Renormalization and Ward identities using complex space-time dimension

Eugene R. Speer

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
(Received 27 July 1973)

Complex-dimensional renormalization is defined for an arbitrary Feynman amplitude and shown to be equivalent to BPH renormalization. Using quantum electrodynamics as an example, Ward identities are proved; here Carlson's theorem extends the identities from integer to complex dimension. Both complex dimensional and analytic regularization are necessary at intermediate stages.

In Ref. 1, t'Hooft and Veltman propose a renormalization method based on the generalization of the dimension of space–time to a complex number. (Such a generalization was previously proposed by Regge[2-5] to discuss analytic properties of Feynman amplitudes.) In this paper we verify that the method indeed gives a renormalization (proved in Ref. 1 for graphs continuing up to two loops; see also Ref. 6) and use Carlson's theorem to show that it maintains Ward identities. (We will actually discuss Ward identities in the case of QED, but the methods used are in no way special.) We define an intermediate regularization which includes both complex dimension and the usual $\boldsymbol{\lambda}$ regularization of analytic renormalization, to give amplitudes which are well defined in all integer dimensions; when all the λ parameters are set to 1, the regularization of Ref. 1 is recovered.

1. REGULARIZATION

We suppose given a connected Feynman graph G with m vertices $V_1, \ldots, V_m$ and L lines $\{l_1, \ldots, l_L\} = \mathcal{L}$; $h = L - m + 1$ will denote the number of loops of G. [When other graphs enter the discussion we write $m(G), \mathcal{L}(G')$, etc.] The line $l \in \mathcal{L}$ has initial vertex V_{i_l} and final vertex V_{f_l}, and we associate with l a complex variable λ_l; the point $\lambda^\circ \in \mathbb{C}^L$ is specified by $\lambda_l = 1$, all l. (Multiples of variables are denoted by boldface.) Finally, n-dimensional Minkowski space M_n uses Lorentz inner product

$$p \cdot q = p^\mu q^\nu g_{\mu\nu} = p^\circ q^\circ - \sum_1^{n-1} p^i x^i$$

and Fourier transform

$$\tilde{f}(p) = \int_{M_n} \frac{d^n p}{(2\pi)^{n/2}} f(x) e^{ip\cdot x};$$

if $\mathbf{p} = (p_1, \ldots p_m)$, we let $s_{ij} = p_i \cdot p_j$, $\quad 1 \le i \le j \le m$.

We begin by discussing scalar particles, so that each line l has propagator $\Delta^l(\lambda_l) \in \mathcal{S}'(\mathbb{R}^n)$

$$\tilde{\Delta}^l(\lambda_l) = \frac{i}{(m^2 - p^2 - i0)^{\lambda_l}}.$$

The physical propagator is obtained by setting $\lambda_l = 1$. For $\mathrm{Re}\lambda_l > n/2$, Δ^l is a continuous function of[7] $x \in \mathbb{R}^n$, and hence the Feynman amplitude $\mathcal{T}(\boldsymbol{\lambda}, n) \in \mathcal{S}'(\mathbb{R}^{nm})$, given by

$$\mathcal{T}(\boldsymbol{\lambda}, n)(\mathbf{x}) = \prod_{\mathcal{L}} \Delta^l(\lambda_l)(x_{f_l} - x_{i_l}),$$

is well defined. Its Fourier transform is easily calculated[7]:

$$\tilde{\mathcal{T}}(\boldsymbol{\lambda}, n)(\mathbf{p}) = f_G(\boldsymbol{\lambda}, n)\Gamma\left(\sum \lambda_l - \frac{hn}{2}\right)$$

$$\times \delta\left(\sum_1^m p_i\right) \int_0^1 \cdots \int_0^1 \left(\prod_{\mathcal{L}} \alpha_l^{\lambda_l - 1} d\alpha_l\right)$$

$$\times \delta\left(1 - \sum \alpha_l\right) d(\alpha)^{-n/2} \left(\sum \alpha_l m_l^2 - \frac{D(\alpha, s)}{d(\alpha)}\right)^{(hn/2 - \Sigma\lambda_l)}, \tag{1.1}$$

where

$$f_G(\boldsymbol{\lambda}, n) = ((2\pi)^{n/2}(-i)^{m-1}/2^{(nh/2)}\prod_{\mathcal{L}}\Gamma(\lambda_l)),$$

and $d(\alpha), D(\alpha, s)$ are the Symanzik polynomials for G. We sometimes write $\mathcal{T} = \lim_{\epsilon \to 0^+} \mathcal{T}_\epsilon$, where $\mathcal{T}_\epsilon$ is defined by replacing each m_l^2 in (1.1) by $m_l^2 - i\epsilon$.

Definition 1.1: The (analytically and complex-dimensionally) regularized amplitude $\mathcal{T}(\boldsymbol{\lambda}, \nu)$ for the graph G is obtained from (1.1) by replacing n by the complex variable ν.

We remark that this definition is equivalent to that of Ref. 1 (except for the presence of the λ's); in particular, it is obtained by applying the formulas of Ref. 1, Appendix A, to a p-space Feynman integral. $\mathcal{T}(\boldsymbol{\lambda}, \nu)$ may be considered as an element of $\mathcal{S}'(R^{n\cdot m})$ for any n, since it depends only on the invariants s_{ij}.

Theorem 1.2: $\mathcal{T}(\boldsymbol{\lambda}, \nu)$ may be analytically continued to a meromorphic function of $(\boldsymbol{\lambda}, \nu) \in \mathbb{C}^{L+1}$, having simple poles on the linear varieties

$$\nu_H = \sum_{\mathcal{L}(H)} \lambda_l - h(H)\frac{\nu}{2} = -k, \tag{1.2}$$

for each irreducible subgraph H of G and positive integer k (H is irreducible if it is connected and cannot be disconnected by removing a line or vertex).

Proof: We introduce into (1.1) the scaling transformations of the α variables used in the theory of analytic renormalization.[7] Then $\mathcal{T}(\boldsymbol{\lambda}, \nu)$ becomes a sum of integrals of the form

$$\delta(\textstyle\sum p_i)\Gamma(\nu_G) f_G(\boldsymbol{\lambda}, \nu) \int_0^1 \cdots \int_0^1 \prod_H t_H^{\nu_H - 1} dt_H E(t)^{-\nu/2}$$

$$\times \left(\sum_{\mathcal{L}} m_l^2 \beta_l(t) - \frac{F(t, s)}{E(t)}\right)^{-\nu_G}, \tag{1.3}$$

where $\prod_H$ is over an "s family" of subgraphs H which are irreducible or consist of a single line. In (1.3), β_l is a monomial in the t_H with $\beta_l = 1$ for some l, E and F are polynomials with E strictly positive in the region of integration. When the factors $t_H^{\nu_H - 1}$ in (1.3) are regarded as distributions,[7,8] (1.3) is well defined for all $\boldsymbol{\lambda}$, ν, and the singularity structure of (1.2) emerges. [Actually, in

(1.3) we have $H \neq G$; the singularity (1.2) for $H = G$ arises from the factor $\Gamma(\nu_G)$. Similarly the apparent poles of (1.3) arising from the cases where $H = \{l\}$ are cancelled by the factors $\Gamma(\lambda_l)^{-1}$ in $f_G(\boldsymbol{\lambda}, \nu)$.]

We now extend the definition to the case of particles with spin 1 or $\frac{1}{2}$; we assume that a particular representation $\{\gamma_\mu \mid \mu = 0, 1, \ldots, n-1\}$ of the Clifford algebra $C(M_n)$ has been chosen in each dimension, such that the trace of any product of an odd number of γ's vanishes. We will actually normalize the trace to satisfy $\mathrm{Tr}(1) = 4$ in all dimensions.[1] Now the amplitude for any process, calculated in n-dimensions, is a linear combination of certain tensor forms

$$p_i{}^\mu,\ g^{\mu\nu},\ \not{p}_i,\ \not{p}_i\gamma^\mu, \text{etc.} \tag{1.4}$$

(with distributions similar to 1.1 as coefficients). How are we to interpret these tensors when n is replaced by the complex variable ν?

The solution proposed in Ref. 1 is threefold: (a) external momenta are always from M_4; and all γ matrices are eliminated before introducing the complex dimension by (b) evaluating the trace for closed loops of spinor lines and (c) inserting projection operators, and then taking the trace, for open spinor lines. For simple graphs this procedure is adequate; however, when recursive subtractions are necessary, difficulty is encountered, particularly with procedure (c). That is, it is unclear how the amplitude (or vertex part) for a subgraph can be inserted recursively into the amplitude for the graph if we have defined only its traces when multiplied by various γ matrices.

For this reason we will treat tensors such as (1.4) as symbolic quantities, which may be interpreted as existing in whichever dimension is necessary at any time (4 for physical renormalization, arbitrary n for recursive subtraction). To regularize an amplitude we therefore express it as a linear combination of these symbolic forms in dimension n, then replace n by ν in the coefficient distributions. As in Ref. 1, an additional polynomial ν dependence of these coefficients is generated by contractions via the relation $g^\mu{}_\mu = n$.

Finally, we define similarly regularized amplitudes for generalized graphs. Let $Q = \{U_1, \ldots U_n\}$ be a partition of $\{V_1, \ldots V_m\}$, with $U_i = \{V_{i1}, \ldots V_{im(i)}\}$; let $G(U_i)$ be the subgraph of G formed by all lines joining vertices in U_i; and let $\tilde{G}$ be the graph obtained from G by contracting the subgraphs $G(U_i)$. Suppose we are given vertex parts

$$\mathfrak{X}(U_i) = \begin{cases} 1, & \text{if } m_i = 1, \\ 0, & \text{if } G(U_i) \text{ not IPI}, \\ \delta(\sum p_{ia}) D_i(p_{ia}) & \text{otherwise}, \end{cases} \tag{1.5}$$

with D_i a Lorentz-invariant polynomial, of degree at most equal to

$$\mu(G(U_i)) = \sum_{\mathcal{L}(G(U_i))} (r_l + 2) - 4(m_i - 1), \tag{1.6}$$

where r_l is the degree of the polynomial in the numerator of the lth propagator. (Note that this superficial divergence is computed in dimension 4.) The coefficients of D may depend on $\boldsymbol{\lambda}, \nu$. $\mathfrak{X}(U_i)$ may be interpreted as defining a vertex part in any particular dimension, according to the interpretation above.

In dimension n we form the amplitude

$$\mathcal{T}_{Q,\mathfrak{X}}(\boldsymbol{\lambda}, n) = \prod_{\mathcal{L}(\tilde{G})} \Delta^l(\lambda_l) \prod_{i=1}^{M} \mathfrak{X}(U_i). \tag{1.7}$$

Equation (1.7) is again a linear combination of invariant tensors with coefficients similar to (1.1), so that we may define $\mathcal{T}(\boldsymbol{\lambda}, \nu)$ for any ν as above. Symanzik rules for these amplitudes are worked out in Appendix C; for future use we note that each coefficient is a sum of terms of the form (compare 1.3)

$$f_G(\boldsymbol{\lambda}, \nu)\Gamma(\nu_G - j_G)\delta(\textstyle\sum p_i)\int_0^1 \cdots \int_0^1 \prod_H t_H^{\nu_H - j_H - 1} dt_H E_{\tilde{G}}(t)^{-\nu/2 - j} \times P(t, s, \nu)[\textstyle\sum m_l^2 \beta_l(t) - F_{\tilde{G}}(t, s)/E_{\tilde{G}}(t)]^{-(\nu_G - j_G)}, \tag{1.8}$$

where j, j_H are positive integers and P is a polynomial.

2. RENORMALIZATION

The regularized amplitudes are renormalized[1] by a slight variation of the standard BPH scheme of recursive subtractions. In the α-space context of this paper we can use the (minimal) counterterm structure associated with generalized vertices of the graph, i.e., we make subtractions only for those divergent subgraphs consisting of *all* lines connecting a given subset of the vertices. (In a p-space formulation as in Ref. 1 subtractions for additional divergent loop integrations are necessary.)

In this section we use dimensional regularization only, i.e., we set $\boldsymbol{\lambda} = \boldsymbol{\lambda}^\circ$ at all times, and will therefore omit the λ dependence of the amplitudes.

Definition 2.1: If $f(\nu)$ has an isolated singularity at $\nu = 4$, let Kf be the singular part, defined by

$$Kf(\nu) = \int_{|\nu' - 4| = r} \frac{f(\nu')}{\nu' - \nu} d\nu'$$

for $|\nu - 4| < r$.

Definition 2.2: Let $\mathcal{Y}(\nu)(V_i) = 1$, and suppose inductively that we have defined $\mathcal{Y}(\nu)(V'_1, \ldots, V'_{r'})$ for all generalized vertices $\{V'_1, \ldots, V'_{r'}\} \subset \{V_1, \ldots, V_m\}$, with $r' < r$. Then

$$\bar{\mathcal{P}}(\nu)(V'_1, \ldots, V'_r) = \sum_Q \mathcal{T}_{Q,\mathcal{Y}}(\nu)(V'_1, \ldots, V'_r), \tag{2.1}$$

$$\mathcal{Y}(\nu)(V'_1, \ldots, V'_r) = -K\bar{\mathcal{P}}(\nu)(V'_1, \ldots, V'_r), \tag{2.2}$$

$$\mathcal{P}(\nu)(V'_1, \ldots, V'_r) = \bar{\mathcal{P}}(\nu)(V_1, \ldots, V_r) + \mathcal{Y}_\nu(V'_1, \ldots, V'_r). \tag{2.3}$$

In (2.1) the sum is over all partitions Q of $\{V'_1, \ldots, V'_r\}$ into at least two generalized vertices. $\mathcal{P}(\nu)(V_1, \ldots, V_m)$ *is then the renormalized amplitude for the graph* G.

We prove below that $\mathcal{Y}(\nu)(V'_1, \ldots, V'_r)$ is in fact a vertex part; this was shown in Ref. 1 for graphs containing up to two loops. [Actually, it is necessary to know this inductively for $r' < r$ in order that (2.1) be well defined.] Formulas (2.1)–(2.3) exactly parallel the BPH scheme except that renormalization is effected by discarding a pole in the complex dimension ν rather than discarding low order terms of a Taylor series.

We wish to compare (2.1)–(2.3) with the BPH $\mathfrak{R}$ operation, and will follow the notation of Ref. 9; in particular, if $\mathcal{W}(\nu)$ is an amplitude associated with some generalized vertex $\{V'_1, \ldots, V'_r\}$, and $\mathcal{W}(\nu) = \delta(\sum p'_i)F(\nu, \mathbf{p})$, then $M\mathcal{W} = \delta(\sum p'_i)G(\nu, p)$, with G the Maclaurin series for F in $\mathbf{p}$ up to order $\mu(V'_1, \ldots, V'_r)$. Define finite vertex parts by

$$\hat{\mathfrak{X}}(\nu)(V'_1, \ldots, V'_r) = \begin{cases} 1 & \text{if } r = 1 \\ 0 & \text{if } G(V'_1, \ldots, V'_r)\text{IPR} \\ (1 - K)M\bar{\mathcal{P}}(\nu)(V'_1, \ldots, V'_r) & \text{otherwise.} \end{cases} \tag{2.4}$$

[$\hat{\mathfrak{X}}$ is certainly a vertex part, by definition of M; it is "finite" because the $(1-K)$ factor removes the singularity at $\nu = 4$.] Let $\mathfrak{R}', \mathfrak{X}', \overline{\mathfrak{R}}'$ be the BPH quantities defined using this finite renormalization and complex dimensional regularization:

$$\overline{\mathfrak{R}}'(\nu)(V_1', \ldots, V_r') = \sum_{\mathcal{Q}} \mathcal{T}_{\mathcal{Q},\mathfrak{X}'}(\nu)(V_1', \ldots, V_r'), \tag{2.5}$$

$$\mathfrak{X}'(\nu)(V_1', \ldots, V_r') = -M\overline{\mathfrak{R}}'(\nu)(V_1', \ldots, V_r') + \hat{\mathfrak{X}}(\nu)(V_1', \ldots, V_r'), \tag{2.6}$$

$$\mathfrak{R}(\nu)(V_1', \ldots, V_r') = \overline{\mathfrak{R}}(\nu)(V_1', \ldots, V_r') + \mathfrak{X}'(\nu)(V_1', \ldots, V_r'), \tag{2.7}$$

with $\sum_{\mathcal{Q}}$ as in (2.1).

Theorem 2.3: For any $\{V_1', \ldots, V_r'\} \subset \{V_1, \ldots, V_m\}$,

$$\mathfrak{X}'(\nu)(V_1', \ldots, V_r') = \mathcal{Y}(\nu)(V_1', \ldots, V_r'), \tag{2.8}$$

$$\overline{\mathfrak{R}}'(\nu)(V_1', \ldots, V_r') = \overline{\mathcal{P}}(\nu)(V_1', \ldots, V_r'), \tag{2.9}$$

$$\mathfrak{R}'(\nu)(V_1', \ldots, V_r') = \mathcal{P}(\nu)(V_1', \ldots, V_r'). \tag{2.10}$$

Corollary 2.4: The complex-dimensional renormalization of Definition 2.2 belongs to the class of BPH renormalizations.

Proof: This is precisely the content of (2.10).

Corollary 2.5: $\mathcal{Y}(\nu)$ is a vertex part.

Proof: This follows from (2.8), since $\mathfrak{X}'(\nu)$ is a vertex part.

The crucial lemma is

Lemma 2.6: $\mathfrak{R}'(\nu)(V_1', \ldots, V_r')$ is analytic at $\nu = 4$.

Since we expect the $\mathfrak{R}$ operation to remove all divergences, Lemma 2.6 is intuitively reasonable. The proof, however, is complicated by the complex-dimensional regularizations; we relegate it to Appendix A.

Proof of Theorem 2.3: Formulas (2.8)–(2.10) certainly hold if $r = 1$; suppose inductively that they hold for all $r < r_0$. Then from (2.1) and (2.5), using (2.8) for $r < r_0$, (2.9) holds for $r = r_0$. Thus $\hat{\mathfrak{X}}(\nu)(V_1', \ldots, V_{r_0}') = (1-K)M\overline{\mathfrak{R}}'(\nu)(V_1', \ldots, V_{r_0}')$, and from (2.6), $\mathfrak{X}'(\nu)(V_1', \ldots, V_{r_0}') = -KM\overline{\mathfrak{R}}'(\nu)(V_1', \ldots, V_{r_0}')$. Since $K^2 = K$, $K\mathfrak{X}'(\nu)(V_1', \ldots, V_{r_0}') = \mathfrak{X}'(\nu)(V_1', \ldots, V_{r_0}')$. But from Lemma 2.6,

$$\begin{aligned} 0 &= K\mathfrak{R}'(\nu)(V_1', \ldots, V_{r_0}') \\ &= K\overline{\mathfrak{R}}(\nu)(V_1', \ldots, V_{r_0}') + K\mathfrak{X}'(\nu)(V_1', \ldots, V_{r_0}') \\ &= -\mathcal{Y}(\nu)(V_1', \ldots, V_{r_0}') + \mathfrak{X}'(\nu)(V_1', \ldots, V_{r_0}'), \end{aligned}$$

proving (2.8). Equation (2.10) follows immediately from (2.8) and (2.9).

3. WARD IDENTITIES

We will use QED as an example in this section, but the arguments given are quite general. Consider then a particular Ward identity, e.g., for the vacuum polarization tensor:

$$k_\mu \Pi^{\mu\sigma}(k) = 0. \tag{3.1}$$

We wish to prove a regularized version

$$k_\mu \Pi^{\mu\sigma}(\nu; k) = 0. \tag{3.2}$$

Then the recursive subtractions of (2.1) preserve (3.2) (the proof is the same as for any gauge-invariant regularization), and the $(1-K)$ operation in (2.3) yields a renormalized amplitude which also satisfies the Ward identity.

There are three difficulties in establishing (3.2): (a) in complex dimension the contraction over the index μ is meaningless (as explained in Sec. 1 we do not take external vectors as four-dimensional); (b) even in integer dimensions divergent quantities are involved, and the formal proofs are therefore suspect; and (c) the usual manipulations to establish (3.1) are based on the p-space integral form, which is not available to us when ν is complex. We will treat these difficulties in turn.

The problem of contractions in nonintegral dimension is handled by regarding Ward identities as relations between the coefficients of various tensors. In (3.2), for example, write $\Pi^{\mu\sigma}(\nu; k) = A(\nu, k^2)g^{\mu\sigma} + B(\nu; k^2)k^\mu k^\sigma$, so that the Ward identity becomes

$$A(\nu, k^2) + k^2 B(\nu, k^2) = 0. \tag{3.3}$$

Now (3.3) makes sense for all values of ν (and we prove it below). All Ward identities may be interpreted in this sense. (It is necessary to first choose a linear basis for all the tensor forms, and express all amplitudes in terms of this basis.) We will usually not mention this explicitly in what follows.

Before proceeding we introduce the following notation. For any QED graph G, with l a Fermion line incident on an external photon vertex V_i, let $G_{l,i}$ be the graph obtained from G by replacing l with a scalar particle, and removing the γ matrix associated with V_i; let $\tilde{G}_{l,i}$ be obtained from $G_{l,i}$ by contracting l. Then the amplitudes for these graphs are related by

Lemma 3.1:

$$\mathcal{T}_{G_{l,i}}(\boldsymbol{\lambda}, \nu)|_{\lambda_l = 0} = i\mathcal{T}_{\tilde{G}_{l,i}}(\boldsymbol{\lambda}, \nu).$$

Proof: With ν an integer, the lemma follows immediately from the p-space Feynman integral, using $\tilde{\Delta}^l|_{\lambda_l=0} = i$. For nonintegral ν we argue directly from (1.1), treating the factor $\alpha_l^{\lambda_l - 1}$ $[= (\alpha_l)_+^{\lambda_l - 1}]$ as a distribution and using[8] $(\alpha_l^{\lambda_l - 1}/\Gamma(\lambda_l))|_{\lambda_l = 0} = \delta(\alpha_l)$. By holding $\mathrm{Re}\,\lambda_{l'} \gg 0$, $l' \neq l$, divergence difficulties are avoided; the result extends to all $\boldsymbol{\lambda}$ by analytic continuation.

To treat difficulty (b) we prove a modified identity involving the λ's. The propagator $\tilde{S}(\lambda, p) = i(\not{p} + m)(m^2 - p^2 - i0)^{-\lambda}$ satisfies a generalization of the usual Ward–Takahashi identity:

$$\tilde{S}(\lambda_a, p)k\tilde{S}(\lambda_b, p+k) = -\tilde{S}(\lambda_a, p)\tilde{\Delta}(\lambda_b - 1, p+k) + \tilde{\Delta}(\lambda_a - 1, p)\tilde{S}(\lambda_b, p+k), \tag{3.4}$$

where Δ is the scalar propagator of mass m. Inserting (3.4) into the p-space Feynman integral for $\mathcal{T}_G$ immediately proves the integer dimension case of

Theorem 3.2: Suppose that G is a QED graph with V_1 an external photon vertex and a, b the fermion lines incident on V_1. Then

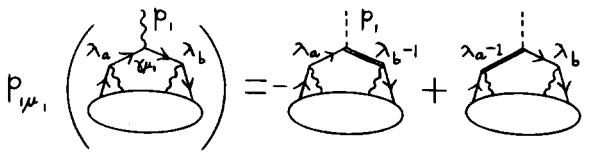

FIG. 1. A λ-regularized identity.

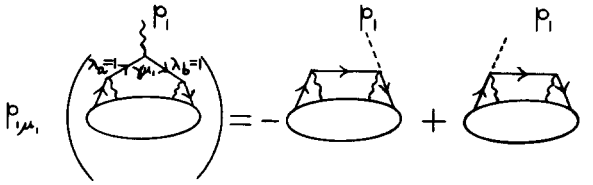

FIG. 2. The identity with λ regularization removed.

$$(p_1)_{\mu_1}\mathcal{T}_G^{\mu_1\mu_2\cdots}(\boldsymbol{\lambda},\nu)(p_1,\ldots,p_m) - \mathcal{T}_{G_{b,1}}^{\mu_2\cdots}(\lambda_a,\lambda_b-1,\ldots;\nu)(\mathbf{p}) + \mathcal{T}_{G_{a,1}}^{\mu_2\cdots}(\lambda_a-1,\lambda_b,\ldots;\nu)(\mathbf{p}). \tag{3.5}$$

Remark: This identity is indicated pictorially in Fig. 1, where the double line denotes a scalar particle, and the dotted line an external momentum (no longer particularly a photon). If in (3.5) we set $\lambda_l = 1$ for all l, and use Lemma 3.1, we obtain

$$(p_1)_{\mu_1}\mathcal{T}_G^{\mu_1\cdots}(\nu)(\mathbf{p}) = -\mathcal{T}_{G_{b,1}}^{\mu_2\cdots}(\nu)(\mathbf{p}) + \mathcal{T}_{G_{a,1}}^{\mu_2\cdots}(\nu)(\mathbf{p}) \tag{3.6}$$

(see Fig. 2). But (3.6) is precisely the relation needed to establish Ward identities (after inserting an external photon vertex into a diagram in all possible ways). Thus there remains only to prove Theorem 3.2 for noninteger ν.

We wish to apply Carlson's theorem[10,11] to a suitable function; to avoid complicated analytic continuations we work in a region of $(\boldsymbol{\lambda},\nu)$ space in which there is no ultraviolet divergence. (It is here that we use critically the λ regularization.) The necessary estimate comes from

Lemma 3.3: Let G be an arbitrary Feynman graph, $\mathcal{T}'_\epsilon(\boldsymbol{\lambda},\nu,\mathbf{s})$ the coefficient of some tensor form in the Feynman amplitude for G, and m_0 the minimal mass occurring in G. Then there exist positive constants a, b, k such that for $|s_{ij}| < a$ and $\mathrm{Re}\lambda_l > \frac{1}{2}\mathrm{Re}\nu > b$,

$$\left|f_G^{-1}(\boldsymbol{\lambda},\nu)\Gamma(\nu_G)^{-1}\left(\frac{m_0^2}{2}\right)^{\nu_G}\mathcal{T}'_\epsilon(\boldsymbol{\lambda},\nu,\mathbf{s})\right| \le K\exp\frac{(2\epsilon L|\nu_G|)}{m_0^2}. \tag{3.7}$$

Proof: According to (1.8) the function to be estimated is a sum of terms of the form

$$\left(\prod_{i=1}^{j_G}(\nu_G - i)^{-1}\right)\left(\frac{m_0^2}{2}\right)^{j_G}\int_0^1\cdots\int_0^1\prod t_H^{\nu_H-j_H-1}dt_H P(\mathbf{t},\mathbf{s},\nu) \times E(\mathbf{t})^{-\nu/2-j}X^{-\nu_G+j_G}, \tag{3.8}$$

where

$$X = \left[\left(\frac{F(\mathbf{t},\mathbf{s})}{E(\mathbf{t})} + \sum_l(m_l^2 - i\epsilon)\beta_l\right)\frac{2}{m_0^2}\right],$$

with P a polynomial and j_H, j fixed positive integers.

Take $b = \max\{j_H, j\} + 1$ ($H = G$ is included). Then $|\nu_G - i|^{-1} \le 1$, for $i \le j_G$; using (1.2), $\mathrm{Re}\nu_H \ge \mathrm{Re}\nu/2 > j_H + 1$, so $|t_H^{\nu_H-j_H-1}| \le 1$; and, since $E(t) \ge 1$, $|E(t)^{-\nu/2-j}| < 1$. The polynomial ν dependence is dominated by the exponential in (3.7). Finally, for a sufficiently small, $\mathrm{Re}X > 1$ (since $\beta_l \ge 0$ and $\beta_{l_0} = 1$ for some l_0) and $\mathrm{Im}X \le 2L\epsilon/m_0^2$. Hence

$$0 \ge \arg X \ge \tan^{-1}\left(-\frac{2\epsilon L}{m_0^2}\right) \ge \frac{-2\epsilon L}{m_0^2},$$

and

$$|X^{-(\nu_G-j)_G}| = |X|^{j_G}\exp(-\mathrm{Re}\nu_G\ln|X| + \mathrm{Im}\nu_G\arg X) \le K'\exp(2\epsilon L|\nu_G|/m_0^2).$$

Inserting these estimates into (3.8) yields (3.7) immediately.

Proof of Theorem 3.2: Let the coefficient of some tensor form in (3.5) (with ϵ dependence added) be $g_\epsilon(\boldsymbol{\lambda},\nu,\mathbf{s})$; we wish to show

$$g_\epsilon(\boldsymbol{\lambda},\nu,\mathbf{s}) = 0. \tag{3.9}$$

Fixing real numbers η_l, with $\eta_l > \frac{1}{2}$, and setting $\lambda_l = \eta_l\nu$, we have $\nu_G = \beta\nu$ with $\beta = \beta(\eta) > 1$. For ϵ sufficiently small and $|s_i| < a$, we may by Lemma 3.5 apply Carlson's theorem to

$$h(\nu) = f(\boldsymbol{\lambda},\nu)^{-1}\Gamma(\nu_G)^{-1}\left(\frac{m_0^2}{2}\right)^{\nu_G}g_\epsilon(\boldsymbol{\lambda},\nu,\mathbf{s})\Big|_{\lambda_l=\eta_l\nu},$$

to find that $h(\nu) = 0$. Since g_ϵ is analytic in $\boldsymbol{\lambda}$ and real analytic in $\mathbf{s}$, (3.9) follows for all $(\boldsymbol{\lambda},\mathbf{s})$; the $\epsilon \to 0$ limit gives (3.5).

APPENDIX A: THE BPH $\mathcal{R}$ OPERATION WITH COMPLEX-DIMENSIONAL REGULARIZATION

Our purpose is to sketch the proof of Lemma 2.6. We wish to use, with as little additional machinery as possible, the convergence estimates which Hepp[9] has given for the $\mathcal{R}$ operation. Complications arise because Hepp's methods rely heavily on the product structure of the regularized Feynman amplitude, a structure which is destroyed by the complex-dimensional regularization.

We first rewrite the $\mathcal{R}'$ operation in terms of the $\mathcal{R}$ operation (which involves no finite renormalization) for generalized graphs:

$$\mathcal{R}'\mathcal{T}(\boldsymbol{\lambda},\nu) = \sum_Q\mathcal{R}\mathcal{T}_{Q,\hat{\mathfrak{X}}}(\boldsymbol{\lambda},\nu), \tag{A1}$$

where the sum is over all partitions Q of $\{V_1,\ldots,V_m\}$ and $\mathcal{T}_{Q,\hat{\mathfrak{X}}}$ is defined as in (1.7), but starting from the vertex parts $\hat{\mathfrak{X}}$ of (2.4). The usual proof[12] of (A1) does not involve the product structure of $\mathcal{T}$, but only the multilinearity in the vertex parts of the amplitude for a generalized graph; this continues to hold for complex-dimensional regularization. It therefore suffices to prove

Lemma A1: For any finite vertex parts $\hat{\mathfrak{X}}(\boldsymbol{\lambda},\nu)$, the amplitude

$$\mathcal{R}\mathcal{T}_{Q,\hat{\mathfrak{X}}}(\boldsymbol{\lambda},\nu) \tag{A2}$$

is analytic at $(\boldsymbol{\lambda}^0, 4)$.

Proof: For simplicity we discuss only the partition

$Q = \{\{V_1\}, \ldots, \{V_m\}\}$. Hepp (Lemma 3.1 of Ref. 9) gives a representation of (A2) which extends immediately to n dimensions and the inclusion of λ regularization:

$$\mathcal{R}\mathcal{T}(\boldsymbol{\lambda}, n) = \sum_{\pi} \sum_{T} \mathcal{F}_T^{\pi}(\boldsymbol{\lambda}, n). \tag{A3}$$

Here $\sum_{\pi}$ runs over permutations π of $\mathcal{L}$, and $\sum_T$ over certain "trees"; $\mathcal{F}_T^{\pi}$ is the Feynman amplitude for the tree T integrated over the region $\mathcal{D}^{\pi} = \{\alpha \mid \alpha_{l_{\pi(1)}} \leq \cdots \leq \alpha_{l_{\pi(L)}}\}$. Moreover, $\mathcal{F}_T^{\pi}$ is shown to be given a Feynman-like integral; following through the proof, we find that the $\boldsymbol{\lambda}, n$ dependence enters in four ways: an overall factor $f(\boldsymbol{\lambda}, n)\Gamma(\nu_G - j_G)^{-1}$, as in (1.8); possible polynomial n-dependence (from $g^{\mu}{}_{\mu} = n$); factors $\alpha_l^{\lambda_l - 1}$ in the integrand, and a replacement of Hepp's $D^{I'} = \prod_{l \in \mathcal{L}(I') - \mathfrak{M}} D_l^{-2}$ by $\prod_{l \in \mathcal{L}(I') - \mathfrak{M}} D_l^{-n/2}$. Thus we may define $\mathcal{F}_T^{\pi}(\boldsymbol{\lambda}, \nu)$ for noninteger ν; (A3) for general ν now follows from Carlson's theorem (the argument is similar to those of Sec. 3). Finally, the estimates given by Hepp in Lemma 3.4 show that $\mathcal{F}_T^{\pi}(\boldsymbol{\lambda}, \nu)$ is analytic at $(\boldsymbol{\lambda}_0, 4)$.

APPENDIX B: SYMANZIK RULES FOR GENERALIZED GRAPHS

In Ref. 7 (see also Ref. 13) we have given "Symanzik" rules for arbitrary Feynman graphs, which allow the α-space Feynman integral to be written down directly. We here record the corresponding rules for generalized graphs.

Thus let G be a Feynman graph as in Sec. 1, with propagators

$$\tilde{\Delta}_l(\lambda_l)(p) = \frac{iZ_l(p)}{(m^2 - p^2 - i0)^{\lambda_l}}, \tag{B1}$$

where Z_l is a polynomial. Let Q be a partition of $\{V_1, \ldots, V_m\}$ into generalized vertices $U_i = \{V_{i1}', \ldots, V_{im(i)}'\}$, $i = 1, \ldots, M$, and $\tilde{G}$ the corresponding contracted graph. The incidence matrix for G is written

$$e_{ia}^l = \begin{cases} 1, & \text{if } V_{f_l} = V_{ia}' \\ -1, & \text{if } V_{i_l} = V_{ia}' \\ 0, & \text{otherwise,} \end{cases}$$

so that $e_i^l = \sum_a e_{ia}^l$ is the incidence matrix of G. Given vertex parts $\mathfrak{X}(U_i)$ [(1.5)], we wish to calculate $\mathcal{T}_{Q,\mathfrak{X}}$ as given by (1.7).

Writing

$$Z_l(p) = Z_l\left(\frac{1}{i}\frac{\partial}{\partial u_l}\right) e^{ip \cdot u_l}\Big|_{u_l = 0},$$

(B1) becomes $\Delta_l = \lim_{\epsilon \to 0+} \Delta_{l,\epsilon}$, where

$$\Delta_{l,\epsilon}(x) = \frac{\exp[\frac{1}{2}\pi i(\lambda_l - n/2)]}{2^{n/2}\Gamma(\lambda_l)} Z_l\left(\frac{1}{i}\frac{\partial}{\partial u_l}\right) \int_0^{\infty} \alpha_l^{\lambda_l - 1 - n/2} d\alpha_l$$

$$\times \exp\{i[-(1/4\alpha_l)(e_{ia}^l x_{ia} + u_l)^2 - (m^2 - i\epsilon)\alpha_l]\}. \tag{B2}$$

Similarly,

$$\mathfrak{X}(\mathbf{x}_i) = (2\pi)^{(n/2)(m_i - 2)} D_i\left(\frac{1}{i}\frac{\partial}{\partial s_{ia}}\right)$$

$$\times \prod_{a=1}^{m_i - 1} \delta(x_{ia} + s_{ia} - x_{i,a+1} - s_{i,a+1})\Big|_{s_{ia} = 0}. \tag{B3}$$

We insert (B2) and (B3) into (1.7) and take the Fourier transform. It is most convenient to change integration variables to $y_{ia} = x_{ia} + s_{ia}$. The $y_{i2}, \ldots, y_{im(i)}$ integrations are then done using the δ functions; the remaining Gaussian integration is virtually identical with that encountered in calculating an amplitude for $\tilde{G}$. Evaluating this as in Ref. 7 (or Ref. 12) gives

$$\tilde{\mathcal{T}}_{Q,\mathfrak{X}}(\boldsymbol{\lambda}, n) = g_Q(\boldsymbol{\lambda}, n)\delta(\textstyle\sum p_{ia}) \int_0^{\infty} \cdots \int_0^{\infty} (\prod \alpha_l^{\lambda_l - 1} d\alpha_l)\, d(\alpha)^{-n/2}$$

$$\times \prod D_i\left(\frac{1}{i}\frac{\partial}{\partial s_{ia}}\right) \prod Z_l\left(\frac{1}{i}\frac{\partial}{\partial u_l}\right) \exp\left[i\left(\sum_{i,j=1}^{M} r_i r_j \frac{D_{ij}^k(\alpha)}{d(\alpha)}\right.\right.$$

$$\left.\left. + \textstyle\sum p_{ia} s_{ia} - \sum_l [t_l^2/4\alpha_l + (m^2 - i\epsilon)\alpha_l]\right)\right] \Big|_{\mathbf{s} = \mathbf{u} = 0}. \tag{B4}$$

Here

$$g_Q(\boldsymbol{\lambda}, n) = \prod_{\mathcal{L}(\tilde{G})} \left(\frac{\exp[(\pi i/2)(\lambda_l - n/2)]}{2\Gamma(\lambda_l)}\right) \left(\frac{e^{n\pi i/4}}{\pi i}\right)^{M-1},$$

$$t_l = u_l - \sum_{i,a} e_{ia}^l s_{ia},$$

$$r_i = q_i + \sum_l \alpha_l^{-1} e_i^l t_l$$

$$q_i = \sum_{a=1}^{m(i)} p_{ia},$$

and $d(\alpha), D_{ij}^k(\alpha)$ are the Symanzik functions for[7] $\tilde{G}$:

$$d(\alpha) = \sum_T \prod_{l \notin T} \alpha_l,$$

$$D_{ij}^k = \sum_{T_2} \prod_{l \notin T_2} \alpha_l;$$

the sums are respectively over all trees in $\tilde{G}$ and over all 2-trees in $\tilde{G}$ which separate U_k from U_i and U_j. $k \in \{1, \ldots, M\}$ is arbitrary.

Now recall[12] the following definition: Given a set of quantities $\{X_i\}$ and associated pairwise contractions $\{\overline{X_i X_j}\}$, the τ *product* of a monomial in $\mathbf{X}$ is defined by summing over all contractions, precisely as in Wick's theorem for a T product; the τ product is extended to polynomials in $\mathbf{X}$ by linearity. If we evaluate the u, s derivatives in (B4) and set $\mathbf{u} = \mathbf{s} = 0$, the integrand will contain a factor (see Ref. 12)

$$\tau[\prod Z_l(X^l) \prod D_i(Y_{ia})],$$

where

$$X^l = \sum_{i,j} e_i^l D_{ij}^k q_j / \alpha_l\, d(\alpha), \tag{B5}$$

$$Y_{ia} = p_{ia} - \sum_l e_{ia}^l X^l, \tag{B6}$$

$$\overline{X_{\mu}^l X_{\nu}^{l'}} = \tfrac{1}{2}\left(\sum_{i,j} \frac{e_i^l e_j^l D_{ij}^k}{\alpha_l \alpha_{l'} d(\alpha)} - \frac{\delta_{l,l'}}{\alpha_l}\right) g_{\mu\nu}, \tag{B7}$$

and other contractions are calculated from (B6), (B7), and the relation $\overline{X^l p_{ia}} = 0$.

[The momenta X^l have an intrinsic characterization. For fixed $\{\alpha_l\}$ and $\{q_i\}$, let $\{k^l \mid l \in \mathcal{L}(\tilde{G})\}$ be arbitrary n-vectors satisfying momentum conservation in $\tilde{G}$. $k^l = X^l$ is then a stationary point for the function $\sum \alpha_l (k^l)^2$;

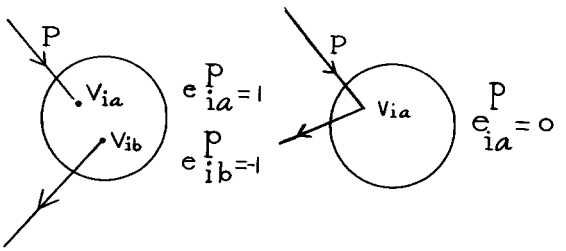

FIG. 3. Incidence of paths on generalized vertices.

Y_{ia} is the total momentum flowing out of V_{ia} in this momentum configuration. See Ref. 5, Sec. 4. 2.]

More explicit Symanzik rules depend on the idea of incidence of an oriented path or circuit P in $\tilde{G}$ on lines of $\tilde{G}$ and on vertices of G. For $l \in \mathcal{L}(\tilde{G})$,

$$e_l^P = \begin{cases} 1, & \text{if } l \in P, \text{ and the orientations of } l \text{ and } P \text{ coincide,} \\ 0, & \text{if } l \notin P, \\ -1, & \text{otherwise.} \end{cases}$$

For a vertex V_{ia}, $e_{ia}^P = \sum_l e_{ia}^l e_l^P$, i.e.,

$$e_{ia}^P = \begin{cases} 1, & \text{if precisely one line of } P \text{ is incident on } V_{ia}, \text{ and } P \text{ is oriented into } V_{ia}, \\ -1, & \text{if one line of } P \text{ is incident, oriented out,} \\ 0, & \text{otherwise.} \end{cases}$$

See Fig. 3. For each tree T of $\tilde{G}$ let $P_{kj}(T)$ be the path in T joining U_k to U_j, oriented from U_k to U_j. Then

$$X^l = d(\alpha)^{-1} \sum_T \sum_j e^{P_{kj}(T)} \left(\prod_{l \notin T} \alpha_l \right) q_j, \tag{B8}$$

$$Y_{ia} = p_{ia} - d(\alpha)^{-1} \sum_T \sum_j e_{ia}^{P_{kj}(T)} \left(\prod_{l \notin T} \alpha_l \right) q_j \tag{B9}$$

(the result is independent of k). Let T^* denote a set of M lines in $\tilde{G}$ containing precisely one circuit $C(T^*)$, which is given an arbitrary orientation. Then

$$\overline{X_\mu^l X_\nu^{l'}} = d(\alpha)^{-1} \sum_{T^*} \left(\prod_{l'' \notin T^*} \alpha_{l''} \right) e_l^{C(T^*)} e_{l'}^{C(T^*)} g_{\mu\nu}, \tag{B10}$$

$$\overline{X_\mu^l Y_{ia\nu}} = -\, d(\alpha)^{-1} \sum_{T^*} \left(\prod_{l'' \notin T^*} \alpha_{l''} \right) e_l^{C(T^*)} e_{ia}^{C(T^*)} g_{\mu\nu}, \tag{B11}$$

$$\overline{Y_{ia\mu} Y_{jb\nu}} = d(\alpha)^{-1} \sum_{T^*} \left(\prod_{l \notin T^*} \alpha_l \right) e_{ia}^{C(T^*)} e_{jb}^{C(T^*)} g_{\mu\nu}. \tag{B12}$$

(B8)–(B12) are the desired rules.

The final form of (B4) is thus

$$\tilde{T}_{Q,\mathfrak{T}}(\boldsymbol{\lambda}, n) = g_Q(\boldsymbol{\lambda}, n) \delta(\textstyle\sum p_{ia}) \int_0^\infty \cdots \int_0^\infty (\prod \alpha_l^{\lambda_l - 1}\, d\alpha_l)$$
$$\times d_{\tilde{G}}(\alpha)^{-n/2} \tau[\prod D_i(Y_{ia}) \prod Z_l(X^l)]$$
$$\times \exp\left[i \left(D_{\tilde{G}}(\alpha, s)/d_{\tilde{G}}(\alpha) - \sum_l (m_l^2 - i\epsilon)\alpha_l \right) \right]. \tag{B13}$$

At this point the variable scalings corresponding to s families in $\tilde{G}$ may be introduced. The result is of the form (1. 8).

[1] G. t'Hooft and M. Veltman, Nucl. Phys. B44, 189 (1972).
[2] G. Ponzano, T. Regge, E. Speer, and M. J. Westwater, Commun. Math. Phys. 15, 83 (1969).
[3] G. Ponzano, T. Regge, E. Speer, and M. J. Westwater, Commun. Math. Phys. 18, 1 (1970).
[4] T. Regge, E. Speer, and M. J. Westwater, Fortsch. Phys. 20, 365 (1972).
[5] E. Speer and M. J. Westwater, Ann. Inst. H. Poincaré 14A, 1 (1971).
[6] H. J. Vega and F. A. Schaposnik, "Dimensional Renormalization", preprint, Universidad Nacional de la Plata.
[7] E. Speer, "Lectures on Analytic Renormalization", University of Maryland Technical Report 73-067.
[8] I. M. Gelfand and G. E. Shilov, *Generalized Functions* (New York, Academic, 1964), Vol. I.
[9] K. Hepp, Commun. Math. Phys. 2, 301 (1966).
[10] E. J. Squires, *Complex Angular Momenta and Particle Physics* (Benjamin, New York, 1963).
[11] E. C. Titchmarsh, *Theory of Functions* (Oxford U.P., New York, 1939).
[12] E. Speer, *Generalized Feynman Amplitudes* (Princeton U.P., Princeton, 1969).
[13] N. Nakanishi, *Graph Theory and Feynman Integrals* (Gordon and Breach, New York, 1971).

On a complex representation of Lorentzian random variables

T. P. Eggarter

Laboratoire de Physique des Solides, Université Paris-Sud, Centre d'Orsay, 91 Orsay, France*
(Received 19 July 1973)

A Lorentzian random variable X is characterized by its mean value a and its width γ. We point out some interesting properties of the map $X \mapsto a + i\gamma$. We illustrate the usefulness of this map by solving a particular problem in the theory of disordered materials.

I. PROPERTIES OF LORENTZIAN RANDOM VARIABLES

A random variable is said to be Lorentzian if its probability density is of the form

$$p_X(x) = \frac{1}{\pi} \frac{\gamma}{(x-\alpha)^2 + \gamma^2} \tag{1}$$

In what follows we will use the words "X is (a, γ)" to indicate that X is Lorentzian with the probability density equation (1).

Let X, Y be two independent random variables, which are (a, γ) and (b, δ), respectively; and n a nonnegative real. Then the following properties hold:

(i) nX is $(na, n\gamma)$,

(ii) $n + X$ is $(n + a, \gamma)$,

(iii) $-X$ is $(-a, \gamma)$,

(iv) $X + Y$ is $(a + b, \gamma + \delta)$,

(v) X^{-1} is $\left(\frac{a}{a^2 + \gamma^2}, \frac{\gamma}{a^2 + \gamma^2}\right)$.

The proof of the above properties follows from the fact that if $Z = f(X, Y)$, then

$$p_Z(z) = \langle \delta(z - f(x,y)) \rangle = \iint dx\, dy\, \delta(z - f(x,y)) p_X(x) p_Y(y), \tag{2}$$

One simply has to carry out the last integral in (2) to check (i) to (v). Properties (i)–(iv) mean that the space $\mathcal{L}$ of Lorentzians is closed under linear combinations. This also holds, for example, for Gaussians, although there is one fundamental difference: for Gaussians the squares of the widths are added, $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$, while for Lorentzians it is simply the widths. Therefore there is no central limit theorem in the latter case; the average $X = N^{-1}(X_1 + X_2 + \cdots + X_N)$ of a certain number of equally distributed Lorentzians has the same probability density as each term in the sum. Averages do not become sharp as N increases. Property (v) is remarkable. It is not shared by Gaussians or any other of the common probability laws.

II. THE MAP $X \mapsto a + i\gamma$

The way in which the parameters (a, γ) and (b, δ) combine in properties (i)–(v) immediately suggests the use of complex variables. We define a map M: $\mathcal{L} \to \mathbb{C}^+$ ($\mathbb{C}^+$ = upper half of the complex plane) by the "recipe" that if X is (a, γ) it is sent into the point $z_X = a + i\gamma$:

$$X \mapsto z_X = a + i\gamma \tag{3}$$

Properties (i)–(v) take with this correspondence the following form: Let $X \mapsto z_X$ and $Y \mapsto z_Y$. Let n be a nonnegative real; then

(i) $nX \mapsto nz_X$,

(ii) $n + X \mapsto n + z_X$,

(iii) $-X \mapsto -z_X^*$,

(iv) $X + Y \mapsto z_X + z_Y$,

(v) $X^{-1} \mapsto (z_X^*)^{-1}$.

We see that Eq. 3 maps certain arithmetic operations on random variables into exactly the same operations in the complex plane, with the only slight complication that the complex conjugate must be taken in (iii) and (v). We also note incidentally that ii is a particular case of iv, since a real may be thought of as a random variable with dispersion zero.

III. A PROBLEM IN THE THEORY OF DISORDERED MATERIALS

The Hamiltonian matrix describing an electron on a linear chain of length N in the tight-binding approximation is the tridiagonal matrix[1]

$$H = \begin{pmatrix} \epsilon_1 & 1 & & & & \\ 1 & \epsilon_2 & 1 & & & \\ & 1 & \epsilon_3 & 1 & & \\ & & 1 & \cdot & & \\ & & & & \cdot & 1 \\ & & & & 1 & \epsilon_N \end{pmatrix}. \tag{4}$$

If we look for an eigenvector of the form

$$\psi = \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_N \end{pmatrix}, \tag{5}$$

the eigenvalue equation $H\psi = E\psi$ takes the form

$$r_1 \equiv c_1/c_2 = 1/(E - \epsilon_1), \tag{6a}$$

$$r_j \equiv c_j/c_{j+1} = 1/(E - \epsilon_j - r_{j-1}), \tag{6b}$$

$$r_{N-1} \equiv c_{N-1}/c_N = E - \epsilon_N. \tag{6c}$$

If the ϵ's are independent random variables with a probability density $g(\epsilon)$ and we disregard for the time being the condition (6c), then all r_i's can be expressed as functions of the ϵ_j's. If $p_j(r_j)$ is the probability density for r_j we get from (6b) the recurrence relation

$$p_j(r_j) = r_j^{-2} \int_{-\infty}^{\infty} g(\epsilon) p_{j-1}(E - \epsilon - r_j^{-1}) d\epsilon. \tag{7}$$

It can be shown that, provided $g(\epsilon)$ is reasonably well behaved, $p_j(x)$ converges to a function $p_\infty(x)$ as $j \to \infty$; the convergence, moreover, is uniform in x. Let $E_i(\{\epsilon_j\})$,

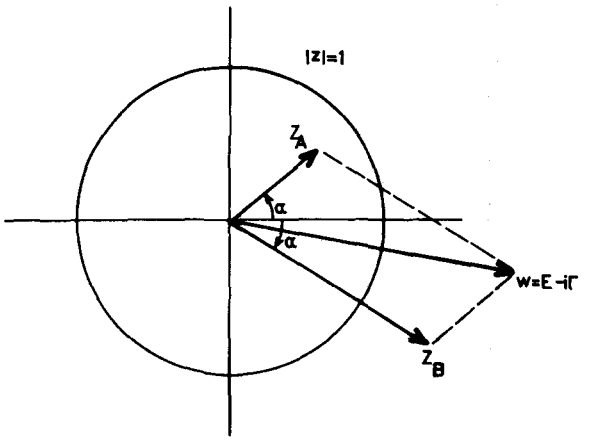

FIG. 1. Location of the roots of Eq. (10).

$i = 1, 2, \ldots, N$ be the eigenvalues of the problem (6a)–(6c). Economou and Papatriantafillou[1] prove that

$$\rho(E) \equiv \lim_{N\to\infty} \left\langle \frac{1}{N} \sum_{i=1}^{N} \delta(E - E_i) \right\rangle = \int_{-\infty}^{\infty} p_\infty(x) p_\infty(1/x) dx \qquad (8)$$

or, in words, that the knowledge of $p_\infty(x)$ permits the computation of the average density of eigenvalues. All this holds for an arbitrary $g(\epsilon)$.

If we consider in particular the case that the ϵ_i's are independent Lorentzians, we see from Eqs. (6a), (6b) and properties (ii)–(v) that all r_j's will also have Lorentzian distributions. That the problem is exactly soluble in this case for the linear chain[1] as well as for a Cayley tree[2] is known. We think nevertheless that the following solution, based on the map Eq. 3, is so simple and elegant that it deserves to be presented even if the result is not new.

Let all ϵ's be $(0, \Gamma)$ and suppose r_{j-1} is (a_{j-1}, γ_{j-1}). From Eq. (6b) and properties (ii)–(v) we have

$$r_{j-1} \;\mapsto\; z_{j-1} = a_{j-1} + i\gamma_{j-1},$$

$$-r_{j-1} \;\mapsto\; -z_{j-1}^*,$$

$$E - \epsilon_j \;\mapsto\; E + i\Gamma,$$

$$E - \epsilon_j - r_{j-1} \;\mapsto\; E + i\Gamma - z_{j-1}^*,$$

$$r_j = 1/(E - \epsilon_j - r_{j-1}) \;\mapsto\; z_j = 1/(E - i\Gamma - z_{j-1}),$$

or, introducing $w \equiv E - i\Gamma$,

$$z_j = 1/(w - z_{j-1}), \qquad (9)$$

while a similar analysis gives from (6a) $z_1 = w^{-1}$. The natural question to ask next is if the iteration of (9) will converge to a limiting point. To answer, we notice first that if a limiting point exists, it must be a solution of

$$z^2 - wz + 1 = 0. \qquad (10)$$

Equation 10 has two roots z_A and z_B, which satisfy $z_A + z_B = w$ and $z_A z_B = 1$. Thus, z_A and z_B must be located in the complex plane as shown in Fig. 1; we have called z_A the root inside the unit circle. To see if the iteration of (9) will eventually lead to z_A or z_B it is most convenient to look at (9) in another complex plane in which the fixed points have been conformally mapped into 0 and ∞. We do so by defining

$$u = (z - z_A)/(z - z_B) \qquad (11)$$

which, when inserted into (9) gives after some elementary steps

$$u_j = z_A^2 u_{j-1}. \qquad (12)$$

Since $|z_A| < 1$ the iteration converges exponentially fast to $u = 0$, corresponding to $z = z_A$. Having located the limiting point, we know p_∞ [related to z_A by Eq. (3)] and we calculate after a simple integration

$$\pi\rho(E) = \mathrm{Im}[1/(w - 2z_A)]. \qquad (13)$$

Another quantity of interest is the localization length $L(E)$ (see Ref. 1 for details) defined by

$$-\frac{1}{L(E)} \equiv \int_{-\infty}^{\infty} \log|x| p_\infty(x) dx. \qquad (14)$$

Again doing a simple integral, we get

$$L(E) = 1/-\log|z_A|. \qquad (15)$$

The whole problem has been reduced by the map Eq. 3 to the solution of a second order algebraic equation (10).

We conclude this work with the remark that the equivalent problem on a Cayley[2] tree of connectivity K (instead of a linear chain) is solved by the same method. The only difference is that (6b) must be replaced by

$$r_j = 1/\left(E - \epsilon_j - \sum_{s=1}^{K} r_s\right) \qquad (16)$$

and consequently (10) is replaced by $Kz^2 - wz + 1 = 0$.

* Laboratoire associé au CNRS.

[1] E. N. Economou and C. Papatriantafillou, Solid State Comm. 11, 197 (1972).

[2] P. W. Anderson, R. Abou Chacra, and D. J. Thouless, Cambridge preprint TCM/35/1972.

An extension of Hamilton's principle to include dissipative systems

R. M. Kiehn

University of Houston, Houston, Texas
(Received 13 April 1973)

The key idea of conservative Hamiltonian systems is the fact that the closed line integral of action is an absolute invariant of the motion. Dissipation effects may be included by considering those systems for which the closed integral of action is a parameter-dependent, conformal invariant of the motion. An application of this idea to hydrodynamics is made, and the conditions required for the validity of the Liouville theorem with respect to conformal Hamiltonian flows are examined.

INTRODUCTION

In 1922 E. Cartan proposed an extension of Hamiltonian mechanics[1] which has yet to fully penetrate the modern literature of fluid or statistical mechanics. The instances where the technique has been exposed are few in number and do not use the full power of the method.[2] Cartan proposed the principle: An (adiabatic) physical system admits a closed integral of action which is a parameter-independent invariant of the motion.

Cartan's extension essentially generalizes Poincare's notion of integral invariance to include integration domains which are not necessarily equal time point sets. However, all classical Hamiltonian methods, including Cartan's, apply only to adiabatic systems. In these classical formulations, the usual Hamiltonian force, $dp_\mu/d\tau$, is always the gradient of a single scalar potential function, and as such yields zero contribution to the cyclic work, $\oint f_\mu dq^\mu$.

Herein, it will be shown that Cartan's basic approach can be extended to include dissipative phenomena. Non-adiabatic systems, as well as systems with time-dependent Hamiltonians, can be treated from the one general principle: A physical system admits a closed integral of action which is a parameter-dependent, conformal invariant of the motion. A conformal invariant is essentially an invariant of an orientation preserving deformation.

Those vector fields which preserve conformal invariance of the closed integral of action will be studied in state space, using the current notation of the mathematics that Cartan invented.[3] A phase-space density will be propagated along such vector fields, and the Liouville question will be asked: Is the phase-space density m a differential invariant of the dissipative, conformal, Hamiltonian vector field? An affirmative answer will require two flow constraints which are not automatically true for conformal Hamiltonian flows.

PARAMETER-INDEPENDENT HAMILTONIAN FLOWS IN THE MANNER OF CARTAN

Cartan's notion of Hamiltonian mechanics is built on the idea that for many physical systems there exists an integral of the c^2 differentiable 1-form of action, $\mathfrak{A} = p_\mu dq^\mu - Hd\tau$, evaluated on a closed chain z of state space variables, q^μ, p_μ, τ; the value of the integral is an invariant of the dynamical vector field of flow $\mathbf{V}$ as long as the points which make up the closed chain z are points along the tube of trajectories of the vector field. This notion creates a relative integral invariant which is slightly more general than the idea invented by Poincare: Cartan's relative integral invariant, $\oint\mathfrak{A}$, may be evaluated by integrating around a closed curve of arbitrary connected points on the system of trajectories of $\mathbf{V}$, while Poincare's relative integral invariant must be evaluated along a closed curve through *equal time* points of the same system of trajectories. Cartan's method is independent of the time parameterization of the trajectories! Cartan's invariant characterizes the homology class of chains confined to a tube of trajectories.

For a contravariant vector field $\mathbf{V}$ the propagator of some mathematical object ω down the trajectories of the vector field is precisely the Lie derivative $\pounds_{\mathbf{V}}$. If an object is invariant with respect to propagation down the trajectories of the vector field, then its Lie derivative vanishes; i.e., $\pounds_{\mathbf{V}}\omega = 0$ implies invariance of ω w.r.t. $\mathbf{V}$.

If the components of the vector field $\mathbf{V}$ are rescaled by some c^2 function γ, then the trajectories of $\mathbf{V}$ are reparameterized by γ, but the solution curves in a Picard sense are the same; i.e., the solution curves of the system of first-order differential equations,

$$\frac{dx^1}{V^1} = \frac{dx^2}{V^2} = \cdots = \frac{dt}{V^n}, \tag{1a}$$

are the same as the solution curves for the system,

$$\frac{dx^1}{\gamma V^1} = \frac{dx^2}{\gamma V^2} = \cdots = \frac{dt}{\gamma V^n}. \tag{1b}$$

Only the time ticks along the solution curves are modified by γ. If an object, ω, is invariant w.r.t. $\gamma\mathbf{V}$ for all γ, then

$$\pounds_{\gamma\mathbf{V}}\omega = 0, \quad \text{all } \gamma, \tag{2}$$

and the object is invariant with respect to the solution curves alone; the result is independent from any parameterization (of time ticks) along the solution curve.

For differential forms, the Lie derivative admits a simple construction in terms of the exterior derivative (d) and the interior product (i) operators:

$$\pounds_{\mathbf{V}}\omega = i(\mathbf{V})d\omega + di(\mathbf{V})\omega. \tag{3}$$

Utilization of this construction permits the Cartan problem to be formulated easily.[4] A Hamiltonian flow is a vector field $\mathbf{V}$ on state space which leaves the closed integral of action an invariant for all parameterizations of the flow:

$$\pounds_{\gamma\mathbf{V}}\oint\mathfrak{A} = \oint\pounds_{\gamma\mathbf{V}}\mathfrak{A} = \pounds_{\gamma\mathbf{V}}\oint(p_\mu dq^\mu - Hd\tau) = 0, \text{ all } \gamma. \tag{4}$$

For a given $\mathfrak{A}$, the requirement (4) imposes constraints on the admissible vector fields, $\mathbf{V}$. Using (3), the constraint (4) becomes

$$\pounds_{\gamma\mathbf{V}}\oint\mathfrak{A} = \oint\pounds_{\gamma\mathbf{V}}\mathfrak{A} = \oint i(\gamma\mathbf{V})d\mathfrak{A} + \oint d(i(\gamma\mathbf{V})\mathfrak{A}) = 0. \tag{5}$$

The second integral vanishes over a cycle z as the integral is a perfect differential. When the function γ is arbitrary, it is necessary and sufficient that the integrand of the first integral vanish,

$$i(\mathbf{V})d\mathfrak{A} = 0, \tag{6}$$

and $\mathbf{V}$ is said to be an extremal field for $\mathfrak{A}$. Consider the $2N+1$ components of $\mathbf{V}$ as $\{v^\mu, f_\mu, 1\}$. Equation (6) may be explicity evaluated as,

$$\left(f_\mu + \frac{\partial H}{\partial q^\mu}\right) dq^\mu + \left(-v^\mu + \frac{\partial H}{\partial p_\mu}\right) dp_\mu + \left(-v^\mu \frac{\partial H}{\partial p_\mu} - f_\mu \frac{\partial H}{\partial q^\mu}\right) d\tau = 0. \tag{7}$$

For arbitrary $dp_\mu, dq^\mu, d\tau$, the bracket factors must vanish to yield Hamilton's equations:

$$v^\mu = \frac{\partial H}{\partial p_\mu}$$
$$f_\mu = -\frac{\partial H}{\partial q^\mu}. \tag{8}$$

Note that no statement is made about $\partial H/\partial \tau$. Cartan's principle yields Hamiltonian dynamics—the extremal field is unique.

The utility of the extremal analysis is limited to adiabatic systems (as $f_\mu = -\partial H/\partial q^\mu$), but certain trajectories may yield an energy change if $\partial H/\partial \tau$ is not zero. The fundamental flow invariant is the closed integral of action, not necessarily the energy, H, and the integration cycle z is not necessarily along a single solution curve of the motion.

A slightly less restrictive system could be established by recognizing that the first integral of (5) would vanish if the integrand were exact; i.e., consider those cases where $\gamma i(\mathbf{V})d\mathfrak{A} = -dP$. Then a search is made for those many trajectories $\mathbf{V}$ that leave $i(\mathbf{V})d\mathfrak{A}$ integrable; i.e., of the form $-dP/\gamma$. Those admissible parameterizations are those functions γ which are integrating factors for $i(\mathbf{V})d\mathfrak{A}$. Such systems may be dissipative and will be discussed collectively with another type of dissipation in that which follows.

ASSOCIATED AND EXTREMAL FIELDS

Before Cartan's concept is extended, few words are necessary about extremal and associated fields.[5] An extremal field for a form ω is a contravariant vector field $\mathbf{V}$ that satisfies the equation $i(\mathbf{V})d\omega = 0$. An associated vector field, for a form ω is a contravariant vector field $\mathbf{V}$ that satisfies the equation $i(\mathbf{V})\omega = 0$. If a vector field is both an associated and an extremal field for the form ω, then ω is a differential invariant (and, therefore, an *open* integral invariant) w.r.t. the flow $\gamma\mathbf{V}$ for all parameterizations γ; i.e.,

$$\pounds_{\gamma\mathbf{V}}\omega = i(\gamma\mathbf{V})d\omega + di(\gamma\mathbf{V})\omega = 0 + 0 = 0. \tag{9}$$

Such vector fields are of interest because the integration path of the action integral need not be closed, and yet the action integral is still an invariant of the flow, independent from any parameterization of the flow. The integral is an invariant of the homology class of chains constrained to the same system of trajectories. If the field is extremal only, then the integration end points must be fixed.

For the action one-form $\mathfrak{A}$, the Lagrange function is defined to be $L = i(\mathbf{V})\mathfrak{A}$. The additional constraint that the field be associated to the action implies that the Lagrange function L vanishes, i.e.,

$$L \triangleq (p_\mu v^\mu - H) = 0, \tag{10}$$

which in view of (6) states that

$$\left(p_\mu \frac{\partial H}{\partial p_\mu} - H\right) = 0. \tag{11}$$

Hence, a Hamiltonian vector field which is both extremal and associated to the action $\mathfrak{A}$ implies that the Hamiltonian is homogeneous of degree one in the momenta p_μ, a situation typical of the relativistic Hamiltonian for a free particle.

For such cases, the natural volume element on $2N+1$ space, $\Omega = \mathfrak{A} \wedge d\mathfrak{A} \wedge \cdots \wedge d\mathfrak{A} = \{p_\mu \partial H/\partial p_\mu - H\} dq^\mu \wedge \cdots \wedge dp_\mu \wedge \cdots \wedge d\tau$, is degenerately zero. (It is ordinarily assumed that $d\mathfrak{A} \wedge \cdots \wedge d\mathfrak{A}$ is not zero for N factors.)

Although the relativistic free particle leads to a Hamiltonian vector field which is both extremal and associated, it should be recognized that the usual variational approach for the nonrelativistic particle leads to an extremal field which is not associated. Hamilton's equations define an extremal vector field for the action, $\mathfrak{A}$; this classical Hamiltonian field may or may not be an associated vector field.

Again, a slightly less restrictive situation arises if one searches for all vector fields that leave $i(\mathbf{V})d\mathfrak{A}$ integrable, for then the admissible parameterizations are those such that $\gamma i(\mathbf{V})d\mathfrak{A} = -d(\gamma L)$.

AN EXTENSION OF CARTAN'S METHOD: CONFORMAL INVARIANCE

The previously mentioned classical cases of extremal, and associated vector field systems are special cases of broader class of vector fields characterized by the constraint relations,

$$i(\gamma\mathbf{V})d\mathfrak{A} = \Gamma\mathfrak{A} - dP, \tag{12}$$

and

$$i(\gamma\mathbf{V})\mathfrak{A} = \gamma L. \tag{13}$$

In order to justify interest in (12), extend Cartan's argument to include those vector fields $\mathbf{V}$ which leave the closed integral of action conformally invariant for a given parameterization γ. Then the definition of conformal invariance,

$$\pounds_{\gamma\mathbf{V}} \oint \mathfrak{A} = \oint \pounds_{\gamma\mathbf{V}}\mathfrak{A} = \oint \Gamma\mathfrak{A}, \tag{14}$$

requires that

$$\oint \{i(\gamma\mathbf{V})d\mathfrak{A} - \Gamma\mathfrak{A}\} = 0. \tag{15}$$

By de Rham's theorem, it is both necessary and sufficient that the integrand be a perfect differential, $-dP$; i.e.,

$$i(\gamma\mathbf{V})d\mathfrak{A} - \Gamma\mathfrak{A} = -dP, \tag{16}$$

which is the same as Eq. (12). (Of particular interest for some branches of physics are those homothetic variations where Γ is a rational constant, or perhaps a Hermitian function.)

For the vector field $\mathbf{V} = \{v^\mu, f_\mu, 1\}$, explicit evaluation

of (16) yields a modified set of Hamiltonian equations:

$$v^\mu = \frac{\partial H}{\partial p_\mu} + \left(\frac{1}{\gamma}\right)\left(\frac{\partial P}{\partial p_\mu}\right), \tag{17}$$

$$f_\mu = -\frac{\partial H}{\partial q^\mu} - \left(\frac{1}{\gamma}\right)\left(\frac{\partial P}{\partial q^\mu}\right) + \left(\frac{\Gamma}{\gamma}\right)p_\mu, \tag{18}$$

$$0 = \Gamma\left\{p_\mu \frac{\partial H}{\partial p_\mu} - H\right\} - \left(\frac{\partial P}{\partial \tau} + \{P,H\}\right), \tag{19}$$

where $\{P,H\}$ is the Poisson bracket

$$\{P,H\} = \left(\frac{\partial P}{\partial q^\mu}\right)\left(\frac{\partial H}{\partial p_\mu}\right) - \left(\frac{\partial P}{\partial p_\mu}\right)\left(\frac{\partial H}{\partial q^\mu}\right). \tag{20}$$

The trajectories in $2N+1$ state space are determined by the integration of the usual system of first-order equations,

$$\frac{dq^\mu}{v^\mu} = \frac{dp_\mu}{f_\mu} = \frac{d\tau}{1}, \tag{21}$$

but now the system admits nonadiabatic phenomena, for the force is not exact; in fact, the curl components of the force do not necessarily vanish. The three-dimensional space components of the curl of the force are given by the Gibbs' expression,

$$\mathrm{curl}\mathbf{f} = +(1/\gamma^2)\,\mathrm{grad}\gamma \times \mathrm{grad}P + \mathrm{grad}(\Gamma/\gamma) \times \mathbf{p}. \tag{22}$$

A coordinate free expression for the complete force 1-form, f, can be constructed from (12) and (13):

$$f = (\Gamma/\gamma)\mathfrak{A} - (1/\gamma)dP + dL. \tag{23}$$

The nonclosed components of f are given by the exterior derivative of (23),

$$df = d(\Gamma/\gamma) \wedge \mathfrak{A} + (\Gamma/\gamma)d\mathfrak{A} + (1/\gamma^2)d\gamma \wedge dP, \tag{24}$$

where the spacelike parts of (24) reduce to (22). The fundamental equation of motion is given by the coordinate free expression:

$$\pounds_{\mathbf{V}}\,\mathfrak{A} = f. \tag{25}$$

Following Cartan, if any propagation invariants are known, others may be obtained by sequentially operating on the Lie derivative with the exterior derivative (raising) operator and the interior product (lowering) operator. From (25) one obtains by exterior differentiation the first fundamental extension:

$$d\pounds_{\mathbf{V}}\mathfrak{A} = \pounds_{\mathbf{V}}(d\mathfrak{A}) = df. \tag{26}$$

From (25) one obtains by interior products the second fundamental extension, which incidentally measures the variation of the Lagrange function down the flow:

$$i(\mathbf{V})\pounds_{\mathbf{V}}\mathfrak{A} = \pounds_{\mathbf{V}}(L) = \Gamma/\gamma(L) - (1/\gamma)i(\mathbf{V})dP + i(\mathbf{V})dL. \tag{27}$$

From the definition of the Lie derivative, (27) yields an invariant relation between Γ and P, equivalent to (19):

$$i(\mathbf{V})\{\Gamma\mathfrak{A} - dP\} = 0. \tag{28}$$

It is also of some interest to compute the Lie derivative of the Hamiltonian function with respect to those nonadiabatic flows that satisfy (16). The variation of H down the flow becomes

$$\pounds_{\mathbf{V}}(H) = \left\{\frac{\Gamma(H)}{\gamma} + \frac{1}{\gamma}\left(\frac{\partial P}{\partial \tau}\right) + \frac{\partial H}{\partial \tau}\right\}, \tag{29}$$

which demonstrates that the Hamiltonian varies along conformal Hamiltonian flows (16) via three mechanisms: through the local time variation of H, itself—the classical result; through the local time variation of the pressure function P; and through the conformal dissipation factor, Γ/γ.

If Eqs. (26) and (28) are not evanescent, the raising and lowering processes may be continued. Of particular interest are those combinations which lead to the second Lie derivative of the action $\mathfrak{A}$:

$$\pounds_{\mathbf{V}}\pounds_{\mathbf{V}}\oint\mathfrak{A} = \oint\pounds_{\mathbf{V}}\pounds_{\mathbf{V}}\mathfrak{A} = \oint i(\mathbf{V})df + \oint d(i(\mathbf{V})dL). \tag{30}$$

Note that the only contributions to the second Lie derivative of the closed integral of $\mathfrak{A}$ come from the nonexact components of the projection of the nonclosed components of the force; i.e., from $i(\mathbf{V})df$.

The units of action are joule-sec; the first Lie derivative of the closed integral of action yields the cyclic work, with units of joules; the second Lie derivative of the closed integral of action yields the radiated power, with units of joules/sec. Equation (30) is a measure of the radiated power.

It is to be noted that special cases of (12) are well-studied problems in classical mechanics.

If γ = constant, if dP vanishes, and if $\Gamma = 0$, then (12) implies that $\mathbf{V}$ is the generator of a contract transformation of the third kind (a homogeneous contact transformation).

If γ = constant and if $\Gamma = 0$, then (12) implies that $\mathbf{V}$ is the generator of a contact transformation of the second kind (a restricted nonhomogeneous contact transformation).

If $\Gamma = -i(\mathbf{V})d\ln\gamma$, then (12) implies that $\mathbf{V}$ is the generator of a contact transformation of the first kind (an inhomogeneous contact transformation).

The conformal case (12) is a slightly broader situation than these classical systems mentioned above.

A HYDRODYNAMIC APPLICATION

In order to reinforce the physical content of the preceding theory, it is of some interest to give an example. The fundamental equation of constraint (25) that must be satisfied is essentially a coordinate free representation of Newton's equations of motion, valid in any reference system. It is analogous to the Navier–Stokes equations of hydrodynamics, but in order to easily grasp the correspondence it is necessary to alter the viewpoint slightly, such that the Lagrange–Euler equations per unit mass result from (25). Reconsider the action 1-form $\mathfrak{A}$ to be redefined as action per unit mass. The momenta per unit mass define the covariant components of velocity v_μ; the Hamiltonian per unit mass, H, becomes the sum of the kinetic energy, $\frac{1}{2}v^\mu v_\mu$, and potential energy, Φ, per unit mass—essentially defining the concept of potential energy in terms of the Hamiltonian. The Lagrange function per unit mass L becomes the difference between the kinetic and potential energies per unit mass: $L = (\frac{1}{2}v^\mu v_\mu - \Phi)$. For a Euclidean metric, the Lie derivative on the action per unit mass yields spatial components of the form,

$$\pounds_{\mathbf{V}}\mathfrak{A} = \frac{\partial \mathbf{v}}{\partial \tau} + \tfrac{1}{2}\mathrm{grad}(\mathbf{v}\cdot\mathbf{v}) - \mathbf{v}\times \mathrm{curl}\mathbf{v} + \mathrm{grad}\Phi + \mathrm{grad}L. \tag{31}$$

Equating this expression to the rhs of (25) and rearranging terms leads to the Navier–Stokes format for the spacelike components:

$$\frac{\partial \mathbf{v}}{\partial t} + \tfrac{1}{2}\,\mathrm{grad}(\mathbf{v}\cdot\mathbf{v}) - \mathbf{v}\times \mathrm{curl}\mathbf{v} = -\,\mathrm{grad}\Phi - \mathrm{grad}\frac{P}{\rho} - \frac{\Gamma}{\rho}\mathbf{v}. \tag{32}$$

The parameterization γ has been relabeled ρ and is interpreted as a mass density distribution for the action per unit mass $\mathfrak{A}$.

The recognition that the spatial components of $d\mathfrak{A}$ yield the components of vorticity allows the first fundamental extension (26) to be interpreted as Helmholtz' constraint on the conservation of vorticity. The spatial restriction of (26) yields

$$\pounds_{\mathbf{V}}(d\mathfrak{A}) = 1/\rho^2\ \mathrm{grad}\rho \times \mathrm{grad}P + \mathrm{curl}(\Gamma/\rho\mathbf{v}). \tag{33}$$

For a barytropic $[P = P(\rho)]$ fluid, which is dissipation free in the sense that $\Gamma = 0$, the rhs of (33) [and (26)] vanishes, implying that the vorticity is an invariant of the motion. Note that a substance need not be dissipative in the conformal sense ($\Gamma = 0$), and yet vorticity may be created and destroyed, so long as the equation of state is not barytropic ($d\rho \wedge dP \neq 0$). The meteorological atmosphere is an example where such effects are dominant.

The Bernoulli theorem arises from the second fundamental extension, (27) and (28), for when (28) is combined with (29), and it is assumed that the fluid is isentropic ($dP/\rho = d\psi$), a fundamental invariant (Bernoulli's constant) arises for conformally nondissipative ($\Gamma = 0$), steady ($\partial P/\partial\tau = 0$, $\partial H/\partial\tau = 0$) flows:

$$\pounds_{\mathbf{V}}\{H + \Psi\} = \left\{\frac{\Gamma}{\rho}\,\frac{v_\mu v^\mu}{2} + \frac{1}{\rho}\frac{\partial P}{\partial \tau} + \frac{\partial H}{\partial \tau}\right\}. \tag{34}$$

The Bernoulli invariant is the usual enthalpic combination of $\{\frac{1}{2}v_\mu v^\mu + \Phi + \Psi\}$, which is an invariant of the flow if the rhs of (34) vanishes. It is remarkable that the Helmholtz theorem must be violated, if power is to be radiated from the system; that is, (30) implies that vorticity must be created or destroyed if there is to exist any net radiated power from the system. The failure of Helmholtz's theorem is necessary if power is to be radiated, but not sufficient, for there may exist trajectories $\mathbf{V}$ along which the projections of df are exact.

HOMOTOPIES AND CHARACTERISTICS

At this point in the general theory presented above, no constraint on parameterization has been made, but if one asserts that (14) be true for all γ, then it may be argued that P must vanish; however, in that which follows, P will be retained for completeness. A particularly interesting physical case may be made for those situations where P is defined to be the function, $-\gamma L$, for then (12) becomes the defining equation for an equivalence class of vector fields which may be viewed as generators of a homotopy; $i(\gamma\mathbf{V})$ is the homotopy operator. It is this correspondence which gives credence to the concept that a conformal invariant is essentially an invariant of an orientation preserving deformation.

When an equation of state of the form $P - \gamma L = 0$ exists, then $\gamma\mathbf{V}$ is a strictly conformal vector field for the action $\mathfrak{A}$; i.e., $\pounds_{\mathbf{V}}\mathfrak{A} = \Gamma\mathfrak{A}$. If the additional constraint $i(\mathbf{V})\mathfrak{A} \equiv L = 0$ is imposed, then $\mathbf{V}$ is a characteristic vector field which by definition must be both conformal and associated. Characteristic vectors are integrable in the sense that the Lie bracket of two characteristic vectors of any form ω is an associated vector for the form ω. Note that the vector field of the relativistic free particle is a special case of the characteristic system. That is, an extremal field which is also associated, is a characteristic field with a conformal factor Γ equal to zero.

For a characteristic system of the action $\mathfrak{A}$, Eqs. (10) and (17) yield the relation

$$\left\{p_\mu \frac{\partial H}{\partial p_\mu} - H\right\}\gamma = -\,p_\mu \frac{\partial P}{\partial p_\mu}. \tag{35}$$

As long as P is a function of momenta, the Hamiltonian need not be homogeneous of degree one in momenta, a result to be compared with the *free* relativistic particle case given by (11).

The characteristic vector fields of the action $\mathfrak{A}$ are a special subclass of vector fields covered by the Eqs. (12) and (13). They are of particular interest, geometrically speaking, for from (19) it may be determined that P (and hence γL) is a *singular* surface function in state space; i.e., for the characteristic case, both $P = 0$ and $dP/d\tau = 0$.

THE LIOUVILLE THEOREM: MASS INVARIANCE

Consider now a density on phase space.

$$m = \mu(q^\sigma, p_\sigma, \tau)dq^1 \wedge \cdots \wedge dq^N \wedge dp_1 \wedge \cdots \wedge dp_N, \tag{36}$$

and the extension of the Liouville question: "Is m a differential invariant with respect to the conformal Hamiltonian flow?" For invariance of m w.r.t. any parameterized flow, $\gamma\mathbf{V}$, it follows that

$$\pounds_{\gamma\mathbf{V}}(m) = \beta(m) + \left\{\frac{\partial(\mu V^\sigma)}{\partial \tau}\right\} d\tau \wedge dq^1 \wedge \cdots \wedge \widehat{dq^\sigma} \wedge \cdots \wedge dp_N = 0, \tag{37}$$

where β is defined to be

$$\beta = \left\{\frac{\gamma}{\mu}\frac{d\mu}{d\tau} + \frac{\partial(\gamma v^\sigma)}{\partial q^\sigma} + \frac{\partial(\gamma f_\sigma)}{\partial p_\sigma}\right\}. \tag{38}$$

In order for the Liouville theorem to be true, two conditions must be satisfied:

(a) The equation of continuity must be satisfied; $\beta = 0$.
(b) The flow must be autonomous in state space; $\partial(\gamma\mathbf{V}/\partial\tau) = 0$.

For a flow which satisfies (14), the constraint of continuity requires that:

$$\beta = 0 = \left[\frac{\gamma}{\mu}\frac{d\mu}{d\tau} + \{\gamma, H\} + \left(\Gamma + p_\sigma \frac{\partial \Gamma}{\partial p_\sigma}\right)\right], \tag{39}$$

which surprisingly does not include P or L, explicity. For the classical case ($\Gamma = 0$), it is usually assumed that the Poisson bracket of the parameterization γ and the Hamiltonian vanishes; then the classical result is retrieved from (39): if the flow is continuous ($\beta = 0$), then the density μ is constant, and conversely.

In the general case, continuity does not imply that the phase-space density is constant. For a simple case where it is assumed that $\{\gamma, H\} = 0$ and $p_\sigma\,\partial\Gamma/\partial p_\sigma = 0$, (39) implies that the phase-space density decreases in time in an exponential manner, where the decay constant is precisely the conformality factor per unit parameterization, Γ/γ:

$$\frac{d(\ln\mu)}{d\tau} = \frac{-\Gamma}{\gamma}. \tag{40}$$

The dissipative components of the force are related by the same factor Γ/γ to the components of momentum [see (23)]; the Hamiltonian changes via the same factor [see (26)]. All of these results are in direct correspondence to the classical results[6] which assume a velocity-dependent viscous dissipation.

The conclusion is reached that dissipative effects can be accommodated in Hamiltonian mechanics by studying systems which leave the closed integral of action conformally invariant.

Now consider the second condition required for the validity of the Liouville theorem. The second condition implies that the admissible flow fields are generators of a group of motions in $2N$ space, parameterized by τ; i.e., the functions v^σ and f_σ just admit factorization into products of the functions of p_σ and q^σ and a function of τ. If the factorization is possible, then γ may be chosen such as to annihilate the time factor, and the resulting expression will be independent of τ. From Lie's theorem, the factorization concept is both necessary and sufficient for the vector field to admit group properties, and from (37) it is observed that the group property is thus a necessary condition for Liouville's theorem to be true.

It is interesting to see that the $2N$ equations required for the second condition,

$$\mu\frac{\partial(\gamma v^\sigma)}{\partial\tau} = 0, \qquad \mu\frac{\partial(\gamma f_\sigma)}{\partial\tau} = 0, \tag{41}$$

may be combined to yield a single necessary equation in terms of a scalar function Φ:

$$\Phi = \gamma\mu\ f_\sigma\left\{\frac{\partial v^\sigma}{\partial\tau} - v^\sigma\frac{\partial f_\sigma}{\partial\tau}\right\} = 0. \tag{42}$$

The function Φ is in some sense a measure of the lack of the group property for the flow; note that Φ involves more than the local time variation of the power density, $f_\sigma v^\sigma$. For the classical case, where the Hamiltonian is steady, Φ vanishes automatically in virtue of (8); if $\partial H/\partial\tau$ is not zero, Φ could still vanish through subtraction, orthogonality, or both. For the steady dissipative case, Φ may also vanish. The first situation is the case for potential flow, and the second situation is the case for viscous laminar flow.

It is conjectured that $\Phi \neq 0$ may be a measure of turbulence; in this sense the lack of validity of the Liouville theorem indicates a turbulent regime. ($\beta \neq 0$ might be interpreted as cavitation.)

For the conformal flows given by (17) and (18) a rather formidable expression, equivalent to (42), may be derived for the function Φ in terms of γ, H, P, μ, and Γ. If $\Phi(\gamma, H, P, \mu, \Gamma)$ vanishes, then the necessary condition that the system admits the group property in time is satisfied. If the flow is continuous, then the function $\beta(\gamma, H, \mu, \Gamma)$ must vanish. If the Liouville theorem is to be valid, both constraints must be satisfied. Note that Φ depends on the pressure function, P, where β does not.

NUMBER INVARIANCE

It is of interest to ask if the requirements of integral invariance of the phase-space density w.r.t. a parameter independent flow are different from the two conditions obtained above for differential invariance. It is possible to distinguish between the two concepts of total mass and total number. Total mass M will be defined as the integral of m over an open $2N$-dimensional integration chain. Total number, N, will be defined as the integral of m over a $2N$-dimensional closed integration chain. These concepts are indeed distinct because they have different invariance structures. Certainly, the Liouville requirements for differential invariance imply both invariance of total mass M and total number N. However, invariance of total number does not imply invariance of total mass.

Specifically, ask: Is the integral of m over a $2N$ cycle an invariant with respect to a parameterized flow, $\gamma\mathbf{V}$? It follows that

$$\pounds_{\gamma\mathbf{V}}\oint m = \oint\pounds_{\gamma\mathbf{V}}m = \oint i(\gamma\mathbf{V})dm = 0, \tag{43}$$

for invariance, which implies that

$$\frac{\partial\mu}{\partial\tau}\, i(\gamma\mathbf{V})\Omega = dg; \tag{44}$$

Ω is defined to be the $2N+1$ volume $d\tau\wedge dq^1\wedge\cdots\wedge dp_N$. From (44), for invariance of N, it is necessary that $\gamma\partial\mu/\partial\tau$ be an integrating factor for the flow $\mathbf{V}$ in the sense that the flow $(\gamma\partial\mu/\partial\tau)\mathbf{V}$ is divergence free; i.e., for invariance of N,

$$\frac{\partial(\gamma(\partial\mu/\partial\tau)v^\sigma)}{\partial q^\sigma} + \frac{\partial(\gamma(\partial\mu/\partial\tau)f_\sigma)}{\partial p_\sigma} + \frac{\partial(\gamma(\partial\mu/\partial\tau))}{\partial\tau} = 0. \tag{45}$$

These conditions of total number invariance are *not* equivalent to the two conditions required by the Liouville theorem. In fact, the simple case for which $\partial\mu/\partial\tau = 0$ leads to invariance of total number, N, for any flow, while the total mass may not be conserved by the flow at all.

In the sense conjectured above, a turbulent flow does not preserve total mass, but it could preserve total number.

SUMMARY

By considering those trajectories which allow the closed integral of action to vary in a conformal manner, an extension of the classical Hamiltonian formalism to include dissipation and pressure effects has been achieved. The results have been utilized to show that continuity is explicitly dependent on the dissipation (conformality) function Γ but is explicitly independent of pressure P. On the other hand, the flow group property (no turbulence) depends on both P and Γ. The pressure function P is intimately connected with a choice of parameterization and must be zero if the conformal invariance of the relative integral of action is absolute (parameter independent). An equation of motion, valid in inertial or non-inertial frames, is obtained for the conformal Hamiltonian flows, and frame-independent realizations of Bernoulli's, Helmholtz's, and the radiated power theorems have been constructed.

[1]E. Cartan, *Lecons sur les invariants integraux* (Hermann, Paris, 1958).

[2]W. Slebodzinski, Rocz. RPTMA Polskiego Towarzystwa Matematycznego, Ser. 1: Prace Matematyczne 14, 1 (1970).

[3]L. H. Loomis and S. Sternberg, *Advanced Calculus* (Addison-Wesley, Reading, Mass., 1968); H. Flanders, *Differential Forms* (Academic, New York, 1963); R. L. Bishop and S. Goldberg, *Tensor Analysis on Manifolds* (Macmillan, New York, 1968).

[4]The action of the Lie derivative with respect to a vector field $\mathbf{V}$ on an integral of a differential form ω over a chain c is defined by the equation, $\pounds_{\mathbf{V}}\int_c\omega \triangleq \lim u\to 0\{[\int_{c(u)}\omega x(u) - \int_{c(o)}\omega x(o)]/u\}$, which may be used to prove that $\pounds_{\mathbf{V}}\int_c\omega = \int_c\pounds_{\mathbf{V}}\omega$; see Refs, 3 and 5. For closed integration chains, the symbol $\oint$ is utilized herein. Closed chains as well as open chains may be deformed by a flow.

[5]J. Klein, Ann. Inst. AIFUB Fourier, Grenoble 12, 1 (1962).

[6]J. Jeans, *Dynamic Theory of Gases* (Dover, New York, 1954), p. 68.

Debye potentials in Riemannian spaces

H. Stephani

Fachbereich Theoretische Physik, Friedrich-Schiller-Universität, Jena, Germany
(Received 18 April 1972)

By means of Debye potentials it is possible to get all solutions of source-free Maxwell equations in vacuum from a single scalar equation. Sufficient conditions for the existence of Debye potentials in a given four-dimensional Riemannian space have been found. Some examples of metrics are given, including plane gravitational waves, metrics with spherical symmetry, and cosmological models. The method is generalized to Maxwell fields with sources and Maxwell fields in dielectric media.

1. INTRODUCTION

The integration of the source-free Maxwell equations

$$F^{in}{}_{;n} = 0, \quad \tilde{F}^{in}{}_{;n} = 0, \qquad F^{in} \equiv \tfrac{1}{2}\epsilon^{inab}F_{ab}, \quad i,n = 1,\ldots,4 \tag{1}$$

in curvilinear coordinates or in a given gravitational field is rather a complicated procedure, because (1) is a strongly coupled system of differential equations. Introduction of the potentials

$$F_{am} = A_{m,a} - A_{a,m} \tag{2}$$

does not alter this essentially. Using Debye potentials, the Maxwell equations are fully decoupled and proved to be equivalent to a single scalar equation. This was done for flat space[1,2] and for static gravitational fields with spherical symmetry.[3] The aim of this paper is to give a covariant formulation of this method and to give an invariant characterization of all spaces admitting this method.

The conform invariance of Maxwell equations ensures that the four-potential A_n of the general solution in a metric ds^2 also is the four potential of the general solution in all conformally equivalent metrics $d\bar{s}^2 = M^2(x^i)ds^2$. All calculations and the invariant characterizations of the following chapters are assumed to concern ds^2, whereas for applications in a given metric $d\bar{s}^2$ we have to look for a suitable factor M connecting both.

2. DEBYE POTENTIALS OF SOURCEFREE MAXWELL FIELDS

A. Four-potential and fields

Performing a gauge transformation, the potential of an arbitrary field may be written as

$$A_a = \Pi_{,n}(u^n v_a - v^n u_a) + \epsilon_a{}^{bin}\Phi_{,b}v_i u_n + Pu_a, \tag{3}$$

if the vector fields u^i and v^i fulfill

$$v_i = v_{,i}, \quad u_i = u_{,i}, \quad u_n v^n = 0. \tag{4}$$

The corresponding fields are

$$\begin{aligned} F_{am} = {} & (\Pi_{,n}u^n)_{,a}v_m - (\Pi_{,n}v^n)_{,m}v_a - (\Pi_{,n}v^n)_{,a}u_m \\ & + (\Pi_{,n}v^n)_{,m}u_a + P_{,a}u_m - P_{,m}u_a \\ & + (\epsilon_m{}^{bin}\Phi_{,ba} - \epsilon_a{}^{bin}\Phi_{,bm})v_i u_n. \end{aligned} \tag{5}$$

It may be easily proven, that the v^i component of the "electric" field $F_{ab}v^a u^b$ does not depend on Φ, and the v^i component of the "magnetic" field $F_{ab}v^a u^b$ not on Π and P.

B. Conditions for existence of Debye potentials

Source-free Maxwell fields have only two (transversal) degrees of freedom; so one would try to omit the function P. If this is possible and the remaining functions Π and Φ are subject to two decoupled differential equations, Π and Φ are the Debye potentials. Because we are interested in decoupled equations only, we, first, may put Φ equal to zero and analyze the remaining Maxwell equations

$$\begin{aligned} 0 = {} & F_a{}^m{}_{;m} \\ = {} & (\Pi_{,n}u^n)_{,a}v^m{}_{;m} - (\Pi_{,n}u^n)_{,}{}^{,m}{}_{;m}v_a - 2(\Pi_{,n}u^n)_{;}{}^{m}v_{a;m} \\ & + (\Pi_{,n}v^n - P)_{,}{}^{,m}{}_{;m}u_a - (\Pi_{,n}v^n - P)_{,a}u^m{}_{;m} \\ & + 2(\Pi_{,n}v^n - P)_{,}{}^{m}u_{a;m} \\ & + [(\Pi_{,n}u^n)_{,m}v^m - (\Pi_{,n}v^n - P)_{,m}u^m]_{,a}. \end{aligned} \tag{6}$$

These four differential equations (6) reduce to one, as we will prove later, if (6) has the structure

$$0 = Kv_a + Lu_a + N_{,a} \tag{7}$$

for all functions Π. With the help of the conform factor M, we can make the vector field $u_i = u_{,i}$ satisfy $u^i u_{i;n} = 0$. Then (7) is valid only if the fields u_i and v_i obey either

$$\begin{aligned} & v_{i;n} = \tfrac{1}{2}\theta h_{in}, \quad u_{i;n} = \tfrac{1}{2}Vh_{in}, \quad u_i u^i = \epsilon_3 = \pm 1, \\ & h_{in} = g_{in} - \epsilon_2 v_i v_n - \epsilon_3 u_i u_n, \quad v_i v^i = \epsilon_2 = \pm 1, \end{aligned} \tag{8a}$$

which means the metric may be written as

$$ds^2 = H^2(x,y,u,v)[dx^2 + \epsilon_1 F^2(x,y)dy^2] + \epsilon_2 dv^2 + \epsilon_3 du^2, \quad \epsilon_1\cdot\epsilon_2\cdot\epsilon_3 = -1, \tag{8b}$$

or

$$\begin{aligned} & v_{i;n} = Bv_i v_n, \quad v_i v^i = 0, \\ & v_{i;n} = Cv_i v_n, \quad u_i u^i = 1, \end{aligned} \tag{9a}$$

with the adapted form of the metric

$$ds^2 = du^2 + dy^2 + 2dzdv + 2H(u,y,z,v)dv^2. \tag{9b}$$

C. Derivation of Debye equation

We now substitute (8a) in (6). The calculations with vector fields of type (9a) run along the same lines, the corresponding results will be given in Sec. 2D.

Equation (6) now reads

$$\begin{aligned} 0 = {} & u_a[(\Pi_{,n}v^n - P)_{,}{}^{,m}{}_{;m} - \epsilon_3 V(\Pi_{,n}v^n - P)_{,m}u^m \\ & + \epsilon_3\theta(\Pi_{,n}u^n)_{,m}u^m] + v_a[-(\Pi_{,n}u^n)_{,}{}^{,m}{}_{;m} \\ & + \epsilon_2\Theta(\Pi_{,n}u^n)_{,m}v^m - \epsilon_2 V(\Pi_{,n}v^n - P)_{,m}v^m] \\ & + (P_{,m}u^m)_{,a}. \end{aligned} \tag{10}$$

It shows that $P_{,m}u^m$ is a function of u and v only. With help of a gauge transformation

$$\bar{A}_a = A_a + h(u,v)_{,a}, \tag{11}$$

which changes Π but does not alter Φ, we can achieve $P = 0$. This result remains true if we start with the combination Φ, P, putting Π equal to zero: P can be omitted completely. Maxwell equations including Π and Φ then take the form

$$[(u^a v^m - v^a u^m) D(\Pi)]_{;m} - \epsilon^{ampq} D(\Phi)_{,m} v_p u_q = 0, \tag{12}$$

where the differential operator D is defined by

$$D(\Pi) \equiv \Pi^{,n}_{;n} - \epsilon_2 \theta \Pi_{,n} v^n - \epsilon_3 V \Pi_{,n} u^n. \tag{13}$$

According to (12), $D(\Phi)$ is a function of v and u:

$$D(\Phi) = f(u,v). \tag{14}$$

A special solution of this inhomogeneous differential equation is a function Φ_0 depending on u and v; but this Φ_0 gives no contribution to the potential A_i. Combining this with the analogous result for Π, we see that the general solution of Maxwell equations is described by Debye potentials Π and Φ satisfying the Debye equation $D(\Pi) = D(\Phi) = 0$. The only possible exceptions are fields with $D(\Pi) = \text{const} H^2$ or $D(\Phi) = \text{const}$, which occur if the x–y resp. u–v surfaces are closed.

