(j))x\rangle$ is a completely monotonic numerical function, and by the properties of completely monotonic functions (see Appendix A, Proposition A.1) we can write inequality (4.2), from which the completion of the proof follows.

In general there are other singularities besides the one at $j=\bar{j}$, on the axis $\mathrm{Re}j=\bar{j}$. In order to exclude this possibility we have to strengthen the hypotheses on $A(j)$. Let, for example, $|A(j)|$ exist, $A(j)$ be compact, $A(\mathrm{Re}j)$ and $A(\mathrm{Re}j) - |A(\mathrm{Re}j + i\,\mathrm{Im}j)|$ be quasi-interior, for $\mathrm{Re}j > j_0$. Then the hypotheses of Proposition 3.7 are fulfilled and we can conclude that the only singularity of $R(\lambda_0, A(j))$ on the axis $\mathrm{Re}j=\bar{j}$ is placed at $j=\bar{j}$. An application of this fact is found in a physical problem,[12] where $A(j)$ is a compact completely monotonic operator for $\mathrm{Re}j > j_0$ corresponding to a kernel $K(t,t',j)$ from L^p to L^q, satisfying the inequality $|K(t,t',\mathrm{Re}j + i\,\mathrm{Im}j)| < K(t,t',\mathrm{Re}j)$, and such that $K(t,t',j) > 0$ for real $j > j_0$. Bearing in mind what has been said at the end of Sec. 2, one can realize that all the above hypotheses are verified and, therefore, the previous conclusion holds.

In the following sections we shall limit ourselves to the analysis of the rightmost singularity on the real j axis and nothing will be said about other singularities on $\mathrm{Re}j=\bar{j}$, apart from the fact that the singularity at $j=\bar{j}$ is always isolated. The reason is briefly explained in Appendix A. We shall call the singularity at $j=\bar{j}$, the leading j-plane singularity.

5. LEADING SINGULARITY OF THE RESOLVENT

In this section we shall consider an operator-valued function $A(j)$, analytic in a half-plane $\mathrm{Re}j > j_0$, acting in a complex Banach space $\mathcal{B}^c$. The spectrum of $A(j)$ will vary in general in an extremely complex way as j varies. Given the lack of general mathematical results, we only analyze the case in which the rightmost singularity in the λ plane is a pole of the resolvent $R(\lambda, A(j))$. This choice is due on the one hand to the fact that this property characterizes wide classes of operators (compact and quasicompact operators, operators with compact resolvent, ···) and on the other hand to the simple manageable structure taken up in this case by the spectrum on the j plane.

As a first step we state precisely the mathematical problem we have chosen to study. Let us suppose that for $j=j_1$, the spectrum of $A(j_1)$ can be separated into two spectral sets[5] $\sigma_1 = \sigma_1(A(j_1))$ and $\sigma_2 = \sigma_2(A(j_1))$. We take a Jordan curve γ_1 surrounding σ_1 and outside σ_2 (remember that σ_1 and σ_2 are closed separated sets). One shows[11] that, in a sufficiently small neighborhood U of j_1, $\sigma_1(A(j))$ and $\sigma_2(A(j))$ are still spectral sets, while γ_1 is in the resolvent set of $A(j)$ and still separates them.

Let us consider the projection operator

$$P_{\sigma_1}(j) = \frac{1}{2\pi i}\int_{\gamma_1} R(\lambda, A(j))\,d\lambda, \tag{5.1}$$

where the integration path is anticlockwise oriented. It is an analytic operator-valued function in U. Let $\mathcal{M}_1(j) = P_{\sigma_1}(j)\mathcal{B}^c$ be the range of $P_{\sigma_1}(j)$, that is the subspace associated to the spectral set $\sigma_1(j)$. One can show[13] that in a neighborhood U_1 of j_1, $U_1 \subseteq U$, the dimension of $\mathcal{M}_1(j)$ is constant as j varies.

Let A_1 be the restriction of the operator A to the subspace $\mathcal{M}_1$, and A_2 the analogous restriction to $\mathcal{M}_2 = (1 - P_{\sigma_1})\mathcal{B}^c$ (here we understand the variable j which is inessential for the discussion). Then the spectrum of A_1 (A_2) is given by σ_1 (σ_2),[5,11] and the eigenvalue problem for A is the same as the one for A_1, as long as the eigenvalues contained in σ_1 are concerned. Moreover the resolvent of A commutes with P_{σ_1} and therefore $R(\lambda, A) = R(\lambda, AP_{\sigma_1}) + R(\lambda, AP_{\sigma_2})$. This implies that $R(\lambda, A_2)$ is regular in the region limited by γ_1 and conversely $R(\lambda, A_1)$ is regular outside it. Since we shall be interested in the structure of the singularities contained in a particular region of the λ plane, when γ_1 contains this region we can limit ourselves to the analysis of $R(\lambda, A_1)$. The true resolvent differs from this only by a regular part.

Now let us suppose that $\mathcal{M}_1$ has dimension $N < \infty$. Then $A_1(j)$ may be represented by means of a finite order matrix and the eigenvalue problem becomes the problem of the solution of an algebraic equation in λ, whose coefficients are analytic functions of j in the set U_1. The solution is known[11,14]: The roots (that is, the eigenvalues) are the branches $\lambda_k(j)$ of one or several (locally) algebraic functions, $k = 1, 2, \ldots, r$ and $r \leq N$. More precisely, the branches $\lambda_k(j)$ group together in cycles $\{\lambda_{k_1}(j)\}, \ldots, \{\lambda_{k_s}(j)\}$, $k_1 = 1, \ldots, r_1, \ldots, k_s = r_{s-1} + 1, \ldots, r$. If a cycle contains a single element $\lambda_{\bar{k}}(j)$, this is an analytic function of j in the neighborhood U_1 of j_1. Instead, if a cycle contains p (> 1) elements, e.g., $\lambda_1(j), \lambda_2(j), \ldots, \lambda_p(j)$, then they represent the branches of a single algebraic function. In this case they are representable with series in fractional powers of j of the following type:

$$\lambda_r(j) = \alpha_0 + \alpha_1 \omega^{r-1} j^{1/p} + \alpha_2 \omega^{2(r-1)} j^{2/p} + \cdots, \quad r = 1, \ldots, p, \quad \omega = \exp(2\pi i/p), \tag{5.2}$$

and j in the right member means the principal value of j.

From the point of view of the j plane, for fixed j there is a finite number ($\leq N$) of roots. This number is N if the roots are all distinct. The points (in the j plane) in which the number of roots changes, are called critical points. One shows that in any compact subset of U_1 the number of critical points is finite. The branch points of the functions forming a cycle are always situated at a critical point. An example is given by the point $j=0$ for the set of functions in Eq. (5.2). It is important to observe that the branches of a cycle are continuous at the critical points.

The projection operator (see Eq. 3.5) associated with each eigenvalue $\lambda_k(j)$ is definite and analytic in the re-

gion where $\lambda_k(j)$ is. Instead, in correspondence to a critical point, the projection may have a discontinuity. Moreover if $\lambda_k(j)$ has a branch point, the corresponding projection operator goes to infinity (in norm).[11] For example, the projection operators corresponding to the eigenvalue functions of the cycle (5.2) will be given by the formulas

$$P_r(j)=\sum_{l=-k}^{\infty} P_l j^{l/p}\omega^{(r-1)l}, \tag{5.3}$$

where $P_{-k}\neq 0$, $k>0$, and the symbols are as above. Any eigenvalue function and the corresponding projection operator are on the same Riemann sheet. The λ plane poles associated to the branches of a cycle are of the same order (see Appendix B).

Now, from the fact that the eigenvalues of a matrix are poles, from the preceding considerations about the separation of $R(\lambda,A(j))$ into two pieces and from Eq. (3.4), we may write in a neighborhood of j_1,

$$R(\lambda,A(j))=\sum_{k=1}^{r}\left[\frac{P^{(k)}(j)}{(\lambda-\lambda_k(j))}+\frac{D_2^{(k)}(j)}{(\lambda-\lambda_k(j))^2}+\cdots+\frac{D_{m_k}^{(k)}(j)}{(\lambda-\lambda_k(j))^{m_k}}\right]+R_0(\lambda,j), \tag{5.4}$$

where $R_0(\lambda,j)$ is holomorphic for λ contained in γ_1 and in $\rho(A(j))$. A simple, though important observation, in this context concerns the nature of the singularities in the j plane for the resolvent given above when λ is fixed: The singularities can only be poles, the cuts disappear, as one can verify by direct computation. Indeed this is a consequence of subdividing the spectrum into two parts and of the fact that the projection operator corresponding to the part under consideration has a finite dimensional range and is an analytic operator.

We are now able to state the main problem of this section. The symbols have the same meaning as above.

(H) Let $A(j)$ be for real $j>j_0$ a completely monotonic operator acting in a real Banach lattice $\mathcal{B}$. Let the spectrum of $A(j_1)$ for real $j_1>j_0$, be separable into two spectral sets σ_1 and σ_2, such that the projection operator associated with σ_1 has finite range and σ_1 contains the rightmost singularity in the λ plane.[15] We wish to find the rightmost singularity in the j plane in a neighborhood of j_1 and give its explicit contribution to the resolvent $R(\lambda,A(j))$.

In order to avoid repetitions, from now on j will be a real variable unless otherwise specified. First we observe that by Proposition 3.3 and by the positivity of $A(j)$ (j real !), the rightmost singularity in the λ plane is situated on the real λ axis at a distance from the origin equal to the spectral radius. From what has been said above this is true in a suitable (real) neighborhood of j_1, so that in this neighborhood the spectral radius $r(j)$ is formed by various pieces of the eigenvalue functions $\lambda_k(j)$, coinciding for fixed j with only one of them except in the critical points at which it coincides with two or more of them. By the properties of the functions $\lambda_k(j)$ it follows that $r(j)$ is continuous everywhere in the same neighborhood and differentiable except perhaps at the critical points.

From now on let $\lambda_0\neq 0$ be a point of the positive real λ axis. The equation $\lambda_0=r(j)$ admits by Proposition 4.2 only one solution $j=\bar{j}$[16] [we exclude the constancy of $r(j)$, as it will be justified later]. On the basis of what was said at the end of Sec. 4, $j=\bar{j}$ is the leading j plane singularity of $R(\lambda_0,A(j))$. Our next task is to study this singularity and to give it an explicit representation in the resolvent. We proceed first to analyze a simple case.

(a) The leading singularity in λ is a first order pole and $j=\bar{j}$ is not a critical point.

Then, according to what has just been said, $r(j)$ coincides in an open (real) neighborhood V of $\bar{j}$ with only one eigenvalue function, for example $\lambda(j)$, and therefore it is differentiable. We intend to show that

$$\left.\frac{d\lambda(j)}{dj}\right|_{j=\bar{j}}\equiv\lambda'(\bar{j})<0. \tag{5.5}$$

This result implies, as we shall see, that the leading singularity in the j plane is also a first order pole.

As a first step we observe that, as $\lambda(j)$ corresponds to a first order pole, we have

$$(A(j)-\lambda(j))P(j)=0, \tag{5.6}$$

where $P(j)$ is the projection operator corresponding to $\lambda(j)$. We differentiate this equation and multiply it from the left by $P(j)$. Taking into account that $P(j)$ commutes with $A(j)$, we get

$$P(j)A'(j)P(j)-\lambda(j)P(j)=0. \tag{5.7}$$

Next we build a positive eigenvector corresponding to $\lambda(j)$. By Proposition 3.5 it follows that the vector $y(j)=P(j)x$, for some $x\in\mathbb{C}$, is an eigenvector of $A(j)$ belonging to the cone $\mathbb{C}$ for any $j\in V$. We stress that this happens because in V, $\lambda(j)$ is in fact the spectral radius. As $y(\bar{j})\neq 0$ [17] we may find a functional $x^*\in\mathbb{C}^*$ such that $\langle x^*,y(\bar{j})\rangle>0$. Let us call $y^*(j)$ the functional $P^*(j)x^*$, where $P^*(j)$ is the tranpose of $P(j)$. Then from Eq. (5.7) we obtain

$$\lambda'(j)=\frac{\langle y^*(j),A'(j)y(j)\rangle}{\langle y^*(j),y(j)\rangle}. \tag{5.8}$$

This equation holds in a suitable neighborhood of $j=\bar{j}$, as $\langle y^*(\bar{j}),y(\bar{j})\rangle=\langle x^*,y(\bar{j})\rangle\neq 0$ and the denominator function is continuous. Now $\langle y^*(j),A'(j)y(j)\rangle_{j=\bar{j}}=\langle y^*(\bar{j}),A'(j)y(\bar{j})\rangle_{j=\bar{j}}$, and it is the derivative at $j=\bar{j}$ of the function $\langle y^*(\bar{j}),A(j)y(\bar{j})\rangle=\langle x^*,A(j)y(\bar{j})\rangle$. This is a completely monotonic function for $j>j_0$ (see Sec. 4), different from zero at $j=\bar{j}$. Therefore, its derivative cannot be zero, otherwise it must be zero everywhere[18] (see Appendix A). Finally we can conclude that $\lambda'(\bar{j})\neq 0$. The same proof can easily be extended to any point $j\in V$. $\lambda'(j)$ is obviously negative.[19]

We may rewrite $\lambda(j)$ in a neighborhood of $\bar{j}$, as

$$\lambda(j)=\lambda_0+\lambda'(\bar{j})f(j)(j-\bar{j}), \tag{5.9}$$

where $f(j)$ is an analytic function such that $f(\bar{j})=1$. In that case the resolvent [see Eq. (5.4)] takes the following form:

$$R(\lambda_0,A(j))=\frac{P(j)}{\lambda'(\bar{j})f(j)(\bar{j}-j)}+R_0(\lambda,j). \tag{5.10}$$

Equation (5.5) implies that the leading pole in the j plane is first order. Equation (5.10) holds in a suitable

complex neighborhood of $\bar{j}$.

The initial hypothesis of this subsection (that $j=\bar{j}$ is not a critical point) is certainly satisfied if the leading pole in λ has one-dimensional eigenspace. For the consequent one-dimensionality of the range of the projection operator $P(j)$, prevents the eigenvalue splitting. This situation occurs if, for example, the operator $A(j)$ is quasi-interior for real j (see Sec. 3).

(b) The leading singularity in λ is a first order pole and $j=\bar{j}$ is a critical point but not a branch point.

The statement of the problem deserves some explanation. Let $\lambda_1(j)$ and $\lambda_2(j)$ be the only two-eigenvalue functions that meet at $j=\bar{j}$, and let them be regular and corresponding to first order poles. Then they are representable by means of the series

$$\lambda_1(j)=\lambda_0+a_{11}(j-\bar{j})+a_{12}(j-\bar{j})^2+\cdots,$$
$$\lambda_2(\bar{j})=\lambda_0+a_{21}(j-\bar{j})+a_{22}(j-\bar{j})^2+\cdots. \tag{5.11}$$

which hold in a neighborhood of $\bar{j}$. Let $P_1(j)$ and $P_2(j)$ be the corresponding projection operators. They are analytic operator-valued functions except perhaps at $j=\bar{j}$, where they may have a discontinuity (see above). Instead, the projection operator $P(j)=P_1(j)+P_2(j)$, which corresponds to the spectral set containing $\lambda_1(j)$ and $\lambda_2(j)$, is analytic everywhere in the previously mentioned neighborhood. As one can immediately check, we may give the following representation of P_1 and P_2:

$$P_1(j)=\frac{(A(j)-\lambda_2(j))P(j)}{\lambda_1(j)-\lambda_2(j)},\quad P_2(j)=\frac{(A(j)-\lambda_1(j))P(j)}{\lambda_2(j)-\lambda_1(j)}. \tag{5.12}$$

From these equations it is evident that the two projection operators in general go to infinity as $j\to\bar{j}$. Using representation (5.12) we may write the analog of Eq. (5.4) in following way:

$$R(\lambda,A(j))=\left[\frac{P_1(j)}{\lambda-\lambda_1(j)}+\frac{P_2(j)}{\lambda-\lambda_2(j)}\right]+R_0(\lambda,j)$$
$$=\left[\frac{P}{\lambda-\lambda_1}+\frac{(A-\lambda_1)P}{(\lambda-\lambda_1)(\lambda-\lambda_2)}\right]+R_0. \tag{5.13}$$

Therefore, though $\lambda_1(j)$ and $\lambda_2(j)$ are first order poles of the resolvent for $j\neq\bar{j}$, $R(\lambda,A(\bar{j}))$ will have in general a second order pole at $\lambda=\lambda_0$. The preceding analysis may easily be extended to the general case in which n regular eigenvalue functions, corresponding to first order poles, meet at $j=\bar{j}$ and take the value λ_0. We use the formula

$$P_i(j)=\prod_{\substack{k=1\\k\neq i}}^{n}\frac{(A(j)-\lambda_k(j))P(j)}{\lambda_i(j)-\lambda_k(j)},\quad i=1,\ldots,n,$$
$$P=\sum_{i=1}^{n}P_i. \tag{5.14}$$

Generally the $P_i(j)$ are infinite at $\bar{j}$ and the resolvent of $A(\bar{j})$ may have *a priori* an nth order pole in λ at $\lambda=\lambda_0$. In general we are not able to go ahead with the analysis as in subsection (a). However, we can do it in the special case in which the projection operator corresponding to the leading eigenvalue for $j\geq\bar{j}$ [let it be $\lambda_1(j)$] is bounded uniformly in a neighborhood of j. From now on, in this subsection, it is understood that we assume this hypothesis. Then it is easy to see that $P_1(j)$, given by Eq. (5.14), is regular for $j=\bar{j}$; but $P_1(\bar{j})$ is no longer the projection operator corresponding to $\lambda_1(\bar{j})=\lambda_0$, the relevant projection being now $P(\bar{j})$. That is, the projection operator associated with $\lambda_1(j)$ has at $j=\bar{j}$ a removable discontinuity. Remember that $P_1(j)$ is a positive operator for $j>\bar{j}$, as it corresponds to the spectral radius. Therefore, also $P_1(\bar{j})$ is positive. At this point one can repeat step by step the proof given in subsection (a) [for example, Eq. (5.8) holds for $j\geq\bar{j}$], and conclude that $\lambda_1'(j)<0$.

It can happen that the spectral radius $r(j)$ coincides with $\lambda_1(j)$ in a right neighborhood of $\bar{j}$ and with another eigenvalue in a left neighborhood of $\bar{j}$; $r(j)$ may have an angular point at $j=\bar{j}$. But also the left derivative is negative. For, we observe that $(d/dj)(\mathrm{Re}\lambda_k(j)|_{j=\bar{j}}\leq\lambda_1'(\bar{j})$, $k=2\cdots n$, otherwise $\lambda_i(j)$ would not be the leading eigenvalue in a right neighborhood of $\bar{j}$.

Now we are able to write down the explicit representation of the resolvent $R(\lambda_0,j)$. For brevity we give it only when λ_0 is a first order pole of $R(\lambda,A(\bar{j}))$ namely $(A(\bar{j})-\lambda_0)P(\bar{j})=0$,

$$R(\lambda_0,j)=\sum_{k=1}^{n}\frac{P_{0k}}{(-\lambda_k'(\bar{j}))}\frac{1}{(j-\bar{j})}+\text{regular part},^{20} \tag{5.15}$$

where P_{0k} is the coefficient corresponding to the zero power in the Laurent expansion of $P_k(j)$ near $j=\bar{j}$.

For further details see the end of Appendix C.

(c) The leading singularity in λ is a first order pole and $j=\bar{j}$ is a branch point.

We have at least a cycle of solutions $\{\lambda_1(j),\ldots,\lambda_p(j)\}$ representable in the form given by Eq. (5.2) (here $\alpha_0=\lambda_0$). Then we show that $\alpha_1=\alpha_2=\cdots=\alpha_{p-1}=0$. Suppose first that one of the eigenvalue functions [e.g., $\lambda_1(j)$] coincides with the spectral radius in a right neighborhood of $\bar{j}$. Then the coefficients of development (5.2) are real. As $r(j)$ is a nonincreasing function of j, we must have $\alpha_1\leq 0$. But, then, between the coefficients $\alpha_1\omega^{r-1}$, $r=1,\ldots,p$, there is at least one, specified for example by the index k, such that $\mathrm{Re}(\alpha_1\omega^{k-1})\geq 0$. Therefore, in a right neighborhood of $\bar{j}$ we would have $\mathrm{Re}(\lambda_k(j))\geq\lambda_1(j)$, against the hypothesis that $\lambda_1(j)$ is the leading eigenvalue, unless $\alpha_1=0$. The same argument may be repeated to show that $\alpha_2=0$ and so on; on the contrary, it is inapplicable to α_p.

However, generally, the leading eigenvalue may not belong to the cycle at issue and there may be several cycles. The preceding argument applies to all cycles meeting at $j=\bar{j}$, with the only difference that α_1 will be a complex number (with a nonpositive real part). This detail does not modify the pattern of the proof given above. Therefore, the general form of the eigenvalue functions belonging to a cycle and assuming the leading value λ_0 at $j=\bar{j}$, is

$$\lambda_r(j)=\lambda_0+\alpha_p(j-\bar{j})+\alpha_{p+1}(j-\bar{j})^{1+1/p}+\cdots,$$
$$r=1,\ldots,p. \tag{5.16}$$

We now want to give an explicit representation of the leading term in the j plane of the resolvent $R(\lambda_0,j)$. Let

us suppose that p eigenvalue functions, belonging to the same cycle, meet in λ_0 (the leading eigenvalue for $j=\bar{j}$). One of these coincides obviously with the spectral radius. The total projection operator $P(j)=\sum_{r=1}^{p}P_r(j)$ [see Eq. (5.3)] is regular at $j=\bar{j}$ and $P(\bar{j})$ is the projection operator corresponding to λ_0. Therefore, $(A(\bar{j})-\lambda_0)P(\bar{j})=0$ and $P(\bar{j})$ is a positive operator (see Proposition 3.4). We show in Appendix C that $\alpha_p<0$. Consequently the resolvent $R(\lambda_0,j)$ has the following form:

$$R(\lambda_0,j)=\frac{Q}{(-\alpha_p)(j-\bar{j})}+\text{regular part}, \tag{5.17}$$

where Q is a linear combination of the operators $P_0,P_{-1},\ldots,P_{-k}$ that appear in Eq. (5.3), with complicated coefficients built with $\alpha_p,\alpha_{p+1},\ldots,\alpha_{p+k}$.

(d) The leading singularity in λ is a higher order pole.

In this case the resolvent has a more complex form, Eq. (5.4), and the projection operator corresponding to the leading eigenvalue is not necessarily positive. We can no longer apply the procedure outlined in (a) in order to show that the first derivative of the leading eigenvalue does not vanish. One conclusion still holds: if $j=\bar{j}$ is a branch point for the roots assuming the leading value λ_0, then they have in a neighborhood of $\bar{j}$ the development (5.16).

In the following section we consider a particular class of operators $A(j)$, for which the problem can be fully solved.

6. LEADING SINGULARITY OF THE RESOLVENT: A SPECIAL CASE

Let $A(j)$ be an analytic operator satisfying the hypotheses (H) and moreover let

$$[A(j_1),\ A(j_2)]=0, \tag{6.1}$$

for any j_1 and j_2. The first consequence of this new condition is that all projections are constant. Indeed, using Eqs. (3.3) and (4.2) we have

$$\frac{dP(j)}{dj}=\left(-\frac{dA(j)}{dj}\right)\frac{1}{2\pi i}\int_\gamma\frac{dR(\lambda,A(j))}{d\lambda}d\lambda=0, \tag{6.2}$$

where γ is a Jordan curve surrounding the singularities at issue. Thus the projections, as functions of j, can have only removable discontinuities corresponding to the critical points. They cannot go to infinity as j approaches a critical point. Thus there are no branch points. All eigenvalue functions are everywhere analytic Another consequence is that $A(j)$ commutes with $A'(j)$.

Let $\lambda(j)$ be the only root taking the value λ_0 at $j=\bar{j}$, and corresponding to a m-order pole in the resolvent. Then,

$$(A(j)-\lambda(j))D_m(j)=0, \tag{6.3}$$

where $D_m(j)=(A(j)-\lambda(j))^{m-1}P$ and P is the projection operator associated to $\lambda(j)$. Deriving both sides of Eq. (6.3) we get

$$(A'(j)-\lambda'(j))D_m(j)=0. \tag{6.4}$$

By Proposition (3.4), $D_m(j)$ is a positive operator. Unless $D_m(j)$ has a zero for $j=\bar{j}$ (see Appendix D), we can find at least one vector $x\in\mathbb{C}$ and one functional $x^*\in\mathbb{C}^*$, such that $\langle x^*,D_m(\bar{j})x\rangle>0$. Then the equation

$$\lambda'(j)=\frac{\langle x^*,A'(j)D_m(j)x\rangle}{\langle x^*,D_m(j)x\rangle}, \tag{6.5}$$

holds in a neighborhood of $\bar{j}$. The number $\langle x^*,A'(\bar{j})D_m(\bar{j})x\rangle$ is the value taken up by the derivative of the function $g(j)=\langle x^*,A(j)D_m(\bar{j})x\rangle$. But $g(j)$ is a completely monotonic function, therefore it does not vanish at $j=\bar{j}$, from Eq. (6.3). So $g'(\bar{j})\neq0$. Finally

$$\lambda'(\bar{j})<0. \tag{6.6}$$

Now let $j=\bar{j}$ be a critical point. Namely, at $j=\bar{j}$ two or more eigenvalues meet. The situation is the same as in Sec. 5(c). So we conclude that Eq. (6.6) holds for any root, at least in those points (if they exist) where the root coincides with the rightmost eigenvalue in the λ plane.