D. Summary

Decoupling of Maxwell equations by means of Debye potentials is possible in spaces conformally equivalent to spaces characterized either by (8a) and (8b), with Debye operator

$$D(\Pi) = \Pi^{,n}_{;n} - \epsilon_2 \theta \Pi_{,n} v^n - \epsilon_3 V \Pi_{,n} u^n, \tag{15a}$$

or by (9a) and (9b) with Debye operator

$$D(\Pi) = \Pi^{,n}_{;n} + 2B \Pi_{,n} v^n. \tag{15b}$$

To get the general solution of Maxwell equations, one has to solve the Debye equations

$$D(\Pi) = D(\Phi) = 0. \tag{16}$$

with the exceptions mentioned above. Four-potential

$$A_a = \Pi_{,n}(u^n v_a - v^n u_a) + \epsilon_a{}^{bin} \Phi_{,b} v_i u_n \tag{17}$$

and fields then follow by mere differentiation. For static fields (independent of u), $\Pi_{,n} v^n$ and $\Phi_{,n} v^n$ are the electrostatic and the scalar magnetic potential, respectively.

3. EXAMPLES OF SPACES

We now give some examples of important spaces admitting the method of Debye potentials. According to the preceding chapter, their metric can be written in either of the forms:

$$\begin{aligned} d\bar{s}^2 &= M^2 ds^2 = M^2(x,y,u,v)[H^2(x,y,u,v) \\ &\quad \times \{dx^2 + \epsilon_1 F^2(x,y) dy^2\} + \epsilon_2 dv^2 + \epsilon_3 du^2], \\ d\bar{s}^2 &= M^2 ds^2 = M^2(z,y,v,u)[du^2 + dy^2 + 2dzdv \\ &\quad + 2H(z,y,v,u) dv^2]. \end{aligned} \tag{18}$$

From the symmetry of these line elements, one can easily see that if a space belongs to these classes at all, there are at least two pairs of vectors satisfying (8a) or (8b) with two pairs of Debye potentials and two possibly different Debye equations, namely the vectors attached to the u–v and x–y hypersurfaces, resp. u–v and y–v directions. In addition, it should be kept in mind that the Debye equation is a covariant one. So—the vector fields u^i and v^i being given—they may be solved in any coordinate system of the space ds^2.

A. Flat space

Cartesian coordinates

$$d\bar{s}^2 = ds^2 = dx^2 + dy^2 + dv^2 - du^2. \tag{19}$$

Spherical coordinates

$$d\bar{s}^2 = ds^2 = v^2(dx^2 + \sin^2 x dy^2) + dv^2 - du^2, \tag{20a}$$

$$d\bar{s}^2 = M^2 ds^2 = M^2[M^{-2}(dr^2 - dt^2) + dv^2 + du^2], \quad M = r[2e^u/1 + e^{2u})] \tag{20b}$$

Cylindrical coordinates

$$d\bar{s}^2 = M^2 ds^2 = e^{v2}[e^{-2v}(dz^2 - dt^2) + dv^2 + du^2]. \tag{21}$$

B. Spaces conformally flat

Gravitational fields which are conformally flat[4] include the important cosmological models with Robertson–Walker metric, which are conformally equivalent to the Einstein cosmos

$$d\bar{s}^2 = ds^2 = \sin^2 v(dx^2 + \sin^2 x dy^2) + dv^2 - du^2. \tag{22}$$

By investigating properties of electromagnetic fields in this cosmos, Infeld and Schild[5] found as an "accident" that the potential can be constructed from the solution of a scalar equation which is exactly the Debye equation.

C. Plane gravitational waves

They are a special case of the line element (9b) resp. (18).

D. Spaces conformally decomposible

These spaces are characterized by the line element

$$d\bar{s}^2 = M^2(x^i)[dx^2 + F^2(x,y) dy^2 + dv^2 - K^2(u,v) du^2]. \tag{23}$$

Vacuum solutions of Einstein equations belonging to this class are discussed in detail by Petrov.[6] Special cases are the static degenerated solutions[7] including the exterior Schwarzschild solution.

The class of nonvacuum solutions of this type contains all metrics with spherical symmetry (static or not), e.g., Reissner–Weyl and interior Schwarzschild, and many others, e.g. the cosmological model[8]

$$d\bar{s}^2 = t ds^2 = [\sqrt{t}\, A(z) + B(z)]^2 dz^2 - dt^2 + t(dv^2 + du^2). \tag{24}$$

4. DEBYE POTENTIALS OF ELECTROMAGNETIC FIELDS IN DIELECTRICS

The medium does not alter the degrees of freedom of the field. So it may be possible to describe the field with help of only two scalar functions satisfying two, in general, different equations closely related to Debye equations, if the medium fulfills certain conditions.

In the "physical" space $d\bar{s}^2$ the medium will be described by its 4-velocity $\bar{u}^i = M^{-1}u^i$ and the dielectric constant $\bar{\epsilon}^{in} = M^{-2}\epsilon^{in}$. In the metric $ds^2 = M^{-2}d\bar{s}^2$, to which all further calculations are referred, Maxwell equations then read

$$\tilde{F}^{mn}{}_{;n} = 0, \quad H^{mn}{}_{;n} = 0, \qquad H^{mn} = F^{mn} + F_{ia}u^a[u^m(\epsilon^{in} - g^{in}) - u^n(\epsilon^{im} - g^{im})]. \tag{25}$$

An analysis in the metric (8a) and (8b) shows that ϵ^{in} should fulfill the condition

$$\epsilon^{in} = a(x^k)h^{in} + b(x^k)v^iv^n, \quad a_{,n}h^{in} = 0. \tag{26}$$

This means that the medium is an inhomogeneous and anisotropic one, the preferred direction coinciding with v^i. The four potential is found to be

$$A_i = \Pi_{,n}(u^nv_i - v^nu_i) + \epsilon_i{}^{bmn}\Phi_{,b}v_mu_n + (a-1)a^{-1}\Pi_{,n}v^nu_i, \tag{27}$$

and the modified Debye equations are

$$D(\Pi) + (b-1)\Pi_{,in}h^{in} - ba^{-2}a_{,i}v^i\Pi_{,n}v^n = 0, \qquad D(\Phi) + [(1-a)\Phi_{,n}u^n]_{,i}u^i = 0. \tag{28}$$

Whether in flat space Debye potentials have been used in dielectric media of type (26), is not within the knowledge of the author. Dielectric media with $a = b$ in Schwarzschild space have been treated by Mo and Papas.[3]

5. DEBYE POTENTIALS OF FIELDS WITH SOURCES

Up to now we were concerned with source-free fields or source-free simply connected regions of space. If one of the two vector fields, e.g., u^i, is covariant constant (in ds^2, not in the physical space $d\bar{s}^2$)

$$v_{i;n} = \tfrac{1}{2}\theta h_{in}, \qquad u_{i;n} = 0, \tag{29}$$

our method can be generalized to include fields with current density j^m

$$\tilde{F}^{ma}{}_{;a} = 0, \quad F^{ma}{}_{;a} = j^m. \tag{30}$$

To do this we have to start from the potential (3), P not being omissible because of the new degree of freedom, and write the current in the form

$$j^m = [(v_mu^n - u_mv_m^n)T]_{;n} + \epsilon_m{}^{bin}S_{,b}v_iu_n - (W_{,n}u^n)_{,m} + W^{,n}{}_{;n}u_m \tag{31}$$

which ensures the validity of the equation of continuity. Maxwell equations then turn out to be equivalent to

$$D(\Pi) = -T, \quad D(\Phi) = -S, \quad P = -W. \tag{32}$$

The first and the second are, of course, the generalized Debye equations and the third one is the generalization of Poisson equation

$$P = j_mu^m.$$

To extract T, S, and W from a given current density, one has to solve

$$[(W_{,a}u^a)_{,i}h^{in}]_{;n} = [j_m(g^{mn} - h^{mn})]_{;n}, \tag{33a}$$

$$T_{,a}u^a = \epsilon_2 j_mv^m + W_{,am}u^av^m, \tag{33b}$$

$$(Tv^a)_{;a} = -\epsilon_3 j_mu^m - \epsilon_3 W_{,am}u^au^m + W^{,a}{}_{;a}, \tag{33c}$$

$$S_{,a}h^{ia} = \epsilon^{mrsa}(j_m + W_{,im}u^i)u_rv_s \tag{33d}$$

in the following way: Take the general solution W of (33a), which contains an arbitrary function $W_0(u, v)$, and determine W_0 together with T from (33b) and (33c). As in flat space,[2] this is rather a complicated procedure.

ACKNOWLEDGMENTS

I have to thank Dr. G. Neugebauer and Dr. D. Kramer for valuable suggestions and discussions.

[1]H. Bremmer, in *Handbuch der Physik* edited by S. Flügge (Springer, Berlin, 1958), Bd. XVI, p. 433.
[2]W. Franz, *Theorie der Beugung elektromagnetischer Wellen* (Springer, Berlin, 1958).
[3]T. C. Mo and C. H. Papas, Phys. Rev. D 3, 1708 (1971).
[4]H. Stephani, Commun. Math. Phys. 5, 337 (1967).
[5]L. Infeld and A. Schild, Phys. Rev. 70, 410 (1946).
[6]A. S. Petrow, *Einsteinräume.* (Akademie Verlag, Berlin, 1964).
[7]J. Ehlers and W. Kundt, in *Gravitation,* edited by L. Witten (Wiley, New York, 1962).
[8]H. Stephani, Commun. Math. Phys. 9, 53 (1968).

Invariant imbedding and the resolvent of Fredholm integral equations with variable parameters*

R. Bellman

Departments of Mathematics, Electrical Engineering, and Medicine, University of Southern California, Los Angeles, California 90007

S. Ueno

Department of Electrical Engineering, University of Southern California, Los Angeles, California 90007

(Received 10 April 1972)

In recent years Bellman and Krein have given independently a geometrical invariant imbedding of the resolvent $K(t, y, x)$ $(0 \le t, y \le x)$ of Fredholm integral equations with continuous kernels. In the case of symmetric kernels and constant parameters, the Cauchy systems for the symmetric resolvents have been discussed by several authors. In this paper, when the parameter depends upon the independent variable, we show how to solve exactly the Fredholm integral equation with the aid of invariant imbedding. In other words, by reducing the kernel to symmetric kernel, the Bellman-Krein formula for the Fredholm resolvent permits us to reduce the two-point boundary value problem to the initial value problem.

1. INTRODUCTION

It is well known that the resolvent kernel $K(t, y, x)$ $(0 \le t, y \le x)$ plays a significant role in the theory of linear integral equations (cf. Courant and Hilbert[1]). Bellman[2] and Krein[3] have provided independently a geometrical invariant imbedding of the resolvent K of a Fredholm integral equation with constant parameter and continuous kernel. Furthermore, several authors have shown how to use effectively an initial value method for the determination of the Fredholm integral equations with constant parameters and displacement kernels (cf. Sobolev[4]; Ueno[5]; Bellman, Kalaba, and Ueno[6]; Kagiwada and Kalaba[7,8]; Hummer and Rybicki[9]; Bellman, Kalaba, and Ueno[10]).

The aim of the present paper is to show how the initial-value method for resolvents of Fredholm integral equations with kernels reducible to symmetric kernels provides a feasible numerical solution using high-speed digital computers, without referring to a system of functional equations for the global functions due to the polarity.

2. BELLMAN-KREIN FORMULA

Consider the integral equation for the function $u(t, x)$,

$$u(t, x) = g(t) + \lambda(t) \int_0^x k(t, y)u(y, x)dy, \tag{1}$$

where $g(t)$ is a given forcing function, $0 < \lambda(t) \le 1$ is an inhomogeneous parameter, and $k(t, y)$ is a positive displacement kernel; e.g.,

$$k(t, y) = k(|t - y|), \tag{2}$$

and

$$k(\gamma) = \int_a^b e^{-\gamma/z}\, dw(z), \qquad \gamma \ge 0, \ 0 \le a < b. \tag{3}$$

This case is of great importance in the study of radiative transfer in terrestrial, planetary, and stellar atmospheres. The solution $u(t, x)$ is uniquely defined for x sufficiently small.

On multiplying both sides of Eq. (1) by $1/\sqrt{\lambda}(t)$ and introducing a new function

$$V(t, x) = u(t, x)/\sqrt{\lambda}(t), \tag{4}$$

we obtain the integral equation

$$V(t, x) = G(t) + \int_0^x L(t, y)V(y, x)dy, \tag{5}$$

where

$$G(t) = g(t)/\sqrt{\lambda}(t), \tag{6}$$

and L is the symmetric kernel

$$L(t, y) = [\lambda(t)\lambda(y)]^{1/2}k(t, y). \tag{7}$$

With the aid of the Fredholm resolvent $K(t, y, x)$, the solution of Eq. (5) is expressed as

$$V(t, x) = G(t) + \int_0^x K(t, y, x)G(y)dy. \tag{8}$$

The resolvent is symmetric, i.e.,

$$K(t, y, x) = K(y, t, x), \tag{9}$$

because $L(t, y)$ is symmetric with respect to t and y.

In what follows, we shall derive the Cauchy system for the resolvent $K(t, y, x)$. On differentiating Eq. (8) with respect to x, we have

$$V_x(t, x) = K(t, x, x)G(x) + \int_0^x K_x(t, y, x)G(y)dy. \tag{10}$$

On the other hand, differentiating Eq. (5) with respect to x, we obtain

$$V_x(t, x) = L(t, x)V(x, x) + \int_0^x L(t, y)V_x(y, x)dy \tag{11}$$

Letting $\Phi(t, x)$ be the solution of the integral equation

$$\Phi(t, x) = L(t, x) + \int_0^x L(t, y)\Phi(y, x)dy \tag{12}$$

and comparing Eqs. (11) and (12), we have

$$V_x(t, x) = \Phi(t, x)V(x, x) = \Phi(t, x)\left(G(x) + \int_0^x K(x, y, x) \times G(y)dy\right). \tag{13}$$

A comparison of Eqs. (10) and (13) yields

$$K(t, x, x) = \Phi(t, x) \tag{14}$$

and

$$K_x(t, y, x) = \Phi(t, x)K(x, y, x). \tag{15}$$

Taking into account the symmetric behavior of the

Fredholm resolvent $K(t, y, x)$, we obtain the required Bellman–Krein formula

$$K_x(t, y, x) = \Phi(t, x)\Phi(y, x), \tag{16}$$

where for the sake of definiteness we suppose max $(t, y) = y$.

Equation (16) is to be solved subject to the initial condition

$$K(t, y, y) = \Phi(t, y), \qquad 0 \le y \le x \le x_1. \tag{17}$$

Recalling Eqs. (5) and (8), we note that the Fredholm resolvent $K(t, y, x)$ satisfies the equations

$$K(t, y, x) = L(t, y) + \int_0^x L(t, z)K(z, y, x)dz, \tag{18}$$

$$K(t, y, x) = L(t, y) + \int_0^x K(t, z, x)L(z, y)dz. \tag{19}$$

3. CAUCHY SYSTEM FOR THE AUXILIARY FUNCTIONS

On making use of Eqs. (7) and (14), Eq. (12) reduces to

$$\Phi(t, x) = [\lambda(t)\lambda(x)^{1/2}k(t, x) + \int_0^x [\lambda(t)\lambda(z)]^{1/2}k(t, z) \times K(z, x, x)dz. \tag{20}$$

Introduce $B(t, x, v)$ as the solution of the integral equation

$$B(t, x, v) = e^{-(x-t)/v}\sqrt{\lambda(t)} + \int_0^x [\lambda(t)\lambda(z)]^{1/2} \times k(t, z)B(z, x, v)dz. \tag{21}$$

On making use of the Fredholm resolvent $K(t, y, x)$, Eq. (21) can be rewritten in the form

$$B(t,x,v)=e^{-(x-t)/v}\sqrt{\lambda(t)}+\int_0^x \sqrt{\lambda(z)}e^{-(x-z)/v}K(t, z, x)dz. \tag{22}$$

Differentiating Eq. (22) with respect to x, we get

$$B_x(t, x, v)= -\frac{1}{v}\left(e^{-(x-t)/v}\sqrt{\lambda(t)} + \int_0^x \sqrt{\lambda(z)}e^{-(x-z)/v} K(z, t, x)dz\right) + \sqrt{\lambda(x)}K(t, x, x) + \int_0^x \sqrt{\lambda(z)}e^{-(x-z)/v} K_x(z, t, x)dz. \tag{23}$$

By recalling the Bellman–Krein formula for the Fredholm resolvent (16), Eq. (23) reduces to

$$B_x(t, x, v) = -B(t, x, v)/v + \Phi(t, x)\left(\sqrt{\lambda(x)} + \int_0^x \sqrt{\lambda(z)}e^{-(x-z)/v}\Phi(z, x)dz\right) = -B(t, x, v)/v + \Phi(t, x)B(x, x, v), \tag{24}$$

where $\Phi(t, x)$ may be expressed as

$$\Phi(t, x) = \sqrt{\lambda(x)}\int_a^b B(t, x, v)dw(v). \tag{25}$$

In the theory of radiative transfer, the derivation of the integro-differential equation (24) from the auxiliary equation (21) relies upon the superposition principle (cf. Busbridge[11]). In this paper, however, the application of the Bellman–Krein formula permitted us to find Eq. (24) without using the theorem concerning the trivial solution of homogeneous Fredholm integral equation (cf. Bellman, Kalaba, and Ueno[10]).

Introduce a new function $R(v, u; x)$:

$$R(v, u; x) = \int_0^x e^{-(x-y)/v}\sqrt{\lambda(y)}B(y, x, u)dy. \tag{26}$$

Changing the order of integrations in Eq. (21) and putting $t = x$, we rewrite it in the form

$$B(x, x, v) = \sqrt{\lambda(x)}\left(1 + \int_0^x \sqrt{\lambda(z)} \times B(z, x, v)dz \int_a^b e^{-(x-z)/u}dw(u)\right) = \sqrt{\lambda(x)}\left(1 + \int_a^b R(u, v; x)dw(u)\right). \tag{27}$$

Differentiate Eq. (26) with respect to x. Then

$$R_x(v, u; x) = \sqrt{\lambda(x)}B(x, x, u) - \frac{R(v, u; x)}{v} + \int_0^x e^{-(x-y)/v}\sqrt{\lambda(y)}B_x(y, x, u)dy. \tag{28}$$

Recalling Eq. (24), after some manipulations, Eq. (28) becomes

$$R_x(v, u; x) = -\left(\frac{1}{v} + \frac{1}{u}\right)R(v, u; x) + \lambda(x)\left(1 + \int_a^b R(v', u; x)dw(v')\right) \times \left(1 + \int_a^b R(v, u'; x)dw(u')\right). \tag{29}$$

Equation (29) is a desired invariant imbedding equation governing the reflection function $R(v, u; x)$, together with the initial condition

$$R(v, u; 0) = 0. \tag{30}$$

Furthermore, it can be shown that R function is symmetric with respect to v and u,

$$R(v, u; x) = R(u, v; x), \tag{31}$$

because

$$\int_0^x B(t, x, v)B(t, x, u)dt = R(v, u; x) + \int_0^x B(t, x, u) \times B(t, x, v)dt - R(u, v; x). \tag{32}$$

The R function is similar to the reflection function of radiative transfer and neutron diffusion. Recall that $B(x, x, v)$ is expressible in terms of R vis Eq. (27). Then, Eq. (24) may be used to compute $B(t, x, v)$. Thus, for the computation of the resolvent kernel $K(t, y, x)$, we use the Bellman–Krein formula (16) in terms of $\Phi(t, x)$ given by Eq. (25), which is determined by $B(t, x, v)$.

Let us summarize the set of equations for the calculation of $K(t, y, x)$ with the aid of an initial-value method. The integro-differential equations are

$$R_x(v, u; x) = -\left(\frac{1}{v} + \frac{1}{u}\right)R(v, u; x) + \lambda(x)\left(1 + \int_a^b R(v, z; x)dw(z)\right) \times \left(1 + \int_a^b R(z, u; x)dw(z)\right), \tag{33}$$

$$B_x(t, x, v) = -\frac{B}{v}(t, x, v) + \lambda(x)\left(1 + \int_a^b R(v, z; x)dw(z)\right) \times \int_a^b B(t, x, z')dw(z'), \tag{34}$$

$$K_x(t,y,x) = \Phi(t,x)\Phi(y,x), \tag{35}$$

where $0 \le t, y \le x$, and

$$\Phi(t,x) = \sqrt{\lambda}(x)\int_a^b B(t,x,v)dw(v). \tag{36}$$

The initial conditions are

$$R(v,u;0) = 0, \tag{37}$$

$$B(t,x,v)|_{x=t} = \sqrt{\lambda}(t)\left(1 + \int_a^b R(v,u;t)dw(u)\right), \tag{38}$$

$$K(t,y,x)|_{x=y} = \Phi(t,y). \tag{39}$$

* Supported by the National Science Foundation under Grant No. GP29049 and the Atomic Energy Commission, Division of Research under Contract No. AT(40-3)-113, Project No. 19.

[1]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953).
[2]R. Bellman, Proc. Amer. Math. Soc. 8, 435 (1957).
[3]M. G. Krein, Dokl. Akad. Nauk SSSR 100, 413 (1955).
[4]V. V. Sobolev, *A Treatise on Radiative Transfer* (Van Nostrand, Princeton, N.J., 1963).
[5]S. Ueno, J. Math. Anal. Appl. 2, 361 (1961).
[6]R. Bellman, R. Kalaba, and S. Ueno, "Invariant Imbedding and a Resolvent of the Photon Diffusion Equation," The RAND Corporation, Memorandum, RM-3937-ARPA (1964).
[7]H. Kagiwada and R. Kalaba, Intern. J. Comp. Math. 2, 143 (1968).
[8]H. Kagiwada and R. Kalaba, J. Math. Phys. Sci. 1, 109 (1967).
[9]D. G. Hummer and G. Rybicki, "Computational Methods for Non-LTE Line Transfer Problems," in *Methods in Computational Physics*, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1967), pp. 53-127.
[10]R. Bellman, R. Kalaba, and S. Ueno, J. Math. Phys. 9, 906 (1968).
[11]Ida W. Busbridge, *The Mathematics of Radiative Transfer* (Cambridge U.P., London, 1960).

On the Lee model with dilatation analytic cutoff function*

R. A. Weder†

Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, Belgium‡
(Received 22 May 1973)

The spectrum of the Lee model Hamiltonian with Y interaction is studied; first of all we extend the work of Kato and Mugibayashi about the eigenvalues outside the essential spectrum and about the essential spectrum itself. Furthermore, it is proved that the singular-continuous spectrum is not present; properties of resonances and eigenvalues embedded in the continuous spectrum are obtained.

INTRODUCTION

In this work the spectral properties of the Hamiltonian of the Lee model with Y interaction[1] are studied. As in Refs. 2 and 3 the relativistic form of the free energy part is considered.

In Sec. I we define the model and prove that the Hamiltonian is a self-adjoint operator.

In Sec. II we prove (Theorem 1) that the spectrum of the Hamiltonian, in the (1, 1) sector, for a square-integrable cutoff function consists of

(a) the essential spectrum, denoted by $\sigma_e(H)$, given by $\sigma_e(H) = [M_0, \infty)$, where $M_0 = \min\{M, m + \mu\}$; M, m, μ are the bare masses of the particles.
(b) The set of isolated bound state energies, denoted by $\sigma_d^\gamma(H)$, is a bounded set of isolated finite-dimensional eigenvalues smaller than M_0, which is the only possible accumulation point.

Furthermore, it is proved (Theorem 2) that the set $\sigma_d^\gamma(H)$ is not empty for values of the coupling constant λ, larger than some constant λ_0; and also it is proved that if the Hilbert Schmidt norm of the integral equation kernel of the energy eigenvalue problem is uniformly bounded by the constant K_c for energies $E < M_0$,[4] then the set $\sigma_d^\gamma(H)$ is empty for all values of the coupling constant λ smaller than $1/\sqrt{K_c}$.

We remark that Theorem 1 and Theorem 2 extend the results of Ref. 2 on the eigenvalues outside the essential spectrum and about the essential spectrum itself. They describe quite completely the spectrum of the Hamiltonian. However there are two other questions that remain unanswered:

(a) what are the properties of eigenvalues embedded in the continuous spectrum and (b) is the singular-continuous spectrum present or not. In order to answer these questions we introduce in Sec. III "the method of dilatation invariance" developed in Refs. 5 and 6 in the mathematical theory of the N-body Schrödinger operators, and extended in Refs. (7) and (8) to relativistic quantum theory.

In Sec. III we extend the free Hamiltonian to a self-adjoint analytic family in the strip $S_{\pi/2} = \{z \in \mathbb{C} \mid |\mathrm{Im} z| < \pi/2\}$ of the complex plane $\mathbb{C}$, and we investigate its spectral properties. We define the class of dilatation analytic cutoff functions and we prove that for these cutoff functions the spectrum of the Hamiltonian consists of (see Theorem 3): (a) the continuous spectrum which is in fact absolutely continuous and given by the real interval $[M_0, \infty)$; (b) a bounded set of finite dimensional eigenvalues, different from M and $m + \mu$, and accumulating at most at M and $m + \mu$ which may be eigenvalues. Properties of resonances and bound state wavefunctions are investigated.

A last property of the discrete spectrum of the Hamiltonian is proved in Theorem 4.

Finally, for the definitions of vector-valued and operator valued analytic functions, and the classification of spectrum we refer to Ref. 9.

I. THE MODEL

The Lee model with Y interaction[1] describes the interaction of three particles, denoted, respectively, by V, N, and θ. The allowed processes are summarized by the transitions $V \rightleftarrows N + \theta$. The particles V and N are fermions; θ is a boson particle.

The Hilbert space of states denoted by $\mathcal{H}$ is the direct sum of the Fock spaces[10] of the three particles in the model.

The free Hamiltonian is given by the following operator:

$$H_0 = \int E_v(\bar{p})\, V^x(\bar{p}) V(\bar{p}) d^3p + \int E_N(q) N^x(\bar{q}) N(\bar{q})\, d^3q + \int E_\theta(k)\, \theta^x(\bar{k})\, \theta(\bar{k})\, d^3k.$$

The operators $V, V^x, N, N^x, \theta, \theta^x$ are annihilation and creation operators for the V, N, and θ particles, respectively; they satisfy the usual (anti-) commutation relations:

$$\{V(\bar{p}), V^x(\bar{p}')\} = \{N(\bar{p}), N^x(\bar{p}')\} = [\theta(\bar{p}), \theta^x(\bar{p}')] = \delta^3(\bar{p} - \bar{p}'),$$

the other (anti-) commutators are zero.

Following Refs. 2 and 3 we consider the relativistic expression for the energy functions

$$E_v(p) = (p^2 + M^2)^{1/2}, \quad E_N(q) = (q^2 + m^2)^{1/2}, \quad E_\theta(k) = (k^2 + \mu^2)^{1/2},$$

where M, m, and μ are, respectively, the masses of the V, N, and θ particles.

It can be proved[2] that the free Hamiltonian is a self-adjoint, positive, operator with a dense domain [denoted by $D(H_0)$].

The interaction Hamiltonian is given by

$$H_I(\rho) = \lambda \int \rho(\bar{p}, \bar{q}, \bar{k})\, d^3p\, d^3q\, d^3k \{V^x(\bar{p}) N(\bar{q}) \theta(\bar{k}) + N^x(\bar{q}) \theta^x(\bar{k}) V(\bar{p})\},$$

where $\rho(\bar{p}, \bar{q}, \bar{k})$ is a real,[11] square-integrable, cutoff function:

$$\int |\rho(\bar{p}, \bar{q}, \bar{k})|^2\, d^3p\, d^3q\, d^3k < \infty;$$

and λ is the coupling constant.

One of the characteristic properties of the Lee model, which makes it solvable, is the existence of two conserved quantities: the "baryonic" number $N_1 = N_v + N_N$

and the "charge" $Q = N_V + N_\theta$; here N_V, N_N, and N_θ are total number operators for N, V and θ particles, respectively:

$$N_V = \int V^+(\bar{p})\, V(\bar{p})\, d^3p,$$
$$N_N = \int N^x(\bar{q}) N(\bar{q})\, d^3q,$$
$$N_\theta = \int \theta^x(\bar{k})\, \theta(\bar{k})\, d^3k.$$

This property implies that $\mathfrak{K}$ decomposes in the following way:

$$\mathfrak{K} = \bigoplus_{n_1, n_2=0}^{\infty} \mathfrak{K}(n_1, n_2),$$

where $\mathfrak{K}_{(n_1, n_2)}$ is the subspace of $\mathfrak{K}$ corresponding to the eigenvalues n_1, and n_2 of the operators N_1 and N_2 (n_1, and n_2 are arbitrary natural numbers); and that the total Hamiltonian

$$H(\rho) = H_0 + H_I(\rho),$$

can be written in the following way:

$$H(\rho) = \bigoplus_{n_1, n_2=0}^{\infty} H(\rho) \mid \mathfrak{K}_{(n_1, n_2)},$$

where $H(\rho) \mid \mathfrak{K}_{(n_1, n_2)}$ is the restriction of $H(\rho)$ to $\mathfrak{K}_{(n_1, n_2)}$.

In each $\mathfrak{K}_{(n_1, n_2)}$ the particle number representation[10] has a finite number of components; thus $H_I(\rho) \mid \mathfrak{K}_{(n_1, n_2)}$ is a bounded operator[2]; then $H(\rho)$ is self-adjoint on the domain

$$D(H(\rho)) = \{\psi \in \mathfrak{K} : \psi_{(n_1, n_2)} \in D(H_0) \mid \sum_{n_1, n_2=0}^{\infty} \| H(\rho) \psi_{(n_1, n_2)} \|^2 < \infty\}.$$

In the following sections we will investigate the spectrum of the total Hamiltonian in the sector (1, 1) (which is the first nontrivial sector).

II. THE SPECTRUM OF THE TOTAL HAMILTONIAN FOR A SQUARE-INTEGRABLE CUTOFF FUNCTION

In the particle representation[10] a general element of $\mathfrak{K}_{(1,1)}$ is expressed as[12]

$$\psi = \{a(\bar{p}), b(\bar{q}, \bar{k})\},$$

and its norm is given by

$$\| \psi \|^2 = \int | a(\bar{p}) |^2 d^3p + \int | b(\bar{q}, \bar{k}) |^2 d^3q\, d^3k < \infty.$$

The first component correspond to the single V state, and the second to the $N + \theta$ state.

The spectral properties of the Hamiltonian in this sector are given in the following theorems.

Theorem 1: If the cutoff function, $\rho(\bar{p}, \bar{q}, \bar{k})$, is square-integrable the spectrum of the total Hamiltonian consists of: (a) The essential spectrum $\sigma_e(H(\rho)) = [M_0, \infty)$, where $M_0 = \min\{M, m + \mu\}$, (b) The set of isolated bound state energies ($\sigma_d^r(H(\rho))$: is a bounded set of isolated finite-dimensional eigenvalues smaller than M_0 (this is the only possible accumulation point).

Proof: It follows from the definition of $H_I(\rho)$ that

$$H_I(\rho)\psi = \{\int d^3q\, d^3k\, \rho(\bar{p}, \bar{q}, \bar{k})\, b(\bar{q}, \bar{k}),\ \int d^3p\, \rho(\bar{p}, \bar{q}, \bar{k})\, a(\bar{p})\},$$

where $\psi = \{a(\bar{p}), b(\bar{q}, \bar{k})\} \in \mathfrak{K}$.

Thus, $H_I(\rho)$ is a compact operator (Ref. 13, p. 176).

It is well known[14] that any compact perturbation leaves invariant the essential-spectrum of a self-adjoint operator; thus

$$\sigma_e(H(\rho)) = \sigma_e(H_0).$$

It is easy to see that H_0 has a purely continuous spectrum which consists of the half-line $[M_0, \infty)$.

The fact that $\sigma_d^r(H(\rho))$ is a bounded set follows from Ref. 9, p. 291.

Theorem 2: The set $\sigma_d^r(H(\rho))$, defined above, is not empty for every λ larger than some λ_0; and if there exists a positive constant, K_c, such that the following inequality is satisfied[4]

$$(K(E))^2 = \int d^3p\, d^3p' \left| \int \frac{d^3q\, d^3k \rho(\bar{p}, \bar{q}, \bar{k}) \rho(\bar{p}, \bar{q}, \bar{k})}{[E_v(p) - E]^{1/2} [E_N(q) + E_\theta(k) - E][E_v(p') - E]^{1/2}} \right|^2 < (K_c)^2 < \infty,$$

for all $E < M_0$,

the set $\sigma_d^r(H(\rho))$ is empty for every λ smaller than

$$\lambda_c = 1/\sqrt{K_c}.$$

Proof: It can be proved[2] that the eigenvalue equation $(H - E)\psi = 0$, for $E < M_0$, is equivalent to the following integral equation:

$$\phi(\bar{p}) = \lambda^2 \int K(\bar{p}, \bar{p}', E)\, \phi(\bar{p}')\, d^3p',$$

where $K(\bar{p}, \bar{p}', E)$

$$= \int \frac{d^3q\, d^3k\, \rho(\bar{p}, \bar{q}, \bar{k})\, \rho(\bar{p}', \bar{q}, \bar{k})}{[E_v(p) - E]^{1/2} [E_N(q) + E_\theta(k) - E][E_v(p') - E]^{1/2}},$$

$$\phi(\mathbf{p}) = [E_v(p) - E]^{1/2}\, a(\bar{p}).$$

The function $\phi(\bar{p})$ is square integrable because ψ belongs to $D(H_0)$.

The kernel $K(\bar{p}, \bar{p}', E)$ is square integrable, thus, the equation is of Fredholm type.[15]

As the kernel is symmetric the integral equation has solutions, at least for one E, for every λ larger than some λ_0 (see Ref. 2 and Ref. 15, p. 73), and this proves the first part of the theorem. It is clear that (if the constant K_c exists)

$$\| \lambda^2 \int K(\bar{p}, \bar{p}', E)\, \phi(\bar{p}')\, d^3p' \| \leq \lambda^2 K_c \| \phi \|,$$

then, if $\lambda < 1/\sqrt{K_c}$ the integral equation has no solution for $E < M_0$ (see Ref. 13, p. 151). QED

These two theorems describe quite completely the spectrum of the total Hamiltonian; but there are two questions that remain unanswered:

(a) which are the properties of eigenvalues embedded in the continuous spectrum, and (b) is the singular-continuous spectrum, empty or not.

To give an answer to these questions we will intro-

duce the method of dilatation invariance developed by Refs. 5 and 6 in the mathematical theory of N-body Schrödinger operators, and extended to relativistic quantum theory by Refs. 7 and 8.

III. THE SPECTRUM OF THE TOTAL HAMILTONIAN FOR A DILATATION ANALYTIC CUTOFF FUNCTION

In this section we will study the properties of eigenvalues embedded in the continuous spectrum, resonances, and we will prove that the singular-continuous spectrum is empty, for a class of dilatation analytic cutoff function.

Let $U(\gamma), \gamma \in \mathbf{R}$, be the strongly continuous unitary representation on $\mathfrak{K}$ of the dilatation group defined by

$$\{U(\gamma)\psi\} = \{e^{-3\gamma/2}\, a(e^{-\gamma}\,\bar{p}), e^{-3\gamma}\, b(e^{-\gamma}\bar{q}, e^{-\gamma}\,\bar{k})\}, \gamma \in \mathbf{R},$$

where

$$\psi = \{a(\bar{p}), b(\bar{q}, \bar{k})\} \in \mathfrak{K}.$$

Then, we have the following lemma:

Lemma 1: The family of operators $H_0(\gamma)$ defined by

$$H_0(\gamma) = U(\gamma) H_0 U(-\gamma), \gamma \in \mathbf{R},$$

extends to a self-adjoint, analytic family in the strip of the complex-plane $S_{\pi/2} = \{\gamma \in \mathbf{C} \parallel \mathrm{Im}\gamma | < \pi/2\}$, with domain $[D(H_0(\gamma))]$ equal to the domain of H_0.

Proof: It follows from the definition of $H_0(\gamma)$ that

$$\{H_0(\gamma)\psi\} = \{E_v(\gamma, p)\, a(\bar{p}), (E_N(\gamma, q) + E_\theta(\gamma, k))\, b(\bar{q}, \bar{k})\}, \gamma \in \mathbf{R},$$

where $\psi = \{a(\bar{p}), b(\bar{q}, \bar{k})\} \in D(H_0(\gamma))$, and

$$E_v(\gamma, p) = \sqrt{e^{-2\gamma} p^2 + M^2},$$
$$E_N(\gamma, q) = \sqrt{e^{-2\gamma} q^2 + m^2},$$

and

$$E_\theta(\gamma, k) = \sqrt{e^{-2\gamma} k^2 + \mu^2}.$$

Thus, $H_0(\gamma), \gamma \in \mathbf{R}$, extends to a family of operators for $\gamma \in S_{\pi/2}$.

It can be proved that $E_v(\gamma, p)$, and $[E_N(\gamma, q) + E_\theta(\gamma, k)]$ are differents from zero for all $p, q, k \in [0, \infty)$ if $\gamma \in S_{\pi/2}$; thus, there exists four functions of γ $[M_1(\gamma), M_2(\gamma), M_3(\gamma), M_4(\gamma)]$ such that

$$0 < \left| \frac{E_v(0,p)}{E_v(\gamma,p)} \right| < M_1(\gamma) < \infty,$$

$$0 < \left| \frac{E_v(\gamma,p)}{E_v(0,p)} \right| < M_2(\gamma) < \infty,$$

$$0 < \left| \frac{E_N(0,q) + E_\theta(0,k)}{E_N(\gamma,q) + E_\theta(\gamma,k)} \right| < M_3(\gamma) < \infty,$$

and

$$0 < \left| \frac{E_N(\gamma,q) + E_\theta(\gamma,k)}{E_N(0,q) + E_\theta(0,k)} \right| < M_4(\gamma) < \infty,$$

$$p, q, k \in [0, \infty), \text{ and } \gamma \in S_{\pi/2}.$$

This implies that

$$\| H_0\psi \| \leq (\max\{M_1(\gamma), M_3(\gamma)\}) \| H_0(\gamma)\psi \| \; H_0(\gamma)\psi \|, \quad \psi \in D(H_0(\gamma)),$$

$$\| H_0(\gamma)\psi \| \leq (\max\{M_2(\gamma), M_4(\gamma)\}) \| H_0\psi \|, \quad \psi \in D(H_0).$$

Then $D(H_0(\gamma)) = D(H_0), \gamma \in S_{\pi/2}$, i.e., the domain of $H_0(\gamma)$ is independent of γ for $\gamma \in S_{\pi/2}$.

Previously[7] we proved that

$$\left| \frac{E_v(\gamma_2, p) - E_v(\gamma_1, p)}{\gamma_2 - \gamma_1} \right| \leq 2 \left(\frac{\epsilon + e^{-2R_e\gamma_1}}{\cos \mathrm{Im}\gamma_1} \right) M_1(\gamma_1) | E_v(0, \bar{p}) |,$$

where $\gamma_1, \gamma_2 \in S_{\pi/2}$, $| \gamma_2 - \gamma_1 | < \eta(\gamma_1), \epsilon > 0$, and $\eta(\gamma_1) > 0$.

This implies (together with a similar estimation for the N and θ particles), by the Lebesgue's dominated convergence theorem, that

$$\left\{ \frac{H_0(\gamma_2) - H_0(\gamma_1)}{\gamma_2 - \gamma_1} \psi \right\} \xrightarrow[\gamma_2 \to \gamma_1]{} \left\{ \left(\frac{d}{d\gamma} E_v(\gamma, p) \right)_{\gamma = \gamma_1} a(\bar{p}), \left(\frac{d}{d\gamma} (E_N(\gamma, q) + E_\theta(\gamma, k)) \right)_{\gamma = \gamma_1} b(\bar{q}, \bar{k}) \right\},$$

in the strong topology in $\mathfrak{K}$ for any ψ in the domain of H_0.

Thus, $H_0(\gamma), \gamma \in S_{\pi/2}$, is an analytic family.

Since $H_0(\gamma)$ is self-adjoint for $\gamma \in \mathbf{R}$ we have that

$$H_0(\bar{\gamma}) = H_0^*(\gamma) \quad \text{for all } \gamma \in S_{\pi/2}. \qquad \text{QED}$$

It is clear that $H_0(\gamma), \gamma \in S_{\pi/2}$, can be written as follows:

$$H_0(\gamma) = H_{0,v}(\gamma) \oplus (H_{0,N}(\gamma) \otimes 1 + 1 \otimes H_{0,\theta}(\gamma))^{\sim};$$

where $H_{0,v}(\gamma)$ is defined as

$$(H_{0,v}(\gamma)\Phi)(\bar{p}) = (e^{-2\gamma} p^2 + M^2)^{1/2} \Phi(\bar{p}), \quad \gamma \in S_{\pi/2},$$

in the domain of all Φ in $\mathfrak{L}^2(\mathbf{R}^3)$ such that $(\sqrt{e^{-2\gamma} p^2 + M^2}\, \Phi(\bar{p}))$ is again in $\mathfrak{L}^2(\mathbf{R}^3)$ (and a similar definition for $H_{0,N}(\gamma)$, and $H_{0,\theta}(\gamma)$). The operator $(H_{0,N}(\gamma) \otimes 1 + 1 \otimes H_{0,\theta}(\gamma))^{\sim}$ is the closure of the operator $(H_{0,N}(\gamma) \otimes 1 + 1 \otimes H_{0,\theta}(\gamma))$.

Then, we have the following lemma.

Lemma 2: The spectrum of $H_0(\gamma)$ [denoted by $\sigma(H_0(\gamma))$ for $\gamma \in S_{\pi/2}$] is given by

$$\sigma(H_0(\gamma)) = \sigma(H_{0,v}(\gamma)) \, U(\sigma(H_{0,N}(\gamma)) + \sigma(H_{0,\theta}(\gamma)).$$

Proof: It follows from the definition of the spectrum and the expression of $H_0(\gamma)$ given above that

$$\sigma(H_0(\gamma)) = \sigma(H_{0,v}(\gamma)) \, U(\sigma(H_{0,N}(\gamma) \otimes 1 + 1 \otimes H_{0,\theta}(\gamma))^{\sim}).$$

In a previous paper[7] we proved that

$$\sigma(H_{0,N}(\gamma)) \subset \{Z \in \mathbf{C} \parallel \arg z \mid \leq | \mathrm{Im}\gamma |\},$$
$$\sigma(H_{0,\theta}(\gamma)) \subset \{Z \in \mathbf{C} \parallel \arg z \mid \leq | \mathrm{Im}\gamma |\}, \gamma \in S_{\pi/2}.$$

The operators $H_{0,N}(\gamma)$ and $H_{0,\theta}(\gamma)$ are normal, thus[9]

$$\| (H_{0,N}(\gamma) - Z)^{-1} \| = \{\mathrm{dist}[Z, \sigma(H_{0,N}(\gamma))]\}^{-1}, Z \in \mathbf{C} \setminus \sigma(H_{0,N}(\gamma)),$$

$$\|(H_{0,\theta}(\gamma) - Z)^{-1}\| = \{\mathrm{dist}[Z, \sigma(H_{0,\theta}(\gamma))]\}^{-1},\ Z \in \mathbb{C} \setminus \sigma(H_{0,\theta}(\gamma));$$

but[7]

$$\mathrm{dist}(Z, \sigma(H_{0,N}(\gamma))) \geq |Z| \sin(|\arg z| - |\mathrm{Im}\gamma|),$$
$$\mathrm{dist}(Z, \sigma(H_{0,\theta}(\gamma))) \geq |Z| \sin(|\arg z| - |\mathrm{Im}\gamma|),$$

for $|\arg z| > |\mathrm{Im}\gamma|$.

Then, by a lemma of Ichinose,[16] we have that

$$\sigma(H_{0,N}(\gamma) \otimes 1 + 1 \otimes H_{0,\theta}(\gamma))^{\sim} = \sigma(H_{0,N}(\gamma)) + \sigma(H_{0,\theta}(\gamma)).$$

QED

Remark 1: It follows from Lemma 2 that if the mass of the θ particle is equal to zero the spectrum of $H_0(\gamma)$ is equal to

$$\sigma(H_0(\gamma)) = \{Z \in \mathbb{C} \mid Z = (e^{-2\gamma}p^2 + M^2)^{1/2}, \text{or } Z = (e^{-2\gamma}q^2 + m^2)^{1/2} + e^{-\gamma}k,$$

for some $p, q, k \in [0, \infty)\}$,

and it has the simple shape shown in Fig. 1.

Remark 2: In the original Lee model[1] the recoil of the V and N particles is neglected. In that case the free Hamiltonian is equal to

$$\{H_0\psi\} = \{Ma(\bar{p}), [m + (k^2 + \mu^2)^{1/2}]b(\bar{q}, \bar{k})\},$$
$$\psi = \{a(\bar{p}), b(\bar{q}, \bar{k})\},$$

in the domain, $D(H_0)$, of all ψ in $\mathfrak{K}$ such that

$$\{Ma(\bar{p}); [m + (k^2 + \mu^2)^{1/2}]b(\bar{q}, \bar{k})\} \text{ is again in } \mathfrak{K}.$$

Clearly this Hamiltonian extends to a self-adjoint, analytic family in $S_{\pi/2}$, and its spectrum is equal to

$$\sigma(H_0(\gamma)) = \{M\ U\ m + \sigma(H_{0,\theta}(\gamma))\}, \quad \gamma \in S_{\pi/2},$$

and has the shape shown in Fig. 2.

We will define now a class of cutoff functions which allows us to extend the total Hamiltonian to a self-adjoint analytic family in a strip $S_b, 0 < b < \pi/2$, of the complex plane.

Definition: A dilatation analytic cutoff function is a real square-integrable function $\rho(\bar{p}, \bar{q}, \bar{k})$, having the following property: the family of real, square-integrable functions

$$(U(\gamma)\rho)(\bar{p}, \bar{q}, \bar{k}) = e^{-q\gamma/2}\, \rho(e^{-\gamma}\bar{p}, e^{-\gamma}\bar{q}, e^{-\gamma}\bar{k}), \gamma \in \mathbb{R},$$

has an extension to an analytic family[9] of square-integrable functions [they will be denoted by $\rho(\gamma, \bar{p}, \bar{q}, \bar{k})$][17] for γ in an open connected domain of the complex plane.

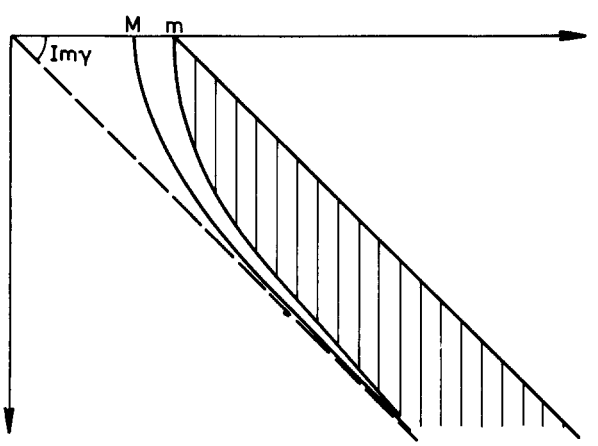

FIG. 1.

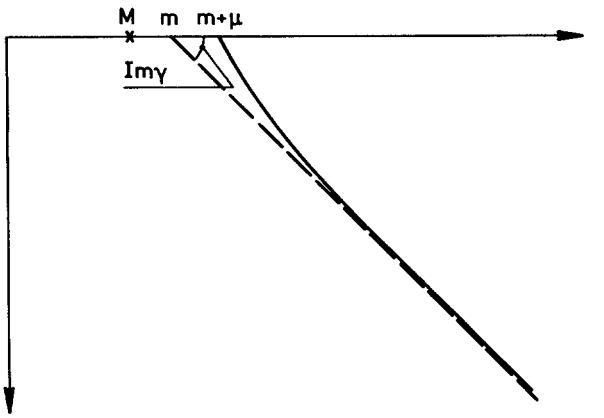

FIG. 2.

Lemma 3: If the cutoff function $\rho(\bar{p}, \bar{q}, \bar{k})$ is dilatation analytic in a strip $S_a, a > 0$, the family of operators $U(\gamma)(H_0 + H_I(\gamma))U(-\gamma), \gamma \in \mathbb{R}$, extends to a self-adjoint, analytic family for $\gamma \in S_b$, where $b = \min\{a, \pi/2\}$, with domain $D(H(\gamma))$ equal to the domain of H_0.

Proof: It follows from the definition of $H_I(\rho)$, and of

$U(\gamma), \gamma \in \mathbb{R}$, that

$$U(\gamma)H_I(\rho)U(-\gamma) = H_I(U(\gamma)\rho), \gamma \in \mathbb{R}.$$

Clearly the family

$$H_I(\rho, \gamma) = H_I(\rho(\gamma)), \gamma \in S_a,$$

is an extention to S_a of the family $U(\gamma)H_I(\rho)U(-\gamma)$, $\gamma \in \mathbb{R}$.

The argument given in Theorem 1 implies that $H_I(\rho, \gamma)$, $\gamma \in S_a$, is a family of compact operator. It follows from the following estimation:

$$\left\| \left[\frac{H_I(\rho, \gamma_2) - H_I(\rho, \gamma_1)}{\gamma_2 - \gamma_1} - H_I\left(\left(\frac{d}{d\gamma}\rho(\gamma)\right)\gamma = \gamma_1\right)\right]\psi \right\| \leq \left\| \frac{\rho(\gamma_2) - \rho(\gamma_1)}{\gamma_2 - \gamma_1} - \left(\frac{d}{d\gamma}\rho(\gamma)\right)_{\gamma=\gamma_1} \right\| \|\psi\|,$$

valid for $\gamma_1, \gamma_2 \in S_a$, and $\psi \in \mathfrak{K}$ that the family is analytic. This implies (together with Lemma 1)[9] that the family $U(\gamma)[H_0 + H_I(\rho)]U(-\gamma)$ extends to a self-adjoint analytic family in the strip S_b, where $b = \min\{a, \pi/2\}$, with $D(H(\rho, \gamma)) = D(H_0)$. QED

Lemma 4: If the cutoff function $\rho(\bar{p}, \bar{q}, \bar{k})$ is dilatation analytic in a strip $S_a, a > 0$, the spectrum of the total Hamiltonian, $H(\rho, \gamma)$, for $0 < |\mathrm{Im}\gamma| < b$, where $b = \min\{a, \pi/2\}$, consists of[18]

(a) The essential spectrum: $\sigma_e(H(\rho, \gamma)) = \sigma(H_0(\gamma))$.
(b) The set of bound state energies is a bounded set of isolated, finite-dimensional real eigenvalues, independent of γ, with M and $m + \mu$ as the only possible accu-

mulation points.
(c) The set of nonreal resonance energies is a set of isolated finite dimensional eigenvalues. A given resonance energy is independent of γ as long as it is not absorbed in the essential spectrum.

Proof: We have proved in Lemma 3, that the operators $H_I(\rho,\gamma)$ are compact (thus, bounded); this implies (together with the fact that the $H_0(\gamma)$ are normal operators) that

$$\| H_I(\rho,\gamma)(H_0(\gamma) - Z)^{-1} \| \leq \| H_I(\rho,\gamma) \| [\mathrm{dist}(Z, \sigma(H_0(\gamma)))],$$
$$Z \in \mathbb{C} \setminus \sigma(H_0(\gamma)).$$

The remainder of the proof goes along the same lines as in Refs. 5–8; we omit the proof here. QED

Theorem 3: If the cutoff function $\rho(\bar{p},\bar{q},\bar{k})$ is dilatation analytic in a strip $S_a, a > 0$, the spectrum of the total Hamiltonian $H(\rho)$ consists of:

(a) The continuous spectrum $[M_0,\infty)$. This spectrum is absolutely continuous, i.e., $\mathcal{K}_{s.c.} = \phi$, and $\mathcal{K} = \mathcal{K}_{a.c.} \oplus \mathcal{K}_p$.

(b) A bounded set of finite dimensional eigenvalues, λ, different from M, and $m + \mu$ [which are precisely the real eigenvalues, different from M and $m + \mu$, of $H(\rho,\gamma), 0 < |\mathrm{Im}\gamma| < b$, where $b = \min\{a, \pi/2\}$] and accumulating, at most, at M, and $m + \mu$, which are possibly eigenvalues.

The projection operators $P(\gamma,\lambda), \gamma \in S_b$, on the eigenspace of $H(\rho,\gamma)$ corresponding to a fixed real eigenvalue, λ, form a self-adjoint analytic family in S_b. The eigenvectors Φ, of $H(\rho)$, corresponding to such eigenvalues are in the dense set, D_b, of analytic vectors[19] in S_b, and the analytic extention $\Phi(\gamma)$ of Φ are eigenvectors of $H(\rho,\gamma)$ corresponding to the same eigenvalue.

Proof: This theorem follows, by standard arguments, from Lemmas 1, 2, 3, 4 and from Lemma 4.4, and Theorems 4.5 and 4.6 of Ref. 20 (see also Refs. 5–8). QED

Theorem 4: If the cutoff function $\rho(\bar{p},\bar{q},\bar{k})$ is dilatation analytic in a strip $S_a; a > 0$, and if there exists a positive constant K_c such that

$$[K(\gamma_0,E)]^2 = \int d^3p\, d^3p' \left| \int \frac{d^3p\, d^3k\, \rho(\gamma_0,\bar{p},\bar{q},\bar{k})\rho(\gamma_0,\bar{p}',\bar{q},\bar{k})}{[E_v(\gamma_0,p) - E]^{1/2}[E_N(\gamma_0,q) + E_\theta(\gamma_0,k) - E][E_v(\gamma_0,p') - E]^{1/2}} \right|^2 < (K_c)2 < \infty,$$

for all $E \in \mathbb{R} \setminus \{M, m + \mu\}$, and some γ_0, such that, $0 < |\mathrm{Im}\gamma_0| < b$, where $b = \min\{a, \pi/2\}$, then, the total Hamiltonian $H(\rho)$ can have eigenvalues only at the points M and $m + \mu$, for every λ smaller than $\lambda_c = 1/\sqrt{K_c}$.

Proof: We can prove by the argument developed in Theorem 2 that (if the constant K_c exists) the total Hamiltonian $H(\rho,\gamma_0)$ has no eigenvalues different from M and $m + \mu$ for every λ smaller than $\lambda_c = 1\sqrt{K_c}$; but this implies, by Theorem 4, that the same result is valid for $H(\rho)$. QED

ACKNOWLEDGMENT

I am very indebted to Professor A. Verbeure, for interesting discussions and suggestions, and for reading the manuscript.

* Work supported by A.G.C.D. (Belgium).
† On leave of absence from Universidad de Buenos Aires (Argentina).
‡ Postal address: Instituut voor Theoretische Fysica, Celestijnenlaan 200 D, B-3030 Heverlee, Belgium.

[1]T. D. Lee, Phys. Rev. 95, 1329 (1954).
[2]Y. Kato and N. Mugibayashi, Progr. Theor. Phys. (Kyoto) 30, 1, 103 (1963).
[3]F. Y. Yndurain, J. Math. Phys. 7, 1133 (1966); 9, 1423 (1968).
[4]This condition has been proved to be unnecessary in Ref. 2 (Lemma 6.1). However their proof seems unclear.
[5]J. Aguilar and J. M. Combes, Comm. Math. Phys. 22, 269 (1971).
[6]E. Balslev and J. M. Combes, Comm. Math. Phys. 22, 280 (1971).
[7]R. A. Weder, "Spectra Properties of One-Body Relativistic Spin-Zero Hamiltonians", Preprint UCL (February 1973).
[8]R. A. Weder, "Spectral Properties of the Dirac Hamiltonian", Preprint (KUL) (April 1973).
[9]T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[10]J. M. Cook, Trans. Am. Math. Soc. 74, 22 (1953).
[11]The restriction to a real cutoff function is not essential. It can be eliminated by considering the interaction $\int\{\rho(\bar{p},\bar{q},\bar{k})V^x(\bar{p})N(\bar{q})\theta(\bar{k}) + \rho^x(\bar{p},\bar{q},\bar{k})N^x(\bar{q})\theta^x(\bar{k})V(\bar{p})\}d^3p\,d^3q\,d^3k$, where $\int|\rho(\bar{p},\bar{q},\bar{k})|^2\,d^3p\,d^3q\,d^3k < \infty$.
[12]In what follows for simplicity, we write, $\mathcal{K}$, $H(\rho)$, H_0, and $H_I(\rho)$ instead of $\mathcal{K}_{(1,1)}$, $H(\rho)|_{\mathcal{K}_{(1,1)}}$, $H_0|_{\mathcal{K}_{(1,1)}}$, and $H_I(\rho)|_{\mathcal{K}_{(1,1)}}$.
[13]F. Riesz and B. Sz. Nagy, *Leçons d'analyse fonctionnelle* (Hungarian Acadêmy of Sciences, Budapest, 1955).
[14]F. Wolff; Comm. Pure Appl. Math. 12, 211 (1959).
[15]S. G. Mikhlin, *Integral Equations* (Pergamon, London, 1957).
[16]T. Ichinose, "On Operators on Tensor Products of Banach spaces", preprint, Nagoya University.
[17]It is clear that the analyticity domain of $\rho(\gamma,\bar{p},\bar{q},\bar{k})$ can always be extended to a complex strip $S_a = \{\gamma \in \mathbb{C} \mid |\mathrm{Im}\gamma| < a, a > 0\}$.
[18]The spectrum of $H(\gamma)$ is independent of $Re\gamma$ because $H(\gamma_1)$ and $H(\gamma_2)$ are unitary equivalents for $\mathrm{Im}\gamma_1 = \mathrm{Im}\gamma_2$.
[19]E. Nelson, Ann. Math. 70, 3 (1959).
[20]E. Balslev, "Spectral Theory of Schrödinger Operators of Many-Body Systems" Lecture Notes, Leuven (1971).

Unitary representations of $SO(n, 1)$

M. K. F. Wong

Fairfield University, Fairfield, Connecticut 06430
(Received 7 May 1973)

All unitary representations of $SO(n, 1)$ have been obtained in the group chain $SO(n, 1) \supset SO(n) \supset SO(n-1) \supset \cdots \supset SO(2)$. The branching laws have been explicitly formulated. These results follow from the observation that the matrix elements of the "noncompact" generators of $SO(n, 1)$ differ from the corresponding matrix elements of the same generators of $SO(n+1)$ by a factor of $\sqrt{-1}$. The branching laws then follow from the unitarity condition. It is also observed that the invariants of $SO(n, 1)$ have the same eigenvalues as the invariants of $SO(n+1)$. Finally we show that the normalized raising and lowering operators in $SO(n+1)$, obtained by Pang and Hecht in graphs, and by Wong in algebraic form, can be similarly defined and applied to $SO(n, 1)$.

INTRODUCTION

The noncompact groups $SO(n, 1)$ have had important applications in physics. In the simplest case, $SO(2, 1)$ is the little group, for a spacelike momentum, of the quantum mechanical Poincaré group. It is the group upon which the theory of complex angular momentum is built. It has been investigated exhaustively by Bargmann.[1] Next comes the Lorentz group $SO(3, 1)$ whose unitary representations were studied by Gel'fand and Naimark.[2] There is no doubt that the Lorentz group has had a vast influence on relativistic physics. The next group, $SO(4, 1)$, is the De Sitter group, whose unitary representations were investigated by Thomas[3] in as early as 1941. In elementary particle physics, the group $SO(4, 1)$ has been applied to the hydrogen atom by Barut *et al.*[4] The group $SO(5, 1)$ has been investigated by Kuriyan *et al.*[5]

In order to understand fully the group $SO(n, 1)$ it is useful to consider properties which are common to the group as a whole, i.e., for any n. Fortunately, these properties do exist. The representations of $SO(n, 1)$ have been considered by Hirai, Ottoson, Chakrabarti, and Schwarz.[6] Our results agree with theirs, but are presented in a different, and, we believe more complete, way. The discrete and degenerate representations of $O(p, q)$ have been considered by Nikolov.[7]

In Sec. 1 we mention briefly the general properties of $SO(n, 1)$. In Sec. 2 we present the general formulas for the matrix elements of the generators of the group in the decomposition $SO(n, 1) \supset SO(n) \supset \cdots \supset SO(2)$. In Sec. 3 we classify the unitary representations of the Lie algebra $SO(n, 1)$, obtaining the so-called branching laws. One of the results of the classification is that there are no "discrete series" for $SO(2k-1, 1)$. Another result is that in the Lie algebra of the Lorentz group $SO(3, 1)$ there is a new (continuous) representation which has not been considered by Gel'fand and Naimark. Our results therefore include all the previous results obtained in the particular cases of $SO(2, 1)$, $SO(3, 1)$, $SO(4, 1)$, and $SO(5, 1)$, but are more complete and are valid for general n. In Sec. 4 we show that the invariants of $SO(n, 1)$ have the same eigenvalues as the invariants of $SO(n+1)$.

Recently Patera[8] has shown that the raising and lowering operators in $U(p+1)$ obtained by Nagel and Moshinsky[9] can be similarly applied to the noncompact group $U(p, 1)$. We show in Sec. 5 that this is also true for $SO(n, 1)$, i.e., the normalized raising and lowering operators in $SO(n+1)$ obtained by Pang and Hecht[10] and by Wong[11] can be similarly applied to $SO(n, 1)$.

I. GENERAL PROPERTIES OF $SO(n,1)$

$SO(n, 1)$ is the group of unimodular transformations in an $(n+1)$-dimensional real space which leave the following form invariant:

$$x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 - x_{n+1}^2$$

The generators of the Lie algebra of $SO(n, 1)$ are represented by antisymmetric J_{ij}, $i, j = i, 2, 3, \ldots, n+1$. $J_{ij} = -J_{ji}$. They satisfy the following commutation relations:

$$[J_{ab}, J_{cd}] = i(g_{ac}J_{bd} - g_{ad}J_{bc} + g_{bd}J_{ac} - g_{bc}J_{ad}), \tag{1}$$

where $g_{11} = g_{22} = \cdots = g_{nn} = -g_{n+1,n+1} = 1$, $g_{ij} = 0$ for $i \neq j$.

If one denotes the generators of $SO(n+1)$ by L_{ij}, then one finds that

$$J_{ij} = L_{ij} \quad \text{for } i, j < n+1, \tag{2}$$

$$J_{n+1,j} = iL_{n+1,j}. \tag{3}$$

It is easy to see that the invariants of $SO(n, 1)$ are[12]:

Racah invariants: $$I_{2k}(R) = L_{i_1 i_2} L_{i_2 i_3} L_{i_3 i_4} \cdots L_{i_k i_1} \tag{4}$$

Louck invariants: $$I_{2k}(L) = \left(\sum_{i_s=1}^{n} (-1)^P L_{i_1 i_2} L_{i_3 i_4} \cdots L_{i_{k-1} i_k} \right)^2 \tag{5}$$

with substitutions from Eqs. (2) and (3) for J_{ij}. For example, in $SO(2, 1)$, the Casimir invariant is

$$I_2 = J_{12}^2 - J_{13}^2 - J_{23}^2 \tag{6}$$

which can be obtained from the Casimir invariant of $SO(3)$,

$$I_2 = L_{12}^2 + L_{13}^2 + L_{23}^2, \tag{7}$$

by interchanging J_{12} with L_{12}, J_{13} with iL_{13}, and J_{23} with iL_{23}.

The rank of $SO(2k, 1)$ is k; the rank of $SO(2k-1, 1)$ is also k. In general, a representation of a group of rank r is labeled by r numbers, corresponding to the r independent invariants of the group.

The unitary representations of $SO(2, 1)$, $SO(3, 1)$, $SO(4, 1)$, and $SO(5, 1)$ have been studied by Bargmann,[1] Gel'fand and Naimark,[2] Thomas,[3] and Kuriyan *et al.*,[5] respectively. They are, of course, all infinite-dimensional. Here we wish to study the group $SO(n, 1)$ as a whole. We start our investigation by determining the matrix elements of the inifinitesimal generators of the group $SO(n, 1)$.

II. MATRIX ELEMENTS OF THE GENERATORS OF $SO(n,1)$

We shall state our results in the form of a theorem.

Theorem 1: The matrix elements of the noncompact generators of $SO(n,1)$, i.e., $J_{n+1,i}$, are equal to $\sqrt{-1}$ times the matrix elements of the corresponding generators of $SO(n+1)$ in the decomposition $SO(n,1) \supset SO(n) \supset SO(n-1) \supset \cdots \supset SO(2)$.

Proof: The noncompact generators of $SO(n,1)$, i.e., $J_{n+1,i}$, where $i = 1, 2, \ldots, n$, can be classified as the n components of a tensor operator in $SO(n)$, satisfying the definition

$$[J_{ij}, T_a] = \sum_b \langle b \mid J_{ij} \mid a \rangle T_b , \tag{8}$$

where $i,j = i, 2, \ldots, n$. T_a, T_b are tensor operators labeled by the Gel'fand states of $SO(n)$, and are linear combinations of $J_{n+1,i}$. $\langle b \mid J_{ij} \mid a \rangle$ is the Gel'fand Zetlin matrix elements of the generators of $SO(n)$.

These tensor operators transform in exactly the same way as the generators of $SO(n+1)$. For example, if in $SO(5)$, $\frac{1}{2}(L_{52} - iL_{15})$ transforms as the tensor operator $T^{1/2\ 1/2}_{-1/2-1/2}$ in $SO(4) = SO(3) \otimes SO(3)$, then in $SO(4,1)$, $\frac{1}{2}(J_{52} - iJ_{15})$ also transforms as the tensor operator $T^{1/2\ 1/2}_{-1/2-1/2}$ in $SO(4)$. This is because in Eq. (8) the commutation relations between J_{ij} and T_a are not affected by the metric $g_{n+1,n+1} = -1$ in Eq. (1).

Using the Wigner–Eckart theorem, one can then write down the matrix elements of these generators as the reduced matrix element times the Clebsch–Gordan coefficients in $SO(n)$:

$$\left\langle \begin{matrix} c \\ \gamma \end{matrix} \middle| \begin{matrix} T_a \\ \alpha \end{matrix} \middle| \begin{matrix} b \\ \beta \end{matrix} \right\rangle = \langle c \| T_a \| b \rangle \begin{pmatrix} b & a & c \\ \beta & \alpha & \gamma \end{pmatrix}, \tag{9}$$

where a, b, c are irreducible representation labels for $SO(n)$ and α, β, γ are state labels for $SO(n-1) \supset SO(n-2) \supset \cdots \supset SO(2)$. The last term on the right-hand side of Eq. (9) is the Clebsch–Gordan coefficient of $SO(n)$. The reduced matrix elements, which must be functions of r additional variables, where r is the rank of the group, are to be solved by the equations from (1):

$$[T_a, T_b] = g_{n+1,n+1}\, k\, J_{ij} = -k\, J_{ij} , \tag{10}$$

where k is a constant, $i,j = 1, 2, \ldots, n$.

In almost all the previous investigations of the $SO(n,1)$ groups, i.e., in $SO(2,1)$, $SO(3,1)$, $SO(4,1)$, by Bargmann, Gel'fand and Naimark, and Thomas, and in $SO(n,1)$ by Ottoson, much effort has been spent on solving Eq. (10). However, we wish to point out that it is not necessary to solve (10), as it were, from the beginning. The solution is known. All one has to do is to compare it with the corresponding commutation relations in $SO(n+1)$. One then finds that in that case

$$[T_a, T_b] = +k\, J_{ij}. \tag{11}$$

Therefore, if one substitutes the reduced matrix elements of the generators in the compact group by the same value multiplied by $\sqrt{-1}$ for the reduced matrix elements of the noncompact group, one can satisfy Eq. (10) in all cases. This proves our theorem, since the Clebsch–Gordan coefficients of $SO(n)$ cannot be different in $SO(n,1) \supset SO(n)$, and in $SO(n+1) \supset SO(n)$.

It is easy to check that the theorem is true for $SO(2,1)$, $SO(3,1)$, $SO(4,1)$, and $SO(5,1)$, since the matrix elements of these groups are all known. We shall therefore write down the matrix elements of the generators $J_{2k,2k-1}$ for $SO(2k-1,1)$ and $J_{2k+1,2k}$ for $SO(2k,1)$. Matrix elements of the other generators can be obtained from these by repeated application of the commutation relations in (1).

With

$$l_{2k,a} = m_{2k,a} + k - a, \quad l_{2k-1,a} = m_{2k-1,a} + k - a,$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} + 1 \\ m_{2k-2,j} \end{matrix} \middle| J_{2k,2k-1} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \right\rangle \equiv A_{2k,2k-1}$$

$$= -i \left| \frac{-\prod_{a=1}^{k-1}(l^2_{2k-2,a} - l^2_{2k-1,j}) \prod_{b=1}^{k}(l^2_{2k,b} - l^2_{2k-1,j})}{l^2_{2k-1,j}(4l^2_{2k-1,j} - 1) \prod_{a\neq j}^{k-1}(l^2_{2k-1,a} - l^2_{2k-1,j})[(l_{2k-1,a} - 1)^2 - l^2_{2k-1,j}]} \right|^{1/2}, \tag{12}$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \middle| J_{2k,2k-1} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \right\rangle \equiv B_{2k,2k-1} = i \frac{\prod_{a=1}^{k-1} l_{2k-2,a} \prod_{a=1}^{k} l_{2k,a}}{\prod_{a=1}^{k-1} l_{2k-1,a}(l_{2k-1,a} - 1)}, \tag{13}$$

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} + 1 \\ m_{2k-1,j} \end{matrix} \middle| J_{2k+1,2k} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle \equiv C_{2k+1,2k}$$

$$= \frac{-i}{2} \left| \frac{\prod_{a=1}^{k-1} -(l_{2k-1,a} - l_{2k,j} - 1)(l_{2k-1,a} + l_{2k,j}) \prod_{b=1}^{k}(l_{2k+1,b} - l_{2k,j} - 1)(l_{2k+1,b} + l_{2k,j})}{\prod_{a\neq j}^{k}(l^2_{2k,a} - l^2_{2k,j})[l^2_{2k,a} - (l_{2k,j} + 1)^2]} \right|^{1/2}. \tag{14}$$

Thus a representation of $SO(2k,1)$ is represented by the Gel'fand pattern

$$\begin{pmatrix} m_{2k+1,1} & m_{2k+1,2} & \cdots & m_{2k+1,k} \\ m_{2k,1} & m_{2k,2} & \cdots & m_{2k,k} \\ & \cdots & & \\ & m_{21} & & \end{pmatrix}$$

and a representation of $SO(2k-1,1)$ is represented by the Gel'fand pattern

$$\begin{pmatrix} m_{2k,1} & m_{2k,2} & \cdots & m_{2k,k} \\ m_{2k-1,1} & \cdots & m_{2k-1,k-1} & \\ & \cdots & & \\ & m_{21} & & \end{pmatrix}.$$

The next step is to determine the relation between the various m's, or what is sometimes called the branching law. In order to do this, we have to use the unitarity condition.

III. CLASSIFICATION OF THE UNITARY REPRESENTATIONS OF $SO(n,1)$

The unitarity condition states that $J_{2k,2k-1}$ and $J_{2k+1,2k}$ are Hermitian matrices. This requirement leads to two conditions. The first condition is that the diagonal matrix elements for $J_{2k,2k-1}$ must be real. This means $B_{2k,2k-1}$ or

$$i\,\frac{\prod_{a=1}^{k-1} l_{2k-2,a}\prod_{a=1}^{k} l_{2k,a}}{\prod_{a=1}^{k-1} l_{2k-1,a}(l_{2k-1,a}-1)}$$

is real.

The second condition is that the product under the square root in Eqs. (12) and (14) must be positive. These two conditions enable us to classify all unitary irreducible representations of $SO(n,1)$. The results are as follows.

A. Discrete series: All m's either discrete integers or half-integers, excluding 0

1. For SO(2k-1,1)

There are no discrete series.

The reason is that from (13) it necessarily follows that one of the l's must be either pure imaginary or zero. This is incompatible with the definition of a discrete series.

2. For SO(2k,1)

We obtain the following branching laws. All m's are either integers or half-integers:

$$m_{2k+1,1} \geq m_{2k+1,2} \geq \cdots \geq m_{2k+1,2k}, \tag{15}$$

$$m_{i+1,j} \geq m_{i,j} \geq m_{i+1,j+1}, \quad j \leq i < 2k, \tag{16}$$

$$S = (0,1,2,\ldots,2k), \tag{17}$$

$$m_{2k+1,i-1} + 2 > m_{2k,i} > m_{2k+1,i} \quad \text{for all } i \leq S, \tag{18}$$

$$m_{2k+1,i+1} > m_{2k,i} > m_{2k+1,i+2} - 2 \quad \text{for all } i > S, \tag{19}$$

$$m_{2k+1,0} = \infty, \quad m_{2k+1,k+1} = -\infty. \tag{20}$$

Note that (15) and (16) are the same as the branching laws in $SO(2k+1)$, but that (17), (18), (19), (20) are different. However, they are similar to the branching laws of $U(2k,1)$. It is easy to see that the dimension of the representation is always infinite.

B. Continuous Series

1. For SO(2k-1,1)

1.1:

$$m_{2k,k} = ic, \quad c \text{ real}, \tag{21}$$

$$m_{2k,1} \geq m_{2k,2} \geq \cdots \geq m_{2k,k-1}, \tag{22}$$

$$m_{i+1,j} \geq m_{i,j} \geq m_{i+1,j+1}, \quad j \leq i < 2k-1, \tag{23}$$

$$m_{2k,i-1} + 2 > m_{2k-1,i} > m_{2k,i} \quad \text{for all } i \leq S, \tag{24}$$

$$m_{2k,i+1} > m_{2k-1,i} > m_{2k,i+2} - 2 \text{ for all } S < i < k-1, \tag{25}$$

$$S = (0,1,2,\ldots,k-1), \tag{26}$$

$$m_{2k,0} = \infty, \quad m_{2k,k-1} > -\infty. \tag{27}$$

We shall call this the principal proper series.

1.2: $m_{2k,k} = 0$. All others from (22) to (27) the same as above. We shall call this the supplementary improper series.

1.3: $m_{2k,k-1} = -1$. (22) to (27) the same. We shall call this the supplementary proper series.

These names are derived from consideration of the Lorentz group $SO(3,1)$. According to Gel'fand and Naimark, the Lorentz group is characterized by two numbers $(l_0 l_1)$ or $(k_0 c)$. They correspond to our notation as follows:

$$l_0 = k_0 = m_{41} + 1, \quad l_1 = c = m_{42}. \tag{28}$$

The reduced matrix elements are given by

$$\begin{aligned} A &= il_0 l_1[l(l+1)]^{-1}, \\ c &= i(l)^{-1}(l^2 - l_0^2)^{1/2}(l^2 - l_1^2)^{1/2}(4l^2 - 1)^{-1/2}. \end{aligned} \tag{29}$$

Thus 1.1 and 1.3 belong to the so-called principal and supplementary series, respectively. However, in the case of $SO(3,1)$ it can be seen from the matrix elements of the generators that the role of l_0 and l_1 can be interchanged. If one does so for the supplementary series, one obtains the supplementary improper series, 1.2. What happens when one interchanges l_1 and l_0 for the principal series? Here the situation is more complicated. But before we proceed, we wish to remark first that, as long as we are dealing with the representations of the Lie algebra, the restriction of l_0 to integers is really not necessary in the derivation of the matrix elements of the generators. (This follows from the proof of Theorem 1.) Thus we shall remove the restriction. It then follows that there are two more cases to be considered. The first one we shall call

1.4: Continuous-discrete series, or principal improper series, where $m_{2k,k-1} = -1 + ic$, c real, $m_{2k-1,k-1} \geq |m_{2k,k}|$, and all others from (22) to (27) the same.

The second one we shall call

1.5: Continuous-continuous series, where $m_{2k,k-1} = -1 + ic$, c real, $0 < m_{2k,k} < 1$, and all others from (22) to (27) the same.

One should note that 1.5 also has a counterpart, which is

1.6: $m_{2k,k} = ic$, $-1 < m_{2k,k-1} < 0$. All others from (22) to (27) remain the same. It is suggested that 1.6 might be called the proper continuous continuous series, and 1.5 the improper continuous continuous series. It should be noted that 1.5 and 1.6 have not been considered by Gel'fand and Naimark.

Let us examine further why this representation has not been included in Gel'fand and Naimark's list. If one follows Naimark's derivation[2] of the principal series, one finds that (on p. 148, Eq. 27), by virtue of the single-valued nature of the function $\chi(\lambda)$

$$\chi_2(\phi + 2\pi) = \chi_2(\phi)$$

and since $\chi_2(\phi) = e^{im\phi}$, m, which corresponds to our l_0, must be an integer. Thus the principal series, where m is an integer, corresponds to the single-valued nature of the function $\chi(\lambda)$.

If, however, one relaxes the restriction of the single-valued nature of $\chi(\lambda)$, and allows the function to be multi-valued, then it is no longer necessary to restrict m to integers, as long as $0 < m < 1$. The representation is well-defined. It is not equivalent to either the principal or the supplementary series, in that its matrix elements cannot be reproduced by either the principal or the supplementary series, and the space $L^2(z)$ can be regarded as the set of all measurable functions $f(z)$ of the parameter $z = x + iy$ satisfying the condition $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(z)|^2\, dx\, dy < \infty$. Thus it is a multivalued, integrable representation of the Lie algebra $SO(3,1)$.

To sum up the results for $SO(3,1)$: There are generally speaking six unitary irreducible representations.

1. Principal proper which is the same as the principal series of Gel'fand and Naimark.

2. Supplementary proper which is the same as the supplementary series of Gel'fand and Naimark.

3. Principal improper, which is obtained from principal proper by interchanging l_0 and l_1.

4. Supplementary improper which is obtained from supplementary proper by interchanging l_0 and l_1.

5. Proper continuous continuous, which has not been considered by Gel'fand and Naimark, our 1.6.

6. Improper continuous continuous, obtained from 1.6 by interchanging l_0 and l_1, our 1.5.

However, in $SO(3,1)$ the improper series eventually reproduce the same matrix elements as the proper series. Therefore, there are really only three distinct unitary representations in $SO(3,1)$. The new representation can be obtained from the Gel'fand–Naimark results by removing the restriction of l_0 to integers, which is feasible from Theorem 1. This, however, changes the representation from being single valued to multivalued, which may be the reason why it is omitted from the Gel'fand Naimark list.

For $SO(5,1)$ our results agree with those of Kuriyan, Mukunda, and Sudarshan.[5] For general k, one notices that our branching laws (22) and (23) seem to be different, in appearance, from those obtained by Hirai or Schwarz. However, this is only in appearance, not in substance, because of the interchangeability of the m_{ij} and m_{ik}. However, we do think that (24) and (25) are a better and more complete description of the branching law of $SO(n,1)$.

2. For $SO(2k,1)$

2.1: Principal series for $k > 1$:

$$m_{2k+1,1} = z + \tfrac{1}{2} - k/2, \tag{30}$$

$$m_{2k+1,k} = z^* - \tfrac{1}{2} + k/2, \tag{31}$$

where z is a complex number, and z^* its complex conjugate. All others from (15) to (20) remain the same.

In the case of $SO(2,1)$ where $k = 1$, then $m_{31} = -\frac{1}{2} + ic$, c real.

2.2: Supplementary series

(a)

$$m_{2k+1,1} = -c, \quad c \text{ real}, \quad 0 < c < 1, \tag{32}$$

$$m_{2k+1,k} = k - c - 1, \tag{33}$$

all other m's integers. (15) to (20) the same as above.

(b)

$$m_{2k+1,1} = \tfrac{1}{2} - c, \quad 0 < c < \tfrac{1}{2}, \tag{34}$$

$$m_{2k+1,k} = k - \tfrac{1}{2} - c, \tag{35}$$

all other m's half integers. (15) to (20) the same.

In the case of $SO(2,1)$, the discrete series consists of two parts: the positive series where

$$m_{21} = m_{31}, \quad m_{31} + 1, \ldots, \infty$$

and the negative series where

$$m_{21} = -m_{31}, \quad -m_{31} - 1, \ldots, -\infty.$$

The principal series yields

$$m_{31} = -\tfrac{1}{2} + ic,$$

m_{21} all integers or all half-integers.

The supplementary series yields what is called by Bargmann the exceptional interval. It also consists of two parts.

$$m_{31} = -c, \quad 0 < c < 1,$$

m_{21} all integers

or

$$m_{31} = \tfrac{1}{2} - c, \quad 0 < c < \tfrac{1}{2},$$

m_{21} all half-integers.

The classification of the unitary representations of $SO(4,1)$ has been done by Thomas and Newton.[3] However, they used different variables. Thomas used p, q, j_1, and j_2. Newton used Q, W, and j_1, j_2. These are related to our notation as follows:

$$Q = -q(q+1) - (p-1)(p+2) = 5/2 - (q + \tfrac{1}{2})^2 - (p + \tfrac{1}{2})^2, \tag{36}$$

$$W = -p(p+1)q(q+1) = -[(q + \tfrac{1}{2})^2 - \tfrac{1}{4}][(p + \tfrac{1}{2})^2 - \tfrac{1}{4}], \tag{37}$$

$$m_{51} = p - 1, \quad m_{52} = q, \tag{38}$$

$$m_{41} = j_1 + j_2, \quad m_{42} = j_1 - j_2. \tag{39}$$

It can be seen that all the features of the representations considered by Thomas and Newton are included in our classification.

IV. INVARIANTS

With the substitution (2) and (3) it can easily be seen that the invariants $I_{2k}(R)$ and $I_{2k}(L)$ commute with all the generators of $SO(n,1)$. Moreover, since the matrix elements of $SO(n,1)$ and $SO(n+1)$ are the same except with the change of (3), it follows that the invariants of $SO(n,1)$ have the same eigenvalues as those of $SO(n+1)$. This can be easily checked in the particular cases of $SO(2,1)$, $SO(3,1)$, and $SO(4,1)$ where everything is known. Let us just cite one example.

The fourth order invariant in $SO(4,1)$ is, according to Thomas,

$$I_4 = (LX + MY + NZ)^2 - (LU + MV + NW)^2 - (LT + WY - UZ)^2 - (MT + UZ - WX)^2 - (NT + VX - UY)^2$$

which has eigenvalue $p(p+1)q(q+1) = (m_{51}+1)(m_{51}+2)m_{52}(m_{52}+1)$. This is the same as the eigenvalue of the fourth order Louck invariant in $SO(5)$.[13]

V. NORMALIZED RAISING AND LOWERING OPERATORS OF $SO(n,1)$

The raising and lowering operators in $SO(n,1)$ are defined in such a way that they increase (decrease) the state labels of $SO(n)$ contained in $SO(n,1)$ by one, all the way down to the subgroup $SO(2)$. Thus following Wong's notation[11] we can define:

For $SO(2k-1,1)$

1. Lowering operator $L_{2k,p}$, $p = 1,2,\ldots,k-1$,

$$[J_{2\alpha,2\alpha-1}, L_{2k,p}] = -\delta_{\alpha p} L_{2k,p}, \quad 0 < \alpha \leq k-1, \tag{40}$$

$$[D_{p'+1,p'}, L_{2k,p}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0, \quad p' = 1,2,\ldots,k-2, \tag{41}$$

$$[E_{2k-1,k-1}, L_{2k,p}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0. \tag{42}$$

All generators A,B,C,D,E,F, are defined exactly the same as in Ref. 11. Equation (3) never enters in the definition of the lowering (raising) operators. The basis in (41), (42) is a "restricted" basis, where $SO(2k-1)$ is in its highest weight.

2. Raising operators $R_{2k,p}$, $p = 1,2,\ldots,k-1$,

$$[J_{2\alpha,2\alpha-1}, R_{2k,p}] = \delta_{\alpha p} R_{2k,p}, \quad 0 < \alpha \leq k-1, \tag{43}$$

$$[D_{p'+1,p'}, R_{2k,p}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0, \; p' = 1,2,\ldots,k-2, \tag{44}$$

$$[E_{2k-1,k-1}, R_{2k,p}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0. \tag{45}$$

3. "Zero-step" operator $O_{2k,k}$,

$$[J_{2\alpha,2\alpha-1}, O_{2k,k}] = 0, \quad 0 < \alpha \leq k-1, \tag{46}$$

$$[D_{p'+1,p'}, O_{2k,k}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0, \; p' = 1,2,\ldots,k-2, \tag{47}$$

$$[E_{2k-1,k-1}, O_{2k,k}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0, \tag{48}$$

$$[R_{2k,p} O_{2k,k}] \left| \begin{matrix} m_{2k,p} \\ m_{2k-1,p} \end{matrix} \right\rangle = 0, \quad p = 1,2,\ldots k-1. \tag{49}$$

For $SO(2k,1)$

1. Lowering operator $L_{2k+1,p}$, $p = 1,2,\ldots,k$,

$$[J_{2\alpha,2\alpha-1}, L_{2k+1,p}] = -\delta_{\alpha p} L_{2k+1,p}, \; \alpha = 1,2,\ldots,k, \tag{50}$$

$$[D_{p'+1,p'}, L_{2k+1,p}] \left| \begin{matrix} m_{2k+1,p} \\ m_{2k,p} \end{matrix} \right\rangle = 0, \quad p' = 1,2,\ldots,k-1, \tag{51}$$

$$[A_{k,k-1}, L_{2k+1,p}] \left| \begin{matrix} m_{2k+1,p} \\ m_{2k,p} \end{matrix} \right\rangle = 0. \tag{52}$$

2. Raising operator $R_{2k+1,p}$, $p = 1,2,\ldots,k$,

$$[J_{2\alpha,2\alpha-1}, R_{2k+1,p}] = \delta_{\alpha p} R_{2k+1,p}, \tag{53}$$

$$[D_{p'+1,p'}, R_{2k+1,p}] \left| \begin{matrix} m_{2k+1,p} \\ m_{2k,p} \end{matrix} \right\rangle = 0, \quad p' = 1,2,\ldots,k-1, \tag{54}$$

$$[A_{k,k-1}, R_{2k+1,p}] \left| \begin{matrix} m_{2k+1,p} \\ m_{2k,p} \end{matrix} \right\rangle = 0. \tag{55}$$

With the substitution from (2) and (3), these raising and lowering operators are the same as those given by Pang and Hecht, or by Wong [in Ref. 11, Eqs. (36) to (40)].

Moreover, the normalization constants $N^{q_m}_{q_{m-1}}$ and $N^{q_m}_{q_{m+1}}$ can be defined in the same way as done by Nagel and Moshinsky, i.e., for $SO(n-1,1)$

$$\left| \begin{matrix} h_1 \ldots h_m \ldots h_k \\ q_1 \ldots q_m - 1 \ldots \end{matrix} \right\rangle = \mathcal{L}^m_n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle = (N^{q_m}_{q_m-1})^{-1} L^m_n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle, \tag{56}$$

$$\left| \begin{matrix} h_1 \ldots h_m \ldots h_k \\ q_1 \ldots q_m + 1 \ldots \end{matrix} \right\rangle = \mathcal{R}^m_n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle = (N^{q_m}_{q_m+1})^{-1} R^m_n \left| \begin{matrix} h \\ q \end{matrix} \right\rangle. \tag{57}$$

Accordingly, the normalization constants $N^{q_m}_{q_{m-1}}$ and $N^{q_m}_{q_{m+1}}$ in $SO(n,1)$ have the same values as those in $SO(n+1)$. These two statements, however, still remain to be proved.

The first statement is that the raising and lowering operators obtained by Pang and Hecht in graphs and by Wong in algebraic form with the substitution from (2) and (3) do satisfy the definitions (40)–(55). The proof is quite simple. Equations (40)–(55) are all commutation relations acting on states which are the highest weight states in $SO(n)$ contained in $SO(n,1)$. Since the states in $SO(n)$ contained in $SO(n,1)$ are exactly the same as those contained in $SO(n+1)$, and since $J_{n+1,j}$ occur at most once in each term in the expression of the lowering and raising operators, the commutation relations (40)–(55) must be satisfied with the substitution from (2) and (3).

The second statement is that the normalization constants have the same value as given by equations (41)–(44) of Ref. 11. To prove this we reverse the process used in Ref. 11. In Ref. 11 we used the normalization constants to obtain the matrix elements of the generators. We now reverse the procedure and use the matrix elements to obtain the normalization constants. The fact is there is a definite relationship between the nor-

malization constants and the matrix elements of the generators.

First, for $SO(n-1,1)$, we use the Wigner–Eckart theorem to factor the matrix elements of $J_{n,n-1}$ into two parts: the reduced matrix element independent of $m_{n-2,i}$ and the matrix element of a vector operator V in $SO(n-1)$ which is dependent on $m_{n-2,i}$. As far as V goes everything is exactly the same in $SO(n-1,1)$ and $SO(n)$. So one has to look at the reduced matrix elements only. But in Ref. 11, we have explicitly

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j}+1 \\ m_{2k-1,j} \end{matrix} \middle| J_{2k,2k-1} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = \frac{-iN^{(2k)} \begin{cases} (h_j = m_{2k,j}) \\ q_j (= m_{2k-1,j}) \\ q_j + 1 \end{cases}}{\sqrt{2}\, N^{(2k-1)} \begin{cases} (h_j = m_{2k-1,j}+1) \\ q_j (= m_{2k-1,j}) \\ q_j + 1 \end{cases}}. \tag{58}$$

In the same way,

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k,j} \end{matrix} \middle| J_{2k+1,2k} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k,j} \end{matrix} \right\rangle = \frac{-iN^{(2k+1)} \begin{cases} h_j (= m_{2k+1,j}) \\ q_j (= m_{2k,j}) \\ q_j + 1 \end{cases}}{N^{(2k)} \begin{cases} h_j (= m_{2k,j}+1) \\ q_j (= m_{2k,j}) \\ q_j + 1 \end{cases}}. \tag{59}$$

In terms of $SO(n-1,1)$ we would have, for the left side of (58) and (59), $\langle\, |-iJ|\, \rangle$. However, in view of Theorem 1, this has to be multiplied by another i when the matrix element is written in terms of the m's. Thus the N's are exactly the same functions of the m's as before.

One cannot fail, of course, to notice the similarity between this result for $SO(n,1)$ and the result recently obtained by Patera for $U(p,1)$. However, one has to note that the raising and lowering operators are defined differently in the two cases. The main difference is that in $U(n,1)$ there is a one-to-one correspondence between the Gel'fand pattern and weight, while in $SO(n,1)$ this one-to-one correspondence does not exist.

The normalized raising and lowering operators can now be used directly in the discrete series of $SO(2k,1)$, where, in a sense, there exists a highest weight (for a negative discrete series) and a lowest weight (for a positive discrete series). Thus one can generate all basis functions by these operators, in terms of, for example, the boson operators.

For the continuous series of $SO(n,1)$ it is not possible to define either a highest weight or a lowest weight. However, one can select an arbitrary state, assume it to be normalized, and apply the raising (lowering) operators to generate the other states. The quantities to be calculated are eventually to be referred to the initially chosen normalized state. In this sense it is still possible to apply the boson technique to $SO(n,1)$.

CONCLUSION

We have tried to understand the representations of $SO(n,1)$ as a whole. Even though the application of $SO(n,1)$ for $n>4$ to physics may still take a long time to come, it is always gratifying to be able to say we know mathematically how to use it in case it is needed. We think therefore that the study of $SO(n,1)$ for general n is not a waste of time. As a side product, we have obtained a new representation in $SO(3,1)$ not considered by Gel'fand and Naimark. We have also attempted to make a complete general classification of all the unitary representations of $SO(n,1)$. One of the general results is that there are no discrete series in $SO(2k-1, 1)$ where all m's are either integers or half-integers (except for the special case where one of the l's is 0).

It is surprising to see what a great change a mere factor of $\sqrt{-1}$ can make. In essence, it changes the branching laws in $U(p)$ and $SO(p)$ to the different branching laws in $U(p-1,1)$ and $SO(p-1,1)$. This in turn is the reason why the unitary representations of compact groups are finite dimensional while the unitary representations of noncompact groups are infinite-dimensional.

[1] V. Bargmann, Ann. Math. 48, 568 (1947).

[2] I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups* (Pergamon, New York, 1963). M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964).

[3] L. H. Thomas, Ann. Math. 42, 113 (1941). T. D. Newton, Ann. Math. 51, 730 (1950). See also J. Dixmier, Bull. Soc. Math. France 89, 9 (1961) and S. Strom, Arkiv Fysik 30, 455 (1965).

[4] A. O. Barut, P. Budini, and C. Fonsdal, Proc. Roy. Soc. (London) A291, 106 (1966). R. Musto, Phys. Rev. 148, 1274 (1966). R. H. Pratt and T. F. Jordan, Phys. Rev. 148, 1276 (1966). M. Bander and C. Itzykson, Rev. Mod. Phys. 38, 330 (1966); 38, 346 (1966).

[5] J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. 8, 204 (1968).

[6] T. Hirai, Proc. Japan Acad. 38, (83), 258 (1962). U. Ottoson, Commun. Math. Phys. 8, 228 (1968). A. Chakrabarti, J. Math. Phys. 9, 2087 (1968). F. Schwarz, J. Math. Phys. 12, 131 (1971).

[7] A. V. Nikolov, Funct. Anal. Appl. 3, 337 (1969).

[8] J. Patera, J. Math. Phys. 14, 279 (1973).

[9] J. G. Nagel and M. Moshinsky, J. Math. Phys. 6, 682 (1965).

[10] S. C. Pang and K. T. Hecht, J. Math. Phys. 8, 1233 (1967).

[11] M. K. F. Wong, J. Math. Phys. 8, 1899 (1967).

[12] G. Racah, Atti Acad. Nazl. Lincei, Rend., Classe Sci. Fis., Mat. e Nat. 8, 108 (1950). J. D. Louck, Los Alamos Scientific Laboratory, Report LA 2451, 1960.

[13] See J. D. Louck, Ref. 12, or M. K. F. Wong, J. Math. Phys. 12, 1530, (1971). Note the misprint in Eq. (51). The fifth term on the right-hand side should read $12\, m_{31}^2$ instead of $12\, m_{32}^2$.

Scattering of photons by an external current

G. Reents

Physikalisches Institut, Universität Würzburg, Germany
(Received 14 May 1973; revised manuscript received 8 August 1973)

Considering the interaction of a quantized electromagnetic field with a prescribed external c-number current, we show that, for each current, there is a representation of the field operators in a separable Hilbert space and a unitary S operator transforming the free incoming field into the free outgoing one. The representation, a generalized product representation, depends on the current so that, in general, for two different currents the corresponding representations are unitarily inequivalent. Photon states and asymptotic observables are defined for the case of an infrared divergent current.

1. INTRODUCTION

In connection with infrared divergences Kibble[1] has discussed the interaction of a classical current with a quantized electromagnetic field. Generalized coherent states are used to construct an S matrix which, in general, is unitary only in a nonseparable Hilbert space.

There are essentially two problems one has to deal with. The first one occurs when the current, more precisely, the transverse part of its Fourier transform on the mass shell is not square integrable. The corresponding S operator maps coherent state representations of the canonical commutation relations (hereafter referred to as CCR's) onto coherent state representations. But restricted to separable subspaces these representations are unitarily inequivalent. This is the reason why Kibble needs a nonseparable Hilbert space. The second problem comes in by the possibly infinite Coulomb phase due to the selfenergy of the current. Considering a wider class of representations of the CCR's we want to show that the S operator can be defined as a unitary operator in a separable Hilbert space.

Let $J_\mu(x)$ be the external current which is required to be conserved:

$$\partial_\mu J^\mu(x) = 0. \tag{1.1}$$

We use a "timelike" metric so that the scalar product in Minkowski space becomes $kx = k^0x^0 - \mathbf{kx}$. By $\mathbf{j}(x)$ we denote the divergence-free space part of the current. Then, in the radiation gauge Maxwell's equation reads

$$\partial_\mu \partial^\mu \mathbf{A}(x) = \mathbf{j}(x). \tag{1.2}$$

The field $\mathbf{A}(\mathbf{x}, x^0)$ and its time derivative at a certain time, e.g., $x^0 = 0$, may be defined in the usual way by their Fourier decomposition

$$\begin{aligned}\mathbf{A}(\mathbf{x}, 0) &= (2\pi)^{-3/2}\int d\mu(\mathbf{k})[\mathbf{a}(\mathbf{k})e^{i\mathbf{kx}} + \mathbf{a}^*(\mathbf{k})e^{-i\mathbf{kx}}],\\ \dot{\mathbf{A}}(\mathbf{x}, 0) &= (2\pi)^{-3/2}\int d\mu(\mathbf{k})i|\mathbf{k}|[-\mathbf{a}(\mathbf{k})e^{i\mathbf{kx}} + \mathbf{a}^*(\mathbf{k})e^{-i\mathbf{kx}}],\end{aligned} \tag{1.3}$$

the integration measure being $d\mu(\mathbf{k}) = d^3\mathbf{k}/2|\mathbf{k}|$. The smeared operators $\mathbf{af}^* = \int d\mu(\mathbf{k})\mathbf{a}(\mathbf{k})\mathbf{f}^*(\mathbf{k})$ and $\mathbf{a}^*\mathbf{f} = \int d\mu(\mathbf{k})\mathbf{a}^*(\mathbf{k})\mathbf{f}(\mathbf{k})$, where the test functions are required to be transverse, i.e., $\mathbf{kf}(\mathbf{k}) = 0$, satisfy the CCR's

$$\begin{aligned}&[\mathbf{a}^*\mathbf{f}, \mathbf{a}^*\mathbf{g}] = 0 = [\mathbf{af}^*, \mathbf{ag}^*],\\ &[\mathbf{af}^*, \mathbf{a}^*\mathbf{g}] = \mathbf{f}^*\mathbf{g} = \int d\mu(\mathbf{k})\mathbf{f}^*(\mathbf{k})\mathbf{g}(\mathbf{k}).\end{aligned} \tag{1.4}$$

As a formal solution of (1.2) we have, with free fields $\mathbf{A}_{\text{in}}(x)$ and $\mathbf{A}_{\text{out}}(x)$,

$$\mathbf{A}(x) = \mathbf{A}_{\substack{\text{in}\\\text{out}}}(x) + \int d^4y D_{\substack{\text{ret}\\\text{adv}}}(x - y)\mathbf{j}(y), \tag{1.5}$$

$D_{\substack{\text{ret}\\\text{adv}}}(x)$ being the usual retarded or, respectively, advanced zero-mass Green's function. So, the free in and out fields are related by

$$\mathbf{A}_{\text{out}}(x) = \mathbf{A}_{\text{in}}(x) + \int d^4y D(x - y)\mathbf{j}(y), \tag{1.6}$$

with

$$D(x) = i(2\pi)^{-3}\int d\mu(\mathbf{k})(e^{-ikx} - e^{ikx}), \tag{1.7}$$

and $k^0 = |\mathbf{k}|$. Thus,

$$\mathbf{a}_{\text{out}}(\mathbf{k}) = \mathbf{a}_{\text{in}}(\mathbf{k}) + i\mathbf{j}(\mathbf{k}), \tag{1.8}$$

where $\mathbf{j}(\mathbf{k})$ is the Fourier transform of $\mathbf{j}(x)$ on the mass shell, namely

$$\mathbf{j}(\mathbf{k}) = (2\pi)^{-3/2}\int d^4x e^{i(|\mathbf{k}|x^0 - \mathbf{kx})}\mathbf{j}(x). \tag{1.9}$$

From (1.8) one sees that the S operator must be of the form

$$S(J) = \exp[i(\mathbf{a}_{\text{in}}^*\mathbf{j} + \mathbf{a}_{\text{in}}\mathbf{j}^*)]\exp(i\sigma). \tag{1.10}$$

The phase factor $e^{i\sigma}$ can be computed by use of Schwinger's variational derivative technique. It comes out as[1]

$$\begin{aligned}\sigma &= \int d\mu(\mathbf{k})\sigma(\mathbf{k}),\\ \sigma(\mathbf{k}) &= \tfrac{1}{2}\int dk^0 J^\mu(k)J_\mu^*(k)\left(\frac{1}{k^0 - |\mathbf{k}|} - \frac{1}{k^0 + |\mathbf{k}|}\right).\end{aligned} \tag{1.11}$$

The last integral has to be taken as a principal value integral and $J(k)$ is the four-dimensional Fourier transform of the current, $J(k) = (2\pi)^{-2}\int d^4x e^{ikx}J(x)$.

Our aim is to give a representation of the CCR's so that an operator of the form (1.10) becomes a unitary operator in a separable Hilbert space. This means, neglecting the phase factor, that a transformation $\mathbf{a}(\mathbf{k}) \to \mathbf{a}(\mathbf{k}) + i\mathbf{j}(\mathbf{k})$ is unitarily implementable whatever the current may be. The only restrictions will be that

$$\begin{aligned}&\mathbf{j}(\mathbf{k}) \text{ and } \sigma(\mathbf{k}) \text{ are measurable with respect to}\\ &\mu(\mathbf{k}) \text{ and are finite } \mu\text{-almost everywhere.}\end{aligned} \tag{1.12}$$

We want to emphasize that we are not trying to diagonalize the Hamiltonian at some finite time, or to find that representation where this Hamiltonian has a normalizable ground state. For linear interactions, similar to that one in our model, this problem has been discussed by Verbeure and Verboven.[2]

In Sec. 2 we define a class of representations of the CCR's which, as we show in Sec. 3, is rich enough to contain representations so that the S operator (1.10)

becomes unitary. Section 4 deals with an infrared divergent current. Photon states and asymptotic observables are defined for this case, and the description of scattering experiments is outlined.

After all we should note that the electromagnetic field serves as an example only. It can be replaced by other Bose fields with external sources.

2. GENERALIZED PRODUCT REPRESENTATIONS OF THE CCR's

Instead of representing the CCR's in the form (1.4), one usually passes to the unitary operators

$$W(\mathbf{f}) = \exp(\mathbf{a}^*_{\text{in}}\mathbf{f} - \mathbf{a}_{\text{in}}\mathbf{f}^*). \tag{2.1}$$

(We use now $\mathbf{a}^*_{\text{in}}, \mathbf{a}_{\text{in}}$ because we want to express the S operator in terms of these.) By a formal application of (1.4) we obtain the multiplication law

$$W(\mathbf{f})W(\mathbf{g}) = W(\mathbf{f}+\mathbf{g})e^{i(\mathbf{f},\mathbf{g})}, \tag{2.2}$$

and, as a consequence of it, the CCR's

$$W(\mathbf{f})W(\mathbf{g}) = W(\mathbf{g})W(\mathbf{f})e^{2i(\mathbf{f},\mathbf{g})}, \tag{2.3}$$

where $(\mathbf{f},\mathbf{g})$ is the real skew-symmetric bilinear form given by

$$(\mathbf{f},\mathbf{g}) = (1/2i)\int d\mu(\mathbf{k})[\mathbf{f}(\mathbf{k})\mathbf{g}^*(\mathbf{k}) - \mathbf{f}^*(\mathbf{k})\mathbf{g}(\mathbf{k})]. \tag{2.4}$$

We call representation of the CCR's a map of the test functions $\mathbf{f} \to W(\mathbf{f})$ into unitary operators $W(\mathbf{f})$ satisfying (2.2). To guarantee the existence of the field operators by Stone's theorem it is required that $W(\lambda\mathbf{f})$ is weakly continuous in the real parameter λ for each fixed $\mathbf{f}$ of the test function space V.

According to (1.12) we can find a partition of R^3 into μ-measurable subsets I_r with $I_r \cap I_{r'} = \emptyset$ for $r \neq r'$, $\cup_r I_r = R^3$ such that

$$\int_{I_r} d\mu(\mathbf{k})\mathbf{j}(\mathbf{k})\mathbf{j}^*(\mathbf{k}) < \infty, \quad \left|\int_{I_r} d\mu(\mathbf{k})\sigma(\mathbf{k})\right| < \infty \tag{2.5}$$

for all r. $\mu(I_r)$ need not be finite. Having fixed this partition we restrict the test functions in the same way, i.e.,

$$V = \{\mathbf{f} | \int_{I_r} d\mu(\mathbf{k})\mathbf{f}(\mathbf{k})\mathbf{f}^*(\mathbf{k}) < \infty \text{ for all } r\}. \tag{2.6}$$

The partition of R^3 induces a decomposition of V into—with respect to the bilinear form $(\mathbf{f},\mathbf{g})$ orthogonal—subspaces V_r, $V_r = \{\mathbf{f} | \mathbf{f} \in V, \operatorname{supp}\mathbf{f} \subset I_r\}$. Let H_r be the Fock space generated by the $\mathbf{a}_{\text{in}}(\mathbf{k}), \mathbf{a}^*_{\text{in}}(\mathbf{k})$ with test functions $\mathbf{f}_r \in V_r$. Because the $\mathbf{f}_r(\mathbf{k})$ are square integrable with respect to $d\mu(\mathbf{k})$, $W_r(\mathbf{f}_r) = \exp(\mathbf{a}^*_{\text{in}}\mathbf{f}_r - \mathbf{a}_{\text{in}}\mathbf{f}^*_r)$ are unitary operators in H_r which of course satisfy (2.2). Consider a subspace $V^0 \subset V$ defined by

$$V^0 = \{\mathbf{f} | \mathbf{f} \in V, \mathbf{f}_r(\mathbf{k}) = 0 \text{ for almost all } r\}. \tag{2.7}$$

Let $H = \otimes_r (H_r, \phi_r)$ be the incomplete tensor product[3] of the H_r with reference vector $\otimes_r \phi_r$, $\phi_r \in H_r$, $\|\phi_r\| = 1$ for all r. For test functions $\mathbf{f} \in V^0, \mathbf{f} = \sum_r \oplus\, \mathbf{f}_r$, we define in H a representation of the CCR's by

$$W(\mathbf{f}) = \otimes_r W_r(\mathbf{f}_r). \tag{2.8}$$

Let us call this a generalized product representation (GPR) of the CCR's. It differs from the often studied direct product and partial-tensor-product representations[4-7] because the subspaces V_r have infinite dimension. Irreducibility follows from the irreducibility of the Fock representation in each factor. Theorem 2.1 of Ref. 7 is immediately carried over, i.e., two GPR with reference vectors $\otimes_r \phi_r$ and $\otimes_r \psi_r$, respectively, are unitarily equivalent if and only if $\sum_r(1 - |\langle\phi_r, \psi_r\rangle|) < \infty$. This means for instance that a GPR $(\otimes_r \phi_r)$ is equivalent to the Fock representation if $\sum_r(1 - |\langle\phi_r, \Omega_r\rangle|) < \infty$, where Ω_r denotes the Fock vacuum in each H_r. If there exists a sequence of coherent states $|\mathbf{f}_r\rangle = W_r(\mathbf{f}_r)\Omega_r$ such that $\sum_r(1 - |\langle\phi_r|\mathbf{f}_r\rangle|) < \infty$ then the GPR $(\otimes_r \phi_r)$ is equivalent to a (generalized) coherent state representation. Clearly, each coherent state representation is a GPR. However, the coherent states form a total but not dense set in each H_r and, therefore, there exist vectors $\phi_r \in H_r$ such that $\sum_r(1 - |\langle\phi_r|\mathbf{f}_r\rangle|)$ diverges for each sequence of coherent states $|\mathbf{f}_r\rangle$.

The test function space V^0 is incomplete in the sense that there exist sequences of test functions $\{\mathbf{f}^{(n)}\}, \mathbf{f}^{(n)} \in V^0$, such that $W(\mathbf{f}^{(n)})$ converges strongly to a unitary $W(\{\mathbf{f}^{(n)}\})$ and there is no $\mathbf{f} \in V^0$ with $W(\mathbf{f}) = W(\{\mathbf{f}^{(n)}\})$. It has been shown[8,9] that the representation itself, at first defined for test function $\mathbf{f} \in V^0$, determines the completion of V^0. For product representations characterized by $\otimes_r \phi_r$ the admissible test functions are just those $\mathbf{f} = \sum_{r=1}^{\infty} \oplus\, \mathbf{f}_r$ for which $\otimes_r W_r(\mathbf{f}_r)\phi_r$ is weakly equivalent to $\otimes_r \phi_r$.[9,10] If one wants the completion to be a linear vector space one has to impose on $\mathbf{f}$ a slightly more restrictive condition. The corresponding operators are of the form $W(\mathbf{f}) = \otimes_r W_r(\mathbf{f}_r)e^{i\alpha_r}$ with certain real numbers α_r. We shall show in the next section that in a suitably chosen GPR the S operator becomes one of these operators.

3. THE REPRESENTATION FOR A GIVEN CURRENT

Let $\mathbf{j}_r(\mathbf{k}) = \mathbf{j}(\mathbf{k})\chi_r(\mathbf{k})$ and $\sigma_r = \int d\mu(\mathbf{k})\sigma(\mathbf{k})\chi_r(\mathbf{k})$ where $\chi_r(\mathbf{k})$ is the characteristic function of I_r. We define the truncated current and phase by $\mathbf{j}^{(n)}(\mathbf{k}) = \sum_r \mathbf{j}^{(n)}_r(\mathbf{k})$ and $\sigma^{(n)} = \sum_r \sigma^{(n)}$:

$$\mathbf{j}^{(n)}_r(\mathbf{k}) = \begin{cases} \mathbf{j}_r(\mathbf{k}), & r \leq n \\ 0, & r > n \end{cases}, \qquad \sigma^{(n)}_r = \begin{cases} \sigma_r, & r \leq n \\ 0, & r > n \end{cases}.$$

In the GPR defined in (2.5)–(2.8) the corresponding S operator becomes

$$S(J^{(n)}) = e^{i\sigma^{(n)}} \otimes_r W_r(i\mathbf{j}^{(n)}_r). \tag{3.1}$$

Certainly, $S(J^{(n)})$ is unitary in each incomplete tensor product space because $W_r(i\mathbf{j}^{(n)}_r) = 1_r$ for almost all r.

Lemma 3.1: (Araki, Woods): There exists a subsequence n_s, $s = 1, 2, \cdots, n_s > n_{s-1}$ for all s (where $n_0 = 0$) such that $\lim_{s\to\infty} e^{i\sigma^{(n_s)}}$ exists.

Proof: Consider the sequence $\sigma_r, r = 1, 2, \cdots$. There exist integers m_k and a sequence n_k as above such that[9]

$$\sum_{k=1}^{\infty} |2\pi m_k - \sum_{r=n_{k-1}+1}^{n_k} \sigma_r| \leq \pi.$$

Let

$$2\pi m_k - \sum_{r=n_{k-1}+1}^{n_k} \sigma_r = \epsilon_k.$$

Then

$$e^{i\sigma(n_s)} = e^{i\sum_{r=1}^{n_s}\sigma_r} = e^{i\sum_{k=1}^{s}\epsilon_k}$$

which converges because $\sum_{k=1}^{s}|\epsilon_k|$ does so.

Let us therefore assume that the original partition I_r has been chosen in such a way that $\lim_{n\to\infty} e^{i\sigma^{(n)}}$ exists.

In our model $e^{i\sigma}$ is an over-all phase factor which has no physical meaning. It could have been omitted if we only require the S matrix to transform $\mathbf{a}_{\text{in}}(\mathbf{k})$ into $\mathbf{a}_{\text{out}}(\mathbf{k})$. Anyhow, the preceding Lemma shows how one can get rid of such an infinite phase.

Lemma 3.2: Let $U_r, r = 1, 2, \cdots$, be unitary operators in Hilbert spaces H_r. Let $U_r^{(n)} = U_r$ for $r \leq n$, $U_r^{(n)} = 1_r$ for $r > n$ and define $U^{(n)} = \bigotimes_r U_r^{(n)}$. If

$$\sum_r |1 - \langle\phi_r, U_r\phi_r\rangle| < \infty, \quad \phi_r \in H_r, \quad \|\phi_r\| = 1, \tag{3.2}$$

then

(i) $U = \bigotimes_r U_r$ is unitary in $H = \bigotimes_r (H_r, \phi_r)$.

(ii) $U = \text{s-lim}_{n\to\infty} U^{(n)}$.

Proof: (i) The product vectors $\psi = \bigotimes_r \psi_r$ with $\psi_r = \phi_r$ for almost all r form a total set in H. Equation (3.2) implies $\sum_r |1 - \langle\phi_r, U_r\psi_r\rangle| < \infty$, thus[3] $U\psi = \bigotimes_r U_r\psi_r \in H$. Hence U is an operator in H and it is unitary because (3.2) implies $\sum_r |1 - \langle\phi_r, U_r^*\phi_r\rangle| < \infty$.

(ii) It suffices [see (i)] to show that $\lim_{n\to\infty} \|(U^{(n)} - U)\bigotimes_r \phi_r\| = 0$.

But

$$\|(U^{(n)} - U)\bigotimes_r \phi_r\|^2 = 2(1 - \text{Re}\prod_r \langle U_r^{(n)}\phi_r, U_r\phi_r\rangle)$$
$$= 2(1 - \text{Re}\prod_{r>n}\langle\phi_r, U_r\phi_r\rangle) \underset{n}{\to} 0,$$

for (3.2) implies that $\Pi_r\langle\phi_r, U_r\phi_r\rangle$ converges.

In our case H_r is the Fock space with test functions having support in I_r and $U_r = W_r(i\mathbf{j}_r)$. It depends on the spectrum of the W_r whether one can find a sequence of vectors $\phi_r \in H_r$ satisfying (3.2). Let $W_F(\mathbf{f})$ be the Fock representation of the CCR's with test functions $\mathbf{f} \in V_F$. It is known[11] that V_F can be taken as a Hilbert space with inner product $\langle\mathbf{f}, \mathbf{g}\rangle = \int d\mu(\mathbf{k})\mathbf{f}^*(\mathbf{k})\mathbf{g}(\mathbf{k})$. The bilinear form (2.4) becomes $(\mathbf{f}, \mathbf{g}) = -\text{Im}\langle\mathbf{f}, \mathbf{g}\rangle$. For any $\mathbf{g} \in V_F$, $\mathbf{g} \neq 0$, let $\mathbf{A}_F\mathbf{g}$ denote the self-adjoint generator of the one parameter unitary group

$$\{W_F(\lambda\mathbf{g}); -\infty < \lambda < \infty\}.$$

Lemma 3:3: The spectrum of $\mathbf{A}_F\mathbf{g}$ is purely continuous and covers the whole real line.

Proof: This is clear because the Fock representation when restricted to one degree of freedom becomes unitarily equivalent to a direct sum of Schrödinger representations. The representation space may be written as $H = H_{\mathbf{g}} \otimes H', H_{\mathbf{g}} = L^2(R^1)$, such that $W_F(\mathbf{g})$ becomes $W_F(\mathbf{g}) = e^{i\gamma x} \otimes 1'$ with a real γ.

Lemma 3.4: There exists a sequence of vectors $\phi_r \in H_r$, (H_r Fock spaces) $\|\phi_r\| = 1$ for all r, such that

$$\sum_r |1 - \langle\phi_r, W_r(i\mathbf{j}_r)\phi_r\rangle| < \infty. \tag{3.3}$$

Proof: If $\mathbf{j}_r = 0$, then $\langle\phi_r, W_r(i\mathbf{j}_r)\phi_r\rangle = 1$. For $\mathbf{j}_r \neq 0$ we write $W_r(i\mathbf{j}_r) = \int e^{i\lambda} dE_\lambda^{(r)}$ where $E_\lambda^{(r)}$ is the resolution of the identity for the infinitesimal generator of $W_r(i\mathbf{j}_r)$. Let $\epsilon_r > 0$ be a sequence of positive real numbers so that $\sum_r \epsilon_r < \infty$. According to Lemma 3.2 $(E_{\epsilon_r}^{(r)} - E_{-\epsilon_r}^{(r)})H_r = H_{r,\epsilon_r}$ is a nontrivial subspace of H_r. Take a unit vector $\phi_r \in H_{r,\epsilon_r}$. Then

$$\sum_r |1 - \langle\phi_r, W_r(i\mathbf{j}_r)\phi_r\rangle| = \sum_r |\int (1 - e^{i\lambda}) d\langle\phi_r, E_\lambda^{(r)}\phi_r\rangle|$$
$$\leq \sum_r \int_{-\epsilon_r}^{\epsilon_r} |1 - e^{i\lambda}| d\langle\phi_r, E_\lambda^{(r)}\phi_r\rangle$$
$$\leq \sum_r \max_{|\lambda|\leq\epsilon_r} 2\left|\sin\frac{\lambda}{2}\right| \leq \sum_r \epsilon_r < \infty.$$

So, we have according to Lemma 3.1

$$\text{s-lim}_{n\to\infty} S(J^{(n)}) = \text{s-lim}_{n\to\infty} e^{i\sigma^{(n)}} \bigotimes_r W_r(i\mathbf{j}_r^{(n)})$$

exists and is unitary in $H = \otimes_r(H_r, \phi_r)$. Let us summarize these results as

Theorem 3.1: For a given current there exists an irreducible representation of the CCR's in a separable Hilbert space so that an S operator of the form (1.10) can be given a meaning as a unitary operator in this space.

It is clear from this construction that different currents lead to different and, in general, unitarily inequivalent representations. Even for one current the construction is not unique. It is not hard to find a sequence of vectors $\chi_r \in H_r$, also fulfilling the conditions of Lemma 3.3 and such that $\sum_r(1 - |\langle\phi_r, \chi_r\rangle|)$ does not converge. This means that the representations in $H_1 = \otimes_r(H_r, \phi_r)$ and $H_2 = \otimes_r(H_r, \chi_r)$ are unitarily inequivalent. However, as we want to show now, any two such representations are physically equivalent in the sense that they lead to the same predictions for photon scattering.

4. PHOTON STATES AND ASYMPTOTIC OBSERVABLES

The above formalism was mainly developed in order to investigate infrared problems which arise if $\int d\mu(\mathbf{k})\mathbf{j}^*(\mathbf{k})\mathbf{j}(\mathbf{k})$ is divergent at $\mathbf{k} = 0$ (we assume this to be the only point of divergence). The subsets $I_r, r = 1, 2, \cdots$, which form a partition of $\mathbf{k}$ space might be chosen as spherical shells, i.e.,

$$I_r = \{\mathbf{k} | d_r \leq |\mathbf{k}| < d_{r-1}\}, \tag{4.1}$$

where the radius d_r goes to zero and I_1 may be given by $I_1 = \{\mathbf{k} | 1 \leq |\mathbf{k}| < \infty\}$.

In order to define physically meaningful asymptotic observables and corresponding physical photon states, we take into account that for any given scattering experiment (preparation of an incoming state and detection of outgoing photons) there is an energy threshold Λ below which single photons cannot be detected. Accordingly, we define the asymptotic observables corresponding to such an experiment to be functions of $\mathbf{a}_{\text{ex}}^*(\mathbf{k})$ and $\mathbf{a}_{\text{ex}}(\mathbf{k})$ with $|\mathbf{k}| \geq \Lambda$, where "ex" means "in" or "out." This threshold distinguishes "hard" (= detectable) photons, with $|\mathbf{k}| \geq \Lambda$, from the remaining field excitations which might be called "soft photons" (although the latter term is somewhat misleading since, e.g., there is no total number operator for "soft photons" in our state space H). Of course, the energy threshold Λ, and thus the distinction between "hard" and "soft" photons, depends on the experiment considered. We will show, however, that our formalism is applicable for any $\Lambda > 0$, and is thus able

to describe scattering experiments with any finite energy resolution.

Referring to the partition I_r (4.1) there is for every $\Lambda > 0$ a finite $r(\Lambda)$ such that

$$d_{r(\Lambda)} \leqslant \Lambda < d_{r(\Lambda)-1}. \tag{4.2}$$

Consider the space

$$H' = \overset{r(\Lambda)}{\underset{r=1}{\otimes}} H_r$$

which is clearly the Fock space of photons with an energy $|\mathbf{k}| \geqslant d_{r(\Lambda)}$. We may factorize H' as $H' = H_1(\Lambda) \otimes H''$, $H_1(\Lambda)$ being the Fock space of photons with momenta $|\mathbf{k}| \geqslant \Lambda$, so that the total Hilbert space

$$H = \underset{r}{\otimes} (H_r, \phi_r) = H' \otimes \{\underset{r>r(\Lambda)}{\otimes} (H_r, \phi_r)\}$$
$$= H_1(\Lambda) \otimes H'' \otimes \{\underset{r>r(\Lambda)}{\otimes} (H_r, \phi_r)\}$$

may be written as

$$H = H_1(\Lambda) \otimes H_2(\Lambda). \tag{4.3}$$

Certainly, the scattering operator admits a similar decomposition, i.e.,

$$S = S_1(\Lambda) \otimes S_2(\Lambda), \tag{4.4}$$

where up to a finite phase factor $S_1(\Lambda)$ is given by

$$S_1(\Lambda) = \exp(i \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \{\mathbf{a}^*_{\text{in}}(\mathbf{k})\mathbf{j}(\mathbf{k}) + \mathbf{a}_{\text{in}}(\mathbf{k})\mathbf{j}^*(\mathbf{k})\}) \tag{4.5}$$

which is unitary in $H_1(\Lambda)$ and $S_2(\Lambda)$ acts unitarily on $H_2(\Lambda)$. Photon states are described by special unit vectors $\Phi \in H, \Phi = \Phi_1 \otimes \Phi_2$ with $\Phi_i \in H_i(\Lambda)$ and $\|\Phi_i\| = 1$ for $i = 1, 2$. According to the foregoing considerations the asymptotic observables are of the form $A = A_1 \otimes 1_2$ so that the second factor Φ_2 of Φ cannot be determined by preparing or measuring processes. Let $\Phi_1 \in H_1(\Lambda)$ be an initially prepared state. If we ask for the probability of finding the state $\Psi_1 \in H_1(\Lambda)$ after scattering, we have to compute

$$\langle \Phi | S(P(\Psi_1) \otimes 1_2) S^{-1} | \Phi \rangle$$
$$= \langle \Phi_1 | S_1(\Lambda) P(\Psi_1) S_1^{-1}(\Lambda) | \Phi_1 \rangle \langle \Phi_2 | S_2(\Lambda) S_2^{-1}(\Lambda) | \Phi_2 \rangle$$
$$= |\langle \Phi_1 | S_1(\Lambda) \Psi_1 \rangle|^2, \tag{4.6}$$

where $P(\Psi_1)$ is the projection on to the one-dimensional subspace of $H_1(\Lambda)$ generated by Ψ_1. The result (4.6) obviously does not depend on the special representation of the CCR's constructed in Sec. 3 because only the Fock states Φ_1, Ψ_1 enter. It is also independent of the "soft photon" part Φ_2 of the initial state Φ, which was not fixed anyway by the preparation of Φ and must thus be considered as completely arbitrary. To discuss an example we consider the n-particle bremsstrahlung, i.e., we put $\Phi_1 = |0_{\text{in}}\rangle$ and ask for the probability ω_n to find n photons after scattering. We have, therefore, to replace $P(\Psi_1)$ of Eq. (4.6) by the projection P_n onto the n-particle subspace of $H_1(\Lambda)$.

Writing P_n as

$$P_n = \sum \frac{1}{n!} (\mathbf{a}^*_{\text{in}}\mathbf{f}_{l_1})(\mathbf{a}^*_{\text{in}}\mathbf{f}_{l_2}) \cdots (\mathbf{a}^*_{\text{in}}\mathbf{f}_{l_n}) |0_{\text{in}}\rangle\langle 0_{\text{in}}| (\mathbf{a}_{\text{in}}\mathbf{f}_{l_n}) \cdots (\mathbf{a}_{\text{in}}\mathbf{f}_{l_1})$$

with a (in the region $|\mathbf{k}| \geqslant \Lambda$) complete system of orthonormal vectors $\mathbf{f}_l(\mathbf{k})$, one obtains after some elementary calculations

$$\omega_n = \frac{1}{n!} \left(\int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \mathbf{j}^*(\mathbf{k}) \mathbf{j}(\mathbf{k}) \right)^n \exp\left(- \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \mathbf{j}^*(\mathbf{k}) \mathbf{j}(\mathbf{k}) \right). \tag{4.7}$$

This is just the expected Poissonian distribution because the mean number of detectable photons is given by

$$\bar{n} = \langle 0_{\text{in}} | S_1(\Lambda) \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \mathbf{a}^*_{\text{in}}(\mathbf{k}) \mathbf{a}_{\text{in}}(\mathbf{k}) S_1^{-1}(\Lambda) | 0_{\text{in}} \rangle$$
$$= \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \mathbf{j}^*(\mathbf{k}) \mathbf{j}(\mathbf{k}). \tag{4.8}$$

So, what we have arrived at is a representation of the CCR's in a separable Hilbert space and a scattering operator S acting unitarily on this space. Neither S nor the representation depend on the experimental situation characterized by a finite resolution limit Λ. Actually this parameter partly determines the asymptotic observables and the form of those vectors that correspond to physical states. Equation (4.6) shows that the results are the same as obtained by formal computation in Fock space.

We want to conclude with a remark concerning perturbation series. From Eq. (4.7) one sees that ω_n, especially $\omega_1 = \bar{n}(\Lambda) e^{-\bar{n}(\Lambda)}$ tends to zero when Λ approaches zero. However, expanding the S operator (4.5) as

$$S_1(\Lambda) = 1 + i \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \{\mathbf{a}^*_{\text{in}}(\mathbf{k})\mathbf{j}(\mathbf{k}) + \mathbf{a}_{\text{in}}(\mathbf{k})\mathbf{j}^*(\mathbf{k})\} + \cdots,$$

the first order term for ω_1 is given by

$$\omega_1^{(1)} = \int_{|\mathbf{k}| \geqslant \Lambda} d\mu(\mathbf{k}) \mathbf{j}^*(\mathbf{k}) \mathbf{j}(\mathbf{k}) = \bar{n}(\Lambda)$$

which diverges with Λ going to zero. Thus our model, although mathematically consistent as well as physically reasonable if solved rigorously, is still infrared divergent in perturbation theory. Discussing infrared problems of quantum electrodynamics one should take into consideration that there might be divergences which exclusively originate in perturbation expansion.

ACKNOWLEDGMENT

My thanks are due to K. Kraus for many valuable comments, for his constant readiness to discuss the related problems, and for a critical reading of the manuscript.

[1] T. W. B. Kibble, J. Math. Phys. 9, 315 (1968).
[2] A. F. Verbeure and E. J. Verboven, Physica 37, 23 (1967).
[3] J. von Neumann, Comp. Math. 6, 1 (1938).
[4] L. Streit, Commun. Math. Phys. 4, 22 (1967).
[5] J. R. Klauder, J. McKenna, and E. J. Woods, J. Math. Phys. 7, 822 (1966).
[6] G. C. Hegerfeldt, J. Math. Phys. 10, 1681 (1969).
[7] G. Reents, J. Math. Phys. 11, 1961 (1970).
[8] G. C. Hegerfeldt and J. R. Klauder, Commun. Math. Phys. 16, 329 (1970).
[9] E. J. Woods, Commun. Math. Phys. 17, 1 (1970).
[10] G. Reents, "On infinite direct products of continuous unitary one-parameter groups" (to be published).
[11] H. Araki and E. J. Woods, J. Math. Phys. 4, 637 (1963).

On the approach to equilibrium in kinetic theory

Miroslav Grmela

Department of Mathematics, Carleton University, Ottawa, Canada
(Received 6 September 1972; final revised manuscript received 1 June 1973)

Definition, existence and properties of the equilibrium states, existence and uniqueness of the approach to the equilibrium states, the Onsager reciprocity relations, and existence of hydrodynamic stage in the approach to the equilibrium states are the problems which are discussed in this paper for dynamics defined by the family of kinetic equations of the Enskog–Vlasov type. A modification of the Chapman–Enskog method for quantitative calculations involving the kinetic equations considered is suggested.

I. INTRODUCTION

It is well understood that as the time evolution of dilute gases with short-range intermolecular forces proceeds the kinetic stage is reached, where the state of the gases at time t is adequately described by one function $f(\mathbf{r}, \mathbf{v}, t)$, where $\mathbf{r}$ and $\mathbf{v}$ are position and velocity vectors, respectively, and the time development of f is governed by the kinetic equation

$$\frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial t} = Rf(\mathbf{r}, \mathbf{v}, t), \tag{1}$$

where $R = \mathbf{v}(\partial/\partial \mathbf{r}) + R_B$, R_B is the Boltzmann collision operator. The subsequent time evolution of dilute gases can be deduced from Eq. (1). The particularly interesting problems are the following: (α) study of time independent solutions of Eq. (1) which can be identified with the thermodynamic equilibrium states, (β) problem of the existence and properties of the time approach to the equilibrium states studied in (α), (γ) problem of existence of hydrodynamic stage in the approach to the equilibrium states. Discussions of these problems on the basis of the Boltzmann equation have proved to be a source of inspiration for discussions of these problems and other dynamical models including Hamiltonian dynamics.

Our purpose is to study problems (α), (β), (γ) on the base of dynamics defined by the family of kinetic equations of the type (1) where R is an element of the class $\mathfrak{R}$ of kinetic operators defined in Sec. II [Eq. (2)]. The operator $-\mathbf{v}(\partial/\partial \mathbf{r}) + R_B$ is now only one element of $\mathfrak{R}$. An essentially new feature, with respect to the discussion based on the Boltzmann equation only, appears in the problem (α) (Sec. III). In contrast to a discussion based on the Boltzmann equation only, there arise here many possible equilibrium states which may be physically interpreted as indicating phase transitions (a recent review of the literature where this idea was laid down is given in Ref. 1). The single-phase, thermodynamically stable equilibrium states are defined and a Hilbert space, having a direct thermodynamical meaning, is attached to each of them. The time dependence governed by the corresponding linearized kinetic operators is studied in these spaces (Sec. IV). The main results are formulated in three theorems. The first theorem deals with existence and uniqueness problem, in the second theorem a symmetry relation for Green's function (an extension of Onsager's reciprocity relation in nonequilibrium thermodynamics) is proved; the third theorem (related to the Chapman–Enskog method and existence of hydrodynamic stage) discusses the spectrum of the Fourier transform of the linearized kinetic operators. If our discussion is reduced to the Boltzmann equation only, then the single-phase, thermodynamically stable equilibrium states are all possible equilibrium states; the Hilbert space attached to them is the $\mathfrak{L}_2$ space with the Maxwell distribution in velocities as a weight function in the inner product (introduced by Grad[3]) and the three theorems are known.[4,5]

II. KINETIC EQUATIONS OF THE ENSKOG-VLASOV TYPE

The class $\mathfrak{R}$ of operator R in Eq. (1) consists of linear combinations of four operators R_i, $i = 1, 2, 3, 4$, defined below with real coefficients $c_i \geq 0$ for $i = 1, 3, 4$ and $c_2 > 0$, i.e. $\mathfrak{R} \ni R_{\{c_i\}} = c_1R_1 + c_2R_2 + c_3R_3 + c_4R_4$. The basic operators R_i are defined as follows: (the summation convention $\mathbf{a} \cdot \mathbf{b} = a_\alpha b_\alpha = a_1b_1 + a_2b_2 + c_3b_3$, where $\mathbf{a}, \mathbf{b}$ are three-dimensional vectors, is used)

$$R_1 f = -v_\alpha \frac{\partial f}{\partial r_\alpha},$$

$$R_2 f = \eta\{\mathbf{r}, t\} R_B f = \eta\{\mathbf{r}, t\}\sigma^2 \times \int d^2\boldsymbol{\kappa} \int_{(g_\alpha \kappa_\alpha) > 0} d^3\mathbf{v}_1 (g_\alpha \kappa_\alpha)(f'(\mathbf{r}) f_1'(\mathbf{r}) - f(\mathbf{r}) f_1(\mathbf{r})), \tag{2}$$

where $\eta[\mathbf{r}, t] \equiv \eta\{n(\mathbf{r}, t)\}$ is a functional of $n(\mathbf{r}, t) = \int d^3\mathbf{v} f(\mathbf{r}, \mathbf{v}, t)$ (its physical meaning is explained below), σ is the diameter of the hard-core repulsive potential between two particles in the system considered, $\mathbf{g} = \mathbf{v}_1 - \mathbf{v}$, $\mathbf{v}' = \mathbf{v} + \boldsymbol{\kappa}(g_\alpha \kappa_\alpha)$, $\mathbf{v}_1' = \mathbf{v}_1 - \boldsymbol{\kappa}(g_\alpha \kappa_\alpha)$, $\boldsymbol{\kappa}$ is a unit vector directed from the center of the sphere with velocity $\mathbf{v}$ to the center of the sphere with $\mathbf{v}'$, the standard abbreviation $f_1'(\mathbf{r}) \equiv f(\mathbf{r}, \mathbf{v}_1', t) \cdots$ etc. is used,

$$R_3 f = \sigma^3 \int_{(g_\alpha \kappa_\alpha) > 0} d^3\mathbf{v}_1 (g_\gamma \kappa_\gamma) \times \kappa_\alpha \left\{ \eta[\mathbf{r}, t] \left(f'(\mathbf{r}) \frac{\partial f_1'(\mathbf{r})}{\partial r_\alpha} + f(\mathbf{r}) \frac{\partial f_1(\mathbf{r})}{\partial r_\alpha} \right) + \frac{1}{2} \frac{\partial \eta[\mathbf{r}, t]}{\partial r_\alpha} (f'(\mathbf{r}) f_1'(\mathbf{r}) + f(\mathbf{r}) f_1(\mathbf{r})) \right\}, \tag{2'}$$

$$R_4 f = \frac{1}{m} \frac{\partial f(\mathbf{r})}{\partial v_\alpha} \int d^3\mathbf{r}_1 \int d^3\mathbf{v}_1 \frac{\partial V(|\mathbf{r} - \mathbf{r}_1|)}{\partial r_\alpha} f_1(\mathbf{r}_1),$$

where m is the mass of one particle, $V(|\mathbf{r} - \mathbf{r}_1|)$ is the long-range attractive part of the two particle potential function. The functions η and V are arbitrary except we impose the mathematical requirement that all expressions containing these functions have a sense; some further requirements will come later in the text (Lemma 1).

Let us recall the physical meaning of some of the kinetic equations (1) with $R \in \mathfrak{R}$. The operator R_1 is the Knudsen operator. $R_1 f$ gives the rate of change of f due to the fact that the particles have a finite velocity, and consequently, change their position with time. The operator R_2 is the Boltzmann collision operator and

$R_2 f$ gives the rate of change of f due to the fact that the particles change their velocities with time because of the hard-core potential between them (binary collisions). The operator R_3 is the part of the Enskog operator proportional to σ^3. $R_3 f$ gives a correction to $R_2 f$ due to the fact that the hard-core potential has a finite range σ. The function η occurs in order to include approximately the two particle space correlations ($\eta \equiv 1$ if there is no such correlation). The operator R_4 is the Vlasov operator. It represents a mean field type influence of the long-range potential between particles on the rate of change of f.

Our class $\mathfrak{R}$ of operators in Eq. (1) does not include all kinetic operators known now, but it hopefully includes the physically most important ones. The physical significance of various kinetic operators is measured by comparing some qualitative and quantitative properties of solutions to Eq. (1) with results of observation. In order to know which kinetic operator in a given situation is the physically significant one, it is first necessary to know how to get the qualitative and quantitative properties of solutions of the kinetic equations and which of these properties are related to particular observations. An example of a physical system whose kinetic stage seems to be reasonably described by Eq. (1) with $R \in \mathfrak{R}$ is a van der Walls gas.[2,1]

In order to eliminate problems related to boundary conditions, we shall assume the periodic boundary conditions on the boundary of a finite volume $\Omega = \{-a \leq r_i \leq a,\ i = 1,2,3\}$, where a is a sufficiently large positive number. Equation (1) becomes mathematically meaningful if a space of functions serving as the domain of $R \in \mathfrak{R}$ (denoted by $\mathfrak{D}$) is specified by giving its, topological structure. In principle, the structure of $\mathfrak{D}$ together with the kinetic equation should come as a result of its derivation from Hamiltonian dynamics. We want to start our study of the approach to equilibrium with the kinetic equation (1), therefore a structure of $\mathfrak{D}$ must be discussed inside kinetic theory. We shall show that the established correspondence between the study of the time-independent solutions to Eq. (1) which are, moreover, invariant with respect to the transformation $\mathbf{v} \to -\mathbf{v}$ (called the equilibrium states) and thermodynamics allows us to make a physically reasonable classification of equilibrium states leaving the problem of the structure of $\mathfrak{D}$ undecided. This discussion of equilibrium states is also used to propose some local properties of $\mathfrak{D}$ at the equilibrium states needed for our study of the approach to the equilibrium states.

III. THE EQUILIBRIUM STATES

We shall define the equilibrium states as the time independent solutions of Eq. (1) that are invariant with respect to the transformation $\mathbf{v} \to -\mathbf{v}$ denoted hereafter as $T_{\mathbf{v}}$. It is easy to see that all $R \in \mathfrak{R}$ with $c_2 = 0$ change sign if the transformation $T_{\mathbf{v}}$ is applied, the operators $R \in \mathfrak{R}$ with $c_1 = c_3 = c_4 = 0$ remain unchanged. The following notation is used:

$$\mathfrak{R}^{(+)} = \{R \in \mathfrak{R}; c_1 = c_3 = c_4 = 0\}\ \mathfrak{R}^{(-)} = \{R \in \mathfrak{R}; c_2 = 0\}.$$

Evidently, we have $\mathfrak{R} = \mathfrak{R}^{(+)} + \mathfrak{R}^{(-)}$. Elements of $\mathfrak{R}^{(+)}$ are denoted by $R^{(+)}$, elements of $\mathfrak{R}^{(-)}$ are denoted by $R^{(-)}$. A direct consequence of the definition of the equilibrium states are the following two equations determining them:

$$R^{(+)} f = 0, \tag{3}$$

$$R^{(-)} f = 0. \tag{4}$$

Let us introduce a set $\mathfrak{F} = \{f \in \mathfrak{D}; R^{(+)} f = 0, T_{\mathbf{v}} f = f\}$. Elements of $\mathfrak{F}$ are denoted by f_e. Since Eq. (3) is equivalent to $R_B f = 0$ (we assume that $\eta \neq 0$), the Boltzmann classical result gives

$$f_e = n(\mathbf{r}) \exp[\beta(\mathbf{r}) v^2], \tag{5'}$$

where $n(\mathbf{r})$ and $\beta(\mathbf{r})$ are arbitrary functions of $\mathbf{r}$ such that $f_e \in \mathfrak{D}$. Inserting f_e given by Eq. (5') into Eq. (4), an equation for $n(\mathbf{r})$ and $\beta(\mathbf{r})$ is obtained. Because $n(\mathbf{r})$ and $\beta(\mathbf{r})$ are independent of v, it is easy to see that $\beta(\mathbf{r})$ must be a constant independent of $\mathbf{r}$. We shall use hereafter

$$f_e = n(\mathbf{r}) M_\beta(v), \tag{5}$$

where $M_\beta(v) = (m/2\pi)^{3/2}\beta^{3/2} \exp(-\frac{1}{2}m\beta v^2)(\int d^3\mathbf{v} f_e = n(\mathbf{r}))$ and β is a positive real number. If we now insert f_e given by Eq. (5) into Eq. (4), an equation for $n(\mathbf{r})$ is obtained:

$$R^{(-)} f_e = n(\mathbf{r})\Big(- v_\alpha \frac{\partial}{\partial r_\alpha} M_\beta(v)[c_1 \ln n + c_4 \beta G_{1,\mathbf{r}} n + c_3 e(\eta(n) n + H(n))]\Big) = 0, \tag{6}$$

where $e = \frac{2}{3}\pi\sigma^3$,

$$G_{1,\mathbf{r}} n = \int_\Omega d^3\mathbf{r} V(|\mathbf{r} - \mathbf{r}_1|) n(\mathbf{r}_1), \tag{7}$$

$$H(n) = \int \eta(n) dn.$$

Equation (6) can be written also in another form as the condition for an extremum of a functional

$\ln \Xi_{eq}(n(\mathbf{r}); \alpha_{eq}, \beta, a, \{c\})$, where $\{c\} \equiv c_1, c_2, c_3, c_4$, i.e.,

$$\frac{\delta \ln \Xi_{eq}}{\delta n} = 0, \tag{6'}$$

where

$$\ln \Xi_{eq}(n(\mathbf{r}); \alpha_{eq}, \beta, a\{c\}) = - c_1 \int_\Omega d^3\mathbf{r} n(\mathbf{r}) \ln n(\mathbf{r}) - c_3 e \int_\Omega d^3\mathbf{r} n(\mathbf{r}) H(n) - c_4 \frac{\beta}{2} \int\int_{\Omega\times\Omega} d^3\mathbf{r} d^3\mathbf{r}' V(|\mathbf{r} - \mathbf{r}'|) n(\mathbf{r}) n(\mathbf{r}') + c_1 \alpha_{eq} \int_\Omega d^3\mathbf{r} n(\mathbf{r}), \tag{8}$$

α_{eq} is an arbitrary constant.

The function $n(\mathbf{r})$ satisfying Eq. (6) will be denoted by $n_{eq}(\mathbf{r})$, the function $f_{eq}(\mathbf{r}, \mathbf{v}) = n_{eq}(\mathbf{r}) M_\beta(v)$ is called an equilibrium state. The following notation is introduced: $\mathfrak{F}_{eq} = \{f \in \mathfrak{D}; T_{\mathbf{v}} f = f,\ R^{(+)} f = 0,\ R^{(-)} f = 0\}$. It is evident that $\mathfrak{F}_{eq} \subset \mathfrak{F}$. The equilibrium states f_{eq} are elements of $\mathfrak{F}_{eq}$. A finer classification of states inside $\mathfrak{F}_{eq}$ will be needed.

The functions $f_{eq} \in \mathfrak{F}_{eq}$ satisfying, moreover,

$$\int\int_{\Omega\times\Omega} d^3\mathbf{r} d^3\mathbf{r}' \frac{\delta^2 \ln \Xi_{eq}}{\delta n(\mathbf{r})\, \delta n(\mathbf{r}')} \chi(\mathbf{r}) \chi(\mathbf{r}') < 0, \tag{9}$$

where $\chi(\mathbf{r})$ is an arbitrary function such that the integral in (9) exists, will be denoted by $f_{eqs}(\mathbf{r}, \mathbf{v}) = n_{eqs}(\mathbf{r}) M_\beta(v)$ and called the thermodynamically stable equilibrium states. We define $\mathfrak{F}_{eqs} = \{f \in \mathfrak{F}_{eq}; f$ satisfies Eq. (9)$\}$. If $n_{eqs}(\mathbf{r})$ is a constant independent of $\mathbf{r}$

$$n_{eqs}(\mathbf{r}) = n_0, \tag{10}$$

then $f_0(\mathbf{r},\mathbf{v}) = n_0 M_\beta(v)$ are called single-phase thermodynamically stable equilibrium states. All such f_0 form a set $\mathfrak{F}_0 \subset \mathfrak{F}_{eqs} \subset \mathfrak{F}_{eq} \subset \mathfrak{F} \subset \mathfrak{D}$. Since $\ln\Xi_{eq}$ depends on $\alpha_{eq}, \beta, a, \{c\}$ and on the two free functions η, V in the definition of $\mathcal{R}$, the conditions defining the classification $\mathfrak{F}_0 \subset \mathfrak{F}_{eqs} \subset \mathfrak{F}_{eq} \subset \mathfrak{F} \subset \mathfrak{D}$ represent relations among these quantities which are in general different for different $\mathfrak{D}$. The relationship between equilibrium states in the sense of Eqs. (3)–(10) and equilibrium states studied in thermodynamics is based on the following interpretation of $\ln\Xi_{eq}$:

$$\ln\Xi_{eq}(n_{eq}(\mathbf{r}); \alpha_{eq}, \beta, a, \{c\}) = \ln\Xi_{tk}(\alpha_{eq}, \beta, \Omega) = -\beta U[T, \mu, \Omega], \qquad (11)$$

where U is the grand canonical thermodynamic potential[6] ($dU = -SdT - Pd\,\Omega - Nd\,\mu$), $T = (k_B\beta)^{-1}$, k_B is the Boltzmann constant, $\mu = \alpha_{eq}/\beta$.

IV. APPROACH TO THE SINGLE-PHASE THERMO DYNAMICALLY STABLE EQUILIBRIUM STATES

Only local dynamical properties of single phase thermodynamically stable equilibrium states $f_0 \in \mathfrak{F}_0$ will be studied. An appropriate Hilbert space $\mathcal{K}_\alpha = \{\varphi(\mathbf{r},\mathbf{v})$ are Lebesgue measurable, $(\varphi_1\varphi)_\alpha < \infty\}$ which will be specified later is attached to each $f_0 \in \mathfrak{F}_0$. The Hilbert space $\mathcal{K}_\alpha$ is of course in general different for different f_0, i.e., it depends on $\alpha_{eq}, \beta, \Omega, \{c\}, \eta(n), V(r)$. The time development of φ is assumed to be governed by a linear part of R denoted by P. In other words, $P\varphi$ is obtained as the linearization of Rf around f_0. The linear operator P depends on $\alpha_{eq}, \beta, a, \{c\}$, $\eta(n)$, $V(r)$. The class of all operators P corresponding to all $R \in \mathcal{R}$ and all $f_0 \in \mathfrak{F}_0$ is denoted as $\mathcal{R}_0$.

Simple calculations give a general form of $P \in \mathcal{R}_0$ in the linear equation

$$\frac{\partial\varphi}{\partial t} = P\varphi, \qquad (12)$$

$$P\varphi = c_1 D_1\varphi + c_2\eta_0 R_{B,l}\varphi + c_3[e\eta_0(D_1G_2 - G_2D_1)\varphi + \eta_1 e D_1G_2 + \eta_0(\sigma^3/4)D_2]\varphi + c_4\beta D_1G_1G_2\varphi, \qquad (13)$$

where $G_1\varphi = G_{1,\mathbf{r}} I_{\mathbf{v}}\varphi = \int_\Omega d^3\mathbf{r}_1 V(|\mathbf{r}-\mathbf{r}_1|)\varphi(\mathbf{r}_1,\mathbf{v})$, $I_{\mathbf{v}}$ is the identity transformation with respect to the dependence of φ on $\mathbf{v}$,

$$D_1\varphi = -v_\alpha \frac{\partial}{\partial r_\alpha}\varphi,$$

$$G_2\varphi = G_{2,\mathbf{v}} I_{\mathbf{r}}\varphi = \int d^3\mathbf{v}_1 f_0(v_1)\varphi(\mathbf{r},\mathbf{v}_1),$$

$$D_2\varphi = \int d^2\kappa \int d^3\mathbf{v}_1 f_0(v_1)(v_1 - v)_\gamma \kappa_\gamma \kappa_\alpha \frac{\partial}{\partial r_\alpha}(\varphi(\mathbf{r},\mathbf{v}_1') - \varphi(\mathbf{r},\mathbf{v}_1)),$$

$R_{B,l}$ is the linearized Boltzmann operator,[3] $R_{B,l} = K\varphi - \nu(v)\varphi$, the explicit form of the integral operator K and the function $\nu(v)$ are given for example in Ref. 3, $\eta_0 = \eta(n_0)$, $\eta_1 = n_0(\delta\eta(n)/\delta n)_{n_0}$. We shall assume that $0 < \eta_0 < \infty$, $\eta_1 < \infty$. The domain of $P \in \mathcal{R}_0$ is denoted as $\mathfrak{D}_{0,\alpha} = \{\varphi \in \mathcal{K}_\alpha; P\varphi \in \mathcal{K}_\alpha, \varphi$ is absolutely continuous on $\mathbf{r} \in \Omega$ and $-\infty \le v_i \le +\infty$, $i = 1,2,3$, and $\varphi(-a, r_2, r_3, \mathbf{v}) = \varphi(a, r_2, r_3, \mathbf{v})$, $\varphi(r_1, -a, r_3, \mathbf{v}) = \varphi(r_1, r_2, -a, \mathbf{v}) = \varphi(r_1, r_2, a, \mathbf{v})\}$. The finite volume Ω and the periodic boundary conditions in the definition of $\mathfrak{D}_{0,\alpha}$ can be replaced by the whole coordinate space $\mathbb{R}^3$ together with the condition that the φ tends to zero sufficiently rapidly as $|\mathbf{r}| \to \infty$ without effecting the discussion which will follow. In contrast, any global discussion dealing with whole $\mathfrak{D}$, including the discussion of equilibrium states in Sec. III, is in general sensitive to this change.

Lemma 1: Let $\mathcal{K}_\alpha = \mathcal{K}_1 = \mathcal{L}_2$ [i.e., $\psi, \varphi)_\alpha = (\psi,\varphi)_1 = \int_\Omega d^3\mathbf{r} \int d^3\mathbf{v}\,\psi(\mathbf{r},\mathbf{v})\varphi(\mathbf{r},\mathbf{v})$]. Then

(a) G_1 is a bounded self-adjoint operator, i.e., $G_1 = G_1^{\dagger_1}$ ($\dagger_1$ means the adjoint with respect to the inner product $(,)_1$;

(b) G_2 is a bounded operator and $f_0G_2 = (f_0G_2)^{\dagger_1}$;

(c) the integral operator K in $R_{B,l}$ is a bounded and compact operator; moreover, $f_0K = (f_0K)^{\dagger_1}$, $\nu(v) \sim v$ for $v \to \infty$ and $(\varphi, f_0R_{B,l}\varphi)_1 \le 0$, where the equality holds only for $\varphi = 1, \mathbf{v}, v^2$;

(d) the operator D_1 is a closed and $D_1 = -(D_2)^{\dagger_1}$ (domains of D_1 and $D_1^{\dagger_1}$ are identical);

(e) similarly D_2 is a closed operator and $f_0D_2 = -(f_0D_2)^{\dagger_1}$;

(f) $\int d^3\mathbf{r}\int d^3\mathbf{v}\, f_0 P\varphi = 0$ for all $\varphi \in \mathfrak{D}_{0,1}$,

(g) $\int d^3\mathbf{v}\, f_0 \mathcal{P}_{c_1}\varphi = 0$, where $\mathcal{P}_{c_1} = \{P \in \mathcal{R}_0; c_1 = 0\}$;

(h) $\mathfrak{D}_{0,1}$ is dense in $\mathcal{K}_1$.

Proof: (a), (b), (d) and (e) follow from the definitions of the operators G_1, G_2 [Eqs. (7), (12)] and D_1 and D_2 [Eq. (12)] and from the standard considerations (e.g., Ref. 7). The boundedness of G_1 represents a limitation (physically very natural) on functions $V(|\mathbf{r}-\mathbf{r}_1|)$ we shall consider. (c) is a summary of the well-known results about the linear Boltzmann operator,[3] (f) is a consequence of the property $\int d^3\mathbf{r}\int d^3\mathbf{v}\,Rf = 0$ which can be obtained easily by using the definitions of $R \in \mathcal{R}$ [Eq. (12)]; (g) is a consequence of the property $\int d^3\mathbf{v} R_{c_1} f = 0$ where $R_{c_1} = \{R \in \mathcal{R}; c_1 = 0\}$ which can be again easily obtained by using the definitions $R \in \mathcal{R}$ (Eq. (2)). (h) follows from standard considerations (e.g., Ref. 7).

We shall now find an appropriate $\mathcal{K}_\alpha$ denoted as $\mathcal{K}_2$. A natural extension of the function $\ln\Xi_{eq}(n(\mathbf{r}); \alpha_{eq}, \beta, a, \{c\})$ which has been introduced in Eq. (8) is the function $\ln\Xi(f(\mathbf{r},\mathbf{v},t); \alpha, \beta, a, \{c\})$ defined as follows:

$$\begin{aligned}\ln\Xi(f;\alpha,\beta,a,\{c\}) &= -c_1\int_\Omega d^3\mathbf{r}\int d^3\mathbf{v} f(\mathbf{r},\mathbf{v},t)\ln f(\mathbf{r},\mathbf{v},t) \\ &\quad - c_3 e\int_\Omega d^3\mathbf{r}\int d^3\mathbf{v} f(\mathbf{r},\mathbf{v},t)H(n) \\ &\quad + c_1\alpha\int_\Omega d^3\mathbf{v}\int d^3\mathbf{v} f(\mathbf{r},\mathbf{v},t) \\ &\quad - c_1\beta\int d^3\mathbf{r}\int d^3\mathbf{v}\,\tfrac{1}{2}mv^2 f(\mathbf{r},\mathbf{v},t) \\ &\quad - c_4\tfrac{1}{2}\beta\int\int_{\Omega\times\Omega} d^3\mathbf{r}d^3\mathbf{r}_1\int\int d^3\mathbf{v}d^3\mathbf{v}_1\, V(|\mathbf{r}-\mathbf{r}_1|) \\ &\quad \times f(\mathbf{r},\mathbf{v},t)f(\mathbf{r}_1,\mathbf{v}_1,t). \end{aligned} \qquad (14)$$

These are the properties of $\ln\Xi$:

$$\ln\Xi(f_e;\alpha,\beta,a,\{c\}) = \ln\Xi_{eq}(n(\mathbf{r});\alpha_{eq},\beta,a,\{c\}),$$

where $\alpha_{eq} = \alpha - \frac{3}{2}\ln(\beta[m/2\pi])$,

$$\frac{\delta\ln\Xi_{eq}}{\delta n} = \frac{\delta\ln\Xi}{\delta f}\Big|_{f_e}. \qquad (15)$$

The functional Taylor expansion of $\ln\Xi$ gives

$$\ln\Xi(f;\alpha,\beta,a,\{c\}) = \ln\Xi_{eq}(n_0;\alpha,\beta,a,\{c\})$$

$$+ \int_\Omega d^3\mathbf{r} \int d^3\mathbf{v} \frac{\delta \ln \Xi}{\delta f(\mathbf{r},\mathbf{v},t)}\Big|_{f_0} f_0(v)\varphi(\mathbf{r},\mathbf{v},t)$$

$$+ \tfrac{1}{2}\int\int_{\Omega\times\Omega} d^3\mathbf{r}d^3\mathbf{r}_1 \int\int d^3\mathbf{v}d^3v_1 \frac{\delta^2 \ln \Xi}{\delta f(\mathbf{r},\mathbf{v},t)\,\delta f(\mathbf{r}_1,\mathbf{v}_1,t)}\Big|_{f_0}$$

$$\times f_0(v)f_0(v_1)\varphi(\mathbf{r},\mathbf{v},t)\varphi(\mathbf{r}_1,\mathbf{v}_1,t), \tag{16}$$

where the terms $\sim \varphi^k, k \geq 3$ were neglected (we are discussing the linearized kinetic equation only). The second term on the rhs of Eq. (16) equals to zero. It follows from the definition of the equilibrium states in Eq. (6′) and from Eq. (15). The third term on the rhs of Eq. (16) can be written as

$$-\tfrac{1}{2}(\varphi, A\varphi)_1, \tag{17}$$

where $A = f_0\tilde{A}$,

$$\tilde{A} = -c_1 - c_4\beta G_1 G_2 - c_3 e(\eta_1 G_2 + 2\eta_0 G_2).$$

So that Eq. (9) defining the thermodynamic stability can now be written as

$$(\chi(\mathbf{r}), A_{eq}\chi(\mathbf{r}))_1 > 0, \tag{9′}$$

where $A_{eq} = c_1 + c_4\beta G_{1,\mathbf{r}} - c_3 e(\eta_1 + 2\eta_0)$. Inserting $\chi(\mathbf{r}) = \int f_0\varphi d^3\mathbf{v}$, one obtains

$$(\varphi, A\varphi)_1 \geq 0, \tag{18}$$

where equality holds only for $\varphi = 0$. Thus if $\alpha, \beta, a, \{c\}$, n, v are such that f_0 is a single phase thermodynamically stable equilibrium state then the linear operator A is a bounded, self-adjoint and positive operator defined everywhere on $\mathfrak{L}_2$. It follows from the theory of square roots of linear operators[8,9] that there exists a linear operator $A^{1/2}$ which is bounded, self-adjoint and defined everywhere on $\mathfrak{L}_2$ such that $(A^{1/2})^2 = A$. Hence,

$$(\psi, A\varphi)_1 = (\psi, (A^{1/2})^2\varphi)_1 = (A^{1/2}\psi, A^{1/2}\varphi)_1$$
$$= (\psi_{1/2}, A^{1/2}\varphi)_1 = (\psi_{1/2}, \varphi_{1/2})_1 \overset{\text{def.}}{=} (\psi, \varphi)_2. \tag{19}$$

It follows from Eqs. (16) and (19) that

$$\ln \Xi = \ln \Xi_{eq} - \tfrac{1}{2}(\varphi, \varphi)_2 \tag{16′}$$

and by using Eq. (12) we have

$$\frac{\partial \ln \Xi}{\partial t} = -(\varphi, AP\varphi)_1 = -(\varphi, P\varphi)_2. \tag{20}$$

It is also clear from (19) that the following are equivalent: The study of AP in $\mathfrak{K}_1$, the study of P in $\mathfrak{K}_2$ and the study of $P_{1/2} = A^{1/2}P(A^{1/2})^{-1}$ in $(\mathfrak{K}_1)_{1/2}$, where $(\mathfrak{K}_1)_{1/2}$ indicates that the domain consists of the functions $\varphi_{1/2} = A^{1/2}\varphi$. In accordance with the usual physical terminology, we can call the operator $A^{1/2}$ the renormalization operator.

Lemma 2: (a) The operator AP is densely defined and closed in $\mathfrak{K}_1$;

(b) the domains of $AP^{(\pm)} = (AP^{(\mp)})^{\dagger_1}$ are identical;

$$\text{(c)}\quad AP^{(+)} = (AP^{(+)})^{\dagger_1}, \quad (\varphi, AP^{(+)}\varphi)_1 \leq 0, \tag{21}$$

where the equality holds only for $\varphi = 1, \mathbf{v}, v^2$;

$$\text{(d)}\quad AP^{(-)} = -(AP^{(-)})^{\dagger_1}, \quad (\varphi, AP^{(-)}\varphi)_1 = 0; \tag{22}$$

$$\text{(e)}\quad (AP)^{\dagger_1} = T_{\mathbf{v}}(AP). \tag{23}$$

Proof: (a) and (b) follows from Lemma 1 and Eq. (17). In order to prove (c) and (d), we shall write $(\psi, AP\varphi)_1$ explicitly by using Eqs. (12) and (17),

$$(\psi, AP\varphi)_1 = (f_0[c_1 + c_4\beta G_1 G_2 + c_3 e(\eta_1 G_2 + 2\eta_0 G_2)]\psi,$$
$$\{c_1 D_1 + c_2\eta_0 R_{B,l} + c_3[e\eta_0(D_1 G_2 - G_2 D_1)$$
$$+ \eta_1 e D_1 G_2 + \eta_0(\sigma^3/4)D_2] + c_4\beta D_1 G_1 G_2\}\varphi)_1. \tag{24}$$

The only terms giving a nonzero contribution are those proportional to $c_1^2, c_1c_2, c_1c_4, c_1c_3$. All others give zero as a consequence of the Lemma 1(g). After some calculations and by using explicit definition of P and the Lemma 1, one obtains

$$(\psi, AP\varphi)_1 = (\psi, \{c_1^2 f_0 D_1 + c_1c_2 f_0\eta_0 R_{B,l}$$
$$+ c_1c_4\beta[f_0 D_1 G_1 G_2 - (f_0 D_1 G_1 G_2)^{\dagger_1}]$$
$$+ c_1c_3\eta_1[f_0 D_1 G_2 - (f_0 D_1 G_2)^{\dagger_1}]$$
$$+ c_1c_3 e\eta_0[f_0 D_1 G_2 - (f_0 D_1 G_2)^{\dagger_1}]$$
$$+ c_1c_3\eta_0(\sigma^3/4) f_0 D_2\}\varphi)_1; \tag{25}$$

(c) and (d) follow now easily from (25). (e) is a direct consequence of the definition of $P^{(+)}$ and $P^{(-)}$ and (c), (d) of this lemma.

The next three sections discuss mathematical and physical consequence of the Lemma 2.

V. EXISTENCE AND UNIQUENESS OF THE APPROACH TO THE SINGLE-PHASE THERMODYNAMICALLY STABLE EQUILIBRIUM STATES

Theorem 1: There exists one and only one function φ_t defined for all $t \geq t_0$ (t_0 is fixed) with properties:

(1) φ_t is continuously differentiable with respect to t and $\varphi_t \in \mathfrak{D}_{0,2}$,

(2) Eq. (12) is satisfied and $\varphi_t = U(t, t_0)\varphi_{t_0}$ where $U(t, t_0)$ is a contraction semigroup of class C^0,

(3) the strong limit of φ_t for $t \to t_0$ equals to φ_{t_0} equals to φ_{t_0}.

This theorem is an immediate consequence of Lemma 2 and the Hille–Yoshida theory of semigroups.[9,10] The fact that the Hille–Yoshida theory is a natural tool in kinetic theory has been recognized in transport theory[11] and in the theory of the linearized Boltzmann equation.[8]

VI. ONSAGER RECIPROCITY RELATIONS IN KINETIC THEORY

Interesting physical consequences of results in Sec. IV can be derived from the integral representation of the operator $U(t, t_0)$. Let

$$\varphi(\mathbf{r},\mathbf{v},t) = U(t,t_0)\varphi(\mathbf{r},\mathbf{v},t_0)$$
$$= \int_\Omega d^3\mathbf{r}_0 \int d^3\mathbf{v}_0 G(\mathbf{r},\mathbf{v},t;\mathbf{r}_0,\mathbf{v}_0,t_0)$$
$$\times \varphi(\mathbf{r}_0,\mathbf{v}_0,t_0) \tag{26}$$

be the integral representation of the operator $U(t, t_0)$. The function G in Eq. (26) is called a kernel of the integral representation or also Green's function of Eq. (12).

Since the operator P defining $U(t,t_0)$ is independent of time, we have

$$G(\mathbf{r},\mathbf{v},t;\mathbf{r}_0,\mathbf{v}_0,t_0)=G(\mathbf{r},\mathbf{v};\mathbf{r}_0,\mathbf{v}_0,t-t_0). \tag{27}$$

It follows from Eqs. (19) and (19′) that

$$\frac{\partial \ln\Xi}{\partial t}\Big|_{t_0} = \int\int_{\Omega\times\Omega} d^3\mathbf{r}d^3\mathbf{r}_1 \int\int d^3\mathbf{v}d^3\mathbf{v}_1 \frac{\delta^2 \ln\Xi}{\delta f(\mathbf{r},\mathbf{v},t)\,\delta f(\mathbf{r}_1,\mathbf{v}_1,t)}\Big|_{f_0} \times f_0(v)f_0(v_1)\frac{\partial\varphi(r,v,t)}{\partial t}\Big|_{t_0}\varphi(\mathbf{r}_1,\mathbf{v}_1,t_0), \tag{28}$$

or, introducing a shorter notation,

$$\frac{\partial \ln\Xi}{\partial t}\Big|_{t_0} = \int\int_{\Omega\times\Omega} d^3\mathbf{r}d^3\mathbf{r}_1 \int\int d^3\mathbf{v}d^3\mathbf{v}_1 A(\mathbf{r},\mathbf{v};\mathbf{r}_1,\mathbf{v}_1) \times \frac{\partial\varphi(r,v,t)}{\partial t}\Big|_{t_0}\varphi(\mathbf{r}_1,\mathbf{v}_1,t_0), \tag{28′}$$

where

$$A(\mathbf{r},\mathbf{v};\mathbf{r}_1,\mathbf{v}_1)=A(\mathbf{r}_1,\mathbf{v}_1;\mathbf{r},\mathbf{v})=f_0(v)f_0(v_1) \times \delta^2\ln\Xi/\delta f(\mathbf{r},\mathbf{v},t)\,\delta f(\mathbf{r}_1,\mathbf{v}_1,t)|_{f_0}.$$

Using Eq. (26), we can write

$$\frac{\partial\varphi}{\partial t}\Big|_{t_0} = \int_\Omega d^3\mathbf{r}_0 \int d^3\mathbf{v}_0 \frac{\partial}{\partial t} G(\mathbf{r},\mathbf{v};r_0,v_0,t-t_0)|_{t_0} \times \varphi(\mathbf{r}_0,\mathbf{v}_0,t_0). \tag{29}$$

Inserting (29) into (28′), we have

$$\frac{\partial \ln\Xi}{\partial t}\Big|_{t_0} = \int\int_{\Omega\times\Omega} d^3\mathbf{r}d^3\mathbf{r}_1 \int\int d^3\mathbf{v}d^3\mathbf{v}_1 L(\mathbf{r},\mathbf{v};\mathbf{r}_1,\mathbf{v}_1) \times \varphi(\mathbf{r},\mathbf{v},t_0)\,\varphi(\mathbf{r}_1,\mathbf{v}_1,t_0), \tag{30}$$

where

$$L(\mathbf{r},\mathbf{v};\mathbf{r}_1,\mathbf{v}_1) = \int_\Omega d^3\mathbf{r}_0 \int d^3\mathbf{r}_0 A(\mathbf{r}_0,\mathbf{v}_0;\mathbf{r}_1,\mathbf{v}_1) \times \frac{\partial}{\partial t} G(\mathbf{r}_0,\mathbf{v}_0;\mathbf{r},\mathbf{v},t-t_0)\Big|_{t_0}. \tag{31}$$

Theorem 2: The function L defined in (31) satisfies the following relation called the Onsager reciprocity relation in kinetic theory:

$$L(\mathbf{r},\mathbf{v};\mathbf{r}_1,\mathbf{v}_1)=L(\mathbf{r}_1,-\mathbf{v}_1;\mathbf{r},-\mathbf{v}). \tag{32}$$

Proof: Let us write Eq. (12) as

$$\Pi\varphi(\mathbf{r},\mathbf{v},t)=0, \tag{12′}$$

where

$$\Pi=\frac{\partial}{\partial t}-P$$

is a linear operator defined on the Hilbert space $\tilde{\mathcal{H}}_2$,

$$\tilde{\mathcal{H}}_2=\{\varphi(\mathbf{r},\mathbf{v},t) \text{ is Lebesgue measurable; } ((\varphi,\varphi))_2<\infty\}, \tag{33}$$

where $((\psi,\varphi))_2=\int_{-\infty}^{+\infty} dt(\psi,\varphi)_2$ and $\varphi(\mathbf{r},\mathbf{v},t)$ is defined to be zero for $t<t_0$. It follows from the Lemma 2(c), (d) that the domain of Π, denoted by $\tilde{\mathcal{D}}_{0,2}$, is indeed a natural extension of $\mathcal{D}_{0,2}$ in $\mathcal{H}_2$. Using Eq. (23), we obtain

$$(A\pi)^\dagger = A\left(-\frac{\partial}{\partial t}\right) - (AP)^{\dagger_1} = T_v T_t A\pi, \tag{34}$$

where T_t is the transformation $t\to -t$.

From the definition of $U^{\dagger_1}(t,t_0)$ [i.e., $(\varphi(t_0),AU(t,t_0)\varphi(t_0))_1=(AU(t,t_0)^{\dagger_1}\varphi(t_0),\varphi(t_0))_1$] and by using the integral representation of $U(t,t_0)$, we have

$$\int_\Omega d^3\mathbf{r}_1'\int d^3\mathbf{v}_1' A(\mathbf{r}_1',\mathbf{v}_1';\mathbf{r}_1,\mathbf{v}_1)G^\dagger(\mathbf{r}_1',\mathbf{v}_1',t_1;\mathbf{r},\mathbf{v},t) = \int_\Omega d^3\mathbf{r}_1'\int d^3\mathbf{v}_1' A(\mathbf{r}_1',\mathbf{v}_1';\mathbf{r},\mathbf{v})G(\mathbf{r}_1',\mathbf{v}_1',t;\mathbf{r}_1,\mathbf{v}_1,t_1), \tag{35}$$

where $G^\dagger$ is a kernel of the integral representation of $U^{\dagger 1}(t,t_0)$. The relation (35) becomes the familiar relation[12] if written with renormalized quantities

$$G_{1/2}^{\dagger_{1/2}}(\mathbf{r}_1,\mathbf{v}_1,t_1;\mathbf{r},\mathbf{v},t)=G_{1/2}(\mathbf{r},\mathbf{v},t;\mathbf{r}_1,\mathbf{v}_1,t_1). \tag{35′}$$

Now combining Eqs. (34) and (35), one obtains

$$\int d^3\mathbf{r}_1'\int d^3\mathbf{v}_1' A(\mathbf{r}_1,\mathbf{v}_1;\mathbf{r}_1,\mathbf{v}_1)G^\dagger(\mathbf{r}_1',\mathbf{v}_1',t_1;\mathbf{r},\mathbf{v},t) = \int d^3\mathbf{r}_1 d^3\mathbf{v}_1 A(\mathbf{r}_1',-\mathbf{v}_1';r_1-\mathbf{v}_1) \times G(\mathbf{r}_1',-\mathbf{v}_1',-t_1;\mathbf{r},-\mathbf{v},-t). \tag{36}$$

The Onsager reciprocity relation in kinetic theory (32) follows now immediately from Eqs. (31) and (36).

VII. HYDRODYNAMICS

Information about the behavior of solutions to Eq. (12) for large t can be obtained by discussing the spectrum of the linear operator P near $\lambda=0$. Since we follow the approach to f_0 only [f_0 is independent of $\mathbf{r}$, see Eq. (10)] one can expect that the simpler discussion of the spectrum near $\lambda=0$ of the spatial Fourier transformation of P yields a useful information. (More details related to this problem can be found in Ref. 4.) Discussions in this section require complexification of the Hilbert space $\mathcal{H}_2$. The properties (21) and (22) in Lemma 2 will read now as

$$\mathrm{Re}\,(\varphi,AP^{(\dagger)}\varphi)_1\le 0, \tag{21′}$$

$$\mathrm{Re}\,(\varphi,AP^{(-)}\varphi)_1=0. \tag{22′}$$

The following notation is used: $\hat{\varphi}=F\varphi$, $\hat{P}=FPF^{-1}$, where $F=F_{\mathbf{r}}I_{\mathbf{v}}$, $F_{\mathbf{r}}$ is the spatial Fourier transformation, $I_{\mathbf{v}}$ is the identity operator with respect to the dependence on $\mathbf{v}$. We shall assume that $\mathbf{k}$ is fixed and $|k|=k$ is small. The operator $\hat{P}$ and the function φ for k fixed are denoted by $P_{\mathbf{k}}$ and $\varphi_{\mathbf{k}}$, respectively. The operator $P_{\mathbf{k}}$ acts in the Hilbert space $\mathcal{H}_{2,k}$ of functions $\varphi_{\mathbf{k}}$, where the inner product is defined by $\langle\psi_{\mathbf{k}},\varphi_{\mathbf{k}}\rangle 2=\int d^3\mathbf{v}\psi_{\mathbf{k}}(\mathbf{v})A_k\varphi_k(\mathbf{v})$. We shall assume the $\int d^3\mathbf{r}e^{i\mathbf{k}\mathbf{r}}V(\mathbf{r})<\infty$. The explicit form of the operator $P_{\mathbf{k}}$ can be found in Ref. 1. All properties of $P_{\mathbf{k}}$ and $A_{\mathbf{k}}P_{\mathbf{k}}$ which we shall need follow immediately from the properties of P and AP derived in Lemma 2 or they can be deduced from the explicit expressions for $P_{\mathbf{k}}$ and A_k by following step by step the discussion in Lemma 2. The great advantage of the space $\mathcal{H}_{2,k}$ over $\mathcal{H}_{1,k}$ used in Ref. 1 for a study of the spectrum of $P_{\mathbf{k}}$ becomes evident if one compares the complicated and incomplete study of the spectrum $P_{\mathbf{k}}$ in Ref. 1 with the simple and more detailed discussion which will follow. It must be pointed out, however, that in Ref. 2 the approach to any equilibrium state which is independent of $\mathbf{r}$ has been discussed. If we consider $\mathcal{H}_2$, we require moreover that the equilibrium states to which the system is evolving are thermodynamically stable [Eq. (9)]. The condition defining the thermodynamic stability [Eq. (9)] has been obtained in Ref. 2 (by using however only models of $P_{\mathbf{k}}$ and leaving out the study of the residual spectrum) as a condition for absence of the eigenvalues with positive real part (a condition for the linear stability).

Theorem 3: The residual spectrum of the operator P_k is empty; the essential spectrum lies in the region $\mathrm{Re}\lambda \leq -\nu_0$, where $\nu_0 = \lim_{v\to 0} \nu(v)$; the points of the point spectrum can lie only between $\mathrm{Re}\lambda = 0$ and $\mathrm{Re}\lambda = -\nu_0$. Moreover, if k is small the eigenfunctions corresponding to the eigenvalues closest to $\lambda = 0$ are well approximated by linear combinations of five functions, namely constant, $\mathbf{v}$ and v^2.

Proof: Since this theorem is known for the Boltzmann equation[4,5] and its proof is based in fact only on the properties listed in Lemma 2, we can just sketch the main ideas.

The residual spectrum of $A_k P_k$ is empty as follows from Eq. (23) (by using the property that for each point λ of the residual spectrum of a linear operator, the complex conjugate of λ lies in the point spectrum of its adjoint). It is worthwhile to note that the fact that the residual spectrum is empty is a consequence of the same property of the operator P[Eq. (23)] as the Onsager reciprocity relation (32). Hence, we have a link between the residual spectrum and the Onsager reciprocity relation which gives a physical interpretation of the problem of the residual spectrum.

The essential spectrum can be discussed by using exactly the same method as that one used in the theory of the linearized Boltzmann equation (the Weyl–Kato theorem[9] about the influence of a compact perturbation of a closed operator on the essential spectrum). The same method can be even used in both $\mathcal{K}_{2,k}$ and $\mathcal{K}_{1,k}$.[1] The general result is that the essential spectrum lies in the region $\mathrm{Re}\lambda < -\nu_0$, where $\nu_0 = \lim_{v\to 0} \nu(v)$. It means that the essential spectrum does not influence the asymptotic behavior ($t \to \infty$) of $\varphi(\mathbf{r}, \mathbf{v}, t)$. This result gives also a physical meaning to the essential spectrum.

The most important part of the spectrum with respect to the long-time behavior of $\varphi_k(\mathbf{v}, t)$ is that which is close to $\lambda = 0$. We have just shown that only point spectrum occurs in the vicinity of $\lambda = 0$. Equations (21), (22) mean, moreover, that there is no point spectrum with $\mathrm{Re}\lambda > 0$, so that the points of the point spectrum can lie only between $\mathrm{Re}\lambda = 0$ and $\mathrm{Re}\lambda = -\nu_0$. The operator $A_k P_k^{(+)}$ has the eigenvalue $\lambda = 0$, its corresponding eigenfunctions are $1, \mathbf{v}, v^2$ [see Eq. (21)]. The operator $A_k P_k^{(-)}$ is proportional to $i\mathbf{kv}$ for small k and large v. Because we want to get an information about the spectrum of $A_k P_k$ near $\lambda = 0$ if k is small, it is natural to consider $A_k P_k^{(-)}$ as a perturbation of $A_k P_k^{(+)}$. This idea has been used first by McLennan[9] in the theory of the linearized Boltzmann equation. Mathematically, we have the unbounded operator $A_k P_k^{(+)}$ [unbounded due to the multiplication operator $\nu(v)\varphi_k$] perturbed by another unbounded operator $A_k P_k^{(-)}$; ik is considered as a perturbation parameter. Mathematical theory dealing with such a problem is due to Rellich.[9] The basic idea is that a small k causes a small change in eigenvalues and corresponding eigenfunctions if the perturbation operator is relatively bounded with respect to the unperturbed operator (an operator A is relatively bounded with respect to an operator B if $\|Au\| \leq a\|u\| + b\|Bu\|$, where a, b are finite constants). Following McLennan[5] we can show that indeed the operator $A_k P_k^{(-)}$ is relatively bounded with respect to $A_k P_k^{(+)}$; therefore, if k is sufficiently small, the eigenfunctions corresponding to eigenvalues being closest to zero are close to $1, \mathbf{v}, v^2$.

From Theorem 3, we can now expect that for t large the state of the system is adequately described by five quantities only (we shall call them renormalized hydrodynamic state variables):

$$n_0(\mathbf{r}, t) = \int d^3\mathbf{v} A(1 + \varphi(\mathbf{r}, \mathbf{v}, t)), \tag{37}$$

$$n_0(\mathbf{r}, t) u_0(\mathbf{r}, t) = \int d^3\mathbf{v} A\mathbf{v}(1 + \varphi(\mathbf{r}, \mathbf{v}, t)),$$

$$\tfrac{3}{2} n_0(\mathbf{r}, t) k_B T_0(\mathbf{r}, t) = \int d^3\mathbf{v} A \frac{m}{2}(\mathbf{v} - u_0(\mathbf{r}, t))^2(1 + \varphi(\mathbf{r}, \mathbf{v}, t)).$$

We can now also conclude that an appropriate method for obtaining the long-time behavior of the function $f(\mathbf{r}, \mathbf{v}, t)$ in Eq. (1), approaching a single-phase thermodynamically-stable equilibrium state, is the renormalized Chapman–Enskog method. The function $f(\mathbf{r}, \mathbf{v}, t)$ is assumed to be dependent on time only through the renormalized hydrodynamic state variable (37). The first term in the expansion (in the space derivative) of the renormalized hydrodynamic state variable (37) is the corresponding renormalized local equilibrium state. The information about the properties of solutions to Eq. (1) which we have obtained so far implies that this renormalized Chapman–Enskog method is valid for the same reasons (f close to f_0) as the usual Chapman–Enskog method if applied to the Boltzmann equation. By putting $c_1 = c_2 = 1$, $c_3 = c_4 = 0$ into (37) (i.e., only the Boltzmann equation is considered) the renormalized Chapman–Enskog method becomes identical with the usual Chapman–Enskog method.

We believe that the view of kinetic theory developed in this paper, which is basically an extension of nonequilibrium thermodynamics, can be used for a study of further extensions of nonequilibrium thermodynamics which may lead to an exact microscopic dynamics.

ACKNOWLEDGMENTS

This research was partially supported by the NRC Grant A6167. I am indebted to my colleagues from Department of Mathematics and Department of Physics particularly then to Professors V. Dlab, M. Rahman, and M. K. Sundaresan, for their support and discussions. I have benefited also from discussions about this work with Professors G. G. Emch, H. Grad, J. L. Lebowitz, I. Prigogine, R. Rosen, K. M. van Vliet and from correspondence with Professor H. Primas. I would like to thank the referee for his suggestions which helped improve the presentation of the work.

[1] M. Grmela, J. Stat. Phys. 3, 347 (1971).
[2] L. de Sobrino, Can J. Phys. 45, 363 (1967).
[3] H. Grad, in *Rarefied Gas Dynamics* (edited by J. A. Laurmann (Academic, New York, 1963).
[4] G. Scharf, Helv. Phys. Acta 40, 929 (1967).
[5] J. A. McLennan, Phys. Fluids 8, 1580 (1965).
[6] H. B. Callen, *Thermodynamics* (Wiley, New York, London, 1960).
[7] G. Helmberg, *Introduction to Spectra Theory in Hilbert Space* (North Holland, Amsterdam, 1969), Chap. IV.
[8] M. A. Krasnoselskii, *Topological Methods in the Theory of Nonlinear Integral Equations* (Pergamon, New York, 1964), Chap. IV.
[9] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), Chap. VI.
[10] K. Yoshida, *Functional Analysis* (Springer, New York, 1968).
[11] G. M. Wing, *An Introduction to Transport Theory* (Wiley, New York, 1962).
[12] P. M. Morse and F. Feschbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I, Sec. 7.5.

Cluster coefficients for the one-dimensional Bose gas with point interactions

L. R. Dodd and A. M. Gibbs

Department of Mathematical Physics, University of Adelaide, Adelaide, S. Australia 5001
(Received 18 June 1973)

Exact expressions for the cluster coefficients b_2 and b_3 for the one-dimensional Bose gas with repulsive δ-function interactions are found. The calculation depends only on scattering information in the form of two- and three-body scattering wavefunctions. Although not in agreement with a previous calculation by Servadio, our results are shown to be consistent with the work of Yang and Yang which involves the use of periodic rather than scattering boundary conditions.

I. INTRODUCTION

Both the thermodynamic properties[1-3] and the S matrix[4,5] are known for a system of bosons moving in one dimension and interacting through repulsive δ-function potentials. It is interesting to ask whether the statistical mechanics of this system can be deduced from the S-matrix along the lines suggested by the general theory of Dashen, Ma, and Bernstein.[6] In such an approach the solutions of the n-body problem in a finite domain are not needed; only scattering information is used to construct the partition function. In addition to providing a specific illustration of the relationship between the S matrix and statistical mechanics, this type of calculation could be of practical use in finding new results for other one-dimensional gases where the S-matrix or scattering solutions are known[4,5] but where the solutions satisfying periodic boundary conditions are not available.

In this paper an exact expression for the third virial coefficient for the one-dimensional Bose gas with repulsive point interactions is found. The calculation uses scattering information in the form of the two- and three-body scattering wavefunctions. The question of whether on-shell information alone is sufficient to determine the equilibrium thermodynamic properties of the gas is not answered, since our starting point is not the S matrix, but the cluster operator of Lee and Yang.[7] It is not clear whether the cluster coefficients may be expressed in terms of S-matrix elements since the formal limiting processes used in Ref. 3 are not valid for the singular amplitudes of this system. We hope to return to this question later. However, we are able to demonstrate that the third virial coefficient may be calculated from the solutions of the scattering problem without invoking box normalization.

Servadio[8] has also applied the cluster formulation to the one-dimensional Bose gas with point interactions. Although our expression for the second cluster coefficient b_2 agrees with the result stated in Ref. 8, the results for the matrix elements of the third cluster operator U_3 disagree.

An independent check is provided by the work of Yang and Yang[2,3] where the equation of state is found by a different method employing periodic boundary conditions. In the Appendix the results of this paper are shown to be consistent with the theory of Yang and Yang. In view of this agreement we believe that the expression given in Ref. 8 for the third cluster operator is incorrect.

II. DEFINITIONS AND BACKGROUND RESULTS

We consider a system of n bosons interacting through two-body repulsive δ-function potentials of strength c. The scattering wavefunction $\phi_{\{k\}}$ associated with incoming momenta $k_1 \cdots k_n$ is a linear combination of all possible plane-wave types[4,9]:

$$\phi_{\{k\}}(x) = [n!(2\pi)^n]^{-1/2} \Sigma_P a(P) \exp(i \Sigma_i k_{Pi} x_i). \tag{1}$$

The sum runs over all permutations P of order n and the coefficients $a(P)$ are determined in terms of the momenta $k_1 \cdots k_n$ and the strength c of the interaction by

$$a(P) = \exp(\tfrac{1}{2} i \Sigma_{i<j} \psi_{Pi,Pj}), \tag{2}$$

with

$$\psi_{i,j} = 2 \tan^{-1}\left(\frac{c}{k_i - k_j}\right). \tag{3}$$

The wavefunction (1) is only valid in the region $-\infty < x_1 < x_2 < x_3 \cdots < x_n < \infty$, but the wavefunction in all other regions may be obtained from (1) by symmetry considerations.

Gaudin[9] has shown that the scattering wavefunctions (1) satisfy the closure relation

$$\int dk_1 \ldots dk_n \phi^*_{\{k\}}(x) \phi_{\{k\}}(y) = \delta^n(x - y). \tag{4}$$

We recall that the equilibrium pressure p and density ρ are determined in terms of the cluster coefficients b_n, the fugacity z, and the inverse temperature β by

$$p = \beta^{-1} \sum_{n=1}^{\infty} b_n z^n, \tag{5a}$$

$$\rho = \sum_{n=1}^{\infty} n b_n z^n. \tag{5b}$$

In the coordinate representation and in the infinite volume limit the coefficients b_n may be calculated from the expression[7]

$$b_n = \frac{1}{n!} \int \langle 0, x_2, x_3, \ldots, x_n | U_n | 0, x_2, \ldots, x_n \rangle dx_2 \ldots dx_n. \tag{6}$$

The cluster operator U_n is the connected part of the operator $W_n = e^{-\beta H_n}$, where H_n is the n-particle Hamiltonian. Explicitly,

$$\langle x_1' | U_1 | x_1 \rangle = \langle x_1' | W_1 | x_1 \rangle, \tag{7a}$$

$$\langle x_1' x_2' | U_2 | x_1 x_2 \rangle = \langle x_1' x_2' | W_2 | x_1 x_2 \rangle - \langle x_1' | W_1 | x_1 \rangle \langle x_2' | W_1 | x_2 \rangle, \tag{7b}$$

$$\begin{aligned} \langle x_1' x_2' x_3' | U_3 | x_1 x_2 x_3 \rangle = {} & \langle x_1' x_2' x_3' | W_3 | x_1 x_2 x_3 \rangle \\ & - \langle x_1' | W_1 | x_1 \rangle \langle x_2' x_3' | W_2 | x_2 x_3 \rangle \\ & - \langle x_2' | W_1 | x_2 \rangle \langle x_1' x_3' | W_2 | x_1 x_3 \rangle \\ & - \langle x_3' | W_1 | x_3 \rangle \langle x_1' x_2' | W_2 | x_1 x_2 \rangle \\ & + 2 \langle x_1' | W_1 | x_1 \rangle \langle x_2' | W_1 | x_2 \rangle \\ & \times \langle x_3' | W_1 | x_3 \rangle, \text{ etc.} \end{aligned} \tag{7c}$$

With the help of the complete set of scattering states of Eq. (4), the matrix elements of W_n may be expressed as

$$\langle x'|W_n|x\rangle = \int dk_1 \ldots dk_n \phi^*_{\{k\}}(x') \exp[-\beta(k_1^2 + k_2^2 + \cdots + k_n^2)\phi_{\{k\}}(x). \quad (8)$$

Our aim is to calculate the coefficients b_2 and b_3 directly from Eqs. (7) using Eq. (8). The third virial coefficient is then given by

$$a_3 = (4b_2^2 - 2b_3)/b_1^3.$$

III. THE CLUSTER COEFFICIENT b_2

The calculation of b_2 from Eq. (7b) is quite straightforward but we include here some of the details which are repeated in the more complex calculation of b_3 in Sec. IV. The calculation is simplified by considering the difference between the coefficient b_2 and its value $b_2^0 = (4\pi\beta)^{-1/2}2^{-3/2}$ in the ideal Bose limit, where the interaction strength $c \to 0$. In this limit the coefficients $a(P)$ of the plane waves in Eq. (1) are unity.

Using Eq. (8) we have

$$\begin{aligned}\langle x_1x_2|e^{-\beta H_2} - e^{-\beta H_2^0}|x_1x_2\rangle &= \tfrac{1}{2}(2\pi)^{-2}\Sigma_P\Sigma_Q \int dk_1 \int dk_2 e^{-\beta(k_1^2+k_2^2)} e^{i(k_{P1}-k_{Q1})} \\ &\times e^{i(k_{P2}-k_{Q2})x_2}(a(P)a^*(Q) - 1) \\ &= \frac{1}{(2\pi)^2}\int dk_1 \int dk_2 e^{-\beta(k_1^2+k_2^2)} e^{i(k_1-k_2)(x_1-x_2)}(e^{i\psi_{12}} - 1). \end{aligned} \quad (9)$$

We note that the disconnected parts, i.e., the terms which are independent of x_1 and x_2, cancel. From Eq. (3) the factor

$$\{21\} = e^{i\psi_{12}} - 1$$

in Eq. (9) is a simple pole which may be written as

$$\begin{aligned}\{21\} &= 2ic(k_1 - k_2 - ic)^{-1} \\ &= -2c\int_0^\infty e^{-i(k_1-k_2-ic)s}ds. \end{aligned} \quad (10)$$

With the help of this representation, the integrations over the momenta in Eq. (9) may be performed yielding

$$\begin{aligned}\langle x_1x_2|U_2 - U_2^0|x_1x_2\rangle &= \frac{-2c}{(2\pi)^2}\left(\frac{\pi}{\beta}\right)\int_0^\infty e^{-(1/2\beta)(s+x_2-x_1)^2 - cs}ds \\ &= \frac{-2c}{(2\pi)^2}\frac{\pi}{\beta}\, e^{c(x_2-x_1)}\int_{x_2-x_1}^\infty e^{-s^2/2\beta - cs}ds. \end{aligned} \quad (11)$$

The expression (11) is only valid for $x_1 < x_2$. However, it is easy to see from their definition that the matrix elements of U are symmetric in the coordinates, so that in order to calculate b_2 from Eq. (6), we take

$$b_2 = \tfrac{1}{2}\int_0^\infty \langle 0x_2|U_2|0x_2\rangle dx_2 + \tfrac{1}{2}\int_{-\infty}^0 \langle x_20|U_2|x_20\rangle dx_2. \quad (12)$$

From Eqs. (11) and (12), we find, on integrating by parts,

$$\begin{aligned}b_2 - b_2^0 &= \frac{-2c}{(2\pi)^2}\frac{\pi}{\beta}\int_0^\infty dx e^{cx}\int_x^\infty e^{-s^2/2\beta - cs}ds \\ &= \frac{1}{2\pi\beta}\int_0^\infty ds e^{-s^2/2\beta}(e^{-cs} - 1). \end{aligned} \quad (13)$$

The results (11) and (13) have already been obtained by Servadio.