We now give the expression of the resolvent $R(\lambda_0,j)$ when $\bar{j}$ is not a critical point, the extension to critical points being immediate,

$$R(\lambda_0,j)=\frac{D_m(\bar{j})}{(-\lambda'(\bar{j}))^m(j-\bar{j})^m}+\text{nonleading terms}. \tag{6.7}$$

In physical problems we need to know quantities such as $R^k(\lambda_0,j)$. Using Eq. (3.3) we see that the leading term is given by the formula

$$R^k(\lambda_0,j)=\binom{m+k-2}{k-1}\frac{D_m(\bar{j})}{[-\lambda'(\bar{j})(j-\bar{j})]^{m+k-1}}$$

$$+\text{nonleading terms}. \tag{6.8}$$

In the second part of this section we deal with a problem relevant to the multiproduction theory.[1] We are faced with the question of the coincidence between the leading singularity of $A(j)$ and the leading singularity of $R(\lambda_0,A(j))$. The problem is different from the one studied so far, in the sense that in Sec. 5 we set $\bar{j}>j_0$, whereas here we examine the consequences of the condition $\bar{j}=j_0$.

Let $A(j)$ be a completely monotonic operator for $j>j_0$, with a q-order pole at $j=j_0$. Therefore, $A(j)$ can be developed in a Laurent series in a neighborhood of j_0.

$$A(j)=\sum_{i=-q}^{\infty}A_i(j-j_0)^i, \tag{6.9}$$

where A_i $(i=-q,-q+1,\cdots)$ are bounded operators in $\mathcal{B}$. However let the hypotheses (H) and condition (6.1) still hold. $A(j)$ is not defined at $j=j_0$; therefore we cannot define the spectral radius, the eigenvalues, $\cdots$. We resort to the subsidiary operator $\hat{A}(j)$, so defined

$$\hat{A}(j)=A(j)(j-j_0)^q. \tag{6.10}$$

It is clear that $\hat{A}(j)$ is analytic in a neighborhood of j_0 and positive in a right neighborhood of j_0, therefore also at j_0. We observe that $\hat{A}(j)$ is in general not completely monotonic. The relation between the spectral radii $r(j)$ and $\hat{r}(j)$ (where they exist!) is given by

$$\hat{r}(j)=|j-j_0|^q r(j). \tag{6.11}$$

This relation obviously also holds for complex j. As $\hat{A}(j)$ is positive for $j\geq j_0$, $\hat{r}(j)$ coincides with one of the analytic roots of $\hat{A}(j)$, e.g., $\hat{\lambda}_1(j)$, in a right neighbor-

hood of j_0 (including j_0). We recall that $r(j)$ is continuous for $j > j_0$ and nonincreasing, by Proposition 4.2. Therefore $\lim_{j \to j_0^+} r(j)$ exists.

Three possibilities may occur:

(i) $\lim_{j \to j_0^+} r(j) > \lambda_0$

The leading singularity of $R(\lambda_0, A(j))$ lies on the right of j_0. This is the situation previously studied.

(ii) $\lim_{j \to j_0^+} r(j) < \lambda_0$

The leading singularity in $R(\lambda_0, A(j))$ comes from the pole of $A(j)$ at $j = j_0$.[21]

(iii) $\lim_{j \to j_0^+} r(j) = \lambda_0$

$R(\lambda_0, A(j))$ has a leading singularity at j_0, coming from two singularities of $R(\lambda, j)$, one at $j = j_0$ and the other at $\lambda = r(j)$, that overlap as $\lambda \to \lambda_0$.

As we said above, we shall study the last case.

Let us define the analytic function $\lambda_1(j)$ by means of

$$\lambda_1(j) = (j - j_0)^{-q} \hat{\lambda}_1(j). \tag{6.12}$$

It coincides with the leading eigenvalue of $A(j)$ for $j \neq j_0$. In fact, if $\hat{\lambda}_1(j)$ corresponds to an m-order pole of the resolvent, that is, if the analog of Eq. (6.3) holds [we call $\hat{D}_m(j) = (\hat{A}(j) - \hat{\lambda}(j))^{m-1} P$, the projection P being the same both for $A(j)$ and for $\hat{A}(j)$], we have

$$(A(j) - \lambda(j))\hat{D}_m(j) = 0, \tag{6.13}$$

and, from Eq. (6.4)

$$(A'(j) - \lambda'(j))\hat{D}_m(j) = 0. \tag{6.14}$$

Equations (6.13) and (6.14) hold in a right neighborhood of j_0 (including j_0), as $A(j)\hat{D}_m(j)$ and $A'(j)\hat{D}_m(j)$ are finite at $j = j_0$, because of continuity. So, if $\hat{D}_m(j)$ does not vanish at $j = j_0$ (otherwise, see Appendix D), we may find a vector $x \in \mathbb{C}$ and a functional $x^* \in \mathbb{C}^*$, such that $\langle x^*, \hat{D}_m(j)x \rangle > 0$, and we can write

$$\lambda'(j) = \frac{\langle x^*, A'(j)\hat{D}_m(j)x \rangle}{\langle x^*, \hat{D}_m(j)x \rangle}. \tag{6.15}$$

Equation (6.15) holds for $j \geq j_0$ close enough to j_0. Using it, we can show that $\lambda'(j) < 0$ (see Appendix D).

In order to give a representation of the resolvent $R(\lambda, A(j))$ analogous to Eq. (6.7) we refer to the resolvent $R(\hat{\lambda}, \hat{A}(j))$ where $\hat{\lambda} = (j - j_0)^q \lambda$. Due to the resolvent homogeneity, we can write

$$R(\lambda, A(j)) = (j - j_0)^q R(\hat{\lambda}, \hat{A}(j)). \tag{6.16}$$

This relation holds for $|\hat{\lambda}| > \hat{r}(j)$, that is, for $|\lambda| > r(j)$ according to Eq. (6.11), and for $\mathrm{Re} j > j_0$. If $\hat{\lambda}(j)$ is the only root such that $\lambda(j) = \lambda_0$ (the generalization to a critical point is straightforward) we can write, by Eq. (5.4),

$$\begin{aligned} &R(\lambda, A(j)) \\ &= (j - j_0)^q \left\{ \frac{P}{\hat{\lambda} - \hat{\lambda}(j)} + \frac{\hat{D}_2(j)}{(\hat{\lambda} - \hat{\lambda}(j))^2} + \cdots + \frac{\hat{D}_m(j)}{(\hat{\lambda} - \hat{\lambda}(j))^m} \right. \\ &\quad \left. + R_0(\hat{\lambda}, j) \right\} \\ &= \frac{P}{\lambda - \lambda(j)} + \cdots + \frac{\hat{D}_m(j)}{(j - j_0)^{(m-1)q}(\lambda - \lambda(j))^m} + R_0'(\lambda, j). \end{aligned} \tag{6.17}$$

The functions R_0 and R_0' are holomorphic for $\mathrm{Re} j > j_0$ and for λ belonging to a suitable neighborhood of λ_0 (including λ_0!). For $j = j_0$, R_0' may have a singularity independent of λ.

In the last part of this section we briefly discuss the consequences of the results so far obtained in the framework of the multiproduction theory. For a more detailed account we refer the reader to Ref. 1.

Observe that $PR_0' = 0$ and P completely reduces $A(j)$.[5] Roughly speaking $A(j)$ is separated into two parts, one of which gives rise to the principal part of the resolvent and the other to R_0'. This is the general mathematical formulation of those models that in the physical literature are called two-component models. Precisely the part giving rise to R_0' is called the "properly diffractive component, the other is the "properly multiperipheral" component. R_0' may have for $j = j_0$ a higher order pole (or a more complex singularity) than the one contained in the principal part for $\lambda = \lambda_0$. This corresponds to a purely diffractive model, and from the multipheripheral point of view it is uninteresting. So we must suppose that the singularity contained in R_0' is at most as strong as the one contained in the principal part. The highest order of the pole in Eq. (6.17) for $\lambda = \lambda_0$ is $(m-1)q + m$. In this case the coincidence of the initial pole with the final one is possible only if $q = 1$, $m = 1$. However the order of the pole can be lowered by the presence of a zero in $D_m(j)$ for $j = j_0$.[1] In this way the coincidence of the poles is always possible, as one may verify by explicity constructing models. An example is given in Ref. 22.

ACKNOWLEDGMENTS

We thank Professor A. Bassetto and Professor M. Toller for helpful discussions.

APPENDIX A

A real-valued function $F(\xi)$ is called completely monotonic in the real axis interval $a < \xi < b$ if the following conditions are satisfied.

$$(-1)^k \frac{d^k F(\xi)}{d\xi^k} > 0 \quad \text{for } a < x < b \text{ and } k = 0, 1, 2, \cdots. \tag{A1}$$

One shows[23] that the function $F(z) = F(\xi + i\eta)$ is analytic in every circular domain with center c, with $a < c < b$, and radius smaller than $c - a$. If $b = +\infty$, the function $F(z)$ is analytic in the half-plane $\xi > a$.

In this paper we repeatedly use the following property: If a function $F(\xi)$, completely monotonic in $a < \xi < b$, vanishes at a point of this open interval, then it vanishes everywhere. This follows immediately from the analy-

ticity of $F(z)$ and from condition (A1). The same property obviously holds for all derivatives.

Proposition A.1: If $F(\xi)$ is completely monotonic for $a<\xi<\infty$, then

$$|F(\xi+i\eta)|\leq F(\xi), \quad a<\xi<b, \quad -\infty<\eta<+\infty.$$

Proposition A.2: A function $F(\xi)$ is completely monotonic in the interval $a<\xi<\infty$ if and only if

$$F(\xi)=\int_0^\infty \exp(-\xi t)\,d\alpha(t),$$

where $\alpha(t)$ is a nondecreasing function of bounded variation.

The former integral is of the Stieltjes type. If $\alpha(t)$ is absolutely continuous, the integral may be written as

$$F(\xi)=\int_0^\infty \exp(-\xi t)f(t)\,dt, \tag{A2}$$

where $f(t)$ is a nonnegative summable function.

The original motivation of this paper was the study of the Laplace transform of a positive operator $\mathcal{A}(t)$. Symbolically

$$A(z)=\int_0^\infty \exp(-zt)\mathcal{A}(t)\,dt. \tag{A3}$$

Leaving summability questions[1] out of consideration here, this means that in a Banach lattice $\mathcal{B}$, for any $x\in\mathbb{C}$ and $x^*\in\mathbb{C}^*$,

$$\langle x^*,A(z)x\rangle=\int_0^\infty \exp(-zt)\,\langle x^*,\mathcal{A}(t)x\rangle\,dt, \tag{A4}$$

are completely monotonic for z greater than some z_0, on the basis of Proposition A.2. Corollary 2.5 says that $A(z)$ is a completely monotonic operator. We stress that in this paper we are mainly interested in finding the consequences of the positivity of $\mathcal{A}(t)$. Therefore, we are satisfied with deriving some abstract theorems about a generic completely monotonic operator. But in view of the applications it is interesting to wonder if, antitransforming a completely monotonic operator, we get a positive bounded operator. No theorem giving necessary and sufficient conditions exists on this subject. We give a typical—and significant—example.

Proposition A.3: Let $F(z)$ be analytic for $\xi>a$, $F(z)\to 0$ for $\xi\geq a+\delta>a$ when $z\to\infty$ in two dimensions, and

$$\int_{-\infty}^{+\infty}|F(\xi+i\eta)|\,d\eta<\infty \quad \text{for any } \xi>a.$$

Then $F(z)$ can be represented as a Laplace transform of $f(t)$, with

$$f(t)=\frac{1}{2\pi i}\int_{\xi-i\infty}^{\xi+i\infty}\exp(tz)F(z)\,dz \quad (\xi>a),$$

where the integral is independent of the choice of ξ in $\xi>a$. Moreover $f(t)$ is continuous in $-\infty<t<+\infty$ and equals zero for $t<0$. If for any $x\in\mathbb{C}$ and $x^*\in\mathbb{C}^*$, all functions $\langle x^*,A(z)x\rangle$, where $A(z)$ is completely monotonic, satisfy the hypotheses of Proposition A.3, then the antitransformed functions $f_{x,x^*}(t)$ are continuous in t, and for fixed t they are bounded bilinear functionals of x and x^*. Therefore, they define a linear continuous operator $\mathcal{A}(t)$, which is positive as a consequence of Proposition A.2.

Finally we justify the choice, made in Sec. 4, of studying the rightmost real j-axis singularity of the resolvent. The reason is that this singularity is connected to the asymptotic behavior of the antitransformed function. Let us consider the following Tauberian theorem:

Proposition A.4: Let $F(\xi)$ be a completely monotonic function in $0<\xi<\infty$ and let α and A be nonnegative constants. Then if

$$F(\xi)\sim\frac{A}{\xi^\alpha} \quad \text{for } \xi\to 0^+,$$

we have

$$\alpha(t)\sim\frac{At^\alpha}{\Gamma(\alpha+1)} \quad \text{as } t\to+\infty.$$

Referring to Eq. (A2), we observe that this theorem gives us the average asymptotic behavior of $f(t)$. In order to have more detailed information about the pointwise asymptotic behavior of $f(t)$ we need to impose some conditions on $f(t)$. For example if $f(t)$ is nondecreasing, then from Proposition A.4 it follows that

$$f(t)\sim\frac{At^{\alpha-1}}{\Gamma(\alpha)} \quad \text{as } t\to+\infty.$$

APPENDIX B

We demonstrate that the poles associated with eigenvalues belonging to the same cycle are of the same order. Let us suppose at first that an eigenvalue, e.g., $\lambda_{\bar r}$, has a simple pole. Then the projection operator $P_{\bar r}(j)$ satisfies

$$(A(j)-\lambda_{\bar r}(j))P_{\bar r}(j)=0. \tag{B1}$$

Introducing in Eq. (B1) the developments (5.2) and (5.3) for $\lambda_{\bar r}(j)$ and $P_{\bar r}(j)$ and the Taylor series for $A(j)$, we obtain

$$\sum_{q=0}^{\infty}\sum_{l=-k}^{\infty}(A_qP_l(j-\bar j)^{(qp+l)/p}\omega^{l(\bar r-1)}-\lambda_qP_l(j-\bar j)^{(q+l)/p}\times\omega^{(q+l)(\bar r-1)}=0. \tag{B2}$$

Remembering that $\omega=\exp(2\pi i/p)$, we can rewrite Eq. (B2) as

$$\sum_{s=-k}^{\infty}(j-\bar j)^{s/p}\omega^{s(\bar r-1)}\left\{\sum_{q=0}^{(s+k)/p}A_qP_{s-qp}-\sum_{q=0}^{s+k}\lambda_qP_{s-q}\right\}=0. \tag{B3}$$

If Eq. (B3) holds for a given value of r, it holds for every r and therefore one gets

$$(A(j)-\lambda_r(j))P_r(j)=0, \quad r=1,2,\ldots,p. \tag{B4}$$

In general, if an eigenvalue $\lambda_{\bar r}$ has an n-order pole, one obtains an equation with the same structure as (B3), whose validity is independent of the particular value of r; hence any pole associated with the eigenvalue λ_r $(r=1,2,\ldots,p)$ is of the same order n. An essential condition for this proof is that any eigenvalue function and the corresponding projection are on the same Riemann sheet (see Sec. 5).

APPENDIX C

Let $\{\lambda_r(j)\}$, $r=1,\ldots,p$, form a cycle of eigenvalues of an operator $A(j)$ satisfying all the hypotheses (H) of

Sec. 5. As shown in Sec. 5 (c), near the branch point $j=\bar{j}$, where they take the real value λ_0 [= the leading eigenvalue of $A(\bar{j})$], they may be represented by the series

$$\lambda_r(j)=\lambda_0+a_1(j-\bar{j})+a_2(j-\bar{j})^{1+1/p}\omega^{(r-1)}$$
$$+a_3(j-\bar{j})^{1+2/p}\omega^{2(r-1)}+\cdots,\quad r=1,\ldots,p. \tag{C1}$$

Let us suppose that the cycle contains the leading singularity in the λ plane in a neighborhood of $j=\bar{j}$. We may take, for example, $\lambda_1(j)$ coinciding with the spectral radius in a right neighborhood of $\bar{j}$. Let us also suppose that no other eigenvalue takes the value λ_0 in $j=\bar{j}$. [Incidentally we observe that this implies p is odd; so that on the left on $\bar{j}$ the spectral radius coincides with $\lambda_{(p-1)/2}(j)$, as is easy to verify. Indeed, for even p, there cannot be two real roots on both sides of $\bar{j}$.]

Let $\lambda_1(j)$ be a first order pole, then (Appendix B) all eigenvalues of the cycle correspond to first order poles. The total projection $P(j)$ is regular, and we must have (by hypothesis, λ_0 corresponds to a first order pole)

$$(A(\bar{j})-\lambda_0))P(\bar{j})=0. \tag{C2}$$

By Proposition (3.4), $P(\bar{j})$ is positive. Referring to Eq. (5.3) we have $P(\bar{j})=pP_0$, so that P_0 is positive too.

On this basis we want to show that $a_1<0$. As the spectral radius is a nonincreasing function of j, we already know that $a_1\leq 0$. Let us now consider the equations

$$(A(j)-\lambda_r(j))P_r(j)=0,\quad r=1,2,\ldots,p. \tag{C3}$$

Summing over r, we get

$$\sum_{r=1}^{p}\lambda_r(j)P_r(j)=A(j)P(j). \tag{C4}$$

Computing the trace[11,24] on both sides (remember that $\mathrm{tr}P_r=m_r$, where m_r is the dimension of the range of P_r, $m=\sum_r m_r=\dim P$) we obtain

$$\sum_{r=1}^{p}\lambda_r(j)m_r=\mathrm{tr}(A(j)P(j)). \tag{C5}$$

The next step is to develop both sides of this equation in powers of $j^{1/p}$. Equating the coefficients at the zero order we have

$$m\lambda_0=\mathrm{tr}(A(\bar{j})P(\bar{j})). \tag{C6}$$

At the first order we get $ma_1=\mathrm{tr}(A'(\bar{j})P(\bar{j})+A(\bar{j})P'(\bar{j}))$. From Eq. (C2) and the properties of the trace, we have $\mathrm{tr}(AP')=\mathrm{tr}(PAP'+P'AP)=\lambda_0\,\mathrm{tr}(PP'+P'P)=2\lambda_0\,\mathrm{tr}(PP'P)=0$, for $PP'P=0$ as one can verify by deriving the equation $P^2=P$. Finally at the first order the following equation holds:

$$ma_1=\mathrm{tr}(A'(j)P(\bar{j})). \tag{C7}$$

Since $P(\bar{j})$ is positive, it is easy to see that in its range $\mathcal{M}$ there exists a basis $\{e_s\}$ $(s=1,\ldots,m)$, formed by positive vectors. (As a matter of fact, in order for this to be true it would be enough for the projected cone to be reproducing in $\mathcal{M}$). Let x be a positive vector belonging to $\mathcal{M}$; then we may write it as $x=\sum_{s=1}^{m}\alpha_s e_s$, where α_s are nonnegative coefficients. We can always find m functionals g_s^* such that $\langle g_s^*, e_k\rangle=\delta_{sk}$. They are positive in the subspace $\bar{\mathcal{M}}$. Then we may write

$$\mathrm{tr}(A'(\bar{j})P(\bar{j}))$$
$$=\sum_{s=1}^{m}\langle g_s^*,A'(\bar{j})P(\bar{j})e_s\rangle$$
$$=\frac{\partial}{\partial j}\sum_{s=1}^{m}\langle g_s^*,A(j)P(\bar{j})e_s\rangle\Big|_{j=\bar{j}}$$
$$=\frac{\partial}{\partial j}\mathrm{tr}(A(j)P(\bar{j}))\Big|_{j=\bar{j}}. \tag{C8}$$

Now $\mathrm{tr}(A(j)P(\bar{j}))$ is a completely monotonic function different from zero for $j=\bar{j}$ [by Eq. (C6)], so its derivative cannot be zero at $j=\bar{j}$. Therefore, $a_1\neq 0$.

This proof is always applicable when we have a bunch of roots with the common value λ_0 at $j=\bar{j}$, splitting up from this point and satisfying these two conditions: First, they must have equal first derivatives at $\bar{j}$; second, the total projection must be positive for $j=\bar{j}$.

This is the case, for example, when λ_0 is the leading eigenvalue and all projections are infinite at $j=\bar{j}$. For it follows from Eq. (5.14) that all first derivatives of the roots at $j=\bar{j}$ are equal.

APPENDIX D

Referring to the problem stated in Sec. 6, let $D_m(j)$, corresponding to the eigenvalue $\lambda(j)$, have a q-order zero for $j=\bar{j}$. We define the operator $\bar{D}_m(j)=D_m(j)/(j-\bar{j})^q$. $\bar{D}_m(j)$ is everywhere defined in a neighborhood of $j=\bar{j}$, and it is positive for $j>\bar{j}$; therefore it is positive also for $j=\bar{j}$.

Now it is easy to verify that Eqs. (6.3) and (6.4) still hold if we replace $D_m(j)$ by $\bar{D}_m(j)$ for $j=\bar{j}$. Because of continuity these equations will also hold for $j=\bar{j}$. From now on the procedure for showing that $\lambda'(\bar{j})<0$, is the same as in Sec. 6. The formulas (6.7) and (6.8) change in an obvious way.

The proof of Eq. (6.6) is rather different, if $j=\bar{j}$ is a singular point for $A(j)$. We must assume that $A(j)$ is the Laplace transform of a bounded positive operator $\mathcal{A}(t)$ (see Appendix A), so that one has

$$\langle x^*,A(j)\hat{D}_m(j)x\rangle=\int_0^\infty \exp(-jt)\langle x^*,\mathcal{A}(t)\hat{D}_m(j)x\rangle\,dt, \tag{D1}$$

$$\langle x^*,A'(j)\hat{D}_m(j)x\rangle=-\int_0^\infty \exp(-jt)t\langle x^*,\mathcal{A}(t)\hat{D}_m(j)x\rangle\,dt, \tag{D2}$$

where x^*, x, and $\hat{D}_m(j)$ have been defined in Sec. 6. From Sec. 6 we know that $\lambda(\bar{j})>0$ and therefore

$$\int_0^\infty \exp(-jt)\langle x^*,\mathcal{A}(t)\hat{D}_m(j)x\rangle\,dt\geq l>0, \tag{D3}$$

for j belonging to a suitable right neighborhood U of $\bar{j}$. From inequality (D3) it follows that also the integral in Eq. (D2) has a positive least upper bound in U. This is evident if

$$\int_1^\infty \exp(-jt)\langle x^*,\mathcal{A}(t)\hat{D}_m(j)x\rangle\,dt\geq l_1>0,$$

for $j \in U$. Otherwise, we have certainly

$$\int_0^1 \exp(-jt)\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle\, dt \geq l_2 > 0, \tag{D4}$$

for $j \in U$. Remembering that $\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle$ is continuous in the two variables, we have for $0 \leq y \leq 1$ and $j \in U$

$$\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle \leq K, \tag{D5}$$

where K is a suitable constant, such that $\delta = l_2/2K < 1$. From the inequalities (D4) and (D5), one gets

$$\int_\delta^1 \exp(-jt)\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle\, dt \geq l_2/2. \tag{D6}$$

Then

$$\begin{aligned}\int_0^1 &\exp(-jt)t\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle\, dt \\ &> \int_\delta^1 \exp(-jt)t\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle\, dt \\ &> \delta\int_\delta^1 \exp(-jt)\langle x^*, \mathcal{A}(t)\hat{D}_m(j)x\rangle\, dt > \frac{l_2^2}{2K}.\end{aligned} \tag{D7}$$

Therefore,

$$\lambda'(\bar{j}) = \lim_{j\to\bar{j}} \langle x^*, A'(j)\hat{D}_m(j)x\rangle < 0. \tag{D8}$$

[1]L. Bonora and P. Pasti, "Spectral Properties of Positive Operators and j Plane Singularities of Multiperipheral Dynamics," to be published in Nuovo Cimento.

[2]See, for example, H.H. Schaefer, *Banach Lattices and Positive Operators* (Springer, Berlin, 1974); M.A. Krasnosel'skii, *Positive Solutions of Operator Equations* (Noordhoff, Groningen, 1964).

[3]We mean linear operator and linear functional, though many of the following properties are shared by nonlinear operators and functionals.

[4]Z. Semadeni, *Banach Spaces of Continous Functions*, Vol. 1 (Warsawa, 1971).

[5]N. Dunford and J.J. Schwartz, *Linear Operators* (Interscience, New York, 1958), Part I; A.E. Taylor, *Introduction to Functional Analysis* (Wiley, New York, 1958).

[6]By this we mean: $K(t,t')$ is measurable with respect to the product measure; for any $f(t) \in L^p$, $K(t,t')f(t')$ is almost everywhere (for t) integrable in t' and the function $\int |K(t,t')f(t')|\, dt'$ belongs to L^q.

[7]H.H. Schaefer, loc. cit.

[8]N. Bourbaki, *Elements de mathematique. Integration* (Hermann, Paris, 1965), Chap. 4.

[9]W.A.J. Luxemburg and A.C. Zaanen, Indag. Math. 33, 422 (1971).

[10]S. Karlin, J. Math. Mech. 8, 907 (1959).

[11]T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1966).

[12]S. Pinsky and W.I. Weisberger, Phys. Rev. D 2, 2365 (1970).

[13]F. Wolf, Math. Ann. 124, 317 (1952).

[14]B.A. Fuchs and V.I. Levin, *Functions of a Complex Variable*, Vol. II (Pergamon, New York, 1961).

[15]See Ref. 18 below for a further (though less important) characterization of $A(j)$.

[16]It is understood that $\bar{j}$ belongs to the neighborhood of j_1 defined above. There is no loss of generality in thinking $j_1 = \bar{j}$.

[17]Remember that by equation $P^2 = P$, it follows that $\|P\| \geq 1$. Thus $P(j)$ cannot have zeroes.

[18]As a matter of fact $\langle y^*(j), A(j)y(\bar{j})\rangle$ could be constant for $j > j_0$, thus invalidating the conclusion. This is however a very special case. From the point of view of Appendix A the Laplace antitransformed object would be a δ function. So we exclude from our consideration operators $A(j)$ such that $\langle x^*, A(j)x\rangle$ are not strictly decreasing for all $x \in \mathbb{C}$ and $x^* \in \mathbb{C}$.

[19]H. Baumgärtel, Math. Nachr. 26, 361 (1964).

[20]One can verify on the basis of Eq. (5.10) that all $\lambda_k'(\bar{j})$, associated to a projection operator going to infinity, are equal.

[21]This situation is typical of a class of multiproduction models called purely diffractive.

[22]A. Bassetto, L. Bonora, and P. Pasti, Nucl. Phys. B 83, 333 (1974).

[23]D.V. Widder, *The Laplace Transform* (Princeton U.P., Princeton, New Jersey, 1946); G. Doetsch, *Introduction to the Theory and Application of the Laplace Transformation* (Springer, Berlin, 1974).

[24]J. Butler, Arch. Math. 10, 21 (1959).

Induced projective representations*

R. Dirl

Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria
(Received 28 September 1976)

The left- (right-) regular projective representation of a finite group G and the corresponding "projective" representation of the left-group algebra are defined for a given standard factor system, and special features of these constructions are discussed. Starting from a given projective unitary irreducible representation of a normal (but not necessarily Abelian) subgroup N of G, we obtain by induction the matrix elements of the projective unitary irreducible representations of G, where the corresponding group algebra is used as aid. These considerations are of interest for the construction of projective unitary irreducible representations of little cogroups of nonsymmorphic space groups. For the present method allows us to construct, for a $\mathbf{q}$ lying on the "surface" of the Brillouin zone, these projective representations out from unitary irreducible representations belonging to $\mathbf{q}$'s of "lower" symmetry. This method is used to determine for all little cogroups of the nonsymmorphic space group $Pn3n$ complete sets of projective unitary irreducible representations.

I. INTRODUCTION

A topic of interest in mathematical[1-3] and physical[4-9] literature is projective representations of finite and continuous groups. The present paper deals mainly with the problem of constructing by means of induction[1,2] all projective unitary irreducible representations (projective unirreps) of a finite group G for a given standard factor system[1,2,9] out from the projective unirreps of a normal (not necessarily Abelian) subgroup N of G. In particular there is a direct application to the problem of constructing in a systematic way all projective unirreps of the little cogroups which are isomorphic to one of the subgroups of the point group of the crystal whose symmetry group is a nonsymmorphic space group. In this connection it is well known that the determination of the projective unirreps of the little cogroups is one of the several methods[4-6,9-14] which are used to calculate the ordinary vector unirreps for the corresponding groups of the $\tilde{\mathbf{q}}$-vectors.

Instead of using standard methods[2,15,16] we prefer to use the corresponding group algebra as an aid to the induction procedure. At the first moment it seems to be a matter of taste to carry out the induction procedure with the aid of group algebra. However, when comparing the present method with the standard method it seems to be justified to prefer the described method, since the induction procedure becomes in this case more apparent and the explicit determination of the matrix elements of the induced projective unirreps of G is achieved in a practicable way. This comes from the fact that the group algebra of the supergroup is especially suited as representation space, if calculating induced representations, since the construction of ideals and their decomposition in further left ideals can be carried out in a very transparent form.

The material is organized as follows: In Sec. II.A we define the unitary left-, respectively right-regular projective representation of a finite group G for a given standard factor system Q, respectively Q^*. In Sec. II.B we summarize the properties of projective unireeps[9] of G for a given standard factor system Q and compare these projective unirreps with ordinary vector unirreps of G. The definition of the so-called "projective" left-group algebra $A(G)$ is given in the following section. In order to be able to carry out the induction procedure we start in Sec. III.A from the assumption that all projective unirreps of a normal subgroup N of G for the given standard factor system Q (whose domain of definition has to be restricted to the direct product group $N\times N$) are well known. In Sec. III.B we define for a given n_μ-dimensional projective unirrep D^μ of N the corresponding little group $N\{\mu\}$. This gives rise to the existence of a n_μ-dimensional projective unirrep B^μ of $N\{\mu\}$ which belongs however to a standard factor system P $(=QK^*)$ which is in general different from the given one. The task to determine the structure of the special standard factor system K is investigated in Sec. III.C. Thereby it can be shown that contrary to the cases $n_\mu>1$ for the special case $n_\mu=1$ the factor system K can be determined quite generally. In Sec. III.D we construct special induced projective representations of $N\{\mu\}$ which in general are reducible. The reduction of these projective representations can be carried out, if the projective unirreps of the factor group $N\{\mu\}/N$ (belonging to the factor system K) are known. The structure of the induced projective unirreps of $N\{\mu\}$ which are composed of the projective unirreps of the factor group $N\{\mu\}/N$ and B^μ, is discussed in Sec. III.E.[2,3] The last step to induce out from projective unirreps of $N\{\mu\}$, the projective unirreps of G, is carried out in Sec. III.F. The special case, where G is a semidirect product group, is discussed shortly in Sec. III.G. In Sec. III.H we discuss difficulties which may arise, if a nonnormal subgroup H of G is chosen with its projective unirreps as starting point for the induction procedure. In Sec. IV we apply the described method in order to determine the vector unirreps of a nonsymmorphic space group. In Sec. IV.A it will be shown that the main problem lies in the determination of the projective unirreps of the little cogroups. In Sec. IV.B we show that the factor systems are satisfying compatibility relations which makes it possible to apply the described induction procedure for

*Partly presented at the 5th International Colloquium on Group Theoretical Methods in Physics, July 1976, Université de Montréal, Canada.

a systematic determination of the projective unirreps of the little cogroups. In Sec. V we calculate quite generally complete sets of projective unirreps for all little cogroups of the nonsymmorphic space group $Pn3n$.

II. PROJECTIVE REPRESENTATIONS

A. Unitary projective regular representations

Throughout this paper it is assumed that G is a finite group of order $|G|$. We call a function

$$Q:\ G\times G\to C \tag{2.1}$$

of modulus one a standard factor system of G if

$$Q(e,x)=Q(x,e)=1, \tag{2.2}$$

$$Q(x,y)^{-1}=Q^*(x,y), \tag{2.3}$$

$$Q(x,y)Q(xy,z)=Q(x,yz)Q(y,z), \tag{2.4}$$

for all $x,y,z\in G$, where e denotes the identity element of G. The definition of the Hilbert space $L^2(G)$ for a finite or a continuous group G, which is essential for the following considerations, is well known.[1,15-18] In case G is finite, $L^2(G)$ is isomorphic to a $|G|$-dimensional Euclidean space. The scalar product for $L^2(G)$ is given by

$$\langle f,g\rangle=\frac{1}{|G|}\sum_{x\in G}f^*(x)g(x). \tag{2.5}$$

In order to define for a given factor system Q the corresponding "projective left- (right-) regular representation" of G, one has to modify the definition of the ordinary vector representation analogously. We define by

$$\begin{aligned}&V:\ x\to V(x),\\ &[V(x)f](z)=Q(x,x^{-1}z)f(x^{-1}z)\\ &\quad\text{for all } x,z\in G \text{ and } f\in L^2(G),\end{aligned} \tag{2.6}$$

the unitary projective left-regular representation $V(G)$ of G,

$$V(G)=\{V(x):\ x\in G\}, \tag{2.7}$$

$$V(x)V(y)=Q(x,y)V(xy)\ \text{ for all } x,y\in G, \tag{2.8}$$

$$\langle V(x)f,V(x)g\rangle=\langle f,g\rangle\ \text{ for all } f,g\in L^2(G). \tag{2.9}$$

The multiplication law (2.8) and the unitarity (2.9) can be readily verified by means of Eqs. (2.6), (2.3), and (2.5). For a fixed standard factor system Q it is obvious that the definition of the right-regular projective representation of G must be similar to that one given by Eq. (2.6). The unitary operators

$$\begin{aligned}&V':\ x\to V'(x),\\ &[V'(x)f](z)=Q(x,x^{-1}z^{-1})f(zx)\\ &\quad\text{for all } x,z\in G \text{ and } f\in L^2(G)\end{aligned} \tag{2.10}$$

defines the unitary projective right-regular representation, since

$$V'(x)V'(y)=Q(x,y)V'(xy)\ \text{ for all } x,y\in G, \tag{2.11}$$

$$\langle V'(x)f,V'(x)g\rangle=\langle f,g\rangle\ \text{ for all } f,g\in L^2(G), \tag{2.12}$$

is satisfied. However the parallelism to ordinary vector representations disappears completely, since the set of operators

$$\{V(x)V'(y):\ (x,y)\in G\times G\} \tag{2.13}$$

does not define a projective representation of the direct product group $G\times G=\{(x,y):\ x,y\in G\}$ except Q is a trivial standard factor system. Such ones are defined by

$$Q(x,y)=\omega(xy)\omega^*(x)\omega^*(y)\ \text{ with } |\omega(x)|=1 \text{ for all } x\in G. \tag{2.14}$$

In case we introduce instead of definition (2.10) the following new one:

$$\begin{aligned}&V'':\ x\to V''(x),\\ &[V''(x)f](z)=Q^*(z,x)f(zx)\\ &\quad\text{for all } x,z\in G \text{ and } f\in L^2(G),\end{aligned} \tag{2.15}$$

it can be easily shown that the set of unitary operators

$$\{T(x,y)=V(x)V''(y):\ (x,y)\in G\times G\} \tag{2.16}$$

forms a projective unitary representation of the direct product group $G\times G$ belonging to the factor system

$$\begin{aligned}&\underline{Q}((x,y),(x',y'))=Q(x,x')Q^*(y,y')\\ &\quad\text{for all } (x,y),(x',y')\in G\times G\end{aligned} \tag{2.17}$$

$$T(x,y)T(x',y')=Q(x,x')Q^*(y,y')T(xx',yy'). \tag{2.18}$$

Note that if you restrict the direct product group $G\times G$ to the Kronecker product $G[x]G=\{(x,x):\ x\in G\}$, the set of operators

$$\{T(x,x):\ x\in G\} \tag{2.19}$$

represents an ordinary vector representation of G.

B. Projective and vector unirreps

It is well known[9] that the set of all matrix elements of the projective unirreps of G for a given standard factor system Q forms an orthogonal but not normalized basis of $L^2(G)$. The elements of the set

$$\{D^\beta_{pq}:\ \beta\in A_{G(Q)},\ p,q=1,2,\ldots,n_\beta\}, \tag{2.20}$$

where $A_{G(Q)}$ denotes the set of all equivalence classes and n_β the dimension of the corresponding projective unirrep, must satisfy the following equations:

$$\sum_q D^\beta_{pq}(x)D^\beta_{qr}(y)=Q(x,y)D^\beta_{pr}(xy), \tag{2.21}$$

$$\frac{1}{|G|}\sum_{x\in G}D^\beta_{pq}{}^*(x)D^\xi_{rs}(x)=n_\beta^{-1}\delta_{\beta\xi}\delta_{pr}\delta_{qs}, \tag{2.22}$$

$$\sum_{\beta pq}n_\beta D^\beta_{pq}{}^*(x)D^\beta_{pq}(y)=|G|\delta_{xy}. \tag{2.23}$$

However, one must be aware that projective unirreps belonging to nontrivial factor systems have properties which differ essentially from those of the vector unirreps. Whereas for vector unirreps the identity representation must always exist, for projective unirreps the identity representation cannot be realized. The theorem of Burnside

$$\sum_\beta n_\beta^2=|G| \tag{2.24}$$

however, still remains valid. Denoting the set of all matrix elements of the ordinary vector unirreps of G by

$$\{D^\alpha_{ij}:\ \alpha\in A_G,\ i,j=1,2,\ldots,n_\alpha\} \tag{2.25}$$

it is well known that they must satisfy equations which are analogous to Eqs. (2.21)–(2.23), except for the first one, where the Q's have to be replaced by 1 for all $x, y \in G$. Because of (2.22) and the corresponding equation for the matrix elements of the vector unirreps both types of functions are not normalized to unity. Furthermore to obtain a more suitable transformation law with respect to the projective (vector) regular representation we introduce the following bases for $L^2(G)$:

$$\{R^{\beta}_{pq} = \sqrt{n_\beta}\, D^{\beta *}_{pq}:\ \beta \in A_{G(Q)},\ p,q = 1,2,\ldots,n_\beta\}, \tag{2.26}$$

$$\{R^{\alpha}_{ij} = \sqrt{n_\alpha}\, D^{\alpha *}_{ij}:\ \alpha \in A_G,\ i,j = 1,2,\ldots,n_\alpha\}. \tag{2.27}$$

Similar to vector unirreps we obtain for the projective unirreps

$$T(x,y)R^{\beta}_{pq} = \sum_{rs} D^{\beta}_{rp}(x) D^{\beta *}_{sq}(y) R^{\beta}_{rs} \tag{2.28}$$

that $L^2(G)$ decomposes under the action of the unitary projective representation T into a direct sum of projective unirreps $D^{\beta\beta *} = \{D^{\beta\beta *}(x,y) = D^{\beta}(x) \otimes D^{\beta *}(y):\ (x,y) \in G \times G\}$ of $G \times G$ which belong to the factor system $\underline{Q}$. Each projective unirrep $D^{\beta\beta *}$ occurs in T only once. The transformation law

$$T(x,y)R^{\alpha}_{ij} = \sum_{kl} D^{\alpha}_{ki}(x) D^{\alpha *}_{lj}(y) R^{\alpha}_{kl}$$

$$\text{with } T(x,y) = V(x)V''(y), \tag{2.29}$$

where $V(x)$, respectively $V''(y)$ denotes the left-, respectively right-regular vector representation of G,

$$[T(x,y)f](z) = [V(x)V''(y)f](z) = f(x^{-1}zy), \tag{2.30}$$

reflects the corresponding situation for the vector unirreps of G. The equivalence of the bases (2.26) and (2.27) implies that there must exist a unitary transformation depending on the factor system Q which connects both bases,

$$R^{\alpha}_{ij} = \sum_{\beta pq} U^{\alpha;\beta *}_{ij,pq} R^{\beta}_{pq}, \tag{2.31}$$

$$R^{\beta}_{pq} = \sum_{\alpha ij} U^{\alpha;\beta}_{ij,pq} R^{\alpha}_{ij}. \tag{2.32}$$

However, the determination of the unitary transformation U is unsolved, except for the trivial case where Q is a trivial factor system. Note a further remarkable difference between vector and projective unirreps of a given group G. Whereas the complex conjugate of a vector unirrep must be equivalent to one of the given set of the vector unirreps, the situation for projective unirreps is quite different (except for trivial factor systems), since the complex conjugate projective unirreps belong to the factor system Q^* which prevents in any case the above mentioned equivalence relation.

C. Projective group algebra

The definition of the group algebra $A(G)$ of a finite (compact continuous) group G and its properties for ordinary vector representations are extensively dealt with in mathematical and physical literature.[15,16,6,8,17,19,20] One can imagine that the definition of a projective (left-) group algebra $A(G)$ for a given standard factor system Q is quite similar to that of ordinary vector representations. Thereby we restrict our considerations to the definition of the left-group algebra which is in the following, briefly called, group algebra $A(G)$. We call the set

$$A(G) = \{F:\ F \in L^2(G)\} \tag{2.33}$$

whose elements are given by

$$F = |G|^{-1} \sum_{x \in G} F(x) V(x), \tag{2.34}$$

the (left-) group algebra $A(G)$ belonging to the standard factor system Q. $A(G)$ is a Hilbert space being isomorphic to $L^2(G)$ if the scalar product in $A(G)$ is defined by

$$\{F,H\} = \langle F,H\rangle \text{ for all } F,H \in L^2(G). \tag{2.35}$$

$A(G)$ becomes a symmetric ring if the product FH is defined in the usual way and the involution is the mapping $F \to F^\dagger$, where $F^\dagger$ is the adjoint operator. We briefly call the symmetric Hilbert ring $A(G)$ the group algebra. However, one must be careful, e.g., the following relation:

$$V(x)^\dagger = Q^*(x,x^{-1})V(x^{-1}) \text{ for all } x \in G \tag{2.36}$$

shows that the adjoint operator $V(x)^\dagger$ [belonging to $A(G)$ since G is finite] is not identical with the operator $V(x^{-1})$ which represents the inverse group element to x. In case we introduce $\{V(x):\ x \in G\}$ as an orthogonal (but not normalized) basis of $A(G)$ it is obvious that

$$D^{\text{reg}} = \{D^{\text{reg}}(x):\ x \in G\}, \tag{2.37}$$

$$D^{\text{reg}}_{jk}(x_i) = Q(x_i, x_k)\delta_{e, x_j^{-1} x_i x_k} \tag{2.38}$$

defines a (left-) regular projective matrix representation of G and therefore also of $A(G)$.

Analogous to ordinary vector representations $A(G)$ is a semisimple algebra[19,20] which decomposes into a direct sum of n_β^2-dimensional, symmetric and simple algebras $A^\beta(G)$,

$$A(G) = \sum_{\beta \oplus} A^\beta(G). \tag{2.39}$$

For every n_β^2-dimensional simple algebra $A^\beta(G)$ there exists a basis

$$\{E^{\beta}_{pq}:\ \beta \in A_{G(Q)},\ p,q = 1,2,\ldots,n_\beta\} \tag{2.40}$$

whose elements are called "units." They must satisfy

$$E^{\beta\,\dagger}_{pq} = E^{\beta}_{qp}, \tag{2.41}$$

$$E^{\beta}_{pq} E^{\xi}_{rs} = \delta_{\beta\xi}\delta_{qr} E^{\beta}_{ps}, \tag{2.42}$$

$$F = \sum_{\beta pq} \{E^{\beta}_{pq}, F\} n_\beta^{-1} E^{\beta}_{pq}, \tag{2.43}$$

$$V(x) E^{\beta}_{pq} = \sum_{r} D^{\beta}_{rp}(x) E^{\beta}_{rq}. \tag{2.44}$$

In case one knows the projective unirreps of G for the given standard factor system Q the units have to be constructed by

$$E^{\beta}_{pq} = n_\beta |G|^{-1} \sum_{x} D^{\beta *}_{pq}(x) V(x), \tag{2.45}$$

$$V(x) = \sum_{\beta pq} D^{\beta}_{pq}(x) E^{\beta}_{pq}. \tag{2.46}$$

On the other hand, if one has constructed elements of the algebra $A(G)$ satisfying Eqs. (2.41) and (2.42), Eq. (2.44) allows us to calculate the matrix elements of the

projective unirreps for the given standard factor system Q. Relation (2.44) is used in the following sections to actually calculate the matrix elements of the projective unirreps of G.

III. INDUCED PROJECTIVE REPRESENTATIONS

A. Normal subgroup N

In the following we restrict our considerations to the construction of induced projective unirreps of G out of projective unirreps of normal subgroups N of G. The main reason for this restriction arises from that, in that only for normal subgroups a systematic induction procedure can be developed, whereas for nonnormal subgroups H of G the induction proceeds (apart from special cases) much less systematically, as shall be discussed shortly by means of an example in Sec. III.H. But one may not infer from these remarks that all nonnormal subgroups (with their corresponding projective unirreps) are excluded to be chosen as the starting point for the described induction procedure, since for any subgroup H of G whose chain of normalizers

$$N^m(H)=\{x:\ xN^{m-1}(H)x^{-1}=N^{m-1}(H), x\in G\},\quad m\geq 1,$$
$$H\equiv N^0(H)\subseteq N^1(H)\equiv N(H)\subseteq N^2(H)$$
$$\equiv N(N(H))\subseteq\cdots\subseteq N^k(H)\subseteq G, \tag{3.1}$$

is ending in G, the developed method can be used to construct step by step the projective unirreps of G, because $N^m(H)$ [being the normalizer of $N^{m-1}(H)$ in G] contains $N^{m-1}(H)$ as normal subgroup.

Now we assume that for a normal but not necessarily Abelian subgroup $N=\{e,n,m,\cdots\}$ of the given group G, the projective unirreps for the given standard factor system Q are already known. This implies that the elements of the set

$$\{D^\mu_{pq}:\ \mu\in A_{N(Q)},\ p,q=1,2,\ldots,n_\mu\} \tag{3.2}$$

must satisfy the following equations:

$$\sum_q D^\mu_{pq}(n)D^\mu_{qr}(m)=Q(n,m)D^\mu_{pr}(nm), \tag{3.3}$$

$$|N|^{-1}\sum_{n\in N} D^{\mu*}_{pq}(n)D^\nu_{rs}(n)=n_\mu^{-1}\delta_{\mu\nu}\delta_{pr}\delta_{qs}, \tag{3.4}$$

$$\sum_{\mu pq} n_\mu D^{\mu*}_{pq}(n)D^\mu_{pq}(m)=|N|\delta_{nm}. \tag{3.5}$$

Thereby $A_{N(Q)}$ denotes the set of all equivalence classes of the projective unirreps of N (for the given factor system Q which has to be restricted to the subgroup $N\times N$ of $G\times G$) and n_μ denotes the dimension of the corresponding projective unirrep D^μ.