IV. THE CLUSTER COEFFICIENT b_3

Again we consider the difference between b_3 and its value $b_3^0 = (4\pi\beta)^{-1/2}3^{-3/2}$ in the ideal Bose limit. From Eq. (7c),

$$\begin{aligned}\langle x_1x_2x_3|U_3 - U_3^0|x_1x_2x_3\rangle &= \langle x_1x_2x_3|e^{-\beta H_3} - e^{-\beta H_3^0}|x_1x_2x_3\rangle \\ &- \langle x_1|e^{-\beta H_1}|x_1\rangle\langle x_2x_3|e^{-\beta H_2} - e^{-\beta H_2^0}|x_2x_3\rangle \\ &- \langle x_2|e^{-\beta H_1}|x_2\rangle\langle x_1x_3|e^{-\beta H_2} - e^{-\beta H_2^0}|x_1x_3\rangle \\ &- \langle x_3|e^{-\beta H_1}|x_3\rangle\langle x_1x_2|e^{-\beta H_2} - e^{-\beta H_2^0}|x_1x_2\rangle. \end{aligned} \quad (14)$$

After using Eq. (8) and exploiting the symmetry in the momentum variables, the right-hand side of Eq. (14) becomes

$$\begin{aligned}\int dk e^{-\beta k^2}([213]\{213\} &+ [132]\{132\} + [321]\{321\} \\ &+ [312]\{312\} + [231]\{231\} - [213]\{21\} - [132]\{32\} \\ &- [321]\{31\}), \end{aligned} \quad (15)$$

where we have introduced the notation

$$[ijk] = e^{i(k_1-k_i)x_1}e^{i(k_2-k_j)x_2}e^{i(k_3-k_k)x_3},$$

$$\{ijk\} = e^{i/2(\psi_{12}+\psi_{13}+\psi_{23}-\psi_{ij}-\psi_{ik}-\psi_{jk})} - 1,$$

and

$$\int dk\, e^{-\beta k^2} = \frac{1}{(2\pi)^3}\int dk_1 dk_2 dk_3\, e^{-\beta(k_1^2+k_2^2+k_3^2)}.$$

First we note that the coefficient $\{213\} - \{21\}$ of the plane wave [213] which represents the scattering of particles 1 and 2 with particle 3 propagating freely, vanishes. Similarly the disconnected part [132] of the cluster integral, which results from the scattering of particle 2 and 3 while particle 1 propagates freely, vanishes due to the cancellation of the three-body term $\{132\}$ and the two-body term $\{32\}$. On the other hand $\{321\}$ and $\{31\}$ are not equal, so that a term involving [321] which is independent of x_2 survives. The different treatment of the scattering of particles 1 and 3 results from the definite ordering of the particles in the initial state. In order for particles 1 and 3 to interchange momenta in three-body processes, they must also scatter from particle 2. It should also be remarked that although the term [321] is independent of x_2, there is no difficulty with convergence when integrating over x_2 in the process of forming the coefficient b_3. Since x_2 is limited by the condition $x_1 < x_2 < x_3$, the x_1 and x_3 dependence is sufficient to produce convergence.

By taking the above cancellations into account and by suitable relabeling of variables, the expression (15) becomes

$$\int dk\, e^{-\beta k^2}(2[231]\{231\} + [321](\{321\} - \{31\})).$$

In terms of the integral representation of the plane-wave coefficients of Eq. (10),

$$\{231\} = \{21\}\{31\} + \{31\} + \{21\}$$

and

$$\begin{aligned}\{321\} - \{31\} &= (1 + \{31\})(\{21\}\{32\} + \{21\} + \{32\}) \\ &= (1 - \{31\} + \tfrac{3}{2}\{\overline{31}\})(\{21\} + \{32\}), \end{aligned}$$

where

$$\{\overline{31}\} = \frac{4ic}{k_1 - k_3 - 2ic} = -4c\int_0^\infty e^{-i(k_1-k_3-2ic)s}ds$$

has been introduced. The identity for $\{321\} - \{31\}$ involving $\{\overline{31}\}$ is not obvious but simplifies considerably the rest of the calculation. With these integral representations of the pole terms, the integrations over the momenta k_1, k_2, and k_3 can now be performed. In terms of the basic integrals,

$$f(x) = -2ce^{cx}\int_x^\infty ds\, e^{-s^2/2\beta - cs}$$

and

$$g_{c'}(x,y) = 4c'ce^{c'x}e^{cy}\int_x^\infty ds\int_y^\infty dt\, e^{-1/2\beta(s^2+t^2+st)}e^{-c's}e^{-ct},$$

the result is

$$\begin{aligned}\langle x_1x_2x_3|U_3 - U_3^0|x_1x_2x_3\rangle &= 2/(2\pi)^3(\pi/\beta)^{3/2}[g_c(x_3-x_2, x_2-x_1)\\ &+ e^{-3/8\beta(x_3-x_2)^2}f(\tfrac{1}{2}(x_3+x_2)-x_1)\\ &+ e^{-3/8\beta(x_1-x_2)^2}f(x_3-\tfrac{1}{2}(x_1+x_2))\\ &+ e^{-3/8\beta(x_3-x_1)^2}f(\tfrac{1}{2}(x_3-x_1))\\ &+ \tfrac{3}{2}g_{2c}(x_3-x_1,0) - g_c(x_3-x_1,0)]. \end{aligned} \quad (16)$$

Only the first term in the above expression is given for the matrix elements of $U_3 - U_3^0$ in Ref. 8. From Eq. (6)

$$\begin{aligned} b_3 - b_3^0 &= \frac{1}{6}\int_{-\infty}^\infty dx_2\int_{-\infty}^\infty dx_3\langle 0, x_2, x_3|U_3 - U_3^0|0, x_2x_3\rangle\\ &= \frac{1}{3}\int_{x_1<x_2<x_3}\langle x_1x_2x_3|U_3 - U_3^0|x_1x_2x_3\rangle(\delta(x_1)\\ &+ \delta(x_2) + \delta(x_3))dx_1dx_2dx_3. \end{aligned} \quad (17)$$

Our final expression for b_3 is obtained from Eqs. (16) and (17). After repeated integration by parts over the coordinates $x_1x_2x_3$, several terms cancel, leading to

$$b_3 - b_3^0 = \frac{6}{(2\pi)^3}\left(\frac{\pi}{\beta}\right)^{3/2}\int_0^\infty ds\int_0^\infty dt e^{-(1/2\beta)(s^2+t^2+st)}(e^{-cs-ct}-1)e^{-cs}. \quad (18)$$

In the Appendix it is shown that this result agrees with the work of Yang and Yang, which is based on the solution of the n-body problem with periodic boundary conditions.

ACKNOWLEDGMENT

The authors wish to thank Professor J. B. McGuire for stimulating their interest in this problem.

APPENDIX

In the theory of Yang and Yang[3] the third cluster coefficient is determined by

$$b_3 - b_3^0 = \frac{1}{2\pi}\int_{-\infty}^\infty dk(a_3 - a_1a_2), \quad (A1)$$

where

$$a_1 = e^{-\beta k^2}, \quad a_2 = e^{-\beta k^2}\underline{O}a_1$$

$$a_3 = e^{-\beta k^2}\tfrac{1}{2}(\underline{O}a_1)^2 - \tfrac{1}{2}\underline{O}a_1^2 + \underline{O}a_2,$$

and $\underline{O}$ is the integral operator

$$\frac{c}{\pi}\int_\infty^\infty dq\frac{1}{c^2+(k-q)^2}.$$

After some simplification, we obtain

$$b_3 - b_3^0 = \frac{3}{4\pi}\int dk_1dk_2dk_3e^{-\beta(k_1^2+k_2^2+k_3^2)}\left[\frac{c}{\pi}\frac{1}{(k_1-k_2)^2+c^2}\right] \times \left[\frac{c}{\pi}\frac{1}{(k_2-k_3)^2+c^2} - \delta(k_2-k_3)\right]. \quad (A2)$$

When the pole terms are expressed as integrals as in Eq. (10) and the integrations over the momenta performed, the right-hand side of Eq. (A2) becomes

$$\frac{3}{(2\pi)^2}\left(\frac{\pi}{\beta}\right)^{3/2}\int_0^\infty ds\int_0^\infty dt[e^{-1/2\beta(s^2+t^2+st)} + e^{-1/2\beta(s^2+t^2-st)}](e^{-cs}-1)e^{-ct}. \quad (A3)$$

Now, if $f(s,t)$ is symmetric in s and t, then

$$\int_0^\infty ds\int_0^\infty dt\, e^{-(1/2\beta)(s^2+t^2-st)}f(s,t) = 2\int_0^\infty ds\int_0^\infty dt\, e^{-(1/2\beta)(s^2+t^2+st)}f(s+t,t). \quad (A4)$$

Using the identity (A4) to simplify (A3), we obtain the desired result, Eq. (18).

[1]E. Lieb and W. Liniger, Phys. Rev. 130, 1605 (1963).
[2]C. N. Yang and C. P. Yang, J. Math. Phys. 10, 115 (1969).
[3]C. P. Yang, Phys. Rev. A 2, 154 (1970).
[4]J. B. McGuire, J. Math. Phys. 5, 622 (1964).
[5]C. N. Yang, Phys. Rev. 168, 1920 (1968).
[6]R. Dashen, S. Ma, and M. J. Bernstein, Phys. Rev. 187, 345 (1969); R. Dashen and S. Ma, J. Math. Phys. 11, 1136 (1970); 12, 689 (1971).
[7]T. D. Lee and C. N. Yang, Phys. Rev. 113, 1165 (1959).
[8]S. Servadio, J. Math. Phys. 12, 2413 (1971).
[9]M. Gaudin, J. Math. Phys. 12, 1674 (1971).

Lienard–Wiechert fields and general relativity

E. T. Newman

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 29 June 1973)

An analogy is established between the Lienard–Wiechert solutions of the Maxwell equations and the Robinson–Trautman solutions of the Einstein equations by virtue of the fact that a principal null vector field of either the Maxwell or Weyl tensor in each case satisfies the following four conditions: (1) The field is a geodesic field, (2) it has nonvanishing divergence, (3) it is shear free, and (4) it is twist (or curl) free.

I. INTRODUCTION

It is the purpose of this note to point out the very strong analogy between the Lienard–Wiechert solutions of the Maxwell equations and the class of solutions of the Einstein equations known as the Robinson–Trautman (RT) metrics.[1] More precisely the analogy is with the regular type II, RT metrics, all other RT metrics having unacceptable singularities.

The RT metrics can be geometrically characterized by the following properties of one of the principal null vector (p.n.v.) fields of the Weyl tensor: (i) the p.n.v. field is the tangent field to a congruence of null geodesics; (ii) the field has nonvanishing divergence; (iii) the shear of the field is zero; (iv) the twist (or curl) of the field is zero. These conditions are both necessary and sufficient for a vacuum metric to be a RT metric.

We will prove that a necessary and sufficient condition for a Maxwell field regular at infinity (in flat space) to be a Lienard–Wiechert field is that one of the p.n.v. fields of the Maxwell tensor satisfies conditions (i)–(iv) thus establishing the analogy.

[We mention (without proof here) that this analogy has a rather remarkable generalization. It is known[2] that the Maxwell equations can be analytically extended into complex Minkowski space. The real solution corresponding to a "Lienard–Wiechert particle" moving in complex Minkowski space has a p.n.v. field which satisfies only conditions (i)–(iii). (For a special case of this, see Ref. 2.) When the "particle" is confined to the real Minkowski space, (iv) is automatically satisfied. The analogy to this, in general relativity, is the class of suitably regular solutions of the Einstein equations satisfying (i)–(iii), namely, the regular twisting type II metrics.[3,4] These solutions can be naturally viewed as being given on real subspaces of a complex manifold in which the source (a Lienard–Wiechert-like particle) is moving in the complex region.]

II. THE LIENARD-WIECHERT FIELD

The simplest way to prove our contention concerning the properties of the p.n.v. field of the Lienard–Wiechert solution is to use the spin-coefficient form of the Maxwell equations with a coordinate and tetrad system associated with the p.n.v. field.

It is known[5] that the most general null vector field satisfying conditions (i)–(iv) is obtained by taking a timelike worldline (actually spacelike or null worldlines could be used, but they lead to unacceptable singularities) and constructing the light cone at each point on the worldline. The tangent vectors to the generators of each cone form the field in question.

The null coordinate system associated with the family of cones can be introduced by the transformation[6]

$$x^\mu = \xi^\mu(\phi) + r l^\mu / v(\phi, \zeta, \bar{\zeta}), \tag{1}$$

with $x^\mu = \xi^\mu(\phi)$ being the parametric form of the world line (ϕ is $\sqrt{2}/2$ times the proper time), with

$$l^\mu = \frac{1}{2\sqrt{2}P_0}\left(1 + \zeta\bar{\zeta}, \zeta + \bar{\zeta}, \frac{\zeta - \bar{\zeta}}{i}, -1 + \zeta\bar{\zeta}\right),$$

$2P_0 = 1 + \zeta\bar{\zeta}$, ζ and $\bar{\zeta}$ being complex stereographic coordinates on the sphere, and $v(\phi, \zeta, \bar{\zeta}) = l^\mu \xi_\mu(\phi)$, $\dot{\xi}_\mu \dot{\xi}^\mu = 2$.

Our vector field in the new coordinate system has the form

$$l_\mu = \delta^\phi_\mu, \qquad l^\mu = \delta^\mu_r,$$

and the Minkowski metric becomes

$$ds^2 = 2(1 - \frac{\dot{v}}{v} r)d\phi^2 + 2dr d\phi - \frac{r^2}{2}\frac{d\zeta d\bar{\zeta}}{P_0^2 v^2}$$
$$= 2(l_\mu n_\nu - m_\mu \bar{m}_\nu) dx^\mu dx^\nu,$$

with

$$l_\mu dx^\mu = d\phi, \qquad n_\mu dx^\mu = [1 - (\dot{v}/v)r]d\phi + dr,$$
$$m_\mu dx^\mu = \frac{-r d\bar{\zeta}}{2P_0} v, \qquad \bar{m}_\mu dx^\mu = \frac{-r d\zeta}{2P_0} v.$$

The tetrad components of the Maxwell tensor[7]

$$\phi_0 = F_{\mu\nu} l^\mu m^\nu, \qquad \phi_1 = \tfrac{1}{2} F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu),$$
$$\phi_2 = F_{\mu\nu} \bar{m}^\mu n^\nu \tag{2}$$

satisfy the equations

$$D\phi_1 + \frac{2\phi_1}{r} = -\frac{v^2}{r}\,\bar{\eth}_0\left(\frac{\phi_0}{v}\right),$$
$$D\phi_2 = \frac{\phi_2}{r} = -\frac{v}{r}\,\bar{\eth}_0\phi_1,$$
$$\dot{\phi}_0 - \left(1 - \frac{\dot{v}}{v}r\right)D\phi_0 - \frac{1}{r}\left(1 - \frac{\dot{v}}{v}r\right)\phi_0 = -\frac{v}{r}\,\eth_0\phi_1, \tag{3}$$
$$\dot{\phi}_1 - \left(1 - \frac{\dot{v}}{v}r\right)D\phi_1 - \frac{2}{r}\phi_1 = -\frac{v^2}{r}\,\eth\left(\frac{\phi_2}{v}\right) + \phi_0 v\,\eth_0\frac{\dot{v}}{v},$$

where the dot and D denote, respectively, the derivative with respect to ϕ and r, and $\eth_0$ and $\bar{\eth}_0$ are angular derivates.[8]

The condition $\phi_0 = 0$ is equivalent to the statement that l_μ is a p.n.v. of the Maxwell field. Our task is thus to show that the regular solutions of

$$D\phi_1 + \frac{2\phi_1}{r} = 0, \qquad D\phi_2 + \frac{\phi_2}{r} = -\frac{v}{r}\,\bar{\eth}_0\phi_1,$$

$$0 = -\frac{v}{r}\eth_0\phi_1,$$
$$\dot{\phi}_1 - \left(1 - \frac{\dot{v}}{v}r\right)D\phi_1 - \frac{2}{r}\phi_1 - \frac{2}{r}\phi_1 = -\frac{v^2}{r}\eth_0\left(\frac{\phi_2}{v}\right) \tag{4}$$

are the Lienard–Wiechert fields.

Integrating the first two equations yields

$$\phi_0 = 0, \quad \phi_1 = \frac{\phi_1^0(\phi, \zeta, \bar{\zeta})}{r^2},$$
$$\phi_2 = \frac{\phi_2^0(\phi, \zeta, \bar{\zeta})}{r} + \frac{v}{r^2}\bar{\eth}_0\phi_1^0, \tag{5}$$

and substitution into the second set yields first $\eth\phi_1^0 = 0$, which implies $\phi_1^0 = \phi_1^0(\phi)$ and

$$(\phi_1^0 v^{-2})^{\cdot} = -\eth_0\frac{\phi_2^0}{v} \tag{6}$$

When Eq. (6) is integrated over the sphere (using the properties of $\eth_0$) we obtain

$$\frac{d}{d\phi}(\phi_1^0 \int v^{-2}d\Omega) = 0.$$

(This is nothing but a special case of the law of conservation of charge.) By explicit integration (or by using more general methods[6]), we find

$$\int v^{-2}d\Omega = 4\pi$$

and we have $\dot{\phi}_1^0 = 0$ or $\phi_1^0 = e$. Thus the final equation to be solved [Eq. (6)] becomes

$$2e\frac{\dot{v}}{v} = v^2\eth_0\left(\frac{\phi_2^0}{v}\right).$$

That a solution is

$$\phi_2^0 = -ev\bar{\eth}_0\left(\frac{\dot{v}}{v}\right) \tag{7}$$

can easily be checked by noting[6] that $v^2\eth_0\bar{\eth}_0 \log P_0 v = 1$ and differentiating with respect to ϕ; and that it is the general regular solution is seen by viewing the differential equation in the instantaneous rest frame of the particle, i.e., when $v = 1$ and thus $\dot{v}$ (because $\ddot{\xi}^\mu$ is orthogonal to ξ^μ) has only $l = 1$ spherical harmonic parts. The equation becomes

$$2e\dot{v} = \eth_0\phi_2^0.$$

One can immediately conclude (from the properties of $\eth_0$ and $\bar{\eth}_0$) that the general regular solution is

$$\phi_2^0 = -e\bar{\eth}_0\dot{v}$$

which is Eq. (7) in the instantaneous rest frame.

Summarizing the results, we have for the tetrad components of the Maxwell tensor

$$\phi_0 = 0,$$
$$\phi_1 = \frac{e}{r^2},$$
$$\phi_2 = -\frac{ev}{r}\bar{\eth}_0\frac{\dot{v}}{v}.$$

By a straightforward but tedious inversion of Eq. (2) one can convert this result into an expression for $F_{\mu\nu}$ and verify that it is the Lienard–Wiechert field. Roughly this can be seen to be true by again viewing the solution in the instantaneous rest frame and noticing that the solution is just the Coulomb field and an electric-dipole radiation field depending linearly on the acceleration. These are known properties of the Lienard–Wiechert field.

[1] I. Robinson and A. Trautman, Proc. R. Soc. Lond. A 265, 463 (1962).
[2] E. T. Newman, J. Math. Phys. 14, 102 (1973).
[3] R. P. Kerr, Phys. Rev. Lett. 11, 237 (1963).
[4] C. J. Talbot, Commun. Math. Phys. 13, 45 (1969).
[5] Though the author does not know a specific reference for this result, it is however common knowledge and can be inferred from the work of Robinson and Trautman.
[6] A. Held, E. T. Newman, and R. Posadas, J. Math. Phys. 11, 3145 (1970).
[7] E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[8] J. N. Goldberg *et al.*, J. Math. Phys. 8, 2155 (1967).

Multipole moments of stationary space-times*

R. O. Hansen†

Department of Physics, University of Chicago, Chicago, Illinois
(Received 19 July 1973)

Multipole moments are defined for stationary, asymptotically flat, source-free solutions of Einstein's equation. There arise two sets of multipole moments, the mass moments and the angular momentum moments. These quantities emerge as tensors at a point Λ "at spatial infinity." They may be expressed as certain combinations of the derivatives at Λ of the norm and twist of the timelike Killing vector. In the Newtonian limit, the moments reduce to the usual multipole moments of the Newtonian potential. Some properties of these moments are derived, and, as an example, the multipole moments of the Kerr solution are discussed.

1. INTRODUCTION

In general relativity, there are many exact solutions for which intuitive pictures are unavailable. One means of constructing such intuitive pictures in Newtonian theory is provided by multipole moments. The multipole moments of solutions of Einstein's equation might similarly suggest interpretations of these fields. For example, if a Newtonian potential exists having the same multipole structure as a given solution of Einstein's equation, that potential could be regarded as the Newtonian analog of the given solution.

One would not expect that all features of Newtonian multipole moments will survive the transition to general relativity, for curvature is likely to destroy the moments in regions near the sources. In fact, the existence of multipole moments in Newtonian theory depends in an essential way on the (conformal) flatness of Euclidean space.[1] However, the moments may reappear as the curvature dies away, i.e., at large distances from the sources. Thus we might expect that multipole moments exist for asymptotically flat space-times, and that these moments should be quantities "at infinity."

A definition of the multipole moments of asymptotically flat, static space-times has been given by Geroch.[2] The prescription for obtaining the moments is as follows: Introduce a 3-surface orthogonal to the timelike Killing vector. On that 3-surface a positive-definite metric and a scalar field ψ are defined. Conformally complete the 3-surface by the addition of a single point "at infinity". The scalar field ψ then defines a smooth function $\tilde{\psi}$ on the conformally completed space. The multipole moments of the space-time are given as certain combinations of the derivatives of $\tilde{\psi}$ "at infinity."

We shall extend the definition of multipole moments to stationary space-times. Two essential difficulties arise in this case. First, a replacement must be found for the surfaces orthogonal to the timelike Killing vector. Second, the field ψ, in terms of which the multipole moments of static space-times were written, is unsuitable for the description of the moments of stationary space-times.

A solution to the first difficulty, i.e., a 3-surface anologous to Euclidean space in Newtonian theory, has already been given[3]: the manifold S of trajectories of the timelike Killing vector. Geroch has also discussed a set of equations on S which are equivalent to Einstein's equation for space-times admitting a timelike Killing vector. These equations will play, in our discussion of multipole moments, the role of Laplace's equation in Newtonian theory.

The resolution of the second difficulty, that of finding suitable fields in terms of which to express the multipole moments, is the subject of Sec. 2 of this paper. There emerge two such fields, ϕ_M and ϕ_J, analogous to the Newtonian mass and angular momentum potentials.[4] These potentials are given by

$$\begin{aligned} \phi_M &= \tfrac{1}{4}\lambda^{-1}(\lambda^2 + \omega^2 - 1), \\ \phi_J &= \tfrac{1}{2}\lambda^{-1}\omega, \end{aligned} \tag{1.1}$$

where λ is the norm and ω the twist of the timelike Killing vector. The prescription given by Geroch[2] is applied to these potentials to yield two sets of multipole moments for each stationary, asymptotically flat solution of Einstein's equation.

In Sec. 3, we use our definition of multipole moments to calculate explicitly the mass and angular momentum moments of the Kerr solution. Some possibilities for the further extension of the definition of multipole moments are discussed in Sec. 4.

2. DEFINITION OF THE MULTIPOLE MOMENTS

The first step in the definition of multipole moments is the introduction of the 3-dimensional space. Let M, g_{ab} be a 4-manifold with metric of signature $(-,+,+,+)$, satisfying the vacuum Einstein equation. Let ξ^a be a timelike Killing vector field on M. Define λ, the norm of ξ^a, by

$$\lambda = -\xi^a\xi_a, \tag{2.1}$$

and ω_a by

$$\omega_a = \epsilon_{abcd}\xi^b\nabla^c\xi^d. \tag{2.2}$$

It follows from the (vacuum) Einstein equation that

$$\nabla_{[a}\omega_{b]} = 0. \tag{2.3}$$

Thus, there exists[5] a scalar field ω (the twist of ξ^a) such that

$$\omega_a = \nabla_a\omega. \tag{2.4}$$

The set S of trajectories of ξ^a in M is (at least locally) a differentiable manifold. It was shown by Geroch[3] that there is a one-to-one correspondence between tensor fields $T'^{a\cdots b}{}_{c\cdots d}$ on S and tensor fields $T^{a\cdots b}{}_{c\cdots d}$ on M satisfying

$$\begin{aligned} &\pounds_\xi T^{a\cdots b}{}_{c\cdots d} = 0, \\ &\xi_a T^{a\cdots b}{}_{c\cdots d} = 0, \cdots, \xi^d T^{a\cdots b}{}_{c\cdots d} = 0. \end{aligned} \tag{2.5}$$

Accordingly, we shall speak of tensor fields on M satisfying (2.5) as tensor fields on S. In particular, λ and ω are tensor fields on S. The tensor field

$$h_{ab} = \lambda g_{ab} + \xi_a\xi_b \tag{2.6}$$

is a positive-definite metric on S.[6] Indices of tensors on S will be raised and lowered using h_{ab}.

We next require that the curvature of S die away at large distances from the sources, i.e., that S be asymptotically flat. We use the definition given by Geroch.[2]

A 3-manifold S with metric h_{ab} is said to be asymptotically flat if there exists a manifold $\tilde{S}$ with metric $\tilde{h}_{ab}$ such that:

(i) $\tilde{S} = S \cup \Lambda$, where Λ is a single point,

(ii) $\tilde{h}_{ab} = \Omega^2 h_{ab}$ is a smooth metric on $\tilde{S}$,

and

(iii) $\Omega|_{\Lambda} = 0$, $\tilde{D}_a\Omega|_{\Lambda} = 0$, $\tilde{D}_a\tilde{D}_b\Omega|_{\Lambda} = 2\tilde{h}_{ab}|_{\Lambda}$, where $\tilde{D}_a$ is the derivative operator associated with $\tilde{h}_{ab}$. For example, Euclidean 3-space is asymptotically flat in this sense. Choose for the conformal factor, e.g., $\Omega = r^{-2}$, where r is the distance from some origin.

Having defined the manifold of trajectories S and its conformal completion $\tilde{S}$, we now introduce the multipole moments of the fields ϕ_M and ϕ_J, defined by (1.1). The symbol ϕ will denote either of these fields.

Let $\tilde{\phi} = \Omega^{-1/2}\phi$, and define recursively a set of tensor fields $P_{a_1\cdots a_s}$ on $\tilde{S}$ by

$$\begin{aligned} &P = \tilde{\phi}, \\ &P_{a_1\cdots a_{s+1}} = \mathcal{C}[\tilde{D}_{a_1}P_{a_2\cdots a_{s+1}} - \tfrac{1}{2}s(2s-1)\tilde{R}_{a_1a_2}P_{a_3\cdots a_{s+1}}], \end{aligned} \tag{2.7}$$

where $\mathcal{C}[T_{a\cdots b}]$ denotes the totally symmetric, trace-free part of $T_{a\cdots b}$, and $\tilde{R}_{ab}$ is the Ricci tensor associated with $\tilde{h}_{ab}$. We define the 2^s moment of ϕ to be the value of $P_{a_1\cdots a_s}$ at Λ. The 2^s moment of ϕ_M will be written $M_{a_1\cdots a_s}$, the 2^s moment of ϕ_J, $J_{a_1\cdots a_s}$.

The formal definition of the multipole moments of a stationary, asymptotically flat space–time is now complete. Given the metric g_{ab} of the space–time and the timelike Killing vector field ξ^a, construct the manifold S of trajectories of the Killing vector and its metric h_{ab}, given by Eq. (2.6). Define the fields ϕ_M and ϕ_J by (1.1). Choose a conformal factor Ω so that $\tilde{S}, \tilde{h}_{ab}$ satisfies conditions (i)–(iii) for asymptotic flatness. Substitute into Eqs. (2.7) to obtain the moments $M_{a_1\cdots a_s}$ and $J_{a_1\cdots a_s}$.

It is natural to ask whether our definition, applied to a Newtonian potential in flat space, yields the usual multipole moments for that system. We now show that this is the case.

Let ϕ be a Newtonian potential, i.e., let ϕ satisfy

$$D^2\phi = 0 \tag{2.8}$$

in Euclidean 3-space. Then ϕ posesses a multipole-moment expansion, i.e., ϕ can be written in the form

$$\phi = r^{-1}Q + r^{-3}Q_a x^a + \tfrac{1}{2}r^{-5}Q_{ab}x^a x^b + \cdots, \tag{2.9}$$

where x^a is the position vector[7] with respect to some origin, and $Q_{a_1\cdots a_s}$, the 2^s moment of ϕ about that origin, is a constant tensor field.

Define the new position vector[7] $\bar{x}^a = r^{-2}x^a$ with respect to the origin "at infinity." The series (2.9), expressed in terms of $\bar{x}^a$, takes the form

$$\phi = r^{-1}Q + r^{-1}Q_a\bar{x}^a + \tfrac{1}{2}r^{-1}Q_{ab}\bar{x}^a\bar{x}^b + \cdots. \tag{2.10}$$

It was noted above that $\Omega = r^{-2}$ is an appropriate choice of conformal factor for completing Euclidean 3-space. Making this choice, and setting $\tilde{\phi} = \Omega^{-1/2}\phi$, we have

$$\tilde{\phi} = Q + Q_a\bar{x}^a + \tfrac{1}{2}Q_{ab}\bar{x}^a\bar{x}^b + \cdots. \tag{2.11}$$

It follows immediately from (2.11) that

$$Q = \tilde{\phi}|_{\Lambda}, \quad Q_a = \tilde{D}_a\tilde{\phi}|_{\Lambda}, \quad Q_{ab} = \tilde{D}_a\tilde{D}_b\tilde{\phi}|_{\Lambda}, \cdots, \tag{2.12}$$

where $\tilde{D}_a$ is the derivative operator associated with $\tilde{h}_{ab}$. Note that the expressions (2.12) coincide, in the case of vanishing curvature, with those of (2.7). Thus our formulation, applied to a Newtonian potential in flat space, yields the correct values for the multipole moments of the system.[8]

We have seen that the multipole moments defined by Eq. (2.7) have the correct values for a Newtonian system. Recall, however, that the usual multipole moments of that system refer to an origin, and depend on the choice of that origin. The behavior of the Newtonian multipole moments under a change of origin is in fact reflected, in our formulation, by the behavior of the multipole moments defined by (2.7) under a change of conformal factor. The role of the choice of conformal factor has been discussed by Geroch.[2] We shall review that discussion briefly.

Conditions (i)–(iii) for asymptotic flatness do not determine the conformal factor uniquely. In fact, one can choose any conformal factor Ω' of the form

$$\Omega' = \Omega w, \tag{2.13}$$

where Ω is a conformal factor satisfying conditions (i)–(iii), and w is any smooth function whose value at Λ is 1. Since arbitrary conformal transformations can be generated by repeated application of infinitesimal ones, it suffices to consider the case in which w differs infinitesimally from 1.

The multipole moments which result from the use of the conformal factor Ω' are related to those associated with Ω by

$$P'_{a_1\cdots a_s}|_{\Lambda} = P_{a_1\cdots a_s}|_{\Lambda} - \tfrac{1}{2}s(2s-1)\mathcal{C}[P_{a_1\cdots a_{s-1}}\tilde{D}_{a_s}w]|_{\Lambda}. \tag{2.14}$$

Notice that the change in the 2^s moment depends only on the 2^{s-1} moment. Geroch[2] has shown that this behavior reflects precisely the usual dependence of Newtonian multipole moments on the choice of origin.

It is remarkable that our definition of multipole moments makes no reference to Einstein's equation. One would expect Einstein's equation to assume the role of Laplace's equation in Newtonian theory. But the prescription of Eqs. (2.7) could be applied, on three-dimensional Euclidean space, to a much larger class of functions than the solutions of Laplace's equation: Any function f such that $\tilde{f} = \Omega^{-1/2}f$ is smooth at Λ could be used. Of course, not every function f on Euclidean 3-space defines a field $\tilde{f}$ which is smooth at Λ. The function must tend to zero "at infinity" at least as fast as r^{-1}. In addition, its "angular oscillation" must die away at infinity.

The solutions of Laplace's equation which tend to zero at infinity do indeed satisfy these requirements: Roughly speaking, the radial and angular dependence of those functions are coupled through their second derivatives. More precisely, recall that Laplace's equation in three dimensions is conformally invariant, that is, if ϕ

satisfies (2.8), we have

$$\tilde{D}^2\tilde{\phi} = 0 \tag{2.15}$$

on $\tilde{S}$. But it is well known that solutions of Laplace's equation either diverge or are smooth at any point. Since $\tilde{\phi}$ is bounded at Λ, its derivatives of all orders exist and are continuous there, i.e., $\tilde{\phi}$ is smooth. Thus Eqs. (2.7) indeed define a collection of tensors at Λ if ϕ is a Newtonian potential.

The point, then, is that Eqs. (2.7) lead to a meaningful definition of multipole moments in general relativity only if ϕ_M and ϕ_J define fields $\tilde{\phi}_M$ and $\tilde{\phi}_J$ which are smooth at Λ. The above discussion for Newtonian fields depends in an essential way on the properties of Laplace's equation. This suggests that Einstein's equation should be used to establish the necessary smoothness of $\tilde{\phi}_M$ and $\tilde{\phi}_J$.

Geroch[3] has given a set of equations on S, equivalent to Einstein's equation for spacetimes admitting a time-like Killing vector. These equations, written in terms of ϕ_M and ϕ_J, take the form[9]

$$\begin{aligned}(D^mD_m - \mathcal{R}/8)\phi_M &= (15/8)\kappa^4\phi_M,\\ (D^mD_m - \mathcal{R}/8)\phi_J &= (15/8)\kappa^4\phi_J,\end{aligned} \tag{2.16}$$

$$\begin{aligned}\mathcal{R}_{ab} = {}& 2[(D_a\phi_M)(D_b\phi M) + (D_a\phi_J)(D_b\phi_J)]\\ & - \tfrac{1}{2}[D_a(1 + 4\phi_M^2 + 4\phi_J^2)^{1/2}][D_b(1 + 4\phi_M^2 + 4\phi_J^2)^{1/2}],\end{aligned} \tag{2.17}$$

where $\mathcal{R}_{ab}$ is the Ricci tensor and $\mathcal{R}$ the scalar curvature associated with h_{ab}, and where κ is defined by

$$\kappa^4 = \tfrac{1}{2}\lambda^{-2}[(D^m\lambda)(D_m\lambda) + (D^m\omega)(D_m\omega)]. \tag{2.18}$$

The arguments which were used to show smoothness for solutions of Laplace's equation suggest that the conformal behavior of these equations will be important. Notice that Eqs. (2.16) are conformally invariant provided that ϕ and κ have dimensions[10] $\sec^{-1/2}$; thus we have

$$\tilde{D}^m\tilde{D}_m\tilde{\phi} = (\tilde{\mathcal{R}}/8)\tilde{\phi} + (15/8)\tilde{\kappa}^4\tilde{\phi}, \tag{2.19}$$

where $\tilde{\mathcal{R}}$ is the scalar curvature associated with $\tilde{h}_{ab}$. Equation (2.19) is an elliptic differential equation on $\tilde{S}$; its coefficients will be smooth on $\tilde{S}$ if $\tilde{\kappa}$ is smooth at Λ. Thus, $\tilde{\phi}$ will be smooth if $\tilde{\kappa}$ is smooth. What remains is to show that $\tilde{\kappa}$ is a smooth function on $\tilde{S}$.

It follows from Eqs. (2.16) and (2.17) that $\tilde{\kappa}$ satisfies the elliptic differential equation

$$\begin{aligned}\tilde{D}^m\tilde{D}_m\tilde{\kappa} = {}& (\tilde{\mathcal{R}}/8)\tilde{\kappa} + (3/8)\tilde{\kappa}^5 + \tilde{\kappa}^{-7}\tilde{\mathcal{B}}^{abc}\tilde{\mathcal{B}}_{abc}\\ & + \tilde{\kappa}^{-7}\tilde{F}(\tilde{\phi}, \tilde{D}_a\tilde{\phi}, \tilde{D}_a\tilde{D}_b\tilde{\phi}, \tilde{\kappa}, \tilde{D}_a\tilde{\kappa}, \tilde{\psi}, \tilde{D}_a\tilde{\psi}, \tilde{D}_a\tilde{D}_b\tilde{\psi}),\end{aligned} \tag{2.20}$$

where $\tilde{\psi} = \frac{1}{4}\lambda^{-1}(\lambda^2 + \omega^2 + 1)$, and $\tilde{\mathcal{B}}_{abc} = \tilde{D}_{[a}(\tilde{\mathcal{R}}_{b]c} - \frac{1}{4}\tilde{h}_{b]c}\tilde{\mathcal{R}})$ is the Bach tensor of $\tilde{h}_{ab}$.[11] The function $\tilde{F}$ (which vanishes for static solutions) is a conformal invariant of dimensions $\sec^{-6}$ provided that $\tilde{\psi}$ is dimensionless.[12] Thus (2.20) is a conformally invariant equation, whose coefficients will be smooth at Λ if $\tilde{\phi}$ and $\tilde{\psi}$ are smooth there. Hence $\tilde{\kappa}$ will be smooth on $\tilde{S}$ if $\tilde{\phi}$ and $\tilde{\psi}$ are smooth.

The function $\tilde{\psi}$ satisfies the conformally invariant elliptic differential equation

$$\begin{aligned}\tilde{D}^m\tilde{D}_m\tilde{\psi} = {}& 6\tilde{\kappa}^4\tilde{\psi} + 2(\tilde{\phi}_M^2 + \tilde{\phi}_J^2)^{-1}\tilde{D}^m\tilde{\psi}(\tilde{\phi}_M\tilde{D}_m\tilde{\phi}_M + \tilde{\phi}_J\tilde{D}_m\tilde{\phi}_J)\\ & + 2\tilde{\psi}(4\tilde{\psi}^2 - 1)(\tilde{\phi}_M^2 + \tilde{\phi}_J^2)^{-2}(\tilde{\phi}_M\tilde{D}^m\tilde{\phi}_J\\ & - \tilde{\phi}_J\tilde{D}^m\tilde{\phi}_M)(\tilde{\phi}_M\tilde{D}_m\tilde{\phi}_J - \tilde{\phi}_J\tilde{D}_m\tilde{\phi}_M).\end{aligned} \tag{2.21}$$

The coefficients of (2.21) will be smooth on $\tilde{S}$ if $\tilde{\kappa}$ and $\tilde{\phi}$ are smooth. Thus, $\tilde{\psi}$ will be smooth if $\tilde{\phi}$ and $\tilde{\kappa}$ are smooth. But now our argument has become circular. The smoothness of $\tilde{\phi}$ follows from that of $\tilde{\kappa}$; the smoothness of $\tilde{\kappa}$ follows from that of $\tilde{\phi}$ and $\tilde{\psi}$; and the smoothness of $\tilde{\psi}$ follows from that of $\tilde{\phi}$ and $\tilde{\kappa}$.

Surprisingly enough, such a circular argument is precisely what is needed. Consider Eqs. (2.19), (2.20), and (2.21) as a set of coupled, second-order elliptic differential equations for $\tilde{\phi}, \tilde{\kappa}$, and $\tilde{\psi}$. If the coefficients of these equations are of class C^n, their solutions are either discontinuous at Λ or of class C^{n+2} (cf. Ref. 13). But $\tilde{\phi}, \tilde{\kappa}$, and $\tilde{\psi}$ are known to be continuous at Λ. It follows inductively that they are C^∞ functions at Λ.[14]

Thus Eqs. (2.7) indeed define multipole moments for the stationary gravitational field. The procedure for obtaining the moments is as follows: The conformally completed manifold $\tilde{S}$ is constructed from the manifold of trajectories, S. The potentials ϕ_M and ϕ_J then define smooth fields $\tilde{\phi}_M$ and $\tilde{\phi}_J$ on $\tilde{S}$, from which, via the recursive relations (2.7), the multipole moments $M_{a_1\cdots a_s}$ and $J_{a_1\cdots a_s}$ are calculated. In the next Sec., we shall use this procedure to calculate the multipole moments of the Kerr solution.

3. THE MULTIPOLE MOMENTS OF THE KERR SOLUTION

In recent years, there has been considerable interest in those solutions of Einstein's equation which might represent the exterior field of a collapsed object. In particular, it appears that the Kerr solution[15] is the only candidate for the exterior field of a stationary black hole.[16] We shall here calculate the multipole moments of the Kerr solution, using the definition of Sec. 2.

Certain simplifications occur for the multipole moments of any axisymmetric space-time. In that case, there exists a Killing vector η^a in S, i.e., a vector field satisfying

$$\pounds_\eta h_{ab} = 0, \qquad \pounds_\eta\phi_M = \pounds_\eta\phi_J = 0. \tag{3.1}$$

The conformal factor may be chosen so that $\tilde{\eta}^a$ $(= \eta^a)$ is also a Killing vector on $\tilde{S}$:

$$\pounds_{\tilde{\eta}}\tilde{h}_{ab} = 0; \tag{3.2}$$

and so that the axis vector

$$\tilde{z}^a = 2\tilde{\epsilon}^{abc}\tilde{D}_b\tilde{\eta}_c \tag{3.3}$$

is a unit vector at Λ:

$$\tilde{z}^a\tilde{z}_a\big|_\Lambda = 1. \tag{3.4}$$

The axis of the axial Killing vector passes through Λ. Thus, the action of the axial Killing vector defines rotations on the tensors at Λ, under which the multipole moments must be invariant. Since the only tensors at Λ invariant under the action of the axial Killing vector are outer products of the metric and the axis vector, the 2^s multipole moments are necessarily multiples of $\mathcal{C}[\tilde{z}_{a_1}\cdots\tilde{z}_{a_s}]\big|_\Lambda$, the symmetric, trace-free outer product of the axis vector with itself. Thus the 2^s moments are completely determined by the numbers M_s and J_s, defined by

$$\begin{aligned}M_s &= (s!)^{-1}M_{a_1\cdots a_s}\tilde{z}^{a_1}\cdots\tilde{z}^{a_s}\big|_\Lambda,\\ J_s &= (s!)^{-1}J_{a_1\cdots a_s}\tilde{z}^{a_1}\cdots\tilde{z}^{a_s}\big|_\Lambda.\end{aligned} \tag{3.5}$$

Stationary, axisymmetric solutions have, in addition to the axial Killing vector, a reflection symmetry, whose action reverses the sign of the axis vector and of ϕ_J, but leaves ϕ_M invariant. Thus the mass moments $M_{a_1\cdots a_s}$ are invariant under the action of this symmetry, and the angular momentum moments $J_{a_1\cdots a_s}$ are reversed in sign. But the moments are multiples of $\mathcal{C}[z_{a_1}\cdots z_{a_s}]|_\Lambda$, so only even mass and odd angular momentum moments will occur. Furthermore, the Kerr solution depends on only two parameters (m and a). Thus the multipole moments M_s and J_s are functions only of these two parameters:

$$M_s = M_s(m, a), \quad J_s = J_s(m, a). \tag{3.6}$$

We can compute the multipole moments of the Kerr solution directly from (2.7). In Boyer–Lindquist coordinates,[15] the spatial metric takes the form

$$\begin{aligned} d\sigma^2 = {} & (r^2 - 2mr + a^2\cos^2\theta)(r^2 - 2mr + a^2)^{-1}dr^2 \\ & + (r^2 - 2mr + a^2\cos^2\theta)d\theta^2 \\ & + (r^2 - 2mr + a^2)\sin^2\theta d\phi^2. \end{aligned} \tag{3.7}$$

The norm and twist of the timelike Killing vector are given by

$$\lambda = (r^2 - 2mr + a^2\cos^2\theta)(r^2 + a^2\cos^2\theta)^{-1} \tag{3.8}$$

and

$$\omega = 2ma\cos\theta(r^2 + a^2\cos^2\theta)^{-1}. \tag{3.9}$$

Defining a new radial variable $\bar{R}$ by

$$r = \bar{R}^{-1}[1 + m\bar{R} + \tfrac{1}{4}(m^2 - a^2)\bar{R}^2], \tag{3.10}$$

and choosing for the conformal factor

$$\Omega = \bar{R}^2\{[1 - \tfrac{1}{4}(m^2 - a^2)\bar{R}^2]^2 - a^2\bar{R}^2\sin^2\theta\}^{-1/2}, \tag{3.11}$$

we obtain for the conformally transformed metric

$$\begin{aligned} d\tilde{\sigma}^2 = {} & d\bar{R}^2 + \bar{R}^2 d\theta^2 + \{1 - a^2\bar{R}^2\sin^2\theta[1 \\ & - \tfrac{1}{4}(m^2 - a^2)\bar{R}^2]^{-2}\}^{-1}\bar{R}^2\sin^2\theta d\phi^2. \end{aligned} \tag{3.12}$$

It is easily verified (by introducing Cartesian coordinates, for example), that this metric and conformal factor satisfy conditions (i)–(iii) for asymptotic flatness, with Λ at the point $\bar{R} = 0$. The fields $\tilde{\phi}_M$ and $\tilde{\phi}_J$ are given by

$$\begin{aligned} \tilde{\phi}_M = {} & -m[1 + \tfrac{1}{4}(m^2 - a^2)\bar{R}^2]\{[1 - \tfrac{1}{4}(m^2 - a^2)\bar{R}^2]^2 \\ & - a^2\bar{R}^2\sin^2\theta\}^{-3/4} \end{aligned} \tag{3.13}$$

$$\tilde{\phi}_J = ma\bar{R}\cos\theta\{[1 - \tfrac{1}{4}(m^2 - a^2)\bar{R}^2]^2 - a^2\bar{R}^2\sin^2\theta\}^{-3/4}.$$

Substituting (3.13) into Eq. (2.7), we obtain[17] for the multipole moments of the Kerr solution

$$\begin{aligned} & M_{2s} = (-1)^{s+1}ma^{2s}, \quad M_{2s+1} = 0, \\ & J_{2s} = 0, \quad J_{2s+1} = (-1)^s ma^{2s+1}. \end{aligned} \tag{3.14}$$

Hernandez[18] has given these multipole moments in the weak-field approximation, where the problem reduces to finding the moments of a Newtonian potential in flat space. Since our moments turn out to be linear in m, one might expect to obtain the same result in this case from the weak-field approximation. Our moments also agree with the formulas given by Newman and Janis[19] for the lowest moments of the Kerr solution. The agreement here is more surprising, since the definition of multipole moments used by Newman and Janis involves fields at null infinity.

4. DISCUSSION AND CONCLUSIONS

In this section, we shall discuss some features of the potentials ϕ_M and ϕ_J, and suggest some possible future developments for multipole moments in general relativity.

We shall first show that, in an appropriate limit, ϕ_M and ϕ_J reduce to Newtonian potentials in a Euclidean 3-space. Consider a one-parameter family $M, g_{ab}(\mu)$ of stationary solutions of Einstein's equation (without sources), such that for $\mu = 0, g_{ab}(0)$ is a flat metric on M. Without loss of generality, we may take the Killing vector ξ^a to be independent of μ.[20] Then S is also independent of μ as a manifold, but its metric $h_{ab}(\mu)$ is not. Of course, the fields ϕ_M and ϕ_J in general depend on μ.

If S is to be asymptotically flat, the Killing vector ξ^a must reduce, in the flat-space limit, to a vector field which generates time translations in Minkowski space; hence we have $\phi_M(0) = 0$, $\phi_J(0) = 0$, and $h_{ab}(0)$ is a Euclidean metric. Then Eqs. (2.16) and (2.17), written to first order in μ, become

$$\begin{aligned} & D^m D_m \dot{\phi}_M = 0, \\ & D^m D_m \dot{\phi}_J = 0, \end{aligned} \tag{4.1}$$

and

$$\dot{R}_{ab} = 0, \tag{4.2}$$

where $\dot{T}^{a\cdots b}{}_{c\cdots d}$ is used to denote $d/d\mu(T^{a\cdots b}{}_{c\cdots d}(\mu))$, and where all quantities appearing in (4.1) and (4.2) are evaluated at $\mu = 0$. Note that Eqs. (4.1) are simply Laplace's equation in a Euclidean 3-space. Thus the functions ϕ_M and ϕ_J reduce, in the linearized theory, to Newtonian potentials.[21]

Our choices for the mass and angular momentum potentials may still seem rather arbitrary. Certainly any member of a large class of fields could have been used to define the multipole moments of the space–time. In fact, another possible candidate (in the static case) for the mass potential was given by Geroch.[2] However, this function does not define a smooth field on $\tilde{S}$ for the case in which the Killing vector is not hypersurface-orthogonal; nor does any obvious generalization of this function. Another possible choice, for example, for the mass potential would be the function $(\phi_M^2 + \phi_J^2)^{1/2}$. The smoothness of this field, with dimensions $\sec^{-1/2}$, follows directly from the smoothness of ϕ_M and ϕ_J. Furthermore, this function has the same Newtonian limit as ϕ_M.[21]

Thus the potentials we have used in defining the multipole moments are by no means unique. This ambiguity in the choice of the potentials reflects the ambiguity in the passage from Newtonian theory to general relativity. We have, in fact, taken advantage of this freedom to choose potentials which emphasize the analogy between stationary gravitational fields and stationary electromagnetic fields. The fact that ϕ_M and ϕ_J satisfy the same Eq. (2.16) suggests that mass and angular momentum may be treated "on the same footing" as sources of the gravitational field. In this respect, stationary gravitational fields can be considered analogous to stationary Maxwell fields, in which both the electric and magnetic scalar potentials satisfy Laplace's equation.

A notation arising from the fact that ϕ_M and ϕ_J satisfy the same equation is developed in the Appendix. This notation casts Eqs. (2.16) and (2.17) into a particularly simple form, and considerably reduces the work of performing computations.

Although the presence of curvature introduces ambiguities in the definition of the mass and angular momentum potentials, the lowest-order behavior of the fields is unaffected by curvature. For example, curvature cannot create monopole angular momentum fields in asymptotically flat space-times. We shall sketch a proof that this is the case.

One of the field equations, in the presence of matter, takes the form[3]

$$D^m(\lambda^{-2}\omega_m) = 0, \tag{4.3}$$

so that the expression

$$\int_S \lambda^{-2}\omega_m\, dS^m, \tag{4.4}$$

evaluated over any 2-surface in S enclosing the sources, vanishes. The surface S lies in a region free of sources, so that ω_a is a gradient. The expression (4.4) may then be written entirely in terms of quantities on $\tilde{S}$, i.e., in terms of $\tilde{\phi}_M$, $\tilde{\phi}_J$, $\tilde{h}_{ab}$, and Ω. Shrinking S onto the point Λ, one sees that the vanishing of (4.4) implies the vanishing of the angular momentum monopole:

$$\tilde{\phi}_J \Big|_\Lambda = 0. \tag{4.5}$$

Thus, the presence of curvature does not affect the lowest-order behavior of the angular momentum field.

A number of questions about the multipole moments of stationary space-times remain unanswered. For example, Geroch's[2] conjecture that the multipole moments of a static space-time uniquely determine its structure (at least in some neighborhood of Λ) may also be made for stationary space-times, and remains an open question.

One might also ask if the moments can be expressed in terms of integrals over the source distributions of the stationary field. Whether this can be done is not at all clear, for the multipole moments defined here are quantities "at infinity", while the sources of the field all lie in bounded regions of S. The curvature of S prevents us from transporting the multipole moments to finite regions. Thus the problem of expressing the multipole moments as integrals over the source distribution appears to present serious difficulties.

Can the definition of multipole moments be extended to more general contexts, e.g., to solutions of Einstein's equation admitting no Killing vectors? A straightforward generalization of our approach seems unlikely to succeed. The conformal completion of spacelike surfaces in an arbitrary, asymptotically flat space-time has only a C^0 structure.[22] This lack of sufficient smoothness structure appears to preclude the introduction of multipole moments along the lines we have followed, i.e., as the values at infinity of the derivatives of some potentials. In fact, multipole moments at spatial infinity apparently do not exist, even for electromagnetic fields, except when the field is stationary. A better understanding of the multipole moments of the electromagnetic field would probably shed light on the gravitational case.

The possibility of a definition of multipole moments for nonstationary gravitational fields as quantities at null infinity appears more promising. One definition of this type has been discussed by Janis and Newman.[23] Of course, these moments are in general time-dependent, since radiation escaping from the system is registered at null infinity. It would be interesting to attempt to reformulate the construction of Janis and Newman along the more geometrical lines we have followed here.

Perhaps multipole moments could be introduced for stationary, asymptotically flat Einstein-Maxwell fields. The analogy between stationary gravitational fields and stationary electromagnetic fields discussed above suggests this is likely to be the case. One would expect to obtain four sets of moments, corresponding to the mass, angular momentum, and the electric and magnetic potentials. The formalism given by Geroch[3] can be used to write the field equations for stationary Einstein-Maxwell fields in the form

$$\begin{aligned}
D^m D_m\epsilon &= \lambda^{-1}[(D^m\lambda)(D_m\epsilon) - \omega^m D_m\beta],\\
D^m D_m\beta &= \lambda^{-1}[(D^m\lambda)(D_m\beta) + \omega^m D_m\epsilon],\\
D^m D_m\lambda &= \lambda^{-1}[(D^m\lambda)(D_m\lambda) - \omega^m\omega_m] + [(D^m\epsilon)(D_m\epsilon)\\
&\qquad + (D^m\beta)(D_m\beta)],\\
D^m\omega_m &= 2\lambda^{-1}(D^m\lambda)\omega_m,\\
D_{[a}\omega_{b]} &= 2(D_{[a}\epsilon)(D_{b]}\beta),\\
\mathcal{R}_{ab} &= \tfrac{1}{2}\lambda^{-2}[(D_a\lambda)(D_b\lambda) + \omega_a\omega_b] - \lambda^{-1}[(D_a\epsilon)(D_b\epsilon)\\
&\qquad + (D_a\beta)(D_b\beta)],
\end{aligned} \tag{4.6}$$

where ϵ and β are electric and magnetic scalar potentials. The twist vector ω_a of the timelike Killing vector field is now no longer curl-free, but a corrected version, given by $\omega_a - (\epsilon D_a\beta - \beta D_a\epsilon)$, is. The idea is to find combinations of these fields, which, like ϕ_M and ϕ_J in the source-free case, define smooth fields on $\tilde{S}$. Equations (2.7) would then yield, from these four potentials, the electromagnetic and gravitational multipole moments. More generally, one might try to introduce multipole moments for the stationary gravitational field coupled with spin s massless fields.

ACKNOWLEDGMENTS

I wish to express my deep gratitude to R. Geroch, my thesis advisor, for his guidance and encouragement during the course of this research. I also wish to thank S. Chandrasekhar for his comments on an earlier version of this paper, and on the multipole moments of the Kerr solution; and J. Winicour for a remark which led to the arguments surrounding Equations (2.19), (2.20), and (2.21). This research was supported in part by National Science Foundation Grant number GP-34721X with the University of Chicago.

APPENDIX: NOTATION

The fact that arbitrary linear combinations of ϕ_M and ϕ_J satisfy Eq. (2.16), and the fact that only certain simple combinations of these fields occur in our equations, suggest the introduction of a notation which formalizes these features. In practice, this notation considerably reduces the labor in performing computations.

It can be verified that the solutions of

$$(D^m D_m - \mathcal{R}/8)f = (15/8)\kappa^4 f \tag{A1}$$

which are functions only of λ and ω form a three-dimensional vector space V. The potentials ϕ_M and ϕ_J span a two-dimensional linear subspace of V. Consider V^* (the

dual space of V), the space of linear mappings from V to R. Define tensors over V and V^* as the multilinear mappings from V and V^* to R. It is convenient to denote the elements of V by objects with a Greek superscript, e.g., v^α; and the elements of V^* by objects with a Greek subscript, e.g., μ_α. Tensors of arbitrary rank are then given the appropriate number of Greek subscripts and superscripts to indicate their rank as multilinear mappings over V and V^*. For example, a tensor of second rank over V may be written $T_{\alpha\beta}$. We adopt the usual index conventions for Greek indices. Index substitution and contraction, for example, have their usual meanings.

We define objects with both Greek and Latin indices as the multilinear mappings from V and V^* to tensor fields on S. For example, if ξ^α is an element of V and η_α an element of V^*, we may write

$$T^a{}_{b\alpha}{}^\beta \xi^\alpha \eta_\beta = \lambda^a{}_b,$$

where $\lambda^a{}_b$ is a tensor field on S. In particular, an object with no Latin indices is now a multilinear mapping from V and V^* to scalar fields on S. We define the action of the derivative operator D_a on objects with both Greek and Latin indices in terms of its action when the object is contracted with arbitrary elements of V and V^*, e.g.,

$$\eta_\alpha \xi^\beta D_a T_{bc}{}^\alpha{}_\beta = D_a(T_{bc}{}^\alpha{}_\beta \eta_\alpha \xi^\beta)$$

for all η_α in V^* and ξ^α in V.

We next introduce ϕ_α, a distinguished member of this algebra. Let f be a solution of (A1), and let f^α be the corresponding element of V. Define ϕ_α by

$$\phi_\alpha f^\alpha = f \tag{A2}$$

for all f. It is convenient to introduce a metric $G_{\alpha\beta}$ of signature $(-,+,+)$ on this algebra, such that

$$G^{\alpha\beta}\phi_\alpha\phi_\beta = -1 \tag{A3}$$

and

$$D_a G_{\alpha\beta} = 0, \tag{A4}$$

where $G^{\alpha\beta}$ is the inverse of $G_{\alpha\beta}$.

Equations (2.16) and (2.17) now take the form

$$\begin{aligned} &D^m D_m \phi_\alpha = 2\mathcal{R}\phi_\alpha, \\ &\mathcal{R}_{ab} = \tfrac{1}{2}(D_a\phi^\alpha)(D_b\phi_\alpha), \end{aligned} \tag{A5}$$

where Greek indices are raised and lowered using $G_{\alpha\beta}$. Computations are conveniently carried out using this notation. For example, the function F which appears in Eq. (2.20) takes the form[24]

$$\begin{aligned} F = {} & (D^c\phi^\beta D_c\phi_\beta)(D^aD^b\phi^\alpha D_aD_b\phi_\alpha) \\ & - (D^c\phi^\alpha D_c\phi^\beta)(D^aD^b\phi_\alpha D_aD_b\phi_\beta) \\ & + (D_b\phi^\beta D^aD^b\phi_\alpha)(D^c\phi^\alpha D_aD_c\phi_\beta) \\ & - (D_b\phi^\alpha D^aD^b\phi_\alpha)(D^c\phi^\beta D_aD_c\phi_\beta) \\ & + \tfrac{1}{2}(D^a\phi^\alpha D_a\phi_\alpha)^3 \\ & + \tfrac{1}{2}(D^a\phi^\alpha D_a\phi_\alpha)(D^b\phi^\beta D_b\phi^\gamma)(D^c\phi_\beta D_c\phi_\gamma). \end{aligned} \tag{A6}$$

Recall that ϕ_M and ϕ_J span a two-dimensional subspace of V. In fact, a basis may be introduced in V, in terms of which the components of ϕ_α are (ψ, ϕ_M, ϕ_J). Furthermore, Eqs. (A5) are now seen to be invariant under "Lorentz transformations" in V. Thus, if $F_{\alpha\beta}$ satisfies

$$F_{\alpha\beta} = F_{[\alpha\beta]}, \quad F_{\alpha\gamma}F^{\beta\gamma} = \delta_\alpha{}^\beta, \quad D_a F_{\alpha\beta} = 0, \tag{A7}$$

where $\delta_\alpha{}^\beta$ is the unit tensor over V and V^*, and if we define

$$\phi'_\alpha = F_\alpha{}^\beta \phi_\beta, \tag{A8}$$

then Eqs. (A5), written in terms of ϕ'_α, take the form

$$\begin{aligned} &D^m D_m \phi'_\alpha = 2\mathcal{R}\phi'_\alpha, \\ &\mathcal{R}_{ab} = (D_a\phi'^\alpha)(D_b\phi'_\alpha). \end{aligned} \tag{A9}$$

Hence, application of this Lorentz transformation leads to a new stationary solution of the vacuum Einstein equation.

In fact, the Lorentz transformations on V are precisely the transformations discussed by Geroch in Ref. 3. The $SO(2)$ subgroup of those transformations which cannot be reduced to pure gauge may be realized as the spatial rotations in V, i.e., rotations in the (ϕ_M, ϕ_J) plane. We are thus led to interpret the transformations given by Geroch, in the case of a timelike Killing vector field, as rotations between the mass and angular momentum potentials, analogous to the duality rotations of stationary Maxwell fields.

Unfortunately, except in the case of vanishing mass, at most one of the family of solutions obtained in this way can be asymptotically flat. Under such a rotation, the angular momentum potential will, in general, acquire a monopole component. But, as we have seen, monopole angular momentum fields cannot occur in asymptotically flat spaces. Thus, the transformations (A8) apparently do not preserve the asymptotic flatness of the space-time.

* This research has been presented as a thesis to the Department of Physics, University of Chicago, in partial fulfillment of the requirements for the Ph.D. degree.

† Address as of September, 1973: Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213.

[1] R. Geroch, J. Math. Phys. 11, 1955 (1970).
[2] R. Geroch, J. Math. Phys. 11, 2580 (1970).
[3] R. Geroch, J. Math. Phys. 12, 918 (1971).

[4] In the Newtonian limit, the angular momentum potential does not affect the motion of masses; as a result, it never appears in Newtonian theory. The situation is analogous to nonrelativistic electrodynamics, in which the magnetic interactions between charges are of order $(v/c)^2$.

[5] We shall be concerned with regions of M distant from matter, and will assume that the space-time is asymptotically flat. Then we may excise from M a closed subset surrounding the sources, such that ω is well defined outside this subset.

[6] This metric, denoted by $\tilde{h}_{ab}$ in Ref. 3, is related by a factor of λ, "the redshift with respect to infinity," to the metric obtained by projecting g_{ab} into S. Thus h_{ab} is the metric obtained on S by a field of observers who scale their measuring instruments to agree with the interval between pulses of light emitted "at infinity" with a prearranged frequency.

[7] More precisely, x^a is the position vector (field) with respect to some origin: that (unique) vector field in Euclidean space which satisfies $D_a x_b = h_{ab}$, where h_{ab} is the Euclidean metric, and which vanishes at the origin. The vector field $\bar{x}^a$ bears the same relationship to $\tilde{h}_{ab}$, which is also a Euclidean metric.

[8] Note, however, that the multipole moments defined by Eqs. (2.7) are tensors at Λ, while the usual Newtonian moments are constant tensor fields on Euclidean 3-space. Furthermore, we have not yet shown that the fields ϕ_M and ϕ_J reduce to Newtonian potentials in the appropriate limit. This point will be discussed in Sec. 4.

[9] These equations are written more compactly in Eqs. (A7).

[10] The *dimensions* of a tensor are said to be $\sec^n$, where n is the following number: [power of Ω by which that tensor is to be multiplied under

conformal transformations] − [number of covariant indices] + [number of contravariant indices]. Thus, the metric tensor is dimensionless, and the dimension of a tensor is unchanged by the raising and lowering of indices and by contraction.

[11]The Bach tensor plays, in three dimensions, the role of the Weyl tensor in higher dimensions: its vanishing is necessary and sufficient for conformal flatness. The only property we need, however, is that it is a conformally invariant tensor of dimensions sec^{-3}. This fact may be verified directly from the conformal behavior of $\mathcal{R}_{ab}$.

[12]F is written explicitly in Eq. (A6), using a more compact notation.

[13]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II, p. 336.

[14]The induction may be carried out as follows: Assume that $\tilde{\phi}$ and $\tilde{\psi}$ are C^{n+1}, and that $\tilde{\kappa}$ is C^n. Then the coefficients of (2.19) and (2.21) are C^n, and the coefficients of (2.20) are C^{n-1}. Hence $\tilde{\phi}$ and $\tilde{\psi}$ are C^{n+2} and $\tilde{\kappa}$ is C^{n+1}. The dependence of $\tilde{F}$ on $\tilde{D}_a\tilde{\kappa}$ is linear, as required by the theorem of Ref. 13. Note that if the mass of the system vanishes, i.e., $\tilde{\phi}_{M\hat{\Lambda}} = 0$, the arguments presented here fail; in that case, a more delicate analysis is required.

[15]R. P. Kerr, Phys. Rev. Letters **11**, 237 (1963); R. H. Boyer and R. W. Lindquist, J. Math. Phys. **8**, 265 (1967).

[16]W. Israel, Phys. Rev. **164**, 1776 (1967); B. Carter, Phys. Rev. **174**, 1559 (1968); B. Carter, Phys. Rev. Letters **26**, 331 (1971); S. W. Hawking, Commun. Math. Phys. **25**, 152 (1972).

[17]Although the substitution of ϕ_M and ϕ_J into Eqs. (2.17) is straightforward, the resulting expressions are rather complicated. Some features of the Kerr solution which lead to a more convenient method of calculating the moments will be discussed in a forthcoming paper.

[18]W. C. Hernandez, Phys. Rev. **159**, 1070 (1967).

[19]E. T. Newman and A. I. Janis, J. Math. Phys. **6**, 915 (1965).

[20]There always exists a gauge transformation which takes $\xi^a(\mu)$ to $\xi^a(0)$. Thus, nothing is lost in assuming that ξ^a is a fixed vector field.

[21]To obtain the Newtonian limit, note that, for a system of interacting masses, the angular momentum potential will be of order v/c (cf. the field equations of Ref. 3). Thus, in the limit of small velocities, we have $\phi_J \ll \phi_M$.

[22]R. Geroch, J. Math. Phys. **13**, 956 (1972).

[23]A. I. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965).

[24]Because the notation is not defined in a conformally invariant way, the conformal invariance of F is not manifest when it is written in this notation. A compact notation which is also conformally invariant would be very helpful.

Classical $SU(3)$ gauge field equations*

Alfred C. T. Wu

Department of Physics, University of Michigan, Ann Arbor, Michigan 48104

Tai Tsun Wu

Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts 02138
(Received 14 August 1973)

Admissible forms of the static solutions to the $SU(3)$ gauge field equation are examined. It is shown that by a proper choice of the form of solutions which extricate the $SU(3)$ indices, the set of nonlinear *partial* differential equations is reducible to nonlinear *ordinary* differential equations for the radial functions.

I. INTRODUCTION

In a previous paper coauthored by one of us,[1] some static solutions of the classical $SU(2)$ isotopic gauge field[2] equations were discussed. A crucial feature is the fact that by a judicious choice of the form of solutions which properly extricate the isotopic indices, the set of nonlinear *partial* differential equations is reducible to nonlinear *ordinary* differential equations for the radial functions.

The purpose of the present note is to examine the static case of $SU(3)$ unitary gauge field[3] equations. We find that the above feature also holds for the $SU(3)$ gauge field equations.

II. LOCAL $SU(3)$ GAUGE FIELD

As is well known, the number of the gauge field components is equal to the dimension of the regular representation of the underlying group. For $SU(3)$, this number is eight. The octet gauge field may be arranged in terms of 3×3 matrices.

$$C_\mu = \kappa c_\mu^A \lambda_A, \quad \text{summed over } A = 1, .., 8, \tag{1}$$

where κ is a scale factor and λ_A are the set of eight 3×3 Gell-Mann matrices[4] satisfying the commutation relation

$$[\lambda_A, \lambda_B] = ig_{AB}^C \lambda_C. \tag{2}$$

Let

$$F_{\mu\nu} = \kappa f_{\mu\nu}^A \lambda_A, \tag{3}$$

where

$$f_{\mu\nu}^A = c_{\mu,\nu}^A - c_{\nu,\mu}^A - \kappa\epsilon g_{BE}^A c_\mu^D c_\nu^E, \tag{4}$$

in which ϵ is the coupling constant and the comma with respect to μ, ν is a short-hand notation for the differentiation

$$c_{\mu,\nu}^A \equiv \partial_\nu c_\mu^A. \tag{5}$$

Away from sources, we take as the free Lagrangian

$$\mathcal{L}_0 = -\tfrac{1}{4} f_{\mu\nu A} f_{\mu\nu}^A. \tag{6}$$

The equation of motion reads

$$f_{\mu\nu,\nu A} = -\kappa\epsilon g_{AB}^D c_\nu^B f_{\mu\nu D}. \tag{7}$$

Substitution of Eq. (4) then gives

$$c_{\mu,\nu\nu}^A + \kappa\epsilon g_{BD}^A (2c_{\mu,\nu}^D - c_{\nu,\mu}^D) c_\nu^B - (\kappa\epsilon)^2 g_{BD}^A g_{EF}^D c_\nu^B c_\mu^E c_\nu^F = 0. \tag{8}$$

From here on, we choose the scale so that

$$\kappa\epsilon = 1. \tag{9}$$

III. STATIC CASE

We shall primarily be interested in the static situation where

$$c_4^A = 0, \tag{10a}$$

$$c_{i,4}^A = 0. \tag{10b}$$

Thus the only nonvanishing components are the spatial ones.

Furthermore, the subsidiary condition holds:

$$c_{\mu,\mu}^A = 0. \tag{11}$$

Instead of the label $A = 1, \ldots, 8$, we find it convenient to adopt double indices lm, each running from 1 to 3. The traceless condition would give us still eight independent components D_μ^{lm}. Under such a correspondence, we make the following transcription.

For the indices:

$$A \to lm, B \to pq, D \to rs, E \to uv, F \to xy; \tag{12}$$

for the field

$$c_\mu^A \to D_\mu^{lm}; \tag{13}$$

and for the structure constants

$$g_{ABD} \to \mathcal{G}_{lm,pq,rs}. \tag{14}$$

Explicitly, the structure constants read

$$\mathcal{G}_{lm,pq,rs} = \delta_{mp}\delta_{qr}\delta_{sl} - \delta_{lq}\delta_{mr}\delta_{ps}. \tag{15}$$

Equation (8) becomes then for the static case

$$D_{i,jj}^{lm} + \mathcal{G}_{pq,rs}^{lm}(2D_{i,j}^{rs} - D_{j,i}^{rs})D_j^{pq} - \mathcal{G}_{pq,rs}^{lm}\mathcal{G}_{uv,xy}^{rs} D_j^{pq} D_i^{uv} D_j^{xy} = 0, \tag{16}$$

where the comma with respect to the spatial components i or j *only* denotes a differentiation.

IV. FORM OF STATIC SOLUTIONS

We look for the static solutions to the field equations (16) in the following form:

$$D_i^{lm} = (\epsilon_{ilj} x_j x_m + \epsilon_{imj} x_j x_l) f + \epsilon_{lmk}\epsilon_{kin} x_n h, \tag{17}$$

where f and h are two functions of the radial distance r alone, $r = (x_1^2 + x_2^2 + x_3^2)^{1/2}$.

It should be emphasized that for Eq. (16) to be satisfied for every internal index, severe restrictions must prevail on the admissible forms of the solutions. It is rather remarkable that the choice (17) indeed gives a consistent solution. In other words, the symmetric combination $\epsilon x x$ (in front of f) and the antisymmetric combination $\epsilon\epsilon x$ (in front of h) become jointly preserved under the operations indicated on the left-hand side of (16). After a straightforward calculation, their coefficients can be collected. The vanishing of these coefficients gives a pair of coupled nonlinear ordinary differential equations for f and h. The result is

$$f'' + 6r^{-1}f' - 14fh + 7r^2fh^2 - r^4f^3 = 0, \tag{18a}$$

$$h'' + 4r^{-1}h' + 7r^2f^2 - 3h^2 - 7r^4f^2h + r^2h^3 = 0. \tag{18b}$$

Or, in terms of a (F, H) pair defined as

$$F \equiv r^3 f, \tag{19a}$$

$$H \equiv 1 - r^2 h, \tag{19b}$$

Eq. (18) reads

$$F'' + r^{-2}F(-13 + 7H^2 - F^2) = 0, \tag{20a}$$

$$H'' + r^{-2}[4 - (5 + 7F^2)H + H^3] = 0. \tag{20b}$$

We shall not attempt to discuss the solutions to Eqs. (20) here except by making the following obvious remarks.

(i) Equations (20) possess three real singular points located at

$$\begin{pmatrix} F \\ H \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ \frac{1}{2}(-1 \pm \sqrt{17}) \end{pmatrix}. \tag{21}$$

(ii) In (17), the f part solutions cannot exist by itself, while the h part can, i.e., when $h = 0, f = 0$; however, when $f = 0, h$ has nonzero solutions.

* Work supported in part by the U.S. Atomic Energy Commission.

[1] T. T. Wu and C. N. Yang, in *Properties of Matter under unusual Conditions* (in honor of E. Teller's 60th birthday), edited by H. Mark and S. Fernbach (Interscience, New York, 1969), pp. 349–354.
[2] C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).
[3] For extensions of Yang–Mills formalism to higher rank groups, see e.g., R. Utiyama, Phys. Rev. **101**, 1597 (1956); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961); and others.
[4] M. Gell-Mann, Caltech Report No. CTSL-20 (1961); Phys. Rev. **125**, 1067 (1962); see collection in M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (Benjamin, New York, 1964).

Exact calculation of the energy and heat capacity for the triangular lattice with three different coupling constants

Gordon W. Cunningham

Federal Systems Center, IBM, Gaithersburg, Maryland 20760

Paul H. E. Meijer

Catholic University of America, Washington, D.C. 20017

National Bureau of Standards, Washington, D.C. 20234

(Received 19 July 1973)

Calculation of the internal energy and heat capacity of the general anisotropic triangular Ising lattice is derived from the double integral form of the partition function. The principal result is the reduction of the elliptic integrals to a standard form for three arbitrary coupling constants. Both the standard form and the method of reduction are due to Legendre. The method of reduction is one involving two linear transformations. A straightforward reduction of elliptic integrals to standard form could not be used in this application. This is because of the functional dependence of the two linear transformations on the many combinations and permutations of the signs and relative magnitudes of the coupling energies of the lattice. A relatively simple formulation is presented in which the many combinations and permutations previously mentioned are reduced to only two distinct cases. An independent numerical solution was calculated directly from the partition function as a means of verifying the formulation presented in this paper.

I. INTRODUCTION

Several independent solutions for the energy and heat capacity of the triangular Ising lattice appeared in 1950. Wannier[1] and independently Husimi and Syozi[2] solved the isotropic case. Houtappel,[3] Newell,[4] and Temperley[5] solved the general problem of anisotropic coupling (except for the final calculation of the thermodynamic quantities) in which the nearest neighbor coupling energies J_i in each of the three principal directions in the lattice are distinct (i.e., $J_1 \neq J_2 \neq J_3$). In doing so, all three derived a double integral expression for the partition function in the general anisotropic case. Houtappel[3] and Temperley[5] then took the isotropic case ($J_1 = J_2 = J_3$) in formulating the energy and heat capacity, while Newell[4] presented the corresponding results for a particular anisotropic case in which two coupling energies are equal and the third is independent. This paper presents the general anisotropic formulation of energy and heat capacity.

The principal activity in the development is the reduction of elliptic integrals to standard forms. The standard forms are those of Legendre. The particular reduction used (the reduction not being unique[6]) is one introduced by Legendre and further developed by Cayley[7] involving two linear transformations. Although the reduction of elliptic integrals to standard form is well-known and is usually a straightforward matter, the case here is tedious and complicated by the large number of possible combinations and permutations of the signs and relative magnitudes of the three coupling energies, and the influence they have on both the choice of the linear transformations and the specific representation of the elliptic integrals in standard form. The probability of error or of an unnecessarily cumbersome resultant formulation is very high, especially when compared to the much simpler isotropic case in which incidents of both have been noted in the literature.[8-10] Therefore, the analytic development presented in this paper was verified in the same manner as most complicated analytic integration formulas: an independent numerical solution was made for comparison. In this independent solution, the energy and heat capacity were computed directly from the integral form of the partition function via numerical quadrature.

II. FORMULATION

The starting point of the formulation is the integral expression for the partition function of the general anisotropic triangular Ising lattice; [1,3,4]

$$\ln\left(\frac{\lambda}{2}\right) = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{2\pi} \ln[\cosh 2K_1 \cosh 2K_2 \cosh 2K_3 + \sinh 2K_1 \sinh 2K_2 \sinh 2K_3 - \sinh 2K_1 \cos\omega_1 - \sinh 2K_2 \cos\omega_2 - \sinh 2K_3 \cos(\omega_1 + \omega_2)]\, d\omega_1 d\omega_2, \tag{1}$$

where $K_i = J_i/kT$, J_i is the coupling energy in the ith principal direction in the lattice, k is Boltzmann's constant, and T is the absolute temperature. The desired results are closed-form expressions for (i) the expectation value of the normalized lattice energy and (ii) the lattice heat capacity. These quantities are given below as derivatives of the logarithm of the partition function:

$$\langle E \rangle = -\sum_{j=1}^{3} J_i \frac{\partial \ln\lambda}{\partial K_j}, \tag{2}$$

$$C = k\sum_{i=1}^{3}\sum_{j=1}^{3} K_i K_j \frac{\partial^2 \ln\lambda}{\partial K_i K_j}. \tag{3}$$

Expression (1) cannot be completely integrated in closed form.[1,3,4] Therefore, after some changes of variables, rearrangements, and integrations, expression (2) will be applied, thus allowing a complete closed-form integration. Let

$$a = \cosh 2K_1 \cosh 2K_2 \cosh 2K_3 + \sinh 2K_1 \sinh 2K_2 \sinh 2K_3, \tag{4}$$

$$b = -\sinh 2K_1, \tag{5}$$

$$c = -\sinh 2K_2, \tag{6}$$

$$d = -\sinh 2K_3. \tag{7}$$

Expression (1) becomes, after the changes of variables,

rearrangements, and integrations,

$$\ln\left(\frac{\lambda}{2}\right) = \frac{1}{8\pi^2}\left[\int_0^{2\pi}\int_0^{2\pi} \ln(a + b\cos\omega_1)d\omega_1 d\omega_2 + \int_0^{2\pi}\int_0^{2\pi} \ln(1 + u\sin\omega_2 + v\cos\omega_2)d\omega_1 d\omega_2\right], \quad (8)$$

where

$$u = -d\sin\omega_1/(a + b\cos\omega_1),$$

$$v = (c + d\cos\omega_1)/(a + b\cos\omega_1).$$

The first term integrates immediately for ω_2 and from tables for ω_1. The integration over ω_2 in the second term is listed in the integral tables of Bierens DeHaan.[11] Applying (2) to the result of these integrations and simplifying, we find

$$\langle E\rangle = -\sum_{i=1}^{3} J_i\left[\partial_i \ln(a + \sqrt{a - b^2})^{1/2} - \frac{1}{8\pi}\int_0^{2\pi}\frac{(g\partial_i f - f\partial_i g)}{fg}d\omega + \frac{1}{8\pi}\int_0^{2\pi}\frac{(g\partial_i f - f\partial_i g)}{fg\sqrt{1 - g/f}}d\omega\right], \quad (9)$$

where the subscript has been dropped from ω_1, and where

$$\partial_i \equiv \frac{\partial}{\partial K_i},$$

$$f = (a + b\cos\omega)^2, \quad (10)$$

$$g = c^2 + d^2 + 2cd\cos\omega. \quad (11)$$

The first integral in (9) is an elementary form. The second integral is a combination of elliptic integrals. Before going into their reduction to standard form, the first integral is combined with the first term and simplified:

$$-J_l \coth 2K_l,$$

where

$$l = \begin{cases} 2, & J_2^2 > J_3^2 \\ 3, & J_2^2 < J_3^2. \end{cases} \quad (12)$$

Rewriting the second integral by changing integration variables, noting that the integrand is an even function about $\omega = \pi$, and substituting for f and g from (10) and (11), we have

$$\langle E\rangle = -J_l \coth 2K_l - \sum_{i=1}^{3} J_i\left(\frac{1}{2\pi}\int_{-1}^{1}\frac{\partial_i(a + bx)}{\Delta}dx - \frac{1}{2\pi}\int_{-1}^{1}\frac{[bx^2\partial_i(cd) + [a\partial_i(cd) + b\partial_i\tfrac{1}{2}(c^2 + d^2)]x + a\partial_i\tfrac{1}{2}(c^2 + d^2)]dx}{\Delta(c^2 + d^2 + 2cdx)}\right), \quad (13)$$

where

$$\Delta = [(1 - x^2)(b^2x^2 + 2(ab - cd)x + a^2 - c^2 - d^2)]^{1/2} \quad (14)$$

and l is as defined in (12).

The reduction of the elliptic integrals in (13) will now be made and will follow the method of Legendre as presented and developed further by Cayley.[7] The elliptic integrals in (13) are of the general form

$$\int\frac{R(x)dx}{\sqrt{X}}, \quad (15)$$

where $R(x)$ is a real rational function of x, and X is a real quartic function with real coefficients of the form

$$X = (\zeta + 2\eta x + \theta x^2)(\lambda + 2\mu x + \nu x^2). \quad (16)$$

Two linear real transformations are involved. In the first, x is replaced by

$$\frac{p + qx}{1 + x}. \quad (17)$$

In the second, x^2 is replaced by

$$\frac{A + Bx^2}{C + Dx^2}. \quad (18)$$

Application of (17) to (15) reduces it to

$$\int\frac{R'dx}{[\pm(1 \pm mx^2)(1 \pm nx^2)]^{1/2}} \quad (19)$$

and application of (18) to (19) further reduces it to standard form

$$\int\frac{R''dx}{[(1 - x^2)(1 - k^2x^2)]^{1/2}}, \quad (20)$$

where R' and R'' are the singly and doubly transformed rational functions of x, respectively, m and n are real, positive numbers; and the modulus k is a real function of m and n such that $k^2 < 1$.

Transformation (17) is determined by the two conditions

$$\zeta + \eta(p + q) + \theta pq = 0, \quad (21)$$

$$\lambda + \mu(p + q) + \nu pq = 0. \quad (22)$$

By (13) and (16),

$$\zeta = 1, \quad \eta = 0, \quad \theta = -1, \quad (23)$$

$$\lambda = a^2 - c^2 - d^2, \quad \mu = ab - cd, \nu = b^2. \quad (24)$$

Combining (21), (22), and (23), we find

$$q = p^{-1},$$

$$p = -\left(\frac{\lambda + \nu}{2\mu}\right)\left\{1 \pm \left[1 - \left(\frac{2\mu}{\lambda + \nu}\right)^2\right]^{1/2}\right\}. \quad (26)$$

The transformation must be real. Therefore, p and q must be real, and thus we must have

$$\left(\frac{2\mu}{\lambda+\nu}\right)^2 \leq 1.$$

Analysis of $[2\mu/(\lambda+\nu)]$ for all values of J_1, J_2, J_3, and T show this to be the case. It will be convenient to leave p as a parameter in the transformation, thereby allowing greater flexibility [by being able to vary the sign of the radical in (26)] in choosing a simpler end result, as will be seen later.

Continuing with the first transformation, by combining (25) with (17) and applying the result to (15), postponing the transformation of R for the moment, and using the limits of integration in (13), we get

$$\int_{-1}^{1}\frac{R\,dx}{\Delta} = Q\int_{-1}^{1}\frac{R'\,dx}{\Delta'}, \tag{27}$$

where Δ' is defined as

$$\Delta' = [(1-x^2)(1+m'x^2)]^{1/2}, \tag{28}$$

$$m' = (\nu + 2\mu p + \lambda p^2)/(\lambda + 2\mu p + \nu p^2), \tag{29}$$

$$Q = \left(\frac{1-p^2}{\lambda+2\mu p+\nu p^2}\right)^{1/2}.$$

We now have the form (19) with $n = 1$ and m not yet determined.

This completes the first transformation on the integrals in (13), except for the rational functions R. These functions do not necessarily influence the linear transformations. They do determine, however, the specific standard forms taken by the elliptic integrals. They will be transformed after the second transformation has been accomplished.

For the second transformation, Eq. (20), there are five distinct possibilities.[7] As a result of (28) and the limits of integration in (13), only two are of interest.

Case I (5° in Ref. 7, p. 318):

The denominator of (19) becomes

$$[(1-nx^2)(1-mx^2)]^{1/2}, \tag{30}$$

with $m > n$ and x from 0 to $1/\sqrt{n}$, or from $1/\sqrt{m}$ to ∞.

In this case, (18) becomes

$$y^2/m\,. \tag{31}$$

Case II (2° in Ref. 7, p. 318):

The denominator of (19) is

$$[(1-nx^2)(1+mx^2)]^{1/2} \tag{32}$$

with $x^2 < 1/n$.

In this case (18) is

$$(1/n)(1-x^2). \tag{33}$$

By comparing (28) with (30) and (32), and by the limits of integration in (27), we have, in terms of m',

$-1 < m' < 0 \Rightarrow$ case I with $m = 1, n = -m'$, and the second transformation is the identity transformation, (34)

$m' > 0 \Rightarrow$ case II with $n = 1, m = m'$, and the second transformation is x^2 replaced with $(1-x^2)$. (35)

Analysis of m' as a function of the four independent variables J_1, J_2, J_3, and T over their full ranges shows that if the sign of the radical in (26) is not fixed, then the condition

$$-\infty < m' < -1 \tag{36}$$

can always be avoided by choosing convenient signs. The last condition to be satisfied is that Q in (27) be real. Analysis of Q, in the same manner as that of m', shows that the sign of the radical in (26) can always be chosen to make Q real, and in doing so, the condition in (36) is never encountered. Therefore the second transformation will always be either case I or case II.

With the first and second transformations determined, we now apply these transformations to (13), including the function R. The result, after extensive simplification, is

$$\langle E\rangle = -J_l \coth 2K_l - \sum_{i=1}^{3} J_i[C_{i1}F(k) + C_{i2}\Pi(r,k) + C_{i3}\Pi(t,k)], \tag{37}$$

where $F(k)$ is the complete elliptic integral of the first kind of modulus k, and where $\Pi(r,k)$ and $\Pi(t,k)$ are complete elliptic integrals of the third kind of modulus k and parameters r and t, respectively. The modulus k and the parameters r, t are real and satisfy $k^2 < 1$ and $-\infty < r, t < \infty$. The definitions of k, t, and r depend on which form the second transformation takes.

Case I:

$$k^2 = -m', \tag{38}$$

$$r = p^2, \tag{39}$$

$$t = \left(\frac{e+p}{1+ep}\right)^2, \tag{40}$$

where

$$e = \frac{2cd}{c^2+d^2}.$$

Case II:

$$k^2 = m'/(1+m'), \tag{41}$$

$$r = p^2/(p^2-1), \tag{42}$$

$$t = \frac{(e+p)^2}{(1-e^2)(p^2-1)}. \tag{43}$$

The coefficients C_{ij} are not in final form, since a specific representation of the complete elliptic integrals of the third kind has not yet been determined. They will therefore be given later.

Each of the two elliptic integrals of the third kind can be expressed as combinations of complete and incomplete elliptic integrals of the first and second kinds, and more compactly in terms of Heuman's λ function,[12] defined as

$$\Lambda_0(\theta,k) = (2/\pi)\,[E(k)\,F(\theta,k') + F(k)\,E(\theta,k') - F(k)F(\theta,k')],$$

where

$$k' = \sqrt{1-k^2},$$

and $E(k), F(\theta, k')$, and $E(\theta, k')$ are the complete elliptic integral of the second kind, and the incomplete elliptic integrals of the first and second kinds, respectively.

There are four possible representations of the elliptic integrals of the third kind. Without listing the specific forms,[12] the conditions that determine them are, in terms of the modulus k and the parameters r and t

$$\begin{aligned} &r,t<0,\\ &0<r,t<k^2,\\ &k^2<r,t<1,\\ &1<r,t. \end{aligned} \tag{44}$$

In addition to being functions of J_1, J_2, J_3, and T, the quantities k, r, and t also depend on the form of the second transformation. An extensive analysis of these quantities was made over the full range of their independent variables, including variations of the second transformations. It was determined that of the four possible conditions in (44), the third always holds in case I and the first always holds in case II. Applying this result to (37) and noting that the linear independence of the coupling energies, the linear independence of any pair of the three kinds of elliptic integrals,[13] and the linear independence of the elliptic integrals of the third kind in (37) having different parameters, all combine to preclude a further simplification in the number of independent terms. As a result, we have the final form for the expectation value of the normalized lattice energy:

$$\langle E\rangle = -J_l \coth 2K_l - \sum_{i=1}^{3}\sum_{j=1}^{3} J_i C_{ij}\psi_j, \tag{45}$$

where l is given by (12), and

$$\psi_1 = F(k),$$

$$\psi_2 = \Lambda_0(\theta_r, k),$$

$$\psi_3 = \Lambda_0(\theta_t, k),$$

and k, θ_r, θ_t, and the coefficients C_{ij} depend on which form the second transformation takes; k is given by (38) and (41).

Case I:

$$\theta_r = \sin^{-1}\left(\frac{r-k^2}{r(1-k^2)}\right)^{1/2}, \quad \text{with } r \text{ given by (39)},$$

$$\theta_t = \sin^{-1}\left(\frac{t-k^2}{t(1-k^2)}\right)^{1/2}, \quad \text{with } t \text{ given by (40)},$$

$$C_{11} = \frac{Q}{2\pi p}(b' + a_1 p),$$

$$C_{21} = \frac{Q}{2\pi p}\left(a_2 p - \frac{(b+ap)(d+cp)c'}{(c^2+d^2)(e+p)}\right),$$

$$C_{31} = \frac{Q}{2\pi p}\left(a_3 p - \frac{(b+ap)(c+dp)d'}{(c^2+d^2)(e+p)}\right),$$

$$C_{12} = -\frac{Q}{4}\left(\frac{1-p^2}{p^2-k^2}\right)^{1/2} b' \operatorname{sgn}(p),$$

$$C_{22} = -\tfrac{1}{2}\coth 2K_2 \tanh 2K_1 C_{12},$$

$$C_{32} = \coth 2K_3 \tanh 2K_2 C_{22},$$

$$C_{13} = 0,$$

$$C_{23} = -\frac{Q}{8}\frac{(ae-b)}{(1+ep)}\left(\frac{1-p^2}{t-k^2}\right)^{1/2} \times \coth 2K_2 \operatorname{sgn}(c^2-d^2) \operatorname{sgn}[2cd + p(c^2+d^2)],$$

$$C_{33} = -\coth 2K_3 \tanh 2K_2 C_{23}.$$

Case II:

$$\theta_r = \sin^{-1}\left(\frac{r}{r-k^2}\right)^{1/2}, \quad \text{with } r \text{ given by (42)},$$

$$\theta_t = \sin^{-1}\left(\frac{t}{t-k^2}\right)^{1/2}, \quad \text{with } t \text{ given by (43)},$$

$$C_{11} = \frac{Q}{2\pi\sqrt{1+m'}}\left(a_1 + \frac{b'p}{k^2(1-p^2)+p^2}\right),$$

$$C_{21} = \frac{Q}{2\pi p\sqrt{1+m'}}\left(a_2 p - \frac{(b+ap)(d+pc)c'}{(c^2+d^2)(e+p)} + \frac{k^2(1-p^2)bc'}{2(k^2(1-p^2)+p^2)c} - \frac{k^2 p}{k^2-t} \cdot \frac{(ae-b)(1+ep)dc'}{(c^2-d^2)(e+p)e}\right),$$

$$C_{31} = \frac{Q}{2\pi p\sqrt{1+m'}}\left(a_3 p - \frac{(b+ap)(c+dp)d_3}{(c^2+d^2)(e+p)} + \frac{k^2(1-p^2)bd'}{2(k^2(1-p^2)+p^2)d} + \frac{k^2 p(ae-b)(1+ep)cd'}{(k^2-t)(c^2-d^2)(e+p)e}\right),$$

$$C_{12} = -\frac{Qb'}{4\sqrt{1+m'}}\left(\frac{1-p^2}{k^2(1-p^2)+p^2}\right)^{1/2} \operatorname{sgn}(p),$$

C_{22}, C_{32}, and C_{13} same as for case I,

$$C_{23} = -\frac{Q}{4\sqrt{1+m'}}\frac{c'd(ae-b)}{e(c^2-d^2)\sqrt{k^2-t}} \operatorname{sgn}(e+p),$$

C_{33} same as case I,

where

$$\operatorname{sgn}(x) \equiv \begin{cases} 1 & x>0 \\ 0 & x=0 \\ -1 & x<0 \end{cases}$$

and

$$a_i \equiv \partial_i a, \qquad b' = -2\cosh 2K_1,$$
$$c' = -2\cosh 2K_2, \qquad d' = -2\cosh 2K_3.$$

For the heat capacity, we combine (2) and (3), and apply the result to (45):

$$C = -\frac{1}{T}\sum_{\epsilon=1}^{3} K_\epsilon\left(\frac{2J_\epsilon \delta_{l\epsilon}}{\sinh^2 2K_\epsilon} - \sum_{i=1}^{3}\sum_{j=1}^{3} J_i(C_{ij}\partial_\epsilon\psi_j + \partial_\epsilon(C_{ij})\psi_j)\right), \tag{46}$$

where

$$\partial_\epsilon \psi_1 = \frac{\partial F}{\partial k}\,\partial_\epsilon k,$$

$$\partial_\epsilon \psi_2 = \frac{\partial \Lambda_0}{\partial \theta_r}\,\partial_\epsilon \theta_r + \frac{\partial \Lambda_0}{\partial k}\partial_\epsilon k,$$

$$\partial_\epsilon \psi_3 = \frac{\partial \Lambda_0}{\partial \theta_t}\,\partial_\epsilon t + \frac{\partial \Lambda_0}{\partial k}\partial_\epsilon k,$$

and $\delta_{i\epsilon}$ is the Kronecker δ function.

The partial derivatives

$$\frac{\partial F}{\partial k}, \frac{\partial \Lambda_0}{\partial k}, \frac{\partial \Lambda_0}{\partial \theta_r}, \frac{\partial \Lambda_0}{\partial \theta_t}$$

can be found in Ref. 12. The derivatives $\partial_\epsilon k, \partial_\epsilon \theta_r, \partial_\epsilon t$, as well as $\partial_\epsilon C_{ij}$ are straightforward but tedious, numbering 60 in all, and therefore will not be given here.

Taking the isotropic case for (45), we get

$$\langle E \rangle = -J \coth 2K + \alpha F(k), \tag{47}$$

where

$$\alpha = QJ/\pi\sqrt{1+m'}\,(\operatorname{csch} 2K - e^{4K}\sinh 2K).$$

Houtappel's corresponding equation has three different forms for α and k; one for $J < 0, 0 < T < \infty$, one for $J > 0, T < T_c$ (critical temperature), and one for $J > 0, T > T_c$. In (47), there are only two forms of α and k, one for $J < 0$ (case I) and one for $J > 0$ (case II).

Assuming Wannier's corrected formula[10,14] corresponding to (47), his is simpler still, having only one form for both $J < 0$ and $J > 0$. It is unlikely, however, that this one formulation would hold for the general anisotropic lattice. Thus the two cases presented here are a likely minimum formulation.

[1]G. H. Wannier, Phys. Rev. 79, 357 (1950).
[2]K. Husimi and I. Syozi, Prog. Theor. Phys. 5, 177, 341 (1950).
[3]R. M. F. Houtappel, Physica 16, 425 (1950).
[4]G. F. Newell, Phys. Rev. 79, 876 (1950).
[5]H. N. V. Temperly, Proc. R. Soc. A 202, 202 (1950).
[6]Bateman Manuscript Project, 1953, Vol. 2, p. 304.
[7]A. Cayley, *An Elementary Treatise on Elliptic Functions* (Dover, New York, 1961), pp. 311-318.
[8]C. Domb, Advan. Phys. 9, 213 (1960).
[9]P. H. E. Meijer, and G. W. Cunningham (unpublished).
[10]G. H. Wannier, Phys. Rev. B 7, 5017 (1973).
[11]D. Bierens DeHaan, *Table of Integrals* 2, 332, p. 472
[12]P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Physicists*, (Springer-Verlag, Berlin, 1954).
[13]Reference 7, pp. 1-5.
[14]The author pointed out to us that his erratum contained an error. In the first column there are two equivalent forms given for k. The first is correct. The second should have in the numerator as the second factor $3 + |\mu|$, not $3 - |\mu|$.

The Lie algebra $so(N)$ and the Duffin–Kemmer–Petiau ring*

Ephraim Fischbach

Physics Department, Purdue University, Lafayette, Indiana 47907

James D. Louck and Michael Martin Nieto

Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

C. K. Scott†

Physics Department, McMaster University, Hamilton, Ontario, Canada

(Received 22 March 1973; revised manuscript received 21 June 1973)

An explicit expression is given for the unit element E of the ring generated by the Duffin–Kemmer–Petiau (DKP) operators β_μ. The relation of E to the unit operator I (unit matrix in a matrix representation) is also derived. It is pointed out that one must be careful to distinguish E from I. Bhabha's observation that one may use the irreducible representations (irreps) of the Lie algebra $so(5)$ to obtain the irreps of the Dirac, DKP, and other algebras is given a concise and general setting in terms of a relation between the Lie algebra $so(n+1)$ and a family of semisimple operator rings. We emphasize that for the case $n+1=5$ this means that there is an underlying relationship between the physical DKP and Dirac algebras and wave equations.