Because of the fact that G is assumed to be finite, $A(N)$ must be a subset of $A(G)$. Likewise for $A(G)$ the semisimple algebra $A(N)$ decomposes into a direct orthogonal sum of n_μ^2-dimensional simple algebras $A^\mu(N)$. For every simple algebra $A^\mu(N)$ we take the elements

$$E^\mu_{pq}=n_\mu|N|^{-1}\sum_{n\in N} D^{\mu*}_{pq}(n)V(n),\quad p,q=1,2,\ldots,n_\mu \tag{3.6}$$

as a basis which satisfies the relations

$$E^{\mu\dagger}_{pq}=E^\mu_{qp}, \tag{3.7}$$

$$E^\mu_{pq}E^\nu_{rs}=\delta_{\mu\nu}\delta_{qr}E^\mu_{ps}, \tag{3.8}$$

$$V(n)E^\mu_{pq}=\sum_r D^\mu_{rp}(n)E^\mu_{rq}. \tag{3.9}$$

B. Little group $N\{\mu\}$

Starting from the assumption that N is a normal subgroup of G we realize that

$$V(y)^\dagger V(n)V(y)=Q(n,y)Q^*(y,y^{-1}ny)V(y^{-1}ny)$$
$$\text{for all } y\in G \text{ and } n\in N \tag{3.10}$$

implies an automorphism for the regular projective representation of the normal subgroup N and therefore also of the group algebra $A(N)$,[17]

$$F\in A(N)\Leftrightarrow V(y)^\dagger FV(y)\in A(N)$$
$$\text{for all } y\in G \text{ and } F\in A(N). \tag{3.11}$$

Consequently we call the set

$$N\{\mu\}=\{x:\ V(x)^\dagger F^\mu V(x)\in A^\mu(N),\ F^\mu\in A^\mu(N)\} \tag{3.12}$$

forming a group, the "little group" which belongs to the projective unirrep D^μ of N. That the group $N\{\mu\}$ is a subgroup G containing itself, N, as a normal subgroup is readily verified. Furthermore, we have to note that every simple algebra $V(y)^\dagger A^\mu(N)V(y)$ (for all $y\in G$) must be identical with one of the simple algebras of $A(N)$ whose dimension is the same as for $A^\mu(N)$ since N is a normal subgroup of G. Denoting the elements of the set $G:N\{\mu\}$ of left-coset representatives (of the left cosets of $N\{\mu\}$ with respect to G) by

$$\underline{z}_j\in G:N\{\mu\}, \tag{3.13}$$

it is obvious that

$$\underline{z}_j N\{\mu\}\underline{z}_j^{-1}=N\{\mu(\underline{z}_j)\},\quad \underline{z}_j\in G:N\{\mu\} \tag{3.14}$$

are conjugate subgroups corresponding to inequivalent projective unirreps of N (provided $\underline{z}_j\neq\underline{z}_k$) which are denoted by $\mu(\underline{z}_j)$ according to the notation $N\{\mu(\underline{z}_j)\}$. This construction decomposes $A_{N(Q)}$ into disjoint subsets

$$\{\mu\}=\{\mu(\underline{z}_j):\ \underline{z}_j\in G:N\{\mu\}\}, \tag{3.15}$$

$$\mu(x)=\mu\Leftrightarrow x\in N\{\mu\}. \tag{3.16}$$

Choosing from each such set one element we obtain the so-called fundamental domain $\Delta A_{N(Q)}$,

$$A_{N(Q)}=\{\mu(\underline{z}_j):\ \mu\in\Delta A_{N(Q)},\underline{z}_j\in G:N\{\mu\}\} \tag{3.17}$$

which has to be determined if one wants to calculate all projective unirreps of G by means of induction out from the projective unirreps $\{D^\mu:\ \mu\in\Delta A_{N(Q)}\}$ of N. Finally we mention that

$$A(N)=\sum_{\mu\in\Delta A_{N(Q)}}{}^{\oplus}A^{\{\mu\}}(N). \tag{3.18}$$

The subalgebras $A^{\{\mu\}}(N)$ appearing in Eq. (3.18) are invariant with respect to the automorphism (3.11) for all $y\in G$.

C. Special projective unirrep of $N\{\mu\}$

To be more concrete the automorphism of the simple algebras $A^\mu(N)$ defined by Eq. (3.12) implies for the elements of the basis (3.6),

$$V(x)^{\dagger}E^{\mu}_{pq}V(x)=n_{\mu}|N|^{-1}\sum_{n}D^{\mu}_{pq}{}^{*}(xnx^{-1})Q^{*}(n,x^{-1})$$
$$\times Q(x^{-1},xnx^{-1})V(n), \tag{3.19}$$

that the projective unirrep

$$D^{\mu}(xnx^{-1})Q(n,x^{-1})Q^{*}(x^{-1},xnx^{-1})\sim D^{\mu}(n)\Leftrightarrow x\in N\{\mu\} \tag{3.20}$$

must be equivalent to itself, supposed $x\in N\{\mu\}$, otherwise

$$D^{\mu}(\underline{z}n\underline{z}^{-1})Q(n,\underline{z}^{-1})Q^{*}(\underline{z}^{-1},\underline{z}n\underline{z}^{-1})$$
$$\sim D^{\mu(\underline{z})}(n)\Leftrightarrow \underline{z}\in G:N\{\mu\}. \tag{3.21}$$

Thereby it is obvious that the standard factor system Q enters essentially into the definition of the little group $N\{\mu\}$. According to Schurs lemma[1-3] there must exist for every $x\in N\{\mu\}$ (up to a unimodular factor forming a trivial factor system) a unitary matrix $B^{\mu}(x)$ which satisfies

$$D^{\mu}(xnx^{-1})Q(n,x^{-1})Q^{*}(x^{-1},xnx^{-1})=B^{\mu}(x)D^{\mu}(n)B^{\mu}(x)^{\dagger}. \tag{3.22}$$

It is well known[1-3] that the set

$$B^{\mu}=\{B^{\mu}(x):\ x\in N\{\mu\}\} \tag{3.23}$$

forms a projective unirrep of $N\{\mu\}$ which belongs however to a standard factor system being in general different to the given one,

$$B^{\mu}(x)B^{\mu}(x')=P(x,x')B^{\mu}(xx'). \tag{3.24}$$

That $B^{\mu}(x)$ (for all $x\in N\{\mu\}$) must be unitary follows from the unitarity of $D^{\mu}(n)$ (for all $n\in N$). Now it is obvious that the new factor system P must be correlated to the given one. If, remembering that the product of two factor systems forms a new one, we make the ansatz

$$P(x,x')=Q(x,x')R^{*}(x,x')\ \text{ for all } x,x'\in N\{\mu\}, \tag{3.25}$$

where because of (3.22) and the assumption

$$B^{\mu}(n)=D^{\mu}(n)\ \text{ for all } n\in N, \tag{3.26}$$

the standard factor system R must satisfy the following condition:

$$R(n,x)=R(x,x^{-1}nx)\ \text{ for all } n\in N \text{ and } x\in N\{\mu\}. \tag{3.27}$$

For the following considerations it is suitable to define a set $N\{\mu\}:N$ of left-coset representatives (of the left cosets of N with respect to $N\{\mu\}$):

$$\underline{x}\in N\{\mu\}:N\Leftrightarrow x=\underline{x}n\ \text{ for all } x\in N\{\mu\}, \tag{3.28}$$

$$N\{\mu\}=\{\underline{x}N:\ \underline{x}\in N\{\mu\}:N\}. \tag{3.29}$$

As a consequence of (3.26) we can choose for the factor system R,

$$R(n,m)=1\ \text{ for all } n,m\in N, \tag{3.30}$$

if we restrict the domain of definition of R from the direct product group $N\{\mu\}\times N\{\mu\}$ to $N\times N$. When considering relation (3.27) it is suggestive that it must be possible to replace the factor system R by a new one which is a constant function on the left cosets $(\underline{x}N\times\underline{x}'N)$ of $N\times N$ with respect to $N\{\mu\}\times N\{\mu\}$. We start from the new factor system K

$$K(x,x')=R(x,x')\omega(xx')\omega^{*}(x)\omega^{*}(x')$$
$$\text{for all } x,x'\in N\{\mu\}, \tag{3.31}$$

where ω shall be a suitable defined trivial standard factor system such that

$$K(x,x')=K(\underline{x}n,\underline{x}'n')=K(\underline{x},\underline{x}')\ \text{ for all } x,x'\in N\{\mu\} \tag{3.32}$$

is satisfied. Equation (3.27) has to be replaced by

$$K(n,x)=K(x,x^{-1}nx)\omega(x^{-1}nx)\omega^{*}(n)$$
$$\text{for all } n\in N \text{ and } x\in N\{\mu\}. \tag{3.33}$$

If choosing

$$\omega(x)=R(c(x)x^{-1},x), \tag{3.34}$$

$$c(x)=c(\underline{x}n)=c(\underline{x})=\underline{x}\ \text{ for all } x\in N\{\mu\}, \tag{3.35}$$

it can be shown by a straightforward calculation using Eqs. (2.2)–(2.4), (3.27), and (3.30) that the factor system

$$K(x_1,x_2)=R(x_1,x_2)R(c(x_1x_2)(x_1x_2)^{-1},x_1x_2)$$
$$\times R^{*}(c(x_1)x_1^{-1},x_1)R^{*}(c(x_2)x_2^{-1},x_2) \tag{3.36}$$

indeed forms a standard factor system of the desired properties (3.32). Therefore, we always assume in the following that the projective unirrep B^{μ} belongs to the factor system

$$P(x_1,x_2)=Q(x_1,x_2)K^{*}(x_1,x_2)\ \text{ for all } x_i\in N\{\mu\}, \tag{3.37}$$

where the standard factor system K satisfies (3.32).

The reason why we consider the special case $n_{\mu}=1$ separately lies in the fact that contrary to the case $n_{\mu}>1$, the factor system K can be determined quite generally. In order to verify this assertion we start from the equivalence relation (3.20) which reduces for this case to an equality,

$$D^{\mu}(xnx^{-1})Q(n,x^{-1})Q^{*}(x^{-1},xnx^{-1})=D^{\mu}(n)$$
$$\text{for all } n\in N \text{ and } x\in N\{\mu\}. \tag{3.38}$$

Obviously

$$B^{\mu}(x)=1\Leftrightarrow P(x,x')=1\ \text{ for all } x,x'\in N\{\mu\} \tag{3.39}$$

must be a solution for (3.38). This implies that P is a trivial standard factor system. If we take the following choice:

$$B^{\mu}(n)=D^{\mu}(n)\ \text{ for all } n\in N, \tag{3.40}$$

$$B^{\mu}(\underline{x})=1\ \text{ for all } \underline{x}\in N\{\mu\}:N, \tag{3.41}$$

we can extend these (one-dimensional) matrices by means of

$$B^{\mu}(\underline{x}n)=Q^{*}(\underline{x},n)D^{\mu}(n)\ \text{ for all } n\in N \text{ and } \underline{x}\in N\{\mu\}:N \tag{3.42}$$

to a projective unirrep of the type (3.24) whereas the condition $K(x,n)=1$ (for all $x\in N\{\mu\}$ and $n\in N$) is already taken into account. Since P must be a trivial

standard factor system we obtain

$$P(x,x') = B^{\mu}(x)B^{\mu}(x')B^{\mu *}(xx') \quad \text{for all } x,x' \in N\{\mu\} \tag{3.43}$$

which has, because of (3.37), as consequence that K must take the following special form:

$$K(x_1,x_2) = Q(x_1,x_2)P^*(x_1,x_2)$$
$$= K(\underline{x}_1,\underline{x}_2) = Q(\underline{x}_1,\underline{x}_2)Q^*(\underline{x}_{12},n_{12})D^{\mu}(n_{12}), \tag{3.44}$$

$$\underline{x}_1\underline{x}_2 = \underline{x}_{12}n_{12}, \quad \underline{x}_{12} \in N\{\mu\}:N, \quad \text{and } n_{12} \in N. \tag{3.45}$$

However, contrary to the case $n_{\mu}=1$ we must confess that for the cases $n_{\mu}>1$ the difficult task to determine quite generally the explicit form of the standard factor system K remains unsolved, i.e., one has to calculate by means of Eqs. (3.25)–(3.27), (3.34), and (3.35) for every group the factor system K separately.

D. Induced projective representations of $N\{\mu\}$

In order to obtain induced projective representations of $N\{\mu\}$ which belong to the original standard factor system Q we start from (3.19) and use (3.22). Because of

$$V(x)^{\dagger}E^{\mu}_{pq}V(x) = \sum_{rs} B^{\mu}_{rp}(x)^{\dagger}E^{\mu}_{rs}B^{\mu}_{qs}(x) \quad \text{for all } x \in N\{\mu\} \tag{3.46}$$

it is suggestive to introduce in the $|N\{\mu\}:N|n_{\mu}^2$-dimensional subalgebra $A(D^{\mu}\uparrow N\{\mu\})$ of $A(G)$ instead of the obvious basis,

$$\{V(\underline{x})E^{\mu}_{pq}: \ \underline{x} \in N\{\mu\}:N; \ p,q=1,2,\ldots,n_{\mu}\}, \tag{3.47}$$

a new basis by means of the following definition:

$$\{F^{\mu;\underline{x}}_{pq} = \sum_{k=1}^{n_{\mu}} B^{\mu}_{kp}(\underline{x})^{\dagger}V(\underline{x})E^{\mu}_{kq}:$$
$$\underline{x} \in N\{\mu\}:N, \ \ p,q=1,2,\ldots,n_{\mu}\}. \tag{3.48}$$

This new basis of $A(D^{\mu}\uparrow N\{\mu\})$ has as a consequence that $A(D^{\mu}\uparrow N\{\mu\})$ decomposes, with respect to the left regular projective representation of $N\{\mu\}$ in n_{μ}, left ideals of the dimension $n_{\mu}|N\{\mu\}:N|$, where $|N\{\mu\}:N|$ denotes the order of the set of the left-coset representatives of N with respect to $N\{\mu\}$. The last assertion can be seen from

$$V(n)F^{\mu;\underline{x}}_{pq} = \sum_{r} D^{\mu}_{rp}(n)F^{\mu;\underline{x}}_{rq}, \tag{3.49}$$

$$V(\underline{x}_1)F^{\mu;\underline{x}_2}_{pq} = K(\underline{x}_1,\underline{x}_2)\sum_{k} B^{\mu}_{kp}(\underline{x}_1)F^{\mu;\underline{x}_1\underline{x}_2}_{kq}, \tag{3.50}$$

$$F^{\mu;\underline{x}_1\underline{x}_2}_{kq} = F^{\mu;\underline{x}_{12}}_{kq}, \tag{3.51}$$

which implies for

$$V(\underline{x}_1n)F^{\mu;\underline{x}_2}_{pq} = K(\underline{x}_1,\underline{x}_2)\sum_{k} B^{\mu}_{kp}(\underline{x}_1n)F^{\mu;\underline{x}_{12}}_{kq} \tag{3.52}$$

that the following induced projective representation of $N\{\mu\}$

$$V(\underline{x}_1n)F^{\mu;\underline{x}_2}_{pq} = \sum_{\underline{x}_3 \in N\{\mu\}:N} \ \sum_{k=1}^{n_{\mu}} D^{\mu\uparrow N\{\mu\}}_{\underline{x}_3k,\underline{x}_2p}(\underline{x}_1n)F^{\mu;\underline{x}_3}_{kq} \tag{3.53}$$

belongs to the given standard factor system Q,

$$D^{\mu\uparrow N\{\mu\}}(x)D^{\mu\uparrow N\{\mu\}}(x') = Q(x,x')D^{\mu\uparrow N\{\mu\}}(xx'). \tag{3.54}$$

The matrix elements $D^{\mu\uparrow N\{\mu\}}_{\underline{x}_3k,\underline{x}_2p}(\underline{x}_1n)$ of the reducible unitary projective representation

$$D^{\mu\uparrow N\{\mu\}} = \{D^{\mu\uparrow N\{\mu\}}(\underline{x}n): \ n \in N, \ \underline{x} \in N\{\mu\}:N\} \tag{3.55}$$

can be readily obtained when comparing (3.52) with (3.53).

E. Projective unirreps of $N\{\mu\}/N$ and $N\{\mu\}$

In order to decompose the reducible projective representation (3.55) in its irreducible constituents, one has to construct the projective unirreps of the factor group $N\{\mu\}/N$ for the given factor system $\underline{K}$ (arising from the factor system K), whose domain of definition is restricted to the left-coset representatives and where K is given by Eqs. (3.36). These projective unirreps of $N\{\mu\}/N$ can be extended by means of the inverse canonical homomorphism $(N\{\mu\}\times N\{\mu\}\to N\{\mu\}/N\times N\{\mu\}/N)$ to projective unirreps of $N\{\mu\}$ belonging to the standard factor system K. We denote the set of the matrix elements of all projective unirreps of $N\{\mu\}/N$ by

$$\{D^{\kappa}_{ab}: \ \kappa \in A_{N\{\mu\}/N(\underline{K})}, \ a,b=1,2,\ldots,n_{\kappa}\}. \tag{3.56}$$

In case we identify the elements of $N\{\mu\}/N$ with the corresponding left-coset representatives by means of the following mapping

$$\underline{x}N \to \underline{x} \quad \text{for all } \underline{x} \in N\{\mu\}:N \tag{3.57}$$

we can write for the multiplication law

$$\sum_{b} D^{\kappa}_{ab}(\underline{x}_1)D^{\kappa}_{bc}(\underline{x}_2) = \underline{K}(\underline{x}_1,\underline{x}_2)D^{\kappa}_{ac}(\underline{x}_{12}), \tag{3.58}$$

where $\underline{x}_{12}$ has to be understood as

$$\underline{x}_1N\underline{x}_2N = \underline{x}_{12}N \to \underline{x}_{12}. \tag{3.59}$$

Now the projective unirreps of $N\{\mu\}/N$ belonging to the factor system $\underline{K}$ can be easily extended to projective unirreps of $N\{\mu\}$ belonging to the factor system K by means of

$$D^{\kappa}(\underline{x}n) = D^{\kappa}(\underline{x}),$$
$$D^{\kappa}(n) = 1 \quad \text{for all } n \in N \text{ and } \underline{x} \in N\{\mu\}:N. \tag{3.60}$$

This by no means contradicts with the definition of the factor system K which satisfies (3.32), since

$$D^{\kappa}(\underline{x}_1n_1)D^{\kappa}(\underline{x}_2n_2) = K(\underline{x}_1n_1,\underline{x}_2n_2)D^{\kappa}(\underline{x}_1n_1\underline{x}_2n_2)$$
$$= K(\underline{x}_1,\underline{x}_2)D^{\kappa}(\underline{x}_{12}) = D^{\kappa}(\underline{x}_1)D^{\kappa}(\underline{x}_2) \tag{3.61}$$

is always valid. The knowledge of the projective unirreps D^{κ} of $N\{\mu\}/N$ allows us to construct the corresponding units of $A(D^{\mu}\uparrow N\{\mu\})$,

$$L^{\kappa;\mu}_{ap,bq} = n_{\kappa}|N\{\mu\}/N|^{-1}\sum_{\underline{x}} D^{\kappa *}_{ab}(\underline{x})F^{\mu;\underline{x}}_{pq}, \quad \mu \in A_{N(Q)},$$
$$\kappa \in A_{N\{\mu\}/N(\underline{K})}, \ \ p,q=1,2,\ldots,n_{\mu}, \ \ a,b=1,2,\ldots,n_{\kappa}. \tag{3.62}$$

After a straightforward calculation by using Eqs. (3.62), (3.61), (3.52), and (3.32) we obtain for

$$V(\underline{x}n)L^{\kappa;\mu}_{ap,bq} = \sum_{cr} D^{\kappa}_{ca}(\underline{x})B^{\mu}_{rp}(\underline{x}n)L^{\kappa;\mu}_{cr,bq}$$
$$= \sum_{cr} D^{\kappa;\mu}_{cr,ap}(\underline{x}n)L^{\kappa;\mu}_{cr,bq} \tag{3.63}$$

that the projective unirreps of $N\{\mu\}$ which are induced out from the projective unirrep D^{μ} of N have the following structure:

$$D^{\kappa,\mu} = \{D^{\kappa,\mu}(x) = D^{\kappa,\mu}(\underline{x}n) = D^{\kappa}(\underline{x}) \otimes B^{\mu}(\underline{x}n):\ x \in N\{\mu\}\}. \tag{3.64}$$

Thereby D^{κ} is a projective unirrep of $N\{\mu\}/N$ belonging to the factor system $\underline{K}$ and B^{μ} is the special projective unirrep of $N\{\mu\}$ which belongs to the factor system $P = QK^*$.

Finally we observe that the subduced projective representations

$$D^{\kappa,\mu}\downarrow N = (\oplus n_{\kappa})\, D^{\mu} \tag{3.65}$$

decomposes into the direct sum of projective unirreps of N without any further unitary transformation.

F. Projective unirreps of G

As the last and simplest step we carry out the induction $D^{\kappa,\mu}\uparrow G$. For this purpose we introduce in $A(G)$ a basis whose elements are given by

$$P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq} = V(\underline{z}_1)\, L^{\kappa;\mu}_{ap,bq}\, V(\underline{z}_2)^{\dagger},\quad \mu \in \Delta A_{N(Q)},$$
$$\kappa \in A_{N\{\mu\}/N(\underline{K})},\quad p,q = 1,2,\ldots,n_{\mu},$$
$$a,b = 1,2,\ldots,n_{\kappa},\quad \underline{z}_j \in G:N\{\mu\}. \tag{3.66}$$

It can be shown by straightforward calculations that the elements (3.66) satisfy the necessary and sufficient conditions

$$P^{(\kappa,\mu)\uparrow G\,\dagger}_{\underline{z}_1 ap, \underline{z}_2 bq} = P^{(\kappa,\mu)\uparrow G}_{\underline{z}_2 bq, \underline{z}_1 ap}, \tag{3.67}$$

$$P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq} P^{(\tau,\xi)\uparrow G}_{\underline{z}_3 cr, \underline{z}_4 ds} = \delta_{\kappa\tau}\delta_{\mu\xi}\delta_{bc}\delta_{qr}\delta_{\underline{z}_2,\underline{z}_3} P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_4 ds}, \tag{3.68}$$

$$F = \sum_{\substack{\kappa\mu \\ \underline{z}_1 ap \\ \underline{z}_2 bq}} \{P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq}, F\}\, n^{-1}_{\kappa\mu\uparrow} P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq} \quad \text{for all } F \in A(G), \tag{3.69}$$

$$n_{\kappa\mu\uparrow} = |G:N\{\mu\}|\, n_{\kappa} n_{\mu}, \tag{3.70}$$

to be units of $A(G)$. The proof of the first equation is trivial, of the second one it is less trivial, since one has to use among other things the fact that $\mu(\underline{z}_j)$ with $\underline{z}_j \neq e$ characterizes an unirrep of N which is not equivalent to this one which belongs to μ. The proof of the third equation (completeness) is rather complicated. Because of (2.44) the matrix elements of the projective unirreps of G belonging to the given standard factor system Q which are induced out from the projective unirreps of N are defined by

$$V(y)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq} = \sum_{\underline{z}_3 cr} D^{(\kappa,\mu)\uparrow G}_{\underline{z}_3 cr, \underline{z}_1 ap}(y)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_3 cr, \underline{z}_2 bq}, \tag{3.71}$$

where its dimension $n_{\kappa\mu\uparrow}$ is given by Eq. (3.70). ($|G:N\{\mu\}|$ denotes the order of the set of the left-coset representatives of $N\{\mu\}$ with respect to G.)

$$D^{(\kappa,\mu)\uparrow G}(y)\, D^{(\kappa,\mu)\uparrow G}(y') = Q(y,y')\, D^{(\kappa,\mu)\uparrow G}(yy'). \tag{3.72}$$

The explicit form of the matrix elements of the induced projective unirreps of G can be obtained by comparison of (3.71) with the following formula:

$$V(y_j)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq}$$
$$= V(\underline{z}_j x_j)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq}$$
$$= Q(y_j, \underline{z}_1)\, Q^*(\underline{z}_{j,\,11(j)}, x_{j,\,11(j)} x_{11(j)} x_{j(1)})\, V(\underline{z}_{j,\,11(j)})$$
$$\times \sum_{cr} D^{\kappa;\mu}_{cr,ap}(x_{j,\,11(j)} x_{11(j)} x_{j(1)})\, P^{(\kappa,\mu)\uparrow G}_{ecr, \underline{z}_2 bq}. \tag{3.73}$$

Thereby we have used that any $y \in G$ can be written uniquely as

$$y = \underline{z}x \quad \text{with } \underline{z} \in G:N\{\mu\} \quad \text{and } x \in N\{\mu\} \tag{3.74}$$

and that the product of any two elements of the set of the left-coset representatives (of $N\{\mu\}$ with respect to G) is given by

$$\underline{z}_1 \underline{z}_2 = \underline{z}_{12} x_{12} \quad \text{with } \underline{z}_{12} \in G:N\{\mu\} \quad \text{and } x_{12} \in N\{\mu\}, \tag{3.75}$$

respectively

$$\underline{z}_j^{-1} x_k \underline{z}_j = \underline{z}_{j(k)} x_{k(j)} \quad \text{for all } \underline{z}_j \in G:N\{\mu\} \quad \text{and } x_k \in N\{\mu\},$$

so that

$$\underline{z}_{j(k)} \in G:N\{\mu\} \quad \text{and } x_{k(j)} \in N\{\mu\}, \tag{3.76}$$

since the groups $N\{\mu\}$ are in general not normal subgroups of G. We note that formula (3.73) can be rewritten in the following form:

$$V(y_j)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap, \underline{z}_2 bq}$$
$$= V(\underline{z}_3) V(\underline{z}_3)^{\dagger} V(y_j) V(\underline{z}_1)\, P^{(\kappa,\mu)\uparrow G}_{\underline{e}ap, \underline{z}_2 bq}$$
$$= Q(y_j, \underline{z}_1) Q^*(\underline{z}_3, \underline{z}_3^{-1} y_j \underline{z}_1) \sum_{cr} D^{\kappa,\mu}_{cr,ap}(\underline{z}_3^{-1} y_j \underline{z}_1)\, P^{(\kappa,\mu)\uparrow G}_{\underline{z}_3 cr, \underline{z}_2 bq} \tag{3.77}$$

[where $\underline{z}_3 (= \underline{z}_{j,\,11(j)})$ is uniquely determined by y_j and $\underline{z}_1$], which gives rise by means of the following definitions (Kronecker–Delta whose domain of definitions are left cosets of $N\{\mu\}$ with respect to G):

$$\Delta^{\mu}(y_1, y_2) = \delta_{y_1 N\{\mu\},\, y_2 N\{\mu\}} = \delta_{\underline{z}_1 N\{\mu\},\, \underline{z}_2 N\{\mu\}}, \tag{3.78}$$

to a more convenient form for the matrix elements of induced projective unirreps of G,

$$D^{(\kappa,\mu)\uparrow G}_{\underline{z}_3 cr, \underline{z}_1 ap}(y) = \Delta^{\mu}(\underline{z}_3, y\underline{z}_1) Q(y, \underline{z}_1) Q^*(\underline{z}_3, \underline{z}_3^{-1} y \underline{z}_1)$$
$$\times D^{\kappa;\mu}_{cr,ap}(\underline{z}_3^{-1} y \underline{z}_1). \tag{3.79}$$

Finally we realize that the induction procedure can be seen most clearly by the special structure of the units (3.66).

The "inverse" procedure, namely the subduction

$$D^{(\kappa,\mu)\uparrow G}\downarrow N \sim \sum_{\underline{z} \in G:N\{\mu\}} (\oplus n_{\kappa})\, D^{\mu}(\underline{z}) \tag{3.80}$$

can be easily carried out and shows that the Frobenius' theorem also holds for projective representations.[2] Thereby we find that the reducible projective representations $D^{(\kappa,\mu)\uparrow G}\downarrow N$ decompose without any further unitary transformation (up to equivalence) into the direct sum of the corresponding projective unirreps of N. The last assertion can be readily verified by means of Eqs. (3.73) and (3.21) if y_j is an element of N,

$$V(n)\,P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ap,\underline{z}_2 bq} = Q(n,\underline{z}_1)Q^*(\underline{z}_1,\underline{z}_1^{-1}n\underline{z}_1)\sum_r D^{\mu}_{rp}(\underline{z}_1^{-1}n\underline{z}_1)\,P^{(\kappa,\mu)\uparrow G}_{\underline{z}_1 ar,\underline{z}_2 bq}, \tag{3.81}$$

or can be directly seen from Eq. (3.79) together with (3.78) and (3.64),

$$D^{(\kappa,\mu)\uparrow G}_{\underline{z}_3 cr,\underline{z}_1 ap}(n) = \delta_{\underline{z}_3,\underline{z}_1}\delta_{ca}\,Q(n,\underline{z}_1)Q^*(\underline{z}_1,\underline{z}_1^{-1}n\underline{z}_1) \times D^{\mu}_{rp}(\underline{z}_1^{-1}n\underline{z}_1). \tag{3.82}$$

G. Special case $G = P \circledS N$

In this part we discuss shortly the induction procedure for the special case where G is a semidirect product group. Thereby we use a more appropriate notation for the group elements of the semidirect product group

$$G = P \circledS N = \{zn:\ z\in P,\ n\in N\}, \tag{3.83}$$

$$z^{-1}nz \in N \quad \text{for all } z\in P \text{ and } n\in N, \tag{3.84}$$

$$zz' \in P \quad \text{for all } z,z'\in P, \tag{3.85}$$

which indicates clearly that N is a normal subgroup of G. From the beginning we assume that we have given a standard factor system Q for G, which can be transferred in a more convenient form which takes into account that G is a semidirect product group (see Ref. 2, Theorem 9.4).

As in Sec. III.A we assume that all projective unirreps of N are given. The structure of the little groups must now be of the form

$$N\{\mu\} = P\{\mu\} \circledS N, \tag{3.86}$$

where $P\{\mu\}$ is a subgroup of P. Its elements are defined by

$$z\in P\{\mu\} \Leftrightarrow D^{\mu}(znz^{-1})Q(n,z^{-1})\,Q^*(z^{-1},znz^{-1}) \sim D^{\mu}(n). \tag{3.87}$$

We note that the n_μ-dimensional projective unirrep B^μ of $N\{\mu\}$ belongs to a standard factor system of the type (3.37). For the special case $n_\mu = 1$ the standard factor system K reduces to Q where the domain of definition has to be restricted to $P\{\mu\}\times P\{\mu\}$. According to the general case we have to construct the projective unirreps of $P\{\mu\}$ belonging to the standard factor system K in order to obtain the projective unirreps of $N\{\mu\}$ which belong to the original factor system Q. Summarizing the results we obtain for the projective unirreps of $N\{\mu\}$,

$$D^{\kappa,\mu} = \{D^{\kappa,\mu}(zn) = D^{\kappa}(z)\otimes B^{\mu}(zn):\ z\in P\{\mu\},\ n\in N\} \tag{3.88}$$

where D^κ belongs to K and B^μ to QK^*. The last step, the induction $D^{\kappa,\mu}\uparrow G$ has to be carried out completely in the same way as in Sec. III.F.

H. Induction for nonnormal subgroups H of G

Analogously to Sec. III.A we assume that the matrix elements of a complete set of projective unirreps of a nonnormal subgroup $H = \{e,h,\cdots\}$ of G are already known. This implies that the elements of the set

$$\{D^{\mu}_{pq}:\ \mu\in A_{H(Q)},\ p,q = 1,2,\ldots,n_\mu\} \tag{3.89}$$

must satisfy equations which are completely equivalent to Eqs. (3.3)–(3.5). Thereby $A_{H(Q)}$ denotes the set of all equivalence classes of the projective unirreps D^μ of H, which belong to the factor system Q.

Since G is finite, $A(H)$ must be a subset of $A(G)$ and decomposes into a direct orthogonal sum of n_μ^2-dimensional simple algebras $A^\mu(H)$. For every simple algebra we take the elements

$$E^{\mu}_{pq} = n_\mu |H|^{-1}\sum_h D^{\mu *}_{pq}(h)V(h), \quad p,q = 1,2,\ldots,n_\mu \tag{3.90}$$

as basis.

Choosing a set $G:H$ of left coset representatives $\underline{z}$ (of H with respect to G) it is obvious that the set

$$\{K^{\mu,\underline{z}}_{pq} = V(\underline{z})E^{\mu}_{pq}:\ \underline{z}\in G:H,\ \mu\in A_{H(Q)},\ p,q = 1,2,\ldots,n_\mu\} \tag{3.91}$$

forms a basis of $A(G)$ which engenders a projective matrix representation of G belonging to the factor system Q. This projective matrix representation is equivalent to the left regular projective matrix representation (2.37), (2.38). Since the sets

$$\{K^{\mu,\underline{z}}_{pq}:\ \underline{z}\in G:H,\ p = 1,2,\ldots,n_\mu\} \tag{3.92}$$

form bases of left ideals of $A(G)$, the regular projective representation decomposes into the direct sum of in general reducible projective representations of G,

$$\begin{aligned} V(y)K^{\mu,\underline{z}_1}_{pq} &= \sum_{\underline{z}_2 r} D^{\mu\uparrow G}_{\underline{z}_2 r,\underline{z}_1 p}(y)K^{\mu,\underline{z}_2}_{rq} \\ &= V(\underline{z}_3)V(\underline{z}_3)^{\dagger}V(y)V(\underline{z}_1)K^{\mu,\underline{e}}_{pq} \\ &= Q(y,\underline{z}_1)Q^*(\underline{z}_3,\underline{z}_3^{-1}y\underline{z}_1)V(\underline{z}_3) \\ &\quad\times\sum_r D^{\mu}_{rp}(\underline{z}_3^{-1}y\underline{z}_1)K^{\mu,\underline{e}}_{rq}. \end{aligned} \tag{3.93}$$

Introducing the definition

$$\Delta(y_1,y_2) = \delta_{y_1 H,y_2 H} = \delta_{\underline{z}_1 H,\underline{z}_2 H}, \tag{3.94}$$

where $y_i = \underline{z}_i h_i$ with $\underline{z}_i\in G:H$ and $h_i\in H$ must be taken into account, we obtain immediately the matrix elements

$$D^{\mu\uparrow G}_{\underline{z}_2 r,\underline{z}_1 p}(y) = \Delta(\underline{z}_2,y\underline{z}_1)Q(y,\underline{z}_1)Q^*(\underline{z}_2,\underline{z}_2^{-1}y\underline{z}_1)D^{\mu}_{rp}(\underline{z}_2^{-1}y\underline{z}_1) \tag{3.95}$$

of the projective representations $D^{\mu\uparrow G}$ ($\mu\in A_{H(Q)}$) of G which belong to the standard factor system Q.

In order to be able to decide whether for a fixed μ $D^{\mu\uparrow G}$ forms a projective unirrep of G or, what is completely equivalent to the question, whether the elements

$$V(\underline{z}_j)E^{\mu}_{pq}V(\underline{z}_k)^{\dagger}:\ j,k = 1,2,\ldots,|G||H|^{-1} \tag{3.96}$$

are units of $A(G)$, we investigate the corresponding character relations. According to the general orthogonality relations of characters (which also hold for projective representations) we know that $D^{\mu\uparrow G}$ is irreducible if and only if

$$\sum_y |\chi^{\mu\uparrow G}(y)|^2 = |G| \tag{3.97}$$

is satisfied where

$$\chi^{\mu\uparrow G}(y)=\sum_{\underline{z}}\Delta(\underline{z},y\underline{z})Q(y,\underline{z})Q^*(\underline{z},\underline{z}^{-1}y\underline{z})\,\chi^{\mu}(\underline{z}^{-1}y\underline{z}), \tag{3.98}$$

$$\chi^{\mu}(h)=\sum_{p}D^{\mu}_{pp}(h), \tag{3.99}$$

denotes the character of the projective representation $D^{\mu\uparrow G}$ of G, respectively of the projective unirrep D^{μ} of H. A simple calculation yields

$$\sum_{y}|\chi^{\mu\uparrow G}(y)|^2=|G|+\sum_{\underline{z}_1\neq\underline{z}_2}\sum_{y\in H_{\underline{z}_1,\underline{z}_2}}Q(y,\underline{z}_1)Q^*(\underline{z}_1,\underline{z}_1^{-1}y\underline{z}_1)\times Q^*(y,\underline{z}_2)Q(\underline{z}_2,\underline{z}_2^{-1}y\underline{z}_2)\times\chi^{\mu}(\underline{z}_1^{-1}y\underline{z}_1)\chi^{\mu*}(\underline{z}_2^{-1}y\underline{z}_2). \tag{3.100}$$

Thereby the group $H_{\underline{z}_1,\underline{z}_2}$ is defined as the intersection of the two conjugate subgroups $\underline{z}_1H\underline{z}_1^{-1}$ and $\underline{z}_2H\underline{z}_2^{-1}$. Now $D^{\mu\uparrow G}$ is irreducible if and only if the second term of Eq. (3.100) vanishes. This is the generalization of Johnston's irreducibility criterion[6] to projective representations.

Inspecting the second term of Eq. (3.100) for the simple example where

$$H_{\underline{z}_1,\underline{z}_2}=\{e\}\ \text{ for all }\underline{z}_i\in G:H \tag{3.101}$$

we find immediately

$$\sum_{y}|\chi^{\mu\uparrow G}(y)|^2=|G|+(|G|\,|H|^{-1}-1)|G|\,|H|^{-1}n_{\mu}^2 \tag{3.102}$$

which implies that every projective representation $D^{\mu\uparrow G}$, $\mu\in A_{H(Q)}$ is reducible. Now there remains the difficult task to determine such linear combinations of the elements (3.92) which engenders projective unirreps of G. However this task can only be carried out much less systematically than for normal subgroups, since neither the concept of little groups nor the concept of factor groups can be applied.

IV. NONSYMMORPHIC SPACE GROUPS

A. Unirreps of nonsymmorphic space groups

In this section we apply the described induction procedure in order to determine the vector unirreps of a nonsymmorphic space group. Thereby we shall show that the determination of the allowed vector unirreps of the little groups $G^{\vec{q}}$ [of the $\vec{q}$-vectors belonging to the fundamental domain (ΔBZ) of the first Brillouin zone (BZ)] is the main problem which arises when constructing the vector unirreps of a nonsymmorphic space group G. It is well known that for nonsymmorphic space groups there exist several methods to calculate the allowed vector unirreps of $G^{\vec{q}}$. One method consists in the determination of all projective unirreps of the factor group $G^{\vec{q}}/T$, where T denotes the translation group and where the form of the factor systems are due to the fact that G is not a semidirect product group. The factor groups are isomorphic to one of the subgroups of the point group P ($\simeq G/T$) of the crystal.

A space group G consists of all elements whose corresponding symmetry operations leave the crystal lattice invariant,

$$G=\{(\alpha|\vec{\tau}(\alpha)+\vec{t}):\ \alpha\in P,\vec{t}\in T\}, \tag{4.1}$$

$$(\alpha|\vec{\tau}(\alpha)+\vec{t})(\beta|\vec{\tau}(\beta)+\vec{t}')=(\alpha\beta|\vec{\tau}(\alpha\beta)+\vec{t}(\alpha,\beta)+D(\alpha)\vec{t}'+\vec{t}), \tag{4.2}$$

$$\vec{t}(\alpha,\beta)=\vec{\tau}(\alpha)+D(\alpha)\vec{\tau}(\beta)-\vec{\tau}(\alpha\beta). \tag{4.3}$$

The symbol $\vec{\tau}(\alpha)$ denotes nonprimitive lattice translations which are uniquely determined when the multiplication law of the nonsymmorphic space group is established. The vectors $\vec{t}(\alpha,\beta)$ defined by Eq. (4.3) are (like $\vec{t}$ or $\vec{t}'$) elements of the translation group and their appearance is typical for nonsymmorphic space groups. Furthermore $D=\{D(\alpha):\ \alpha\in P\}$ is a n-dimensional (n = dimension of the crystal lattice) orthogonal representation of the point group P of the crystal.

First of all, we note that the standard factor system Q of G is equal to one, since we are interested in vector unirreps of G. According to the general induction procedure we start from the one-dimensional vector unirreps of the normal subgroup T of G,

$$D^{\vec{q}}=\{D^{\vec{q}}(\vec{t})=e^{-i\vec{q}\cdot\vec{t}}:\ \vec{t}\in T\},\quad \vec{q}\in BZ\,. \tag{4.4}$$

For the next step [see Eqs. (3.20) and (3.38)] we have to determine for any $\vec{q}$ the corresponding little group $G^{\vec{q}}$ which consists of all space group elements leaving $\vec{q}$ invariant modulo a reciprocal lattice vector $\vec{Q}$,

$$G^{\vec{q}}=\{(\alpha|\vec{\tau}(\alpha)+\vec{t}):\ D(\alpha)\vec{q}=\vec{q}+\vec{Q}\{\vec{q}(\alpha)\},(\alpha|\vec{\tau}(\alpha)+\vec{t})\in G\}. \tag{4.5}$$

Thereby we have to note that the reciprocal lattice vectors $\vec{Q}\{\vec{q}(\alpha)\}$ are uniquely determined. Denoting the elements of $G:G^{\vec{q}}$ by

$$(\underline{\sigma}|\vec{\tau}(\underline{\sigma}))\in G:G^{\vec{q}} \tag{4.6}$$

we are now in the position to determine the fundamental domain ΔBZ of the Brillouin zone [see Eq. (3.15)].

According to Eqs. (3.40) and (3.41) we choose

$$B^{\vec{q}}((e|\vec{t}))=D^{\vec{q}}(\vec{t})\ \text{ for all }\vec{t}\in T, \tag{4.7}$$

$$B^{\vec{q}}((\alpha|\vec{\tau}(\alpha)))=1\ \text{ for all }(\alpha|\vec{\tau}(\alpha))\in G^{\vec{q}}:T, \tag{4.8}$$

where $(\alpha|\vec{\tau}(\alpha))$ are the left coset representatives of T with respect to $G^{\vec{q}}$. As a consequence of Eq. (3.44) we obtain for

$$K^{\vec{q}}((\alpha|\vec{\tau}(\alpha)+\vec{t}),(\beta|\vec{\tau}(\beta)+\vec{t}'))=K^{\vec{q}}((\alpha|\vec{\tau}(\alpha)),(\beta|\vec{\tau}(\beta)))=D^{\vec{q}}(\vec{t}(\alpha,\beta))\quad\text{for all }(\gamma|\vec{\tau}(\gamma)+\vec{t})\in G^{\vec{q}}, \tag{4.9}$$

where the standard factor system $K^{\vec{q}}$ must be supplied by the indices $\vec{q}$. This implies that projective representations of $G^{\vec{q}}/T$ enter in a natural way into the construction of vector unirreps of $G^{\vec{q}}$.

Therefore, in order to obtain the vector unirreps of $G^{\vec{q}}$ we have to calculate the projective unirreps of the factor group $G^{\vec{q}}/T$ belonging to the factor system $\underline{K^{\vec{q}}}$, whose domain of definition is restricted to the left coset representatives. Because of the canonical homomorphism we are able to simplify the notation by identifying the elements of the factor groups $G^{\vec{q}}/T$ with those of the groups

$$P^{\vec{q}} = \{\alpha:\ D(\alpha)\vec{q} = \vec{q} + \vec{Q}\{\vec{q}(\alpha)\}\} \tag{4.10}$$

which are hereafter called "little cogroups." Furthermore we write instead of

$$\underline{K}^{\vec{q}}((\alpha|\vec{\tau}(\alpha)),(\beta|\vec{\tau}(\beta))) \equiv R^{\vec{q}}(\alpha,\beta) = e^{-i\vec{q}\cdot\vec{t}(\alpha,\beta)}\,. \tag{4.11}$$

Assuming that all projective unirreps of $P^{\vec{q}}$ belonging to the factor system $R^{\vec{q}}$ are already determined,

$$\{D^{\kappa}_{ab}:\ \kappa \in A_{P^{\vec{q}}(R^{\vec{q}})},\ a,b = 1,2,\ldots,n_{\kappa}\}, \tag{4.12}$$

we can extend by means of the inverse canonical homomorphism these projective unirreps to such ones of $G^{\vec{q}}$ belonging to $K^{\vec{q}}$ [compare with Eqs. (3.60) and (3.61)],

$$D^{\kappa}_{ab}(\alpha) = D^{\kappa}_{ab}((\alpha|\vec{\tau}(\alpha)+\vec{t})) \quad \text{for all } \vec{t}\in T\,. \tag{4.13}$$

Because of Eq. (3.64) we obtain as vector unirreps of $G^{\vec{q}}$

$$D^{\kappa,\vec{q}} = \{D^{\kappa,\vec{q}}((\alpha|\vec{\tau}(\alpha)+\vec{t})) = D^{\kappa}(\alpha)\otimes D^{\vec{q}}(\vec{t}):\ (\alpha|\vec{\tau}(\alpha)+\vec{t}) \in G^{\vec{q}}\}\,. \tag{4.14}$$

As last and simplest step we obtain immediately by means of Eq. (3.79) the matrix elements of the vector unirreps of the nonsymmorphic space group,

$$D^{(\kappa,\vec{q})\uparrow G}_{\underline{\sigma}_3{}^c,\underline{\sigma}_1{}^a}((\beta|\vec{\tau}(\beta)+\vec{t})) = \Delta^{\vec{q}}((\underline{\sigma}_3|\vec{\tau}(\underline{\sigma}_3)),(\beta|\vec{\tau}(\beta)+\vec{t})(\underline{\sigma}_1|\vec{\tau}(\underline{\sigma}_1))) \times D^{\kappa,\vec{q}}_{ca}((\underline{\sigma}_3|\vec{\tau}(\underline{\sigma}_3))^{-1}(\beta|\vec{\tau}(\beta)+\vec{t})(\underline{\sigma}_1|\vec{\tau}(\underline{\sigma}_1)))\,, \tag{4.15}$$

$$\Delta^{\vec{q}}((\alpha|\vec{\tau}(\alpha)+\vec{t}),(\gamma|\vec{\tau}(\gamma)+\vec{t}')) = \delta_{(\alpha|\vec{\tau}(\alpha)+\vec{t})G^{\vec{q}},(\gamma|\vec{\tau}(\gamma)+\vec{t}')G^{\vec{q}}}\,. \tag{4.16}$$

Because of Eqs. (4.2), (4.11), and (4.14) we obtain for

$$D^{\kappa,\vec{q}}((\underline{\sigma}_3|\vec{\tau}(\underline{\sigma}_3))^{-1}(\beta|\vec{\tau}(\beta)+\vec{t})(\underline{\sigma}_1|\vec{\tau}(\underline{\sigma}_1))) = R^{\vec{q}}(\underline{\sigma}_3)(\beta,\underline{\sigma}_1)R^{\vec{q}}(\underline{\sigma}_3)^*(\underline{\sigma}_3,\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1)\,e^{-iD(\underline{\sigma}_3)\vec{q}\cdot\vec{t}}D^{\kappa}_{ca}(\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1)\,. \tag{4.17}$$

Together with the obvious simplifications for the functions $\Delta^{\vec{q}}$

$$\Delta^{\vec{q}}((\alpha|\vec{\tau}(\alpha)+\vec{t}),(\gamma|\vec{\tau}(\gamma)+\vec{t}')) \equiv \Delta^{\vec{q}}(\alpha,\gamma) = \delta_{\alpha P^{\vec{q}},\gamma P^{\vec{q}}}\,, \tag{4.18}$$

we arrive at the result

$$D^{(\kappa,\vec{q})\uparrow G}_{\underline{\sigma}_3{}^c,\underline{\sigma}_1{}^a}((\beta|\vec{\tau}(\beta)+\vec{t})) = \Delta^{\vec{q}}(\underline{\sigma}_3,\beta\underline{\sigma}_1)R^{\vec{q}}(\underline{\sigma}_3)(\beta,\underline{\sigma}_1)R^{\vec{q}}(\underline{\sigma}_3)^*(\underline{\sigma}_3,\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1) \times e^{-iD(\underline{\sigma}_3)\vec{q}\cdot\vec{t}}D^{\kappa}_{ca}(\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1)\,. \tag{4.19}$$

This result shows, as already pointed out, that the main task lies in the determination of the projective unirreps of $P^{\vec{q}}$ which belong to the standard factor system $R^{\vec{q}}$. Thereby it should be noted that for this method the problem of determining allowed representations does not appear, in contrast to other methods where the allowed vector unirreps of larger groups have to be determined.

The determination of the projective unirreps of $P^{\vec{q}}$ belonging to the factor system $R^{\vec{q}}$ can be simplified, this is yet to be seen in the following sections when changing each factor system by a trivial one. In particular, it is convenient to change for each $\vec{q}\in\Delta BZ$ the factor system by means of the following definitions:

$$D^{\kappa}(\alpha) = e^{-i\vec{q}\cdot\vec{\tau}(\alpha)}R^{\kappa}(\alpha),\quad \alpha\in P^{\vec{q}}, \tag{4.20}$$

$$R^{\kappa}(\alpha)R^{\kappa}(\beta) = S^{\vec{q}}(\alpha,\beta)R^{\kappa}(\alpha\beta)\,, \tag{4.21}$$

$$S^{\vec{q}}(\alpha,\beta) = e^{-i\vec{q}\cdot(D(\alpha)-1)\vec{\tau}(\beta)} \quad \text{for all } \alpha,\beta\in P^{\vec{q}}\,, \tag{4.22}$$

which implies that all projective unirreps $R^{\kappa} = \{R^{\kappa}(\alpha):\ \alpha\in P^{\vec{q}}\}$ of $P^{\vec{q}}$ belong to the standard factor system $S^{\vec{q}}$. This new factor system $S^{\vec{q}}$ can be readily transferred in the complete equivalent form

$$S^{\vec{q}}(\alpha,\beta) = e^{i\vec{Q}\{\vec{q}(\alpha)\}\cdot[\vec{\tau}(\alpha\beta)-\vec{\tau}(\alpha)]} \tag{4.23}$$

by means of

$$\vec{Q}\{\vec{q}(\alpha\beta)\} = \vec{Q}\{\vec{q}(\alpha)\} + D(\alpha)\,\vec{Q}\{\vec{q}(\beta)\} \quad \text{for all } \alpha,\beta\in P^{\vec{q}} \tag{4.24}$$

and $\vec{Q}\cdot\vec{t} = 2\pi n$, n = integer. Hence the matrix elements (4.19) take the form

$$D^{(\kappa,\vec{q})\uparrow G}_{\underline{\sigma}_3{}^c,\underline{\sigma}_1{}^a}((\beta|\vec{\tau}(\beta)+\vec{t})) = \Delta^{\vec{q}}(\underline{\sigma}_3,\beta\underline{\sigma}_1)R^{\vec{q}}(\underline{\sigma}_3)(\beta,\underline{\sigma}_1)R^{\vec{q}}(\underline{\sigma}_3)^*(\underline{\sigma}_3,\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1) \times e^{-i\vec{q}\cdot\vec{\tau}(\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1)}e^{-iD(\underline{\sigma}_3)\vec{q}\cdot\vec{t}}R^{\kappa}_{ca}(\underline{\sigma}_3^{-1}\beta\underline{\sigma}_1)\,. \tag{4.25}$$

Finally we mention that the construction of the corresponding units can be carried out immediately. These units are useful if G-adapted functions have to be constructed. G-adapted functions are ones which transform according to vector unirreps of G and are needed for the calculation of energy bands. We obtain by means of Eqs. (4.25), (4.14), (3.66), (3.62), and (3.6) the operators

$$E^{(\kappa,\vec{q})\uparrow G}_{\underline{\sigma}_1{}^a,\underline{\sigma}_2{}^b} = U((\underline{\sigma}_1|\vec{\tau}(\underline{\sigma}_1)))n_{\kappa}|P^{\vec{q}}|^{-1}\sum_{\alpha\in P^{\vec{q}}}e^{i\vec{q}\cdot\vec{\tau}(\alpha)} \times R^{\kappa}_{ab}{}^*(\alpha)\,U((\alpha|\vec{\tau}(\alpha)))E^{\vec{q}}U((\underline{\sigma}_2|\vec{\tau}(\underline{\sigma}_2)))^{\dagger} \tag{4.26}$$

representing the units in the Hilbert space in question. Thereby $E^{\vec{q}}$ denotes the units of T and $U((\beta|\vec{\tau}(\beta)))$ unitary operators (representing space group elements) which must satisfy the multiplication law (4.2).