1. INTRODUCTION AND RÉSUMÉ

There has been a renewed interest in the use of the Duffin[1]–Kemmer[2]–Petiau[3] (DKP) algebra for the description of interactions in meson physics. The details of these physical applications have been described elsewhere.[4–7] The present paper is mathematical in content and is intended to serve three purposes: (a) to demonstrate by explicit construction the existence of a *unit element* E of the algebra generated by the four DKP operators[8] $\beta_1, \beta_2, \beta_3, \beta_4$, where we are careful to note the distinction between E and the *unit operator* I which is customarily adjoined to the algebra, a fact which is often only implicit (hence, is a potential source of confusion); (b) to bring into sharper focus a method for finding the inequivalent irreducible representations (irreps) of the DKP algebra[9] by using the theory of Lie algebras (a method which does not require knowing the number of basis elements in the ring); and (c) to point out the generality of this approach for the representation theory of rings.

For completeness of presentation, let us recall briefly the basic properties of the DKP algebra. The DKP ring R' is the algebra over C which is generated by the four operators $\beta_\mu (\mu = 1, 2, 3, 4)$ which satisfy

$$\beta_\mu\beta_\nu\beta_\lambda + \beta_\lambda\beta_\nu\beta_\mu = \beta_\lambda\delta_{\mu\nu} + \beta_\mu\delta_{\lambda\nu}. \tag{1}$$

The operators $\beta_\mu (\mu = 1, 2, 3, 4)$ occur in the DKP meson field equation and the resulting algebra is of particular interest to physicists. In our discussion we will use the following properties of the β operators which follow from Eq. (1):

$$\begin{aligned} &\beta_\mu^2\beta_\nu + \beta_\nu\beta_\mu^2 = \beta_\nu, \qquad \mu \neq \nu, \\ &\beta_\mu\beta_\nu\beta_\mu = \beta_\mu\delta_{\mu\nu}, \\ &\beta_\mu^2\beta_\nu^2 = \beta_\nu^2\beta_\mu^2 . \end{aligned} \tag{2}$$

The contents of this paper are organized as follows: In Sec. 2 we give the unit element E of the DKP ring R', and demonstrate its relation to the 126 basis elements[2,10] of the ring R obtained by adjoining the *unit operator* I to R'. The standard basis elements of R are listed in Table I. The generalization of these results to n operators[11] satisfying Eq. (1) is also noted. In Sec. 3, we consider some general aspects of the Lie algebra $so(n+1), n = 2, 3, \cdots$, generated by the operators $J_{ab} = -J_{ba}$ $(a, b = 1, 2, \cdots, n+1)$. The principal result which is obtained is the relation between the irreps of this Lie algebra and the irreps of a class of semisimple algebras which includes the generalized Dirac algebra[12] and DKP algebra[11] R_n as special cases. Although this relation is perhaps implicit in Bhabha's work[13] for $n = 4$, the generalization to all $n \geq 2$ is not trivial and appears to be not widely known. For the physically interesting case $n = 4$, it is, of course, the covariance of relativistic wave equations (finite number of components) under Lorentz transformations[14] which accounts for the occurrence of these algebras.[13]

To summarize: The principal new results presented are (a) the explicit relation between the unit elements of the algebras R_n' and R_n [Eq. (14)]; and (b) Theorem 2, which states the relation between a class of semisimple algebras and the Lie algebra $so(n+1)$ for arbitrary $n \geq 2$.

2. THE INDEPENDENT ELEMENTS OF THE DKP RING

It is often not made clear in the literature whether one is dealing with the ring R' generated by the operators $\beta_\mu (\mu = 1, 2, 3, 4)$ alone or with the ring R generated by **I**, $\beta_\mu (\mu = 1, 2, 3, 4)$, where **I** is the unit operator[8] which, *by definition*, has the properties $\mathbf{I}\beta_\mu = \beta_\mu\mathbf{I} = \beta_\mu$ and, more generally, $\mathbf{I}R' = R'\mathbf{I} = R'$. The ring R' has its own unit element which we denote by E, and the point to note is that E and I are distinct operators in the ring R. This distinction[15] becomes quite clear when one considers the particular representation $\beta_\mu \to \mathbf{0}$ (zero operator) in which case $E \to \mathbf{0}$, but **I**, of course, is still the unit operator (represented by the unit matrix in any finite dimensional matrix representation). The purpose of this section is to clarify these differences by exhibiting explicitly the expression for E in terms of the generators β_μ [Eq. (3)]; this result is then used, in turn, to establish (in R) the relation [Eq. (10)] between I and E.

We begin by recalling the definition of the elementary symmetric functions:

$$\varphi_1(x_\mu) = \sum_\mu x_\mu, \qquad \varphi_2(x_\mu) = \sum_{\nu>\mu} x_\mu x_\nu,$$

$$\phi_3(x_\mu) = \sum_{\lambda>\nu>\mu} x_\mu x_\nu x_\lambda, \ldots, \text{etc.} \tag{3}$$

Then we have

Theorem 1: The DKP algebra R' generated by the four operators $\beta_\mu (\mu = 1, 2, 3, 4)$ has a unit element which is

$$E = \sum_\mu \beta_\mu^2 - \sum_{\nu>\mu} \beta_\mu^2\beta_\nu^2 + \sum_{\lambda>\nu>\mu} \beta_\mu^2\beta_\nu^2\beta_\lambda^2 - \sum_{\rho>\lambda>\nu>\mu} \beta_\mu^2\beta_\nu^2\beta_\lambda^2\beta_\rho^2 \quad (4)$$

$$= \varphi_1(\beta_\mu^2) - \varphi_2(\beta_\mu^2) + \varphi_3(\beta_\mu^2) - \varphi_4(\beta_\mu^2).$$

Proof: The theorem can be proven using Eqs. (2) by a somewhat tedious direct verification of $E\beta_\sigma = \beta_\sigma E = E$ (only one value of σ is needed because of the permutational symmetry) and, consequently, $ER' = R'E = R'$. A more elegant proof is given below.

We next consider the expression for the unit element E in terms of the commuting operators η'_μ which are defined by

$$\eta'_\mu = 2\beta_\mu^2 - E \quad \text{in } R'. \quad (5)$$

Note that the operators η'_μ in R' are *not to be confused* with the similarly defined operators

$$\eta_\mu = 2\beta_\mu^2 - \mathbf{I} \quad \text{in } R. \quad (6)$$

Obviously the relation between the two definitions is

$$\eta'_\mu = \eta_\mu + \mathbf{I} - E \quad \text{in } R. \quad (7)$$

The four operators η'_μ generate a commutative subalgebra of the DKP algebra R'. One can establish directly as a corollary to Theorem 1, the expression for the unit element E of R' in terms of the η'_μ operators.

Corollary. The expression for the unit element of the DKP meson algebra R' in terms of the η'_μ operators is

$$E = \sum_\mu \eta'_\mu - \sum_{\nu>\mu} \eta'_\mu\eta'_\nu + \sum_{\lambda>\nu>\mu} \eta'_\mu\eta'_\nu\eta'_\lambda - \sum_{\rho>\lambda>\nu>\mu} \eta'_\mu\eta'_\nu\eta'_\lambda\eta'_\rho \quad (8)$$

$$= \varphi_1(\eta'_\mu) - \varphi_2(\eta'_\mu) + \varphi_3(\eta'_\mu) - \varphi_4(\eta'_\mu).$$

Let us now look at the connection between E and $\mathbf{I}$ when considered in the larger ring R. This will allow us to give a transparent demonstration that E is the unit element in the ring R'. We start by observing that

$$\prod_{\mu=1}^{4} (E - \beta_\mu^2) = \mathbf{0} \quad \text{in } R'. \quad (9)$$

However,

$$\Lambda \equiv \prod_{\mu=1}^{4} (\mathbf{I} - \beta_\mu^2) = \prod_{\mu=1}^{4} \tfrac{1}{2}(\mathbf{I} - \eta_\mu) = \mathbf{I} - E \quad \text{in } R, \quad (10)$$

so that

$$E^2 = (\mathbf{I} - \Lambda)^2 = E. \quad (11)$$

Since it is trivial to verify the relation $\Lambda\beta_\mu = \beta_\mu\Lambda = \mathbf{0}$, we obtain immediately the result,

$$E\beta_\mu = \beta_\mu E = \beta_\mu, \quad (12)$$

which, together with property (11), expresses the fact that E is the unit element in R'.

Let us summarize: The semisimple ring R has $126 = 1^2 + 5^2 + 10^2$ independent basis elements (Table I) and the ring R' has 125 independent basis elements (Table I with $\mathbf{I}$ deleted and η_μ replaced by η'_μ); E and $\mathbf{I}$ are each represented by the unit matrix in the five- and ten-dimensional irreps of the algebra, but, in the one-dimensional irrep $\beta_\mu \to 0$, E corresponds to 0, whereas $\mathbf{I}$ corresponds to 1.

We close this section by observing that all these results can be carried over directly to the n dimensional DKP algebra ($\mu = 1, \ldots n$). Specifically, the unit element E_n of the algebra R'_n is

$$E_n = \sum_{s=1}^{n} (-1)^{s+1}\varphi_s(\beta_\mu^2), \quad (13)$$

and its connection to the unit element $\mathbf{I}_n$ of the algebra R_n is

$$E_n = \mathbf{I}_n - \frac{1}{2^n}\prod_{\mu=1}^{n} (\mathbf{I}_n - \eta_\mu). \quad (14)$$

3. THE GENERALIZED DKP, DIRAC, AND *so*(N) ALGEBRAS

An illuminating method for obtaining all the irreps of the DKP algebra is to consider the irreps of the Lie algebra *so*(5) of the group of real proper orthogonal matrices of dimension five. Since the method applies quite generally to a class of semisimple algebras which may be obtained from the Lie algebra *so*(N) of the group of real proper orthogonal matrices of dimension N, we will describe this more general case[16] thereby obtaining a semisimple algebra, $\mathfrak{G}_n^{(k)}$, which for $k = 1/2$ becomes the generalized Dirac algebra[12] and for $k = 1$ becomes the generalized DKP algebra.[11] The irreps of this algebra are given explicitly (up to an equivalence) by the known irreps of the basis elements of the Lie algebra *so*(n).[17–20]

As stated in the introduction, while the results presented here for $N = 5$ are perhaps already either explicit or implicit in Bhabha's[13] treatment of relativistic wave

TABLE I: The 126 independent elements of the DKP meson algebra R. In column (a) the independent elements are listed in terms of β_μ and β_μ^2. In column (b) the independent elements are listed in terms of β_μ and $\eta_\mu = 2\beta_\mu^2 - \mathbf{I}$.

Elements: (a)	(b)	Number of Elements
$\mathbf{I}$	$\mathbf{I}$	1
β_μ	β_μ	4
$\beta_\mu\beta_\nu$	$\beta_\mu\beta_\nu$	12
$\beta_\mu\beta_\nu\beta_\sigma$	$\beta_\mu\beta_\nu\beta_\sigma$	12
$\beta_\mu\beta_\nu\beta_\sigma\beta_\rho$	$\beta_\mu\beta_\nu\beta_\sigma\beta_\rho$	6
β_μ^2	η_μ	4
$\beta_\mu^2\beta_\nu^2$	$\eta_\mu\eta_\nu$	6
$\beta_\mu^2\beta_\nu^2\beta_\sigma^2$	$\eta_\mu\eta_\nu\eta_\sigma$	4
$\beta_\mu^2\beta_\nu^2\beta_\sigma^2\beta_\rho^2$	$\eta_\mu\eta_\nu\eta_\sigma\eta_\rho$	1
$\beta_\mu^2\beta_\nu$	$\eta_\mu\beta_\nu$	12
$\beta_\mu^2\beta_\nu\beta_\sigma$	$\eta_\mu\beta_\nu\beta_\sigma$	24
$\beta_\mu^2\beta_\nu\beta_\sigma\beta_\rho$	$\eta_\mu\beta_\nu\beta_\sigma\beta_\rho$	12
$\beta_\mu^2\beta_\nu^2\beta_\sigma$	$\eta_\mu\eta_\nu\beta_\sigma$	12
$\beta_\mu^2\beta_\nu^2\beta_\sigma\beta_\rho$	$\eta_\mu\eta_\nu\beta_\sigma\beta_\rho$	12
$\beta_\mu^2\beta_\nu^2\beta_\sigma^2\beta_\rho$	$\eta_\mu\eta_\nu\eta_\sigma\beta_\rho$	4
		126

equations, it nonetheless seems worthwhile to review briefly the basic algebraic relations and to state the basic results for the general case, since the generalization does not follow automatically from the specific case. In this respect, we should point out that there exists other work[21] based on Bhabha's results for $n = 4$, as well as work[22] dealing with related questions.

A basis of an operator realization[8] of the Lie algebra $so(N)$ is given by the set of Hermitian operators

$$J_{ab} = -J_{ba}, \quad a,b = 1,2,\ldots,N \tag{15}$$

which satisfy the commutation relations

$$[J_{ab},J_{cd}] = i(\delta_{ac}J_{bd} + \delta_{bd}J_{ac} - \delta_{bc}J_{ad} - \delta_{ad}J_{bc}). \tag{16}$$

Let us suppose that we are given a set of Hermitian operators $\{\alpha_\mu : \mu = 1,2,\ldots,n\}$ which satisfy the double commutator relations

$$[[\alpha_\mu,\alpha_\nu],\alpha_\lambda] = \alpha_\mu\delta_{\nu\lambda} - \alpha_\nu\delta_{\mu\lambda}, \quad \mu,\nu,\lambda = 1,2,\ldots,n. \tag{17}$$

We may then define Hermitian operators $\{J_{ab}\}$ in the following manner:

$$J_{\mu,n+1} = -J_{n+1,\mu} = \alpha_\mu, \quad J_{\mu\nu} = -i[\alpha_\mu,\alpha_\nu],$$
$$J_{n+1,n+1} = 0. \tag{18}$$

Using these definitions, the assumed double commutator property of the α_μ, and Jacobi's identity for commutators, it is straightforward to prove that the J_{ab} satisfy Eq. (16) where $N = n + 1$. Conversely, given a set of Hermitian operators J_{ab} satisfying the Lie algebra relation (16), then the n Hermitian operators defined by

$$\alpha_\mu = J_{\mu,n+1}, \quad \mu = 1,2,\ldots,n \tag{19}$$

satisfy Eq. (17).

The proof of the following result may now easily be given (and we omit it): *Each inequivalent irrep of the Lie algebra so(n + 1) determines an inequivalent irrep of the operator algebra* (17), *and conversely*. The significance of this result is immediately apparent when we remark that both $\alpha_\mu = \gamma_\mu/2$, where γ_μ is a Dirac operator ($\mu = 1,2,\ldots,n$), and $\alpha_\mu = \beta_\mu$, where β_μ is a DKP operator ($\mu = 1,2,\ldots,n$), are particular examples of Hermitian operators which satisfy the algebraic relation (17). In order to understand more fully the origin of this property, we next discuss several relations relevant to the irreps of $so(N)$

Each irrep of $so(N)$ is specified by a set of ordered numbers

$$\{l^N\} = \{l_1^N l_2^N \cdots l_r^N\}, \tag{20}$$

where $r = N/2$ for N even and $r = (N-1)/2$ for N odd. These numbers are either all integers or all half-integers which satisfy

$$l_1^N \geq l_2^N \geq \cdots \geq l_{r-1}^N \geq |l_r^N| \geq 0, \tag{21}$$

where for odd N the last number l_r^N in the sequence (20) is nonnegative, i.e., $l_r^N \geq 0$ for odd N. The dimension of the irreducible representation specified by $\{l^N\}$ is given by

$$\dim\{l^N\} = \frac{2^{r-1}}{2!4!\cdots(2r-2)!}\prod_{i<j=1}^{r}(\lambda_i^2 - \lambda_j^2) \tag{22}$$

for $N = 2r$ ($r = 1,2,\ldots$), and by

$$\dim\{l^N\} = \frac{1}{1!3!\cdots(2r-1)!}\prod_{i=1}^{r}(2\lambda_i + 1) \times \prod_{i<j=1}^{r}[\lambda_i(\lambda_i+1) - \lambda_j(\lambda_j+1)] \tag{23}$$

for $N = 2r + 1$ ($r = 1,2,\ldots$), where the λ_i are defined by

$$\lambda_i = l_i^N + r - i. \tag{24}$$

[For $r = 1$ the product term of differences is defined to be unity in Eqs. (22) and (23).]

A second important property of the basis (15) of $so(N)$ is the unitary equivalence of each generator J_{ab} to a particular generator, say, J_{12}. This latter generator is itself a member of the triplet $\mathbf{J} = (J_{23},J_{31},J_{12})$ whose components satisfy the commutation relations of an angular momentum **J**. This signifies that an irreducible representation of J_{12} [as an element of the $so(N)$ algebra] is equivalent to a direct sum of diagonal block matrices of various dimensions $2j + 1$ for certain values of j selected from $0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$, where each such block itself has diagonal elements $m = j, j-1, \ldots, -j$. Noting further that in the irreducible representation $\{l^N\}$ the eigenvalues of J_{12} are $m = l_1^N, l_1^N - 1, \ldots, -l_1^N$, we may assert: *Each* α_μ *satisfies the minimal characteristic equation*

$$\prod_{m=-l_1^N}^{l_1^N}(\alpha_\mu - m\mathbf{I}) = \mathbf{0} \tag{25}$$

in the irrep $\{l^N\}$. The proof follows directly from the Cayley–Hamilton theorem for Hermitian matrices and the properties mentioned above.

We are thus led to introduce[8] the following algebra over C for each integer of half-integer k:

Theorem 2: Let $\mathfrak{B}_n^{(k)}$ denote the algebra generated by **I** (unit operator) and $\alpha_1,\alpha_2,\ldots,\alpha_n$ which satisfy

$$[[\alpha_\mu,\alpha_\nu],\alpha_\lambda] = \alpha_\mu\delta_{\nu\lambda} - \alpha_\nu\delta_{\mu\lambda}, \tag{26}$$

$$\prod_{m=-k}^{k}(\alpha_\mu - m\mathbf{I}) = \mathbf{0}. \tag{27}$$

Then:

(a) $\mathfrak{B}_n^{(0)}$ is the trivial algebra with elements **0**, **I**.

(b) $\mathfrak{B}_n^{(1/2)}$ is the Dirac algebra.

(c) $\mathfrak{B}_n^{(1)}$ is the DKP algebra R_n.

(d) All the inequivalent irreducible Hermitian matrix representations of $\mathfrak{B}_n^{(k)}$ may be obtained from the known representations of the Lie algebra $so(n + 1)$ through the identification $\alpha_\mu = J_{\mu,n+1}$. Specifically, the irreps of $\mathfrak{B}_n^{(k)}$ may be designated by $\{l^{n+1}\}$, where k together with $\{l^{n+1}\}$ is any set of integers or half-integers satisfying [cf. Eq. (21)]

$$k \geq l_1^{n+1} \geq l_2^{n+1} \geq \cdots \geq |l_r^{n+1}| \geq 0. \tag{28}$$

(e) The algebra $\mathfrak{B}_n^{(k)}$ is semisimple.

(f) The rank of the algebra $\mathfrak{B}_n^{(k)}$ is given by the sum

$$\mathrm{rank}\,\mathfrak{B}_n^{(k)} = \sum_{\{l^{n+1}\}}(\dim\{l^{n+1}\})^2, \tag{29}$$

where the sum is either over all integers $\{l^{n+1}\}$ *or* all half-integers $\{l^{n+1}\}$ satisfying Eq. (28).

(g) The property (trace α_μ) = 0 holds in each representation (irreducible or reducible) of the α_μ.

The proofs of the above statements (a)-(g) are contained in the Appendix.

The ranks of the Dirac and DKP algebras are easily calculated from Eqs. (22), (23), and (24). Thus, from

$$\dim\{\tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}\} = 2^r \quad \text{for } so(2r+1), \tag{30a}$$

$$\dim\{\tfrac{1}{2}, \tfrac{1}{2}, \ldots, \pm\tfrac{1}{2}\} = 2^{r-1} \quad \text{for } so(2r), \tag{30b}$$

we obtain

$$\operatorname{rank}\mathfrak{B}_n^{(1/2)} = 2^n. \tag{31}$$

Similarly, from[23]

$$\dim\{\dot{1}_q\dot{0}_{r-q}\} = \binom{N}{q} \quad \text{for } so(N), \tag{32a}$$

where $q = 0, 1, \ldots, r$ for N odd and $q = 0, 1, \ldots, r-1$ for N even, and

$$\dim\{1, \ldots, 1, \pm 1\} = \tfrac{1}{2}\binom{2r}{r} \tag{32b}$$

for $so(2r)$, the rank of $\mathfrak{B}_n^{(1)}$ is found to be[24]

$$\operatorname{rank}\mathfrak{B}_n^{(1)} = \binom{2n+1}{n}. \tag{33}$$

We conclude by emphasizing the special physical case of the above results, the algebra $so(5)$. There we are considering in particular the DKP and Dirac algebras used in relativistic wave equations ($\mu = 1, 2, 3, 4$). From Eq. (23) one has that the dimensions of the representations are

$$d_5(l_1, l_2) = \tfrac{1}{6}(2l_1+3)(2l_2+1)[(l_1+1)(l_1+2) - l_2(l_2+1)]. \tag{34}$$

As shown in Table II, the irreps with $l_1 \leq 1$ are the three irreps of the DKP algebra and the Dirac algebra for physical 4-space. That is, by considering the algebra $so(5)$ one has a unifying principle[13] connecting the Dirac and DKP relativistic wave equations. For higher spin ($l_1 > 1$) some work has been done for the cases $S = 3/2$ and 2.[13,25]

APPENDIX

We give the proofs of the statements (a)-(g) in the main text:

(a) Choose $k = 0$ in Eq. (27).

(b) Choose $k = \frac{1}{2}$ in Eq. (27) to obtain $\alpha_\mu^2 = \frac{1}{4}$. Choose $\lambda = \nu$ and $\mu \neq \nu$ in Eq. (26) and use $\alpha_\nu^2 = \frac{1}{4}$ to obtain $\alpha_\nu\alpha_\mu\alpha_\nu = -\alpha_\mu/4$ for $\mu \neq \nu$. Multiply this expression from the right by α_ν to obtain $\alpha_\nu\alpha_\mu + \alpha_\mu\alpha_\nu = 0$ for $\mu \neq \nu$. Thus, $\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}$, where $\gamma_\mu = 2\alpha_\mu$. Conversely, it is easily demonstrated that $\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}$ implies Eqs. (26) and (27) for $\alpha_\mu = \gamma_\mu/2$.

(c) Choose $k = 1$ in Eq. (27) to obtain $\alpha_\mu^3 = \alpha_\mu$. Again choose $\lambda = \nu$ and $\mu \neq \nu$ in Eq. (26) to obtain $\alpha_\mu\alpha_\nu^2 + \alpha_\nu^2\alpha_\mu - 2\alpha_\nu\alpha_\mu\alpha_\nu = \alpha_\mu (\mu \neq \nu)$. Multiplication from the right by α_ν now yields $\alpha_\nu^2\alpha_\mu\alpha_\nu = 2\alpha_\nu\alpha_\mu\alpha_\nu^2 (\mu \neq \nu)$. Multiplication of this result by α_ν from the left and right, respectively, gives the relations $\alpha_\nu\alpha_\mu\alpha_\nu = 2\alpha_\nu^2\alpha_\mu\alpha_\nu^2$ and $\alpha_\nu^2\alpha_\mu\alpha_\nu^2 = 2\alpha_\nu\alpha_\mu\alpha_\nu$ which require $\alpha_\nu\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu$). Equation (26) for $\mu \neq \nu$ now becomes $\alpha_\mu\alpha_\nu^2 + \alpha_\nu^2\alpha_\mu = \alpha_\mu$. Next choose $\mu \neq \nu \neq \lambda$ in Eq. (26), multiply from the left by α_ν, and use $\alpha_\nu\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu$) to obtain $\alpha_\nu^2\alpha_\mu\alpha_\lambda + \alpha_\nu\alpha_\lambda\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu \neq \lambda$), which yields $\alpha_\nu\alpha_\mu\alpha_\lambda + \alpha_\nu^2\alpha_\lambda\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu \neq \lambda$) upon multiplication from the left by α_ν. This last result is now reduced to $\alpha_\nu\alpha_\mu\alpha_\lambda + \alpha_\lambda\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu \neq \lambda$) upon using the already established relations $\alpha_\nu^2\alpha_\lambda = -\alpha_\lambda\alpha_\nu^2 + \alpha_\lambda$ ($\nu \neq \lambda$) and $\alpha_\nu\alpha_\mu\alpha_\nu = 0$ ($\mu \neq \nu$). We have now proved that Eqs. (26) and (27) for $k = 1$ imply the four relations $\alpha_\mu^3 = \alpha_\mu$, $\alpha_\mu\alpha_\nu\alpha_\mu = 0$ ($\mu \neq \nu$), $\alpha_\mu\alpha_\nu^2 + \alpha_\nu^2\alpha_\mu = \alpha_\mu$ ($\mu \neq \nu$), and $\alpha_\mu\alpha_\nu\alpha_\lambda + \alpha_\lambda\alpha_\nu\alpha_\mu = 0$ ($\mu \neq \nu \neq \lambda$) which together comprise the statement of the DKP algebraic properties $\alpha_\mu\alpha_\nu\alpha_\lambda + \alpha_\lambda\alpha_\nu\alpha_\mu = \alpha_\mu\delta_{\nu\lambda} + \alpha_\lambda\delta_{\nu\mu}$. Conversely, using similar methods, it may be demonstrated that these relations imply Eqs. (26) and (27).

(d) For the proof we have already pointed out the one-to-one correspondence between the Hermitian matrix irreps of $so(n+1)$ and the operator algebra (26). The subsidiary relation (27) simply means that we must now select only those irreps satisfying (27), and it follows from Eq. (25) that this will be the case if and only if the label l_1^N satisfies $l_1^N \leq k$ *and* the set of representation labels $\{l^N\}$ is comprised either of all integers or all half-integers according to k.

(e) Each finite-dimensional representation of the Lie algebra $so(n+1)$ ($n \geq 2$) is completely reducible, which, in turn, implies that $\mathfrak{B}_n^{(k)}$ is a two-sided completely reducible ring. Since $\mathfrak{B}_n^{(k)}$ also has an identity, it follows[26] that $\mathfrak{B}_n^{(k)}$ is semisimple.

(f) We use the fact that each irrep of a compact group is equivalent to a unitary irrep, and, hence each irrep of a basis of the Lie algebra $so(n+1)$ is equivalent to a Hermitian irrep. Accordingly, *all* the inequivalent finite-dimensional irreps of $\mathfrak{B}_n^{(k)}$ are obtained from those of $so(n+1)$. The expression for the rank of $\mathfrak{B}_n^{(k)}$ now follows from the theorem of Frobenius and Schur[27] equating the number of independent elements to the sum of squares of the dimensions of the irreps.

(g) This trace property is an immediate consequence of the fact that each α_μ is equivalent (by a nonsingular transformation) to a component of an ordinary angular momentum matrix. Alternatively, the property follows directly from Eq. (26) upon taking the trace.

TABLE II: Irreducible representations of $so(5)$.

Dimension	Algebra	Matrix
$d_5(0,0) = 1$	DKP(trivial)	$\alpha_\mu = \beta_\mu^{(1\times1)}$
$d_5(\frac{1}{2},\frac{1}{2}) = 4$	Dirac	$\alpha_\mu = \gamma_\mu/2$
$d_5(1,0) = 5$	DKP	$\alpha_\mu = \beta_\mu^{(5\times5)}$
$d_5(1,1) = 10$	DKP	$\alpha_\mu = \beta_\mu^{(10\times10)}$

* Supported in part by the U.S. Atomic Energy Commission (E.F., J.D.L. and M.M.N.) and the National Research Council of Canada (C.K.S.).

† Present address: Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87544.

[1] R. J. Duffin, Phys. Rev. **54**, 1114 (1938).
[2] N. Kemmer, Proc. Roy. Soc. A **173**, 91 (1939).
[3] G. Petiau, University of Paris Thesis (1936). Published in Acad. Roy. de Belg., Classe Sci., Mém. in 8° **16**, No. 2 (1936).
[4] E. Fischbach, F. Iachello, A. Lande, M. M. Nieto, and C. K. Scott, Phys. Rev. Letters **26**, 1200 (1971).

[5]E. Fischbach, M. M. Nieto, H. Primakoff, C. K. Scott, and J. Smith, Phys. Rev. Letters 27, 1403 (1972).
[6]E. Fischbach, M. M. Nieto, H. Primakoff, and C. K. Scott, Phys. Rev. Letters 29, 1046 (1972).
[7]E. Fischbach, M. M. Nieto, and C. K. Scott, Phys. Can. 27, 64 (1971); Prog. Theor. Phys. (Kyoto) 48, 574 (1972); Phys. Rev. D 6, 726 (1972); 7, 207 (1973).
[8]For the purposes of this paper, it is sufficient to consider that all operators which are introduced are linear operators which act in an abstract separable Hilbert space $\mathfrak{H}$. The unit operator I is then defined by the action $\mathbf{I}\,\psi = \psi$ for each $\psi \in \mathfrak{H}$, and it may always be adjoined to the other operator sets which we consider.
[9]Throughout this paper we use R' to designate the algebra generated by the β_μ alone and R to designate the algebra R' with I adjoined. We refer to either of these algebras as the DKP algebra.
[10]J. Géhéniau, Acad. Roy. de Belg., Classe Sci., Mém. in 8° 18, No. 1 (1938). In this work the algebra was first decomposed into one-, five-, and ten-dimensional representations.
[11]N. Kemmer, Proc. Camb. Phil. Soc. 39, 189 (1943).
[12]R. Brauer and H. Weyl, Am. J. Math. 57, 425 (1935).
[13]H. J. Bhabha, Rev. Mod. Phys. 17, 200 (1945); 21, 451 (1949).
[14]It is the Lie algebra $so(3,1)$ which enters into the physical problem in place of the subalgebra $so(4)$ of $so(5)$. However, it is well known that a slight change in the reality conditions changes the basis of one of these algebras into a basis for the other, and for the determination of the *finite-dimensional irreps* we are permitted to use the language of the Lie algebras of the orthogonal groups.
[15]In more technical language R is semisimple, whereas R' is not. Perhaps the easiest way to check that R' is not semisimple is to apply the criterion of Pauli and Artin (attributed to them by Kemmer, Ref. 2) to the one-dimensional representation $\beta_\mu \to 0$. See also P. Roman, *Theory of Elementary Particles* (North-Holland, Amsterdam, 1960), Chaps. I and II.
[16]Throughout this paper N will always be related to n through $N = n + 1$.
[17]I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR 71, 1017 (1950).
[18]J. D. Louck, *Theory of Angular Momentum in N-Dimensional Space*, Los Alamos Scientific Laboratory Report LA-2451, 1960 (unpublished).
[19]S. C. Pang and K. T. Hecht, J. Math. Phys. 8, 1233 (1967).
[20]M. K. F. Wong, J. Math. Phys. 8, 1899 (1967).
[21]C. Ryan and E. C. G. Sudarshan, Nucl. Phys. 47, 207 (1963); A. Ramakrishnan, R. Vasudevan, P. S. Chandrasekaran, and N. R. Ranganathan, J. Math. Anal. Appl. 28, 108 (1969).
[22]A. J. Macfarlane, Commun. Math. Phys. 2, 133 (1966); I. Gyuk and H. Umezawa, Phys. Rev. D 3, 898 (1971).
[23]A dot over an integer designates that the integer is repeated a number of times equal to the subscript.
[24]T. Shimpuku, J. Math. Anal. Appl. 27, 181 (1969); Addendum, J. Math. Anal. Appl. 31, 717 (1970).
[25]B. S. Madhava Rao, Proc. Indian Acad. Sci. A 15, 139 (1942).
[26]B. L. van der Waerden, *Modern Algebra*, revised English edition (F. Ungar, New York, 1950), Vol. II, p. 155.
[27]G. Frobenius and I. Schur, Sitzungsber. König. Preuss. Akad. Wissen. Berlin, 186 (1906).

Zeroes of the partition function for the Ising model in the complex temperature plane

H. J. Brascamp* and H. Kunz†

Institut des Hautes Etudes Scientifiques, 91440-Bures-sur-Yvette, France
(Received 8 June 1973)

For special boundary conditions, the zeroes of the partition function of the square Ising model are shown to lie on Fisher's two circles in the complex $\exp(-2\beta J)$ plane. For some more general boundary conditions, the zeroes distribute asymptotically on these circles.

In this note we study the zeroes in the temperature plane of the partition function of the square Ising model. It has been conjectured by Fisher,[1] that the zeroes distribute asymptotically on two circles in the $\exp(-2\beta J)$ plane, given by

$$e^{-2\beta J} = \pm 1 + \sqrt{2}e^{i\varphi}. \tag{1}$$

Numerical investigations, in the case of periodic boundary conditions, by Abe and Katsura,[2] support this conjecture. We confirm it here in the following sense:

(a) For special boundary conditions the zeroes are proved to lie on (1), or, more conveniently, on the unit circle in the $\mathrm{Sh}2\beta J$ plane.

(b) With some other boundary conditions their density is shown to become asymptotically zero in the complement of the circles (1).

The special boundary conditions in (a) are the following (Fig. 1).

(A) Periodic in the horizontal direction, along the upper border of the resulting cylinder (height M, circumference N) a band of fixed spins +, along the lower border a band of alternating spins (so N must be even).

(B) Periodic in a diagonal direction, along the upper border of the cylinder a band of spins +, free along the lower border.

In both cases the symmetry $J \to -J$ (combined with spin reversal on the odd sublattice, and a reflection in case (A) is exact. This implies, that the partition function is an even polynomial in $\mathrm{Sh}2\beta J$ (multiplied by $\mathrm{Ch}2\beta J$ if both M and N are odd).

Moreover, system B is exactly self-dual (i.e., it has high-temperature–low-temperature relations between the partition functions and between the even spin correlation functions). In particular, the partition function is invariant up to a factor under the transformation $\mathrm{Sh}2\beta J \to \mathrm{Sh}^{-1}2\beta J$.

These two symmetries already indicate, that the zeroes might lie on the unit circle in the $\mathrm{Sh}2\beta J$ plane. That this is indeed the case, follows from the exact expressions for the partition functions:

$$Q_A(M,N) = 2^{MN} \prod_{1\le j\le N/2} \prod_{1\le k\le M} [\mathrm{Ch}^2 2\beta J - \mathrm{Sh}2\beta J(\cos\theta_j + \cos\varphi_k)] \ (N \text{ even}), \tag{2}$$

$$Q_B(M,N) = 2^{MN} \prod_{1\le j\le [N/2]} \prod_{1\le k\le M} [\mathrm{Ch}^2 2\beta J - 2\mathrm{Sh}2\beta J \cos\tfrac{1}{2}\theta_j \cos\varphi_k] \times \begin{cases} 1, N \text{ even} \\ \mathrm{Ch}^M 2\beta J, N \text{ odd}. \end{cases} \tag{3}$$

The angles are defined as

$$\theta_j = (2j-1)\pi/N; \ \varphi_k = k\pi/(M+1). \tag{4}$$

Note, that the partition function of A also happens to have a self-dual form. However, the system A has not the full self-duality; the system A* dual to A is the following:

(A*) Periodic in the horizontal direction, free along the upper border of the cylinder, along the lower border an extra imaginary magnetic field $h = \pi i/2\beta$.

In precise terms (cf. Benettin *et al.*[3] for duality in connection with boundary conditions):

$$Q_A(M,N,\beta) = 2^{-1-M/2}(\mathrm{Sh}2\beta J)^{MN+M/2} Q_{A^*}(M+1,N,\beta^*); \quad \mathrm{Sh}2\beta J \ \mathrm{Sh}2\beta^* J = 1. \tag{5}$$

The usefulness of system A^* lies in the fact, that its partition function has been computed before. Using the Pfaffian method, McCoy and Wu[4] found the partition function for an arbitrary magnetic field h along the lower boundary.

The crucial point in the calculation is the following. After the use of the anticyclicity of the Pfaffian (determinant) in the periodic direction, the partition function reduces to a product of N Toeplitz block determinants, parametrized by the angle θ_j, Eq. (4). Each of them leads by a recursion procedure to a quantity of the form

$$\langle \psi_1(\theta) | L(\theta)^M | \psi_2(\theta) \rangle, \tag{6}$$

with a 2×2 "transfer" matrix L. The vectors $\psi_{1,2}$ reflect the boundary conditions [cf. Ref. 4, Eqs. (3.12)–(3.13)].

For the imaginary boundary field under consideration the expression (6) becomes proportional to

$$\lambda_+(\theta)^{M+1} - \lambda_-(\theta)^{M+1} \tag{7}$$

for all θ, where $\lambda_\pm$ are the eigenvalues of L [take $\mathrm{Ch}\beta h = 0, z \ \mathrm{Ch}\beta h = i$ in Ref. 4, Eq. (3.26)]. The result (7) allows one to write the partition function as a double product.

The partition function for case B is found straightforwardly along the same lines. Again the boundary conditions are such, that the computation results in expressions of the form (7).

Let us finally consider the case of periodic boundary conditions in one direction and an arbitrary homogeneous magnetic field h along the lower boundary.[4] It is more convenient then to take $\exp(-2\beta J)$ as a variable, because the partition function is no longer a polynomial in $\mathrm{Sh}2\beta J$.

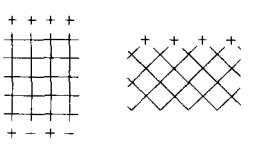

FIG. 1. The special boundary conditions A (left) and B (right). For both, $M \times N = 5 \times 4$ is pictured.

The computation leads to an expression of the form [Ref. 4, Eq. (3.21)]

$$a_+(\theta)\lambda_+(\theta)^M + a_-(\theta)\lambda_-(\theta)^M \tag{8}$$

instead of Eq. (7).

It is convenient to cut the $\exp(-2\beta J)$ plane along the arcs on the circles (1), given by

$$|\lambda_+(\theta)| = |\lambda_-(\theta)|.$$

Then $\lambda_\pm(\theta)$ and $a_\pm(\theta)$ are analytic functions in the cut plane, and

$$\lambda_+(\theta) \neq 0; \quad |\lambda_-(\theta)/\lambda_+(\theta)| < 1.$$

By Rouché's theorem, the zeroes of (8) distribute asymptotically ($M \to \infty$) on the curves

$$a_+(\theta) = 0, \tag{9}$$

$$\lambda_+(\theta) = \lambda_-(\theta)e^{i\varphi}. \tag{10}$$

Equation (10) gives the above-mentioned cuts on the circles (1), and Eq. (9) leads to at most 4 zeroes for each θ; so the density of such zeroes tends to 0 when $M \to \infty$. The density of zeroes of the partition function is therefore completely determined by Eq. (10).

It appears very likely, that the locus of zeroes for the other soluble two-dimensional Ising models is given by the cut, which appears naturally in the analytic expression for the free energy density. It would also be interesting to study the locus of zeroes of other self-dual models, like the Ashkin–Teller–Potts model.

We should like to thank H. Widom and F. Y. Wu for helpful discussions, and N. H. Kuiper for his hospitality at I.H.E.S.

* On leave from the University of Groningen, The Netherlands; supported by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

† Supported by the Swiss National Foundation for Scientific Research.

[1] M. E. Fischer, *Boulder Lectures in Theoretical Physics, 1964* (University of Colorado Press, Boulder 1965), Vol. 7c, p.1.
[2] Y. Abe and S. Katsura, Prog. Theor. Phys. **43**, 1402 (1970).
[3] G. Benettin, G. Gallavotti, G. Jona-Lasinio, and A. L. Stella, Commun. Math. Phys. **30**, 45 (1973).
[4] B. M. McCoy and T. T. Wu, Phys. Rev. **162**, 436 (1967).

Statistical mechanics for velocity dependent interactions*

M. D. Srinivas

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

Y. S. Prahalad

Matscience, Institute of Mathematical Sciences, Madras 20, India

(Received 30 March 1973; revised manuscript received 8 August 1973)

The quantum corrections to the phase space distribution function are obtained for general velocity dependent interactions. It is noted that a study of the thermodynamic properties of bulk nuclear matter may settle the question of velocity dependence of nucleon–nucleon interaction.

Wigner distribution function has been successfully applied by several authors,[1–5] for calculating the quantum corrections to thermodynamic functions. These investigations have been carried out using velocity-independent interactions. The purpose of the present note is to apply this method to a general Hamiltonian in the first instance, and then to the velocity-dependent interaction of the type introduced by Green.[6]

The temperature dependence of an appropriate generalized phase space distribution function (e.g., the Wigner distribution function) is usually taken as the starting point for such an investigation. This consists of a straightforward replacement of the products of quantum mechanical operators by the corresponding products of the c-number functions on which they are mapped. The method leads to corrections which in the Wigner–Weyl correspondence coincide with those obtained by Oppenheim and Ross in the limit of velocity independence. In conclusion it is pointed out that the thermodynamic properties of bulk nuclear matter can help to decide whether nuclear forces are really velocity dependent.

The possibility that the nucleon–nucleon interaction might be velocity dependent has been widely investigated.[6–10] It is well known that a nonlocal potential is equivalent to a velocity-dependent one. Still, the role played by the velocity dependence of the potential is not very clear, and it is interesting to study the thermodynamics of a system with velocity-dependent interactions.

The form of velocity-dependent interaction usually considered is

$$H(\mathbf{x}_1,\mathbf{x}_2,\mathbf{p}_1,\mathbf{p}_2) = f(|\mathbf{x}_1-\mathbf{x}_2|)\frac{\mathbf{p}_1^2+\mathbf{p}_2^2}{2m} + V(|\mathbf{x}_1-\mathbf{x}_2|). \quad (1)$$

Green has shown that it is possible to fit the low-energy data with

$$f(|\mathbf{x}_1-\mathbf{x}_2|) = C\exp(-\alpha|\mathbf{x}_1-\mathbf{x}_2|^2).$$

The calculations that follow could be explicitly carried out for Green's potential as a particular case.

If the two-body Hamiltonian is taken to be

$$H(\mathbf{x}_1,\mathbf{x}_2,\mathbf{p}_1,\mathbf{p}_2) = \frac{\mathbf{p}_1^2}{2m}+\frac{\mathbf{p}_2^2}{2m}+\varphi(\mathbf{x}_1,\mathbf{x}_2)\frac{\mathbf{p}_1^2+\mathbf{p}_2^2}{2m} + U(\mathbf{x}_1,\mathbf{x}_2), \quad (2)$$

then the N-particle Hamiltonian has the form

$$H(\mathbf{x}_1,\cdots,\mathbf{x}_N,\mathbf{p}_1,\cdots,\mathbf{p}_N) = \frac{1}{N-1}\sum_{i<j=1}^{N} g(\mathbf{x}_i,\mathbf{x}_j)\frac{\mathbf{p}_i^2+\mathbf{p}_j^2}{2m} + V(R), \quad (3)$$

where

$$g(\mathbf{x}_i,\mathbf{x}_j) = 1+(N-1)\varphi(\mathbf{x}_i,\mathbf{x}_j) = g(\mathbf{x}_j,\mathbf{x}_i),$$
$$V(R) = \sum_{i<j} U(\mathbf{x}_i,\mathbf{x}_j).$$

When $g(\mathbf{x}_i,\mathbf{x}_j)=1$ we have the usual velocity-independent Hamiltonian. The quantum Hamiltonian that corresponds to (3) in the Weyl ordering will be as given in Eq. (4). [For any classical observable $g(p,q)$, the operator $\hat{g}$ corresponding to it in the Weyl ordering is given by

$$\hat{g} = \int g(p,q)e^{-i\eta q - i\tau p}e^{i\eta\hat{q}+i\tau\hat{p}}d\eta d\tau dq dp.$$

It may also be noted that Green uses the quantum Hamiltonian in the symmetric ordering instead of the Weyl ordering that we have chosen.]

$$\hat{H} = -\frac{\hbar^2}{6m(N-1)}\sum_{i<j}[g(\mathbf{x}_i,\mathbf{x}_j)(\nabla_{\mathbf{x}_i}^2+\nabla_{\mathbf{x}_j}^2) + \nabla_{\mathbf{x}_i}\cdot(g\nabla_{\mathbf{x}_i}) + \nabla_{\mathbf{x}_j}\cdot(g\nabla_{\mathbf{x}_j}) + (\nabla_{\mathbf{x}_i}^2+\nabla_{\mathbf{x}_j}^2)g]. \quad (4)$$

To study the thermodynamic properties of the system, we start with the Bloch equation for the temperature dependence of the canonical ensemble unnormalized density operator:

$$\frac{\partial\hat{\rho}}{\partial\beta} = -\tfrac{1}{2}(\hat{\rho}\hat{H}+\hat{H}\hat{\rho}). \quad (5)$$

Using the definition of the Wigner distribution function,

$$f_w(\mathbf{R},\mathbf{P}) = \left(\frac{1}{\hbar\pi}\right)^{3N}\int e^{(2i/\hbar)\mathbf{P}\cdot\mathbf{Y}}\rho(\mathbf{R}-\mathbf{Y},\mathbf{R}+\mathbf{Y})d\mathbf{Y}, \quad (6)$$

where $\mathbf{R}$, $\mathbf{P}$, and $\mathbf{Y}$ are $3N$-dimensional vectors, we have

$$\frac{\partial f_w}{\partial\beta} = -\tfrac{1}{2}\left(\frac{1}{\hbar\pi}\right)^{3N}\int e^{(2i/\hbar)\mathbf{P}\cdot\mathbf{Y}}[H(\mathbf{R}-\mathbf{Y})+H(\mathbf{R}+\mathbf{Y}) \times\rho(\mathbf{R}-\mathbf{Y},\mathbf{R}+\mathbf{Y})d\mathbf{Y}. \quad (7)$$

Substitution of the Hamiltonian (4) in (7) leads to

$$\frac{\partial f_w}{\partial\beta} = (A+B+C+\Theta)f_w, \quad (8)$$

with

$$\Theta f_w = -\cos[(\hbar/2)\sum_j \nabla_{\mathbf{p}_j}\cdot\nabla_{\mathbf{x}_j}]V(\mathbf{R})f_w(\mathbf{R},\mathbf{P}),$$

where $\nabla_{\mathbf{x}_j}$ operates only on V and

$$Af_w = \frac{\hbar^2}{12m(N-1)}\sum_{i<j}\{(C_{ij}^- + C_{ij}^+)[\nabla_{\mathbf{r}_i}^2+\nabla_{\mathbf{r}_j}^2 - (4/\hbar^2) \times(\mathbf{p}_i^2+\mathbf{p}_j^2)] + (4i/\hbar)(C_{ij}^+ - C_{ij}^-)(\mathbf{p}_i\cdot\nabla_{\mathbf{r}_i}+\mathbf{p}_j\cdot\nabla_{\mathbf{r}_j})\}f_w,$$

$$Bf_w = \frac{\hbar^2}{12m(N-1)} \sum_{i<j} \Big\{ \Big(\nabla_{\mathbf{r}_i} - \frac{2i}{\hbar} \mathbf{p}_i \Big) \cdot \Big[C_{ij}^- \Big(\nabla_{\mathbf{r}_i} - \frac{2i}{\hbar} \mathbf{p}_i \Big) \Big] + \Big(\nabla_{\mathbf{r}_j} - \frac{2i}{\hbar} \mathbf{p}_j \Big) \cdot \Big[C_{ij}^- \Big(\nabla_{\mathbf{r}_j} - \frac{2i}{\hbar} \mathbf{p}_j \Big) \Big] + \Big(\nabla_{\mathbf{r}_i} + \frac{2i}{\hbar} \mathbf{p}_i \Big) \cdot \Big[C_{ij}^+ \Big(\nabla_{\mathbf{r}_i} + \frac{2i}{\hbar} \mathbf{p}_i \Big) \Big] + \Big(\nabla_{\mathbf{r}_j} + \frac{2i}{\hbar} \mathbf{p}_j \Big) \cdot \Big[C_{ij}^+ \Big(\nabla_{\mathbf{r}_j} + \frac{2i}{\hbar} \mathbf{p}_j \Big) \Big] \Big\} f_w,$$

$$Cf_w = \frac{\hbar^2}{12m(N-1)} \sum_{i<j} \Big\{ \Big[\nabla_{\mathbf{r}_i}^2 + \nabla_{\mathbf{r}_j}^2 - \frac{4}{\hbar^2} (\mathbf{p}_i^2 + \mathbf{p}_j^2) \Big] \times (C_{ij}^+ + C_{ij}^-) + \frac{4i}{\hbar} [(p_i \cdot \nabla_{\mathbf{r}_i} + p_j \cdot \nabla_{\mathbf{r}_j}) \times (C_{ij}^+ - C_{ij}^-)] \Big\} f_w,$$

where

$$C_{ij}^{\pm} = g\Big(\mathbf{r}_i \pm \frac{i\hbar}{2} \nabla_{\mathbf{p}_i}, \mathbf{r}_j \pm \frac{i\hbar}{2} \nabla_{\mathbf{p}_j} \Big).$$

In the velocity-independent case $g = 1$, we naturally get the Oppenheim and Ross equation

$$\frac{\partial f_w}{\partial \beta} = \sum_{j=1}^{N} \Big(\frac{\hbar^2}{8m} \nabla_{\mathbf{r}_j}^2 - \frac{\mathbf{p}_j^2}{2m} \Big) f_w - \Theta f_w. \tag{9}$$

In calculating the quantum corrections. Oppenheim and Ross expand f_w as a power series in $\hbar$ and substitute in Eq. (9). Such a calculation would be very difficult to perform directly for Eq. (8).

We will therefore directly transcribe the Bloch equation in phase space by replacing $\hat{\rho}$ by f_w, $\hat{H}$ by the classical Hamiltonian, and the operator products of $\hat{\rho}$ and $\hat{H}$ by the homomorphic phase space products of corresponding quantities. The phase space product $A \times B$ of two classical functions $A(\mathbf{R}, \mathbf{P})$ and $B(\mathbf{R}, \mathbf{P})$ is defined as $A \times B$, with

$$(A \times B)\hat{} = \hat{A}\hat{B}.$$

For the one-particle case, this is given by

$$A(\mathbf{r}, \mathbf{p}) \times B(\mathbf{r}, \mathbf{p}) = \exp[(i\hbar/2)(\nabla_{\mathbf{r}_1} \cdot \nabla_{\mathbf{p}_2} - \nabla_{\mathbf{p}_1} \cdot \nabla_{\mathbf{r}_2})] \times A(\mathbf{r}_1, \mathbf{p}_1) B(\mathbf{r}_2, \mathbf{p}_2) \big|_{\substack{\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r} \\ \mathbf{p}_1 = \mathbf{p}_2 = \mathbf{p}}} \tag{10}$$

Thus the Bloch equation becomes

$$\frac{\partial f_w}{\partial \beta} = -\tfrac{1}{2}(f_w \times H + H \times f_w). \tag{11}$$

Expanding f_w in powers of h, we have

$$f_w = f_{cl}(1 + \hbar\varphi_1 + \hbar^2\varphi_2 + \cdots), \tag{12}$$

where

$$f_{cl} = e^{-\beta H}.$$

Substituting (12) in (10) and equating the coefficients of like powers of h, we have

$$\frac{\partial f_{cl}}{\partial \beta} = -Hf_{cl} \tag{13}$$

$$\frac{\partial \varphi_1}{\partial \beta} = -H\varphi_1. \tag{14}$$

The usual requirement that we obtain the correct high-temperature limit

$$f_w \to f_{cl} \qquad \beta \to 0,$$

shows that $\varphi_1 = 0$ and therefore all φ_{2n+1} vanish identically.

Thus we have

$$f_w = f_{cl}(1 + \hbar^2\chi_1 + \hbar^4\chi_2 + \cdots), \tag{15}$$

for any general Hamiltonian. Substituting (15) in (11), we get [apart from (13)]

$$\frac{\partial \chi_1}{\partial \beta} = -\frac{1}{8f_{cl}} \sum_k \sum_l [(\nabla_{r_{k_1}} \cdot \nabla_{p_{k_2}})(\nabla_{r_{l_1}} \cdot \nabla_{p_{l_2}}) + (\nabla_{r_{k_2}} \cdot \nabla_{p_{k_1}})(\nabla_{r_{l_2}} \cdot \nabla_{p_{l_1}}) - 2(\nabla_{r_{k_1}} \cdot \nabla_{p_{k_2}})(\nabla_{r_{l_2}} \cdot \nabla_{p_{l_1}})] f_{cl}(1) H(2). \tag{16}$$

After a straightforward calculation, we get

$$\frac{\partial \chi_1}{\partial \beta} = -\frac{1}{8f_u} \sum_k \sum_l [(\nabla_{\mathbf{p}_k} \times \nabla_{\mathbf{p}_l} H) \cdot (\nabla_{\mathbf{r}_k} \times \nabla_{\mathbf{r}_l} e^{-\beta H}) + (\nabla_{\mathbf{p}_k} \times \nabla_{\mathbf{p}_l} e^{-\beta H}) \cdot (\nabla_{\mathbf{r}_k} \times \nabla_{\mathbf{r}_l} H) - 2(\nabla_{\mathbf{p}_k} \times \nabla_{\mathbf{r}_l} H) \cdot (\nabla_{\mathbf{r}_k} \times \nabla_{\mathbf{p}_l} e^{-\beta H}]$$
$$-\frac{1}{8f_u} \sum_k \sum_l \{\nabla_{\mathbf{r}_{k_1}} \cdot [(\nabla_{\mathbf{r}_{l_1}} e^{-\beta H(1)} \cdot \nabla_{\mathbf{p}_k}) \nabla_{\mathbf{p}_l} H] + \nabla_{\mathbf{r}_{k_1}} \cdot [(\nabla_{\mathbf{r}_{l_1}} H(1) \cdot \nabla_{\mathbf{p}_k}) \nabla_{\mathbf{p}_l} e^{-\beta H}] - 2\nabla_{\mathbf{r}_{k_1}} [(\nabla_{\mathbf{p}_{l_1}} e^{-\beta H(1)} \cdot \nabla_{\mathbf{p}_k}) \nabla_{\mathbf{r}_l} H]\} \big|_{\substack{\mathbf{r}_{k_1} = \mathbf{r}_k \mathbf{r}_{l_1} = \mathbf{r}_l \cdot \\ \mathbf{p}_{k_1} = \mathbf{p}_k \mathbf{p}_{l_1} = \mathbf{p}_l}} \tag{17}$$

In particular for the Hamiltonian (3)

$$H = \frac{1}{N-1} \sum_{i<j} g(\mathbf{r}_i, \mathbf{r}_j) \frac{\mathbf{p}_i^2 + \mathbf{p}_j^2}{2m} + V(\mathbf{R}),$$

we can integrate Eq. (17) for χ_1 with the initial condition that $\chi_1 = 0$ when $\beta = 0$ (correct high-temperature limit), and obtain

$$\chi_1(\mathbf{R}, \mathbf{P}, \beta) = \frac{\beta^2}{8} \sum_k G_k(\mathbf{R}) \nabla_{\mathbf{r}_k}^2 H$$
$$-\frac{\beta^2}{8} \sum_k \sum_l \{\nabla_{\mathbf{r}_l} G_k(\mathbf{R}) \cdot [(\mathbf{p}_k \cdot \nabla_{\mathbf{r}_k} G_l(\mathbf{R}))\mathbf{p}_l + \mathbf{p}_k \times (\nabla_{\mathbf{r}_k} G_l(\mathbf{R}) \times \mathbf{p}_l)] + [\mathbf{p}_k \times \nabla_{\mathbf{r}_l} G_k(\mathbf{R})] \cdot [\nabla_{\mathbf{r}_k} G_l(\mathbf{R}) \times \mathbf{p}_l]\}$$
$$+\frac{\beta^3}{12} \sum_k \sum_l \{G_l(\mathbf{R})[\mathbf{p}_k \times \nabla_{\mathbf{r}_l} G_k(\mathbf{R})] \cdot (\nabla_{\mathbf{r}_k} H \times \mathbf{p}_l) + \nabla_{\mathbf{r}_l} G_k(\mathbf{R}) \cdot [(\mathbf{p}_k \cdot \nabla_{\mathbf{r}_k} H)\mathbf{p}_l + \mathbf{p}_k \times (\nabla_{\mathbf{r}_k} H \times \mathbf{p}_l]\}$$
$$-\frac{\beta^3}{24} \sum_k \sum_l \{G_k(\mathbf{R}) G_l(\mathbf{R})(\mathbf{p}_k \times \mathbf{p}_l) \cdot (\nabla_{\mathbf{r}_k} \times \nabla_{\mathbf{r}_l} H) + G_l(\mathbf{R}) G_k(\mathbf{R})[(\mathbf{p}_k \cdot \nabla_{\mathbf{r}_k}) \nabla_{\mathbf{r}_l} H] \cdot \mathbf{p}_l + \delta_{kl} G_k(\mathbf{R})(\nabla_{\mathbf{r}_k} H)^2 + \mathbf{p}_k \times (\nabla_{\mathbf{r}_k} \times \nabla_{\mathbf{r}_l} H) \cdot \mathbf{p}_l\}, \tag{18}$$

where

$$G_k(\mathbf{R}) = \frac{1}{m(N-1)} \sum_j [g(\mathbf{r}_j, \mathbf{r}_k) - g(\mathbf{r}_k, \mathbf{r}_k)].$$

When we take $g = 1$ the usual corrections for velocity-independent potentials are obtained. Knowing the correction for the first order, the corrections to higher orders in $\hbar$ could be computed using the recurrence relations

$$\frac{\partial(f_{cl}\chi_n)}{\partial\beta} = -\sum_{\substack{s,m \\ s+m=n}} \frac{(-1)^s}{2^{2s}} \left[\sum_i (\nabla_{\mathbf{r}_{i_1}} \cdot \nabla_{\mathbf{p}_{i_2}} - \nabla_{\mathbf{r}_{i_2}} \cdot \nabla_{\mathbf{p}_{i_1}})\right]^{2s} \times (f_{cl}\chi_m)(1)H(2)|_{1=2}.$$

Knowing the first-order correction for the Wigner distribution function the first nonvanishing corrections to all the phase space distribution functions corresponding to different operator orderings can be calculated.

Knowning the quantum corrections to the distribution function, the quantum corrections to the partition function can immediately be calculated. The partition function Z is given by

$$Z = \int f_w(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N, p_1 \cdots p_N) d\mathbf{r}_1 d\mathbf{r}_2 \cdots dp_N,$$

where $f_w(\mathbf{r}_1 \cdots p_N)$ is the unnormalized distribution function, given by (6). We immediately get an expansion for Z in powers of h^2.

$$Z = Z_{cl}(1 + \hbar^2 p_1 + \hbar^4 p_2 + \cdots),$$

where

$$Z_{cl} = \int f_{cl}(\mathbf{r}_1 \ldots p_N) d\mathbf{r}_1 \ldots dp_N,$$

$$P_s = \int \chi_s f_{cl} d\mathbf{r}_1 \ldots d\mathbf{p}_N = \langle \chi_s \rangle_{cl}.$$

For our case, the first correction to the partition function will be the classical ensemble average of χ_1 [given in (18)]. But these calculations can be explicitly carried out only after the functions $G_k(\mathbf{R})$ are known i.e., when a specific form for the function $g(\mathbf{r}_i, \mathbf{r}_j)$ is chosen.

As has been pointed out, the actual relevance of velocity-dependent potentials for nucleon–nucleon interaction is not very clear, though they have been widely employed in the analysis of N–N scattering, optical and shell model calculations, etc. One also obtains a velocity-dependent potential from some meson field theoretic calculations. In this context, the thermodynamic properties of bulk nuclear matter may throw some light on this problem. The results above could be used to check the relevance of specific velocity dependent potentials such as the one given by Green.

We are grateful to Dr. T. S. Shankara for many stimulating discussions. We also wish to thank Dr. E. Wolf for kindly going through the manuscript and suggesting many improvements.

* Research supported in part by the Air Force Office of Scientific Research.

[1] E. P. Wigner, Phys. Rev. **40**, 749 (1950).
[2] I. Oppenheim and J. Ross, Phys. Rev. **107**, 28 (1957).
[3] K. Imre, E. Ozizmir, M. Rosenbaum and P. F. Zweifel, J. Math. Phys. 8, 1097 (1967).
[4] G. S. Agarwal and E. Wolf, Phys. Rev. D 2, 2161 (1970); 2, 2187 (1970); and 2, 2206 (1970).
[5] M. D. Srinivas, M. S. Sriram, and T. S. Shankara, Int. J. Theor. Phys. 5, 263 (1972).
[6] A. M. Green, Nucl. Phys. **33**, 218 (1962).
[7] C. H. Fay, Phys. Rev. **50**, 560 (1935).
[8] J. A. Wheeler, Phys. Rev. **50**, 643 (1936).
[9] See, for example, A. E. S. Green, T. Sawada, and D. S. Saxon, *Nuclear Independent Particle Model* Academic, New York, 1968), pp. 164–203.
[10] M. Razavy, Phys. Rev. 171, 1201 (1968).

General expressions for the position and spin operators of relativistic systems

M. Lorente* and P. Roman

Boston University, Department of Physics, Boston, Massachusetts 02215
(Received 9 July 1973)

Some general assumptions based on the physical properties of elementary and of composite quantized systems lead in a natural way to general expressions for the position and spin operators of massive particles. All previously proposed position and spin operators, defined in the enveloping algebra of the Poincaré group, appear as special cases. By algebraic arguments a general local position operator is obtained which is proved to coincide under special conditions with the Pryce-Newton-Wigner position operator.

I. INTRODUCTION

During the past 25 years several types of position operators have been proposed as candidates for the observables whose eigenvalues are the measured values of the position of some elementary particle or the center of mass of a composite quantized system. These studies recently acquired special emphasis due to the increased interest in representing the complicated composite nature of hadrons (experimentally manifesting itself in a partonlike structure, Reggeized or dual behaviour) by means of dynamical models with unspecified constituents.

In principle, there are two distinct approaches possible. One may consider the description of the quantized system in the Poincaré framework and seek to identify the covariant position operator in the *enveloping algebra*.[1-7]

Alternatively, one may consider dynamical groups that are larger than the Poincaré group and which contain a covariant position operator in the Lie algebra.[8-10] In this paper we shall consider only the first approach.

The starting point in most of the mentioned papers has been the investigation of the physical properties of the position and spin observables or their commutation relations with other observables. A systematic study of these position operators reveals that all of them correspond to the same algebraic structure and are closely related to the spin operator. It is the aim of our work to derive from the definition of the total angular momentum—as it was proposed by Synge in the classical sense—a general expression for the spin and position operator (Sec. II). Subsequently, we will show that under certain conditions the specific position and spin operators are special cases of the general expressions found by us (Sec. III).

Our systematic approach contains the following assumptions.

(i) In order to construct the position and spin operators, Q_μ and $S_{\mu\nu}$, respectively, we have at our disposal only the generators $J_{\mu\nu}$ and P_μ of the Poincaré group.

(ii) Q_μ and $S_{\mu\nu}$ are Hermitian operators acting on state vectors which belong to some unitary representation of the Poincaré group.

(iii) Q_μ and $S_{\mu\nu}$ are related by the expression

$$J_{\mu\nu} = Q_\mu P_\nu - Q_\nu P_\mu + S_{\mu\nu}, \tag{1}$$

which at the same time is the defining equation for both operators in analogy to the classical picture.[11]

(iv) Some properties of covariance, locality, and tensorial character are not shared by all Q_μ and $S_{\mu\nu}$ and therefore are studied in the special cases.

If the representation of the Poincaré group is realized in a field theory, assumption (ii) does not imply that the field system carries an infinite dimensional representation of $SL(2, C)$; fields with a finite number of components [spanning a nonunitary representation of $SL(2,C)$] are permitted.

Assumption (iii) is the translation to the language of operators of the conservation of total angular momentum. We know from the canonical formalism in field theory that the invariance of the Lagrangian under translations and Lorentz transformations leads to the conservation of energy–momentum and of the total angular momentum, respectively, which are isomorphic to the generators P_μ and $J_{\mu\nu}$. The same is true in the case of first quantization where the four-momentum and the total angular momentum satisfy the same commutation relations as P_μ and $J_{\mu\nu}$, respectively. Therefore in Eq. (1), $J_{\mu\nu}$ stands for the total angular momentum operator $Q_\mu P_\nu - Q_\nu P_\mu$ for the external part or orbital angular momentum operator, and $S_{\mu\nu}$ for the internal part or spin operator.

In this paper we adopt natural units ($\hbar = c = 1$) and the metric tensor is $g_{oo} = -g_{kk} = 1, g_{\mu\nu} = 0$ for $\mu = \nu$. The scalar product of two four-vectors $a^\mu b_\mu$ is denoted by $a.b$.

II. GENERAL EXPRESSIONS FOR THE POSITION AND SPIN OPERATORS

The most general operator Q_μ satisfying assumptions (i–iii) is found in the following way. Any antisymmetric second-rank tensor that belongs to the enveloping algebra of the Poincaré group can be identically rewritten in the form[12]

$$L_{\mu\nu} = (L_{\mu\lambda}P^\lambda P_\nu - L_{\nu\lambda}P^\lambda P_\mu)/P^2 + (L^*_{\mu\lambda}P^\lambda P_\nu - L^*_{\nu\lambda}P^\lambda P_\mu)^*/P^2. \tag{2}$$

(This can be proved by substituting in the second term on the right the explicit expression for the dual tensor and using the properties of the Levi-Civitá symbols). If we take for $L_{\mu\nu}$ the external angular momentum given by (1),

$$L_{\mu\nu} = Q_\mu P_\nu - Q_\nu P_\mu = J_{\mu\nu} - S_{\mu\nu}, \tag{3a}$$

then, since $(Q_\mu P_\nu - Q_\nu P_\mu)^* P^\nu = 0$, we obtain from (2) the relation

$$J_{\mu\nu} - S_{\mu\nu} = (J_{\mu\lambda} - S_{\mu\lambda})P^\lambda P_\nu/P^2 - (J_{\nu\lambda} - S_{\nu\lambda})P^\lambda P_\mu/P^2. \tag{3b}$$

Comparison of (3a) and (3b) gives

$$Q_\mu = (J_{\mu\nu} - S_{\mu\lambda})P^\lambda/P^2 + AP_\mu/P^2, \tag{4}$$

where A is an arbitrary scalar operator. For massive particles we substitute $P^2 = M^2$. (Obviously this equation cannot be used for massless particles.)

Since $J_{\mu\nu}$ and A in general do not commute with P_μ, we must symmetrize (4):

$$Q_\mu = \{J_{\mu\lambda} - S_{\mu\lambda}, P^\lambda/2M^2\} + \{A, P_\mu/2M^2\}. \tag{5}$$

This is the most general expression for the position operator satisfying conditions (i–iii).

In (5) the existence of $S_{\mu\nu}$ is presupposed, and so one needs an expression for $S_{\mu\nu}$ which does not contain Q_μ explicitly. In order to eliminate Q_μ in $S_{\mu\nu}$ given by (1), we take the dual tensor on both sides and contract with P^ν:

$$J^*_{\mu\nu}P^\nu = S^*_{\mu\nu}P^\nu. \tag{6}$$

A particular solution of (6) is obtained with the help of the identity (2) for $L_{\mu\nu} = J_{\mu\nu}$:

$$J_{\mu\nu} = (Y_\mu P_\nu - Y_\nu P_\mu)/P^2 + (W_\mu P_\nu - W_\nu P_\mu)^*/P^2, \tag{7}$$

where $Y_\mu \equiv J_{\mu\lambda}P^\lambda$, (Shirokov tensor),

$$W_\mu \equiv J^*_{\mu\lambda}P^\lambda, \text{(Pauli–Lubanski tensor)}.$$

Now the spin operator

$$S^0_{\mu\nu} = (W_\mu P_\nu - W_\nu P_\mu)^*/P^2 \tag{8}$$

is a solution of (6) as can be shown by taking the dual tensor of (8) and contracting with P^ν. Therefore, the most general solution of (6) will be

$$S_{\mu\nu} = S^0_{\mu\nu} + T_{\mu\nu},$$

with

$$T^*_{\mu\nu}P^\nu = 0. \tag{9}$$

One way to express the general solution of (6) is to parametrize $S_{\mu\nu}$ by choosing an arbitrary timelike four-vector x^ν such that

$$S_{\mu\nu}x^\nu = 0. \tag{10}$$

[The requirement $x^2 > 0$ comes from the condition that $P.x$ in (12) and (15) cannot be zero, which is true if and only if $x^2 > 0$ and $P^2 > 0$.]

In both constraints (6) and (10) for the spin operator $S_{\mu\nu}$, the first equation is implied by the last three. Therefore, we are left with six independent equations which fix completely the components of the antisymmetric tensor $S_{\mu\nu}$. The additional vector x^ν will help us to express Q_μ and $S_{\mu\nu}$ independently. Contracting (4) with x^μ and using (10) we get

$$Q.x = Y.x/M^2 + AP.x/2M^2. \tag{11}$$

Contracting also (1) with x^ν and using (10) and (11), we obtain a general expression for Q_μ independent of $S_{\mu\nu}$, which after symmetrization reads

$$Q_\mu = \{J_{\mu\lambda}, x^\lambda/2P.x\} + \{AP_\mu/2P.x\}, \tag{12}$$

where all scalar operators have been absorbed in A.

Inserting (12) in (1) we obtain, after symmetrization, a general expression for $S_{\mu\nu}$ independent of Q_μ; namely,

$$S_{\mu\nu} = J_{\mu\nu} - \{J_{\mu\lambda}, x^\lambda P_\nu/2P.x\} + \{J_{\nu\lambda}, x^\lambda P_\mu/2P.x\}. \tag{13}$$

We still want an expression for the spin operator of the form $S_{\mu\nu} = S^0_{\mu\nu} + T_{\mu\nu}$, with $S^0_{\mu\nu}$ given by (8). Comparison of (13) with the identity

$$J_{\mu\nu} = \{J_{\mu\lambda}, x^\lambda P_\nu/2P.x\} - \{J_{\nu\lambda}, x^\lambda P_\mu/2P.x\} + (W_\mu x_\nu - W_\nu x_\mu)^*/P.x \tag{14}$$

[which can be proved in the same way as (2)], gives

$$S_{\mu\nu} = (W_\mu x_\nu - W_\nu x_\mu)^*/P.x.$$

Substituting in this expression $W_\mu \equiv J^*_{\mu\lambda}P^\lambda = S^{0*}_{\mu\lambda}P^\lambda$ given by (6), we obtain

$$S_{\mu\nu} = (S^{0*}_{\mu\lambda}P^\lambda x_\nu - S^{0*}_{\nu\lambda}P^\lambda x_\mu)^*.$$

Using the properties of the Levi-Civitá antisymmetric tensor, this becomes

$$S_{\mu\nu} = S^0_{\mu\nu} - S^0_{\mu\lambda}x^\lambda P_\nu/P.x + S^0_{\nu\lambda}x^\lambda P_\mu/P.x. \tag{15}$$

Equations (12) and (15) *are the most general expressions for the position and spin operators, but they depend on the timelike vector* x^μ, which may be taken from the enveloping algebra of the Poincaré group or from without. In the second case we have a violation of covariance, which however will be proved necessary in order to satisfy the requirement of locality (Sec. III).

In order to classify the particular position operators we will follow the classification used by Fleming in Ref. 6.

III. SPECIAL CASES

"The center of inertia"

As we mentioned in the Introduction, the particular position and spin operators are special cases of the general expressions (5) and (15). Choosing $x^\nu = P^\nu$ in (10), we have

$$S_{\mu\nu}P^\nu = 0, \tag{16}$$

which was postulated by Synge[11] in the classical sense, and is interpreted now in the quantum mechanical framework. Substituting $x^\nu = P^\nu$ in (15) and using $S^0_{\mu\nu}P^\nu = 0$, we obtain

$$S_{\mu\nu} = S^0_{\mu\nu} \tag{17}$$

for the spin operator, and substituting (16) in (5) we obtain

$$Q_\mu = Y_\mu/M^2 + \{A, P_\mu/2M^2\} \tag{18}$$

for the position operator.

Since A is still undefined, we can impose some additional conditions.

(i) The requirement $Q_0 = t$, where t is the time variable, is satisfied if $A = (tM^2 - Y_0)/P_0$, so that

$$Q_\mu = Y_\mu/M^2 + tP_\mu/P_0 - \{Y_0, P_\mu/2M^2P_0\}.$$

This is the "center" defined by Pryce[13] in case (d) and corresponds to the "center of mass" (to be defined

later) in the rest frame, and then transformed to any frame by a Lorentz transformation.

(ii) The condition $Q.P = 0$ is satisfied if $A = 0$, hence

$$Q_\mu = Y_\mu/M^2.$$

This is called the "center" by Chakrabarti[14] and Bacry.[15]

(iii) The condition $Q.n = 0$, n being a timelike unit vector is satisfied if $A = -Y.n/P.n$, and we have

$$Q_\mu = Y_\mu/M^2 - \{Y.n, P_\mu/2M^2P.n\},$$

as proposed by Bacry.[16]

(iv) The requirement $Q_0 = t$, when $\mathbf{P} = 0$, gives $A = Mt$, and so

$$Q_k = J_{k0}/M.$$

This is the "center of inertia" defined by Fleming[17] as the operator which describes the point in space-time occupied by the "center of mass" in the rest system. This center transformed to an arbitrary frame gives the expression (18).

From (17) and (18) it can be proved that all position operators (i–iv) transform as four-vector operators of the Poincaré group, and the corresponding spin operators transform as antisymmetric second-rank tensor operators. Expressions (17) and (18) are, therefore, covariant and independent of the frame of reference. The last property (that of being independent of the frame of description), is ascribed by Fleming to "point" operators. But, unfortunately, the "center of inertia" operators (i–iv) do not have self-commuting components, as can be checked, and thus they are not local operators.

"The center of mass"

Taking $x^\nu = n^\nu$ in (10), n^ν being a timelike unit vector representing the hyperplane from where the system is observed, we have

$$S_{\mu\nu}n^\nu = 0, \tag{19}$$

and substituting $x^\nu = n^\nu$ in (12) and (15) we get

$$Q_\mu = \{J_{\mu\lambda}, n^\lambda/2P.n\} + \{A, P_\mu/2P.n\}, \tag{20}$$

$$S_{\mu\nu} = S^0_{\mu\nu} - S^0_{\mu\lambda}n^\lambda P_\nu/P.n + S^0_{\nu\lambda}n^\lambda P_\mu/P.n. \tag{21}$$

Given n^λ, the expression (21) for the spin operator is completely determined, but in (20) we will have an arbitrary scalar operator. As before, special conditions remove this arbitrariness.

(i) $Q_0 = t$, t being the time coordinate, gives

$$Q_\mu = \{J_{\mu\lambda}, n^\lambda/2P.n\} - \{J_{0\lambda}, n^\lambda P_\mu/P_0P.n\} + tP_\mu/P_0.$$

This is the "center of mass" proposed by Pryce,[18] the components of which are the mean values of the coordinates for the constituents of the physical system weighted with their dynamical masses (energies).

(ii) $Q_\mu P^\mu = 0$ gives

$$Q_\mu = \{J_{\mu\lambda}, n^\lambda/2P.n\} + \{Y.n, P_\mu/2M^2P.n\}.$$

This is called "the pseudo-center" by Chakrabarti[19] because it is supposed to coincide with the "center" in the rest system.

(iii) $Q_\mu n^\mu = 0$, gives

$$Q_\mu = \{J_{\mu\lambda}, n^\lambda/2P.n\}.$$

(iv) $Q_0 = t$, for $n = (1, 0, 0, 0)$, gives

$$Q_k = \{J_{k0}, P_0^{-1}\} + tP_k/P_0.$$

This is the "center of mass" derived by Fleming,[20] for a composite system with the help of the total angular momentum and energy–momentum tensors.

Obviously, expressions (20) and (21) depend on the particular hyperplane with the representative unit vector n^λ, and *they do not transform as vector operators or as antisymmetric tensor operators, respectively*.[21] Furthermore, the "center of mass" operators (i–iv) are not local operators because their components do not commute among themselves.[21]

The local position operator

The particular choices of x^ν made above do not lead to a local operator. In order to find such an operator, whose components satisfy

$$[Q_\mu, Q_\nu] = 0, \tag{22}$$

we must find other values for x^ν. Fleming[22] has proved that (22) is true if and only if

$$[S_{\mu\nu}, S_{\rho\sigma}] = i(g_{\nu\rho}S_{\mu\sigma} - g_{\mu\rho}S_{\nu\sigma} - g_{\mu\sigma}S_{\rho\nu} + g_{\nu\sigma}S_{\rho\mu}). \tag{23}$$

The most general form of x^λ is of the form

$$x^\lambda = an^\lambda + bP^\lambda, \tag{24}$$

where a, b are arbitrary constants and n^λ is a timelike unit vector. Inserting (15) into (23) [with x^λ given by (24)] and using the properties

$$[S^0_{\mu\nu}, n_\lambda] = [S^0_{\mu\nu}, P_\lambda] = [S^0_{\mu\nu}, (P.n)^{-1}] = 0,$$

we find

$$[S_{\mu\nu}, S_{\rho\sigma}] = i\left(g_{\nu\rho} - \frac{x_\nu P_\rho + x_\rho P_\nu}{P.x} + \frac{x^2P_\nu P_\rho}{(P.x)^2}\right)S_{\mu\sigma} + \text{antisymmetrizing terms}. \tag{25}$$

Condition (23) is satisfied if and only if

$$\frac{x^2P_\nu P_\rho}{(P.x)^2} - \frac{x_\nu P_\rho + x_\rho P_\nu}{P.x} = 0$$

for all values of ν and ρ. Multiplying both sides in the last expression by $(P.x)^2$ and substituting x_ν from (24), we get

$$(a^2 - b^2M^2)P_\nu P_\rho - a(aP.n + bM^2)(n_\nu P_\rho + n_\rho P_\nu) = 0. \tag{26}$$

For arbitrary P_ν and P_ρ this equation has a nontrivial solution (i.e., different from $a = b = 0$) if

$$a^2 - b^2M^2 = 0 \text{ and } aP.n + bM^2 = 0 \tag{27a}$$

or if

$$a^2 - b^2M^2 = 0 \text{ and } n_\nu P_\rho + n_\rho P_\nu = 0. \tag{27b}$$

The condition $a^2 - b^2M^2 = 0$ gives $b = \pm a/M$. For positive energy state functions, onto which the position operator is applied, the minus sign is excluded by the following argument:

We have shown that $x_\lambda = an_\lambda + bP_\lambda$ is a timelike

four-vector. Inserting the two values of b in this vector, we obtain

$$x_\lambda = a(n_\lambda \pm P_\lambda/M);$$

hence

$$x^2 = 2a^2(1 \pm P.n/M). \tag{28}$$

But $P.n$ is the time component of the four-momentum in the new frame, n^λ being its characteristic vector. Since, for positive energy states,

$$P'_0 = +\sqrt{M^2 + \mathbf{P}'^2} \geq M, \tag{29}$$

it follows from (28) that the minus sign would lead to a spacelike vector $x^2 < 0$.

Inserting the value of $b = a/M$ into the second condition of (27a), we obtain $P.n + M = 0$, which is impossible for positive energy state functions. Therefore we are left with the second choice, Eq. (27b). The first condition gives, as before, $b = a/M$, and the second condition cannot be satisfied simultaneously for all values of ν and ρ. [The solution $n = (0, 0, 0, 0)$ is impossible because n is a unit vector.] The maximal set of equations (26) which satisfy the conditions (27b) is obtained if we restrict ourselves to the consideration of the spatial components $\rho, \nu = 1, 2, 3$ and choose $n = (1, 0, 0, 0)$.

Inserting $x_\lambda = a(n_\lambda + P_\lambda/M)$ into (12) and (15), dividing by a, and using

$$(S^0_{23}, S^0_{31}, S^0_{12}) = (W_0\mathbf{P} - \mathbf{W}P_0)/M^2 \equiv \mathbf{S},$$
$$(S^0_{01}, S^0_{02}, S^0_{03}) = (\mathbf{W} \times \mathbf{P})/M^2,$$
$$(J_{23}, J_{31}, J_{12}) \equiv \mathbf{J},$$
$$(J_{01}, J_{02}, J_{03}) \equiv \mathbf{N},$$

we obtain

$$\mathbf{Q} = -\mathbf{N}M^{-1} + \tfrac{1}{2}\{\tfrac{1}{2}(\mathbf{J} \times \mathbf{P} - \mathbf{P} \times \mathbf{J}), M^{-1}(P_0 + M)^{-1}\} + \tfrac{1}{2}\{A, \mathbf{P}(P_0 + M)^{-1}\}, \tag{30}$$

$$\mathbf{S} = (W_0\mathbf{P} - \mathbf{W}P_0)M^{-2} + (\mathbf{W} \times \mathbf{P}) \times \mathbf{P}(P_0 + M)^{-2}M^{-2}. \tag{31}$$

Expression (30) gives the local position operator, and (31) is the corresponding spin operator satisfying the standard commutation relations

$$[S_j, S_k] = i\epsilon_{jkl}S_l.$$

Again (30) contains some arbitrary scalar operator A. If we take $A = N.P(MP_0)^{-1}$, we obtain from (12) and (30)

$$Q_0 = 0,$$

$$M\mathbf{Q} = -\mathbf{N} + \tfrac{1}{2}\{\tfrac{1}{2}(\mathbf{J} \times \mathbf{P} - \mathbf{P} \times \mathbf{J}), (P_0 + M)^{-1}\} + \tfrac{1}{2}\{\mathbf{N}.\mathbf{P}, \mathbf{P}P_0^{-1}(P_0 + M)^{-1}\}, \tag{32}$$

which coincides with the local "center of mass" proposed by Pryce,[23] provided we symmetrize it and use our convention for the generators J and N. Furthermore, Newton and Wigner[24] derived a position operator from physical considerations of localized elementary particle states which for spin $\frac{1}{2}$ was shown to coincide with Pryce's expression. In Appendix A we sketch a proof which shows the equality of the Pryce and of the Newton–Wigner position operator for arbitrary spin.

It can be checked[25] that the local position operator (32) satisfies the Heisenberg commutation relations and commutes with the spin operator (31). The spatial parts of these two operators transform as a vector and as a second-rank tensor, respectively.[26]

IV. CONCLUDING REMARKS

The initial assumption of Eq. (1) was based principally on two facts: (a) it is the natural covariant generalization of the total angular momentum $J_{\mu\nu}$, (b) the Lie algebra generated by the $J_{\mu\nu}$ [which as defined through (1), belongs to the enveloping algebra of the Poincaré group] and by the linear momentum P_μ is isomorphic to the Lie algebra of the Poincaré group. These assumptions have been proven to be fruitful for the classification of all proposed position and spin operators defined in the enveloping algebra of the Poincaré group. At the same time, we are led to a local position operator from pure algebraic considerations.

Similar assumptions and methods can be applied to groups larger than the Poincaré group, for example, to the inhomogenous deSitter group[27] or the conformal group. In that case, viewed from the restriction to the Poincaré framework, the locality condition is not necessarily in contradiction with the covariance and frame-independence requirements.

ACKNOWLEDGMENTS

The authors want to express their thanks to P. Huddleston for many critical comments and discussions. The first author (M. L.) is indebted to Boston University for the hospitality extended to him and for financial support during the period in which this work was done. He is also indebted to Professor A. Galindo for calling his attention to the problem of uniqueness of the Newton–Wigner position operator.

APPENDIX A: EQUIVALENCE OF THE PRYCE AND NEWTON-WIGNER POSITION OPERATORS FOR ARBITRARY SPIN

The N-W position operator can be written[28]

$$\mathbf{q} = E \prod_{\alpha=1}^{2s} \tfrac{1}{2}(1 + \gamma^0_\alpha) \frac{2^{2s}P_0^{2s+1/2}}{(P_0 + M)^s} \mathbf{X} \frac{P_0^{-1/2}}{(P_0 + M)^s} \times \prod_{\alpha=1}^{2s} \tfrac{1}{2}(1 + \gamma^0_2)E, \tag{A1}$$

where $\prod_{\alpha=1}^{2s} \tfrac{1}{2}(1 + \gamma^0_\alpha)$ has been inserted before the operator E on the right because it is a projection operator. We take the expectation value $(\psi, \mathbf{q}\psi)$, with respect to some state function. Using the commutation relation $[P_0, \mathbf{X}] = i\mathbf{P}/P_0$, we can commute $P_0^{s+1/2}$ to the right side of $\mathbf{X}$, and we will obtain a term proportional to $(\psi, \mathbf{P}\psi)$ and another term of the form

$$\left(\left(\frac{2P_0}{P_0 + M}\right)^s \prod_{\alpha=1}^{2s} \tfrac{1}{2}(1 + \gamma^0_\alpha)E\psi, \mathbf{X}\left(\frac{2P_0}{P_0 + M}\right)^s \times \prod_{\alpha=1}^{2s} \tfrac{1}{2}(1 + \gamma^0_\alpha)E\psi\right), \tag{A2}$$

where we have used the Hermitian property of E and $(1 + \gamma^0)$. Expression (A2) can be decomposed into a sum of $2s$ terms,

$$\sum_{\alpha=1}^{2s} (F_1F_2 \ldots F_{2s}\psi, F_1F_2 \ldots F_{\alpha-1}F_{\alpha+1} \ldots F_{2s}\mathbf{X}F_\alpha\psi), \tag{A3}$$

where

$$F_\alpha \equiv \left(\frac{2P_0}{P_0 + M}\right)^{1/2} \tfrac{1}{2}(1 + \gamma^0_\alpha)E_\alpha.$$

The operator E_α acting on ψ gives the positive energy part, and F_α, in this case, is equivalent to the unitary

transformation used by Foldy and Wouthuysen[29] applied to the positive energy wavefunctions, namely

$$F_\alpha\psi = e^{iS_\alpha}\psi_+ = \left(\frac{2P_0}{P_0+M}\right)^{1/2} \tfrac{1}{2}(1+\gamma_\alpha^0)\psi_+.$$

So we can write (A3) in the form

$$\sum_{\alpha=1}^{2s}(e^{iS_1}e^{iS_2}\cdots e^{iS_{2s}}\psi_+, e^{iS_1}\cdots e^{iS_{\alpha-1}}e^{iS_{\alpha+1}}\cdots e^{iS_{2s}}\mathbf{X}e^{iS_\alpha}\psi_+).$$

But the Foldy–Wouthuysen transformations acting on the operator $\mathbf{X}$ give the Pryce position operator $\mathbf{q}^P$ for spin-$\frac{1}{2}$ particles,[30]

$$e^{-iS}\mathbf{X}e^{iS} = \mathbf{X} + \frac{i\gamma^0\boldsymbol{\alpha}}{2P_0} - \frac{i\gamma^0(\boldsymbol{\alpha}.\mathbf{P})\mathbf{P} + (\boldsymbol{\sigma}\times\mathbf{P})|\mathbf{P}|}{2P_0(P_0+M)|\mathbf{P}|} = \mathbf{q}^P. \quad (A4)$$

Since e^{iS} is a unitary operator, (A3) simplifies to the form

$$\sum_{\alpha=1}^{2s} 1\otimes 1\otimes\cdots\otimes(\psi_+, e^{-iS_\alpha}\mathbf{X}e^{iS_\alpha}\psi_+)\otimes\cdots\otimes 1$$
$$= \sum_{\alpha=1}^{2s} 1\otimes 1\otimes\cdots\otimes(\psi_+, \mathbf{q}^P_{(\alpha)}\psi_+)\otimes\cdots\otimes 1. \quad (A5)$$

Now remembering that in (A4) $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ are the generators of the Lorentz group in the $\frac{1}{2}$ spin representation, we can substitute (A4) into (A5) and we will obtain instead of $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ the generators I^{0k} and I^{ij} for the representation of spin s, in the form given by Bargmann and Wigner,[31]

$$I^{\mu\nu} = \sum_{\alpha=1}^{2s} 1\otimes 1\otimes\cdots\otimes \tfrac{1}{2}i\gamma_\alpha^\mu\gamma_\alpha^\mu\otimes\cdots\otimes 1.$$

But in this way (A5) becomes precisely the expectation value of Pryce's position operator $\mathbf{Q}$ in any arbitrary representation of spin s, as defined in (32). Since however, we disregarded in the above argument the term proportional to $(\psi, \mathbf{P}\psi)$, the Newton–Wigner and the Pryce operators may differ by an expression that corresponds to the last term in our Eq. (30).

APPENDIX B: ON THE UNIQUENESS OF THE NEWTON-WIGNER POSITION OPERATOR

The locality condition (22), when applied to the general expression (4), does not lead to a unique position operator. As we saw in Appendix A, both the Pryce and the Newton–Wigner operators satisfy the locality condition, yet they differ by a term proportional to $\mathbf{P}$.

Using a different approach, Weidlich and Mitra,[32] as well as Galindo[33] have shown that hermiticity, conjugacy to $\mathbf{P}$, locality, correct transformation behavior under rotations, space and time reflections do not uniquely determine the position operator. In fact, all these conditions are satisfied by

$$\mathbf{Q} = B\mathbf{Q}^{(0)}B^{-1}, \quad (B1)$$

where $\mathbf{Q}^{(0)}$ is the Newton–Wigner position operator and B is a unitary operator that commutes with the representation of the Euclidean group E_2.

It is not difficult to show that the position operator given by (B1) is a special case of our general expression (4). Indeed, as we proved in Appendix A, the $\mathbf{Q}^{(0)}$ satisfies (32) and therefore also the spatial part of Eq. (1),

$$\mathbf{J} = \mathbf{Q}^{(0)}\times\mathbf{P} + \mathbf{S}, \quad (B2)$$

is satisfied. If we apply the unitary transformation B onto Eq. (B2), then, using Eq. (B1) and the fact that (by construction) B commutes[34] with $\mathbf{J}$, $\mathbf{P}$, and $\mathbf{S}$, we get

$$\mathbf{J} = \mathbf{Q}\times\mathbf{P} + \mathbf{S}. \quad (B3)$$

Furthermore, $\mathbf{Q}$ obviously satisfies the locality condition (22). Since the most general solution satisfying (1) and (22) is given by (30), it is now clear that the position operator (B1) is a special case of our most general solution. To make (30) equivalent to (B1), it is necessary to demand correct behavior under parity Π and time reversal Θ. It is easy to show[35] that the conditions

$$\Pi\mathbf{Q}\Pi^{-1} = -\mathbf{Q} \text{ and } \Theta\mathbf{Q}\Theta^{-1} = \mathbf{Q}$$

imply

$$\Pi A\Pi^{-1} = A \text{ and } \Theta A\Theta^{-1} = -A,$$

respectively, for the arbitrary scalar operator A in (30). But even these additional conditions do not completely fix the position operator (30). As Galindo[33] has shown, uniqueness of the position operator (B1) is obtained by imposing also a suitable regularity condition on its eigenfunctions. It turns out that the regularity condition invoked by Newton and Wigner[36] renders (B1) unique and indeed leads precisely to the Newton–Wigner position operator.

Note added in proof: We recently became aware of the excellent review article by A. J. Kálnay *[Studies in the Foundations, Methodology and Philosophy of Science,* (Springer, New York, 1971), Vol. 4, p. 93] which discusses the various approaches to the localization problem with great clarity.

* On leave of absence from the University of Madrid, Departamento de Física Teórica, Madrid-3. Sponsored by G.I.F.T.

[1] M. H. L. Pryce, Proc. R. Soc. A **195**, 62 (1948).
[2] T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).
[3] A. Chakrabarti, J. Math. Phys. **4**, 1223 (1963).
[4] H. Bacry, Phys. Lett. **5**, 37, 1963.
[5] H. Bacry, J. Math. Phys. **5**, 109 (1964).
[6] G. N. Fleming, Phys. Rev. B **137**, 188 (1965).
[7] F. Gürsey and S. Orfanidis, Nuovo Cimento A **11**, 225 (1972).
[8] L. Castell, Nuovo Cimento **49**, 285 (1967).
[9] J. E. Johnson, Phys. Rev. **181**, 1755 (1969).
[10] J. J. Aghassi, P. Roman, and R. M. Santilli, Phys. Rev. D **1**, 2753 (1970).
[11] J. L. Synge, *Special Relativity* (North-Holland, Amsterdam, 1965), pp. 219–222. The operators Q_μ and P_μ do not commute but, due to the Heisenberg relations, in most cases Eq. (1) is still Hermitian; otherwise, it should be symmetrized properly.
[12] See, for instance, Ref. 4. The dual tensor $L^*_{\mu\nu}$ is defined by $1/2\ \epsilon_{\mu\nu\rho\sigma}T^{\rho\sigma}$, where $\epsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor with $\epsilon_{0123} = +1$.
[13] See Ref. 1, Eq. (2.11).
[14] See Ref. 3, Eqs. (2.3) for the position operator and (2.4) for the corresponding spin operator.
[15] See Ref. 4, Eq. (15).
[16] See Ref. 5, Eq. (19).
[17] See Ref. 6, Eqs. (4.5) for the position operator and (4.8) for the spin operator.
[18] See Ref. 1, Eq. (2.9).
[19] See Ref. 3, Eq. (2.18).
[20] See Ref. 6, Eqs. (5.7) for the position and (5.11) for the spin operator.
[21] See Ref. 6, Eqs. (5.13), (5.15), and (5.20) for the explicit forms of the commutation relations.
[22] See Ref. 6, pp. 192–193.
[23] See Ref. 1, Eq. (4.1).
[24] See Ref. 2, Eqs. (11) and (22).
[25] See Ref. 6, Eqs. (2.14) and (6.11).
[26] See Ref. 6, Eqs. (3.15) and (6.12).
[27] See R. L. Mallett and G. N. Fleming, J. Math. Phys. **14**, 45 (1973).
[28] See Ref. 2, Eq. (22).
[29] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950), Eq. (16).
[30] See Ref. 29, Eq. (23) and note 7. Also see Ref. 1, Eq. (6.6). The discrepancy between the two formulas can be eliminated by defining properly the arbitrary scalar A in our expression (30).
[31] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. **34**, 211 (1948).
[32] W. Weidlich and A. K. Mitra, Nuovo Cimento **30**, 385 (1963).
[33] A. Galindo, Nuovo Cimento **37**, 413 (1965).
[34] See Ref. 33, Eqs. (8)–(13).
[35] See Ref. 32, Eqs. (1.11) and (1.12).
[36] See Ref. 2, Eq. (8) and p. 404.

Asymptotic solution of neutron transport problems for small mean free paths*

Edward W. Larsen and Joseph B. Keller

Courant Institute of Mathematical Sciences, New York University, New York, New York 10012
(Received 18 June 1973)

A method is presented for solving initial and boundary value problems for the energy dependent and one speed neutron transport equations. It consists in constructing an asymptotic expansion of the neutron density $\psi(\mathbf{r}, \mathbf{v}, \tau)$ with respect to a small parameter ϵ, which is the ratio of a typical mean free path of a neutron to a typical dimension of the domain under consideration. The density ψ is expressed as the sum of an interior part ψ^i, a boundary layer part ψ^b, and an initial layer part ψ^0. Then ψ^i is sought as a power series in ϵ, while ψ^b decays exponentially with distance from a boundary or interface at a rate proportional to ϵ^{-1}. Similarly ψ^0 decays at a rate proportional to ϵ^{-1} with time after the initial time. For a near critical reactor, the leading term in ψ^i is determined by a diffusion equation. The leading term in ψ^b is determined by a half-space problem with a plane boundary. The initial and boundary conditions for the diffusion equation are obtained by requiring ψ^0 and ψ^b to decay away from the initial instant and from the boundary, respectively. The results are illustrated by specializing them to the one speed case. The method may make it possible to treat more realistic and more complex problems than can be handled by other methods.