B. Compatibility relations for factor systems

Before starting to discuss several possibilities of how the projective unirreps for any $P^{\vec{q}}$ belonging to the standard factor system $S^{\vec{q}}$ can be calculated, we investigate in more detail the factor systems $S^{\vec{q}}$. According to Eq. (4.23) we obtain

$$S^{\vec{q}}(\alpha,\beta) = 1 \quad \text{for } \alpha,\beta\in P^{\vec{q}}\,, \tag{4.27}$$

supposing that $\vec{q}$ does not lie on the "surface" of ΔBZ. This implies that for all $\vec{q}$'s not lying on the surface of ΔBZ the projective unirreps reduce to ordinary vector representations[5,21] and only for $\vec{q}$'s lying on the surface do projective unirreps occur. If $\vec{q}_0$ is a point of the surface of ΔBZ it is obvious that the following subgroup relation always holds:

$$P^{\vec{q}_0}\supset P^{\vec{q}},\quad \vec{q} = \epsilon\vec{q}_0,\quad \epsilon\in(0,1)\,. \tag{4.28}$$

Hence we arrive at the result

$$S^{\vec{q}_0}(\alpha,\beta) = S^{\vec{q}}(\alpha,\beta) = 1 \quad \text{for all } \alpha,\beta\in P^{\vec{q}} = P^{\vec{q}_0}\cap P^{\vec{q}}\,. \tag{4.29}$$

Otherwise if $\vec{q}_0$, $\vec{q}$ are two points of the surface of ΔBZ

at which

$$P^{\vec{q}_0} \supseteq P^{\vec{q}} \tag{4.30}$$

is satisfied we obtain for

$$S^{\vec{q}_0}(\alpha,\beta)S^{\vec{q}*}(\alpha,\beta) = \exp[i(\vec{Q}\{\vec{q}_0(\alpha)\} - \vec{Q}\{\vec{q}(\alpha)\}) \cdot (\vec{\tau}(\alpha\beta) - \vec{\tau}(\alpha))] = 1 \quad \text{if } \alpha,\beta \in P^{\vec{q}} = P^{\vec{q}_0} \cap P^{\vec{q}}, \tag{4.31}$$

since the corresponding reciprocal lattice vectors do not depend on the vectors $\vec{q}_0(\alpha) = D(\alpha)\vec{q}_0$, respectively $\vec{q}(\alpha) = D(\alpha)\vec{q}$, presupposing that (4.30) is satisfied. Consequently

$$S^{\vec{q}_0}(\alpha,\beta) = S^{\vec{q}}(\alpha,\beta) \quad \text{for all } \alpha,\beta \in P^{\vec{q}} = P^{\vec{q}_0} \cap P^{\vec{q}}. \tag{4.32}$$

Equation (4.32) contains Eq. (4.29) as a special case if $\vec{q} = \epsilon\vec{q}_0$ [$\epsilon \in (0,1)$] is admitted and is hereafter called "compatibility relations for the standard factor systems." This relation explains why for any element of a line or plane of symmetry of ΔBZ the same projective unirreps of the corresponding little cogroups occur.

Finally we observe that the formulas (4.28) and (4.29), respectively (4.30) and (4.32), offer two possibilities to calculate in a systematic way the projective unirreps of the little cogroups $P^{\vec{q}_0}$. The first possibility consists in using, as a starting point, the vector unirreps of $P^{\epsilon\vec{q}_0}$ [$\epsilon \in (0,1)$] [according to Eqs. (4.28) and (4.29)] for the induction procedure, presupposing that the chain of normalizers of $P^{\epsilon\vec{q}_0}$ is ending in $P^{\vec{q}_0}$. Obviously the second possibility to determining the projective unirreps of $P^{\vec{q}_0}$ consists in starting [according to Eqs. (4.30) and (4.32)] from projective unirreps of $P^{\vec{q}}$. As in the first case the induction procedure can be applied if and only if the chain of normalizers of $P^{\vec{q}}$ is ending in $P^{\vec{q}_0}$. In both cases the compatibility relations for the factor systems guarantees that the induction procedure is applicable.

C. Induced projective unirreps of the little cogroups

In this section we discuss in more detail how the projective unirreps of the little cogroups can be calculated by means of the two methods. Thereby we restrict our considerations to $\vec{q}$'s lying on the surface of ΔBZ.

In the first case we start [according to Eqs. (4.28) and (4.29)] with the definition of a subgroup of the given group $P^{\vec{q}}$ which takes Eq. (4.28) into account,

$$M^{\vec{q}} = \{\sigma : \vec{q}(\sigma) = \vec{q},\ \sigma \in P^{\vec{q}}\}. \tag{4.33}$$

Groups of this kind are hereafter called "trivial little cogroups." Since $\vec{q}$ lies on the surface of ΔBZ, $M^{\vec{q}}$ is a proper subgroup of $P^{\vec{q}}$, however in general not a normal one,

$$M^{\vec{q}} \subset P^{\vec{q}},\quad N(M^{\vec{q}}) \subseteq P^{\vec{q}}. \tag{4.34}$$

According to the preceding section the first method is applicable if and only if the chain of normalizers of $M^{\vec{q}}$ is ending in the given group $P^{\vec{q}}$. Presupposing that this condition is satisfied, we construct the normalizer of $M^{\vec{q}}$ in $P^{\vec{q}}$. The normalizer of $M^{\vec{q}}$ in $P^{\vec{q}}$ is defined by

$$N(M^{\vec{q}}) = N^{\vec{q}} = \{\delta : \vec{Q}\{q(\delta\sigma\delta^{-1})\} = \vec{0},\ \sigma \in M^{\vec{q}},\ \delta \in P^{\vec{q}}\} \tag{4.35}$$

containing $M^{\vec{q}}$ as a normal subgroup [$\vec{Q}\{\vec{q}(\delta)\} \neq \vec{0} \Leftrightarrow \delta \notin N^{\vec{q}}$] and is hereafter called the "trivial extended little cogroup." Now if $N(M^{\vec{q}})$ is a normal subgroup of $P^{\vec{q}}$, i.e., $N(N(M^{\vec{q}})) = N^2(M^{\vec{q}}) = P^{\vec{q}}$, the induction procedure has to be applied once again, otherwise [$N^2(M^{\vec{q}}) \subset P^{\vec{q}}$] the chain of normalizers has to be completed and the induction procedure applied repeatedly.

For the second case we choose for a given $P^{\vec{q}_0}$ a little cogroup $P^{\vec{q}}$ [where $\vec{q}$ is a point of the surface of ΔBZ with lower symmetry than $\vec{q}_0$ and (4.30) is satisfied], such that for the group

$$M^{\vec{q}_0} = \{\alpha : \vec{q}(\alpha) = \vec{q} + \vec{Q}\{\vec{q}(\alpha)\},\ \alpha \in P^{\vec{q}}\} = P^{\vec{q}} \tag{4.36}$$

the chain of normalizers is ending in $P^{\vec{q}_0}$,

$$N^k(M^{\vec{q}_0}) = P^{\vec{q}_0},\quad k \geq 1. \tag{4.37}$$

The further steps are completely the same as for the first case. In particular, the second way is of interest if the first way cannot be chosen.

Finally it should be noted that the so constructed projective unirreps R^κ of $P^{\vec{q}}$ are especially suited for the subduction

$$R^\kappa \downarrow M^{\vec{q}} \sim \sum_{\kappa'} (\oplus m_{\kappa,\kappa'}) R^{\kappa'}, \tag{4.38}$$

$$M^{\vec{q}} = P^{\vec{q}'} \subseteq P^{\vec{q}}, \tag{4.39}$$

which must be carried out when investigating the "compatibility relations",[22,14,23] since the reducible representations (4.38) decompose into the direct sum of the corresponding unirreps of $P^{\vec{q}'}$ without any further unitary transformations. This implies that the compatibility relations which are usually written in a somewhat different form[14] can be replaced by the simple formula (4.38).

V. EXAMPLE: PROJECTIVE UNIRREPS OF THE LITTLE COGROUPS OF $Pn3n$

In this section we determine quite generally for every little cogroup $P^{\vec{q}}$, where $\vec{q}$ lies on the surface of ΔBZ, the projective unirreps R^κ. For this purpose we specify the nonprimitive lattice translations $\vec{\tau}(\alpha)$ which belong to the elements of the point group O_h of a crystal, whose space group is $Pn3n$. In the following we choose the lattice constant as one which can be done without any loss of generality:

$$\vec{\tau}(n) = \vec{0} \quad \text{for all } n \in O, \tag{5.1}$$

$$\vec{\tau}(In) = \vec{\tau}(nI) = (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \quad \text{for all } n \in O. \tag{5.2}$$

Thereby the point-group element I denotes the inversion. In this connection we mention that we hereafter use the notation used in Ref. 6 for the elements of the point group O_h, respectively, elements of ΔBZ.

First of all we have to determine the standard factor system $S^{\vec{q}}$. According to Eq. (4.22) we must calculate for every $\vec{q}$

$$S^{\vec{q}}(\alpha,\beta) = \exp[-i\vec{q}\cdot(D(\alpha) - 1)\vec{\tau}(\beta)] \quad \text{for all } \alpha,\beta \in P^{\vec{q}}, \tag{5.3}$$

where $D(\alpha)$ denotes 3×3 orthogonal matrices, which can be readily obtained from Table 1.4 of Ref. 6. Because of (4.32) the factor systems (5.3) do not change for every $\vec{q}$ which belongs to a line or plane of symmetry. Therefore, it suffices to determine only one point of a line or plane of symmetry for the factor system. Straightforward calculations yield the following tables (which are written in condensed form and in which, for the sake of simplicity, the corresponding indices $\vec{q}$ are omitted):

Point R $[\vec{q}=\pi(1,1,1)]$

$$P^R = O_h = \{E, I\}\times O$$

$$\begin{aligned}
&S(n,n') = 1 = S(In,n') \quad \text{for all } n,n' \in O,\\
&S(I,In) = -1,\\
&S(C_{2w},In) = 1 = -S(IC_{2w},In), \quad w = x,y,z,\\
&S(C_{3k}^{\pm},In) = 1 = -S(IC_{3k}^{\pm},In), \quad k = 1,2,3,4,\\
&S(C_{4w}^{\pm},In) = -1 = -S(IC_{4w}^{\pm},In), \quad w = x,y,z,\\
&S(C_{2u},In) = -1 = -S(IC_{2u},In), \quad u = a,b,c,d,e,f.
\end{aligned} \tag{5.4}$$

Point X $[\vec{q}=\pi(0,1,0)]$

$$P^X = D_{4h} = \{E, I\}\times C_{4v},$$

$$C_{4v} = \{E, C_{2y}, C_{4y}^{\pm}, \sigma_x, \sigma_z, \sigma_{de}, \sigma_{dc}\}$$

$$\begin{aligned}
&S(n,n') = 1 = S(n,In') \quad \text{for all } n,n' \in C_{4v},\\
&S(In,I) = -1,\\
&S(In,C_{4y}^{\pm}) = 1 = -S(In,IC_{4y}^{\pm}),\\
&S(In,C_{2y}) = 1 = -S(In,IC_{2y}),\\
&S(In,\sigma_w) = -1 = -S(In,I\sigma_w), \quad w = x,z,\\
&S(In,\sigma_{du}) = -1 = -S(In,I\sigma_{du}), \quad u = c,e.
\end{aligned} \tag{5.5}$$

Point M $[\vec{q}=\pi(1,1,0)]$

$$P^M = D_{4h} = \{E,\sigma_x\} \circledS (\{E, I\}\times C_{2v}),$$

$$C_{2v} = \{E, C_{2a}, \sigma_z, \sigma_{db}\}$$

$$\begin{aligned}
&S(n,n') = 1 = S(n,\sigma_x n') \text{ for all } n,n' \in \{E, I\}\times C_{2v} = C_{4v},\\
&S(\sigma_x n, C_{2m}) = 1 = -S(\sigma_x n,\sigma_x C_{2m}), \quad m = a,b,x,y,z,\\
&S(\sigma_x n,\sigma_{du}) = 1 = -S(\sigma_x n,\sigma_x\sigma_{du}), \quad u = a,b.
\end{aligned} \tag{5.6}$$

Line T $[\vec{q}=\pi(1,1,z) : z \in (0,1)]$

$$P^T = C_{4v} = \{E,\sigma_x\} \circledS (\{E, C_{2z}\}\times\{E,\sigma_{db}\})$$

$$\begin{aligned}
&S(n,n') = 1 = S(n,\sigma_x n')\\
&\qquad \text{for all } n,n' \in \{E, C_{2z}\}\times\{E,\sigma_{db}\},\\
&S(\sigma_x n,\sigma_x) = -1,\\
&S(\sigma_x n, C_{2z}) = 1 = -S(\sigma_x n,\sigma_x C_{2z}),\\
&S(\sigma_x n,\sigma_{du}) = -1 = -S(\sigma_x n,\sigma_x\sigma_{du}), \quad u = a,b.
\end{aligned} \tag{5.7}$$

Line S $[\vec{q}=\pi(x,1,x) : x \in (0,1)]$

$$P^S = \{E,\sigma_y\}\times\{E,\sigma_{de}\}$$

$$\begin{aligned}
&S(n,n') = 1 = S(n,\sigma_y n') \quad \text{for all } n,n' \in \{E,\sigma_{de}\},\\
&S(\sigma_y n,\sigma_y) = -1,\\
&S(\sigma_y n,\sigma_{de}) = -1 = -S(\sigma_y n,\sigma_y\sigma_{de}).
\end{aligned} \tag{5.8}$$

Line Z $[\vec{q}=\pi(x,1,0) : x \in (0,1)]$

$$P^Z = \{E,\sigma_y\}\times\{E,\sigma_z\}$$

$$\begin{aligned}
&S(n,n') = 1 = S(n,\sigma_y n') \quad \text{for all } n,n' \in \{E,\sigma_z\},\\
&S(\sigma_y n,\sigma_y) = -1,\\
&S(\sigma_y n,\sigma_z) = -1 = -S(\sigma_y n,\sigma_y\sigma_z).
\end{aligned} \tag{5.9}$$

Plane RMX $[\vec{q}=\pi(x,1,z) : x,z \in (0,1),\ x>z]$

$$P^{\text{plane}} = \{E,\sigma_y\}$$

$$S(\sigma_y,\sigma_y) = -1. \tag{5.10}$$

Now we are in the position to define by means of

$$V^{\vec{q}}(\alpha)V^{\vec{q}}(\beta) = S^{\vec{q}}(\alpha,\beta)V^{\vec{q}}(\alpha\beta) \quad \text{for all } \alpha,\beta \in P^{\vec{q}} \tag{5.11}$$

the left-regular projective representation of the corresponding little cogroups $P^{\vec{q}}$ and to carry out for every $\vec{q}$ the induction procedure. For any case we proceed in the following manner: At first we convince ourselves that for a given $P^{\vec{q}}{}_0$ the chain of normalizers of an appropriated chosen starting group $P^{\vec{q}}$ is ending in $P^{\vec{q}}{}_0$. Then we determine by means of Eqs. (3.20), (3.21), and (3.15) for the unirreps of $P^{\vec{q}}$ the little groups $N\{\mu\}$ and the fundamental domain $\Delta A_{P^{\vec{q}}(S^{\vec{q}})}$ with respect to the normalizer $N(P^{\vec{q}})$. For the following step we determine by means of Eq. (3.22) the special projective unirrep B^μ of $N\{\mu\}$ belonging to the factor system (3.37), which fixes the factor system (3.44) for the factor group $N\{\mu\}/N$. Thereby it suffices to investigate (3.22) for an appropriated chosen set of generators of $P^{\vec{q}}$, respectively $N(P^{\vec{q}})$, in order to be able to determine $N\{\mu\}$ and $\Delta A_{P^{\vec{q}}(S^{\vec{q}})}$. In order to obtain the units (3.62) of $A(D^\mu \uparrow N\{\mu\})$ we have to calculate the projective unirreps D^κ of $N\{\mu\}/N$ which belong to the factor system (3.44). Thus formula (3.64) gives the projective unirreps of $N\{\mu\}$ belonging to the original factor system $S^{\vec{q}}{}_0$ (whose domain of definition has to be restricted to $N\{\mu\}\times N\{\mu\}$). By means of (3.66) we obtain the units of $A(N(P^{\vec{q}}))$ and therefrom the projective unirreps of the normalizer $N(P^{\vec{q}})$. Obviously it is sufficient to calculate the projective unirreps for the set of generators of $N(P^{\vec{q}})$ in order to obtain by means of (5.11) the matrix representations of the other elements of $N(P^{\vec{q}})$. For any further step of induction which is necessary according to the chain of normalizers we proceed completely in the same way.

Point R: $P^{\vec{q}} = O \to P^{\vec{q}}{}_0 = O_h$

According to Fig. 4.1 of Ref. 6 we cannot proceed the first way [Eqs. (4.28) and (4.29)], since the chain of normalizers of $P^{\vec{q}} = C_{3v}$ does not end in $P^{\vec{q}}{}_0 = O_h$. Hence we must look for a subgroup of O_h whose projective unirreps can be easily found and whose chain of normalizers is ending in O_h. Because of the special structure of $O_h = \{E, I\}\times O$ and $S(n,n') = 1$ for all $n,n' \in O$ it is obvious to use the vector unirreps of O as starting point for the induction procedure. The vector unirreps of O for the set $\{C_{2x}, C_{2z}, C_{31}^{+}, C_{2d}\}$ of generating elements of O are given by

$$D^\mu: \quad \mu = 0,1$$

$$C_{2x} \to 1, \quad C_{2z} \to 1, \quad C_{31}^{+} \to 1, \quad C_{2d} \to (-1)^\mu; \tag{5.12}$$

D^{μ}: $\mu = 2$

$$C_{2x} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix},$$
$$C_{31}^{+} \to \begin{vmatrix} \omega & 0 \\ 0 & \omega^2 \end{vmatrix}, \quad \omega = e^{-i2\pi/3}, \quad C_{2d} \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}; \tag{5.13}$$

D^{μ}: $\mu = 3, 4$

$$C_{2x} \to \begin{vmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix},$$
$$C_{31}^{+} \to \begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix}, \quad C_{2d} \to (-1)^{\mu} \begin{vmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{vmatrix}. \tag{5.14}$$

According to Eq. (3.22) we have to investigate whether there exists a projective unirrep $\mathrm{B}^{\mu}(I)$ satisfying

$$D^{\mu}(n)S(n, I) = \mathrm{B}^{\mu}(I)D^{\mu}(n)\mathrm{B}^{\mu}(I)^{\dagger}, \quad \mu = 0, 1, 2, 3, 4, \tag{5.15}$$

where $InI = n$ ($n \in O$) and (5.4) are already taken into account. Thereby it suffices to investigate (5.15) for the generating elements of O. Simple calculations yield

$$N\{\mu\} = O \text{ for } \mu = 0, 1 \text{ and } 3, 4, \quad N\{\mu = 2\} = O_h, \tag{5.16}$$

$$\Delta A_O = \{\mu: \mu = 0, 2, 4\}, \tag{5.17}$$

$$\mathrm{B}^{\mu=2}(I) = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix} \text{ with } P(I, I) = S(I, I) = -1 \Leftrightarrow K(I, I) = 1, \tag{5.18}$$

which implies that for $\mu = 2$ vector unirreps of $O_h/O \simeq \{E, I\}$ have to be constructed, since $K(I, I) = 1$. Denoting these vector unirreps by

$$D^{\kappa}(I) = (-1)^{\kappa}, \quad \kappa = 0, 1 \tag{5.19}$$

we obtain immediately by means of (3.64) two inequivalent two-dimensional projective unirreps of O_h, namely

$\mathrm{D}^{(\kappa, \mu=2)\uparrow O_h}$: $\kappa = 0, 1$

$$C_{2x} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix},$$
$$C_{31}^{+} \to \begin{vmatrix} \omega & 0 \\ 0 & \omega^2 \end{vmatrix}, \quad C_{2d} \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad I \to (-1)^{\kappa} \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}. \tag{5.20}$$

For the cases $\mu = 0, 4$ we use the units [compare Eq. (3.66)]

$$\mathrm{P}^{(\mu)\uparrow O_h}_{Ep, Eq} = \mathrm{E}^{\mu}_{pq}, \quad \mathrm{P}^{(\mu)\uparrow O_h}_{Ip, Eq} = \mathrm{V}(I)\mathrm{E}^{\mu}_{pq} \tag{5.21}$$

to determine [by means of Eq. (3.71)] for the generating elements the matrix representations. Thereby E^{μ}_{pq} denote the units of $\mathrm{A}(O)$. The superfluous indices κ (= 0), respectively, $\vec{q}$ of the elements of the left-regular projective representation of O_h are omitted in Eq. (5.21). We obtain immediately for

$\mathrm{D}^{(\mu=0)\uparrow O_h}$:

$$C_{2x} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix},$$
$$C_{31}^{+} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad C_{2d} \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad I \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}; \tag{5.22}$$

$\mathrm{D}^{(\mu=4)\uparrow O_h}$:

$$C_{2x} \to \begin{vmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{vmatrix}, \quad C_{31}^{+} \to \begin{vmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix},$$
$$C_{2d} \to \begin{vmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{vmatrix}, \quad I \to \begin{vmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{vmatrix}, \tag{5.23}$$

which form together with (5.20) the desired projective unirreps of O_h. The matrix representations for the remaining group elements can be easily calculated by means of

$$\mathrm{D}^{(\kappa, \mu)\uparrow O_h}(y_1)\mathrm{D}^{(\kappa, \mu)\uparrow O_h}(y_2) = \mathrm{S}(y_1, y_2)\mathrm{D}^{(\kappa, \mu)\uparrow O_h}(y_1 y_2). \tag{5.24}$$

Point X: $P^{\vec{q}} = P^{\Delta} = C_{4v} \to P^{\vec{q}_0} = P^{X} = \{E, I\} \times C_{4v}$

In this case we choose the first way which is described by the formulas (4.28) and (4.29). The vector unirreps of C_{4v} for a set of generating elements are given by

D^μ: $\mu = 0, 1$

$$C_{4v}^+ \to 1, \quad \sigma_x \to (-1)^\mu ; \tag{5.25}$$

D^μ: $\mu = 2, 3$

$$C_{4v}^+ \to -1, \quad \sigma_x \to (-1)^\mu ; \tag{5.26}$$

D^μ: $\mu = 5$

$$C_{4v}^+ \to \begin{vmatrix} \omega & 0 \\ 0 & \omega^3 \end{vmatrix}, \quad \omega = e^{-i\pi/2}, \quad \sigma_x \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \tag{5.27}$$

According to Eq. (3.22) we have to examine whether there exists a projective unirrep $\mathrm{B}^\mu(I)$ satisfying

$$D^\mu(n)S(I,n) = \mathrm{B}^\mu(I)D^\mu(n)\mathrm{B}^\mu(I)^\dagger , \tag{5.28}$$

where $InI = n$ ($n \in C_{4v}$) and (5.5) are already taken into account. Simple calculations yield

$$N\{\mu\} = C_{4v} \quad \text{for } \mu = 0, 1 \text{ and } 2, 3, \quad N\{\mu = 5\} = P^X , \tag{5.29}$$

$$\Delta A_{C_{4v}} = \{0, 2, 5\}, \tag{5.30}$$

$$\mathrm{B}^{\mu=5}(I) = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix} \quad \text{with } P(I,I) = S(I,I) = -1 \Leftrightarrow K(I,I) = 1 , \tag{5.31}$$

which implies that for $\mu = 5$ vector unirreps of $P^X/C_{4v} \simeq \{E, I\}$ must be calculated. These vector unirreps

$$D^\kappa(I) = (-1)^\kappa, \quad \kappa = 0, 1 \tag{5.32}$$

yield, according to Eq. (3.64), two inequivalent two-dimensional projective unirreps of P^X, namely

$\mathrm{D}^{(\kappa,\mu=5)\uparrow P^X}$: $\kappa = 0, 1$

$$C_{4v}^+ \to \begin{vmatrix} \omega & 0 \\ 0 & \omega^3 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad I \to (-1)^\kappa \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}. \tag{5.33}$$

For the cases $\mu = 0, 2$ we must apply formula (3.66),

$$\mathrm{P}_{E,E}^{(\mu)\uparrow P^X} = \mathrm{E}^\mu, \quad \mathrm{P}_{I,E}^{(\mu)\uparrow P^X} = \mathrm{V}(I)\mathrm{E}^\mu \tag{5.34}$$

in order to be able to determine for the generating elements the matrix representations. The superfluous indices p, q of the units $\mathrm{E}^\mu \in \mathrm{A}(C_{4v})$, respectively κ (= 0), are omitted again. We obtain two further two-dimensional projective unirreps of P^X,

$\mathrm{D}^{(\mu=0)\uparrow P^X}$:

$$C_{4v}^+ \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad I \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}, \tag{5.35}$$

$\mathrm{D}^{(\mu=2)\uparrow P^X}$:

$$C_{4v}^+ \to \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad I \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}, \tag{5.36}$$

which forms together with (5.33) the desired projective unirreps of P^X.