1. INTRODUCTION

Neutron transport theory concerns the determination of $\psi(\mathbf{r}, \mathbf{v}, \tau)$, the density of neutrons at $\mathbf{r}$ with velocity $\mathbf{v}$ at time τ. The major difficulty of the theory is that of solving initial and boundary value problems for the transport equation which ψ satisfies. As a consequence, only relatively simple problems have been solved. Therefore we shall present a new method for solving the energy-dependent and one-speed transport equations which may make it possible to treat more complex, and therefore more realistic, problems. The method is that of constructing an asymptotic expansion of ψ with respect to a small parameter ϵ. The parameter ϵ is the ratio of a typical mean free path of a neutron in the domain D under consideration to a typical dimension of D. This method was used by Matkowsky[1] and Habetler and Matkowsky[2] to treat the one-speed case of a slab with isotropic scattering and nonconstant coefficients.

To find ψ we write it as the sum of four parts; an interior part ψ^i, a boundary layer part ψ^b, an initial layer part ψ^0, and an initial-boundary layer part ψ^{0b}:

$$\psi = \psi^i + \psi^b + \psi^0 + \psi^{0b}. \tag{1.1}$$

We shall see that ψ^b and ψ^0 decay exponentially with distance away from a boundary or interface and with time after the initial time, respectively. The decay rate is rapid, being proportional to ϵ^{-1}. Therefore ψ^b becomes negligible beyond a few mean free paths from a boundary or interface and ψ^0 becomes negligible after the time a neutron requires to travel a few mean free paths. On the other hand, ψ^i is appreciable throughout D, when D is a near critical reactor, while ψ^i is appreciable only in the source region, if any, when D is a subcritical reactor. Furthermore in both cases ψ^i has an asymptotic expansion in powers of ϵ, so it is greater than $\psi^b + \psi^0$ outside the initial and boundary layers. The part ψ^{0b} is important initially near the boundary, and decays exponentially both with time and with distance from the boundary.

The determination of the leading term in the expansion of ψ^i in the near critical case leads to a diffusion equation. The initial and boundary conditions associated with this equation are determined by the requirement that ψ^0 must decay with time and ψ^b must decay with distance from the boundary. The derivation of these conditions seems to resolve an outstanding question regarding the appropriate initial and boundary values for the diffusion equation. In both the subcritical and near- critical cases, the leading term in ψ^b is determined as the solution of a steady half-space problem, in which the boundary near $\mathbf{r}$ is replaced by its tangent plane, and in which time occurs only as a parameter. Similarly ψ^0 is the solution of a spatially independent initial value problem in which position occurs as a parameter. The part ψ^{0b} is the solution of an initial-boundary value problem for a half-space with spatially independent boundary data. Additional boundary layers are needed at edges and corners of boundaries, if there are any, and we show how they can be found.

In Sec. 2 we formulate the problem and in Sec. 3 we apply the method to a subcritical region. In Sec. 4 we treat a near critical region. The results, which are obtained for the energy-dependent case, are specialized to the one-speed case in Sec. 5. In the Appendix the scattering operator K is described precisely, its spectrum is determined in some detail and the pseudoinverse of $\lambda I - K$ is analyzed. These results are used in the other sections and may be of general interest.

We wish to thank Basil Nicolaenko for his helpful advice about the derivation of the results in the Appendix.

2. FORMULATION

Under suitable assumptions, the neutron transport equation is

$$[\partial_\tau + \mathbf{v}\cdot\nabla + v\epsilon^{-1}\sigma(\mathbf{r}, v)]\psi(\mathbf{r}, \mathbf{v}, \tau) = \int v'\psi(\mathbf{r}, \mathbf{v}', \tau)\epsilon^{-1}\sigma_1(\mathbf{r}, \boldsymbol{\Omega}'\cdot\boldsymbol{\Omega}, v' \to v)d\mathbf{v}' + Q(\mathbf{r}, \mathbf{v}, \tau). \tag{2.1}$$

Here we have written the total cross section as $\epsilon^{-1}\sigma$ and we have written $\epsilon^{-1}\sigma_1$ for the cross section for emission of a neutron with velocity v in the direction of the unit vector $\boldsymbol{\Omega}$ due to collision of a neutron with velocity v' in the direction $\boldsymbol{\Omega}'$. This is to emphasize that the mean free path is small of order ϵ. The coefficients σ and σ_1 are assumed to be of the same order as a typical length of the domain D. Q is the rate of emission of neutrons by sources. We wish to solve (2.1) for ψ in D with $\psi = g(\mathbf{r}, \mathbf{v})$ given initially throughout D and with the incoming part of ψ given on the boundary of D for all $t \geq 0$.

To do so we introduce $t = \epsilon\tau$ and we rewrite (2.1) in the form

$$[\epsilon\partial_t + \mathbf{v}\cdot\nabla + \epsilon^{-1}v\sigma(\mathbf{r}, v)]\psi(\mathbf{r}, \mathbf{v}, t) = \epsilon^{-1}K(v\sigma\psi) + Q(\mathbf{r}, \mathbf{v}, t). \quad (2.2)$$

Here we have introduced the integral operator K defined by

$$K\Phi(\mathbf{r}, \mathbf{v}) = \int c(\mathbf{r}, v')f(\mathbf{r}, \Omega'\cdot\Omega, v' \to v)\Phi(\mathbf{r}, \mathbf{v}')d\mathbf{v}'. \quad (2.3)$$

In writing (2.2) we have assumed that ψ and Q are functions of t, and have used the same notation for them. In (2.3) we have introduced $c(\mathbf{r}, v')$ and $f = \sigma_1/\sigma(\mathbf{r}, v')c(\mathbf{r}, v')$, where $c(\mathbf{r}, v')$ is the mean number of neutrons emitted per collision of a neutron at $\mathbf{r}$ with speed v'. In terms of σ_1, c is given by

$$\sigma(\mathbf{r}, v')c(\mathbf{r}, v') = \int \sigma_1(\mathbf{r}, \Omega'\cdot\Omega, v' \to v)d\mathbf{v}. \quad (2.4)$$

From (2.4) and the definition of f, it follows that f satisfies the normalization condition

$$\int f(\mathbf{r}, \Omega'\cdot\Omega, v' \to v)d\mathbf{v} = 1. \quad (2.5)$$

3. THE SUBCRITICAL CASE

When D is subcritical, and ψ has the form (1.1), we assume that ψ^i possesses the asymptotic expansion

$$\psi^i(\mathbf{r}, \mathbf{v}, t, \epsilon) \sim \sum_{n=1}^{\infty} \epsilon^n \psi_n(\mathbf{r}, \mathbf{v}, t). \quad (3.1)$$

Substituting (3.1) into (2.2) and equating coefficients of like powers of ϵ yields

$$v\sigma\psi_n - K(v\sigma\psi_n) = \delta_{n1}Q - \partial_t\psi_{n-2} - \mathbf{v}\cdot\nabla\psi_{n-1}, \quad n = 1, 2, \ldots. \quad (3.2)$$

We call D subcritical when the largest eigenvalue of K is less than one. (See the Appendix.) Then the operator on the left side of (3.2) is uniquely invertible. This is the case, for example, when c is small enough. Then since $\psi_{-1} = \psi_0 = 0$ by definition; (3.2) with $n = 1$ can be solved for ψ_1 with the result

$$\psi_1 = (1/v\sigma)(I - K)^{-1}Q. \quad (3.3)$$

The ψ_n with $n > 1$ are then given recursively by

$$\psi_n = (-1/v\sigma)(I - K)^{-1}(\partial_t\psi_{n-2} + \mathbf{v}\cdot\nabla\psi_{n-1}), n > 1. \quad (3.4)$$

This completes the determination of the expansion (3.1) of ψ^i. It is to be noted that this expansion does not involve the initial or boundary conditions, so ψ^i does not satisfy them. We also note that all the ψ_n vanish at points where $Q = 0$, so $\psi^i = 0$ except within the source region $Q \neq 0$.

To find ψ^0, we introduce the stretched variable $t' = \epsilon^{-2}t$ and the function $\psi'(\mathbf{r}, \mathbf{v}, t', \epsilon)$ defined by

$$\psi'(\mathbf{r}, \mathbf{v}, t', \epsilon) = \psi^0(\mathbf{r}, \mathbf{v}, \epsilon^2 t', \epsilon). \quad (3.5)$$

Let the initial value of ψ be $g(\mathbf{r}, \mathbf{v})$. We choose $\psi^0 + \psi^i = g$ at the initial time and then (3.5) yields

$$\psi'(\mathbf{r}, \mathbf{v}, 0, \epsilon) = g(\mathbf{r}, \mathbf{v}) - \psi^i(\mathbf{r}, \mathbf{v}, 0, \epsilon). \quad (3.6)$$

Now we require ψ^0 to satisfy (2.2) with $Q = 0$, because Q is accounted for by ψ^i. Upon using (3.5) in (2.2) with $Q = 0$ and multiplying the resulting equation by ϵ, we obtain

$$[\partial_{t'} + \epsilon\mathbf{v}\cdot\nabla + v\sigma]\psi' = K(v\sigma\psi'). \quad (3.7)$$

We next set $\epsilon = 0$ in (3.7) and (3.6) and obtain the following initial value problem in which no spatial derivatives occur, so that $\mathbf{r}$ enters as a parameter:

$$\partial_{t'}\psi'(\mathbf{r}, \mathbf{v}, t', 0) = (K - I)(v\sigma\psi'(\mathbf{r}, \mathbf{v}, t', 0)), \quad (3.8)$$

$$\psi'(\mathbf{r}, \mathbf{v}, 0, 0) = g(\mathbf{r}, \mathbf{v}). \quad (3.9)$$

In the subcritical case the largest eigenvalue of K is less than one, so the solution ψ' of (3.8) and (3.9) decays exponentially as t' increases. Since $t' = \epsilon^{-2}t = \epsilon^{-1}\tau$, the decay rate in τ is very large, being proportional to ϵ^{-1}. Thus ψ^0 is vanishingly small outside the initial layer of duration $O(\epsilon)$.

In order to determine ψ^b, we write the equation of the boundary of D in the form $\mathbf{r} = \mathbf{r}_0(\xi)$, where ξ is a two-component vector parameter. We choose a parameter value ξ_0 corresponding to a boundary point $\mathbf{r}_0(\xi_0)$, and introduce the stretched variable $\mathbf{r}'$ and the stretched parameter ξ' by

$$\mathbf{r} = \mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}', \xi = \xi_0 + \epsilon\xi'. \quad (3.10)$$

Then we define $\hat{\psi}(\mathbf{r}', \mathbf{v}, t, \epsilon)$ by

$$\hat{\psi}(\mathbf{r}', \mathbf{v}, t, \epsilon) = \psi^b[\mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}', \mathbf{v}, t, \epsilon]. \quad (3.11)$$

Let the prescribed boundary value of ψ be $h(\xi, \mathbf{v}, t)$ for $\mathbf{v}\cdot\mathbf{n}(\xi) < 0$, where $\mathbf{n}(\xi)$ is the outward normal to the boundary. We choose $\psi^b + \psi^i = h$ on the boundary for $\mathbf{v}\cdot\mathbf{n} < 0$. In terms of $\hat{\psi}$ the boundary condition becomes

$$\hat{\psi}[\epsilon^{-1}\{\mathbf{r}_0(\xi) - \mathbf{r}_0(\xi_0)\}, \mathbf{v}, t, \epsilon] = h(\xi, \mathbf{v}, t) - \psi^i[\mathbf{r}_0(\xi), \mathbf{v}, t, \epsilon], \mathbf{v}\cdot\mathbf{n}(\xi) < 0. \quad (3.12)$$

Upon using (3.10) for ξ in (3.12) and then letting ϵ tend to zero, we obtain

$$\hat{\psi}[\xi'\cdot\nabla_\xi \mathbf{r}_0(\xi_0), \mathbf{v}, t, 0] = h(\xi_0, \mathbf{v}, t), \mathbf{v}\cdot\mathbf{n}(\xi_0) < 0. \quad (3.13)$$

This is just a boundary condition on the tangent plane to D at $\mathbf{r}_0(\xi_0)$, and the boundary value h is independent of position ξ' on this plane.

Next we require ψ^b to satisfy (2.2) with $Q = 0$. When we use (3.11) for ψ^b in (2.2) with $Q = 0$, and multiply the resulting equation by ϵ, we obtain

$$[\epsilon^2\partial_t + \mathbf{v}\cdot\nabla' + v\sigma\{\mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}', v\}]\hat{\psi} = K(v\sigma\{\mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}', v\}\hat{\psi}). \quad (3.14)$$

Now we set $\epsilon = 0$ in (3.14) and get

$$[\mathbf{v}\cdot\nabla' + v\sigma\{\mathbf{r}_0(\xi_0), v\}]\hat{\psi}(\mathbf{r}', \mathbf{v}, t, 0) = K^0(v\sigma\{\mathbf{r}_0(\xi_0), v\}\hat{\psi}(\mathbf{r}', \mathbf{v}, t, 0)), \mathbf{r}'\cdot\mathbf{n}(\xi_0) < 0. \quad (3.15)$$

Here K^0 is obtained from K, defined in (2.3), by using (3.10) for $\mathbf{r}$ in c and f and then setting $\epsilon = 0$, which yields

$$K^0\Phi(\mathbf{r}', \mathbf{v}) = \int c[\mathbf{r}_0(\xi_0), v']f[\mathbf{r}_0(\xi_0), \Omega'\cdot\Omega, v' \to v] \times \Phi(\mathbf{r}', \mathbf{v}')d\mathbf{v}'. \quad (3.16)$$

In deriving (3.16) from (2.3) we have used the fact that the value of $\mathbf{r}$ in c and f can be different from that of $\mathbf{r}$ in Φ. The condition $\mathbf{r}'\cdot\mathbf{n}(\xi_0) < 0$ is a consequence of the fact that $\mathbf{r} = \mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}'$ must lie in D.

We now observe that (3.15) with the boundary condition (3.13) is a half-space problem. The equation does

not contain any t derivative, so the equation is that for a steady state. Time enters the solution only parametrically through the boundary condition. The equation has coefficients independent of position, and the boundary value is also independent of position. Thus the solution is independent of the coordinates which vary along the boundary, and it depends only upon the normal distance from the boundary. In the subcritical case we assume that the solution of (3.15) decays exponentially with this normal distance, so that in the $\mathbf{r}$ variable it decays at a rapid rate proportional to ϵ^{-1}. Therefore ψ^b is negligibly small outside a boundary layer of width $O(\epsilon)$. In order to evaluate ψ^b at a point $\mathbf{r}$ in the boundary layer, we choose $\mathbf{r}_0(\xi_0)$ to be the point on the boundary nearest to $\mathbf{r}$.

We have now shown how to calculate all the terms in the asymptotic expansion of ψ^i and the leading terms in the expansions of ψ^0 and ψ^b. Further terms in ψ^0 and ψ^b can be found by substituting expansions for them into the preceding equations and considering the coefficients of higher powers of ϵ, but we shall not consider them. Instead we shall show why the part ψ^{0b} is needed, and how to determine it. To this end we note from (3.1) that $\psi^i = O(\epsilon)$, while from (3.9) the leading term in ψ^0 satisfies the initial condition to $O(\epsilon)$ and from (3.13) the leading term in ψ^b satisfies the boundary condition to $O(\epsilon)$. Therefore the sum $\psi^i + \psi^0 + \psi^b$ satisfies the initial condition to $O(\epsilon)$, except near the boundary where ψ^b, which is $O(1)$, violates it. Similarly the sum satisfies the boundary condition to $O(\epsilon)$ except near $t = 0$ where ψ^0, which is $O(1)$, violates it. It is to compensate for these $O(1)$ violations of the initial and boundary conditions that ψ^{0b} is needed.

To find ψ^{0b} we introduce both $t' = \epsilon^{-2}t$ and $\mathbf{r}'$ defined by (3.10), as well as ξ'. Then we define $\tilde{\psi}$ by

$$\tilde{\psi}(\mathbf{r}', \mathbf{v}, t', \epsilon) = \psi^{0b}[\mathbf{r}_0(\xi_0) + \epsilon\mathbf{r}', \mathbf{v}, \epsilon^2 t', \epsilon]. \tag{3.17}$$

When we substitute $\tilde{\psi}$ for ψ^{0b} into (2.2) with $Q = 0$, we get (3.14) with $\epsilon^2\partial_t$ replaced by $\partial_{t'}$. Upon setting $\epsilon = 0$ in (3.14) with $\epsilon^2\partial_t = \partial_{t'}$, we obtain

$$\begin{aligned}&[\partial_{t'} + \mathbf{v}\cdot\nabla' + v\sigma\{\mathbf{r}_0(\xi_0), v\}]\tilde{\psi}(\mathbf{r}', \mathbf{v}, t', 0)\\ &\quad = K^0(v\sigma\{\mathbf{r}_0(\xi_0), v\}\tilde{\psi}(\mathbf{r}', \mathbf{v}, t', 0)), \mathbf{r}'\cdot\mathbf{n}(\xi_0) < 0.\end{aligned} \tag{3.18}$$

Next we use (1.1) in the initial condition, make use of (3.5) and (3.9), and recall that $\psi^i = O(\epsilon)$ to get

$$\begin{aligned}&\psi^{0b}(\mathbf{r}, \mathbf{v}, 0, \epsilon)\\ &\quad = g(\mathbf{r}, \mathbf{v}) - \psi^i(\mathbf{r}, \mathbf{v}, 0, \epsilon) - \psi^0(\mathbf{r}, \mathbf{v}, 0, \epsilon) - \psi^b(\mathbf{r}, \mathbf{v}, 0, \epsilon)\\ &\quad = -\psi^b(\mathbf{r}, \mathbf{v}, 0, \epsilon) + O(\epsilon).\end{aligned} \tag{3.19}$$

By setting $\epsilon = 0$ in (3.19) we get $\psi^{0b}(\mathbf{r}, \mathbf{v}, 0, 0) = -\psi^b(\mathbf{r}, \mathbf{v}, 0, 0)$. Then (3.11) and (3.17) yield

$$\tilde{\psi}(\mathbf{r}', \mathbf{v}, 0, 0) = -\hat{\psi}(\mathbf{r}', \mathbf{v}, 0, 0). \tag{3.20}$$

Proceeding similarly with the boundary condition, we get

$$\tilde{\psi}[\xi'\cdot\nabla_\xi \mathbf{r}_0(\xi_0), \mathbf{v}, t', 0] = -\psi'[\mathbf{r}_0(\xi_0), \mathbf{v}, t', 0], \quad \mathbf{v}\cdot\mathbf{n}(\xi_0) < 0. \tag{3.21}$$

The initial-boundary value problem (3.18), (3.20), and (3.21) is that for a half-space in which the coefficients and the boundary value are independent of position. Therefore the solution depends only upon the normal distance from the boundary, as well as upon t' and $\mathbf{v}$. This problem determines the leading term in ψ^{0b}. We assume that the solution decays exponentially with t' and with the normal distance. In the original variables τ and $\mathbf{r}$ the decay rate is $O(\epsilon^{-1})$, so the layer is of thickness $O(\epsilon)$.

If the boundary has an edge or vertex, there is an additional boundary layer term associated with it for the same reason as ψ^{0b} was needed. This term can be found in the same way that we found ψ^b above, but with $\mathbf{r}_0(\xi_0)$ a point on the edge or vertex. Then $\nabla_\xi \mathbf{r}_0(\xi_0)$ has more than one limiting value there. As a consequence, the boundary condition (3.13) holds on the surface of a wedge or cone tangent to the surface at the edge or vertex. Similarly (3.15) holds inside this wedge or cone, so the problem for this term is that of a wedge-shaped or cone-shaped region.

If the domain contains an interface, a boundary layer term like ψ^b must be included on each side of it. These terms can also be found in the same way as ψ^b was.

4. THE NEAR CRITICAL CASE

We shall now show how to construct the asymptotic expansions of the four terms in (1.1) when D is a near critical reactor. We assume that the source strength is small and show this by replacing Q by ϵQ in (2.2). We also assume that $c(\mathbf{r}, v)$ depends upon ϵ and has the asymptotic expansion

$$c(\mathbf{r}, v, \epsilon) \sim \sum_{n=0}^{\infty} \epsilon^n c_n(\mathbf{r}, v). \tag{4.1}$$

However, f will be taken to be independent of ϵ. Since c is given by (4.1), it follows that K, defined by (2.3), has the expansion

$$K \sim \sum_{n=0}^{\infty} \epsilon^n K_n, \tag{4.2}$$

where K_n is given by (2.3) with c replaced by c_n. Then K_n depends parametrically upon $\mathbf{r}$ and acts only on the velocity $\mathbf{v}$, so the eigenvalues of K_n will be functions of $\mathbf{r}$.

By a near critical reactor we mean one for which the largest eigenvalue of the operator K_0 is equal to one throughout all of D, or throughout some subdomain D_c of diameter $O(1)$. Then the largest eigenvalue of K is nearly equal to one, in contrast to the subcritical case in which it is definitely less than one. In the Appendix we show that the largest eigenvalue of K_0, $\lambda_0(\mathbf{r})$, is simple, and that the corresponding eigenfunction ϕ_0 is positive and independent of the direction of $\mathbf{v}$. Thus $\phi_0 = \phi_0(\mathbf{r}, v)$ and when $\lambda_0(\mathbf{r}) = 1$ we have

$$K_0\phi_0 = \phi_0. \tag{4.3}$$

When the near criticality condition is satisfied, we seek ψ^i in the form

$$\psi^i(\mathbf{r}, \mathbf{v}, t, \epsilon) \sim \sum_{n=0}^{\infty} \epsilon^n \psi_n(\mathbf{r}, \mathbf{v}, t). \tag{4.4}$$

Then $\psi^i = O(1)$.

The determination of ψ^0 is similar to that in Sec. 3, Eqs. (3.5)–(3.9) with K replaced by K_0 in (3.8) and g replaced by $g(\mathbf{r}, \mathbf{v}) - \psi_0(\mathbf{r}, \mathbf{v}, 0)$ in (3.9).

If $g - \psi_0(\mathbf{r}, \mathbf{v}, 0)$ contains a component proportional to $\phi_0/v\sigma$, (4.3) and (3.8) show that that component remains in ψ' independent of t'. In order that ψ' decay as t' increases, that component must be absent. This will be the case if ψ_0 has the initial value

$$\psi_0(\mathbf{r}, \mathbf{v}, 0) = a(\mathbf{r})[\phi_0(\mathbf{r}, v)/v\sigma(\mathbf{r}, v)]. \tag{4.5}$$

Here $a(\mathbf{r})$ is defined so that the right side of (4.5) is the component of $g(\mathbf{r}, v)$ along $\phi_0/v\sigma$. When (4.5) holds then ψ' decays as t' increases.

The derivation of ψ^b is similar to that in Sec. 3, equations (3.10)–(3.16). However K^0 must be replaced by K_0^0 in (3.15), where the operator K_0^0 is defined by (3.16) with c replaced by c_0. Also the right side of (3.13) must be replaced by $h(\xi_0, \mathbf{v}, t) - \psi_0[\mathbf{r}_0(\xi_0), \mathbf{v}, t]$. If this expression contains a component proportional to $\phi_0/v\sigma$, then that component remains in the solution $\hat{\psi}$ and is independent of distance from the boundary. In order that $\hat{\psi}$ decay away from the boundary, this component must vanish, which is the case if

$$\psi_0[\mathbf{r}_0(\xi_0), \mathbf{v}, t)] = b[\mathbf{r}_0(\xi_0), t]\,\frac{\phi_0[\mathbf{r}_0(\xi_0), v]}{v\sigma[\mathbf{r}_0(\xi_0), v]}. \tag{4.6}$$

Here b is defined so that the right side of (4.6) is the "half-range" component of h along $\phi_0/v\sigma$. We assume that when (4.6) holds, $\hat{\psi}$ decays away from the boundary.

The derivation of ψ^{0b} is the same as that in Sec. 3 with K^0 replaced by K_0^0 in (3.18). We still expect ψ^{0b} to decay with time and distance from the boundary because its initial and boundary data contain no components along $\phi_0/v\sigma$.

Now we can proceed with the determination of the expansion (4.4) for ψ^i. The initial condition (4.5) and the boundary condition (4.6) have been imposed on the leading term ψ_0. The governing equation for ψ^i is (2.2) with Q replaced by ϵQ. Substituting the series (4.2) and (4.4) into this equation and collecting coefficients of like powers of ϵ yields the following system of equations:

$$(I - K_0)(v\sigma\psi_n) = \sum_{m=1}^{n} K_m(v\sigma\psi_{n-m}) + \delta_{n2}Q - \partial_t\psi_{n-2} - \mathbf{v}\cdot\nabla\psi_{n-1}, \quad n \geq 0, \tag{4.7}$$

where $\psi_{-2} = \psi_{-1} = 0$. The $n = 0$ equation is $(I - K_0)(v\,\sigma\,\psi_0) = 0$, which has the general solution

$$\psi_0(\mathbf{r}, \mathbf{v}, t) = A_0(\mathbf{r}, t)[\phi_0(\mathbf{r}, v)/v\,\sigma(\mathbf{r}, v)]. \tag{4.8}$$

Here A_0 is undetermined. Now the $n = 1$ equation becomes

$$(I - K_0)(v\sigma\psi_1) = A_0(K_1\phi_0) - \boldsymbol{\Omega}\cdot\nabla(A_0\phi_0)/\sigma). \tag{4.9}$$

In the Appendix we show that (4.9) has a solution for $A_0 \not\equiv 0$ only if the right side satisfies a solvability condition which is, in this case,

$$\int(K_1\phi)(\mathbf{r}, v)\phi_0^*(\mathbf{r}, v)v^2dv = 0. \tag{4.10}$$

Here ϕ_0^* is the eigenfunction of the adjoint operator K_0^*, corresponding to the eigenvalue one. Equation (4.10) is a condition on c_1 which can be satisfied, for example, by setting $c_1 = 0$. When (4.10) holds, we can write the general solution of (4.9) as:

$$\psi_1 = (1/v\sigma)[A_1\phi_0 + A_0(L_0K_1\phi_0) - \boldsymbol{\Omega}\cdot L_1\nabla(A_0\phi_0/\sigma)]. \tag{4.11}$$

Here $A_1(\mathbf{r}, t)$ is undetermined, while L_0 and L_1 are operators related to the pseudo-inverse of $I - K_0$ and defined in the appendix. They depend parametrically upon $\mathbf{r}$ and act only on v.

We now use (4.8) and (4.11) in (4.7) with $n = 2$ to obtain an equation for ψ_2. The solvability condition for this equation can be written in the form

$$\begin{aligned} 0 = {} & A_0\int[K_1L_0K_1\phi_0 + K_2\phi_0]\phi_0^*v^2dv \\ & + \int[(4\pi)^{-1}\int Q\,d\boldsymbol{\Omega}]\phi_0^*v^2dv \\ & - \frac{\partial A_0}{\partial t}\int\frac{\phi_0\phi_0^*}{\sigma}v\,dv + \tfrac{1}{3}\int\phi_0^*\nabla\cdot\left[\frac{1}{\sigma}L_1\nabla\,\frac{A_0\phi_0}{\sigma}\right]v^2dv. \end{aligned} \tag{4.12}$$

Here we have used the identity

$$\int(\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}_1)(\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}_2)d\boldsymbol{\Omega} = (4\pi/3)\boldsymbol{\Omega}_1\cdot\boldsymbol{\Omega}_2. \tag{4.13}$$

We now differentiate out the last integrand and collect terms in (4.12) to obtain

$$h_1\frac{\partial A_0}{\partial t} = h_2\Delta A_0 + \mathbf{h}_3\cdot\nabla A_0 + h_4A_0 + Q_0. \tag{4.14}$$

The five coefficients in (4.14) are defined by

$$h_1(\mathbf{r}) = \int\frac{\phi_0\phi_0^*}{\sigma}v\,dv, \tag{4.15}$$

$$h_2(\mathbf{r}) = \int\left(L_1\frac{\phi_0}{\sigma}\right)\frac{\phi_0^*}{\sigma}v^2dv, \tag{4.16}$$

$$\mathbf{h}_3(\mathbf{r}) = \tfrac{1}{3}\int\left[\frac{1}{\sigma}L_1\nabla\frac{\phi_0}{\sigma} + \nabla\left(\frac{1}{\sigma}L_1\frac{\phi_0}{\sigma}\right)\right]\phi_0^*v^2dv, \tag{4.17}$$

$$h_4(\mathbf{r}) = \int\left[K_1L_0K_1\phi_0 + K_2\phi_0 + \tfrac{1}{3}\nabla\cdot\left(\frac{1}{\sigma}L_1\nabla\frac{\phi_0}{\sigma}\right)\right]\phi_0^*v^2\,dv, \tag{4.18}$$

$$Q_0(\mathbf{r}, t) = \int\left[\frac{1}{4\pi}\int Q\,d\boldsymbol{\Omega}\right]\phi_0^*v^2\,dv. \tag{4.19}$$

The diffusion equation (4.14) is the governing equation for A_0. The initial and boundary conditions for A_0 are determined by (4.5), (4.6), and (4.8) to be

$$A_0(\mathbf{r}, 0) = a(\mathbf{r}), \mathbf{r}\in\partial D, \tag{4.20}$$

$$A_0[\mathbf{r}_0(\xi_0), t] = b[\mathbf{r}_0(\xi_0), t]. \tag{4.21}$$

The initial-boundary value problem (4.14), (4.20), and (4.21) possesses a unique solution $A_0(\mathbf{r}, t)$. Thus $\psi_0(\mathbf{r}, v, t)$, the leading term in ψ^i, is completely determined by (4.8). We shall not obtain equations for the other ψ_n, which can be found by continuing the procedure used for ψ_0.

We shall now show that ψ_0 is nonnegative, as it should be since it is a density. First we note that the factors ϕ_0, v, and σ in (4.8) are nonnegative. Since ϕ_0^* is also nonnegative, (4.15) shows that $h_1 > 0$. In the Appendix we show that for physically reasonable functions f, the operator L_1 maps positive functions into positive functions. Then (4.16) shows that $h_2 > 0$. Furthermore, because the initial value g and the boundary value h of ψ are nonnegative, their expansion coefficients a and b are nonnegative. Now A_0 satisfies (4.14) in which h_1, h_2, the source term Q_0, and the initial and boundary values of A_0 are all nonnegative. Therefore A_0 is nonnegative, and thus so is ψ_0.

Finally we shall discuss the criticality of the reactor D by considering the homogeneous boundary value problem (4.14) and (4.20) with $Q_0 = a = 0$. It is convenient to normalize ϕ_0 so that $h_1(\mathbf{r}) \equiv 1$. Then the product solutions $e^{\lambda t}\phi(\mathbf{r})$ of the homogeneous problem are determined by

$$h_2\Delta\phi + \mathbf{h}_3\cdot\nabla\phi + h_4\phi = \lambda\phi, \tag{4.22}$$

$$\phi(\mathbf{r}) = 0, \quad \mathbf{r} \in \partial D. \tag{4.23}$$

If all the eigenvalues λ of this problem are negative, then A_0 will be bounded if b and Q_0 are, and A_0 will tend to zero as t increases if b and Q_0 do. This is a subcritical case, just like that considered in Sec. 3. If one eigenvalue is zero and the rest are negative, the reactor is critical, while if any eigenvalue is positive the reactor is supercritical, since it possesses a neutral or a growing mode in these respective cases.

5. ONE-SPEED TRANSPORT THEORY

We shall now specialize the preceding results to the one-speed transport equation, which is obtained from (2.1) by setting

$$f(\mathbf{r}, \Omega' \cdot \Omega, v' \to v) = f^0(\mathbf{r}, \Omega' \cdot \Omega)(\delta(v' - v)/v^2]. \tag{5.1}$$

We assume that f^0 has the expansion

$$f^0(\mathbf{r}, \Omega' \cdot \Omega) = \sum_{l=0}^{\infty} [(2l+1)/4\pi] f_l^0(\mathbf{r}) P_l(\Omega \cdot \Omega'), \tag{5.2}$$

where the P_l are Legendre polynomials. Then the normalization condition (2.5) shows that $f_0^0(\mathbf{r}) = 1$ throughout D. By using (5.1) and (5.2) in (2.3), we can write K in the form

$$\begin{aligned}(K\phi)(\mathbf{r}, \mathbf{v}) &= [c(\mathbf{r}, v)/4\pi] \int f^0(\mathbf{r}, \Omega' \cdot \Omega) \phi(\mathbf{r}, v\Omega') d\Omega' \\ &= c(\mathbf{r}, v) \sum_{l=0}^{\infty} f_l^0(\mathbf{r}) \sum_{m=-l}^{+l} Y_{lm}(\Omega) \int \phi(\mathbf{r}, v\Omega') Y_{lm}^*(\Omega') d\Omega'.\end{aligned} \tag{5.3}$$

Here Y_{lm} are spherical harmonics and we have used the addition theorem. From (5.3) we see that the eigenvalues of K are $\lambda_{lm}(\mathbf{r}, v) = c(\mathbf{r}, v) f_l^0(\mathbf{r})$ and the corresponding eigenfunctions are $\phi_{lm}(\Omega) = Y_{lm}(\Omega)$. These eigenfunctions are complete in the space of functions which are square integrable in Ω.

When $|\lambda_{lm}| = |c(\mathbf{r}, v) f_l^0(\mathbf{r})| < 1$ for all l and all $\mathbf{r}$ in D, D is subcritical and the results in Sec. 3 apply. In view of (5.3), the inverse in (3.3) can be written explicitly. Then since $\psi^i = \epsilon\psi_1 + 0(\epsilon^2)$, we obtain

$$\begin{aligned}\psi^i(\mathbf{r}, \mathbf{v}, \epsilon) = \frac{\epsilon}{v\sigma(\mathbf{r}, v)} \sum_{l=0}^{\infty} \frac{1}{1 - c(\mathbf{r}, v) f_l^0(\mathbf{r})} \\ \times \sum_{m=-l}^{+l} Y_{lm}(\Omega) \int Q(\mathbf{r}, v\Omega', t) Y_{lm}^*(\Omega') d\Omega' + O(\epsilon^2).\end{aligned} \tag{5.4}$$

The first term in ψ^0 is the solution of (3.8) and (3.9), which can also be written explicitly, and then ψ^0 is given by

$$\begin{aligned}\psi^0(\mathbf{r}, \mathbf{v}, \epsilon) = \sum_{l=0}^{\infty} \exp\{v\sigma(\mathbf{r}, v)[c(\mathbf{r}, v) f_l^0(\mathbf{r}) - 1]t/\epsilon^2\} \\ \times \sum_{m=-l}^{+l} Y_{lm}(\Omega) \int g(\mathbf{r}, v\Omega') Y_{lm}^*(\Omega') d\Omega' + O(\epsilon).\end{aligned} \tag{5.5}$$

The first term in ψ^b is the solution of (3.13) and (3.15), which is a "standard" static half-space problem. The boundary data and the coefficients are independent of position, and time enters the problem only as a parameter in the boundary data. Mika[3] has considered this problem when f^0 is a finite sum of P_l. Therefore if $f^0[\mathbf{r}_0(\xi_0), \Omega \cdot \Omega']$ is a finite sum, then ψ^b is given by

$$\begin{aligned}\psi^b(\mathbf{r}, \mathbf{v}, t) \sim \sum_{n=0}^{\infty} \Bigg[\int_{\nu=0}^{1} b_{n\nu}[\mathbf{r}_0(\xi_0), t] \phi_{n\nu}(\Omega) \\ \times \exp[-|\mathbf{r} - \mathbf{r}_0(\xi_0)|/\epsilon\nu] d\nu \\ + \sum_{m=0}^{M(\xi_0)} b_{nm}[\mathbf{r}_0(\xi_0), t] \phi_{nm}(\Omega) \\ \times \exp[-|\mathbf{r} - \mathbf{r}_0(\xi_0)|/\epsilon\nu_{nm}] \Bigg] + O(\epsilon).\end{aligned} \tag{5.6}$$

Here $\mathbf{r}_0(\xi_0)$ is the point on ∂D closest to $\mathbf{r}$, ν_{nm}, ϕ_{nm}, and $\phi_{n\nu}$ are eigenvalues and eigenfunctions of K at $\mathbf{r}_0(\xi_0)$, while b_{nm} and $b_{n\nu}$ are the appropriate half-range expansion coefficients of the boundary datum $h(\xi_0, \mathbf{v}, t)$.

The problem (3.18), (3.20), and (3.21) for the first term in ψ^{0b} is a standard initial-boundary value problem in a half-space, in which the coefficients and boundary values are independent of position. This problem can be solved by standand methods when f^0 is a finite sum of P_l, but we shall not solve it.

We next consider the near critical case in which the largest eigenvalue of K_0 is equal to one throughout D. Since K_0 is given by (5.3) with c replaced by c_0, the eigenvalues and eigenfunctions of K_0 are $\lambda_{lm}(\mathbf{r}, v) = c_0(\mathbf{r}, v) f_l^0(\mathbf{r})$ and $\phi_{lm}(\mathbf{r}, v) = Y_{lm}(\Omega)$. In the Appendix we show that the largest eigenvalue is positive, simple, and the corresponding eigenfunction is positive. Since only $Y_{00}(\Omega)$ is positive, λ_{00} must be the largest eigenvalue. Therefore the near criticality condition is $c_0 f_0^0 = 1$ and since $f_0^0 = 1$, this condition becomes

$$c_0(\mathbf{r}, v) = 1. \tag{5.7}$$

The problems for the leading terms in ψ^0, ψ^b, and ψ^{0b} are the same as in Sec. 4. In the present case (4.5) and (4.6) become

$$\psi_0(\mathbf{r}, v, 0) = (1/4\pi) \int g(\mathbf{r}, v\Omega') d\Omega', \tag{5.8}$$

$$\psi_0[\mathbf{r}_0(\xi_0), v, t] = \tfrac{1}{2} b_{00}[\mathbf{r}_0(\xi_0), t]. \tag{5.9}$$

Then the initial value of ψ' is $g(\mathbf{r}, v\Omega) - (4\pi)^{-1} \int g(\mathbf{r}, v\Omega') d\Omega'$ and the boundary value of ψ is $h(\xi_0, \mathbf{v}, t) - 2^{-1} b_{00}[\mathbf{r}_0(\xi_0), t]$ since $\phi_{00} = \tfrac{1}{2}$.[4]

The result (4.8) still holds for ψ_0, but (4.14) for A_0 is not applicable because in deriving it we used a solvability condition which involved integration over the speed v. Therefore we shall derive a new equation for A_0 in the present case by following the procedure used to get (4.14). Since now the operator K does not act on the speed v, and since ϕ_0 in (4.8) is a constant, we shall rewrite (4.8) in the form

$$\psi_0(\mathbf{r}, \mathbf{v}, t) = A_0(\mathbf{r}, v, t). \tag{5.10}$$

We next consider the system (4.7) in which K_n is given by (5.3) with c replaced by c_n. For $n = 1$, (4.7) is

$$(I - K_0)(v\sigma\psi_1) = c_1 v\sigma A_0 - v\Omega \cdot \nabla A_0. \tag{5.11}$$

The solvability condition for (5.11) is that the integral over Ω of the right side must vanish. If $A_0 \neq 0$, this implies that

$$c_1(\mathbf{r}, v) = 0. \tag{5.12}$$

Then the solution of (5.11) is easily found to be

$$\psi_1 = A_1 - \sigma^{-1}[1 - \tfrac{1}{3} f_1^0(\mathbf{r})]^{-1} \Omega \cdot \nabla A_0. \tag{5.13}$$

Here $A_1(\mathbf{r}, v, t)$ is undetermined. Now for $n = 2$, (4.7) can be written in the form

$$(I-K_0)(v\sigma\psi_2) = c_2 v\sigma A_0 + Q - \partial_t A_0 - v\Omega\cdot\nabla[A_1 - \sigma^{-1}(1-\tfrac{1}{3}f_1^0)^{-1}\Omega\cdot\nabla A_0]. \quad (5.14)$$

The solvability condition for (5.14) leads to

$$\partial_t A_0 = v\nabla\cdot[(\sigma(3-f_1^0))^{-1}\nabla A_0] + c_2 v\sigma A_0 + Q. \quad (5.15)$$

This is the governing equation for A_0. The initial and boundary conditions for $A_0 = \psi_0$ are stated in (5.8) and (5.9). This initial-boundary value problem determines A_0, which is the leading term in ψ^i. The behavior of A_0 as $t \to \infty$ depends upon the spectrum of the differential operator on the right side of (5.15), so criticality is determined as before. From (5.7) and (5.12) it is necessary that $c(\mathbf{r}, v) = 1 + O(\epsilon^2)$.

In the steady state $\partial_t A_0 = 0$, (5.15) is asymptotically equivalent to the well-known diffusion or P_1 equation. (See, for example, Case and Zweifel, Ref. 5, pp. 196-207. Our derivation of this equation takes into account the scattering moments f_n^0 with $n \geq 2$ and shows that these moments are properly absent from the equation. This is in contrast to previous derivations in which it is assumed at the outset that $f_n^0 = 0$ for $n \geq 2$. Furthermore, except for the slab problem considered in Ref. 2, previous derivations do not lead to boundary conditions, whereas our derivation yields the boundary condition systematically.

APPENDIX

In this appendix we shall explain in detail the properties of the scattering operator K. We begin with a description of the scattering kernel f. Then with K defined on a suitable Banach space, we shall analyze the spectrum of K and prove the existence of a "largest" eigenvalue λ_0 with positive eigenfunction ϕ_0. Next we shall assume that the scattering kernel is rotationally invariant. For this case, we shall partially characterize the pseudoinverse of $\lambda_0 I - K$ and prove that the eigenfunction ϕ_0 is a function only of v. We shall also do this for the one-speed case.

The scattering operator K is defined by

$$(K\phi)(\mathbf{v}) = \int c(v')f(\mathbf{v}'\to\mathbf{v})\phi(\mathbf{v}')d\mathbf{v}', \quad (A1)$$

where we have suppressed $\mathbf{r}$ and where

$$1 = \int f(\mathbf{v}'\to\mathbf{v})d\mathbf{v}. \quad (A2)$$

We now write $cf = b_0(\mathbf{v}'\to\mathbf{v}) + b_1(\mathbf{v}'\to\mathbf{v})$, where b_0 accounts for fission and elastic scattering while b_1 accounts for inelastic scattering. We assume that $b_0 = b_1 = 0$ for $v > v_0$, that b_0 is piecewise continuous, and that

$$0 < a \leq b_0(\mathbf{v}'\to\mathbf{v}) \leq b, \quad v < v_0. \quad (A3)$$

In addition, b_1 is the weighted sum of a finite number of delta functions, each one of which corresponds to a discrete energy loss. The weights are nonnegative piecewise continuous.

Let X be the Banach space of functions $\phi(\mathbf{v})$ such that $\phi(\mathbf{v}) = 0$ for $v > v_0$ and $\|\phi\| = \int|\phi(\mathbf{v})|d\mathbf{v} < \infty$. Then from the above description of the kernel K, it follows that $(K\phi)(\mathbf{v})$ is defined for every $\phi \in X$ and

$$\begin{aligned}\int|K\phi(\mathbf{v})|d\mathbf{v} &\leq \int_{\mathbf{v}}\int_{\mathbf{v}'} c(v')f(\mathbf{v}'\to\mathbf{v})|\phi(\mathbf{v}')|d\mathbf{v}'d\mathbf{v}\\ &\leq \int_{\mathbf{v}'} c(v')|\phi(\mathbf{v}')|\int_{\mathbf{v}} f(\mathbf{v}'\to\mathbf{v})d\mathbf{v}\,d\mathbf{v}'\\ &\leq C\|\phi\|. \end{aligned} \quad (A4)$$

Here we have used equation (A2) and we have defined $C = \sup c(v)$. Since $c(v)$ is the mean number of secondary neutrons produced in the collision of a neutron with speed v, it is bounded and so C exists.

Equation (A4) shows that K is a bounded operator mapping X into X. Let us write $K = B_0 + B_1$, where

$$(B_i\phi)(\mathbf{v}) = \int_{\mathbf{v}'} b_i(\mathbf{v}'\to\mathbf{v})\phi(\mathbf{v}')d\mathbf{v}', \quad i = 0, 1. \quad (A5)$$

Then because of the form of the kernels b_i stated above, it follows that B_0 is compact. Also, B_1 is noncompact, bounded, and there exists a least integer p such that $B_1^p = 0$.

We shall now establish certain "positivity" properties of B_i. The definitions, which we state for convenience, can be found in Krasnoselskii.[6] Let $\mathcal{K} \subset X$ be the cone of real, nonnegative functions. The linear operator A is "positive" if it maps $\mathcal{K}$ into itself. Let $u_0 \in \mathcal{K}$ be defined by $u_0(\mathbf{v}) = 1$ for $v \leq v_0$. Then A is "u_0-positive" if for every $\phi \in \mathcal{K}$, there exist positive numbers α, β and a natural number n such that

$$\alpha \leq (A^n\phi)(\mathbf{v}) \leq \beta, \quad v < v_0. \quad (A6)$$

A is "u_0-bounded above" if the second half of this inequality holds. A is "strictly u_0-positive" or "strictly u_0-bounded above" if either of the above respective definitions holds with $n = 1$. Then we have:

Lemma 1: Let A be a linear, compact, u_0-positive operator. Then A possesses an eigenvector ϕ_0 which lies in the cone $\mathcal{K}$. ϕ_0 is the only eigenvector of A in $\mathcal{K}$. The corresponding eigenvalue λ_0 is positive, simple, and greater in modulus than any other eigenvector of A.

Now consider the operator $K^p = (B_0 + B_1)^p$, which is a sum of products of B_0 and B_1 and each product contains B_0, since $B_1^p = 0$. Since B_0 is compact and B_1 is bounded, it follows that K^p is compact. Since, by (A3), B_0 is strictly u_0 positive, then B_0^p is strictly u_0 positive. Also, since every product of B_0 and B_1 containing B_0 is strictly u_0 bounded above, then K^p is strictly u_0 positive. Let X_0 be the subspace of X consisting of all elements which are piecewise continuous. Then B_0 maps X into X_0 and B_1 maps X_0 into X_0. Therefore every product of B_0 and B_1 containing B_0 maps X into X_0, so K^p maps X into X_0.

Using the results that K^p is compact, strictly u_0 positive, and maps X into X_0, we can prove the following theorem.

Theorem 1: Let K and X be as defined above. Then:

(i) The spectrum of K consists entirely of point eigenvalues of finite multiplicity, except possibly for $\lambda = 0$.

(ii) K possesses a positive, simple eigenvalue λ_0 which is greater in magnitude than all other eigenvalues of K. The eigenfunction ϕ_0 corresponding to λ_0 is strictly positive, and all other eigenfunctions of K are either complex or undergo changes in sign.

(iii) $\phi_0(\mathbf{v})$ is piecewise continuous and bounded.

Proof of (i): Since K^p is compact, the spectrum of K^p consists, except for $\lambda = 0$, entirely of point eigenvalues of finite multiplicity. By the spectral mapping theorem, the spectrum of K consists, except for $\lambda = 0$, of isolated points. These points must all be point eigenvalues of finite multiplicity.

Proof of (ii): We have shown that K^p is compact and u_0-positive. By Lemma 1, K^p possess a unique eigenfunction $\phi_0 \in \mathcal{K}$ with positive eigenvalue which we denote by λ_0^p with $\lambda_0 > 0$. This eigenvalue is simple and

greater in modulus than any other eigenvalue of K^p. Therefore, by the spectral mapping theorem, exactly one of the values of $(\lambda_0^p)^{1/p}$ is a simple eigenvalue of K, greater in modulus than any other eigenvalue of K. Let θ_0 be the eigenfunction of K corresponding to $(\lambda_0^p)^{1/p}$. Then $K^p\theta_0 = \lambda_0^p\theta_0$, and since λ_0^p is simple, θ_0 must be a linear multiple of ϕ_0. This shows that $\phi_0 \in \mathcal{K}$ is an eigenfunction of K and, since K is a positive operator, $\lambda_0 > 0$ is the corresponding simple largest eigenvalue of K. Since every eigenfunction of K is an eigenfunction of K^p and since K^p possesses only one eigenfunction in the cone $\mathcal{K}$, then ϕ_0 must be the only eigenfunction in $\mathcal{K}$.

Proof of (iii): Since K^p maps X into X_0 and since ϕ_0 is in the range of K^p, it follows that $\phi_0 \in X_0$. This completes the proof of the theorem. QED

Let us now assume that the scattering kernel f is rotationally invariant and smooth in its angular variable so that f can be written

$$f(\mathbf{v}' \to \mathbf{v}) = f(\boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}, v' \to v) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}) f_n(v' \to v), \tag{A7}$$

where the P_n are Legendre polynomials. Then K is the operator defined in (A1) with f given by (A7):

$$(K\phi)(\mathbf{v}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \int P_n(\boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}) c(v') f_n(v' \to v) \phi(\mathbf{v}') d\mathbf{v}'. \tag{A8}$$

We shall consider the operator $\lambda_0 I - K$. Let K^* be the adjoint operator of K. K^* has the same essential properties as K so that the adjoint eigenfunction ϕ_0^* corresponding to λ_0 is positive, piecewise continuous, and bounded. Then the solvability condition for the problem

$$(\lambda_0 I - K)\phi(\mathbf{v}) = g(\mathbf{v}) \tag{A9}$$

is

$$0 = \int g(\mathbf{v}) \phi_0^*(\mathbf{v}) d\mathbf{v}. \tag{A10}$$

We define the closed subspaces S_0, S_0', and S_1 of X as follows. S_0 is the subspace of all functions ϕ of the form $\phi = \phi(v)$. S_0' is the subspace of all functions ϕ of the form $\phi = \phi(v)$ which satisfy equation (A10). S_1 is the subspace of all functions ϕ of the form $\phi(\mathbf{v}) = \boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v)$. Here $\boldsymbol{\phi}$ is a vector function whose components lie in S_0. We can now prove the following theorem.

Theorem 2: Let K be defined by equation (A8) and let the closed subspaces S_0, S_0', and S_1 be as defined above. Let ϕ_0 be the positive eigenfunction of Theorem 1. Then:

(i) ϕ_0 and ϕ_0^* are elements of S_0.

(ii) There exists an operator L_0 mapping S_0' into S_0' such that for $\phi \in S_0'$,

$$(\lambda_0 I - K) L_0 \phi = \phi. \tag{A11}$$

(iii) There exist scalar operators T_1 and L_1 mapping S_0 into S_0 such that for $\boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v) \in S_1$,

$$K(\boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v)) = \boldsymbol{\Omega} \cdot (T_1 \boldsymbol{\phi})(v) \tag{A12}$$

and

$$(\lambda_0 I - K)(\boldsymbol{\Omega} \cdot L_1 \boldsymbol{\phi}(v)) = \boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v). \tag{A13}$$

Proof: First let $\phi_0(\mathbf{v})$ be the positive eigenfunction of Theorem 1. Since K is rotationally invariant, any rotation of ϕ_0 will also be an eigenfunction of λ_0. Since λ_0 is simple, ϕ_0 must then be independent of $\boldsymbol{\Omega}$, and so $\phi_0 \in S_0$. For the same reasons, we also have $\phi_0^* \in S_0$.

Next we let $\phi(v) \in S_0$. We introduce ϕ into (A8) to obtain

$$(K\phi)(\mathbf{v}) = \int c(v') f_0(v' \to v) \phi(v') (v')^2 dv' \equiv (T_0\phi)(v). \tag{A14}$$

The operator $\lambda_0 I - K$, restricted to S_0', can then be replaced by $\lambda_0 I - T_0$. By (A9) and (A10), this operator is a one-to-one mapping of S_0' into itself. Therefore the inverse $L_0 = (\lambda_0 I - T_0)^{-1}$ exists and satisfies (A11).

Next we let $\phi(\mathbf{v}) = \boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v) \in S_1$. We introduce ϕ into equation (A8) to obtain

$$K(\boldsymbol{\Omega} \cdot \boldsymbol{\phi}(v)) = \boldsymbol{\Omega} \cdot \int c(v') f_1(v' \to v) \boldsymbol{\phi}(v') (v')^2 dv' \equiv \boldsymbol{\Omega} \cdot T_1 \boldsymbol{\phi}(v). \tag{A15}$$

This verifies (A12). Since ϕ_0^* is independent of $\boldsymbol{\Omega}$, every element $\phi \in S_1$ satisfies (A10). Therefore $\lambda_0 I - K$, restricted to S_1, is a one-to-one mapping of S_1 onto itself. The inverse therefore exists and is defined by (A13), where $L_1 = (\lambda_0 I - T_1)^{-1}$. This completes the proof of the theorem. QED

Now we show that if forward scattering dominates backscattering, then L_1 is a positive operator on S_0. The spectrum of T_1, $\sum(T_1)$, is a subset of $\sum(K)$ and $\lambda_0 \notin \sum(T_1)$. Therefore the spectral radius of T_1 is less than λ_0 and we can write L_1 in the form

$$L_1 = \sum_{n=0}^{\infty} \frac{T_1^n}{\lambda_0^{n+1}}. \tag{A16}$$

From (A8), we obtain

$$f_1(v' \to v) = 2\pi \int_{-1}^{1} \mu f(\mu, v' \to v) d\mu. \tag{A17}$$

In this equation, μ is the cosine of the scattering angle. If forward scattering dominates backscattering, then f_1 will be positive. By (A15), T_1 will then be a positive operator and so by (A16), L_1 will also be positive.

Finally, we consider the one-speed scattering operator

$$(K\phi)(v\boldsymbol{\Omega}) = c(v) \int f(\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}', v) \phi(v\boldsymbol{\Omega}') d\boldsymbol{\Omega}', \tag{A18}$$

where $f(\mu, v)$ is a smoothly varying, strictly positive function of μ and v. If we fix v, then K is a compact, u_0-positive operator on the Banach space of functions absolutely integrable in $\boldsymbol{\Omega}$. Lemma 1 applies to K, and therefore K possesses a "largest" eigenvalue which is positive and simple, and the corresponding eigenfunction is positive.

Finally, we remark that although all of the results proved in this appendix apply to K, they are actually needed in Secs. 4 and 5 for K_0 and K_1. It is easily seen that they do apply. The operators L_0 and L_1 used in Sec. 4 are obtained by requiring $\lambda_0 = 1$ and replacing c by c_0 in (A11) and (A13). The operator T_1 used in section 4 is obtained by setting $c = c_1$ in (A15).

* The research for this paper was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-71-2107.

[1] B. J. Matkowsky, "On the Diffusion Approximation to Transport Theory", 1970, unpublished.
[2] G. J. Habetler and B. J. Matkowsky, "On Boundary Value Problems in Diffusion Theory as Limits of Boundary Value Problems in Transport Theory", 1971, unpublished.
[3] J. R. Mika, NSENA, 11, 415 (1961).
[4] *Developments in Transport Theory*, edited by E. Inonu and P. F. Zweifel (Academic, New York, 1968), p. 329.
[5] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).
[6] M. A. Krasnoselskii, *Positive Solutions of Operator Equations* (Noordhoff, Groningen, 1964).

Dimensional regularization for zero-mass particles in quantum field theory

D. M. Capper and G. Leibbrandt*

International Centre for Theoretical Physics, Trieste, Italy
(Received 8 June 1973)

We show by a suitable redefinition of momentum-space integrals that the technique of dimensional regularization can be extended consistently to include nonnormally ordered theories describing zero-mass particles.

I. INTRODUCTION

Several authors have recently suggested that analytic continuation in the number of space dimensions may be a convenient regularizing technique, especially in the case of gauge theories.[1-3] If this technique is applied to massless particles, however, several difficulties are encountered, such as the usual divergence of on-mass-shell S-matrix elements. This difficulty can, in some theories,[4] be resolved by considering the emission of an infinite number of massless particles.

In this paper we shall be concerned with a somewhat different problem, namely, the elimination of infrared divergences arising from massless tadpoles of the form $\int d^{2\omega}q(q^2)^{-1}$, where 2ω denotes the total number of space-time dimensions (see Fig. 1). If the method of dimensional regularization is applied to such tadpoles *directly* (see Sec. II), it is found that they yield *infinity*. If, on the other hand, the tadpole integral is first rewritten in the form

$$\int \frac{d^{2\omega}q}{q^2} = \int \frac{d^{2\omega}q\, q^2}{q^2(q-p)^2} - 2\int \frac{d^{2\omega}q\, q\cdot p}{q^2(q-p)^2} + p^2 \int \frac{d^{2\omega}q}{q^2(q-p)^2}, \qquad p^2 \neq 0, \quad (1.1)$$

and if each of these integrals is then computed one by one, the value of this integral turns out to be zero. Hence there exists an inconsistency in the value of the tadpole integral depending on the way in which the method is applied. The problem is specifically one of analytic continuation.

It is the purpose of this article to demonstrate that these infrared divergences arising from massless tadpoles can be eliminated *consistently* by a suitable redefinition of the 2ω-dimensional integration over momentum space.[5]

The outline of our paper is as follows. We begin Sec. II with a summary of the method of dimensional regularization and then explain why the present technique is incomplete in the case of massless tadpoles. The central feature is the redefinition, in Sec. III, of the 2ω-dimensional Euclidean integral over momentum space, and the subsequent application of this new definition in Sec. IV to a consistent evaluation of tadpoles.

II. PROBLEMS WITH DIMENSIONAL REGULARIZATION

The main features of dimensional regularization are:

(i) The α parameter representation for momentum-space propagators is employed:

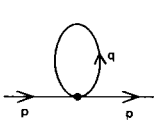

FIG. 1. Lowest-order massless tadpole.

$$\frac{1}{q^2} = \int_0^\infty d\alpha \exp(-\alpha q^2), \qquad q^2 > 0. \quad (2.1)$$

(ii) All inner vector products are defined over a space of 2ω dimensions.

(iii) The momentum space integration is carried out by means of the formula

$$\int d^{2\omega}q \exp(-aq^2 + 2b\cdot q) = (\pi/a)^\omega \exp(b^2/a), \qquad a > 0, \quad (2.2)$$

which defines what is meant by integrating the Gaussian integrand over a 2ω-dimensional Euclidean space.

(iv) The resulting expressions are expanded in a Laurent series about the pole at $\omega = 2$ and, subsequently, continued analytically to four-dimensional space-time.

(v) Pole terms in the Laurent expansion are canceled by means of counterterms[6] in the interaction Lagrangian. The value of the integral is then given by the remaining finite portion of the expansion.

Although the technique of dimensional regularization has been remarkably successful, especially for gauge theories, the above prescription fails in the case of *massless* fields such as gravitation. We shall now clarify where the conventional prescription is defective and why.

The integral I,

$$I = \int \frac{d^{2\omega}q}{q^2} = \int d^{2\omega}q \int_0^\infty \exp(-uq^2)du = \pi^\omega \int_0^\infty du\, u^{-\omega}, \quad (2.3)$$

associated with the tadpole shown in Fig. 1, exhibits an infrared divergence as $\omega \to 2$, the undesirable infinity arising specifically from the lower limit of integration $u = 0$. We also note that, due to this divergence, the interchange of the u- and q-integrations in Eq. (2.3) is strictly not permissible.

It is clearly desirable to find a more meaningful representation of the tadpole integral (2.3) (See Sec. III).

The integral (2.3) may also be written in a slightly different form, i.e.,

$$I = \int \frac{d^{2\omega}q}{q^2}\cdot\frac{(q-p)^2}{(q-p)^2} = \int \frac{d^{2\omega}q\, q^2}{q^2(q-p)^2} - 2\int \frac{d^{2\omega}q\, q\cdot p}{q^2(q-p)^2} + p^2\int \frac{d^{2\omega}q}{q^2(q-p)^2}, \qquad p^2 \neq 0. \quad (2.4)$$

Evaluating the three integrals on the right-hand side of (2.4) separately and applying the technique outlined at the beginning of this section, we obtain

$$I = \{\pi^\omega (p^2)^{\omega-1}\Gamma(1-\omega)[(1-\omega)B(\omega-1,\omega-1) - 2(1-\omega)B(\omega-1,\omega) + (1-\omega)B(\omega-1,\omega+1)]\} + \omega\pi^\omega (p^2)^{\omega-1}\Gamma(1-\omega)B(\omega,\omega), \quad (2.5)$$

where the Γ and β functions are defined by

$$\Gamma(z) = \int_0^\infty dt\, t^{z-1} \exp(-t), \quad \mathrm{Re}z > 0, \tag{2.6}$$

and

$$B(x,y) = \int_0^1 dt\, t^{x-1}(1-t)^{y-1}, \quad \mathrm{Re}x > 0, \quad \mathrm{Re}y > 0, \tag{2.7}$$

respectively. Looked at naively, it would seem from (2.5) that I is zero. To see that this is *not* so, we note that each of the terms in the bracket $\{\cdots\}$ is analytic in the finite strip $\mathfrak{D}_1 : 1 < \mathrm{Re}\omega < 2$, while the last expression, involving $\Gamma(1-\omega)B(\omega,\omega)$, is only analytic in the strip $\mathfrak{D}_2 : 0 < \mathrm{Re}\omega < 1$. Since the domains of analyticity $\mathfrak{D}_1$ and $\mathfrak{D}_2$ *do not overlap* ($\mathfrak{D}_1 \cap \mathfrak{D}_2 = \phi$), there exists no *unique* analytic continuation of the corresponding functions from $\mathfrak{D}_1$ to $\mathfrak{D}_2$ and conversely. This, in turn, rules out any possible cancelation between the terms in $\{\cdots\}$ and $\omega\pi^\omega(p^2)^{\omega-1}\Gamma(1-\omega)B(\omega,\omega)$.

The situation is totally different for massive particles. For instance, one can evaluate the integral

$$\int \frac{d^{2\omega}q}{q^2+m^2} = \pi^\omega (m^2)^{\omega-1}\Gamma(1-\omega), \quad \mathrm{Re}\omega < 1, \quad m^2 \neq 0, \tag{2.8}$$

in either way and still obtain the same result. Unfortunately, the method of inserting a finite mass and letting it approach zero at the end of the calculation is not a useful regulating device. In the first place, the result is ambiguous, depending on the order in which the limits are taken. Secondly, a finite mass will break whatever gauge invariance there might be. Certainly a somewhat different approach is required.

III. REDEFINITION OF GAUSSIAN INTEGRALS IN 2ω-SPACE

A. The function $f(\omega)$

In Sec. II we discussed some of the difficulties which arise in the treatment of massless tadpoles. In this section we shall show how these problems can be overcome by a suitable modification of the gaussian integral (2.2). To this effect let us consider the following

Definition:

$$\int d^{2\omega}q \exp(-aq^2 + 2b\cdot q) \equiv (\pi/a)^\omega \exp[(b^2/a) - af(\omega)], \quad a > 0, \tag{3.1}$$

where the vector b_μ is defined over a 2ω-dimensional space and a behaves like a c-number in that space. The new function $f(\omega)$ satisfies the following five properties:

(i) $f(\omega)$ is a nonzero analytic function of the complex variable $\omega = x + iy$, 2ω being the total number of space-time dimensions;

(ii) $f(\omega) = 0$ for $\omega = \frac{1}{2}n$, $\quad n = 0, 1, 2, \cdots$;

(iii) $\mathrm{Im}f'(\omega) = 0$ for $\omega = \frac{1}{2}n$, $\quad n = 0, 1, 2, \ldots$, where the prime denotes ordinary differentiation with respect to ω;

(iv) $\mathrm{Re}f'(\omega) = 0$ for $\omega = \frac{1}{2}n$, $\quad n = 0, 1, 2, \ldots$;

(v) $[\mathrm{Re}f(\omega)] > 0$ for *any* $\mathrm{Re}\omega$, $\quad \omega \neq \frac{1}{2}n$, and for *some* $\mathrm{Im}\omega$.

Property (i) is required so that we can expand $f(\omega)$ in a Laurent series about $\omega = 2$. Property (ii) is necessary, since (3.1) must agree with (2.2) for 2ω a positive integer. Furthermore, since the original integral (3.1) is *real* (for $\omega = \frac{1}{2}n, n = 0, 1, 2, \ldots$), it is reasonable to require the imaginary part of the *regularized* integral to be zero as well. This leads to property (iii). There does not seem to be any compelling mathematical reason why $\mathrm{Re}f'(\omega)$ should vanish [property (iv)] but, if it can be arranged, it is certainly preferable, since the resulting integrals are then scale invariant. Finally, the application of the definition (3.1), together with property (v), permits a consistent computation of all integrals, *provided* we keep ω complex until all operations involving integrals have been completed. Only then do we invoke the principle of analytic continuation to justify the return to physical four-space, $\omega = 2$.

B. Construction of $f(\omega)$

We now show that there exists at least one function $f(\omega)$ satisfying properties (i)–(v) in Sec. IIIA.

Consider

$$f(\omega) = 1 - \exp[2\pi i g(\omega)], \quad \omega = x + iy, \tag{3.2a}$$

$$g(\omega) = 1 - \cos 2\pi\omega. \tag{3.2b}$$

Property (i) is readily established by recalling the elementary theorem[7] which states that the function of an analytic function is again an analytic function. Since $g(\omega)$ is analytic in the whole ω plane, it follows that $f(\omega)$ is likewise analytic everywhere. Furthermore, since $f(\omega)$ vanishes for $\omega = 0, +\frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$, the second property is also satisfied, while properties (iii) and (iv) follow from the differentiability of $f(\omega)$ [cf. property (i)] which, together with (3.2b), leads to

$$\begin{aligned}\frac{df(\omega)}{d(\omega)} &= 4i\pi^2 \sin\pi\omega \exp[2\pi i g(\omega)] \\ &= 2\pi i g'(\omega) \exp[2\pi i g(\omega)] \\ &= 0 \text{ for } \omega = n/2.\end{aligned} \tag{3.3}$$

It remains to be shown that $\{\mathrm{Re}f(\omega)\} > 0$ for every $\mathrm{Re}\omega \neq n/2, n = 0, 1, 2, \ldots$, and for some $\mathrm{Im}\omega$. Separating $f(\omega)$ into real and imaginary components, with $\omega = x + iy$, we obtain

$$f(\omega) = U + iV, \tag{3.4}$$

where

$$\begin{aligned}U &= 1 - (\cos 2\pi A) \exp(-2\pi B), \\ V &= -(\sin 2\pi A) \exp(-2\pi B),\end{aligned} \tag{3.5}$$

and

$$A = 1 - \cos 2\pi x \cosh 2\pi y, \quad B = \sin 2\pi x \sinh 2\pi y. \tag{3.6}$$

It is now easy to verify that for *any* $x \neq n/2, n = 0, 1, 2, \ldots$, we can find *some* y such that $U > 0$. Of course, for integer or half-integer values of x and $y = 0$, both U and V vanish independently as required by property (ii).

This completes our discussion of the function (3.2). There exist, no doubt, other functions which satisfy properties (i)–(v), but $f(\omega)$ is particularly easy to handle because of its exponential character.

IV. ELIMINATION OF TADPOLES FOR ZERO-MASS PARTICLES

The more general definition (3.1) of the 2ω-dimensional Gaussian integral enables us to:

(i) compute the massless tadpole integral $\int d^{2\omega}q(q^2)^{-1}$ unambiguously;

(ii) establish consistency of (i) above with the evaluation of the integral

$$\int \frac{d^{2\omega}q}{q^2} \cdot \frac{(q-p)^2}{(q-p)^2};$$

(iii) derive in a more concise manner a previous result of the authors concerning $\delta^4(0)$ terms.[8]

A. The tadpole integral

Let us first evaluate, with the aid of Eqs. (2.1) and (3.1), the integral

$$I = \int d^{2\omega}q (q^2)^{-1} = \int d^{2\omega}q \int_0^\infty du\ \exp(-uq^2) \tag{4.1}$$

$$= \pi^\omega \int_0^\infty du\ u^{-\omega} \exp[-uf(\omega)]. \tag{4.2}$$

Since we may choose, by property (v), the real part of $f(\omega)$ to be positive definite for all $\omega \neq n/2, n = 0, 1, 2, \ldots$, we can complete the integration in (4.2) by means of (2.6), obtaining

$$I = \pi^\omega [f(\omega)]^{\omega-1} \Gamma(1-\omega), \tag{4.3}$$

where $\Gamma(1-\omega)$ is analytic in the semi-infinite strip $\mathrm{Re}\omega < 1$. Before $\Gamma(1-\omega)$ can be expanded about $\omega = 2$, it is necessary to continue this function analytically to other values of ω by means of the partial fraction exapnsion[9]

$$\Gamma(1-\omega) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1-\omega)} + \int_1^\infty dt\ t^{-\omega} \exp(-t). \tag{4.4}$$

Substituting

$$\Gamma(1-\omega) = -\left\{ \frac{1}{2-\omega} + \psi(2) + \tfrac{1}{2}(2-\omega)\left[\frac{\pi^2}{3} + \psi^2(2) - \psi'(2)\right] + O((2-\omega)^2)\right\}, \tag{4.5}$$

$$\psi(\omega) = \frac{d}{d\omega} \ln\Gamma(\omega), \tag{4.6}$$

into the right-hand side of Eq. (4.3), and expanding about $\omega = 2$, we find that

$$I = -\pi^\omega [\tfrac{1}{2}(2-\omega) 8i\pi^3 + \tfrac{1}{2}(2-\omega)^2 8i\pi^3 \psi(2) + O((2-\omega)^3)]. \tag{4.7}$$

As $\omega \to 2$, the tadpole integral (4.7) reduces, therefore, formally to zero:

$$\int \frac{d^4q}{q^2} = 0. \tag{4.8}$$

It should be noted that we have not proved that tadpoles are zero (the original integrals are manifestly divergent), but only that within the context of dimensional regularization such diagrams may consistently be put equal to zero.

B. Consistency of Eq. (4.8)

We shall next demonstrate the *consistency* of the result (4.8) with the value obtained from the integral

$$I = \int \frac{d^{2\omega}q}{q^2} = \int \frac{d^{2\omega}q}{q^2} \cdot \frac{(q-p)^2}{(q-p)^2}, \quad p^2 \neq 0, \tag{4.9}$$

$$= p^2 \int \frac{d^{2\omega}q}{q^2(q-p)^2} - 2\int \frac{d^{2\omega}q\ q\cdot p}{q^2(q-p)^2} + \int \frac{d^{2\omega}q\ q^2}{q^2(q-p)^2}. \tag{4.10}$$

We recall from Sec. II that multiplication of the original integrand by $(q-p)^2/(q-p)^2$ is inconsistent for massless particles *if the conventional definition (2.2) is employed.* Using definition (3.1), however, one obtains for the first integral in (4.10),

$$J_1 \equiv \int \frac{d^{2\omega}q}{q^2(q-p)^2}$$

$$= \int_0^\infty \int_0^\infty du\ dv \left[\int d^{2\omega}q\ \exp[-uq^2 - v(q-p)^2]\right] \tag{4.11}$$

$$= \int_0^\infty \int_0^\infty du\ dv \{\pi^\omega (u+v)^{-\omega} \exp[p^2 v^2 (u+v)^{-1} - (u+v) f(\omega)]\} \exp(-vp^2) \tag{4.12}$$

$$= \pi^\omega \int_0^1 d\xi \int_0^\infty d\lambda\ \lambda^{1-\omega} \exp(-\lambda R), \tag{4.13a}$$

where

$$R = p^2 \xi(1-\xi) + f(\omega), \tag{4.13b}$$

so that

$$J_1 = \pi^\omega \Gamma(2-\omega) \int_0^1 d\xi\ R^{\omega-2}, \quad \mathrm{Re}\omega < 2. \tag{4.14}$$

The domain of validity of J_1 is now a *semi-infinite* strip, in contrast to the *finite* strip appearing in Eq. (2.5). To see this we observe that the integral

$$\int_0^1 d\xi [p^2 \xi(1-\xi) + f(\omega)]^{\omega-2}$$

is *not* a β integral in general, since $\{\mathrm{Re} f(\omega)\} \neq 0$ for arbitrary ω.

The other two integrals in (4.10) yield, in similar fashion,

$$J_2 \equiv -2p_\mu \int \frac{d^{2\omega}q\ q_\mu}{q^2(q-p)^2} = -2p^2 \pi^\omega \Gamma(2-\omega) \int_0^1 d\xi (1-\xi) R^{\omega-2}, \tag{4.15}$$

$$J_3 \equiv \int \frac{d^{2\omega}q\ q^2}{q^2(q-p)^2} = \omega\pi^\omega \Gamma(2-\omega) \int_0^1 d\xi\ R^{\omega-1} + p^2 \pi^\omega \Gamma(2-\omega) \int_0^1 d\xi (1-\xi)^2 R^{\omega-2} + \pi^\omega f(\omega) \Gamma(2-\omega) \int_0^1 d\xi R^{\omega-2}, \tag{4.16}$$

where the two Γ functions in Eq. (4.16) impose different restrictions on ω, namely, $\mathrm{Re}\omega < 1$ and $\mathrm{Re}\omega < 2$, respectively. Substituting the right-hand side of Eqs. (4.14)–(4.16) into Eq. (4.10) and simplifying the resulting expressions by means of the identities

$$\int_0^1 d\xi (1-\xi) R^{\omega-2} = \int_0^1 d\xi\ \xi R^{\omega-2} = \tfrac{1}{2} \int_0^1 d\xi\ R^{\omega-2}, \quad \text{for any } \omega, \tag{4.17}$$

$$\int_0^1 d\xi\ R^{\omega-1} = p^2 \int_0^1 d\xi\ \xi R^{\omega-2} - p^2 \int_0^1 d\xi\ \xi^2 R^{\omega-2} + f(\omega) \int_0^1 d\xi\ R^{\omega-2}, \tag{4.18}$$

and with the help of the reduction formula

$$\int_0^1 d\xi\ \xi^2 R^{\omega-2} = [2p^2(1-2\omega)]^{-1} \{2(f(\omega))^{\omega-1} - [\omega p^2 + 2(f(\omega))^2] \int_0^1 d\xi\ R^{\omega-2}\}, \tag{4.19}$$

we find that

$$I = \int \frac{d^{2\omega}q}{q^2} \cdot \frac{(q-p)^2}{(q-p)^2} = \pi^\omega (f(\omega))^{\omega-1} \Gamma(1-\omega). \tag{4.20}$$

Since the right-hand side of (4.20) is identical with (4.2), it follows immediately that

$$\int \frac{d^4q(q-p)^2}{q^2(q-p)^2} = 0, \tag{4.21}$$

which is consistent with the previous result (4.8). In a similar way we can show that $\delta^4(0)$ terms also vanish formally as $\omega \to 2$, in agreement with Ref. 8.

V. CONCLUSION

We have shown that a suitable definition of the 2ω-dimensional Gaussian integral in momentum space enables us to put massless tadpoles consistently equal to zero. The definition we have chosen is, of course, not unique, but this merely corresponds to the usual ambiguity in the choice of subtraction point. However, our choice of definition (3.1) has the advantage of simplicity (e.g., it does not give rise to branch cuts in the ω-plane) and it gives, moreover, consistent results when applied in different ways. The technique described here has been successfully applied to one-loop calculations in quantum gravity, where results consistent with the Slavnov-Ward identities have been obtained.[10]

For higher-order tadpoles the situation is considerably more complicated, although it is *not* necessary to introduce additional regularizing parameters ω_i, $i = 1, 2, \ldots$, as has been described in Ref. 2.

ACKNOWLEDGMENTS

The authors are grateful to Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste. We should also like to thank Dr. John Strathdee for useful discussions. One of us (GL) acknowledges a grant from the National Research Council of Canada.

* On leave of absence from Dept. of Mathematics and Statistics, University of Guelph, Ontario, Canada.

[1] G. 't Hooft and M. Veltman, Nucl. Phys. (B) **44**, 189 (1972).
[2] J. F. Ashmore, Lett. Nuovo Cimento **4**, 289 (1972) and Commun. Math. Phys. **29**, 177 (1973).
[3] C. G. Bollini and J. J. Giambiagi, Nuovo Cimento **12B**, 20 (1972).
[4] S. Weinberg, Phys. Rev. **140**, 516 (1965).
[5] We note that the momentum variable q is defined over a Euclidean space, whereas the number of space dimensions 2ω is, in general, *complex*.
[6] We must be careful that the counterterms are real. The limit $\omega \to 2$ will henceforth be understood to mean: "ω approaches the point 2 from the left along the real axis."
[7] W. Kaplan, *Functions of a Complex Variable* (Addison-Wesley, Cambridge, Mass., 1953).
[8] D. M. Capper and G. Leibbrandt, Lett. Nuovo Cimento **6**, 117 (1973).
[9] A. I. Markushevich, *Theory of Functions of a Complex Variable* (Prentice-Hall, Englewood Cliffs, N.J., 1965), Vol. II, p. 311.
[10] D. M. Capper, G. Leibbrandt, and M. Ramon Medrano, "Calculation of the Graviton Self-Energy Using Dimensional Regularization," Phys. Rev. D. (to appear).

On a conjecture by 't Hooft and Veltman

D. M. Capper

International Centre for Theoretical Physics, Trieste, Italy

G. Leibbrandt

Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada
(Received 8 June 1973)

The conjecture by 't Hooft and Veltman that $\int d^{2\omega}q(q^2)^{\lambda-1} = 0$ (ω, λ complex), within the context of dimensional regularization, is proven in the case of zero-mass particles for *any* ω when $\lambda = 1,2,...$ and for $\lambda = 0$ in the limit $\omega \to 2$.

1. INTRODUCTION

Sparked by the fundamental work of 't Hooft,[1] the renormalization of gauge theories is now one of the most exciting and promising areas in quantum field theory. An important mathematical aspect of the new renormalization program is the method of *dimensional regularization*[2-4] which allows—at least in principle—for a consistent, gauge-invariant computation of Feynman integrals.

The central feature of dimensional regularization is the concept of *analytic continuation* in the number of dimensions. The trick is first to represent all integrals by analytic expressions in a space where these integrals are well-defined and then to continue these expressions analytically to Minkowski four-space. The technique of dimensional regularization has the additional advantage of treating highly divergent Feynman integrals of the type $\int d^4q(q^2)^n, n = 0, 1, 2, \ldots$, in a consistent manner. This was first pointed out in the case of *massive* particles by 't Hooft and Veltman.[5] They observed that within the framework of dimensional regularization no inconsistencies arise, for example, in the various Slavnov-Ward identities,[6] if one assumes that

$$\int d^{2\omega}q(q^2)^{\lambda-1} = 0, \qquad \omega, \lambda \text{ complex}. \tag{1.1}$$

Here 2ω is the total number of dimensions so that $\omega = 2$ corresponds to physical four-space.

The purpose of this note is to prove the 't Hooft–Veltman conjecture (1.1) for $\lambda = 0, 1, 2, \ldots$ in the case of *massless* particles. (For massive particles the proof is straightforward and could be conducted, for example, along the lines suggested in Ref. 7.) We shall specifically show that for integrals over a polynomial, $\lambda = 1, 2, \ldots$, the conjecture (1.1) holds for *any* value of the regulating parameter ω, whereas for $\lambda = 0$ (tadpole integral), it only holds in the limit $\omega \to 2$.

Integrals of the form

$$\int d^{2\omega}q(q^2)^{\lambda-1} = 0, \qquad \lambda = 0, 1, 2, \ldots \tag{1.2}$$

appear naturally in quantum gravity, where they arise in one-loop and two-loop graviton–graviton calculations[6] as well as in the treatment of lower-order and higher-order tadpoles.

2. THE CONTINUITY FUNCTION $f(\omega)$

Following the technique of Ashmore,[4] the first step in the method of dimensional regularization is to define all momentum space integrals over a complex 2ω-dimensional Euclidean space and to parametrize the various propagators in that space according to

$$\frac{1}{q^2} = \int_0^\infty dx \exp(-xq^2), \qquad q^2 > 0. \tag{2.1}$$

The second step consists of integrating over momentum space by means of the generalized Gaussian formula

$$\int d^{2\omega}q \exp[-xq^2 + 2bq] = (\pi/x)^\omega \exp(b^2/x), \qquad x > 0. \tag{2.2}$$

The prescription (2.1), (2.2) works satisfactorily in the case of massive fields, but is *insufficient* for theories involving *massless* particles such as quantum gravity. The difficulty is precisely one of analytic continuation. For gravity it is necessary to replace formula (2.2) by the extended definition[8]

$$\int d^{2\omega}q \exp[-xq^2 + 2bq] = (\pi/x)^\omega \exp[(b^2/x) - xf(\omega)], \qquad x > 0, \tag{2.3}$$

where the *continuity function* $f(\omega)$ satisfies the following four properties:

(i) $f(\omega)$ is a nonzero analytic function of the complex variable $\omega = \sigma + i\tau$;
(ii) $f(\omega) = 0$ for $\omega = \frac{1}{2}n, \quad n = 0, 1, 2, \ldots$;
(iii) $f'(\omega) = 0$ for $\omega = \frac{1}{2}n, \quad n = 0, 1, 2, \ldots$;
(iv) $[\mathrm{Re}f(\omega)] > 0$ for *any* $\mathrm{Re}\omega$, $\omega \neq \frac{1}{2}n$, and *some* $\mathrm{Im}\omega$.

The new definition (2.3) allows us to compute zero mass integrals in a consistent and unambiguous manner, *provided* the regulating parameter ω remains complex until all formal manipulations involving integrals have been executed.

3. PROOF OF CONJECTURE

Let us apply Eqs. (2.1) and (2.3) to the integral $\int d^{2\omega}q(q^2)^{-1}$; then

$$\int d^{2\omega}q(q^2)^{-1} = \pi^\omega \int_0^\infty dx\, x^{-\omega} \exp[-xf(\omega)]. \tag{3.1}$$

Differentiating both sides of the equation

$$\int d^{2\omega}q \exp(-xq^2) = \pi^\omega x^{-\omega} \exp[-xf(\omega)] \tag{3.2}$$

λ times with respect to x, $\lambda = 0, 1, 2, \ldots$, we obtain

$$\int d^{2\omega}q(q^2)^\lambda \exp(-xq^2) = \pi^\omega \sum_{j=0}^{\lambda} \frac{\Gamma(\lambda+1)\Gamma(\omega+j)}{\Gamma(j)\Gamma(\lambda-j+1)\Gamma(\omega)} \times f^{\lambda-j}x^{-\omega-j}\exp(-xf), \qquad \lambda = 0, 1, 2, \ldots, \tag{3.3}$$

so that

$$I \equiv \int d^{2\omega}q(q^2)^{\lambda-1} = \pi^\omega \sum_{j=0}^{\lambda} \frac{\Gamma(\lambda+1)\Gamma(\omega+j)f^{\lambda-j}}{\Gamma(j)\Gamma(\lambda-j+1)\Gamma(\omega)} \times \int_0^\infty dx\, x^{-\omega-j}\exp(-xf), \qquad \lambda = 0, 1, 2, \ldots, \tag{3.4}$$

or in a more suggestive form,

$$I = \pi^\omega [f(\omega)]^{\omega+\lambda-1} \sum_{j=0}^{\lambda} \frac{\Gamma(\lambda+1)\Gamma(\omega+j)}{\Gamma(j)\Gamma(\lambda-j+1)\Gamma(\omega)} \times \int_0^\infty d\alpha\, \alpha^{-\omega-j}\exp(-\alpha), \qquad \lambda = 0, 1, 2, \ldots. \tag{3.5}$$

The last integral is well-defined for all $\mathrm{Im}\omega \neq 0$. We note that had we employed definition (2.2) instead of (2.3) we would have obtained $\int_0^\infty dx\, x^{-\omega-\lambda}$ rather than the integral in (3.4). The former integral contains a divergence as $\omega \to 2$ which makes further operations of this and similar integrals meaningless. Integrating now (3.5) we readily obtain

$$I = \pi^\omega [f(\omega)]^{\omega+\lambda-1} \sum_{j=0}^{\lambda} \frac{\Gamma(\lambda+1)\Gamma(\omega+j)\Gamma(1-\omega-j)}{\Gamma(j+1)\Gamma(\lambda-j+1)\Gamma(\omega)}, \tag{3.6}$$

where we have used Euler's representation of the Γ function:

$$\Gamma(z) = \int_0^\infty dt\, t^{z-1} \exp(-t), \qquad \mathrm{Re}z > 0. \tag{3.7}$$

The right-hand side of Eq.(3.6) may be simplified by observing that

$$\frac{\Gamma(\omega+j)\Gamma(1-\omega-j)}{\Gamma(\omega)} = (-1)^j \Gamma(1-\omega), \tag{3.8}$$

so that Eq.(3.6) reduces to

$$I = \pi^\omega [f(\omega)]^{\omega+\lambda-1} \Gamma(1-\omega) \sum_{j=0}^{\lambda} \frac{(-1)^j \Gamma(\lambda+1)}{\Gamma(j+1)\Gamma(\lambda-j+1)}. \tag{3.9}$$

Since

$$\sum_{j=0}^{\lambda} \frac{(-1)^j \Gamma(\lambda+1)}{\Gamma(j+1)\Gamma(\lambda-j+1)} = \begin{cases} 1 \text{ for } \lambda = 0 & (3.10) \\ 0 \text{ for } \lambda \geqq 1, & (3.11) \end{cases}$$

we see immediately that the left-hand side of Eq.(1.1) becomes:

$$\int d^{2\omega} q (q^2)^{\lambda-1} = \begin{cases} \pi^\omega [f(\omega)]^{\omega+\lambda-1} \Gamma(1-\omega) & \text{for } \lambda = 0 \quad (3.12) \\ 0 & \text{for } \lambda \geqq 1. \quad (3.13) \end{cases}$$

Equation (3.13) shows that the integral over a polynomial vanishes for *any* value of the regulating parameter ω, in particular for $\omega = 2$. For $\lambda = 0$, we recover from Eq.(3.12) the tadpole integral

$$\int d^{2\omega} q (q^2)^{-1} = \pi^\omega [f(\omega)]^{\omega-1} \Gamma(1-\omega), \tag{3.14}$$

which vanishes[8] in the limit $\omega \to 2$ when the right-hand side of Eq.(3.14) is analytically continued to Minkowski four-space:

$$\lim_{\omega \to 2} \int d^{2\omega} q (q^2)^{-1} = 0. \tag{3.15}$$

The limit in Eq.(3.15) is unique. This follows directly from the analyticity of the continuity function $f(\omega)$: Since the limit of $f(\omega)$ exists at $\omega = 2$, that limit has, by definition, a unique value. This completes our proof of the 't Hooft–Veltman conjecture (1.2).

4. DISCUSSION

We have proved in the case of massless particles that integrals over a polynomial, $\lambda = 1, 2, 3, \ldots$ are strictly zero in the context of dimensional regularization. The proof is made possible by employing the extended definition (2.3) of a 2ω-dimensional Gaussian integral. This definition is characterized by the introduction of a nonzero analytic function $f(\omega)$ specifically designed to cope with *massless* particles.

Although integrals of type (1.2) are in principle highly divergent (in Minkowski space), dimensional regularization permits us to equate them formally to zero. If and when this is done, it is found that *no inconsistencies* emerge, for example, in the Slavnov–Ward identities derived for the graviton and fictitious particle contributions to the graviton propagator.[9] The same conclusion holds for the three lowest-order tadpole contributions to the graviton self-energy.[6]

We finally remark that the conjecture (1.2) is extremely useful in the treatment of multiple-loop massless integrals of the form

$$\int \frac{d^4p d^4q}{p^2(q-p)^2q^2} \quad \text{or} \quad \int \frac{d^4p d^4q}{q^2(q-p)^2p^2(k-p)^2}.$$

Such integrals will be discussed in detail elsewhere.

ACKNOWLEDGMENTS

We are grateful to Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste, where most of the work was carried out. One of us (GL) would like to acknowledge an operating grant in aid of research provided by the National Research Council of Canada. We should also like to thank Professor M. Veltman and Dr. G. 't Hooft for acquainting us in a letter with the above problem and for several illuminating discussions.

[1] G. 't Hooft, Nucl. Phys. B 33 173 (1971), 167 (1971).
[2] G. 't Hooft and M. Veltman, Nucl. Phys. B 44, 189 (1972).
[3] C. G. Bollini and J. J. Giambiagi, Nuovo Cimento 12 B, 20 (1972).
[4] J. F. Ashmore, Lett. Nuovo Cimento 4, 289 (1972); Commun. Math. Phys. 29, 177 (1973).
[5] G. 't Hooft and M. Veltman, private communication. See also G. 't Hooft and M. Veltman, CERN preprint TH-1571 (1972).
[6] D. M. Capper, G. Leibbrandt, and M. Ramon Medrano, Phys. Rev. (to appear).
[7] D. M. Capper and G. Leibbrandt, Lett. Nuovo Cimento 6, 117 (1973).
[8] D. M. Capper and G. Leibbrandt, J. Math. Phys. 15, 82 (1974).
[9] See Ref. 6, Eqs. (5.10)–(5.12).

Nonrelativistic current algebra in the N/V limit*

G. A. Goldin

Department of Physics and Graduate School of Education, University of Pennsylvania, Philadelphia, Pennsylvania 19174

J. Grodnik

Burroughs Corp., 200 W. Lancaster Ave., Wayne, Pennsylvania 19087

R. T. Powers† and D. H. Sharp

Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19174
(Received 26 July 1973)

The case of a noninteracting infinite Bose gas at zero temperature is studied in the formalism of local current algebras, using the representation theory of nuclear Lie groups. The class of representations describing such a system is obtained by taking an "N/V limit" of the finite case. These representations can also be determined uniquely from the solutions of a functional differential equation, which follows in turn from a condition on the ground state vector. Finally a system of functional differential equations is formulated for a theory with interactions, using a proposed definition of indefinite functional integration.

1. INTRODUCTION

There are two main reasons that nonrelativistic models based on algebras of local currents have recently drawn the attention of theorists.[1-3] First, they provide an interesting reformulation of ordinary quantum mechanics in terms of observables such as the particle number density $\rho(\mathbf{x})$ and the particle flux density $\mathbf{J}(\mathbf{x})$, rather than the second-quantized field operator $\psi(\mathbf{x})$. In this paper we employ such a reformulation to study the properties of an infinite Bose system. For the case of noninteracting bosons at zero temperature, the local current algebra approach leads to an elegant restatement of known results.[4] When interactions are included, we develop a system of coupled functional differential equations whose solution would describe the properties of an interacting Bose gas. While these equations are not expected to yield explicit solutions to most interacting theories of interest, it is our hope that they will prove susceptible to some method of approximation.

The second reason that such nonrelativistic models are studied is that they may eventually shed light on local relativistic current algebras. As emphasized by Haag and by Wightman, there are many similarities between relativistic quantum field theory and the quantum mechanics of nonrelativistic systems having infinitely many degrees of freedom.[5-6] In particular, the vacuum state in quantum field theory is the analog of the nonrelativistic ground state. It is to be hoped then that the techniques of nonrelativistic current algebra can be carried over and incorporated into the study of relativistic models.[1,7-10]

This paper is concerned with infinite Bose systems in the "N/V limit" or thermodynamic limit, in which the total number of particles N and the volume V of the system become infinite while the average density $\bar{\rho} = N/V$ approaches a finite constant.

In Sec. 2 we review the case of a noninteracting infinite Bose gas at zero temperature, from the standpoint of group representation theory. The group is that obtained by exponentiating the local current commutators. Consequently, the focus of attention is on the properties of the ground state expectation functional

$$L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0). \tag{1.1}$$

In Sec. 3 we show how a condition on the ground state vector,

$$(\nabla\rho + 2i\mathbf{J})(\mathbf{x})\,\Omega_0 = 0, \tag{1.2}$$

uniquely determines the class of representations obtained by other means in Sec. 2. Section 4 reviews the concept of functional differentiation and proposes a specific definition for a companion concept, the indefinite functional integral. The results of Sec. 3 are recast as the derivation and solution of a functional differential equation.

Finally, Sec. 5 formulates a system of such functional differential equations for a theory with interactions.

2. THE FREE BOSE GAS AT ZERO TEMPERATURE

A. Preliminaries[1-3,11]

A second-quantized nonrelativistic Bose field $\psi(\mathbf{x})$ satisfies the canonical commutation relations

$$\begin{aligned}&[\psi(\mathbf{x}),\psi^*(\mathbf{y})] = \delta(\mathbf{x}-\mathbf{y}),\\ &[\psi(\mathbf{x}),\psi(\mathbf{y})] = [\psi^*(\mathbf{x}),\psi^*(\mathbf{y})] = 0.\end{aligned} \tag{2.1}$$

The Fock representation for such a field is defined as follows. Let $\mathcal{K}_n$ be the Hilbert space of complex square integrable functions of n vector variables which are symmetric under the exchange of particle coordinates, and let $\mathcal{K} = \oplus_{n=0}^{\infty}\mathcal{K}_n$ be the direct sum of the $\mathcal{K}_n$. A vector $\Psi \in \mathcal{K}$ has components $\Psi_n \in \mathcal{K}_n$ with $(\Psi,\Psi) = \sum_n(\Psi_n,\Psi_n) < \infty$. The action of the fields $\psi(\mathbf{x})$ and $\psi^*(\mathbf{x})$ in $\mathcal{K}$ is defined by

$$[\psi(\mathbf{x})\Psi]_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) = (n+1)^{1/2}\Psi_{n+1}(\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{x}) \tag{2.2}$$

and

$$[\psi^*(\mathbf{x})\Psi]_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) = n^{-1/2}\sum_{j=1}^{n}\delta(\mathbf{x}-\mathbf{x}_j)\Psi_{n-1}(\mathbf{x}_1,\ldots\hat{\mathbf{x}}_j,\ldots,\mathbf{x}_n). \tag{2.3}$$

Defining the number density of particles as

$$\rho(\mathbf{x}) = \psi^*(\mathbf{x})\psi(\mathbf{x}) \tag{2.4}$$

and the particle flux density (for particles of unit mass) as

$$\mathbf{J}(\mathbf{x}) = (1/2i)[\psi^*(\mathbf{x})\nabla\psi(\mathbf{x}) - (\nabla\psi^*(\mathbf{x}))\psi(\mathbf{x})], \tag{2.5}$$

one obtains the equal time current algebra

$$[\rho(\mathbf{x}), \rho(\mathbf{y})] = 0, \tag{2.6}$$

$$[\rho(\mathbf{x}), J_k(\mathbf{y})] = -i(\partial/\partial x^k)[\delta(\mathbf{x}-\mathbf{y})\rho(\mathbf{x})], \tag{2.7}$$

$$[J_j(\mathbf{x}), J_k(\mathbf{y})] = -i\frac{\partial}{\partial x^k}[\delta(\mathbf{x}-\mathbf{y})J_j(\mathbf{x})] + i\frac{\partial}{\partial y^j}[\delta(\mathbf{x}-\mathbf{y})J_k(\mathbf{y})]. \tag{2.8}$$

Introducing the smeared currents

$$\rho(f) = \int \rho(\mathbf{x})f(\mathbf{x})\,d^3x \tag{2.9}$$

and

$$J(\mathbf{g}) = \int \mathbf{J}(\mathbf{x})\cdot\mathbf{g}(\mathbf{x})\,d^3x, \tag{2.10}$$

we obtain the infinite-dimensional Lie algebra

$$[\rho(f), \rho(g)] = 0, \tag{2.11}$$

$$[\rho(f), J(\mathbf{g})] = i\rho(\mathbf{g}\cdot\nabla f), \tag{2.12}$$

$$[J(\mathbf{f}), J(\mathbf{g})] = iJ(\mathbf{g}\cdot\nabla\mathbf{f} - \mathbf{f}\cdot\nabla\mathbf{g}). \tag{2.13}$$

In Eqs. (2.9)–(2.13) the smearing functions (or their components) belong to Schwartz's space $\mathcal{S}$ of C_∞ functions of rapid decrease.