Point M: $\vec{P^q} = P^\Sigma = C_{2v} \to N(P^\Sigma) = \{E, I\} \times C_{2v}$

$$= C_{4h} \to N^2(P^\Sigma) = P^M = D_{4h} .$$

In this case we choose again the first way [Eqs. (4.28) and (4.29)]. Because of (5.6) we start immediately from the vector unirreps of the normalizer $N(P^\Sigma) = C_{4h}$. The vector unirreps (written in a closed form) are given by

D^μ: $\mu = (k, l, m)$, $k, l, m = 0, 1$

$$C_{2a} \to (-1)^k, \quad \sigma_z \to (-1)^l, \quad I \to (-1)^m. \tag{5.37}$$

According to Eq. (3.22) we examine whether there exists a projective unirrep $\mathrm{B}^\mu(\sigma_x)$ satisfying

$$D^\mu(\sigma_x n \sigma_x)S(\sigma_x, \sigma_x n \sigma_x) = \mathrm{B}^\mu(\sigma_x)D^\mu(n)\mathrm{B}^\mu(\sigma_x)^\dagger$$
$$\text{for all } n \in C_{4h} . \tag{5.38}$$

Thereby we have to take $\sigma_x C_{2a}\sigma_x = C_{2b} = I\sigma_z C_{2a}$, $\sigma_x\sigma_z\sigma_x = \sigma_z$, $\sigma_x I \sigma_x = I$, and (5.6) into account. We obtain for

$$N\{\mu\} = C_{4h} \quad \text{for } \mu = (0,0,0), (0,1,1) \text{ and } (1,0,0), (1,1,1) \text{ and } (0,1,0), (1,0,1) \text{ and } (0,0,1), (1,1,0), \tag{5.39}$$

$$\Delta A_{C_{4h}} = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\} . \tag{5.40}$$

Therefore, we must apply (3.66) for all $\mu = (k, l, m) \in \Delta A_{C_{4h}}$,

$$\mathrm{P}_{E,E}^{(\mu)\uparrow P^M} = \mathrm{E}^\mu, \quad \mathrm{P}_{\sigma_x,E}^{(\mu)\uparrow P^M} = \mathrm{V}(\sigma_x)\mathrm{E}^\mu , \tag{5.41}$$

which gives rise to the following projective unirreps of P^M:

$\mathrm{D}^{(k,l,m)\uparrow P^M}$: $(k, l, m) \in \Delta A_{C_{4h}}$

$$C_{2a} \to \begin{vmatrix} (-1)^k & 0 \\ 0 & (-1)^{k+l+m} \end{vmatrix}, \quad \sigma_z \to \begin{vmatrix} (-1)^l & \\ 0 & (-1)^{l+1} \end{vmatrix},$$
$$I \to \begin{vmatrix} (-1)^m & 0 \\ 0 & (-1)^{m+1} \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \tag{5.42}$$

Line T: $\vec{P^q} = \{E, \sigma_{db}\} \to N(\vec{P^q}) = \{E, C_{2z}\} \times \{E, \sigma_{db}\}$
$$\to N^2(\vec{P^q}) = P^T = \{E, \sigma_x\} \circledS N(\vec{P^q})$$

Like in the previous case we choose the first way [(4.28), (4.29)] and start again from the vector unirreps of the normalizer $N(\vec{P^q})$. These vector unirreps are given by

D^μ: $\mu = 0, 1$

$$\sigma_{db} \to (-1)^\mu, \quad C_{2z} \to 1 , \tag{5.43}$$

D^μ: $\mu = 2, 3$

$$\sigma_{db} \to (-1)^\mu, \quad C_{2z} \to -1 . \tag{5.44}$$

According to (3.22) we examine again whether there exists a projective unirrep $\mathrm{B}^\mu(\sigma_x)$ satisfying

$$D^\mu(\sigma_x n \sigma_x)S(\sigma_x, \sigma_x n \sigma_x) = \mathrm{B}^\mu(\sigma_x)D^\mu(n)\mathrm{B}^\mu(\sigma_x)^\dagger$$
$$\text{for all } n \in N(\vec{P^q}) , \tag{5.45}$$

at which $\sigma_x\sigma_{db}\sigma_x = \sigma_{da}$, $\sigma_x C_{2z}\sigma_x = C_{2z}$, and (5.7) have to be taken into account. We obtain as results

$$N\{\mu\} = N(\vec{P^q}) \quad \text{for } \mu = 0, 1, \quad N\{\mu\} = P^T \quad \text{for } \mu = 2, 3 , \tag{5.46}$$

$$\Delta A_{N(\vec{P^q})} = \{0, 2, 3\} , \tag{5.47}$$

$$\mathrm{B}^\mu(\sigma_x) = 1 \quad \text{for } \mu = 2, 3$$
$$\text{with } P(\sigma_x, \sigma_x) = 1 \Leftrightarrow K(\sigma_x, \sigma_x) = -1 , \tag{5.48}$$

which implies that for $\mu = 2, 3$, projective unirreps of $P^T/N(\vec{P^q}) \simeq \{E, \sigma_x\}$ have to be determined. Denoting these projective unirreps by

$$\mathrm{D}^{\kappa}(\sigma_x) = i(-1)^{\kappa}, \quad \kappa = 0, 1 \tag{5.49}$$

we obtain by means of Eq. (3.64) four inequivalent one-dimensional projective unirreps of P^T, namely

$\mathrm{D}^{(\kappa,\mu)\dagger P^T}$: $\mu = 2, 3$, $\kappa = 0, 1$

$$\sigma_{db} \to (-1)^{\mu}, \quad C_{2z} \to -1, \quad \sigma_x \to i(-1)^{\kappa}. \tag{5.50}$$

For the case $\mu = 0$ we use the units [see Eq. (3.66)]

$$\mathrm{P}^{(0)\dagger P^T}_{E,E} = \mathrm{E}^0, \quad \mathrm{P}^{(0)\dagger P^T}_{\sigma_x,E} = \mathrm{V}(\sigma_x)\mathrm{E}^0 \tag{5.51}$$

to determine the corresponding two-dimensional projective unirrep of P^T [$\mathrm{E}^0 \in \mathbf{A}(N(\vec{P^q}))$],

$\mathrm{D}^{(0)\dagger P^T}$:

$$\sigma_{db} \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad C_{2z} \to \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad \sigma_x \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}, \tag{5.52}$$

which forms together with (5.50) a complete set of projective unirreps of P^T.

Line S: $\vec{P^q} = \{E, \sigma_{de}\} \to N(\vec{P^q}) = P^S = \{E, \sigma_y\} \times \vec{P^q}$

According to the first way [(4.28), (4.29)] we start from the vector unirreps of $\vec{P^q}$,

$$D^{\mu}: \quad \mu = 0, 1 \quad \sigma_{de} \to (-1)^{\mu}, \tag{5.53}$$

and obtain, by inspecting

$$D^{\mu}(n)S(\sigma_y, n) = \mathrm{B}^{\mu}(\sigma_y)D^{\mu}(n)\mathrm{B}^{\mu}(\sigma_y)^{\dagger}, \tag{5.54}$$

the following result:

$$N\{\mu\} = \vec{P^q} \quad \text{for } \mu = 0, 1 \Leftrightarrow \Delta A_{\vec{P^q}} = \{0\}. \tag{5.55}$$

Therefore, we have to use the units [see Eq. (3.66)]

$$\mathrm{P}^{(0)\dagger P^S}_{E,E} = \mathrm{E}^0, \quad \mathrm{P}^{(0)\dagger P^S}_{\sigma_y,E} = \mathrm{V}(\sigma_y)\mathrm{E}^0 \tag{5.56}$$

in order to obtain the two-dimensional projective unirrep of P^S.

$\mathrm{D}^{(0)\dagger P^S}$:

$$\sigma_{de} \to \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \sigma_y \to \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \tag{5.57}$$

Line Z: $\vec{P^q} = \{E, \sigma_z\} \to N(\vec{P^q}) = P^Z = \{E, \sigma_y\} \times \{E, \sigma_z\}$

In this case we proceed completely in the same way as before with the only difference that σ_{de} must be replaced by σ_z.

Plane RMX: $\vec{P^q} = \{E\} \to P^{\mathrm{plane}} = \{E, \sigma_y\}$

In this case we obtain immediately, because of Eq. (5.10), the two one-dimensional projective unirreps of P^{plane},

$$\mathrm{D}^{\kappa}(\sigma_y) = i(-1)^{\kappa}, \quad \kappa = 0, 1. \tag{5.58}$$

Now we are in the position to determine quite generally, by means of Eq. (4.25) together with the projective unirreps of the corresponding little cogroups, a complete set of vector unirreps of the nonsymmorphic space group $Pn3n$.

Finally we recall that the second possibility [as described by Eqs. (4.30) and (4.32)] is equally well suited to determine the projective unirreps of the little cogroups in question. However one must be aware that in general these projective unirreps can only be equivalent to those who are found according to the first way [(4.28), (4.29)] as can be seen from the following example:

Line S: $\vec{P^q} = \{E, \sigma_y\} \to N(\vec{P^q}) = P^S = \{E, \sigma_{de}\} \times \vec{P^q}$

Now we have to start the induction procedure from the projective unirreps (5.58) [where κ is replaced by μ] of P^{plane} and obtain by inspection of

$$\mathrm{D}^{\mu}(n)S(n, \sigma_{de}) = \mathrm{B}^{\mu}(\sigma_{de})\mathrm{D}^{\mu}(n)\mathrm{B}^{\mu}(\sigma_{de})^{\dagger} \quad \text{for all } n \in \{E, \sigma_y\} \tag{5.59}$$

the following result:

$$N\{\mu\} = \vec{P^q} \quad \text{for } \mu = 0, 1 \Leftrightarrow \Delta A_{\vec{P^q}} = \{0\}. \tag{5.60}$$

In this case we have to use the units

$$\mathrm{P}^{(0)\dagger P^S}_{E,E} = \mathrm{E}^0, \quad \mathrm{P}^{(0)\dagger P^S}_{\sigma_{de},E} = \mathrm{V}(\sigma_{de})\mathrm{E}^0 \tag{5.61}$$

in order to be able to determine the two-dimensional projective unirrep of P^S,

$\underline{\mathrm{D}}^{(0)\dagger P^S}$:

$$\sigma_y \to \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}, \quad \sigma_{de} \to \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \tag{5.62}$$

In comparing (5.62) with (5.57) we recognize that these projective unirreps of P^S are only equivalent. A simple calculation yields

$$W^{\dagger}\underline{\mathrm{D}}^{(0)\dagger P^S}(y)W = \mathrm{D}^{(0)\dagger P^S}(y) \quad \text{for all } y \in P^S$$
$$\text{with } W = \frac{1}{\sqrt{2}} \begin{vmatrix} 1 & i \\ -1 & i \end{vmatrix}. \tag{5.63}$$

VI. CONCLUDING REMARKS

It was the aim of this paper to investigate the problem of how the projective unirreps of a finite group G for a given standard factor system Q can be constructed by means of induction from the projective unirreps of a normal subgroup N of G. In order to make the induction procedure more apparent we emphasize the use of group algebraic techniques. Thereby we were able to show that

(i) the projective left- (right-) regular representation, respectively the corresponding representation of the left- (right-) group algebra $\mathbf{A}(G)$ of a finite group G for a given standard factor system Q can be defined in a consistent way;

(ii) the special standard factor system K (being an essential part of the induction procedure) which belongs to the factor group $N\{\mu\}/N$, can be determined quite generally only for the special case $n_{\mu} = 1$; and

(iii) that the units of $\mathbf{A}(G)$ and therefore the matrix elements of the induced projective unirreps of G are obtained quite generally.

However, besides this the key problem (for the cases $n_{\mu} > 1$), namely to determine in full generality the special standard factor system K, remains unsolved. This

implies that the factor system K has to be determined for every case separately.

Finally we recall that the described induction procedure offers an alternative method (and is really used in the example $Pn3n$) to determine by means of induction the projective unirreps of little cogroups of nonsymmorphic space groups out from unirreps of other little cogroups which are subgroups of the original one. These induced projective unirreps have the useful property of decomposing (up to equivalence) immediately into direct sums of projective unirreps of the corresponding subgroup, if the subduction is carried out. Clearly, this is important for physical applications, since such subductions arise, if investigating the "compatibility relations" for space groups.[22,14,23] A further application which we have in mind is that the determination of Clebsch–Gordan coefficients for nonsymmorphic space groups[24,25,5,7,26-28] could become simpler, if one used projective unirreps of the little cogroups which can be found by the proposed method.

ACKNOWLEDGMENTS

I want to thank Professor Simon Altmann for introducing me to this subject by his excellent lectures given in Vienna in April 1975 and for making available to me a preliminary version of his book on induced representations. Furthermore I am indebted to Professor Rhoda Berenson, Professor Joseph Birman, and Professor Jan Morzymas for valuable discussions concerning this subject. Finally I am indebted to Professor van den Broek for calling may attention to an erroneous proposition in the first version of this work.

[1]G. W. Mackey, Acta Mathematica **99**, 265 (1958).
[2]G. W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1968).
[3]Ch. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Wiley, New York, 1962).
[4]O. V. Kovalev, *Irreducible Representations of Space Groups* (Gordon and Breach, New York, 1965).
[5]J. L. Birman, in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).
[6]C. J. Bradley and A. P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[7]R. Berenson and J. L. Birman, J. Math. Phys. **16**, 227 (1975).
[8]M. Hamermesh, *Group Theory and Its Applications to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[9]S. L. Altmann, "Induced Representations in Crystals and Molecules" (Academic, London, 1977, in press).
[10]S. L. Altmann and C. J. Bradley, Rev. Mod. Phys. **37**, 33 (1965).
[11]R. Car, G. Ciucci, and L. Quartabelle, J. Math. Phys. **17**, 1051 (1976).
[12]J. Zak, J. Math. Phys. **1**, 165 (1960).
[13]L. T. Klauder, Jr. and J. G. Gay, J. Math. Phys. **9**, 1488 (1968).
[14]J. F. Cornwell, *Group Theory and Electronic Energy Bands in Solids* (North-Holland, Amsterdam, 1969).
[15]A. J. Coleman, in *Group Theory and Its Applications*, edited by M. Loebl (Academic, New York, 1968).
[16]L. Jansen and M. Boon, *Theory of Finite Groups. Applications in Physics* (North-Holland, Amsterdam, 1967).
[17]R. Dirl and P. Kasperkovitz, *Gruppentheorie, Anwendungen in der Atom—und Festkörperphysik* (Vieweg, Braunschweig, 1977).
[18]R. Dirl, J. Phys. A **9**, 829, 843 (1976).
[19]P. Kasperkovitz and R. Dirl, J. Math. Phys. **15**, 1203 (1974).
[20]D. J. Klein, in *Group Theory and Its Applications*, edited by M. Loebl (Academic, New York, 1975), Vol. II.
[21]J. C. Slater, Rev. Mod. Phys. **37**, 68 (1965).
[22]L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. **50**, 58 (1936).
[23]P. Kasperkovitz and R. Dirl, Acta Phys. Austriaca **42**, 57, 202 (1975).
[24]D. B. Litvin and J. Zak, J. Math. Phys. **9**, 212 (1968).
[25]L. K. Saulevich, D. T. Sviridov, and Yu. F. Smirnov, Sov. Phys.—Crystall. **15**, 355 (1970).
[26]R. Berenson, I. Itzkan, and J. L. Birman, J. Math. Phys. **16**, 236 (1975).
[27]P. Gard, J. Phys. A **6**, 1837 (1973).
[28]A. P. Cracknell and B. L. Davies, in *Lecture Notes in Physics*, edited by A. Janner, T. Boon, and M. Boon (Springer, Berlin, 1976), Vol. 50.

Coherent states associated with the continuous spectrum of noncompact groups*

M. Hongoh

Department of Physics, University of Montreal, Montréal, Québec, Canada
(Received 12 July 1976)

A new, explicit formula is obtained for the coherent states associated with a continuous spectrum of the noncompact group SU(1,1). The method is based on a simple and unified algebraic approach. We briefly discuss its relations to the generalized coherent states of Barut and Girardello, and of Perelomov.

I. INTRODUCTION

The notion of coherent states was first introduced by Weyl for the nilpotent group.[1] Recently it has been extended to any Lie group.[2,3] Barut and Girardello[4] have constructed in particular the coherent states associated with the discrete series of representations of SO(2, 1). They are the eigenstates of the lowering and raising operators of noncompact groups. On the other hand, Perelomov has generalized the coherent states in such a way that a set of coherent states is invariant relative to the action of the group representation operators.

It has been suggested since,[2,4] that one extend the notion of coherent states for the continuous spectrum corresponding to the infinite-dimensional unitary representations of the noncompact groups. But so far no explicit construction seems to have appeared in the literature. The purpose of this paper is to present a general method for constructing such states. In particular we deal with the continuous basis of the simplest semisimple Lie algebra of SU(1, 1). Explicit coherent states are then constructed on this basis, which are somewhat analogous to the Bloch coherent states of SU(2).

In atomic physics, coherent states associated with a system of spins are known as the Bloch coherent states.[5,6] They can be obtained by rotating the lowest angular momentum state (called the lowest Dicke state) in the manifold of the group of the angular momentum; they provide natural and useful bases for various calculations of superradiance and superconductivity.[7,8] Moreover they possess a unique property in that they represent special quantum states most closely approximating classical states, e.g., the uncertainty relation takes its minimum value for these states. The Bloch coherent states of SU(2) form a subsystem of the systems of generalized coherent states of Perelomov. Thus, they are invariant relative to the action of the group representation operators. We might call the generalization of this particular system of coherent states a *Bloch-type* system, and in the text we shall restrict ourselves to this particular system of coherent states. In Sec. II we demonstrate our method by deriving the coherent states for the discrete spectrum of SU(1, 1), and we compare the results with those of Barut and Girardello. In Sec. III, we construct an explicit formula for the coherent states associated with the continuous spectrum, and Sec. IV is devoted to discussion and to some further remarks.

II. COHERENT STATES ASSOCIATED WITH THE DISCRETE SPECTRUM OF SU(1, 1)[17]

In connection with the general scalar coefficients, Van der Waerden[9,10] introduced an invariant form for the compact SU(2) group,

$$W=\prod_i \Lambda_i^{a_i} \tag{2.1}$$

where $\Lambda_i=(\eta_j\xi_k-\eta_k\xi_j)$, $i,j,k=1,2,3$ cyclic. a_i are nonnegative integers. η and ξ are the two fundamental complex variables of the group, and i refers to the three distinct representation spaces. Λ_i are called elementary scalars.

The skew-symmetric form of Λ_i reflects the following theorem[11-14]:

Malcev-Dynkin Theorem: The self contragradient IR (SCIR) of the connected semisimple Lie groups leaves the nondegenerate bilinear form invariant which is symmetric (skew-symmetric) according to whether the SCIR is orthogonal (symplectic).

All the IR's of SU(2) are SC[15] and the smallest SCIR is even-dimensional. Letting $a_1=a_2=0$ and $a_3=2S$, the SU(2) Bloch coherent states can be derived.[16]

We shall now extend the notion of the Van der Waerden invariant for noncompact groups, i.e., let

$$\Lambda_{\text{n.c.}}\equiv(\eta_1\xi_2-\eta_2\xi_1)^{2S}\,,\quad 2S\text{: real number}. \tag{2.2}$$

The Bargmann–Schwinger realization of the SU(1, 1) Lie algebra is well known,

$$\begin{aligned} &L_+=\epsilon\eta\partial_\xi\,,\quad L_-=\epsilon\xi\partial_\eta\,,\\ &L_3=\tfrac{1}{2}(\eta\partial_\eta-\xi\partial_\xi)\,,\quad \epsilon=+\sqrt{-g_{33}}\,, \end{aligned} \tag{2.3}$$

and the basis functions are,[17]

$$g_m(Z)=\frac{(-g_{33})^{\phi-m}\Gamma(2\phi+1)}{\Gamma(\phi+m+1)\Gamma(\phi-m+1)}Z_1^{\phi+m}Z_2^{\phi-m}\,, \tag{2.4}$$

where $g_{33}=-1$ for SU(2) ~ O(3), and $+1$ for SU(1, 1) ~ O(2, 1). One can easily see that $\{Z,Z\}\equiv|Z_1|^2-g_{33}|Z_2|^2$ is preserved by the fundamental representation of the group. Let us consider the following function,

$$F(W,Z)=\sum_{m=0}^{\infty} g_m^*(W)g_m(Z)\,. \tag{2.5}$$

F can be viewed as a scalar product of two basis vectors which belong to the two distinct Hilbert spaces of IR characterized by 2ϕ. In fact,

$$F(W,Z)=\{W,Z\}^{2\phi}\,. \tag{2.6}$$

Comparing Eq. (2.6) and Eq. (2.2) with $2S \to 2\phi$, we establish the correspondences

$$\begin{pmatrix}\eta\\ \xi\end{pmatrix}_1 \to \begin{pmatrix}Z_1\\ Z_2\end{pmatrix}, \quad \begin{pmatrix}\xi\\ \eta\end{pmatrix}_2 \to \begin{pmatrix}W_\perp\\ -g_{33}W_2\end{pmatrix}^* . \tag{2.7}$$

Let us rewrite $W_1^*, W_2^* \to \nu, \mu$ respectively. Then Eq. (2.7) means that the operators in Eq. (2.3), once operated upon $W(\nu, \mu)$ space, would be replaced by,[18]

$$\begin{aligned} &L_+ \to \epsilon\mu\partial_\nu, \quad L_- \to \epsilon\nu\partial_\mu, \\ &L_3 \to \tfrac{1}{2}(\mu\partial\mu - \nu\partial\nu). \end{aligned} \tag{2.8}$$

We are interested in the coherent states defined over the homogeneous factor space $G/H \sim \mathrm{SU}(1,1)/\mathrm{SO}(2)$. Let Z be a vector belonging to the Hilbert space of the set of pure states. This Hilbert space is isomorphic to an upper sheet of the hyperboloid. Without a loss of generality we can set,

$$\{Z, Z\}^{2\phi} = 1. \tag{2.9}$$

The completeness relation for the basis function $g_n(Z) \equiv |\phi n\rangle$ is written

$$F(Z, Z) = \sum_{n=0}^{\infty} |\phi n\rangle\langle\phi n| = 1. \tag{2.10}$$

The homogeneous factor space is a unit disc that can be parametrized by a single complex variable, say ω.[19] We fix our parametrization by letting $w \in G/H$, and $\nu \to \omega$, $\mu \to 1$, respectively. We have

$$F(W, W) = (1 - g_{33}|\omega|^2)^{2\phi}. \tag{2.11}$$

Note that the fundamental length $(1 - g_{33}|\omega|^2)$ is preserved in the operation of the group. Substituting (2.4) for (2.6) with $\phi + m \to N$, $\phi - m \to 2\phi - N$, respectively, and using the binomial expansion,[20] we obtain

$$|\omega\rangle = \sum_{N=0}^{\infty} \left[\frac{(-g_{33})^{2\phi-N}\Gamma(2\phi+1)}{\Gamma(N+1)\Gamma(2\phi-N+1)}\right]^{1/2} \omega^N |\phi N\rangle, \tag{2.12}$$

where the normalization factor is ignored. Note that $|\omega\rangle$ is not identical with the SU(1, 1) coherent states given by Barut and Girardello,[4] in which L_- (L_+) is diagonalized. In fact, their coherent states are

$$|Z\rangle_{\mathrm{B.G.}} = [\Gamma(-2\phi)]^{1/2} \sum_{n=0}^{\infty} \frac{(\sqrt{2}\,Z)^n}{[\Gamma(n+1)\Gamma(-2\phi+n)]^{1/2}} |\phi n\rangle. \tag{2.13a}$$