The action of $\rho(f)$ and $J(\mathbf{g})$ in the Fock representation (2.2)–(2.3) is given by

$$[\rho(f)\Psi]_n = \sum_{j=1}^{n} f(\mathbf{x}_j)\Psi_n \tag{2.14}$$

and

$$[J(\mathbf{g})\Psi]_n = -\tfrac{1}{2}i\sum_{j=1}^{n}[\mathbf{g}(\mathbf{x}_j)\cdot\nabla_j + \nabla_j\cdot\mathbf{g}(\mathbf{x}_j)]\Psi_n. \tag{2.15}$$

The operators $\rho(f)$ and $J(\mathbf{g})$ preserve $\mathcal{H}_n$ as a subspace of $\mathcal{H}$, and restricted to $\mathcal{H}_n$, define the *n-particle representations* of the current algebra (2.11)–(2.13).

A group is obtained by exponentiating the Lie algebra (2.11)–(2.13). Define

$$U(f) = e^{i\rho(f)} \tag{2.16}$$

and

$$V(\varphi_t^{\mathbf{g}}) = e^{iJ(\mathbf{g})}, \tag{2.17}$$

where $\varphi_t^{\mathbf{g}}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the *flow for time t by the vector field* g; i.e.,

$$\frac{\partial\varphi_t^{\mathbf{g}}(\mathbf{x})}{\partial t} = \mathbf{g}(\varphi_t^{\mathbf{g}}(\mathbf{x})), \tag{2.18}$$

and $\varphi_{t=0}^{\mathbf{g}}(\mathbf{x}) = \mathbf{x}$. Then U and V satisfy the group multiplication rules

$$U(f)U(g) = U(f+g), \tag{2.19}$$

$$V(\psi)U(f) = U(f\circ\psi)V(\psi), \tag{2.20}$$

$$V(\varphi)V(\psi) = V(\psi\circ\varphi), \tag{2.21}$$

where $\psi\circ\varphi$ denotes the composition of the flows.

A representation of the group satisfying (2.19)–(2.21) is in fact a representation of the *semidirect product* $\mathcal{S}\wedge\mathcal{K}$, where $\mathcal{S}$ is the group of all f's (under addition) and $\mathcal{K}$ is the group of all ψ's (under composition). The representation theory of such a semidirect product typically focuses attention on the functional

$$L(f) = (\Omega, U(f)\Omega), \tag{2.22}$$

where Ω is a cyclic vector for the $U(f)$'s in the representation.[12]

B. The "N/V" limit

Consider a system of N bosons in a box of volume V. The N-particle representation of the Lie algebra (2.11)–(2.13) describes such a system. Periodic boundary conditions require the smearing functions to be C_∞ functions on the torus T^3, a cube of volume V and of length $2L$ in each spatial direction with corresponding points on opposite boundaries identified. The N-particle representation of the group (2.19)–(2.21) is

$$U_{N,V}(f)\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \exp\left(i\sum_{j=1}^{N}f(\mathbf{x}_j)\right)\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) \tag{2.23}$$

and

$$V_{N,V}(\psi)\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \Psi(\psi(\mathbf{x}_1),\ldots,\psi(\mathbf{x}_N)) \times\left(\det\frac{\partial\psi^k}{\partial x^l}(\mathbf{x})\right)^{1/2} \tag{2.24}$$

where ψ is a C_∞ flow on the torus. The determinant of $\partial\psi^k/\partial x^l$ is the Jacobian of the flow, expressed in the system of local coordinates obtained by the above-mentioned identification of the torus with the cube.

The normalized ground state wave function $\Omega_{N,V}(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ for a system of N free bosons in a box of volume V is

$$\Omega_{N,V}(\mathbf{x}_1,\ldots,\mathbf{x}_N) = (1/\sqrt{V})^N. \tag{2.25}$$

The ground state $\Omega_{N,V}$ is a suitable cyclic vector with which to characterize the representation. Thus we obtain the ground state expectation functional

$$\begin{aligned} L_{N,V}(f) &= (\Omega_{N,V}, U_{N,V}(f)\Omega_{N,V}) \\ &= \int d^3x_1\cdots d^3x_N(1/V)^N\exp\left(i\sum_{j=1}^{N}f(\mathbf{x}_j)\right) \\ &= \left(\frac{1}{V}\int d^3x\,e^{if(\mathbf{x})}\right)^N \end{aligned} \tag{2.26}$$

The functional $L(f)$ in general determines not only the representation of $U(f)$ but also that of $V(\psi)$, at least up to a complex phase "multiplier".[3]

Now it is not possible to take a limit of (2.25) as N and V become infinite, but we can obtain the limit of $L_{N,V}(f)$ as $N, V\to\infty$, with $N/V\to\bar{\rho}$. The constraint $N/V\to\bar{\rho}$, where $\bar{\rho}$ denotes a constant average density, suggests the name "N/V limit" for the procedure used here.

Carrying out this procedure,

$$\begin{aligned} L(f) &= \lim_{\substack{N,V\to\infty \\ N/V\to\bar{\rho}}} L_{N,V}(f) \\ &= \lim_{N\to\infty}\left(1+\frac{\bar{\rho}}{N}\int d^3x[e^{if(\mathbf{x})}-1]\right)^N \\ &= \exp[\bar{\rho}\int(e^{if(\mathbf{x})}-1)\,d^3x]. \end{aligned} \tag{2.27}$$

C. Defining the representation

The Gel'fand–Vilenkin approach to the representation theory of nuclear Lie groups discusses (continuous)

representations of Schwartz's space $\mathcal{S}$ in terms of measures on $\mathcal{S}'$, the continuous dual of $\mathcal{S}$. A functional $L(f)$ is the Fourier transform of a cylindrical measure μ on $\mathcal{S}'$, and thus defines a continuous representation of $\mathcal{S}$, if and only if:

(1) $L(f)$ is continuous with respect to the topology of $\mathcal{S}$,

(2) $L(0) = 1$, and

(3) $L(f)$ is positive definite in the sense that

$$(\forall f_1, \ldots, f_m \in \mathcal{S}, \lambda_1, \ldots, \lambda_m \in \mathfrak{C})$$
$$\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j L(f_j - f_k) \geq 0. \tag{2.28}$$

Under these conditions,

$$L(f) = \int_{F \in \mathcal{S}'} e^{i(F,f)} d\mu(F), \tag{2.29}$$

and the representation of $\mathcal{S}$ may be realized in the Hilbert space $\mathcal{H} = L^2_\mu(\mathcal{S}')$ of μ-square integrable functions on $\mathcal{S}'$, with

$$(U(f)\Psi)(F) = e^{i(F,f)}\Psi(F) \tag{2.30}$$

for $\Psi \in L^2_\mu(\mathcal{S}')$. Furthermore, $\Omega(F) \equiv 1$ is a cyclic vector for the representation.

Therefore, we need to check that Eq. (2.27) satisfies the above three conditions, and indeed defines a representation of $\mathcal{S}$.

Theorem 1: The functional $L(f) = \exp[\bar{\rho} \int (e^{if(\mathbf{x})} - 1) d^3x]$ is the Fourier transform of a cylindrical measure μ on $\mathcal{S}'$, and thus defines a continuous representation U of $\mathcal{S}$, with a cyclic vector Ω such that $L(f) = (\Omega, U(f)\Omega)$.

Proof: (1) $L(f)$ is continuous with respect to the usual topology of $\mathcal{S}$; for if $f_j \to f$ in $\mathcal{S}$ as $j \to \infty$, then $(e^{if_j} - 1) \to (e^{if} - 1)$ in $\mathcal{S}$, and

$$\int (e^{if_j(\mathbf{x})} - 1) d^3x \to \int (e^{if(\mathbf{x})} - 1) d^3x.$$

(2) $L(0) = 1$.

(3) If all of the $f_1, \ldots, f_m$ in Eq. (2.28) have compact support, then $\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j L(f_j - f_k)$ is the N/V limit of the sequence of positive functionals $\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j L_{N,V}(f_j - f_k)$, where V contains the union of the supports of $f_1, \ldots, f_m$. Therefore, Eq. (2.28) holds for functions of compact support. But any $f_1, \ldots, f_m \in \mathcal{S}$ can be approximated arbitrarily closely in $\mathcal{S}$ by C_∞ functions of compact support. Since $L(f)$ is continuous, Eq. (2.28) holds for all $f_1, \ldots, f_m \in \mathcal{S}$. QED

Next we shall explicitly display the representation $U(f)$ defined by the functional $L(f)$ above.[13]

Let $\mathcal{H}$ be the Fock space of a second-quantized canonical nonrelativistic Bose field $\psi(\mathbf{x})$ satisfying Eq. (2.1). Let

$$\psi'(\mathbf{x}) = \psi(\mathbf{x}) + \sqrt{\bar{\rho}}, \qquad \psi'^*(\mathbf{x}) = \psi^*(\mathbf{x}) + \sqrt{\bar{\rho}}. \tag{2.31}$$

Then ψ' and ψ'^* also satisfy canonical commutation relations. The corresponding density is

$$\rho'(\mathbf{x}) = \psi'^*(\mathbf{x})\psi'(\mathbf{x}), \tag{2.32}$$

and $[\rho'(f), \rho'(g)] = 0$.

Theorem 2: With $\rho'(\mathbf{x})$ defined in Fock space by Eq. (2.32), a representation $U_F(f) = e^{i\rho'(f)}$ is obtained for $\mathcal{S}$. The subscript F stands for Fock. The original Fock no-particle state $\Omega_F \in \mathcal{H}$ is cyclic for this representation, and defines an expectation functional $(\Omega_F, U_F(f)\Omega_F) = \exp[\bar{\rho} \int (e^{if(\mathbf{x})} - 1) d^3x]$.

Proof: First let us write $\rho'(f)$ in terms of the original canonical fields:

$$\rho'(f) = \rho(f) + \bar{\rho} \int f(\mathbf{x}) d^3x + \bar{\rho}^{1/2}\psi^*(f) + \bar{\rho}^{1/2}\psi(f), \tag{2.33}$$

where $\rho(f)$ is defined in Eq. (2.4). It is clear that Ω_F is a cyclic vector for the polynomial algebra of operators generated by the identity and the $\rho'(f)$, $f \in \mathcal{S}$. In fact, for a vector Ψ which is an element of $\oplus_{n=0}^{N} \mathcal{H}_n$ in $\mathcal{H}$, $\rho'(f)\Psi \in \oplus_{n=0}^{N+1} \mathcal{H}_n$ in $\mathcal{H}$, with $(\rho'(f)\Psi)_{N+1} = \bar{\rho}^{1/2}\psi^*(f)\Psi_N$. Thus, by the properties of the creation operators $\psi^*(f)$ which follow from Eq. (2.3), if $\oplus_{n=0}^{N} \mathcal{H}_n$ is contained in the closed cyclic subspace generated by applying polynomials in the $\rho'(f)$ to Ω_F, $\oplus_{n=0}^{N+1} \mathcal{H}_n$ is likewise in that subspace. By induction on N, Ω_F is a cyclic vector for the representation.

Next we show that Ω_F is an analytic vector for $\rho'(f)$. In fact, from Eq. (2.33), it is certainly true that for $\Psi \in \oplus_{n=0}^{N} \mathcal{H}_n$ in $\mathcal{H}$,

$$\| \rho'(f)\Psi \| \leq 4(1 + N)(1 + \bar{\rho})(\|f\|_\infty + |\int f(\mathbf{x}) d^3x|)\|\Psi\|, \tag{2.34}$$

where $\|f\|_\infty = \sup_{\mathbf{x} \in \mathbb{R}_3} |f(\mathbf{x})|$. Thus $\sum_{N=0}^{\infty} (t^N/N!)\|\rho'(f)^N\Omega_F\|$ is bounded by $\sum_{N=0}^{\infty} c^N t^N (N+1)$ where c is a constant, and converges for sufficiently small t. Similarly, all elements of $\oplus_{n=0}^{N} \mathcal{H}_n$ are analytic vectors for $\rho'(f)$, for arbitrary N. Having identified a common dense domain of analytic vectors for the $\rho'(f)$, we can now conclude the existence of a unitary representation $U_F(f) = e^{i\rho'(f)}$ in $\mathcal{H}$ with $U_F(f)U_F(g) = U_F(f+g)$.

The cyclicity of Ω_F for the $U_F(f)$ follows immediately from the fact that for Ψ in the domain of $\rho'(f)$, $(1/it)[U_F(tf)\Psi - \Psi] \to \rho'(f)\Psi$ as $t \to 0$.

Finally, it remains for us to evaluate $(\Omega_F, U_F(f)\Omega_F)$. Define the operator-valued distribution

$$A(f) = \bar{\rho} \int f(\mathbf{x}) d^3x + \bar{\rho}^{1/2}\psi^*(f). \tag{2.35}$$

Then

$$\rho'(f)\Omega_F = A(f)\Omega_F, \tag{2.36}$$

and a simple calculation shows that

$$[\rho'(f), A(g)] = A(fg). \tag{2.37}$$

Hence

$$e^{i\rho'(f)}A(g)e^{-i\rho'(f)} = \sum_{n=0}^{\infty} \frac{i^n}{n!} [\mathrm{ad}^n \rho'(f)]A(g) = A(e^{if}g), \tag{2.38}$$

where

$$(\mathrm{ad}X)Y = [X, Y]. \tag{2.39}$$

Now with $L(f) = (\Omega_F, U_F(f)\Omega_F)$,

$$\begin{aligned} \frac{1}{i}\frac{d}{dt} L(tf) &= (\Omega_F, e^{it\rho'(f)}\rho'(f)\Omega_F) \\ &= (\Omega_F, e^{it\rho'(f)}A(f)\Omega_F) = (\Omega_F, A(e^{itf}f)e^{it\rho'(f)}\Omega_F) \\ &= (A^*(e^{itf}f)\Omega_F, e^{it\rho'(f)}\Omega_F) = \bar{\rho} \int e^{itf(\mathbf{x})} f(\mathbf{x}) d^3x\, L(tf). \end{aligned} \tag{2.40}$$

This differential equation in t, when supplemented with the boundary condition $L(0) = 1$, has the unique solution

$$L(tf) = \exp[\bar{\rho} \int (e^{itf(\mathbf{x})} - 1) d^3x]. \tag{2.41}$$

QED

Using the "functional derivative" to be introduced in Sec. 4, Eq. (2.40) may be written

$$\frac{1}{i} \frac{\delta L(f)}{\delta f(\mathbf{x})} = \bar{\rho} e^{if(\mathbf{x})} L(f). \tag{2.42}$$

We have shown that Eq. (2.27) defines a representation of $\mathcal{S}$, the normal subgroup of the semidirect product, and have displayed the representation. Next we show that the full group $\mathcal{S} \wedge \mathcal{K}$ can be represented in the same Hilbert space.

The first step is to anticipate the form of the functional $E(f,\psi) = (\Omega, U(f)\, V(\psi)\, \Omega)$ by taking another N/V limit. Again in the N-particle Fock representation in volume V,

$$\begin{aligned} E_{N,V}(f,\psi) &= (\Omega_{N,V}, U_{N,V}(f)\, V_{N,V}(\psi)\, \Omega_{N,V}) \\ &= \int d^3x_1 \cdots d^3x_N \left(\frac{1}{V}\right)^N \exp\left(i \sum_{j=1}^{N} f(\mathbf{x}_j)\right) \left(\prod_{n=1}^{N} \mathcal{J}_\psi(\mathbf{x}_n)\right)^{1/2}, \end{aligned} \tag{2.43}$$

where

$$\mathcal{J}_\psi(\mathbf{x}) = \det \frac{\partial \psi^k}{\partial x^l}(\mathbf{x}) \tag{2.44}$$

is the Jacobian referred to in Eq. (2.24). Then

$$\begin{aligned} E(f,\psi) &= \lim_{\substack{N,V\to\infty \\ N/V\to\bar{\rho}}} E_{N,V}(f,\psi) \\ &= \exp\{\bar{\rho} \int [e^{if(\mathbf{x})}\sqrt{\mathcal{J}_\psi(\mathbf{x})} - 1]\, d^3x\}. \end{aligned} \tag{2.45}$$

A functional E on a topological group G defines a continuous representation of G, if and only if[14]:

(1) E is continuous,
(2) $E(1) = 1$, and
(3) $\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j E(g_k^{-1} g_j) \geq 0$

$$(\forall g_1, \ldots, g_m \in G, \lambda_1, \ldots, \lambda_m \in \mathbb{C}). \tag{2.46}$$

Now we are ready to prove the next result.

Theorem 3: There exists a representation $U(f)V(\psi)$ of $\mathcal{S} \wedge \mathcal{K}$ in a Hilbert space $\mathcal{H}$, with a cyclic vector $\Omega \in \mathcal{H}$, such that $E(f,\psi) = (\Omega, U(f)V(\psi)\Omega)$ is given by Eq. (2.45).

Proof: We shall show that conditions (1)–(3) above are satisfied by $E(f,\psi)$.

(1) It is necessary to introduce a more careful definition of $\mathcal{K}$.[15] Let $\mathcal{K}_0$ be the group of all C_∞ diffeomorphisms from $\mathbb{R}^3$ onto $\mathbb{R}^3$, having compact support. We topologize $\mathcal{K}_0$ by means of the countable family of metrics

$$\langle\langle \varphi, \psi \rangle\rangle_n = \max_{0 \leq |m| \leq n} \sup_{\mathbf{x} \in \mathbb{R}^3} |(1 + |\mathbf{x}|^2)^n [\varphi^{(m)}(\mathbf{x}) - \psi^{(m)}(\mathbf{x})]|,$$

$$n = 0, 1, 2, \cdots \tag{2.47}$$

for $\varphi, \psi \in \mathcal{K}_0$, where $(m) = (m_1, m_2, m_3)$, $|m| = \sum_{k=1}^{3} m_k$, and $\varphi^{(m)}(\mathbf{x}) = \partial^{|m|} \varphi(\mathbf{x}) / (\partial x^1)^{m_1} (\partial x^2)^{m_2} (\partial x^3)^{m_3}$. $\mathcal{K}$ is the completion of $\mathcal{K}_0$ with respect to this topology. The topology has a countable basis of neighborhoods of each element of $\mathcal{K}$, and is metrizable. The group operations are continuous. $\mathcal{K}$ contains diffeomorphisms which are not of compact support, but which suitably approximate the identity mapping as $|\mathbf{x}| \to \infty$.

Omitting the computations, it follows that if $f_j \to f$ in $\mathcal{S}$ and $\psi_k \to \psi$ in $\mathcal{K}$ as $k, j \to \infty$, then $(e^{if_j}\sqrt{\mathcal{J}_{\psi_k}} - 1) \to (e^{if}\sqrt{\mathcal{J}_\psi} - 1)$ in $\mathcal{S}$, and

$$\int [e^{if_j(\mathbf{x})}\sqrt{\mathcal{J}_{\psi_k}(\mathbf{x})} - 1] d^3x \to \int [e^{if(\mathbf{x})}\sqrt{\mathcal{J}_\psi(\mathbf{x})} - 1] d^3x.$$

Thus $E(f,\psi)$ is continuous.

(2) Clearly $E(0,1) = 1$.
(3) As in the proof of Theorem 1, choose first the elements $(f_1,\psi_1), \ldots, (f_m,\psi_m)$ to have compact support. Then with

$$\begin{aligned} (f_k,\psi_k)^{-1}(f_j,\psi_j) &= (-f_k \circ \psi_k^{-1}, \psi_k^{-1})\,(f_j,\psi_j) \\ &= ([f_j - f_k] \circ \psi_k^{-1}, \psi_j \circ \psi_k^{-1}), \end{aligned} \tag{2.48}$$

the expression

$$\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j E([f_j - f_k] \circ \psi_k^{-1}, \psi_j \circ \psi_k^{-1}) \tag{2.49}$$

is the limit of the sequence

$$\sum_{j,k=1}^{m} \bar{\lambda}_k \lambda_j E_{N,V}([f_j - f_k] \circ \psi_k^{-1}, \psi_j \circ \psi_k^{-1}) \tag{2.50}$$

as $N, V \to \infty$ with $N/V \to \bar{\rho}$, where the volume V contains the union of all of the supports of $f_1, \ldots, f_m$ and $\psi_1, \ldots, \psi_m$. But Eq. (2.50) is positive since $E_{N,V}$ is defined in the N-particle Fock representation in volume V by Eq. (2.43).

Therefore Eq. (2.49) is positive for elements of $\mathcal{S} \wedge \mathcal{K}$ which have compact support. But any element of $\mathcal{S} \wedge \mathcal{K}$ can be approximated arbitrarily closely by elements having compact support, due to the definition of $\mathcal{K}$ as the completion of $\mathcal{K}_0$. Since $E(f,\psi)$ is continuous, Eq. (2.49) is positive for all elements of $\mathcal{S} \wedge \mathcal{K}$. QED

Thus there exists a continuous representation $U(f)V(\psi)$ of $\mathcal{S} \wedge \mathcal{K}$ in a Hilbert space $\mathcal{H}$, with $\Omega \in \mathcal{H}$ cyclic for the $U(f)V(\psi)$, such that

$$E(f,\psi) = (\Omega, U(f)V(\psi)\Omega). \tag{2.51}$$

The next step is to show that Ω is a cyclic vector for the subgroup $\{U(f)\}$. We shall use the following lemma, omitting the proof which is not difficult.[16]

Lemma 1: Let $U(t) = e^{itA}$ be a continuous one-parameter unitary group in $\mathcal{H}$ with A self-adjoint; let $\Psi \in \mathcal{H}$ and let $f(t) = (\Psi, U(t)\Psi)$ be an entire analytic function of t. Then Ψ is an entire analytic vector for A; i.e., the series

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} \| A^n \Psi \| \tag{2.52}$$

is absolutely convergent for all t; and Ψ is in the domain of $U(it)$.

Theorem 4: In the representation of $\mathcal{S} \wedge \mathcal{K}$ defined by Eqs. (2.45) and (2.51), Ω is cyclic for the $\{U(f)\}$.

Proof: Let $h \in \mathcal{S}$. Then with $U(h) = e^{i\rho(h)}$, Eq. (2.45) yields

$$(\Omega, U(th)\Omega) = \exp[\bar{\rho} \int (e^{ith(\mathbf{x})} - 1) d^3x] \tag{2.53}$$

which is an entire analytic function of t. Therefore by Lemma 1, Ω is in the domain of $e^{\rho(h)}$. $e^{\rho(h)}\Omega$ is of course in the closed cyclic subspace generated by the $\{U(f)\Omega\}$.

But it can be shown that $V(\psi)\Omega = e^{\rho(h)}\Omega$, where $h(\mathbf{x}) = \ln\sqrt{\mathcal{J}_\psi(\mathbf{x})}$. In fact,

$$\begin{aligned}
&\| V(\psi)\Omega - e^{\rho(h)}\Omega \|^2 \\
&\quad = 1 - (\Omega, U(-ih)V(\psi)\Omega) - (\Omega, V(\psi^{-1})U(-ih)\Omega) \\
&\qquad + (\Omega, U(-2ih)\Omega) \\
&\quad = 1 - \exp\{\bar\rho \int [e^{h(\mathbf{x})}\sqrt{\mathcal{J}_\psi(\mathbf{x})} - 1]d^3x\} \\
&\qquad - \exp\{\bar\rho \int [e^{h\circ\psi^{-1}(\mathbf{x})}\sqrt{\mathcal{J}_{\psi^{-1}}(\mathbf{x})} - 1]d^3x\} \\
&\qquad + \exp[\bar\rho \int (e^{2h(\mathbf{x})} - 1)d^3x] \\
&\quad = 0, \qquad (2.54)
\end{aligned}$$

after some manipulation of the Jacobians. Thus Ω is cyclic for the $\{U(f)\}$. QED

Now we are ready to represent the full group $\mathcal{S} \wedge \mathcal{K}$ in the Fock space of Theorem 2.

Theorem 5: With $\psi'(\mathbf{x}) = \psi(\mathbf{x}) + \bar\rho^{1/2}$ as in Eq. (2.31), with $\rho'(\mathbf{x})$ given by Eq. (2.32), and with

$$\mathbf{J}'(\mathbf{x}) = (1/2i)[\psi'^*(\mathbf{x})\nabla\psi'(\mathbf{x}) - (\nabla\psi'^*(\mathbf{x}))\psi'(\mathbf{x})] \qquad (2.55)$$

in the Fock space of the nonrelativistic canonical Bose field $\psi(\mathbf{x})$, there exists a continuous unitary representation $U_F(f)$ $V_F(\psi)$ of the group $\mathcal{S} \wedge \mathcal{K}$ such that

$$U_F(f) = e^{i\rho'(f)} \qquad (2.56)$$

and

$$V_F(\varphi_t\mathbf{g}) = e^{i\mathbf{J}'(\mathbf{g})}. \qquad (2.57)$$

Then with $\Omega_F \in \mathcal{K}$ the original Fock vacuum state for $\psi(\mathbf{x})$,

$$E(f,\psi) = (\Omega_F, U_F(f)V_F(\psi)\Omega_F) \qquad (2.58)$$

where $E(f,\psi)$ is given by Eq. (2.45).

Proof: First we assert that the representation of $\mathcal{S} \wedge \mathcal{K}$ obtained in Theorem 3 can be mapped unitarily into the Fock Hilbert space.

Let $\Omega \to \Omega_F$ and $U(f)\Omega \to U_F(f)\Omega_F$, where Ω_F and $U_F(f)$ are as in Theorem 2. Since by Theorem 4, Ω is cyclic for the $U(f)$, this mapping defines a unitary representation not only of $\mathcal{S}$ but of $\mathcal{S} \wedge \mathcal{K}$ in the Fock Hilbert space; we may write $V(\psi) \to V_F(\psi)$, and

$$E(f,\psi) = (\Omega_F, U_F(f)\, V_F(\psi)\Omega_F).$$

It remains only to show that $J'(\mathbf{g})$ as defined by Eq. (2.55) is indeed the infinitesimal generator of the one-parameter unitary group $V_F(\varphi_t\mathbf{g})$. By Stone's theorem, it is sufficient to show that

$$\lim_{t\to 0} \frac{V_F(\varphi_t\mathbf{g}) - I}{it}\Omega_F = J'(\mathbf{g})\Omega_F; \qquad (2.59)$$

the result then follows from the fact that $\rho'(f)$ and $J'(\mathbf{g})$ satisfy the correct algebra of commutation relations on the domain of polynomials in the $\rho'(f)$ applied to Ω_F.

Now by Eq. (2.55),

$$J'(\mathbf{g})\Omega_F = (1/2i)\rho'(\nabla\cdot\mathbf{g})\Omega_F \qquad (2.60)$$

and

$$\begin{aligned}
&\lim_{t\to 0} \left\| \frac{V_F(\varphi_t\mathbf{g}) - I}{it}\Omega_F - \frac{1}{2i}\rho'(\nabla\cdot\mathbf{g})\Omega_F \right\|^2 \\
&\quad = \lim_{s,t\to 0} \left\| \frac{V_F(\varphi_t\mathbf{g}) - I}{it}\Omega_F + \frac{1}{2}\frac{U_F(s\nabla\cdot\mathbf{g}) - I}{s}\Omega_F \right\|^2 \\
&\quad = 0, \qquad (2.61)
\end{aligned}$$

using Eq. (2.45) for $E(f,\psi)$. Thus Eq. (2.59) is demonstrated. QED

To summarize, we have obtained a representation of the group $\mathcal{S} \wedge \mathcal{K}$, the exponentiated nonrelativistic current algebra, in the N/V limit. This was done by means of the expectation functional $E(f,\psi)$. The representation thus obtained was shown to be unitarily equivalent to an explicit representation of the current commutators in a certain Fock space, with the original Fock ground state being the cyclic vector defining the functional $E(f,\psi)$.

In the next section we show how a condition on the cyclic vector which asserts that it is the physical ground state of an infinite free Bose gas uniquely determines the class of representations obtained above; namely those defined by $E(f,\psi)$ for an arbitrarily specified average particle density $\bar\rho$.

3. A CONDITION ON THE GROUND STATE

In this section we study representations of the current algebra, Eqs. (2.11)–(2.13), in which there exists a cyclic vector Ω_0 satisfying the condition

$$[2i\mathbf{J}(\mathbf{x}) + (\nabla\rho)(\mathbf{x})]\Omega_0 = 0. \qquad (3.1)$$

Ω_0 will usually be interpreted as the ground state of the system.

Convincing heuristic arguments that Eq. (3.1) determines the ground state of a noninteracting Bose gas have been given.[11,17–19] Here we shall explore the consequences of this constraint somewhat more systematically. We show that for a system in a box with periodic boundary conditions, Eq. (3.1) implies that the operator $\int\rho(\mathbf{x})d^3x$ has integer eigenvalues. In Sec. 2 we saw that the expectation functional $(\Omega_0, e^{i\rho(f)}\Omega_0)$ is given by Eq. (2.27) in the N/V limit. In this section we show not only that Eq. (2.27) determines a representation satisfying the constraint (3.1), as has been previously indicated,[19] but that it defines the unique class of representations having this property.

Let us investigate the consequences of Eq. (3.1) on the functional

$$L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0), \qquad (3.2)$$

where Ω_0 is a cyclic vector in a representation of the current algebra, satisfying the condition

$$2iJ(\mathbf{g})\Omega_0 = \rho(\nabla\cdot\mathbf{g})\Omega_0 \qquad (3.3)$$

for all real vector functions $\mathbf{g}$ with components in $\mathcal{S}$. Naturally, we shall assume that Ω_0 is in the domain of the operators $J(\mathbf{g})$ and $\rho(f)$ for all $f, \mathbf{g} \in \mathcal{S}$. Actually, for the sake of mathematical rigor we shall assume slightly more. We suppose in addition that the bilinear form $(\rho(f_1)\Omega_0, \rho(f_2)\Omega_0)$ is continuous in f_1 and f_2; i.e., if $f_{1n} \to f_1$ in $\mathcal{S}$, then $(\rho(f_{1n})\Omega_0, \rho(f_2)\Omega_0) \to (\rho(f_1)\Omega_0, \rho(f_2)\Omega_0)$, and similarly for f_2. It then follows that $\| \rho(f_n)\Omega_0 \|^2 \to 0$ if $f_n \to 0$ as $n \to \infty$. This assumption is slightly stronger than assuming continuity of the group representation $U(f)$. It follows readily that $L(f)$

is continuous in f; i.e., if $f_n \to f$ in $\mathcal{S}$, then $L(f_n) \to L(f)$. In fact,

$$\begin{aligned} |L(f_n) - L(f)| &= |(\Omega_0, (e^{i\rho(f_n)} - e^{i\rho(f)})\Omega_0)| \\ &= |(e^{-i\rho(f)}\Omega_0, (e^{i\rho(f_n - f)} - I)\Omega_0)| \\ &\le \| e^{-i\rho(f)}\Omega_0 \| \cdot \| (e^{i\rho(f_n - f)} - I)\Omega_0 \| \\ &\le 1 \cdot \| \rho(f_n - f)\Omega_0 \| \to 0 \end{aligned}$$

as $f_n \to f$.

For the case of a system in a "box" with periodic boundary conditions, assume that Eq. (3.3) holds for all infinitely differentiable periodic vector functions $\mathbf{g}$. The components of such functions will be said to be in $\mathcal{S}_V$, where $\mathcal{S}_V$ has the topology of a nuclear space.

Using Eq. (3.3), we derive a functional equation for $L(f)$ as follows. We have

$$\frac{d}{dt} L(f + t\nabla\cdot\mathbf{g})_{t=0} = i(\Omega_0, e^{i\rho(f)}\rho(\nabla\cdot\mathbf{g})\Omega_0),$$

whence it follows using Eq. (3.3) that

$$\frac{d}{dt} L(f + t\nabla\cdot\mathbf{g})_{t=0} = -2(\Omega_0, e^{i\rho(f)}J(\mathbf{g})\Omega_0). \tag{3.4}$$

Similarly,

$$\begin{aligned} \frac{d}{dt} L(f + t\nabla\cdot\mathbf{g})_{t=0} &= i(\Omega_0, \rho(\nabla\cdot\mathbf{g})\, e^{i\rho(f)}\Omega_0) \\ &= i(\rho(\nabla\cdot\mathbf{g})\Omega_0, e^{i\rho(f)}\Omega_0) \\ &= 2(J(\mathbf{g})\Omega_0, e^{i\rho(f)}\Omega_0) \\ &= 2(\Omega_0, J(\mathbf{g})\, e^{i\rho(f)}\Omega_0). \end{aligned} \tag{3.5}$$

Combining Eqs. (3.4) and (3.5), we have

$$\frac{d}{dt} L(f + t\nabla\cdot\mathbf{g})_{t=0} = (\Omega_0, [J(\mathbf{g}), e^{i\rho(f)}]\Omega_0). \tag{3.6}$$

Now it follows from the current commutation relations that

$$e^{i\rho(f)}J(\mathbf{g})e^{-i\rho(f)} = J(\mathbf{g}) - \rho(\mathbf{g}\cdot\nabla f). \tag{3.7}$$

Combining Eqs. (3.6) and (3.7),

$$\begin{aligned} &\frac{d}{dt} L(f + t\nabla\cdot\mathbf{g})_{t=0} \\ &\quad = (\Omega_0, [J(\mathbf{g}) - e^{i\rho(f)}J(\mathbf{g})e^{-i\rho(f)}]e^{i\rho(f)}\Omega_0) \\ &\quad = (\Omega_0, \rho(\mathbf{g}\cdot\nabla f)\, e^{i\rho(f)}\Omega_0) \\ &\quad = -i\frac{d}{dt} L(f + t\mathbf{g}\cdot\nabla f)_{t=0}. \end{aligned} \tag{3.8}$$

This equation can also be written in the form

$$(\Omega_0, e^{i\rho(f)}[\rho(\nabla\cdot\mathbf{g}) + i\rho(\mathbf{g}\cdot\nabla f)]\Omega_0) = 0 \tag{3.9}$$

for all $f, \mathbf{g} \in \mathcal{S}$ or for all (periodic, infinitely differentiable) $f, \mathbf{g} \in \mathcal{S}_V$ for a system in a "box".

Thus we have obtained a functional equation for $L(f)$ from the original condition on the ground state Ω_0.

Next we shall show that Eq. (3.8) or (3.9) implies that $L(f)$ must be of the form

$$L(f) = F(K(f)), \tag{3.10}$$

where

$$K(f) = \int (e^{if(\mathbf{x})} - 1)\, d^3x$$

and $F(z)$ is a holomorphic function of the complex variable z in the interior of the range of $K(f)$.

In order to prove this result we will need the following two lemmas. We say that the mapping $t \to g_t$ of the interval $[0, 1]$ into $\mathcal{S}$ (respectively $\mathcal{S}_V$) is a *differentiable* mapping of $[0, 1]$ into $\mathcal{S}$ (resp. $\mathcal{S}_V$) with derivative $dg_t/dt = k_t \in \mathcal{S}$ (resp. $\mathcal{S}_V$) if for each $t \in [0, 1]$ we have that $h^{-1}(g_{t+h} - g_t) \to k_t$ as $h \to 0$; where the convergence is in the topology of $\mathcal{S}$ (resp. $\mathcal{S}_V$).

Lemma 2: Suppose that $t \to g_t \in \mathcal{S}$ (resp. $\mathcal{S}_V$) for $0 \le t \le 1$ is a differentiable mapping of $[0, 1]$ into $\mathcal{S}$ (resp. $\mathcal{S}_V$). Furthermore, suppose that

$$K(g_t) = \int (e^{ig_t(\mathbf{x})} - 1)\, d^3x = \text{a constant}; \tag{3.11}$$

that $L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0)$ satisfies (3.8) or equivalently (3.9); and that $L(f)$ is continuous in f with respect to the topology of $\mathcal{S}$ (resp. $\mathcal{S}_V$). Then $L(g_t)$ is a constant independent of t.

Proof: We shall prove the lemma by showing that $(d/dt)\, L(g_t) = 0$. Let $k_t = dg_t/dt$. We begin by showing that

$$\frac{d}{dt} L(g_t) = i(\Omega_0, e^{i\rho(g_t)}\rho(k_t)\Omega_0). \tag{3.12}$$

Now we have

$$\begin{aligned} &h^{-1}[L(g_{t+h}) - L(g_t)] \\ &\quad = h^{-1}(\Omega_0, (e^{i\rho(g_{t+h})} - e^{i\rho(g_t)})\Omega_0) \\ &\quad = h^{-1}(\Omega_0, e^{i\rho(g_t)}(e^{i\rho(g_{t+h} - g_t - hk_t)}\, e^{ih\rho(k_t)} - I)\Omega_0) \\ &\quad = h^{-1}(\Omega_0, e^{i\rho(g_t)}\, e^{i\rho(g_{t+h} - g_t - hk_t)}(e^{ih\rho(k_t)} - I)\Omega_0) \\ &\qquad + h^{-1}(\Omega_0, e^{i\rho(g_t)}(e^{i\rho(g_{t+h} - g_t - hk_t)} - I)\Omega_0). \end{aligned} \tag{3.13}$$

Estimating the second term in Eq. (3.13) as $h \to 0$, we find that

$$\begin{aligned} &|h^{-1}(\Omega_0, e^{i\rho(g_t)}(e^{i\rho(g_{t+h} - g_t - hk_t)} - I)\Omega_0)| \\ &\quad \le \| e^{-i\rho(g_t)}\Omega_0 \| \, \| h^{-1}(e^{i\rho(g_{t+h} - g_t - hk_t)} - I)\Omega_0 \| \\ &\quad \le 1 \cdot \| \rho\left(\frac{g_{t+h} - g_t}{h} - k_t\right)\Omega_0 \| \to 0 \quad \text{as } h \to 0, \end{aligned}$$

since $h^{-1}(g_{t+h} - g_t) - k_t \to 0$ in $\mathcal{S}$ (resp. $\mathcal{S}_V$) as $h \to 0$. Hence, we have

$$\begin{aligned} &\lim_{h\to 0} h^{-1}[L(g_{t+h}) - L(g_t)] \\ &\quad = \lim_{h\to 0} \left(e^{-i\rho(g_t)}e^{-i\rho(g_{t+h} - g_t - hk_t)}\Omega_0, \frac{(e^{ih\rho(k_t)} - I)}{h}\Omega_0\right) \\ &\quad = i(\Omega_0, e^{i\rho(g_t)}\rho(k_t)\Omega_0). \end{aligned} \tag{3.14}$$

Therefore, $L(g_t)$ is differentiable and

$$\frac{dL(g_t)}{dt} = i(\Omega_0, e^{i\rho(g_t)}\rho(k_t)\Omega_0), \tag{3.15}$$

where $k_t = dg_t/dt$.

Next we show that Eq. (3.11) implies

$$(\Omega_0, e^{i\rho(g_t)}\rho(k_t)\Omega_0) = 0. \tag{3.16}$$

Let T be the tempered distribution defined by

$$T(f) = (\Omega_0, e^{i\rho(g)}\rho(f)\Omega_0). \tag{3.17}$$

From Eq. (3.9) we have

$$T(\nabla\cdot\mathbf{h} + i\mathbf{h}\cdot\nabla g) = 0 \tag{3.18}$$

for $f, \mathbf{h} \in S$ (resp. $\mathcal{S}_V$). Let T_g be the distribution defined by

$$T_g(f) = T(e^{-ig} f) \tag{3.19}$$

for $g \in \mathcal{S}$ (resp. $\mathcal{S}_V$). Then we have

$$\begin{aligned} T_g(\nabla \cdot \mathbf{f}) &= T(e^{-ig}\nabla \cdot \mathbf{f}) \\ &= T(\nabla \cdot (e^{-ig}\mathbf{f}) + ie^{-ig}\mathbf{f} \cdot \nabla g) = 0 \end{aligned} \tag{3.20}$$

by Eq. (3.18). Hence,

$$\nabla T_g(\mathbf{x}) = 0 \quad \text{and} \quad T_g(\mathbf{x}) = c_g,$$

where c_g is a constant depending on g. Thus from Eq. (3.19) we have

$$T(\mathbf{x}) = c_g e^{ig(\mathbf{x})}. \tag{3.21}$$

From Eq. (3.11),

$$\frac{d}{dt} \int (e^{ig_t(\mathbf{x})} - 1)\, d^3x = i \int k_t(\mathbf{x})\, e^{ig_t(\mathbf{x})}\, d^3x = 0. \tag{3.22}$$

Therefore, by Eqs. (3.21) and (3.15),

$$\begin{aligned} \frac{d}{dt} L(g_t) &= i(\Omega_0, e^{i\rho(g_t)}\rho(k_t)\Omega_0) \\ &= iT(k_t) = 0. \end{aligned} \qquad \text{QED}$$

The proof of the next lemma is extremely technical; therefore we shall present a mere sketch for the infinite-volume case in the Appendix.

Lemma 3: Suppose $g_1, g_2 \in \mathcal{S}$ (resp. $\mathcal{S}_V$) and

$$\int (e^{ig_1(\mathbf{x})} - 1)\, d^3x = \int (e^{ig_2(\mathbf{x})} - 1)\, d^3x. \tag{3.23}$$

Then for any two neighborhoods N_1 of g_1 and N_2 of g_2 in $\mathcal{S}$ (resp. $\mathcal{S}_V$), there exist functions $h_1 \in N_1$ and $h_2 \in N_2$, and a continuous mapping $t \to f_t$ of $[0,1]$ into $\mathcal{S}$ (resp. $\mathcal{S}_V$), differentiable in the open interval $(0,1)$, such that $f_0 = h_1, f_1 = h_2$ and

$$\int (e^{if_t(\mathbf{x})} - 1)\, d^3x = \text{a constant}.$$

This lemma is easy to prove if f_t is permitted to be complex. The requirement that f_t be real for $0 \le t \le 1$ complicates the proof considerably.

Proof: See Appendix.

Theorem 6: Suppose that $L(g) = (\Omega_0, e^{i\rho(g)}\Omega_0)$, defined for all real $g \in \mathcal{S}$ (resp. $\mathcal{S}_V$), is continuous in g with respect to the topology of $\mathcal{S}$ (resp. $\mathcal{S}_V$). Furthermore, suppose that $L(g)$ satisfies Eq. (3.8) or equivalently Eq. (3.9). Then $L(g)$ is of the form

$$L(g) = F(K(g)),$$

where $K(g) = \int (e^{ig(\mathbf{x})} - 1)\, d^3x$ and where $F(z)$ is a holomorphic function of the complex variable z in the interior of the range of $K(g)$.

Proof: First we show that if $g_1, g_2 \in \mathcal{S}$ (resp. $\mathcal{S}_V$) and

$$\int (e^{ig_1(\mathbf{x})} - 1)\, d^3x = \int (e^{ig_2(\mathbf{x})} - 1)\, d^3x,$$

then $L(g_1) = L(g_2)$. Suppose $\epsilon > 0$. Since $L(g)$ is continuous in g there are neighborhoods N_1 of g_1 and N_2 of g_2 in $\mathcal{S}$ (resp. $\mathcal{S}_V$) such that

$$\begin{aligned} &|L(g_1) - L(h_1)| < \epsilon/2 \quad \text{for all } h_1 \in N_1, \\ &|L(g_2) - L(h_2)| < \epsilon/2 \quad \text{for all } h_2 \in N_2. \end{aligned} \tag{3.24}$$

From Lemma 3 it follows that there exist functions $k_1 \in N_1$ and $k_2 \in N_2$ and a continuous mapping $t \to f_t$ of $[0,1]$ into $\mathcal{S}$ (resp. $\mathcal{S}_V$), differentiable in $(0,1)$, such that $f_0 = k_1, f_1 = k_2$ and

$$\int (e^{if_t(\mathbf{x})} - 1)\, d^3x = \text{a constant}.$$

Then, by Lemma 2, we have that $L(k_1) = L(k_2)$. By the inequality (3.24), it follows that $|L(g_1) - L(g_2)| < \epsilon$. Since $\epsilon > 0$ is arbitrary, $L(g_1) = L(g_2)$. Hence, $L(g)$ depends only on the number

$$K(g) = \int (e^{ig(\mathbf{x})} - 1)\, d^3x.$$

Thus we have $L(g) = F(K(g))$, where $F(z)$ is a complex function defined on the range of $K(g)$ for all $g \in \mathcal{S}$ (resp. $\mathcal{S}_V$). We note that for $\mathcal{S}$ we have that the range of $K(g)$ is $\{z; \mathrm{Re}\, z < 0 \text{ or } z = 0\}$, while for $\mathcal{S}_V$, the range of $K(g)$ is $\{z; |z + V| \le V\}$. Next we show that $F(z)$ is differentiable for z in the interior of the range of $K(g)$.

If $K(g_1)$ is a point in the interior of the range of $K(g)$, then g_1 is not a constant function. Then there are real functions h_1 and h_2 which have the property that, as the point (t_1, t_2) runs over a two-dimensional neighborhood of $(0,0)$, $K(g_1 + t_1h_1 + t_2h_2)$ runs over a complex neighborhood of $K(g)$. The mapping $(t_1, t_2) \to K(g + t_1h_1 + t_2h_2)$ is analytic in t_1 and t_2, and since $L(g_1 + t_1h_1 + t_2h_2)$ is differentiable in t_1 and t_2, it follows from the continuity assumptions on $L(g)$ that $F(z)$ is differentiable in a neighborhood of $K(g_1)$. Hence $F(z)$ is differentiable in the interior of the range of $K(g)$. Since $L(g)$ is continuous in g, it follows that $F(z)$ is continuous on the whole range of $K(g)$.

Next we show that $F(z)$ is holomorphic for z in the interior of the range of $K(g)$. To prove this it is sufficient to show that

$$\frac{\partial}{\partial u} F(u + iv) = -i \frac{\partial}{\partial v} F(u + iv)$$

for $z = u + iv$ in the interior of the range of $K(g)$. Since $F(z)$ is differentiable, we have

$$\begin{aligned} \frac{d}{dt} L(g + th)\Big|_{t=0} &= \frac{\partial F(K(g))}{\partial u} \frac{\partial}{\partial t} \{\mathrm{Re}[K(g + th)]\}\Big|_{t=0} \\ &\quad + \frac{\partial F(K(g))}{\partial v} \frac{\partial}{\partial t} \{\mathrm{Im}[K(g + th)]\}\Big|_{t=0} \\ &= -\frac{\partial F(K(g))}{\partial u} \int h(\mathbf{x})\, \mathrm{sin}g(\mathbf{x}) dx^3 \\ &\quad + \frac{\partial F(K(g))}{\partial v} \int h(\mathbf{x})\, \mathrm{cos}g(\mathbf{x})\, d^3x. \end{aligned} \tag{3.25}$$

Then from Eq. (3.8) we obtain

$$\begin{aligned} &-\frac{\partial F(K(g))}{\partial u} \int (\nabla \cdot \mathbf{f})(\mathbf{x})\, \mathrm{sin}g(\mathbf{x})\, d^3x \\ &+ \frac{\partial F(K(g))}{\partial v} \int (\nabla \cdot \mathbf{f})(\mathbf{x})\, \mathrm{cos}g(\mathbf{x})\, d^3x \\ &+ i\frac{\partial F(K(g))}{\partial u} \int (\mathbf{f} \cdot \nabla g)(\mathbf{x})\, \mathrm{sin}g(\mathbf{x})\, d^3x \\ &+ i\frac{\partial F(K(g))}{\partial v} \int (\mathbf{f} \cdot \nabla g)(\mathbf{x})\, \mathrm{cos}g(\mathbf{x})\, d^3x = 0. \end{aligned} \tag{3.26}$$

From the divergence theorem, we obtain the relationships

$$\begin{aligned} \int (\nabla \cdot \mathbf{f})(\mathbf{x})\, \mathrm{sin}g(\mathbf{x})\, d^3x &= -\int (\mathbf{f} \cdot \nabla g)(\mathbf{x})\, \mathrm{cos}g(\mathbf{x})\, d^3x, \\ \int (\nabla \cdot \mathbf{f})(\mathbf{x})\, \mathrm{cos}g(\mathbf{x})\, d^3x &= \int (\mathbf{f} \cdot \nabla g)(\mathbf{x})\, \mathrm{sin}g(\mathbf{x})\, d^3x. \end{aligned} \tag{3.27}$$

Combining (3.26) and (3.27) we find

$$\left(\frac{\partial F(K(g))}{\partial u} + i\frac{\partial F(K(g))}{\partial v}\right)\int (\nabla\cdot\mathbf{f})(\mathbf{x})\, e^{-ig(\mathbf{x})}\, d^3x = 0.$$

If g is not a constant, one can find an $\mathbf{f}$ with components in $\mathcal{S}$ (resp. $\mathcal{S}_V$) such that

$$\int (\nabla\cdot\mathbf{f})(\mathbf{x})\, e^{-ig(\mathbf{x})}\, d^3x \neq 0.$$

Hence, $\partial F/\partial u = -i(\partial F/\partial v)$, and $F(z)$ is holomorphic for z in the interior of the range of $K(g)$. This completes the proof of the theorem. QED

Next we shall determine the explicit form of $L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0)$ under the further assumption that $U(f)\,V(\psi)$ determines a factor representation of the current algebra. The importance of factor representations lies in the fact that every representation of a C^*-algebra (in particular, the C^*-algebra associated with currents) can be uniquely decomposed into a direct integral of factor representations.[20] Roughly speaking, if one knows all of the factor representations of a C^*-algebra, one can construct all representations by taking direct integrals.

Suppose we have a continuous unitary representation of $\mathcal{S}\wedge\mathcal{K}$. We denote by $\mathfrak{U}$ the *-algebra of polynomials in $U(f)$ and $V(\psi)$, with $f\in\mathcal{S}$ and $\psi\in\mathcal{K}$, and by $\mathfrak{U}'$ the commutant of $\mathfrak{U}$, i.e., the set of all bounded operators which commute with the elements of $\mathfrak{U}$. Finally, we denote by $\mathfrak{U}''$ the bicommutant of $\mathfrak{U}$, i.e., the commutant of $\mathfrak{U}'$. It follows from a theorem of von Neumann[21] that $\mathfrak{U}''$ is the strong closure of $\mathfrak{U}$, i.e., for $A\in\mathfrak{U}''$ and any finite set of vectors $\{\Psi_i; i=1,\ldots,n\}$ in the Hilbert space of the representation, and for $\epsilon>0$, there exists a $B\in\mathfrak{U}$ such that $\|(A-B)\Psi_i\| < \epsilon$. The representation $U(f)\,V(\psi)$ is said to be a factor representation if $\mathfrak{U}'\cap\mathfrak{U}'' = \{\lambda I\}$; i.e., if the only operators common to both $\mathfrak{U}'$ and $\mathfrak{U}''$ are multiples of the identity. Every irreducible representation of the current algebra is a factor representation, since for irreducible representations $\mathfrak{U}' = \{\lambda I\}$.

Let us turn to the question of determining $L(g) = (\Omega_0, e^{i\rho(g)}\Omega_0)$ for a factor representation with a vector Ω_0 satisfying Eq. (3.1) or equivalently Eq. (3.3).

We begin with the case of a system in a box. Since the function $e_0(\mathbf{x})\equiv 1$ is in $\mathcal{S}_V$, we can consider the operator $U(\lambda e_0) = \exp[i\lambda\rho(e_0)]$. Since $U(\lambda e_0)$ commutes with $U(f)V(\psi)$ for all $(f,\psi)\in\mathcal{S}\wedge\mathcal{K}$, $U(\lambda e_0)$ is in the center of the current algebra. Then for a factor representation we must have $U(\lambda e_0) = \omega(\lambda)I$, where $|\omega(\lambda)| = 1$. By the group property and by Stone's theorem, we then have $\exp[i\lambda\rho(e_0)] = \exp[i\lambda Q]I$, and $\rho(e_0) = QI$, where Q is to be interpreted as the total number of particles in the system. But we have already seen that $L(g) = F(K(g))$ where $F(z)$ is analytic for $|z+V|<V$ and continuous for $|z+V|\leq V$.

Now we have

$$F(K(\lambda e_0)) = F(Ve^{i\lambda} - V) = e^{iQ\lambda}. \tag{3.28}$$

Since F is single-valued, $F(Ve^{i\lambda}-V) = F(Ve^{i(\lambda+2\pi)}-V)$. Hence $e^{i2\pi Q} = 1$ and $Q = 0, \pm 1, \pm 2, \cdots$. If $F(z)$ is to be holomorphic for $|z+V|<V$ we must have $Q = 0,1,2,\ldots$; hence it follows that $F(z)$ is of the form

$$F(z) = \left(\frac{z+V}{V}\right)^Q, \quad Q = 0,1,2,\cdots. \tag{3.29}$$

Therefore,

$$L(f) = \left(\int_V e^{if(\mathbf{x})}d^3x\right)^Q. \tag{3.30}$$

Every representation of the current algebra can be expressed as a direct integral of factor representations; thus for an arbitrary representation of the current algebra in a box with the ground state satisfying Eq. (3.1) or (3.3), $L(f)$ is of the form

$$L(f) = \sum_{Q=0}^{\infty} \mu_Q \left(\int_V e^{if(\mathbf{x})}d^3x\right)^Q, \tag{3.31}$$

with $Q = 0,1,2,\cdots$, $\mu_Q\geq 0$, and $\sum_{Q=0}^{\infty}\mu_Q = 1$.

Next we determine the form of $L(f)$ for the case of infinite volume. Again we assume that we have a factor representation, and obtain the general case by taking a direct integral of factor representations.

Consider the expression

$$L(f_1 + f_2^n) = (\Omega_0, e^{i\rho(f_1)}e^{i\rho(f_2^n)}\Omega_0), \tag{3.32}$$

where $f_1\in\mathcal{S}$; $f_2^n(\mathbf{x}) = f_2(\mathbf{x}-n\mathbf{a})$ for $n = 0,1,2,\cdots$ and $f_2\in\mathcal{S}$; and where $\mathbf{a}$ is a vector of unit length.

Now we have

$$L(f_1 + f_2^n) = F(K(f_1 + f_2^n))$$

and

$$K(f_1 + f_2^n) = \int (e^{if_1(\mathbf{x})}e^{if_2(\mathbf{x}-n\mathbf{a})} - 1)d^3x \to K(f_1) + K(f_2) \quad \text{as} \quad n\to\infty. \tag{3.33}$$

Since F is continuous we have

$$L(f_1 + f_2^n) \to F(K(f_1) + K(f_2)) \quad \text{as} \quad n\to\infty.$$

Now the set of all operators of norm not greater than one in a Hilbert space (i.e., the unit ball) is compact in the weak operator topology. Therefore, the sequence $e^{i\rho(f_2^n)}$ has at least one cluster point in the weak operator topology which we shall call G. Since $[e^{iJ(g)}, e^{i\rho(f_2^n)}] \to 0$ strongly as $n\to\infty$ (i.e., the $e^{i\rho(f_2^n)}$ tend to commute with elements of the current algebra as $n\to\infty$), it follows that G is in the commutant of the current algebra. Since G is a cluster point of a sequence of elements of the current algebra, G is also in the weak closure of the current algebra. Hence, by the assumption of a factor representation G is a multiple of the identity, i.e., $G = \lambda I$.

Since we have the existence of the limit

$$L(f_1 + f_2^n) \to F(K(f_1) + K(f_2)) \quad \text{as} \quad n\to\infty$$

and since

$$L(f_2^n) = F(K(f_2)) \quad \text{for all } n,$$

it follows from Eq. (3.32) that

$$L(f_1 + f_2^n) \to \lambda L(f_1) \quad \text{as } n\to\infty$$

and

$$L(f_2^n) = \lambda.$$

Combining these equations, we find

$$F(K(f_1) + K(f_2)) = F(K(f_1))F(K(f_2)) \tag{3.34}$$

for all $f_1, f_2\in\mathcal{S}$. Hence,

$$F(z_1 + z_2) = F(z_1)F(z_2), \tag{3.35}$$

and it follows that $F(z)$ is of the form $F(z) = A\exp\{\bar{\rho}z\}$.

Since $F(0) = 1$ we have $A = 1$ and since $|L(f)| = |F(K(f))| \leq 1$ for all $f \in \mathcal{S}$, we have $|F(z)| \leq 1$ for all z with $\mathrm{Re}\{z\} \leq 0$. Hence $\bar{\rho} \geq 0$ and $F(z) = \exp\{\bar{\rho} z\}$. Thus

$$L(f) = \exp \bar{\rho} \int (e^{if(\mathbf{x})} - 1) d^3x. \tag{3.36}$$

For the general case of a representation with ground state satisfying Eq. (3.1) or (3.3), $L(f)$ is a direct integral of functionals of the above form, i.e.,

$$L(f) = \int_0^\infty \exp[\bar{\rho} \int (e^{if(\mathbf{x})} - 1) d^3x] d\mu(\bar{\rho}), \tag{3.37}$$

where μ is a positive measure on $[0, \infty)$ normalized so that $\int_0^\infty d\mu(\bar{\rho}) = 1$.

We summarize these results as follows:

Theorem 7: Suppose $\mathbf{g} \to J(\mathbf{g})$ and $f \to \rho(f)$ is a *-representation of the nonrelativistic current algebra of Eqs. (2.11)–(2.13) with a cyclic vector Ω_0. Suppose Ω_0 is in the domain of $\rho(f)$ and $J(\mathbf{g})$ for all $f, \mathbf{g} \in \mathcal{S}$, and that $(\rho(f_1)\Omega_0, \rho(f_2)\Omega_0)$ is a continuous bilinear form on $\mathcal{S} \times \mathcal{S}$ (resp. $\mathcal{S}_V \times \mathcal{S}_V$). Finally, suppose that

$$2iJ(\mathbf{g})\Omega_0 = \rho(\nabla \cdot \mathbf{g})\Omega_0 \tag{3.3}$$

for all $\mathbf{g}$ with components in $\mathcal{S}$ (resp. $\mathcal{S}_V$). Then if $L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0)$, $L(f)$ is of the form

$$L(f) = \sum_{Q=0}^{\infty} \mu_Q \left(\int_V e^{if(\mathbf{x})} d^3x \right)^Q \tag{3.31}$$

in a box of volume V, and of the form

$$L(f) = \int_0^\infty \exp[\bar{\rho} \int (e^{if(\mathbf{x})} - 1) d^3x] d\mu(\bar{\rho}) \tag{3.37}$$

in the infinite volume case; where $\mu_Q \geq 0$ for $Q = 0, 1, 2, \cdots, \sum_{Q=0}^\infty \mu_Q = 1$ and where μ is a positive measure on $[0, \infty)$ normalized to unity.

We remark that the form of $L(f, \mathbf{g}) = (\Omega_0, e^{i\rho(f)} e^{iJ(\mathbf{g})}\Omega_0)$ is completely determined by the form of $L(f)$ together with Eq. (3.3).

Theorem 8: Representations corresponding to Eq. (3.30) and Eq. (3.36) respectively are *irreducible*.

Proof: Supposing the contrary, there exists a closed invariant subspace $\mathfrak{M}$ of $\mathfrak{H}$ with $U(f)\mathfrak{M} \subseteq \mathfrak{M}$, $V(\psi)\mathfrak{M} \subseteq \mathfrak{M}$; and we can decompose Ω_0 into $\lambda^{1/2}\Omega_1 + (1-\lambda)^{1/2}\Omega_2$ with $\Omega_1 \in \mathfrak{M}, \Omega_2 \in \mathfrak{M}^\perp$, $0 < \lambda < 1$. Then $U(f)\Omega_2$ and $V(\psi)\Omega_2$ are likewise in $\mathfrak{M}^\perp$. Since Ω_0 is a cyclic vector for the $U(f)$, it follows that $\{U(f)\Omega_1\}$ generates a dense subspace of $\mathfrak{M}$, and $\{U(f)\Omega_2\}$ a dense subspace of $\mathfrak{M}^\perp$. Furthermore, Ω_1 and Ω_2 are in the domains of $\rho(f)$ and $J(\mathbf{g})$ by Stone's theorem, with $\rho(f)\Omega_1 \in \mathfrak{M}$, $\rho(f)\Omega_2 \in \mathfrak{M}^\perp$, etc. Since $\|\rho(f_n)\Omega_0\|^2 \to 0$ if $f_n \to 0$ as $n \to \infty$, $\|\rho(f_n)\Omega_1\|^2 \to 0$ and $\|\rho(f_n)\Omega_2\|^2 \to 0$, whence $(\rho(f_1)\Omega_1, \rho(f_2)\Omega_1)$ and $(\rho(f_1)\Omega_2, \rho(f_2)\Omega_2)$ are continuous in f_1 and f_2. Evidently, $2iJ(\mathbf{g})\Omega_1 = \rho(\nabla \cdot \mathbf{g})\Omega_1$ and similarly for Ω_2.

Thus the functionals $L_1(f) = (\Omega_1, U(f)\Omega_1)$ and $L_2(f) = (\Omega_2, U(f)\Omega_2)$ satisfy all of the assumptions made earlier in this section, with $L(f) = \lambda L_1(f) + (1-\lambda) L_2(f)$. Consequently, L_1 and L_2 must each be of the form of Eq. (3.31) or (3.37), which is impossible unless $L_1 = L_2 = L$. Therefore the representations are irreducible. QED

4. FUNCTIONAL CALCULUS

We have shown that in order to describe a free Bose gas at zero temperature, one takes a representation of the current algebra, Eqs. (2.11)–(2.13), with a cyclic vector Ω_0 satisfying Eq. (3.1). This corresponds to making the assumption that the expectation functional $L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0)$ satisfies a certain functional differential equation.[19] In fact, with

$$\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + 2i\mathbf{J}(\mathbf{x}) \tag{4.1}$$

and

$$(\Omega_0, e^{i\rho(f)}\mathbf{K}(\mathbf{x})\Omega_0) = 0, \tag{4.2}$$

together with the commutation relation

$$[e^{i\rho(f)}, \mathbf{K}(\mathbf{x})] = -2i\nabla f(\mathbf{x})\rho(\mathbf{x})e^{i\rho(f)}, \tag{4.3}$$

one easily obtains

$$-i\nabla f(\mathbf{x})(\Omega_0, e^{i\rho(f)}\rho(\mathbf{x})\Omega_0) + (\Omega_0, \nabla\rho(\mathbf{x})e^{i\rho(f)}\Omega_0) = 0, \tag{4.4}$$

which is the unsmeared form of Eq. (3.9).

Equation (4.4) may be rewritten as a functional differential equation as follows. We use the standard notation for functional derivatives. If $L(f)$ is a continuous functional on Schwartz's space $\mathcal{S}$, we say L has a functional derivative at f if there is a tempered distribution $T_f(\mathbf{x})$ such that

$$\lim_{t \to 0} \frac{L(f + tg) - L(f)}{t} = T_f(g). \tag{4.5}$$

We denote the functional derivative by

$$\frac{\delta L(f)}{\delta f(\mathbf{x})} = T_f(\mathbf{x}). \tag{4.6}$$

It is a consequence of the assumptions we made on $L(f)$ in the beginning of Sec. 3 that L has a functional derivative at all $f \in \mathcal{S}$, and

$$\frac{\delta L(f)}{\delta f(\mathbf{x})} = i(\Omega, e^{i\rho(f)}\rho(\mathbf{x})\Omega). \tag{4.7}$$

Higher functional derivatives are defined in exactly the same fashion.

In this notation, Eq. (4.4) reads

$$-i[\nabla f(\mathbf{x})] \frac{\delta L(f)}{\delta f(\mathbf{x})} + \nabla_{\mathbf{x}} \frac{\delta L(f)}{\delta f(\mathbf{x})} = 0. \tag{4.8}$$

A unique solution to Eq. (4.8) is determined when the following boundary conditions are imposed on $L(f)$:

(i) $L(f)$ is a positive functional in the Bochner sense, Eq. (2.28). This condition is a consequence of the interpretation of $L(f)$ as an inner product in a Hilbert space of positive norm. It establishes that the measures μ_Q and $\mu(\bar{\rho})$ appearing in Eqs. (3.31) and (3.37) are positive.

(ii) $L(0) = 1$. This condition normalizes the inner product to one.

(iii) $|L(f)| \leq 1$. This is a consequence of the unitarity of $U(f)$, Eq. (2.16). This condition guarantees that the average density $\bar{\rho}$ appearing in Eqs. (3.36) and (3.37) is a positive number.

(iv) $L(f)$ is an extremal solution in the sense that it cannot be written as a convex linear combination of two other solutions.

(v) $$\left.\frac{\delta L}{\delta f(\mathbf{x})}\right|_{f=0} = \bar{\rho} = \text{a specified number.}$$

Conditions (i)–(iii) were employed to prove Theorem 7, which implies that Eq. (4.8) has the unique class of solutions specified by Eqs. (3.31) or (3.37). Condition

(iv), as demonstrated in Theorem 8), is used to restrict the general solution to the forms (3.30) and (3.36) defining irreducible representations of the algebra (2.11)–(2.13). Finally, condition (v) selects the particular irreducible representation corresponding to a physical system having a specified value for the average density.

Having written a functional differential equation for $L(f)$ whose solutions describe the infinite free Bose gas at a specified average density, it is natural to seek an equation or system of coupled equations whose solution would describe an infinite Bose gas with an interaction. Such a system is proposed in Sec. 5.

In this section we derive a mathematical relation between two of the quantities which appear in Sec. 5. This relation proves helpful in completing the system of coupled equations, and introduces the concept of indefinite functional integration.

Define

$$R_{ij}(f,\mathbf{x},\mathbf{y}) = (\Omega_0, K_i^*(\mathbf{x})\, e^{i\rho(f)} K_j(\mathbf{y})\,\Omega_0), \tag{4.9}$$

where $\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + 2i\mathbf{J}(\mathbf{x})$ is an operator-valued distribution, and Ω_0 is a cyclic vector for the $\rho(f)$.

Consider the expression

$$\begin{aligned} N'_{ij}(f,\mathbf{x}) &= \left(\Omega_0, K_i^*(\mathbf{x})\, e^{i\rho(f)}\, \frac{1}{\rho(\mathbf{x})}\, K_j(\mathbf{x})\,\Omega_0\right) \\ &= \left(\Omega_0, e^{i\rho(f)} K_i^*(\mathbf{x})\, \frac{1}{\rho(\mathbf{x})}\, K_j(\mathbf{x})\,\Omega_0\right) \\ &\quad - 2i(\partial_i f)(\mathbf{x})\,(\Omega_0, e^{i\rho(f)} K_j(\mathbf{x})\,\Omega_0). \end{aligned} \tag{4.10}$$

One way to define such an expression has previously been proposed.[11] Here we shall, roughly speaking, functionally integrate R_{ij} to obtain N'_{ij}.

In representations of nonrelativistic systems of physical interest one usually has that $\rho(f) \geq 0$ if $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$. This corresponds to the fact that ρ usually describes the number density for a single species of particle. Let us for the moment pretend that $\rho(\mathbf{x})$ is a well-defined self-adjoint operator at each point $\mathbf{x}$, with positive spectrum.

Then we could write

$$\frac{1}{\rho(\mathbf{x})}\, e^{i\rho(f)} = \int_0^\infty dt\, e^{-t\rho(\mathbf{x})}\, e^{i\rho(f)} = \int_0^\infty dt\, e^{i\rho(f+it\delta_{\mathbf{x}})}, \tag{4.11}$$

where $\delta_{\mathbf{x}}(\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y})$, the Dirac delta function. The relation between $R_{ij}(f,\mathbf{x},\mathbf{y})$ and $N'_{ij}(f,\mathbf{x})$ would be given by

$$N'_{ij}(f,\mathbf{x}) = \int_0^\infty dt\, R_{ij}(f + it\delta_{\mathbf{x}}, \mathbf{x}, \mathbf{x}). \tag{4.12}$$

Now in general $\rho(\mathbf{x})$ is not well defined as an operator at a point, and an expression such as $R_{ij}(f + it\delta_{\mathbf{x}}, \mathbf{x}, \mathbf{x})$ is not well defined. In fact, let us compute R_{ij} for an infinite free Bose gas at zero temperature of average density $\bar\rho$, with Ω_0 the ground state. We have $L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0)$ given by Eq. (3.36). Suppose that $h \in \mathcal{S}$ and $\Omega_h = e^{i\rho(h)}\Omega_0$. We shall compute $R^{(h)}_{ij}(f,\mathbf{x},\mathbf{y})$ with respect to the cyclic vector Ω_h:

$$\begin{aligned} R^{(h)}_{ij}(f,\mathbf{x},\mathbf{y}) &= (\Omega_h, K_i^*(\mathbf{x})\, e^{i\rho(f)} K_j(\mathbf{y})\,\Omega_h) \\ &= (K_i(\mathbf{x})\, e^{i\rho(h)}\Omega_0, e^{i\rho(f)} K_j(\mathbf{y})\, e^{i\rho(h)}\Omega_0). \end{aligned} \tag{4.13}$$

From Eq. (4.3) together with $K_i(\mathbf{x})\Omega_0 = 0$, we have

$$K_i(\mathbf{x})\, e^{i\rho(h)}\Omega_0 = 2ie^{i\rho(h)}(\partial_i h)(\mathbf{x})\rho(\mathbf{x})\Omega_0 \tag{4.14}$$

or

$$K_i(\mathbf{x})\Omega_h = 2i(\partial_i h)(\mathbf{x})\rho(\mathbf{x})\Omega_h. \tag{4.15}$$

Hence,

$$\begin{aligned} R^{(h)}_{ij}(f,\mathbf{x},\mathbf{y}) &= 4(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{y})(\Omega_0, e^{i\rho(f)}\rho(\mathbf{x})\rho(\mathbf{y})\Omega_0) \\ &= -4(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{y})\frac{\delta^2 L(f)}{\delta f(\mathbf{x})\,\delta f(\mathbf{y})}. \end{aligned} \tag{4.16}$$

A straightforward computation of the functional derivatives yields

$$\begin{aligned} &R^{(h)}_{ij}(f,\mathbf{x},\mathbf{y}) \\ &\quad = 4(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{y})\,[\bar\rho^2\, e^{if(\mathbf{x})} e^{if(\mathbf{y})} \\ &\qquad + \bar\rho\delta(\mathbf{x}-\mathbf{y})\, e^{if(\mathbf{x})}]\, L(f). \end{aligned} \tag{4.17}$$

It is clear that $R^{(h)}_{ij}(f + it\,\delta_{\mathbf{x}}, \mathbf{x}, \mathbf{x})$ is ill defined, since it contains exponentials of delta functions, as well as a delta function evaluated at zero.

Instead, we propose to interpret Eq. (4.11) as follows. Let $\delta^n_{\mathbf{x}}$ be a sequence of functions in $\mathcal{S}$ which converges to $\delta_{\mathbf{x}}$ in the sense of a distribution, i.e., for all $f \in \mathcal{S}$,

$$\lim_{n\to\infty} \int \delta^n_{\mathbf{x}}(\mathbf{y}) f(\mathbf{y})\, d^3y = f(\mathbf{x}).$$

We now interpret Eq. (4.12) by means of the limiting procedure

$$N'_{ij}(f,\mathbf{x}) = \lim_{n\to\infty} \int_0^\infty dt\, R_{ij}(f + it\,\delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}). \tag{4.18}$$

We shall show that this definition works for the case of $R^{(h)}_{ij}$ in Eq. (4.17). Notice that $R^{(h)}_{ij}(f,g,k)$ can be extended from functions $f,g,k \in \mathcal{S}$ to bounded Borel functions which decrease faster than any polynomial in $\mathbf{x}$ at infinity. Then we can take our approximating sequence for a delta function to be the more convenient set of functions

$$\delta^n_{\mathbf{x}}(\mathbf{y}) = \begin{cases} n^3 & \text{if } |x_i - y_i| \leq 1/2n \text{ for } i = 1,2,3 \\ 0 & \text{otherwise.} \end{cases} \tag{4.19}$$

Computing $R^{(h)}_{ij}(f + it\delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}})$, we find

$$\begin{aligned} &R^{(h)}_{ij}(f + it\delta_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}) \\ &\qquad = A_1(t,n)\, L(f + it\delta^n_{\mathbf{x}}) + A_2(t,n)\, L(f + it\delta^n_{\mathbf{x}}), \end{aligned}$$

where

$$\begin{aligned} A_1(t,n) = 4\bar\rho^2 \int_{V_n} d^3y \int_{V_n} d^3z\, n^6 (\partial_i h)(\mathbf{y})(\partial_j h)(\mathbf{z}) \\ \times e^{if(\mathbf{y})} e^{if(\mathbf{z})} e^{-2tn^3} \end{aligned}$$

and

$$A_2(t,n) = 4\bar\rho \int_{V_n} d^3y\, n^6 (\partial_i h)(\mathbf{y})(\partial_j h)(\mathbf{y})\, e^{if(\mathbf{y})} e^{-tn^3},$$

and where $V_n = \{\mathbf{y};\ |y_i - x_i| \leq 1/2n \text{ for } i = 1,2,3\}$. A straightforward computation shows that

$$|L(f + it\delta^n_{\mathbf{x}}) - L(f)| \leq 1 - e^{-\bar\rho/n^3}$$

for all $\mathbf{x} \in \mathbb{R}^3$, $t \geq 0$ and $f \in \mathcal{S}$. Hence, $L(f + it\delta^n_{\mathbf{x}})$ converges uniformly to $L(f)$ as $n \to \infty$. Since $A_1(t,n)$ and $A_2(t,n)$ converge in the L^1 topology to absolutely integrable functions, it follows that

$$\begin{aligned} &\lim_{n\to\infty} \int_0^\infty dt\, R^{(h)}_{ij}(f + it\delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}) \\ &\qquad = (\lim_{n\to\infty} \int_0^\infty [A_1(t,n) + A_2(t,n)]\, dt)\, L(f). \end{aligned} \tag{4.20}$$

Now performing the integration over t, we have

$$\int_0^\infty A_1(t,n)\,dt = 2\bar{\rho}\; n^3 \int_{V_n} d_y^3 \int_{V_n} d^3 z\,(\partial_i h)(\mathbf{y})(\partial_j h)(\mathbf{z})\, e^{if(\mathbf{y})} e^{if(\mathbf{z})} \to 0 \tag{4.21}$$

as $n \to \infty$, since the square of the volume of V_n goes as n^{-6}. Furthermore,

$$\int_0^\infty A_2(t,n)\,dt = 4\bar{\rho} n^3 \int_{V_n} (\partial_i h)(\mathbf{y})(\partial_j h)(\mathbf{y})\, e^{if(\mathbf{y})} d^3 y \to 4\bar{\rho}(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{x})\, e^{if(\mathbf{x})}. \tag{4.22}$$

Hence we have

$$\lim_{n\to\infty} \int_0^\infty dt R_{ij}^{(h)}(f + it\delta_{\mathbf{x}}^n, \delta_{\mathbf{x}}^n, \delta_{\mathbf{x}}^n) = 4\bar{\rho}(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{y})\, e^{if(\mathbf{x})} L(f). \tag{4.23}$$

Next we compute $N_{ij}'^{(h)}(f,\mathbf{x})$ directly, using the interpretation of $1/\rho(\mathbf{x})$ proposed earlier by Goldin and Sharp.[11] They interpret $1/\rho(\mathbf{x})$ as the map $1/\rho(\mathbf{x})$: $\mathcal{U} \times \mathcal{U} \to \mathcal{S}'$, where $\mathcal{S}'$ is the continuous dual of $\mathcal{S}$ and $\mathcal{U}$ is the linear span of the vector-valued distributions $\{g(\mathbf{x})\rho(\mathbf{x})\Phi \mid \Phi \in D, g \in \mathcal{O}_M\}$, with D a dense invariant domain for $\rho(f)$, $f \in \mathcal{S}$, and $\mathcal{O}_M$ the real-valued C_∞ functions which together with all derivatives are polynomially bounded at ∞.

Now, using Eq. (4.15)

$$\left(\Omega_h, e^{i\rho(f)} K_i^*(\mathbf{x}) \frac{1}{\rho(\mathbf{x})} K_j(\mathbf{x})\Omega_h\right) = \left(K_i(\mathbf{x})\Omega_{h-f}, \frac{1}{\rho(\mathbf{x})} K_j(\mathbf{x})\Omega\right)$$
$$= 4\left((\partial_i h - \partial_i f)(\mathbf{x})\rho(\mathbf{x})\Omega_{h-f}, \frac{1}{\rho(\mathbf{x})}(\partial_j h)(\mathbf{x})\rho(\mathbf{x})\Omega_h\right)$$
$$= 4(\partial_i h - \partial_i f)(\mathbf{x})(\partial_j h)(\mathbf{x})(\Omega_0, e^{i\rho(f)}\rho(\mathbf{x})\Omega_0), \tag{4.24}$$

and

$$N_{ij}^{(h)'}(f,\mathbf{x}) = \left(\Omega_h, e^{i\rho(f)} K_i^*(\mathbf{x}) \frac{1}{\rho(\mathbf{x})} K_j(\mathbf{x})\Omega_h\right) - 2i(\partial_i f)(\mathbf{x})(\Omega_h, e^{i\rho(f)} K_j(\mathbf{x})\Omega_h)$$
$$= 4(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{x})(\Omega_0, e^{i\rho(f)}\rho(\mathbf{x})\Omega_0)$$
$$= -4i(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{x}) \frac{\delta L(f)}{\delta f(\mathbf{x})}$$
$$= 4\bar{\rho}(\partial_i h)(\mathbf{x})(\partial_j h)(\mathbf{x})\, e^{if(\mathbf{x})} L(f), \tag{4.25}$$

where $L(f)$ is given by Eq. (3.36) in evaluating the functional derivative. Hence we see that Eq. (4.18) gives the correct relation between R_{ij} and N_{ij}' for the case at hand. We leave unanswered at this time the important problem of determining a general set of sufficient conditions to be imposed on $R_{ij}(f,\mathbf{x},\mathbf{y})$, in order to ensure that the limiting procedure of Eq. (4.18) leads to a well-defined expression.

5. DETERMINING $L(f)$ WHEN THE PARTICLES INTERACT

The preceding work has shown how the functional $L(f)$ of Eq. (3.36), which determines an irreducible representation of the local current algebra (2.11)–(2.13), can be defined uniquely as the solution to a functional differential equation satisfying the appropriate boundary conditions [conditions (i)–(v) following Eq. (4.8)].

These results apply only to noninteracting bosons. Next we ask whether the same pattern of results persists when interactions are included. Can one find a set of functional differential equations which, when supplemented with suitable boundary conditions, determine a ground state expectation functional $L(f)$? In this section we suggest the possibility of an affirmative answer to this question.

The functional equation which defined $L(f)$ for noninteracting bosons was Eq. (4.8), obtained from the condition (3.1),

$$\mathbf{K}(\mathbf{x})\Omega_0 = (\nabla\rho + 2i\mathbf{J})(\mathbf{x})\Omega_0 = 0.$$

The first step towards deriving a corresponding set of functional equations in the interacting case is to find conditions replacing Eq. (3.1), since the latter correctly expresses the action of the Hamiltonian in a representation only for noninteracting bosons. These conditions take the form of equations which relate, and ultimately determine, the following quantities:

$$L(f) = (\Omega_0, e^{i\rho(f)}\Omega_0), \tag{5.1}$$

$$\mathbf{M}(f,\mathbf{x}) = (\Omega_0, e^{i\rho(f)}\mathbf{K}(\mathbf{x})\Omega_0), \tag{5.2}$$

$$R_{ij}(f,\mathbf{x},\mathbf{y}) = (\Omega_0, K_i^*(\mathbf{x})\, e^{i\rho(f)} K_j(\mathbf{y})\Omega_0), \tag{5.3}$$

$$N_{ij}(f,\mathbf{x}) = \tfrac{1}{2}\left(\Omega_0, e^{i\rho(f)}\left[K_i^*(\mathbf{x})\frac{1}{\rho(\mathbf{x})} K_j(\mathbf{x}) + K_j^*(\mathbf{x})\frac{1}{\rho(\mathbf{x})} K_i(\mathbf{x})\right]\Omega_0\right). \tag{5.4}$$

We assume that the particles interact through a central two-body potential $V(|\mathbf{x}-\mathbf{y}|)$ and write the Hamiltonian (for particles of unit mass) as[1]

$$H = \tfrac{1}{8}\int d^3x\, K_i^*(\mathbf{x})\frac{1}{\rho(\mathbf{x})} K_i(\mathbf{x}) + \tfrac{1}{2}\int\int d^3x\, d^3y\, \rho(\mathbf{x}) V(|\mathbf{x}-\mathbf{y}|)\rho(\mathbf{y}), \tag{5.5}$$

where the repeated index i is summed over $i = 1, 2, 3$. For the Hamiltonian to be well defined it may be necessary to subtract from Eq. (5.5) an infinite constant corresponding to its ground state expectation value, thus establishing a zero of the energy, $(H - E_0)\Omega_0 = 0$.

It should also be noted that we have no guarantee that Eq. (5.4) defining $N_{ij}(f,\mathbf{x})$ makes sense as it stands. Nevertheless, there is reason to hope that the ensuing system of equations ultimately lends itself to a meaningful interpretation and we shall proceed as though the quantities under discussion are all well defined.

1. The first condition replacing (3.1) follows from the requirement that the cyclic vector be an eigenvector of the energy operator,

$$(H - E_0)\Omega = 0, \tag{5.6}$$

which we write in the form

$$(\Omega_0, e^{i\rho(f)} H\Omega_0) - E_0 L(f) = 0. \tag{5.7}$$

To write Eq. (5.7) as a relationship between functionals, we introduce $N_{ij}(f,\mathbf{x})$, Eq. (5.4), and note that

$$\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}\frac{1}{i}\frac{\delta}{\delta f(\mathbf{y})} L(f) = (\Omega_0, \rho(\mathbf{x})\rho(\mathbf{y})\, e^{i\rho(f)}\Omega_0) \tag{5.8}$$

to obtain

$$\frac{1}{8}\sum_{i=1}^{3}\int d^3x\, N_{ii}(f,\mathbf{x}) - \tfrac{1}{2}\int\int d^3x\, d^3y\, V(|\mathbf{x}-\mathbf{y}|)\frac{\delta^2 L(f)}{\delta f(\mathbf{x})\,\delta f(\mathbf{y})} - E_0 L(f) = 0. \tag{5.9}$$

2. A second equation follows from the requirement that Ω_0 be invariant under time reversal,

$$T\Omega_0 = \Omega_0; \tag{5.10}$$

where T is the antiunitary time reversal operator satisfying $(T\Psi, T\Phi) = (\Phi, \Psi)$. To derive the desired equation from (5.10), consider

$$(\Omega_0, e^{i\rho(f)}\mathbf{J}(\mathbf{x})\Omega_0) = (Te^{i\rho(f)}\mathbf{J}(\mathbf{x})\Omega_0, T\Omega_0). \tag{5.11}$$

Since $T\rho(f)T^{-1} = \rho(f)$, $T\mathbf{J}(\mathbf{x})^{T^{-1}} = -\mathbf{J}(\mathbf{x})$, $TiT^{-1} = -i$, and $T\Omega_0 = \Omega_0$, we find

$$(\Omega_0, e^{i\rho(f)}\mathbf{J}(\mathbf{x})\Omega_0) = -(\Omega_0, \mathbf{J}(\mathbf{x})e^{i\rho(f)}\Omega_0). \tag{5.12}$$

Recalling that $e^{i\rho(f)}\mathbf{J}(\mathbf{x})e^{-i\rho(f)} = \mathbf{J}(\mathbf{x}) - \nabla f(\mathbf{x})\rho(\mathbf{x})$, Eq. (3.7), we may write Eq. (5.12) in the form

$$(\Omega_0, e^{i\rho(f)}\mathbf{K}(\mathbf{x})\Omega_0) = (\Omega_0, e^{i\rho(f)}\nabla\rho(\mathbf{x})\Omega_0) - i(\Omega_0, e^{i\rho(f)}\rho(\mathbf{x})\nabla f(\mathbf{x})\Omega_0). \tag{5.13}$$

Finally, we may introduce $\mathbf{M}(f, \mathbf{x})$, Eq. (5.2), and the appropriate functional derivatives, to find

$$\mathbf{M}(f, \mathbf{x}) = \nabla_{\mathbf{x}}\left(\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})}\right) - [\nabla f(\mathbf{x})]\frac{\delta L(f)}{\delta f(\mathbf{x})}. \tag{5.14}$$

All of the dynamical information about a system of interacting bosons is expressed in Eqs. (5.9) and (5.14), when these equations are supplemented with suitable boundary conditions. However, the two equations relate three unknown functionals. The additional relationships among the functions (5.1)–(5.4) needed to complete the system of equations are obtained entirely from consideration of the mathematical properties of the functionals.

3. One of the remaining equations we need has been derived in Sec. 4. It relates the indefinite functional integral of $R_{ij}(f, \mathbf{x}, \mathbf{y})$, Eq. (5.3), to the quantity $N'_{ij}(f, \mathbf{x})$, Eq. (4.10). We can write $N_{ij}(f, \mathbf{x})$, Eq. (5.4), in terms of $R_{ij}(f, \mathbf{x}, \mathbf{y})$ and $\mathbf{M}(f, \mathbf{x})$ as follows:

$$N_{ij}(f, \mathbf{x}) = \tfrac{1}{2}[N'_{ij}(f, \mathbf{x}) + N'_{ji}(f, \mathbf{x})] + i(\partial_i f)(\mathbf{x})M_j(f, \mathbf{x}) + i(\partial_j f)(\mathbf{x})M_i(f, \mathbf{x}), \tag{5.15}$$

where as in Eq. (4.18)

$$N'_{ij}(f, \mathbf{x}) = \lim_{n\to\infty}\int_0^\infty R_{ij}(f + it\delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}})dt. \tag{5.16}$$

4. The final equation relates $R_{ij}(f, \mathbf{x}, \mathbf{y})$ to $\mathbf{M}(f, \mathbf{x})$. Referring to Sec. 2C, we may write $L(f)$, $\mathbf{M}(f, \mathbf{x})$, and $R_{ij}(f, \mathbf{x}, \mathbf{y})$ as

$$L(f) = \int_{\mathcal{S}'} e^{i(F,f)}d\mu(F), \tag{5.17}$$

$$\mathbf{M}(f, \mathbf{x}) = \int_{\mathcal{S}'} \overline{\Omega_0(F)}e^{i(F,f)}(\mathbf{K}(\mathbf{x})\Omega_0)(F)\,d\mu(F), \tag{5.18}$$

and

$$R_{ij}(f, \mathbf{x}, \mathbf{y}) = \int_{\mathcal{S}'} \overline{(K_i(\mathbf{x})\Omega_0)(F)}e^{i(F,f)}[K_j(\mathbf{y})\Omega_0](F)\,d\mu(F), \tag{5.19}$$

where $\mathcal{S}'$ is the continuous dual of Schwartz's space and μ is a cylindrical measure on $\mathcal{S}'$ uniquely determined by $L(f)$.

We now define $\tilde{\mathbf{M}}(F, \mathbf{x})$ to be the inverse Fourier transform of $\mathbf{M}(f, \mathbf{x})$; i.e.,

$$\mathbf{M}(f, \mathbf{x}) = \int_{\mathcal{S}'} e^{i(F,f)}\tilde{\mathbf{M}}(F, \mathbf{x})\,d\mu(F). \tag{5.20}$$

It is not difficult to establish the existence of $\tilde{\mathbf{M}}(F, \mathbf{x})$ using standard methods in the Gel'fand–Vilenkin approach.[3,12] One may prove first that $\mathbf{M}(f, \mathbf{x})$ in Eq. (5.2) is the Fourier transform of a (not necessarily positive) measure μ_1 on $\mathcal{S}'$; then that every set of measure zero in μ is of measure zero in μ_1; and conclude that the Radon–Nikodym derivative $d\mu_1(F)/d\mu(F)$ exists and defines $\tilde{\mathbf{M}}(F, \mathbf{x})$. The assumptions needed to carry through these arguments amount to the statement that the ground state vector Ω_0 is in the domain of $\mathbf{K}(\mathbf{x})$.

Similarly, define the inverse Fourier transform $\tilde{R}_{ij}(F, \mathbf{x}, \mathbf{y})$ of $R_{ij}(f, \mathbf{x}, \mathbf{y})$ by

$$R_{ij}(f, \mathbf{x}, \mathbf{y}) = \int_{\mathcal{S}'} e^{i(F,f)}\tilde{R}_{ij}(F, \mathbf{x}, \mathbf{y})\,d\mu(F). \tag{5.21}$$

Then, since $\Omega_0(F) \equiv 1$ almost everywhere,

$$\tilde{R}_{ij}(F, \mathbf{x}, \mathbf{y}) = \overline{\tilde{M}_i(F, \mathbf{x})}\tilde{M}_j(F, \mathbf{y}) \tag{5.22}$$

almost everywhere, or

$$R_{ij}(f, \mathbf{x}, \mathbf{y}) = \int_{\mathcal{S}'} e^{i(F,f)}\overline{\tilde{M}_i(F, \mathbf{x})}\tilde{M}_j(F, \mathbf{y})\,d\mu(F). \tag{5.23}$$

To summarize, we have the following system of coupled functional equations:

1. "Schrödinger equation" $[(H - E_0)\Omega_0 = 0]$

$$\frac{1}{8}\sum_{i=1}^{3}\int d^3x\, N_{ii}(f, \mathbf{x}) - \tfrac{1}{2}\iint d^3x\, d^3y\, V(|\mathbf{x} - \mathbf{y}|)\frac{\delta^2 L(f)}{\delta f(\mathbf{x})\,\delta f(\mathbf{y})} - E_0 L(f) = 0. \tag{5.9}$$

2. Time-reversal invariance $(T\Omega_0 = \Omega_0)$

$$\mathbf{M}(f, \mathbf{x}) = \nabla_{\mathbf{x}}\left(\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})}\right) - [\nabla f(\mathbf{x})]\frac{\delta L(f)}{\delta f(\mathbf{x})}. \tag{5.14}$$

3. Indefinite functional integration relationship

$$N_{ij}(f, \mathbf{x}) = \tfrac{1}{2}[N'_{ij}(f, \mathbf{x}) + N'_{ji}(f, \mathbf{x})] + i(\partial_i f)(\mathbf{x})M_j(f, \mathbf{x}) + i(\partial_j f)(\mathbf{x})M_i(f, \mathbf{x}), \tag{5.15}$$

where

$$N'_{ij}(f, \mathbf{x}) = \lim_{n\to\infty}\int_0^\infty R_{ij}(f + it\delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}}, \delta^n_{\mathbf{x}})dt.$$

4. Fourier transform relationship

$$R_{ij}(f, \mathbf{x}, \mathbf{y}) = \int_{\mathcal{S}'} e^{i(F,f)}\overline{\tilde{M}_i(F, \mathbf{x})}\tilde{M}_j(F, \mathbf{y})\,d\mu(F), \tag{5.23}$$

where $L(f) = \int_{\mathcal{S}'} e^{i(F,f)}d\mu(F)$ and $\mathbf{M}(f, \mathbf{x}) = \int_{\mathcal{S}'} e^{i(F,f)}\tilde{\mathbf{M}}(F, \mathbf{x})\,d\mu(F)$.

Equations (5.15) and (5.23) together express $N_{ij}(f, \mathbf{x})$ in terms of $\mathbf{M}(f, \mathbf{x})$ and the measure μ of which $L(f)$ is the Fourier transform. Then, substituting for $N_{ii}(f, \mathbf{x})$, Eqs. (5.9) and (5.14) relate the two functionals $L(f)$ and $\mathbf{M}(f, \mathbf{x})$.

The above system of equations can be expected to determine $L(f)$ uniquely only if it is supplemented by appropriate boundary conditions, just as in the case of Eq. (4.8) which defined the free system. The boundary conditions which applied to $L(f)$ in the free case clearly apply in the interacting case as well. We do not know at this point whether these five boundary conditions suffice to determine a unique solution to Eqs. (5.9), (5.14), (5.15), and (5.23) or whether additional boundary conditions are necessary.

In contrast to the non-interacting case discussed in Secs. 3 and 4, we have no means of obtaining a solution to the above system of equations, nor do we have techniques to demonstrate that a solution exists or, if it exists, that it is unique.

There are other ways to supply some of the additional information needed to complete the system of equations

begun with (5.9) and (5.14). For instance one can use the equation of motion for $\mathbf{J}(\mathbf{x})$ in the form

$$(\Omega_0, [e^{i\rho(f)}\mathbf{J}(\mathbf{x}), H]\Omega_0) = 0 \tag{5.24}$$

and the condition that the ground state be rotationally invariant

$$\mathcal{L}\Omega_0 = \Omega_0, \quad \mathcal{L} = \int \mathbf{x} \times \mathbf{J}(\mathbf{x}) d^3x, \tag{5.25}$$

in the form

$$(\Omega_0, [e^{i\rho(f)}N_{ij}(f, \mathbf{x}), \mathcal{L}]\Omega_0) = 0. \tag{5.26}$$

Thus one obtains equations[19] which can be solved so as to express $N_{ij}(f, \mathbf{x})$ in terms of $R_{ij}(f, \mathbf{x}, \mathbf{y})$, $\mathbf{M}(f, \mathbf{x})$, and $L(f)$; and Eqs. (5.24) and (5.26) can replace Eq. (5.15).

In whatever fashion one chooses to complete Eqs. (5.9) and (5.14), one can be sure that the resulting set of equations will not be amenable to exact solution for $L(f)$ in most situations of practical interest. Therefore, one would like to have techniques for its approximate determination. The approach via functional differential equations is most inviting because it is suggestive of such techniques. An approximate functional $L(f)$ would be one which was an approximate solution in some well-defined way to a system of equations whose exact solution defined an irreducible representation of a local current algebra. This is one sense in which it might have meaning to talk about an "approximate representation" of a Lie algebra of local currents.

Finally, we would like to mention that it is possible to develop systems of functional equations whose solutions determine representations of the canonical commutation relations, as has been done in Refs. 19 and 22.

APPENDIX

In this appendix we sketch a proof of Lemma 3, for $g_1, g_2 \in \mathcal{S}$. We believe that Lemma 3 is also valid for $\mathcal{S}_V$, but it appears the proof would be still more involved. Throughout this section we let $K(g)$ denote the functional K evaluated at $g \in \mathcal{S}$, where

$$K(g) = \int (e^{ig(\mathbf{x})} - 1) d^3x.$$

Lemma 3: Suppose $g_1, g_2 \in \mathcal{S}$ and $K(g_1) = K(g_2)$. Then for any two neighborhoods N_1 of g_1 and N_2 of g_2 in $\mathcal{S}$, there exist functions $h_1 \in N_1, h_2 \in N_2$ and a continuous mapping $t \to f_t$ of $[0, 1]$ into $\mathcal{S}$, differentiable in $(0, 1)$, such that $f_0 = h_1, f_1 = h_2$ and $K(f_t) =$ a constant.

Our sketch of a proof consists of a sequence of lemmas stated without proof.

Lemma A1: Suppose $g_1, g_2 \in \mathcal{S}(\mathbb{R}^3)$ with $K(g_1) = K(g_2)$, and N_1 and N_2 are neighborhoods of g_1 and g_2, respectively, in the Schwartz space topology of $\mathcal{S}$. Then there exist functions $h_1 \in N_1$ and $h_2 \in N_2$ such that $K(h_1) = K(h_2)$ and h_1 and h_2 have compact support.