It can be seen easily that the following holds for $|Z\rangle_{\mathrm{B.G.}}$,

$$L_-|Z\rangle_{\mathrm{B.G.}} = Z|Z\rangle_{\mathrm{B.G.}}. \tag{2.13b}$$

On the other hand, we have

$$L_-|\omega\rangle = \sum_N \left[\frac{(-g_{33})^{2\phi-N}\Gamma(2\phi+1)}{\Gamma(N+1)\Gamma(2\phi-N+1)}\right]^{1/2} \omega^{N+1}(2\phi - N)|\phi N\rangle, \tag{2.14}$$

i.e., $|\omega\rangle$ fully retains the property of the Bloch coherent states. Equation (2.11) gives the norm of $|\omega\rangle$, and the nonorthogonality is written

$$\langle\omega'|\omega\rangle = (1 - g_{33}\omega'^*\omega)^{2\phi}. \tag{2.15}$$

With the resolution of unity,

$$\int d\sigma(\omega)|\omega\rangle\langle\omega| = \sum_{N=0}^{\infty} |\phi N\rangle\langle\phi N| = 1, \tag{2.16}$$

one can write

$$\langle f|i\rangle = \int d\sigma(\omega)\langle f|\omega\rangle\langle\omega|i\rangle, \tag{2.17}$$

where $|f\rangle$ and $|i\rangle$ belong to the Hilbert space of IR's of the vectors of G/H. Letting $d\sigma(\omega) = \sigma(r)rdrd\phi$, $|\omega| = r$, the rhs is,

$$2\pi \int_0^{\infty} dr\,\sigma(r)r^{2N+1} \times \sum_N \frac{\Gamma(-2\phi+1)}{\Gamma(N+1)\Gamma(-2\phi+N+1)} \langle f|\phi N\rangle\langle\phi N|i\rangle. \tag{2.18}$$

Due to Eq. (2.16), $\sigma(r)$ satisfies

$$\int_0^{\infty} dr\,\sigma(r)r^{2N+1} = \frac{\Gamma(N+1)\Gamma(-2\phi+N+1)}{2\pi\Gamma(-2\phi+1)}, \tag{2.19}$$

and via the Mellin transform, we obtain

$$\sigma(r) = \frac{2r^{-2\phi}K_\phi(2r)}{\pi\Gamma(-2\phi+1)} > 0, \quad r > 0, \tag{2.20a}$$

where

$$K_\phi(2r) = \frac{\pi}{2} \cdot \frac{I_{-\phi}(2r) - I_{+\phi}(2r)}{\sin(\phi\pi)} \tag{2.20b}$$

and

$$I_\phi(2r) = \sum_{m=0}^{\infty} \frac{r^{\phi+2m}}{\Gamma(m+1)\Gamma(\phi+m+1)}. \tag{2.20c}$$

$K_\phi(2r)$ is the modified Bessel function of the third kind, and $I_\phi(2r)$ is the modified Bessel function of the first kind. Equations (2.20) are the counterparts of the formulas given in Ref. 4. Equations (2.12), (2.16), and (2.15) completely specify the system of the Bloch-type coherent states.

III. COHERENT STATES ASSOCIATED WITH THE CONTINUOUS SPECTRUM OF SU(1, 1)

Let us consider the SU(1, 1) algebra in which L_{23} is diagonalized.[21]

$$\begin{aligned} L_{12} &= (i/\sqrt{2})(\partial_d - d\partial_u), \\ L_{13} &= (\epsilon/\sqrt{2})(u\partial_d + d\partial_u), \\ L_{23} &= (\epsilon/\sqrt{2})(u\partial_u - d\partial_d), \end{aligned} \tag{3.1}$$

where $u = (1/\sqrt{2})(\eta + i\xi)$, $d = (1/\sqrt{2})(\xi + i\eta)$. The basis functions are,

$$f^\alpha(Z) = A_\alpha u^{\phi+\epsilon\alpha} d^{\phi-\epsilon\alpha}, \quad Z(u, d), \tag{3.2}$$

where α is in general a complex number. Our task is to construct the coherent states defined over $G/H \sim \mathrm{SU}(1,1)/\mathrm{O}(1,1)$ as a certain linear combination of $f^\alpha(Z)$. Clearly it suffices to consider rotations around the third axis,

$$R = e^{i\theta L_{12}} = \begin{pmatrix} e^{i\theta/2} & \\ & e^{-i\theta/2} \end{pmatrix} \tag{3.3}$$

which preserve the form $\{|u|^2 + |d|^2\}$ for both compact and noncompact cases. Defining the state conjugate to $|\phi\alpha\rangle \equiv f^\alpha(Z)$,

$$\langle\phi\alpha| = A_\alpha(u^*)^{\phi+\epsilon\alpha}(d^*)^{\phi-\epsilon\alpha}, \tag{3.4}$$

the orthogonality of the basis function can be expressed as follows:

$$\langle \phi\alpha' | \phi\alpha \rangle = \delta(\alpha' - \alpha). \tag{3.5}$$

The normalization constant A_α can be obtained by using the complex binomial expansion of the invariant form,

$$|A_\alpha|^2 = \frac{1}{2\pi} \frac{\Gamma(-\phi+\epsilon\alpha)\Gamma(-\phi-\epsilon\alpha)}{\Gamma(-2\phi)} \geq 0. \tag{3.6}$$

Let us introduce a function

$$\mathcal{J}(W,Z) = \int_{-(i/\epsilon)\infty}^{+(i/\epsilon)\infty} d\alpha f^\alpha(W)^* f^\alpha(Z), \tag{3.7}$$

which is the continuum version of Eq. (2.5). One can see immediately that

$$\mathcal{J}(Z,Z) = \{|u|^2 + |d|^2\}^{2\phi}. \tag{3.8}$$

We now consider the Hilbert space which corresponds to an upper half of the hyperboloid lying along the first axis. Let $Z(u,d)$ belong to this Hilbert space, and we set

$$\{|u|^2 + |d|^2\} = 1. \tag{3.9}$$

Then from Eqs. (3.2) and (3.7) we obtain,

$$\mathcal{J}(W,Z) = \int d\alpha |A_\alpha|^2 (W_1^*)^{\phi+\epsilon\alpha} (W_2)^{\phi-\epsilon\alpha} | \phi\alpha \rangle. \tag{3.10}$$

Due to the correspondences (2.6), we replace w_1^*, $w_2^* \to \theta, \rho$ respectively; where ρ, θ are complex variables similar to u, d. The factor space may be projected onto a unit disc perpendicular to the first axis, which can be parametrized by a single complex variable. This amounts to having $\theta \to \theta$, $\rho \to 1$, and the coherent states are,

$$|\theta\rangle = \int_{-(i/\epsilon)\infty}^{+(i/\epsilon)\infty} d\alpha \left[\frac{\Gamma(-\phi+\epsilon\alpha)\Gamma(-\phi-\epsilon\alpha)}{2\pi\Gamma(-2\phi)} \right]^{1/2} \theta^{\phi+\epsilon\alpha} | \phi\alpha \rangle. \tag{3.11}$$

To the best of our knowledge, this expression has not appeared in the literature before. The nonorthogonality of $|\theta\rangle$ is,

$$\langle \theta' | \theta \rangle = \frac{1}{2\pi} \int d\alpha \frac{\Gamma(-\phi+\epsilon\alpha)\Gamma(-\phi-\epsilon\alpha)}{\Gamma(-2\phi)} (\theta'^*\theta)^{\phi+\epsilon\alpha}. \tag{3.12}$$

$\Gamma(-\phi+\epsilon\alpha)$ and $\Gamma(-\phi-\epsilon\alpha)$ have poles at $\alpha = y + i(N-x)$ and at $-y + i(x-N)$ respectively (x and y are the real and imaginary parts of ϕ).

Integrating clockwise along the contour and summing contributions from poles of $\Gamma(-\phi-\epsilon\alpha)$, we obtain

$$\langle \theta' | \theta \rangle = \frac{1}{\Gamma(-2\phi)} \sum_{N=0}^{\infty} \frac{(-)^N \Gamma(-2\phi+N)}{\Gamma(N+1)} (\theta'^*\theta)^N (\rho'^*\rho)^{2\phi-N} \Big|_{\rho,\rho'=1}. \tag{3.13}$$

Note the rhs is the complex binomial expansion. Thus

$$\langle \theta' | \theta \rangle = (1 + \theta'^*\theta)^{2\phi}, \quad |\theta'^*\theta| < 1, \quad |\arg(-\theta'^*\theta)| < \pi. \tag{3.14}$$

In particular the norm of $|\theta\rangle$ is,

$$\| |\theta\rangle \|^2 = (1 + |\theta|^2)^{2\phi}. \tag{3.15}$$

Dividing $|\theta\rangle$ by the square root of its norm, we define

$$|\tilde{\theta}\rangle = (1 + |\theta|^2)^{-\phi} |\theta\rangle.$$

Then the completeness relation can be written as

$$\langle f | i \rangle = \int d\sigma(\theta) \langle f | \tilde{\theta} \rangle \langle \tilde{\theta} | i \rangle. \tag{3.16}$$

With $d\sigma(\theta) = \sigma(r) r dr d\phi$ we have,

$$\langle f | i \rangle = \int d\alpha \langle f | \phi\alpha \rangle \langle \phi\alpha | i \rangle$$
$$= \frac{2\pi}{(1+|\theta|^2)^{2\phi}} \int_0^\infty \sigma(r) dr$$
$$\times \int_{-\infty}^{+\infty} d\alpha |A_\alpha|^2 r^{2(\phi+\epsilon\alpha)} \langle f | \phi\alpha \rangle \langle \phi\alpha | i \rangle. \tag{3.17}$$

Inserting the beta function for $|A_\alpha|^2$

$$B(-\phi+\epsilon\alpha, -\phi-\epsilon\alpha) = \int_0^\infty dr \frac{2r^{-2\phi-2\epsilon\alpha-1}}{(1+r^2)^{-2\phi}}, \tag{3.18}$$

we obtain

$$2\int_0^\infty dr\, \sigma(r) \frac{1}{r} \int_{-\infty}^{+\infty} d\alpha \langle f | \phi\alpha \rangle \langle \phi\alpha | i \rangle = \langle f | i \rangle, \tag{3.19}$$

where $\sigma(r)$ is,

$$\sigma(r) = (r/2)\delta(r), \quad r \geq 0, \tag{3.20a}$$

or

$$= \frac{r}{\sqrt{2}} \exp(-r^2), \quad r \geq 0. \tag{3.20b}$$

Equation (3.20) together with Eqs. (3.11) and (3.16) completely specify the system of the coherent states associated with the continuous spectrum.

IV. CONCLUDING REMARKS

A simple method is introduced for constructing the explicit *Bloch-type* coherent states for the UIR of noncompact Lie groups. In particular we have studied the coherent states associated with both the discrete and the continuous spectra of SU(1, 1). The method might be useful in constructing the general coherent state representations for other noncompact groups which appear frequently in the applications to physical problems. This will be discussed elsewhere.

ACKNOWLEDGMENTS

The author wishes to thank Professor J. Patera for drawing his attention to the Malcev-Dynkin theorem, and Professor J. Destry for reading the manuscript. It is also a pleasure to express his gratitude to the Centre de Recherches Mathématiques, Université de Montreal for their hospitality, where the present work was partially completed.

*Work supported in part by the National Research Council of Canada.

[1] H. Weyl, *Gruppentheorie und Quantenmechanik* (Leipzig, S. Hirzel, 1928).
[2] A. M. Perelomov, Commun. Math. Phys. **26**, 222 (1972).
[3] E. Onofri, J. Math. Phys. **16**, 1078 (1975).
[4] A. O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1971).
[5] J. M. Radcliffe, J. Phys. A (London) **4**, 313 (1971).
[6] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6**, 2211 (1972).
[7] F. T. Hioe, Phys. Rev. A **8**, 1440 (1973).

[8]F.T. Hioe, J. Math. Phys. **15**, 445 (1974).
[9]B. L. Van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Berlin, Springer, 1932).
[10]M. Hongoh, R.T. Sharp, and D.E. Tilley, J. Math. Phys. **15**, 782 (1974).
[11]A.I. Malcev, Am. Math. Soc. Transl. Ser. 1, **9**, 172 (1962).
[12]E.B. Dynkin, Dokl. Akad. Nauk SSSR (N.S.) **71**, 221 (1950).
[13]E.B. Dynkin, Dokl. Akad. Nauk SSSR (N.S.) **76**, 629 (1951).
[14]A.K. Bose and J. Patera, J. Math. Phys. **11**, 2231 (1970).
[15]The two vectors of Λ_i can be non-SCIR, since they belong to two distinct representation spaces. For example, the smallest nontrivial orthogonal SCIR of SU(3) is $D(1,1)$ (octet) and the symmetric Van der Waerden invariant consists of vectors of $D(10)$ and $D(01)$, which separately are not SCIR.
[16]M. Hongoh, notes on the SU(2) Bloch coherent states, Univ. of Montreal preprint (unpublished).
[17]Throughout the work we shall adopt the notation of A.O. Barut, *Lectures in Theorectical Physics* (Gordon and Breach, New York, 1967), Vol. IXA, pp. 125–71.
[18]To adopt the phase convention used in Z space, the operators acting in w space must be of this form.
[19]For the lowest state, $|\phi n=0\rangle$, the stabilizer is SO(2) and the corresponding orbit is SU(1,1)/SO(2).
[20]The use of the (complex) binomial expansion imposes a certain restriction on the values that ϕ and E_0 can take. While this causes no trouble for discrete cases $D^{\pm}(\phi)$, the general cases can be handled by going to the covering group SU(1,1) $\otimes Z$. See Ref. 9.
[21]A. O. Barut and E. C. Phillips, Commun. Math. Phys. **8**, 52 (1968).

Comments on "Periods on manifolds, quantization, and gauge" by R. M. Kiehn [J. Math. Phys. 18, 614 (1977)]

E. J. Post

8606–6 Pershing Drive, Plaza del Rey, California 90291
(Received 20 April 1977)

In a recent paper entitled "Periods on manifolds, quantization, and gauge," Kiehn[1] suggests an interpretation of flux, charge, and angular momentum quantization in terms of a set of, in principle, independent concepts relating to the periods of one-, two-, and three-dimensional cyclic integrals in space–time. In addition to the very interesting aspects cited by the author, permit me to delineate, what seems, an emerging interrelation with past and ongoing work and the consequences thereof:

I. Since periods are topological invariants having the properties of numbers, the magnitude of these numbers must be taken to be totally independent of the choice of space–time references. It then follows that there is no meaning to an inquiry with the objective of establishing a time and position dependence of these periods. Hence the periods, i.e., the quantum of action and the quantum of electric charge, are so declared to be metric independent quantities.

II. The central premise of topological invariance, for the laws of nature so expressed in terms of period integrals, lends a new perspective to the earlier work of Kottler,[2] Cartan,[3] and van Dantzig[4] on the metric independent invariances of certain laws of nature. A more detailed evaluation of this interrelation is given in a forthcoming paper entitled "Uncertainty and metric structure."[5]

III. The elevation of flux quantization to a truly independent fundamental law rules out the hypothesis of magnetic charge as incompatible with such change in basic assumptions. The new hypothesis in fact establishes a perspective on the recent work of Jehle[6] on flux quantization and particle structure.

IV. While the occurrence of one- and two-dimensional period integrals in physics have been a matter of awareness for some time, Kiehn adds a new item in the form of a three-dimensional period integral related to a differential 3-form of action density and action flux. The need for such a "three index" field was earlier established by Belinfante[7] in local field theory, and in the context of continuum dynamics by Mindlin.[8] Recently the local space–time aspects of that field have been studied in their relation to spaces with torsion by Hehl[9] and co-workers. It seems unavoidable that a concomitant of the latter field should be related to Kiehn's differential 3-form S.

V. The introduction of nontrivial three-dimensional periods raises the question of a possible decomposition, if the three-dimensional cyclic domains C_3 in question can be viewed as the topological product of physically meaningful lower dimensional cyclic domains C_1 and C_2, Kiehn's relation (19),

$$\oint_{C_3} A \wedge H = \oint_{C_1} A \oint_{C_2} H,$$

where

A = 1-form of flux: vector four-potential,

H = 2-form of electric charge: fields **D** and **H**,

thus permits a conceivable cross check with Jehle's fundamental assumptions on flux quantization; provided the topological configurations of C_1, C_2, and C_3 can be known.

[1]R. M. Kiehn, J. Math. Phys. **18**, 614 (1977).
[2]F. Kottler, Sitz. Akad. Wiss. Wein IIa **131**, 119 (1922).
[3]E. Cartan, Ann. Ec. Norm. Super. **41**, 1 (1924).
[4]D. van Dantzig, Proc. Cambridge Philos. Soc. **30**, 421 (1934); Amsterdam Ac. **37**, 521, 526, 644, 825 (1934).
[5]E. J. Post, "Uncertainty and Metric Structure" Scientia, Milan, forthcoming.
[6]H. Jehle, Phys. Rev. D **3**, 306 (1971); **6**, 441 (1972); **11**, 2147 (1975).
[7]F. J. Belinfante, Physica **6**, 887 (1939); **7**, 449 (1940).
[8]R. D. Mindlin, Arch. Rat. Mech. Ann. **16**, 51 (1964).
[9]F. W. Hehl, Rep. Math. Phys. **9**, 55 (1976); F. W. Hehl, P. von der Heyde, G. O. Kerlick, and J. M. Nester, Rev. Mod. Phys. **48**, 393 (1976).

[8]F.T. Hioe, J. Math. Phys. **15**, 445 (1974).
[9]B. L. Van der Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Berlin, Springer, 1932).
[10]M. Hongoh, R.T. Sharp, and D.E. Tilley, J. Math. Phys. **15**, 782 (1974).
[11]A.I. Malcev, Am. Math. Soc. Transl. Ser. 1, **9**, 172 (1962).
[12]E.B. Dynkin, Dokl. Akad. Nauk SSSR (N.S.) **71**, 221 (1950).
[13]E.B. Dynkin, Dokl. Akad. Nauk SSSR (N.S.) **76**, 629 (1951).
[14]A.K. Bose and J. Patera, J. Math. Phys. **11**, 2231 (1970).
[15]The two vectors of Λ_i can be non-SCIR, since they belong to two distinct representation spaces. For example, the smallest nontrivial orthogonal SCIR of SU(3) is $D(1,1)$ (octet) and the symmetric Van der Waerden invariant consists of vectors of $D(10)$ and $D(01)$, which separately are not SCIR.
[16]M. Hongoh, notes on the SU(2) Bloch coherent states, Univ. of Montreal preprint (unpublished).
[17]Throughout the work we shall adopt the notation of A.O. Barut, *Lectures in Theorectical Physics* (Gordon and Breach, New York, 1967), Vol. IXA, pp. 125–71.
[18]To adopt the phase convention used in Z space, the operators acting in w space must be of this form.
[19]For the lowest state, $|\phi n=0\rangle$, the stabilizer is SO(2) and the corresponding orbit is SU(1,1)/SO(2).
[20]The use of the (complex) binomial expansion imposes a certain restriction on the values that ϕ and E_0 can take. While this causes no trouble for discrete cases $D^{\pm}(\phi)$, the general cases can be handled by going to the covering group SU(1,1) $\otimes Z$. See Ref. 9.
[21]A. O. Barut and E. C. Phillips, Commun. Math. Phys. **8**, 52 (1968).

Comments on "Periods on manifolds, quantization, and gauge" by R. M. Kiehn [J. Math. Phys. 18, 614 (1977)]

E. J. Post

8606–6 Pershing Drive, Plaza del Rey, California 90291
(Received 20 April 1977)

In a recent paper entitled "Periods on manifolds, quantization, and gauge," Kiehn[1] suggests an interpretation of flux, charge, and angular momentum quantization in terms of a set of, in principle, independent concepts relating to the periods of one-, two-, and three-dimensional cyclic integrals in space–time. In addition to the very interesting aspects cited by the author, permit me to delineate, what seems, an emerging interrelation with past and ongoing work and the consequences thereof:

I. Since periods are topological invariants having the properties of numbers, the magnitude of these numbers must be taken to be totally independent of the choice of space–time references. It then follows that there is no meaning to an inquiry with the objective of establishing a time and position dependence of these periods. Hence the periods, i.e., the quantum of action and the quantum of electric charge, are so declared to be metric independent quantities.

II. The central premise of topological invariance, for the laws of nature so expressed in terms of period integrals, lends a new perspective to the earlier work of Kottler,[2] Cartan,[3] and van Dantzig[4] on the metric independent invariances of certain laws of nature. A more detailed evaluation of this interrelation is given in a forthcoming paper entitled "Uncertainty and metric structure."[5]

III. The elevation of flux quantization to a truly independent fundamental law rules out the hypothesis of magnetic charge as incompatible with such change in basic assumptions. The new hypothesis in fact establishes a perspective on the recent work of Jehle[6] on flux quantization and particle structure.

IV. While the occurrence of one- and two-dimensional period integrals in physics have been a matter of awareness for some time, Kiehn adds a new item in the form of a three-dimensional period integral related to a differential 3-form of action density and action flux. The need for such a "three index" field was earlier established by Belinfante[7] in local field theory, and in the context of continuum dynamics by Mindlin.[8] Recently the local space–time aspects of that field have been studied in their relation to spaces with torsion by Hehl[9] and co-workers. It seems unavoidable that a concomitant of the latter field should be related to Kiehn's differential 3-form S.

V. The introduction of nontrivial three-dimensional periods raises the question of a possible decomposition, if the three-dimensional cyclic domains C_3 in question can be viewed as the topological product of physically meaningful lower dimensional cyclic domains C_1 and C_2, Kiehn's relation (19),

$$\oint_{C_3} A \wedge H = \oint_{C_1} A \oint_{C_2} H,$$

where

A = 1-form of flux: vector four-potential,

H = 2-form of electric charge: fields **D** and **H**,

thus permits a conceivable cross check with Jehle's fundamental assumptions on flux quantization; provided the topological configurations of C_1, C_2, and C_3 can be known.

[1]R. M. Kiehn, J. Math. Phys. **18**, 614 (1977).
[2]F. Kottler, Sitz. Akad. Wiss. Wein IIa **131**, 119 (1922).
[3]E. Cartan, Ann. Ec. Norm. Super. **41**, 1 (1924).
[4]D. van Dantzig, Proc. Cambridge Philos. Soc. **30**, 421 (1934); Amsterdam Ac. **37**, 521, 526, 644, 825 (1934).
[5]E. J. Post, "Uncertainty and Metric Structure" Scientia, Milan, forthcoming.
[6]H. Jehle, Phys. Rev. D **3**, 306 (1971); **6**, 441 (1972); **11**, 2147 (1975).
[7]F. J. Belinfante, Physica **6**, 887 (1939); **7**, 449 (1940).
[8]R. D. Mindlin, Arch. Rat. Mech. Ann. **16**, 51 (1964).
[9]F. W. Hehl, Rep. Math. Phys. **9**, 55 (1976); F. W. Hehl, P. von der Heyde, G. O. Kerlick, and J. M. Nester, Rev. Mod. Phys. **48**, 393 (1976).

ERRATA

Erratum: Calculation of the Molien generating function for invariants of space groups [J. Math. Phys. 18, 1459 (1977)]

Marko V. Jarić and Joseph L. Birman

Physics Department, City College-CUNY, New York, New York 10031
(Received 18 July 1977)

The Molien generating function for irreducible representations *X3 and *X4, given in Table III should read:

$$\frac{1+2z^4+5z^6+11z^8+9z^{10}+11z^{12}+6z^{14}+3z^{16}}{(1-z^2)(1-z^4)^3(1-z^6)^2}.$$

Thus the Molien function for *X3 and *X4 is identical to that for *R4.

CUMULATIVE AUTHOR INDEX

All authors published so far in the current volume are listed alphabetically with the issue and page numbers following the dash. A cumulative author and subject index covering each volume is published annually. An (E) *after the page number indicates an Erratum.*