Lemma A2: Suppose $t \to z(t)$ is a differentiable mapping of the closed interval $[0, 1]$ into the left half complex z plane. Suppose that $z(0) = z(1) = 0$ and that there is a $\delta > 0$ such that $|\operatorname{Re}\{z(t)\}| \geq \delta |\operatorname{Im}\{z(t)\}|$ for all $t \in [0, 1]$. Then there is a continuous mapping $t \to k_t$ of $[0, 1]$ into $\mathcal{S}$, differentiable in $(0, 1)$, such that $K(k_t) = z(t)$ for all $t \in [0, 1]$, and $k_0 = k_1 = 0$. Furthermore, the functions k_t may all be chosen to have support in a single compact region of $\mathbb{R}^3$.

Lemma A3: Suppose $h_1, h_2 \in \mathcal{S}(\mathbb{R}^3)$ and $K(h_1) = K(h_2)$. Then there is a real-valued differentiable function $s(t)$ for $t \in [0, 1]$ with $s(0) = s(1) = 1$ and $s(t) > 0$ for all $t \in [0, 1]$, such that if $g_t(\mathbf{x}) = (1 - t)h_1(s(t)\mathbf{x}) + th_2(s(t)\mathbf{x})$ then $z(t) = K(h_1) - K(g_t)$ satisfies the hypotheses of Lemma A2.

Proof of Lemma 3: Suppose $g_1, g_2 \in \mathcal{S}$ and $K(g_1) = K(g_2)$. Let N_1 and N_2 be neighborhoods of g_1 and g_2, respectively. By Lemma A1, we can choose $h_1 \in N_1$ and $h_2 \in N_2$ with $K(h_1) = K(h_2)$ and with h_1 and h_2 of compact support. By Lemma A3 there exists a differentiable function $s(t)$ on the interval $[0, 1]$, allowing us to construct

$$g_t(\mathbf{x}) = (1 - t)h_1(s(t)\mathbf{x}) + th_2(s(t)\mathbf{x}),$$

$$\text{with } z(t) = K(h_1) - K(g_t)$$

satisfying the hypotheses of Lemma A2. Since h_1 and h_2 have compact support, the g_t all have support in some compact region S. By Lemma A2, there is a continuous mapping $t \to k_t$ of $[0, 1]$ into $\mathcal{S}(\mathbb{R}^3)$, differentiable in the open interval $(0, 1)$, such that $K(k_t) = z(t) = K(h_1) - K(g_t)$.

By translating the functions k_t we can ensure that the functions k_t and g_t have disjoint supports, without changing the values of $K(k_t)$. Then let $f_t = g_t + k_t$; we have $K(f_t) = K(g_t + k_t) = K(g_t) + K(k_t) = K(h_1)$. Since $f_0 = h_1$ and $f_1 = h_2$, the lemma is proved.

* Work supported in part by U.S. Atomic Energy Commission under Contract No. AT (30-1)-2171.

† Alfred P. Sloan Foundation Fellow.

[1] R. F. Dashen and D. H. Sharp, Phys. Rev. **165**, 1857 (1968).
[2] J. Grodnik and D. H. Sharp, Phys. Rev. D **1**, 1531 (1970).
[3] G. Goldin, J. Math. Phys. **12**, 462 (1971).
[4] H. Araki and E. J. Woods, J. Math. Phys. **4**, 637 (1963).
[5] R. Haag, "Particles and Cross Sections in a Theory of Local Observables" in *Recent Developments in Particle Physics* edited by M. J. Moravcsik (Gordon and Breach, New York, 1966).
[6] A. S. Wightman, "The Problem of Existence of Solutions in Quantum Field Theory" in *Proceedings of the Fifth Annual Eastern Theoretical Physics Conference* edited by D. Feldman (W. A. Benjamin, New York, 1967).
[7] D. H. Sharp, Phys. Rev. **165**, 1867 (1968).
[8] H. Sugawara, Phys. Rev. **170**, 1659 (1968).
[9] C. Sommerfield, Phys. Rev. **176**, 2019 (1968).
[10] A. Dicke and G. Goldin, Phys. Rev. D **5**, 845 (1972).
[11] G. Goldin and D. H. Sharp, "Lie Algebras of Local Currents and their Representations" in *1969 Battelle Rencontres: Group Representations*, edited by V. Bargmann (Springer-Verlag, Berlin, 1970).
[12] I. Gel'fand and N. Vilenkin, *Generalized Functions, Application of Harmonic Analysis* (Academic, New York, 1964), Vol. 4.
[13] The authors wish to thank Professor R. Haag for helpful suggestions concerning this point.
[14] See, for example, M. A. Naimark, *Normed Rings* (P. Noordhoff N. V., Groningen, 1964).
[15] G. Goldin, Ph.D. Thesis, Princeton University (1968) (unpublished).
[16] E. Nelson, Ann. Math. **70**, 572 (1959).
[17] W. J. Pardee, L. Schlessinger, and J. Wright, Phys. Rev. **175**, 2140 (1968).
[18] J. Grodnik and D. H. Sharp, Phys. Rev. D **1**, 1546 (1970).
[19] G. Goldin and D. H. Sharp, "Functional Differential Equations Determining Representations of Local Current Algebras" in *Magic Without Magic: John Archibald Wheeler*, edited by J. R. Klauder (Freeman, San Francisco, 1972).
[20] J. Dixmier, *Les C*-algèbres et leurs représentations* (Gauthier-Villars, Paris, 1964).
[21] J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien* (Gauthier-Villars, Paris, 1957).
[22] J. Grodnik, Phys. Rev. D **3**, 2955 (1971).

Scattering of scalar waves from a random irregular interface

G. G. Zipfel Jr.*

Naval Research Laboratory, Washington, D. C. 20375
(Received 14 May 1973)

The problem of a scalar wave scattering from a random interface between two isotropic homogeneous media is solved in terms of a diagrammatic series expansion. It is assumed that the interface is single valued and that it has multivariate Gaussian statistics. Partial summation of the series in terms of linear integral equations is briefly discussed.

I. INTRODUCTION

The problem of waves scattering from a random rough surface is one of considerable practical interest. For example, the ensemble statistics for scattering from a stationary random rough surface should well approximate time-averaged scattering from a time-varying surface such as the sea surface or the time-averaged scattering from a stationary random rough surface which is being scanned. In treating this problem it is important to have a systematic procedure for evaluating the statistics of the scattered fields. Such a procedure has been developed for scalar waves scattering from a surface with Neumann boundary conditions.[1] The result is a diagrammatic expansion for the correlation functions of the scattered field analogous to those for propagation through a random medium[2] and quantum field theory.[3]

While this problem is physically realizable and of intrinsic interest, it is somewhat contrived since most surfaces occurring in nature are not impenetrable and thus represent a more complicated boundary problem. In the present paper a set of diagrammatic rules are developed for the scattering of scalar waves from a random rough interface between two media which, except for the presence of the interface, are homogeneous and isotropic. The media are assumed to have differing impedances and velocities of propagation. The rules which have been developed are not unique, but have been chosen to be the simplest commensurate with the following desirable properties:

(i) The lowest-order diagram, i.e., the Born approximation, is exact when the surface roughness vanishes.

(ii) The lowest-order diagram is exact for any degree of roughness when the properties of the two media are identical, that is, the interface vanishes.

(iii) When the velocity of propagation is zero or the impedance is infinite in the medium not containing the source, the diagrammatic expansion is well-defined and depends only on the properties of the medium containing the source. In these limits the boundary condition becomes the Dirichlet condition and the surface is impenetrable.

(iv) When the impedance of the medium not containing the source is allowed to become zero, the diagrammatic expansion is well-defined and depends only on the properties of the medium containing the source. In this limit the boundary condition becomes the Neumann condition and the surface is impenetrable.

It should be pointed out that finding one diagrammatic expansion which has all these properties is not trivial. Many alternate expansions can be found for which partial summation of the expansions must be performed in order to examine the limiting cases (ii), (iii), and (iv).

Properties (i) and (ii) are desirable because they constitute the limit of two cases where a truncated perturbation series is a useful approximation. Clearly, the diagrammatic expansion corresponds to a perturbation series in these cases only if properties (i) and (ii) are true. Properties (iii) and (iv) are desirable since scattering from an impenetrable Dirichlet or Neumann surface is considerably simpler than scattering from an interface, and this simplicity should not be hidden in an infinite series if the treatment of cases which are almost impenetrable is to be facilitated.

One interesting result that comes out of the diagrammatic expansion developed in the next section is that the lowest-order (Born) term does not correspond to the Kirchhoff[4,5] approximation except for the limits of a Dirichlet or Neumann surface, but is a simpler approximation.

In Sec. II the integral equations for scattering from a single-valued deterministic interface are developed in a form which satisfies properties (i)–(iv) and where the surface height enters only through the function $\exp\{ik_z h(\mathbf{x}_\perp)\}$. This form is very desirable since subsequent statistical averages can be expressed in terms of the characteristic functions associated with the surface height probability distributions and these characteristic functions are well-known if the surface statistics are multivariate Gaussian.

In Sec. III the diagrammatic rules for scattering from a deterministic surface are determined from the integral equations constructed in Sec. II. The diagrammatic rules for the moments of the field scattered from a random rough surface are then derived from these rules and the statistical properties of the surface. It is assumed that the surface has multivariate Gaussian statistics. Partial summation of the series for the averaged field and bilinear moment of the field is discussed in terms of linear integral equations.

Section IV contains some concluding remarks.

II. SURFACE INTEGRAL EQUATION

Figure 1 shows a schematic representation of the situation under consideration. The upper region (I) is filled with an isotropic homogeneous medium which supports scalar waves with a phase velocity of v_+ and whose impedance is Z_+. Similarly, the lower region (II) is filled with an isotropic homogeneous medium which supports scalar waves with a phase velocity of v_- and whose impedance is Z_-. These two media are joined by an irregular interface specified by

$$z = h(\mathbf{x}_\perp)$$

where $\mathbf{x}_\perp$ is a two-dimensional vector in the plane perpendicular to the z axis, and $h(\mathbf{x}_\perp)$ is assumed to be a

single-valued function with zero mean value and multivariate Gaussian statistics.

The Green's function for a harmonic source located at x′ satisfies the partial differential equations

$$\left(\nabla^2 - v_+^{-2}\frac{\partial^2}{\partial t^2}\right) G(\mathbf{x}\mid\mathbf{x}';t) = \delta_3(\mathbf{x}-\mathbf{x}')e^{-i\omega t}, \quad z > h(\mathbf{x}_\perp),$$

$$\left(\nabla^2 - v_-^{-2}\frac{\partial^2}{\partial t^2}\right) G(\mathbf{x}\mid\mathbf{x}';t) = \delta_3(\mathbf{x}-\mathbf{x}')e^{-i\omega t}, \quad z < h(\mathbf{x}_\perp),$$

plus boundary conditions at the interface and at infinity. Since the interface is assumed to be stationary, the time variation is easily separated by letting

$$G(\mathbf{x}\mid\mathbf{x}';t) = G(\mathbf{x}\mid\mathbf{x}')e^{-i\omega t}.$$

The partial differential equations that $G(\mathbf{x}\mid\mathbf{x}')$ satisfies are

$$(\nabla^2 + k_+^2)G(\mathbf{x}\mid\mathbf{x}') = \delta_3(\mathbf{x}-\mathbf{x}'), \quad z > h(\mathbf{x}_\perp), \tag{1a}$$

$$(\nabla^2 + k_-^2)G(\mathbf{x}\mid\mathbf{x}') = \delta_3(\mathbf{x}-\mathbf{x}'), \quad z < h(\mathbf{x}_\perp), \tag{1b}$$

where

$$k_+ = \omega/v_+, \quad k_- = \omega/v_-.$$

The boundary conditions on $G(\mathbf{x}\mid\mathbf{x}')$ at the interface, corresponding to the continuity of pressure and normal velocity for an acoustic field, are

$$Z_+^{-1/2}\lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_+} G(\mathbf{x}\mid\mathbf{x}') = Z_-^{-1/2}\lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_-} G(\mathbf{x}\mid\mathbf{x}),$$

$$Z_+^{1/2}\lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_+} \mathbf{n}(\mathbf{x}_\perp)\cdot\nabla G(\mathbf{x}\mid\mathbf{x}') = Z_-^{1/2}\lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_-} \mathbf{n}(\mathbf{x}_\perp)\cdot\nabla G(\mathbf{x}\mid\mathbf{x}').$$

The + or − above indicates whether the limit is from above or below the interface, respectively. The vector

$$\mathbf{x}_s(\mathbf{x}_\perp) = h(\mathbf{x}_\perp)\hat{i}_z + \mathbf{x}_\perp$$

is the position vector for the point on the surface whose x,y coordinates are given by $\mathbf{x}_\perp$. The vector $\mathbf{n}(\mathbf{x}_\perp)$ is normal to the surface at this point. At infinity the radiation boundary condition corresponding to outgoing scattered waves is assumed. The interface boundary conditions have been chosen so as to guarantee that the interface is free of sources or sinks. This was accomplished by requiring that the normal component of the conserved current associated with (1) be continuous across the interface, that is

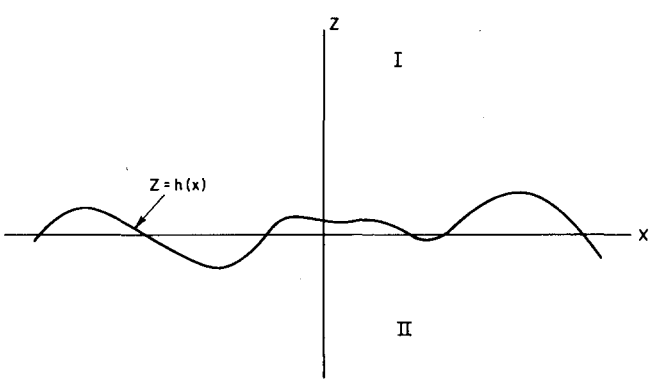

FIG. 1. A cross section through a representative interface.

$$\lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_+} \mathbf{n}(\mathbf{x}_\perp)\cdot\mathbf{j}(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_s(\mathbf{x}_\perp)_-} \mathbf{n}(\mathbf{x}_\perp)\cdot\mathbf{j}(\mathbf{x}),$$

where

$$\mathbf{j}(\mathbf{x}) = \mathrm{Im}\,G^*(\mathbf{x}\mid\mathbf{x}')\nabla G(\mathbf{x}\mid\mathbf{x}'), \quad \nabla\cdot\mathbf{j}(\mathbf{x}) = 0,$$

$\mathbf{j}(\mathbf{x})$ is proportional to the energy flux.

The problem of finding $G(\mathbf{x}\mid\mathbf{x}')$ can be reformulated in terms of an integral equation which, after some manipulation, can be written in a form that is suitable for statistical averaging.

The point of departure is to derive two integral relations for $G(\mathbf{x}\mid\mathbf{x}')$; one following from the application of Green's theorem to the upper region and the other from the application of Green's theorem to the lower region. For the upper region the integral relation[1,6,7] is

$$\begin{aligned} G_+(\mathbf{x}'\mid\mathbf{x}'')\,\theta(z'-h(\mathbf{x}'_\perp)) &= G_{0+}(\mathbf{x}'-\mathbf{x}'')\,\theta(z''-h(\mathbf{x}''_\perp)) \\ &+ \int d^2\mathbf{x}_\perp\, n_\alpha(\mathbf{x}_\perp)\left[G_{0+}(\mathbf{x}'-\mathbf{x}_s(\mathbf{x}_\perp))\,G_{+,\alpha}(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')\right. \\ &\left. + \frac{\partial}{\partial x'_\alpha}\,G_{0+}(\mathbf{x}'-\mathbf{x}_s(\mathbf{x}_\perp))\,G_+(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')\right], \end{aligned} \tag{2a}$$

where

$$G_{+,\alpha}(\mathbf{x}'\mid\mathbf{x}'') = \frac{\partial}{\partial x'_\alpha}\,G_+(\mathbf{x}'\mid\mathbf{x}'').$$

Equation (2a) follows directly from (1a). The Green's function $G_{0+}(\mathbf{x})$ is a solution of

$$(\nabla^2 + k_+^2)G_{0+}(\mathbf{x}) = \delta_3(\mathbf{x})$$

for all $\mathbf{x}$, and it satisfies the radiation boundary condition as $|\mathbf{x}| \to \infty$. Summation over repeated Greek indices is assumed, and $n_\alpha(\mathbf{x}_\perp)$, $\alpha = 1,2,3$, are the components of the vector

$$\mathbf{n}(\mathbf{x}_\perp) = \hat{i}_z - \nabla_\perp h(\mathbf{x}_\perp),$$

which is normal to the surface, but not of unit length. The function $\theta(x)$ is the unit step function

$$\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}$$

The Green's function $G_+(\mathbf{x}'|\,\mathbf{x}'')$ is equal to $G(\mathbf{x}'\mid\mathbf{x}'')$ in region I, but it continues smoothly across the interface into region II, where it does not equal $G(\mathbf{x}'\mid\mathbf{x}'')$. This is the reason for introducing the distinguishing notation subscript +. Thus $G_{+,\alpha}(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')$ and $G_+(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')$ are the upper boundary values of grad G and G, respectively.

Similarly, for region II the integral relation is

$$\begin{aligned} G_-(\mathbf{x}'\mid\mathbf{x}'')\,\theta(-z'+h(\mathbf{x}'_\perp)) &= G_{0-}(\mathbf{x}'-\mathbf{x}'')\,\theta(-z''+h(\mathbf{x}''_\perp)) \\ &- \int d^2\mathbf{x}_\perp\, n_\alpha(\mathbf{x}_\perp)\left[G_{0-}(\mathbf{x}'-\mathbf{x}_s(\mathbf{x}_\perp))\,G_{-,\alpha}(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')\right. \\ &\left. + \frac{\partial}{\partial x'_\alpha}\,G_{0-}(\mathbf{x}'-\mathbf{x}_s(\mathbf{x}_\perp))\,G_-(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')\right]. \end{aligned} \tag{2b}$$

The Green's function $G_{0-}(\mathbf{x})$ is a solution of

$$(\nabla^2 + k_-^2)G_{0-}(\mathbf{x}) = \delta_3(\mathbf{x})$$

for all $\mathbf{x}$, and it satisfies the radiation boundary condition as $|\mathbf{x}| \to \infty$. The Green's function is equal to $G_-(\mathbf{x}'\mid\mathbf{x}'')$ in region II but not in region I and $G_{-,\alpha}(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')$ and $G_-(\mathbf{x}_s(\mathbf{x}_\perp)\mid\mathbf{x}'')$ are the lower boundary values of grad G and G, respectively.

Integral equations for $G_{\pm}(x_s(x_\perp) \mid x'')$ and $n_\alpha(x_\perp) G_{\pm,\alpha}(x_s(x_\perp) \mid x'')$ can be derived from (2a) and (2b) by taking the surface limit of (2a) and (2b) and the gradient of (2a) and (2b). These limits must be performed carefully since both sides of Eqs. (2a) and (2b) possess singularities in x' as x' approaches $x_s(x'_\perp)$. The discontinuities are particularly bad in the gradients of (2a) and (2b). One way of dealing with this is to rewrite (2a) and (2b) and the normal derivatives of (2a) and (2b) in such a way that the singularities present manifestly cancel to such an extent that the surface limit can be properly taken. First (2a) and (2b) will be considered. Let

$$\theta(z) = \tfrac{1}{2} + E(z)$$

where

$$E(z) = \begin{cases} +\tfrac{1}{2}, & z > 0, \\ -\tfrac{1}{2}, & z < 0 \end{cases}$$

With no loss of generality it can be assumed that the source is in region I. Equations (2a) and (2b) can then be rewritten as

$$G_+(x' \mid x'') = 2G_{0+}(x' - x'') - 2G_+(x' \mid x'')E(z' - h(x'_\perp)) + 2\int d^2x_\perp n_\alpha(x_\perp)\left[G_{0+}(x' - x_s(x_\perp))G_{+,\alpha}(x_s(x_\perp) \mid x'') + \frac{\partial}{\partial x'_\alpha} G_{0+}(x' - x_s(x_\perp))G_+(x_s(x_\perp) \mid x'')\right], \tag{3a}$$

$$G_-(x' \mid x'') = 2G_-(x' \mid x'')E(z' - h(x'_\perp)) - 2\int d^2x_\perp n_\alpha(x_\perp)\left[G_{0-}(x' - x_s(x_\perp))G_{-,\alpha}(x_s(x_\perp) \mid x'') + \frac{\partial}{\partial x'_\alpha} G_{0+}(x' - x_s(x_\perp))G_-(x_s(x_\perp) \mid x'')\right]. \tag{3b}$$

As already mentioned $G_+(x' \mid x'') = G(x' \mid x'')$ in region I and $G_-(x' \mid x'') = G(x' \mid x'')$ in region II. If the interface is sufficiently smooth, then a continuation of $G_+(x' \mid x'')$ through the interface into region II will not be singular until x' is a finite distance below the interface, since any singularities encountered correspond to images of the source below the interface. Similarly, if $G_-(x' \mid x'')$ is continued upward through the interface, then no singularities will be encountered until x' is a finite distance above the interface. For this reason the surface singularities due to the $G_\pm(x' \mid x'')E(z' - h(x'_\perp))$ term on the right-hand sides of (3a) and (3b) must be canceled by similar singularities in the integral term, leaving a function that can be continued through the interface. In particular this cancelation will be done in such a way that the terms on the right-hand sides of (3a) and (3b) are replaced by terms that are individually continuous, so that the surface limit can then be taken on each term separately. The normal derivative of (3a) and (3b), that is $n(x_\perp)$ dotted into the gradients of (3a) and (3b), will be treated in a similar manner. This partial removal of the singularities and performance of the surface limit yields four coupled surface integral equations which, when combined with the boundary conditions at the interface, will be the basis for subsequent developments.

The first step in the partial cancelation of the singularities in (3a) and (3b) is to rewrite the free Green's function. This Green's function $G_{0\pm}(x)$ is given by

$$G_{0\pm}(x) = (2\pi)^{-3}\int d^3k\, e^{ik\cdot x}(k_\pm^2 - k^2 + i\epsilon)^{-1} = -\frac{e^{ik_\pm|x|}}{4\pi|x|}.$$

Integrating over k_z only gives

$$G_{0\pm}(x) = (2\pi)^{-2}\int d^2k_\perp\, e^{ik_\perp\cdot x_\perp}\frac{e^{i\sqrt{k_\pm^2 - k_\perp^2 + i\epsilon}\,|z|}}{2i\sqrt{k_\pm^2 - k_\perp^2 + i\epsilon}}. \tag{4}$$

Taking the z derivative of $G_{0\pm}(x)$ gives

$$\frac{\partial}{\partial z} G_{0\pm}(x) = (2\pi)^{-2}\int d^2k_\perp\, e^{ik_\perp\cdot x_\perp} e^{i\sqrt{k_\pm^2 - k_\perp^2}\,|z|}E(z),$$

which in the neighborhood of $z = 0$ becomes

$$\frac{\partial}{\partial z} G_{0\pm}(x) \to \delta_2(x_\perp)E(z).$$

Because of this behavior, the most important singularity in the integral term in (3a) and (3b) must be contained in the $n_\alpha(\partial/\partial x'_\alpha)G_{0\pm}$ part of the integrand. This singularity can be removed by letting

$$\frac{\partial}{\partial x_\alpha} G_{0\pm}(x) \triangleq G_{0\pm\alpha}(x) + \delta_z(x_\perp)\cos(k_\pm z)E(z)\delta_{\alpha 3}. \tag{5}$$

This choice is not unique, but has been chosen to be simple and to improve the behavior of the $k_\pm \to \infty$ limit. Thus,

$$G_{0\pm\alpha}(x) = (2\pi)^{-3}\int d^3k\, e^{ik\cdot x}\left(\frac{k_\alpha}{k_\pm^2 - k^2 + i\epsilon} - P\frac{k_z\delta_{\alpha 3}}{k_\pm^2 - k_z^2}\right), \tag{6}$$

where P indicates the Cauchy principal value distribution. Substitution of (5) into (3a) and (3b) gives the integral relations

$$G_+(x' \mid x'') = 2G_{0+}(x' - x'') - 2\{G_+(x' \mid x'') - G_+(x_s(x'_\perp) \mid x'')\cos[k_+(z' - h(x'_\perp))]\}E(z' - h(x'_\perp)) + 2\int d^2x_\perp n_\alpha(x_\perp)[G_{0+}(x' - x_s(x_\perp))G_{+,\alpha}(x_s(x_\perp) \mid x'') + G_{0+\alpha}(x' - x_s(x_\perp))G_+(x_s(x_\perp) \mid x'')] \tag{7a}$$

and similarly

$$G_-(x' \mid x'') = 2\{G_-(x' \mid x'') - G_-(x_s(x'_\perp) \mid x'')\cos[k_-(z' - h(x'_\perp))]\}E(z' - h(x'_\perp)) - 2\int d^2x_\perp n_\alpha(x_\perp)[G_{0-}(x' - x_s(x_\perp))G_{-,\alpha}(x_s(x_\perp) \mid x'') + G_{0-\alpha}(x' - x_s(x_\perp))G_-(x_s(x_\perp) \mid x'')]. \tag{7b}$$

In (7a) and (7b) the coefficient of $E(z' - h(x'_\perp))$ vanishes linearly as $x' \to x_s(x'_\perp)$ and thus this term is continuous (but not differentiable) there. The surface limits of (7a) and (7b) can now be legitimately taken on a term by term basis, giving the integral equations

$$G_+(x_s(x_\perp) \mid x'') = 2G_{0+}(x_s(x'_\perp) - x'') + 2\int d^2x_\perp[G_{0+}(x_s(x'_\perp) - x_s(x_\perp))G_{+n}(x_s(x_\perp) \mid x'') + n_\alpha(x_\perp)G_{0+\alpha}(x_s(x'_\perp) - x_s(x_\perp))G_+(x_s(x_\perp) \mid x'')], \tag{8a}$$

$$G_-(x_s(x'_\perp) \mid x'') = -2\int d^2x_\perp[G_{0-}(x_s(x'_\perp) - x_s(x_\perp))G_{-n}(x_s(x_\perp) \mid x'') + n_\alpha(x_\perp)G_{0-\alpha}(x_s(x'_\perp) - x_s(x_\perp))G_-(x_s(x_\perp) \mid x'')], \tag{8b}$$

where

$$G_{\pm n}(x \mid x'') = n_\alpha(x_\perp)G_{\pm,\alpha}(x \mid x'')$$

In a similar way, integral relations for $G_{\pm n}$ can be found. Taking the gradient of (3a) gives

$$\frac{\partial}{\partial x'_\alpha} G_+(x' \mid x'') = 2\frac{\partial}{\partial x'_\alpha} G_{0+}(x' - x'') - 2\left(\frac{\partial}{\partial x'_\alpha} G_+(x' \mid x'')\right) E(z' - h(x'_\perp)) - 2G_+(x' \mid x'') n_\alpha(x'_\perp)\,\delta(z' - h(x'_\perp))$$
$$+ 2\int d^2x_\perp \left[\frac{\partial}{\partial x'_\alpha} G_{0+}(x' - x_s(x_\perp))\, G_{+n}(x_s(x_\perp) \mid x'') + n_\beta(x_\perp)\left(\frac{\partial^2}{\partial x'_\beta \partial x'_\alpha} G_{0+}(x' - x_s(x_\perp))\right) G_+(x_s(x_\perp) \mid x'')\right] \tag{9a}$$

and similarly the gradient of (3b) gives

$$\frac{\partial}{\partial x'_\alpha} G_-(x' \mid x'') = 2\left(\frac{\partial}{\partial x'_\alpha} G_-(x' \mid x'')\right) E(z' - h(x'_\perp)) + 2G_-(x' \mid x'') n_\alpha(x'_\perp)\,\delta(z' - h(x'_\perp))$$
$$- 2\int d^2x_\perp \left[\frac{\partial}{\partial x'_\alpha} G_{0-}(x' - x_s(x_\perp))\, G_{-n}(x_s(x_\perp) \mid x'') + n_\beta(x_\perp)\left(\frac{\partial^2}{\partial x'_\beta \partial x'_\alpha} G_{0-}(x' - x_s(x_\perp))\right) G_-(x_s(x_\perp) \mid x'')\right]. \tag{9b}$$

As before, the singularities must be partially canceled before the surface limit can be legitimately performed. This cancelation can be accomplished by letting

$$\frac{\partial}{\partial x_\alpha} G_{0\pm}(x) \triangleq G_{0\pm\alpha}(x) + \delta_2(x_\perp)\cos(k_\pm z)\, E(z)\delta_{\alpha 3}, \tag{10}$$

$$\frac{\partial^2}{\partial x_\beta \partial x_\alpha} G_{0\pm}(x) \triangleq G_{0\pm\beta\alpha}(x) + \delta_2(x_\perp)[\delta(z) - k_\pm \sin(k_\pm z)\, E(z)]\,\delta_{\beta\alpha}. \tag{11}$$

Equation (6) gives an explicit expression for $G_{0\pm\alpha}(x)$. The corresponding expression for $G_{0\pm\beta\alpha}(x)$ is

$$G_{0\pm\beta\alpha}(x) = -(2\pi)^{-3}\int d^3k\, e^{i\mathbf{k}\cdot\mathbf{x}}\left(\frac{k_\beta k_\alpha}{k_\pm^2 - k^2 + i\epsilon} - P\frac{k_z^2\delta_{\beta\alpha}}{k_\pm^2 - k_z^2}\right). \tag{12}$$

Letting

$$G_{\pm n}(x' \mid x'') = n_\alpha(x'_\perp)\frac{\partial}{\partial x'_\alpha} G_\pm(x' \mid x'').$$

and combining (10) and (11) with (9a) gives

$$G_{+n}(x' \mid x'') = 2n_\alpha(x'_\perp)\frac{\partial}{\partial x'_\alpha} G_{0+}(x' - x'')$$
$$- 2\{G_{+n}(x' \mid x'') - G_{+n}(x_s(x'_\perp) \mid x'')\cos[k_+(z' - h(x'_\perp))]$$
$$+ k_+ G_+(x_s(x_\perp) \mid x'')n^2(x'_\perp)\sin[k_+(z' - h(x'_\perp))]\}\, E(z' - h(x'_\perp))$$
$$- 2[G_+(x' \mid x'') - G_+(x_s(x'_\perp) \mid x'')]n^2(x'_\perp)\,\delta(z' - h(x_\perp))$$
$$+ 2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0+\alpha}(x' - x_s(x_\perp))\, G_{+n}(x_s(x_\perp) \mid x'')$$
$$+ n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0+\alpha\beta}(x' - x_s(x_\perp))\, G_+(x_s(x_\perp) \mid x'')]. \tag{13a}$$

Similarly, combining (10) and (11) with (9b) gives

$$G_{-n}(x' \mid x'') = 2\{G_{-n}(x' \mid x'') - G_{-n}(x_s(x'_\perp) \mid x'')\cos[k_-(z' - h(x'_\perp))]$$
$$+ k_- G_-(x_s(x'_\perp) \mid x'')\, n^2(x'_\perp)$$
$$\times \sin[k_-(z' - h(x'_\perp))]\}\, E(z' - h(x'_\perp))$$
$$+ 2\{G_-(x' \mid x'') - G_-(x_s(x') \mid x'')\}\, n^2(x'_\perp)\,\delta(z' - h(x'_\perp))$$
$$- 2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0-\alpha}(x' - x_s(x_\perp))\, G_{-n}(x_s(x_\perp) \mid x'')$$
$$+ n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0-\alpha\beta}(x' - x_s(x_\perp))\, G_-(x_s(x_\perp) \mid x'')]. \tag{13b}$$

The coefficients of $\delta(z' - h(x'_\perp))$ and $E(z' - h(x'_\perp))$ in (13a) and (13b) vanish linearly as $z' \to h(x'_\perp)$, thus the $\delta(z' - h(x'_\perp))$ can be dropped and the $E(z' - h(x'_\perp))$ term is continuous at the interface. Taking the surface limits of (13a) and (13b) gives

$$G_{+n}(x_s(x'_\perp) \mid x'') = 2G_{0+n}(x_s(x'_\perp) \mid x'')$$
$$+ 2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0+\alpha}(x_s(x'_\perp) - x_s(x'_\perp))\, G_{+n}(x_s(x_\perp) \mid x'')$$
$$+ n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0+\alpha\beta}(x_s(x'_\perp) - x_s(x_\perp))\, G_+(x_s(x_\perp) \mid x'')], \tag{14a}$$

$$G_{-n}(x_s(x'_\perp) \mid x'') = -2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0-\alpha}(x_s(x'_\perp) - x_s(x_\perp)\, G_{-n}(x_s(x_\perp) \mid x'')$$
$$+ n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0-\alpha\beta}(x_s(x'_\perp) - x_s(x_\perp))\, G_-(x_s(x_\perp) \mid x'')]. \tag{14b}$$

The boundary conditions can be used to relate G_- and G_{-n} to G_+ and G_{+n} evaluated on the interface, and thus (8b) and (14b) can be written in terms of G_+ and G_{+n}. Carrying this out and introducing the compact notation

$$x_s(x_\perp) = x_s, \quad x_s(x'_\perp) = x'_s, \text{ etc.}$$

and letting

$$R = Z_+/Z_-$$

gives

$$G_+(x'_s \mid x'') = -2\int d^2x_\perp [R\, G_{0-}(x'_s - x_s)\, G_{+n}(x_s \mid x'')$$
$$+ n_\alpha(x_\perp) G_{0-\alpha}(x'_s - x_s)\, G_+(x_s \mid x'')] \tag{15a}$$

$$G_{+n}(x'_s \mid x'') = -2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0-\alpha}(x'_s - x_s)\, G_{+n}(x_s \mid x'')$$
$$+ R^{-1} n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0-\alpha\beta}(x'_s - x_s)\, G_+(x_s \mid x'')]. \tag{15b}$$

Rewriting (8a) and (14a) gives

$$G_+(x_s \mid x'') = 2G_{0+}(x'_s - x'') + 2\int d^2x_\perp [G_{0+}(x'_s - x_s)\, G_{+n}(x_s \mid x'')$$
$$+ n_\alpha(x_\perp) G_{0+\alpha}(x'_s - x_s) G_+(x_s \mid x'')] \tag{16a}$$

$$G_{+n}(x_s \mid x'') = 2G_{0+n}(x_s \mid x'')$$
$$+ 2\int d^2x_\perp [n_\alpha(x'_\perp) G_{0+\alpha}(x'_s - x_s)\, G_{+n}(x_s \mid x'')$$
$$+ n_\alpha(x'_\perp) n_\beta(x_\perp) G_{0+\alpha\beta}(x'_s - x_s)\, G_+(x_s \mid x'')]. \tag{16b}$$

Equations (15a), (15b), (16a), and (16b) are the basis for the next developments.

In order to be able to solve for $G_+(x_s \mid x'')$ and $G_{+n}(x_s \mid x'')$ two coupled integral equations are needed, one expressing $G_+(x_s \mid x'')$ in terms of $G_+(x_s \mid x'')$ and $G_{+n}(x_s \mid x'')$ and another expressing $G_{+n}(x_s \mid x'')$ in terms of $G_+(x_s \mid x'')$ and $G_{+n}(x_s \mid x'')$. At first sight it might appear that either (15a) and (15b) or (16a) and (16b) could be used, but this would be incorrect. The solutions of (16a) and (16b) can not be unique since they were derived from the Green's integral relation

for the upper region, $z > h(x_\perp)$, only. In other words, (16a) and (16b) are equivalent to saying that there is a point source of x″ and the scattered wave for $z > h(x_\perp)$ is outgoing. There are many field configurations that satisfy this condition. This nonuniqueness means that the resolvent of (16a) and (16b) is singular, and (16a) and (16b) are redundant. A similar argument applies to (15a) and (15b) since they were derived from Green's integral relation for the lower region only. Thus the resolvent of (15a) and (15b) is also singular. However, a linear combination of (15a) and (16a) together with a linear combination of (15b) and (16b) constitutes a pair of coupled integral equations which will, in general, have a well-defined resolvent and thus a unique solution. This procedure can be heuristically justified by a matrix analogy that embodies the essential features of the problem.

Assume that there are two matrices, each of which maps the same vector space A onto different subspaces B and C. Because the mapping is not one to one, the matrices are not invertable. Also assume that the subspaces span the original vector space, i.e., that $A = B \oplus C$. Thus the mapping generated by a matrix which is a linear combination of the matrices or a linear combination of the matrices multiplied by nonsingular matrices will in general be one to one, and the matrix will be invertable. The lack of uniqueness is very analogous to the gauge arbitrariness of electrodynamics.

Equations (15a) and (16a) will be combined as follows; multiply (16a) by $f_{+\mu\nu}(x, x'_s)n_\nu(x'_\perp)$, (15a) by $f_{-\mu\nu}(x, x'_s) \times n_\nu(x'_\perp)$, add the results and integrate over $x'_\perp$. This gives

$$\begin{aligned}\int d^2x'_\perp &[f_{+\mu\nu}(x, x'_s) + f_{-\mu\nu}(x, x'_s)]n_\nu(x'_\perp)\,G_+(x'_s|x'') \\ &= 2\int d^2x'_\perp\, f_{+\mu\nu}(x, x'_s)n_\nu(x'_\perp)\,G_{0+}(x'_s - x'') \\ &\quad + 2\int d^2x'_\perp\, d^2x_{1\perp}\{[f_{+\mu\nu}(x, x'_s)n_\nu(x'_\perp)\,G_{0+}(x'_s - x_{1s}) \\ &\quad - Rf_{-\mu\nu}(x, x'_s)n_\nu(x'_\perp)\,G_{0-}(x'_s - x_{1s})]G_{+n}(x_{1s} \mid x'') \\ &\quad + [f_{+\mu\nu}(x, x'_s)n_\nu(x'_\perp)\,G_{0+\alpha}(x'_s - x_{1s}) \\ &\quad - f_{-\mu\nu}(x, x'_s)n_\nu(x'_\perp)G_{0-\alpha}(x'_s - x_{1s})]n_\alpha(x_{1\perp})\,G_+(x_{1s}|x'')\}.\end{aligned} \tag{17a}$$

Similarly combining (15b) and (16b) by multiplying (16b) by $g_+(x, x'_s)$, (15b) by $g_-(x, x'_s)$, adding the results, and integrating over $x'_\perp$ gives

$$\begin{aligned}\int d^2x'_\perp &[g_+(x, x'_s) + g_-(x, x'_s)]G_{+n}(x'_s \mid x'') \\ &= 2\int d^2x'_\perp\, g_+(x, x'_s)n_\alpha(x'_\perp)\,G_{0+\alpha}(x'_s - x'') \\ &\quad + 2\int d^2x'_\perp\, d^2x_{1\perp}\{n_\alpha(x'_\perp)[g_+(x, x'_s)\,G_{0+\alpha}(x'_s - x_{1s}) \\ &\quad - g_-(x, x'_s)\,G_{0-\alpha}(x'_s - x_{1s})]G_{+n}(x_{1s} \mid x'') \\ &\quad + n_\alpha(x'_\perp)[g_+(x, x'_s)\,G_{0+\alpha\beta}(x'_s - x_{1s}) \\ &\quad - R^{-1}g_-(x, x'_s)\,G_{0-\alpha\beta}(x'_s - x_{1s})]n_\beta(x_{1\perp})\,G_+(x_{1s} \mid x'')\}.\end{aligned} \tag{17b}$$

At the present juncture the functions $f_{\pm\alpha\beta}(x, x_s)$ and $g_\pm(x, x_s)$ are rather arbitrary. Equations (17a) and (17b) as they presently stand are not suitable for iteration. One condition that will be imposed on the choice of $f_{\pm\alpha\beta}(x, x_s)$ and $g_\pm(x, x_s)$ is that (17a) and (17b), when transformed to k space, have a form that is suitable for iteration. Also, it is desirable that the full Green's function be easily constructed. It should be noted that (17a) and (17b) are three dimensional while (15a), (15b), (16a), and (16b) are two dimensional. The reason for expressing everything in terms of three-space rather than two-space is that the Fourier transforms of the kernels are simpler and later statistical averaging is made much easier. The passage from the two-dimensional surface integral equation to a three-dimensional one implies nonuniqueness in addition to the nonuniqueness due to going from four doubly redundant integral equations to two nonredundant ones. This arbitrariness is similar to the gauge arbitrariness of electrodynamics, which is also the result of expanding the number of degrees of freedom of the system in order to obtain a simpler set of equations. The choice of particular forms for $f_{\pm\alpha\beta}(x, x_s)$ and $g_\pm(x, x_s)$ corresponds to the various choices of gauges in electrodynamics, and the choice which ultimately is made is dictated by the problem under consideration. For the present problem it is felt that the particular choice, which will shortly be discussed, is the simplest commensurate with the desired properties outlined in the introduction.

It is convenient to transform (17a) and (17b) to k space. The transformation of the Green's function to k space is defined by

$$Z(x' \mid x'') = (2\pi)^{-6}\int d^3k'\, d^3k''\, e^{i(k'\cdot x' - k''\cdot x'')}Z(k' \mid k''),$$

where $$Z(x' \mid x'') = G_+(x' \mid x'') \quad \text{or} \quad G_{+n}(x' \mid x''). \tag{18}$$

Integrals of the form

$$\mathcal{J} = \int d^2x_\perp\, F(x_s)\,n(x_\perp)\cdot V(x_s)$$

occur frequently in (17a) and (17b), so rewriting $\mathcal{J}$ in terms of **k** space will make the transformation of (17a) and (17b) easier. The transformation of F and V to **k** space is defined by

$$F(x_s) = (2\pi)^{-3}\int d^3k\, F(k)\,e^{-ik\cdot x_s}, \tag{19}$$

$$V(x_s) = (2\pi)^{-3}\int d^3k\, V(k)\,e^{-ik\cdot x_s}. \tag{20}$$

Using (19) and (20) together with

$$n(x_\perp) = \hat{i}_z - \nabla_\perp h(x_\perp),$$

$$x_s(x_\perp) = x_\perp + \hat{i}_z h(x_\perp)$$

gives

$$\begin{aligned}F(x_s)n(x_\perp)\cdot V(x_s) = (2\pi)^{-6}&\int d^3k_1 d^3k_2\, F(k_1)\,e^{-i(k_{1\perp} - k_{2\perp})\cdot x_\perp} \\ &\times\left(i_z - i\,\frac{\nabla_\perp}{k_{1z} - k_{2z}}\right)\cdot V(k_2)\,e^{-i(k_{1z} - k_{2z})h(x_\perp)}.\end{aligned}$$

Integrating over $x_\perp$, interchanging the order of the integrations, and integrating by parts in $x_\perp$ gives

$$\begin{aligned}\mathcal{J} = (2\pi)^{-6}\int d^3k_1 d^3k_2\, F(k_1)\Big\{&A(k_1 - k_2)\,\frac{\mathbf{k}_1 - \mathbf{k}_2}{k_{1z} - k_{2z}}\cdot V(k_2) \\ &- \frac{i}{k_{1z} - k_{2z}}\oint_c ds_\perp \cdot V(k_2) \\ &\times \exp[i(\mathbf{k}_{1\perp} - \mathbf{k}_{2\perp})\cdot x_\perp + i(k_{1z} - k_{2z})h(x_\perp)]\Big\},\end{aligned}$$

where

$$A(k) = \int d^2x_\perp\, e^{-k\cdot x_s(x_\perp)}. \tag{21}$$

The last term in this expression for $\mathcal{J}$ is a line integral in the plane over the contour C which is to be removed to infinity. The contour integral, as a function of $\mathbf{k}_{1\perp} - \mathbf{k}_{2\perp}$, oscillates at an increasing rate as C goes to infinity. It will be assumed that $F(k_1)$ or $V(k_2)$ is sufficiently well-behaved that the integral over $\mathbf{k}_1$ and $\mathbf{k}_2$ of these rapid oscillations vanishes, that is, the surface term can be dropped. Thus

$$\mathcal{J} = (2\pi)^{-6} \int d^3k_1 d^3k_2 F(\mathbf{k}_1) A(\mathbf{k}_1 - \mathbf{k}_2)[(\mathbf{k}_1 - \mathbf{k}_2)\cdot \mathbf{V}(\mathbf{k}_2)/(k_{1z} - k_{2z})]\,. \quad (22)$$

Combining (17a) with (18) and (22) gives

$$\begin{aligned}
&\int d^3k_1 d^3k_2 [f_{+\mu\nu}(\mathbf{k}_1, \mathbf{k}_2) + f_{-\mu\nu}(\mathbf{k}_1, \mathbf{k}_2)] \frac{k_{1\nu} - k_{2\nu}}{k_{1z} - k_{2z}} \\
&\quad A(\mathbf{k}_1 - \mathbf{k}_2) G_+(\mathbf{k}_2 \mid \mathbf{k}'') \\
&\quad - 2(2\pi)^3 \int d^3k_1 f_{+\mu\nu}(\mathbf{k}, \mathbf{k}_1) \frac{k_{1\nu} - k''_\nu}{k_{1z} - k''_z} A(\mathbf{k}_1 - \mathbf{k}'') G_{0+}(\mathbf{k}'') \\
&\quad - 2(2\pi)^{-3} \int d^3k_1 d^3k_2 d^3k_3 \Big([f_{+\mu\nu}(\mathbf{k}, \mathbf{k}_1) G_{0+}(\mathbf{k}_2) \\
&\quad - R f_{-\mu\nu}(\mathbf{k}, \mathbf{k}_1) G_{0-}(\mathbf{k}_2)] \\
&\quad \times \frac{k_{1\nu} - k_{2\nu}}{k_{1z} - k_{2z}} A(\mathbf{k}_1 - \mathbf{k}_2) A(\mathbf{k}_2 - \mathbf{k}_3) G_{+n}(\mathbf{k}_3 \mid \mathbf{k}'') \\
&\quad + [f_{+\mu\nu}(\mathbf{k}, \mathbf{k}_1) G_{0+\alpha}(\mathbf{k}_2) \\
&\quad - f_{-\mu\nu}(\mathbf{k}, \mathbf{k}_1) G_{0-\alpha}(\mathbf{k}_2)] \frac{k_{1\nu} - k_{2\nu}}{k_{1z} - k_{2z}} \\
&\quad \times \frac{k_{2\alpha} - k_{3\alpha}}{k_{2z} - k_{3z}} A(\mathbf{k}_1 - \mathbf{k}_2) A(\mathbf{k}_2 - \mathbf{k}_3) G_+(\mathbf{k}_3 \mid \mathbf{k}'') \Big) = 0,
\end{aligned} \quad (23)$$

where

$$G_{0\pm}(\mathbf{k}) = (k_\pm^2 - \mathbf{k}^2 + i\epsilon)^{-1}, \quad (24)$$

$$G_{0\pm\alpha}(\mathbf{k}) = k_\alpha G_{0\pm}(\mathbf{k}) - P[k_z/(k_\pm^2 - k_z^2)]\delta_{\alpha 3}, \quad (25)$$

$$G_{0\pm\alpha\beta}(\mathbf{k}) = -k_\alpha k_\beta G_{0\pm}(\mathbf{k}) + P[k_z^2/(k_\pm^2 - k_z^2)]\delta_{\alpha\beta}. \quad (26)$$

Similarly, combining (17b) with (18) and (22) gives

$$\begin{aligned}
&\int d^3k_1 d^3k_2 [g_+(\mathbf{k}, \mathbf{k}_1) + g_-(\mathbf{k}, \mathbf{k}_1)] A(\mathbf{k}_1 - \mathbf{k}_2) G_{+n}(\mathbf{k}_2 \mid \mathbf{k}'') \\
&\quad - 2i(2\pi)^3 \int d^3k_1 g_+(\mathbf{k}, \mathbf{k}_1) A(\mathbf{k}_1 - \mathbf{k}'') G_{0+}(\mathbf{k}'') \\
&\quad - 2(2\pi)^{-3} \int d^3k_1 d^3k_2 d^3k_3 \\
&\quad \times \Big([g_+(\mathbf{k}, \mathbf{k}_1) G_{0+\alpha}(\mathbf{k}_2) - g_-(\mathbf{k}, \mathbf{k}_1) G_{0-\alpha}(\mathbf{k}_2)] \\
&\quad \times \frac{k_{1\alpha} - k_{2\alpha}}{k_{1z} - k_{2z}} A(\mathbf{k}_1 - \mathbf{k}_2) A(\mathbf{k}_2 - \mathbf{k}_3) G_{+n}(\mathbf{k}_3 \mid \mathbf{k}'') \\
&\quad + [g_+(\mathbf{k}, \mathbf{k}_1) G_{0+\alpha\beta}(\mathbf{k}_2) - R^{-1} g_-(\mathbf{k}, \mathbf{k}_1) G_{0-\alpha\beta}(\mathbf{k}_2)] \\
&\quad \times \frac{k_{1\alpha} - k_{2\alpha}}{k_{1z} - k_{2z}} \frac{k_{2\beta} - k_{3\beta}}{k_{2z} - k_{3z}} A(\mathbf{k}_1 - \mathbf{k}_2) A(\mathbf{k}_2 - \mathbf{k}_3) \\
&\quad \times G_+(\mathbf{k}_3 \mid \mathbf{k}'') \Big) = 0.
\end{aligned} \quad (27)$$

As mentioned before, it is desirable to choose $g_\pm$ and $f_{\pm\alpha\beta}$ is such a way that the integral equations are suitable for iteration. Thus, choose

$$g_-(\mathbf{k}, \mathbf{k}') = \delta_3(\mathbf{k} - \mathbf{k}') - g_+(\mathbf{k}, \mathbf{k}'), \quad (28)$$

$$f_{-\alpha\beta}(\mathbf{k}, \mathbf{k}') = \delta_3(\mathbf{k} - \mathbf{k}')\delta_{\alpha\beta} - f_{+\alpha\beta}(\mathbf{k}, \mathbf{k}'). \quad (29)$$

Also, it is desirable for the series generated by the iteration of (23) and (27) to be simple for the statistical averaging which must be performed later. In particular, it is desirable to choose $g_\pm$ and $f_{\pm\alpha\beta}$ so that they are functionally independent of $A(\mathbf{k})$. If this is true then the n^{th} term in the series contains a product of the form

$$\prod_{j=1}^{n} A(\mathbf{k}_j),$$

and this is the only way the surface height $h(\mathbf{x}_\perp)$ enters. This product, since it is easy to average in terms of a cluster expansion, means that the whole series of products of series will be amenable to statistical averaging. It is worth noting here that if it is required that the lowest order diagrams, i.e., the Born approximation, corresponds to the Kirchhoff approximation, then $g_\pm$ and $f_{\pm\alpha\beta}$ cannot be chosen to be independent of $A(\mathbf{k})$, and therefore the lowest-order approximation will not be the Kirchhoff approximation. The reason why expansion around the Kirchhoff approximation is undesirable can be understood by recalling that the Kirchhoff approximation involves replacing the field and its normal derivative on the interface by their values assuming the surface is an infinite plane which is locally tangent. For an impenetrable surface these are proportional to those of the incident field, but for an interface complications arise because the reflection coefficient is a function of the slope of the surface resulting in terms that are difficult to handle statistically. It is also convenient to choose g_+ and $f_{+\alpha\beta}$ to be of the form

$$g_+(\mathbf{k}, \mathbf{k}') = g(\mathbf{k})\delta_3(\mathbf{k} - \mathbf{k}'), \quad (30)$$

$$f_{+\alpha\beta}(\mathbf{k}, \mathbf{k}') = f(\mathbf{k})\delta_{\alpha\beta}\delta_3(\mathbf{k} - \mathbf{k}'). \quad (31)$$

Introducing the new quantities

$$T_\alpha(\mathbf{k} \mid \mathbf{k}'') G_{0+}(\mathbf{k}'') = (2\pi)^{-6} \int d^3k_2 [(k_\alpha - k_{2\alpha})/(k_z - k_{2z})] \times A(\mathbf{k} - \mathbf{k}_2) G_+(\mathbf{k}_2 \mid \mathbf{k}'')], \quad (32)$$

$$T_n(\mathbf{k} \mid \mathbf{k}'') G_{0+}(\mathbf{k}'') = (2\pi)^{-6} \int d^3k_2 A(\mathbf{k} - \mathbf{k}_2) G_{+n}(\mathbf{k}_2 \mid \mathbf{k}''), \quad (33)$$

and combining (28)–(33) with (23) and (27) gives the coupled integral equations

$$\begin{aligned}
T_\alpha(\mathbf{k} \mid \mathbf{k}'') &= 2(2\pi)^{-3} f(\mathbf{k}) \frac{k_\alpha - k''_\alpha}{k_z - k''_z} A(\mathbf{k} - \mathbf{k}'') \\
&\quad + 2(2\pi)^{-3} \int d^3k_1 \frac{k_\alpha - k_{1\alpha}}{k_z - k_{1z}} A(\mathbf{k} - \mathbf{k}_1) \\
&\quad \times \{\{f(\mathbf{k})[G_{0+}(\mathbf{k}_1) + R G_{0-}(\mathbf{k}_1)] - R G_{0-}(\mathbf{k}_1)\} T_n(\mathbf{k}_1 \mid \mathbf{k}'') \\
&\quad + \{f(\mathbf{k})[G_{0+\beta}(\mathbf{k}_1) + G_{0-\beta}(\mathbf{k}_1)] - G_{0-\beta}(\mathbf{k}_\perp)\} T_\beta(\mathbf{k}_1 \mid \mathbf{k}'')\}.
\end{aligned} \quad (34)$$

$$\begin{aligned}
T_n(\mathbf{k} \mid \mathbf{k}'') &= 2i(2\pi)^{-3} g(\mathbf{k}) \frac{(k_\alpha - k''_\alpha) k''_\alpha}{k_z - k''_z} A(\mathbf{k} - \mathbf{k}'') \\
&\quad + 2(2\pi)^{-3} \int d^3k_1 \frac{k_\alpha - k_{1\alpha}}{k_z - k_{1z}} A(\mathbf{k} - \mathbf{k}_1) \\
&\quad \times \{\{g(\mathbf{k})[G_{0+\alpha}(\mathbf{k}_1) - G_{0-\alpha}(\mathbf{k}_1)] \\
&\quad - G_{0-\alpha}(\mathbf{k}_1)\} T_n(\mathbf{k}_1 \mid \mathbf{k}'') \\
&\quad + \{g(\mathbf{k})[G_{0+\alpha\beta}(\mathbf{k}_1) + R^{-1} G_{0-\alpha\beta}(\mathbf{k}_1)] \\
&\quad - R^{-1} G_{0-\alpha\beta}(\mathbf{k}_1)\} T_\beta(\mathbf{k}_1 \mid \mathbf{k}'')\}.
\end{aligned} \quad (35)$$

It is still necessary to choose a form for $g(\mathbf{k})$ and $f(\mathbf{k})$. This will be done by requiring that the lowest-order term in an iterative solution of (34) and (35) be exact when $h(\mathbf{x}_\perp) = 0$. The first step in this direction is to write the full Green's function in terms of T_n and T_α. Equation (2a) gives for region I, i.e., for $z' > h(\mathbf{x}'_\perp)$

$$\begin{aligned}
&G(\mathbf{x}' \mid \mathbf{x}'') \\
&\quad = G_{0+}(\mathbf{x}' - \mathbf{x}'') + \int d^2x_\perp \Big(G_{0+}(\mathbf{x}' - \mathbf{x}_s) G_{+n}(\mathbf{x}_s \mid \mathbf{x}'') \\
&\quad + n_\alpha(\mathbf{x}_\perp) \frac{\partial}{\partial x'_\alpha} G_{0+}(\mathbf{x}' - \mathbf{x}_s) G_+(\mathbf{x}_s \mid \mathbf{x}'') \Big).
\end{aligned} \quad (36)$$

Introducing the Fourier transforms for the factors in the integrand and using (22) gives

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_{0+}(\mathbf{x}' - \mathbf{x}'') + (2\pi)^{-9} \int d^3k' d^3k d^3k'' e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} \times G_{0+}(\mathbf{k}') A(\mathbf{k}' - \mathbf{k}) \{ G_{+n}(\mathbf{k} \mid \mathbf{k}'') + ik'_\alpha [(k'_\alpha - k_\alpha)/(k'_z - k_z)] G_+(\mathbf{k} \mid \mathbf{k}'') \}.$$

Thus for

$$z' > h(\mathbf{x}'_\perp), \quad z'' > h(\mathbf{x}''_\perp),$$

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_{0+}(\mathbf{x}' - \mathbf{x}'') + (2\pi)^{-3} \int d^3k' d^3k'' e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} \times G_{0+}(\mathbf{k}') [T_n(\mathbf{k}' \mid \mathbf{k}'') + ik'_\alpha T_\alpha(\mathbf{k}' \mid \mathbf{k}'')] G_{0+}(\mathbf{k}''). \quad (37)$$

Similarly from (2b) one gets for region II

$$G(\mathbf{x}' \mid \mathbf{x}'') = -R^{-1/2} \int d^2x_\perp \left(R G_{0-}(\mathbf{x}' - \mathbf{x}_s) G_{+n}(\mathbf{x}_s \mid \mathbf{x}'') + n_\alpha(\mathbf{x}_\perp) \frac{\partial}{\partial x'_\alpha} G_{0-}(\mathbf{x}' - \mathbf{x}_s) G_+(\mathbf{x}_s \mid \mathbf{x}'') \right) \quad (38)$$

or in k space

$$G(\mathbf{x}' \mid \mathbf{x}'') = -R^{-1/2}(2\pi)^{-3} \int d^3k' d^3k'' e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} G_{0-}(\mathbf{k}') \times [R T_n(\mathbf{k}' \mid \mathbf{k}'') + ik'_\alpha T_\alpha(\mathbf{k}' \mid \mathbf{k}'')] G_{0+}(\mathbf{k}''). \quad (39)$$

It is interesting to note that (36) and (38) guarantee that $G(\mathbf{x}' \mid \mathbf{x}'')$ satisfies the partial differential Eqs. (1a) and (1b), and the radiation condition. If the upper and lower surface limits are taken, then (34) and (35) are just another way of writing the boundary condition at the surface, one much more useful for present purposes.

In order to determine $g(\mathbf{k})$ and $f(\mathbf{k})$, the Green's function for a flat interface is needed. The Green's function can be written in the form

$$G(\mathbf{x}' \mid \mathbf{x}'') = (2\pi)^{-2} \int d^2k_\perp e^{i\mathbf{k}_\perp \cdot (\mathbf{x}'_\perp - \mathbf{x}''_\perp)} \mathcal{G}(z' \mid z''), \quad (40)$$

where $\mathcal{G}(z' \mid z'')$ satisfies

$$\left(\frac{d^2}{dz'^2} + K_+^2 \right) \mathcal{G}(z' \mid z'') = \delta(z' - z''), \quad z' > 0, \quad (41)$$

$$\left(\frac{d^2}{dz'^2} + K_-^2 \right) \mathcal{G}(z' \mid z'') = 0, \quad z' < 0, \quad (42)$$

where

$$K_\pm = \sqrt{k_\pm^2 - \mathbf{k}_\perp^2 + i\epsilon}$$

and where z'' is assumed to be greater than zero. Also $\mathcal{G}(z' \mid z'')$ satisfies the boundary conditions

$$\lim_{z' \to 0+} \mathcal{G}(z' \mid z'') = R^{-1/2} \lim_{z' \to 0-} \mathcal{G}(z' \mid z''),$$

$$\lim_{z' \to 0+} \frac{d}{dz'} \mathcal{G}(z' \mid z'') = R^{1/2} \lim_{z' \to 0-} \frac{d}{dz'} \mathcal{G}(z' \mid z'').$$

It is easy to verify that (40) satisfies (1a) and (1b) for a flat surface. The solution to (40) and (41) that satisfies the boundary conditions is, for $z'' > z' > 0$

$$\mathcal{G}(z' \mid z'') = \frac{e^{iK_+(z''-z')}}{2iK_+} + \frac{RK_+ - K_-}{RK_+ + K_-} \frac{e^{iK_+(z'+z'')}}{2iK_+} \quad (43)$$

and for $z'' > 0 > z'$

$$\mathcal{G}(z' \mid z'') = R^{1/2} \frac{e^{i(K_+ z'' - K_- z')}}{i(RK_+ + K_-)}. \quad (44)$$

The procedure to be followed is to compare (40), with $\mathcal{G}(z' \mid z'')$ as given by (43) and (44), with $G(\mathbf{x}' \mid \mathbf{x}'')$ as calculated from (37) and (39) using the lowest-order or Born term for $T_n(\mathbf{k}' \mid \mathbf{k}'')$ and $T_\alpha(\mathbf{k}' \mid \mathbf{k}'')$. The Born term is the first term on the right side of (34) and (35). Substitution of this term into (37) and using the fact that

$$A(\mathbf{k}) = (2\pi)^2 \delta_2(\mathbf{k}_\perp),$$

when

$$h(\mathbf{x}_\perp) = 0$$

gives for $z'' > z' > 0$,

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_{0+}(\mathbf{x}' - \mathbf{x}'') + 2i(2\pi)^{-4} \int d^2k'_\perp e^{i\mathbf{k}'_\perp \cdot (\mathbf{x}'_\perp - \mathbf{x}''_\perp)} \times \int dk'_z dk''_z e^{-i(k'_z z' - k''_z z'')} G_{0+}(\mathbf{k}') [g(\mathbf{k}') k''_z + f(\mathbf{k}') k'_z] \times G_{0+}(\mathbf{k}'_\perp, k''_z) \quad (45)$$

and for $z'' > 0 > z'$,

$$G(\mathbf{x}' \mid \mathbf{x}'') = -R^{-1/2} 2i(2\pi)^{-4} \int d^2k'_\perp e^{i\mathbf{k}'_\perp \cdot (\mathbf{x}'_\perp - \mathbf{x}''_\perp)} \times \int dk'_z dk''_z e^{i(k'_z z' - k''_z z'')} G_{0-}(\mathbf{k}') [Rg(\mathbf{k}') k''_z + f(\mathbf{k}') k'_z] G_{0+}(\mathbf{k}'_\perp, k''_z). \quad (46)$$

It will now be assumed that $f(\mathbf{k})$ and $g(\mathbf{k})$ are independent of k_z. This means that the k'_z and k''_z integrations in (45) and (46) can be performed. Using (4) and integrating the remaining terms over k'_z and k''_z gives for $z'' > z' > 0$,

$$G(\mathbf{x}' \mid \mathbf{x}'') = (2\pi)^{-2} \int d^2k_\perp e^{i\mathbf{k}_\perp \cdot (\mathbf{x}'_\perp - \mathbf{x}''_\perp)} \times \left(\frac{e^{iK_+(z''-z')}}{2iK_+} + [f(\mathbf{k}_\perp) - g(\mathbf{k}_\perp)] \frac{e^{iK_+(z''-z')}}{2iK_+} \right) \quad (47)$$

and for $z'' > 0 > z'$,

$$G(\mathbf{x}' \mid \mathbf{x}'') = R^{-1/2}(2\pi)^{-2} \int d^2k_\perp e^{i\mathbf{k}_\perp \cdot (\mathbf{x}'_\perp - \mathbf{x}''_\perp)} \times \frac{e^{i(K_+ z'' - K_- z')}}{2iK_+K_-} [Rg(\mathbf{k}_\perp)K_+ + f(\mathbf{k}_\perp)K_-]. \quad (48)$$

Comparing (47) and (48) with (40), (43), and (44) shows that if

$$f(\mathbf{k}_\perp) = \frac{RK_+}{RK_+ + K_-}, \quad (49)$$

$$g(\mathbf{k}_\perp) = \frac{K_-}{RK_+ + K_-}, \quad (50)$$

with

$$K_\pm = \sqrt{k_\pm^2 - \mathbf{k}_\perp^2 + i\epsilon}, \quad (51)$$

then the lowest-order term in the series expansion for T_n and T_α reduces to the exact result for a flat interface. This means that the solution for an interface of small roughness but with an arbitrarily strong discontinuity will deviate only slightly from the Born term.

In the Introduction it was pointed out that the integral equation for T_n and T_α, whose formulation is now complete, have additional desirable properties which will now be discussed. One nice property is that the Born term is also exact when the media become identical, independent of the roughness of the interface. This can be shown by direct calculation.

When the media are identical

$$R = 1, \quad k_+ = k;$$

thus

$$f(\mathbf{k}_\perp) = g(\mathbf{k}_\perp) = \tfrac{1}{2}.$$

Examination of the kernels appearing in (34) and (35) shows that they are identically zero in this limit and so the Born term is exact. It is not immediately obvious from the resulting expression for the Born term that this is true, so it is explicitly calculated in the Appendix. The result is

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_0(\mathbf{x}' - \mathbf{x}''),$$

where

$$G_0(\mathbf{x}) = (2\pi)^{-3} \int d^3k \, \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{k^2 - \mathbf{k}^2 + i\epsilon}.$$

This means that the higher-order terms are small both for interfaces which have strong discontinuities with small roughness and any roughness with weak discontinuities.

Another desirable property is that the effects of propagation in region II manifestly disappear in the limits which should yield a surface with Dirichlet or Neumann boundary conditions.

The first limit of this type which will be examined is

$$R = Z_+/Z_- \to \infty,$$

which corresponds to waves in region I scattering from a Neumann surface. In this limit $f(\mathbf{k}_\perp) \to 1$ and $g(\mathbf{k}_\perp) \to 0$, such that

$$Rg(\mathbf{k}_\perp) = R[f(\mathbf{k}_\perp) - 1] \to -K_-(\mathbf{k}_\perp)/K_+(\mathbf{k}_\perp)$$

Thus (35) becomes

$$T_n(\mathbf{k} \mid \mathbf{k}'') = O(R^{-1}) - 2(2\pi)^3 \int d^3k_1 [(k_\alpha - k_{1\alpha})/(k_z - k_{1z})] \times A(\mathbf{k} - \mathbf{k}_1) G_{0-\alpha}(\mathbf{k}_1) T_n(\mathbf{k}_1 \mid \mathbf{k}'').$$

The solution of this equation vanishes as R^{-1} since the resolvent is nonsingular. Thus (34) becomes

$$T_\alpha(\mathbf{k} \mid \mathbf{k}'') = 2(2\pi)^{-3} \frac{k_\alpha - k''_\alpha}{k_z - k''_z} A(\mathbf{k} - \mathbf{k}'') + 2(2\pi)^{-3} \int d^3k_1 \frac{k_\alpha - k_{1\alpha}}{k_z - k_{1z}} \times A(\mathbf{k} - \mathbf{k}_1) G_{0+\alpha}(\mathbf{k}_1) T_\alpha(\mathbf{k}_1 \mid \mathbf{k}''). \quad (52)$$

The Green's function for $z' > h(\mathbf{x}'_\perp)$ is given by (37), which becomes in this limit

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_{0+}(\mathbf{x}' - \mathbf{x}'') + (2\pi)^{-3} \int d^3k' d^3k'' e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} G_{0+}(\mathbf{k}') \times ik'_\alpha T_\alpha(\mathbf{k}' \mid \mathbf{k}'') G_{0+}(\mathbf{k}''). \quad (53)$$

Similarly the Green's function for $z' < h(\mathbf{x}'_\perp)$ is given by (39). Using the result that $T_n = O(R^{-1})$, (39) gives

$$G(\mathbf{x}' \mid \mathbf{x}'') = 0, \quad z' < h(\mathbf{x}'_\perp).$$

Thus, as expected, there is not any propagation into region II in this limit.

The next limit which will be examined is

$$R = Z_+/Z_- \to 0,$$

which corresponds to waves in region I scattering from a Dirichlet surface. In this limit $g(\mathbf{k}_\perp) = 1$ and $f(\mathbf{k}_\perp) \to 0$, such that

$$R^{-1} f(\mathbf{k}_\perp) = R^{-1}[g(\mathbf{k}_\perp) - 1] \to -K_+(\mathbf{k}_\perp)/K_-(\mathbf{k}_\perp).$$

Thus (34) becomes

$$T_\alpha(\mathbf{k} \mid \mathbf{k}'') = O(R) - 2(2\pi)^3 \int d^3k_1 [(k_\alpha - k_{1\alpha})/(k_z - k_{1z})] \times A(\mathbf{k} - \mathbf{k}_1) G_{0-\beta}(\mathbf{k}_1) T_\beta(\mathbf{k}_1 \mid \mathbf{k}'').$$

The solution of this equation vanishes as R; thus (35) becomes

$$T_n(\mathbf{k} \mid \mathbf{k}'') = 2i(2\pi)^{-3} \frac{k_\alpha - k''_\alpha}{k_z - k''_z} A(\mathbf{k} - \mathbf{k}'') + 2(2\pi)^{-3} \int d^3k_1 \frac{k_\alpha - k_{1\alpha}}{k_z - k_{1z}} \times A(\mathbf{k} - \mathbf{k}_1) G_{0+\alpha}(\mathbf{k}_1) T_n(\mathbf{k}_1 \mid \mathbf{k}''). \quad (54)$$

The Green's function for $z' > h(\mathbf{x}'_\perp)$ is given by (37), which becomes in this limit

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_{0+}(\mathbf{x}' - \mathbf{x}'') + (2\pi)^{-3} \int d^3k' d^3k'' e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}')} \times G_{0+}(\mathbf{k}') T_n(\mathbf{k}' \mid \mathbf{k}'') G_{0+}(\mathbf{k}''). \quad (55)$$

Similarly the Green's function for $z' < h(\mathbf{x}'_\perp)$ is given by (39). Using the result that $T_\alpha = O(R)$, (39) gives

$$G(\mathbf{x}' \mid \mathbf{x}'') = 0, \quad z' < h(\mathbf{x}'_\perp). \quad (56)$$

Again, as expected, there is not any propagation into region II in this limit.

The last limit which will be examined is R finite and $k_-^2 \to \infty$. This also yields a Dirichlet surface. In this limit

$$f(\mathbf{k}_\perp) = 0, \quad g(\mathbf{k}_\perp) = 1. \quad (57)$$

Examination of (34) shows that the kernels and the inhomogeneous term vanish in this limit and that

$$T_\alpha(\mathbf{k} \mid \mathbf{k}'') \to 0 \text{ as } k_-^2 \to 0 \quad (58)$$

if the integrations indicated in (34) are well-defined in this limit. This limit is rather tricky, and this is really only a reasonable argument for (57). Because of (57) and (58), this case reduces to the previous case where $R = 0$, and Eqs. (54)–(56), therefore, are also the limiting equations when $k_-^2 \to \infty$.

This concludes the basic formulation of the problem. The important results which will be used in the next section are Eqs. (34), (35), (37), (39), (49), and (50).

III. DIAGRAMMATIC RULES

The integral equations (34) and (35) can be written in a compact form which makes transparent the construction of the diagrammatic rules for an arbitrary term in the series resulting from the iteration of (34) and (35). It is convenient to define several new quantities. Let

$$\tau_l(\mathbf{k} \mid \mathbf{k}') = T_l(\mathbf{k} \mid \mathbf{k}'), \quad l = 1, 2, 3,$$
$$\tau_4(\mathbf{k} \mid \mathbf{k}') = T_n(\mathbf{k} \mid \mathbf{k}'), \quad (59)$$

$$\mathcal{P}^{\frac{1}{2}\frac{1}{2}}_{lq}(\mathbf{k}) = \begin{cases} G_{0\pm l q}(\mathbf{k}) f(\mathbf{k}_\perp), & l, q = 1, 2, 3 \\ R^{-1/2\pm 1/2} G_{0\pm q}(\mathbf{k}) f(\mathbf{k}_\perp), & l = 4, q = 1, 2, 3, \\ G_{0\pm l}(\mathbf{k}) g(\mathbf{k}_\perp), & l = 1, 2, 3, q = 4, \\ R^{1/2\mp 1/2} G_{0\pm}(\mathbf{k}) g(\mathbf{k}), & l = q = 4, \end{cases} \quad (60)$$

where the double superscripts on the left are associated with the double signs on the right of (60);

$$\mathcal{P}^{12}_{lq}(\mathbf{k}) = \begin{cases} -G_{0\pm lq}(\mathbf{k})g(\mathbf{k}_\perp), & l,q = 1,2,3, \\ -R^{-1/2\pm 1/2}\,G_{0\pm q}(\mathbf{k})g(\mathbf{k}_\perp), & l = 4, q = 1,2,3, \\ -G_{0\pm l}(\mathbf{k})f(\mathbf{k}_\perp), & l = 1,2,3, q = 4, \\ -R^{1/2\mp 1/2}\,G_{0\pm}(\mathbf{k})f(\mathbf{k}_\perp), & l = q = 4, \end{cases} \tag{61}$$

$$V_{4l}(\mathbf{k}) = V_{l4}(\mathbf{k}) = 2(2\pi)^{-3}k_l/k_z, \quad l = 1,2,3, \tag{62}$$

with all other components of $V_{lq}(\mathbf{k})$ being zero, let

$$\Psi^1_{q} = \begin{cases} \pm iR^{1/4\pm 1/4}k_q, & q = 1,2,3, \\ \pm R^{1/4\mp 1/4}, & q = 4, \end{cases} \tag{63}$$

and finally let

$$\begin{aligned} \mathfrak{F}^1_{lq}(\mathbf{k}) &= \begin{cases} f(\mathbf{k}_\perp)\delta_{lq}, & l,q = 1,2,3, \\ g(\mathbf{k}_\perp) = 1 - f(\mathbf{k}_\perp), & l = 4, q = 4, \end{cases} \\ \mathfrak{F}^1_{4q}(\mathbf{k}) &= \mathfrak{F}^1_{q4}(\mathbf{k}) = 0, \qquad q = 1,2,3, \\ \mathfrak{F}^2_{lq}(\mathbf{k}) &= \delta_{lq} - \mathfrak{F}^1_{lq}(\mathbf{k}). \end{aligned} \tag{64}$$

With these definitions (34) and (35) become

$$\tau_l(\mathbf{k} \mid \mathbf{k}') = \sum_{p,q=1}^{4} \sum_{i=1}^{2} \mathfrak{F}^i_{lp}(\mathbf{k})\,\mathfrak{D}^{i1}_{pq}(\mathbf{k} \mid \mathbf{k}')\Psi^1_q(\mathbf{k}'), \tag{65}$$

where $\mathfrak{D}^{ij}_{pq}$ satisfies the integral equation

$$\begin{aligned} \mathfrak{D}^{ij}_{pq}(\mathbf{k} \mid \mathbf{k}'') &= V_{pq}(\mathbf{k} - \mathbf{k}'')\,\delta^{ij}A(\mathbf{k} - \mathbf{k}'') \\ &+ \sum_{s,l=1}^{4} \sum_{m,n=1}^{2} \int d^3k_1\, V_{ps}(\mathbf{k} - \mathbf{k}_1)\delta^{im}A(\mathbf{k} - \mathbf{k}_1) \\ &\times \mathcal{P}^{mn}_{sl}(\mathbf{k}_1)\,\mathfrak{D}^{nj}_{lq}(\mathbf{k}_1 \mid \mathbf{k}''). \end{aligned} \tag{66}$$

The connection between τ_l and the full Green's function can also be written in a compact form. Equations (37) and (39) together with the definitions just introduced give for region I,

$$\begin{aligned} G(\mathbf{x}' \mid \mathbf{x}'') &= G_{0+}(\mathbf{x}' - \mathbf{x}'') + (2\pi)^{-3}\int d^3k'\,d^3k''\,e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} \\ &\times G_{0+}(\mathbf{k}')\,S^1(\mathbf{k}' \mid \mathbf{k}'')\,G_{0+}(\mathbf{k}''), \end{aligned} \tag{67}$$

and for region II,

$$\begin{aligned} G(\mathbf{x}' \mid \mathbf{x}'') &= (2\pi)^{-3}\int d^3k'd^3k''\;e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} \\ &\times G_{0-}(\mathbf{k}')\,S^2(\mathbf{k}' \mid \mathbf{k}'')\,G_{0+}(\mathbf{k}''), \end{aligned} \tag{68}$$

where

$$S^j(\mathbf{k} \mid \mathbf{k}'') = \sum_{l,p,q=1}^{4} \sum_{i=1}^{2} \Psi^j_l(\mathbf{k})\,\mathfrak{F}^i_{lp}(\mathbf{k})\,\mathfrak{D}^{i1}_{pq}(\mathbf{k} \mid \mathbf{k}'')\Psi^1_q(\mathbf{k}''). \tag{69}$$

Given the form of (66), the construction of the diagrammatic rules for $\mathfrak{D}$ is straightforward. The rules follow from iterating (66), and they are shown in Fig. 2. Summation of the set of diagrams shown in Fig. 3 gives $\mathfrak{D}^{ij}_{pq}(\mathbf{k} \mid \mathbf{k}'')$.

While the Green's function contains all the information needed to construct the solution corresponding to an arbitrary distribution of sources in region I, it is often true that the sources and detectors are far enough removed from the interface that they can be taken to be at infinity. In this case (67) and (68) simplify considerably since the problem is reduced to that of a superposition of plane waves incident on the interface. The reduction formula which connects the asymptotic scattered field with the incident field follows directly from rewriting (67) and (68) and then using (4). Equation (67) can be written in coordinate space as

$$\begin{aligned} G(\mathbf{x}' \mid \mathbf{x}'') &= G_{0+}(\mathbf{x}' - \mathbf{x}'') \\ &+ (2\pi)^3 \int d^3x_1 d^3x_2\, G_{0+}(\mathbf{x}' - \mathbf{x}_1)\,S^1(\mathbf{x}_1 \mid \mathbf{x}_2)\,G_{0+}(\mathbf{x}_2 - \mathbf{x}''). \end{aligned}$$

Using (4) with $z'' > z' > h(\mathbf{x}'_\perp)$, gives for region I,

$$\begin{aligned} G(\mathbf{x}' \mid \mathbf{x}'') &= \int d^2k'_\perp d^2k''_\perp\, e^{i\mathbf{k}'_\perp\cdot\mathbf{x}'_\perp} \\ &\times [e^{-ik'_z z'}\,\delta_2(\mathbf{k}'_\perp - \mathbf{k}''_\perp) - i\pi e^{ik'_z z'}(1/k'_z)\,S^1(\mathbf{k}' \mid \mathbf{k}'')] \\ &\times i\,e^{i\mathbf{k}''\cdot\mathbf{x}''}/8\pi^2 k''_z, \end{aligned} \tag{70}$$

where

$$\begin{aligned} k'_z &= \sqrt{k^2_+ - \mathbf{k}'^2_\perp + i\epsilon}, \\ k''_z &= -\sqrt{k^2_+ - \mathbf{k}''^2_\perp + i\epsilon}. \end{aligned} \tag{71}$$

Similarly (68) can be written in coordinate space as

$$\begin{aligned} &G(\mathbf{x}' \mid \mathbf{x}'') \\ &= (2\pi)^3 \int d^3x_1 d^3x_2\, G_{0-}(\mathbf{x}' - \mathbf{x}_1)\,S^2(\mathbf{x}_1 \mid \mathbf{x}_2)\,G_{0+}(\mathbf{x}_2 - \mathbf{x}''). \end{aligned}$$

Using (4) with $z'' > h(\mathbf{x}'_\perp) > z'$ gives for region II,

$$\begin{aligned} G(\mathbf{x}' \mid \mathbf{x}'') &= \int d^2k'_\perp\, d^2k''_\perp\, e^{i\mathbf{k}'_\perp\cdot\mathbf{x}'_\perp} \\ &\times [i\pi e^{-ik'_z z'}(1/k'_z)\,S^2(\mathbf{k}' \mid \mathbf{k}'')]\,i\,e^{i\mathbf{k}''\cdot\mathbf{x}''}/8\pi^2 k''_z, \end{aligned} \tag{72}$$

where

$$\begin{aligned} k'_z &= \sqrt{k^2_- - \mathbf{k}'^2_\perp + i\epsilon}, \\ k''_z &= \sqrt{k^2_+ - \mathbf{k}''^2_\perp + i\epsilon}. \end{aligned} \tag{73}$$

The scattering matrix $S^j(\mathbf{k}_{out} \mid \mathbf{k}_{in})$ which connects an asymptotic incoming plane wave in region I with asymptotic outgoing plane waves in regions I or II ($j = 1$ or 2, respectively) can be directly read from (70) and (72). For asymptotic states with the normalization

$$\int d^2x_\perp\, \varphi^*_{\mathbf{k}'_\perp}(\mathbf{x}_\perp, z)\,\varphi_{\mathbf{k}''_\perp}(\mathbf{x}_\perp, z) = \delta_2(\mathbf{k}'_\perp - \mathbf{k}''_\perp) \tag{74}$$

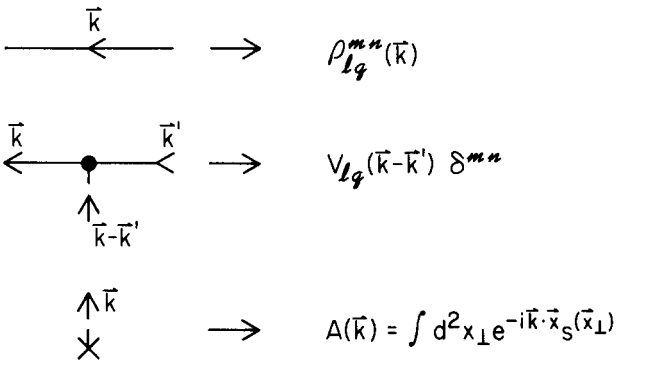

FIG. 2. The propagator $\mathcal{P}^{mn}_{lq}(\mathbf{k})$ is defined by (60) and (61). In (60) and (61) the quantities $G_{0\pm}$, $G_{0\pm\alpha}$, $G_{0\pm\alpha\beta}$, f, and g appear and they are defined in (24)–(26) and (49)–(51). The vertex function $V_{lq}(\mathbf{k})$ is defined in (63). Multiplication of the factors constituting a diagram, integration over the internal momenta associated with $\mathcal{P}^{mn}_{lq}$, and summation over internal indices yields the value of the diagram.

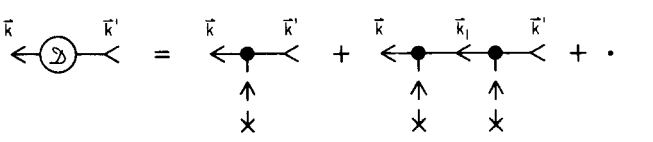

FIG. 3. Diagrammatic representation of the iterative solution of (66) for $\mathfrak{D}$. The full Green's function can be calculated from $\mathfrak{D}$ using (67)–(69), (63), and (64).

the scattering matrix in region I is

$$S^1(\mathbf{k}_{out} \mid \mathbf{k}_{in}) = -i\pi(\mathbf{k}_{out} \cdot \hat{i}_z)^{-1} S^1(\mathbf{k}_{out} \mid \mathbf{k}_{in}), \quad (75)$$

where

$$\mathbf{k}_{in} = \mathbf{k}_{\perp in} - \sqrt{k_+^2 - \mathbf{k}_{\perp in}^2}\,\hat{i}_z,$$
$$\mathbf{k}_{out} = \mathbf{k}_{\perp out} + \sqrt{k_+^2 - \mathbf{k}_{\perp out}^2}\,\hat{i}_z; \quad (76)$$

and in region II is

$$S^2(\mathbf{k}_{out} \mid \mathbf{k}_{in}) = i\pi(\mathbf{k}_{out} \cdot \hat{i}_z)^{-1} S^2(\mathbf{k}_{out} \mid \mathbf{k}_{in}) \quad (77)$$

where

$$\mathbf{k}_{out} = \mathbf{k}_{\perp out} - \sqrt{k_+^2 - \mathbf{k}_{\perp out}^2}\,\hat{i}_z. \quad (78)$$

That is, if the incident wave is of the form

$$\varphi_{in}(\mathbf{x}) = \int d^2k_\perp\, a_{in}(\mathbf{k}_\perp)\, e^{i(\mathbf{k}_\perp \cdot \mathbf{x}_\perp - \sqrt{k_+^2 - \mathbf{k}_\perp^2}\, z)} \quad (79)$$

and the scattered wave in region I is of the form

$$\varphi^1_{out}(\mathbf{x}) = \int d^2k_\perp\, a^1_{out}(\mathbf{k}_\perp)\, e^{i(\mathbf{k}_\perp \cdot \mathbf{x}_\perp + \sqrt{k_+^2 - \mathbf{k}_\perp^2}\, z)} \quad (80)$$

and in region II is of the form

$$\varphi^2_{out}(\mathbf{x}) = \int d^2k_\perp\, a^2_{out}(\mathbf{k})\, e^{i(\mathbf{k}_\perp \cdot \mathbf{x}_\perp - \sqrt{k_-^2 - \mathbf{k}_\perp^2}\, z)}, \quad (81)$$

then

$$a^j_{out}(\mathbf{k}_\perp) = S^j_{+\epsilon}(\mathbf{k} \mid \mathbf{k}')\, a_{in}(\mathbf{k}'), \quad j = 1, 2, \quad (82)$$

where $+\epsilon$ indicates an outgoing propagator, and

$$k_z = \pm\sqrt{k_\pm^2 - \mathbf{k}_\perp^2}, \quad j = \{\tfrac{1}{2},$$

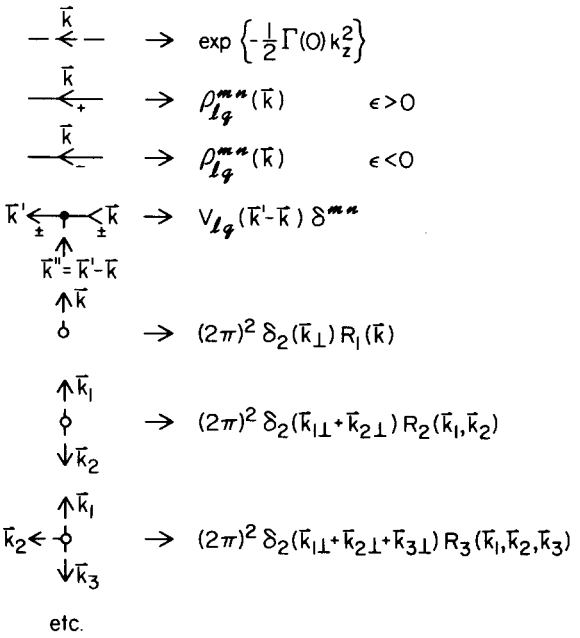

FIG. 4. Diagrammatic rules for $\langle \pi \mathfrak{D} \rangle$. Each diagram in the series has equal weight and is constructed by multiplying the indicated factors associated with the lines and vertices of the diagrams. Integration over three momenta k associated with internal $\mathcal{P}$ lines and summation over internal indices completes the construction of the term in the series associated with the diagram. The propagator $\mathcal{P}^{mn}_{lq}(\mathbf{k})$ is defined by (60) and (61). In (60) and (61) the quantities $G_{0\pm}$, $G_{0\pm\alpha}$, $G_{0\pm\alpha\beta}$, f and g appear and they are defined in (24)–(26) and (49)–(51). The vertex function $V_{lq}(\mathbf{k})$ is defined in (63). The vertex function $R_n(\{\mathbf{k}\}_n)$ is defined in (88)–(90) and explicitly displayed in (81)–(83) for $n = 1, 2, 3$.

$$k'_z = -\sqrt{k_-^2 - \mathbf{k}_\perp'^2}.$$

Since the complex conjugate solution $G^*(\mathbf{x} \mid \mathbf{x}')$ is needed in order to be able to calculate intensities, it is necessary to state the rules for construction of $G^*(\mathbf{x} \mid \mathbf{x}')$. Examination of (1a) and (1b) shows that $G^*(\mathbf{x} \mid \mathbf{x}')$ and $G(\mathbf{x} \mid \mathbf{x}')$ satisfy the same partial differential equation. The only difference is in the radiation condition at infinity. Thus the integral equation for $G^*(\mathbf{x} \mid \mathbf{x}')$ is the same as that for $G(\mathbf{x} \mid \mathbf{x}')$ except that ϵ is replaced by $-\epsilon$. It is also convenient to change the sign convention on the Fourier transform. Thus (18) becomes

$$Z^*(\mathbf{x}' \mid \mathbf{x}'') = (2\pi)^{-6} \int d^3k' d^3k''\, e^{i(\mathbf{k}' \cdot \mathbf{x}' - \mathbf{k}'' \cdot \mathbf{x}'')} \times Z(-\mathbf{k}' \mid -\mathbf{k}'') \quad \text{as} \quad \epsilon \to -\epsilon, \quad (83)$$

and similarly for (19) and (20). With these changes the manipulations leading to the series expansion for $G^*(\mathbf{x}' \mid \mathbf{x}'')$ are identical to those already developed for $G(\mathbf{x}' \mid \mathbf{x}'')$. This completes the discussion of scattering from a deterministic surface.

The statistical properties of the fields scattered from a random interface will now be considered. It will be assumed that $h(\mathbf{x}_\perp)$ has multivariate Gaussian statistics and that averages over $\mathbf{x}_\perp$ can be replaced by ensemble averages. The statistical quantities which are of interest are the correlation functions of the scattered field. Since the scattered field can be expressed in terms of $S^j(\mathbf{k} \mid \mathbf{k}')$ via (67) and (68) or (75) and (77), it is only necessary to consider ensemble averages of the form

$$\langle \prod_j S^j_{+\epsilon}(\mathbf{k}_l \mid \mathbf{k}_m) \prod_{j'} S^{j'}_{-\epsilon}(\mathbf{k}_{l'} \mid \mathbf{k}_{m'}) \rangle,$$

where the $\pm\epsilon$ indicates the propagator to be used. A cluster expansion for these expectation values can be easily found because $S^j(\mathbf{k} \mid \mathbf{k}')$ is expressible as a series where the only functional dependence on $h(\mathbf{x}_\perp)$ arises through $A(\mathbf{k})$ whose cluster expansion is well-known.[1,8] Due to the structural similarity of the series expansion for $\mathfrak{D}^{ij}_{pq}(\mathbf{k} \mid \mathbf{k}')$ to the series expansion for scattering from a Neumann surface,[1] the diagrammatic rules for moments of $\mathfrak{D}^{ij}_{pq}(\mathbf{k} \mid \mathbf{k}')$ can be easily written down and are shown in Fig. 4. The vertex functions R_1, R_2, and R_3 are given by

$$R_1(\mathbf{k}) = 1, \quad (84)$$

$$R_2(\mathbf{k}_1, \mathbf{k}_2) = \int d^2y_\perp \exp(-i\mathbf{k}_{2\perp} \cdot \mathbf{y}_\perp) \times \{\exp[-\Gamma(\mathbf{y}_\perp) k_{1z} k_{2z}] - 1\}, \quad (85)$$

$$R_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \int d^2y_{1\perp} d^2y_{2\perp} \exp[-i(\mathbf{k}_{1\perp} \cdot \mathbf{y}_{1\perp} + \mathbf{k}_{2\perp} \cdot \mathbf{y}_{2\perp})] \times \{\exp[-\Gamma(\mathbf{y}_{1\perp} - \mathbf{y}_{2\perp}) k_{1z} k_{2z} - \Gamma(\mathbf{y}_{1\perp}) k_{1z} k_{3z} - \Gamma(\mathbf{y}_{2\perp}) k_{2z} k_{3z}] - \exp[-\Gamma(\mathbf{y}_{1\perp} - \mathbf{y}_{2\perp}) k_{1z} k_{2z}] - \exp[-\Gamma(\mathbf{y}_{1\perp}) k_{1z} k_{3z}] - \exp[-\Gamma(\mathbf{y}_{2\perp}) k_{2z} k_{3z}] + 2\}, \quad (86)$$

where

$$\Gamma(\mathbf{x}_{1\perp} - \mathbf{x}_{2\perp}) = \langle h(\mathbf{x}_{1\perp}) h(\mathbf{x}_{2\perp}) \rangle \quad (87)$$

is the height correlation function for the surface. The general expression for R_n is[1]

$$R_n(\{\mathbf{k}\}_n) = \exp\left[\tfrac{1}{2}\Gamma(0) \sum_{j=1}^n k_{jz}^2\right] \times \int d^2y_{1\perp} \cdots d^2y_{n\perp}\, \delta_2(y_{n\perp}) \exp\left(-i \sum_{j=1}^n \mathbf{k}_{j\perp} \cdot \mathbf{y}_{j\perp}\right) \times K_n(\{k_z, \mathbf{y}_\perp\}_n), \quad (88)$$

where $\{k\}_n$ indicates a set with n members, and where K_n is defined recursively by

$$K_n(\{k_z, x_\perp\}_n) = F_n(\{k_z, x_\perp\}_n) - \sum_{j \text{ perm}} \sum_{M=2}^{n} \sum_{\{m_i\}_M} \prod_{i=1}^{M} K_{m_i}(\{k_{jz}, x_{j\perp}\}_{m_j}), \quad (89)$$

where

$$F_n(\{k_z, x_\perp\}_n) \exp\left[-\tfrac{1}{2} \sum_{j,l=1}^{n} k_{jz} \Gamma(x_{j\perp} - x_{l\perp}) k_{lz}\right]. \quad (90)$$

The sum $\sum_{\{m_i\}_M}$ is over all unordered M element sets $\{m_i\}_M$ such that $\sum_{i=1}^{M} m_i = n$ and $\sum_{j \text{ perm}}$ denotes a sum over all different labelings j of the unordered sets $\{k_{jz}, x_{j\perp}\}_{m_i}$ with $j = 1, \cdots, n$. The vertex function R_n can be rather complicated for large n. In practice only a few lowest-order vertex functions would be retained.

Once the moments of $\mathfrak{D}^{ij}_{pq}(k,k')$ are known, the moments of the Green's function can be found. It is necessary to insert a word of caution about this procedure. In the ensemble of multivariate Gaussian surfaces which is being considered here, there always can be found a surface whose excursions are very large. This means that on ensemble averaging the condition that the source be in region I will be violated to a degree depending on the distance from the surface to the source and detector. If the source and detector are removed far from the surface then this will occur with a vanishingly small probability. In most experimental situations the source and detector are completely in one or the other regions. This apparent difference is due to the finite statistical sample inherent in an experiment, and any comparison which is sensitive to this difference should be ignored.

As an example of the application of these rules the incoherent back-scattered and transmitted intensity in the asymptotic region will be calculated in the lowest-order approximation. The incident wave normalization will be chosen to be

$$\int d^2x_\perp \, \varphi^*_{(\text{in})k_\perp}(x) \, \varphi_{(\text{in})k'_\perp}(x) = \delta_2(k_\perp - k'_\perp).$$

This means that

$$a_{\text{in}}(k_\perp) = (2\pi)^{-1} \delta_2(k_\perp - k_{(\text{in})\perp})$$

in Eq. (79). Equations (81) and (82) then give

$$\langle \varphi^{j*}_{\text{out}}(x) \, \varphi^{j}_{\text{out}}(x') \rangle = (2\pi)^{-2} \int d^2k_\perp d^2k'_\perp \times e^{i(k\cdot x - k'\cdot x')} \langle \mathcal{S}^{j}_{+\epsilon}(k \mid k_{\text{in}}) \, \mathcal{S}^{j}_{\epsilon}(-k \mid -k_{\text{in}}) \rangle. \quad (91)$$

Now $\mathcal{S}$ can be related to $\mathfrak{D}$ through (75), (77), and (69), thus

$$\langle \mathcal{S}^{j}_{+\epsilon}(k \mid k_{\text{in}}) \, \mathcal{S}^{j}_{\epsilon}(-k' \mid -k_{\text{in}}) \rangle = \frac{\pi^2}{k_z^2} \sum_{l,p,q=1}^{4} \sum_{i=1}^{2} \Psi^{j}_{l}(k) \, \mathfrak{F}^{i}_{lp}(k) \, \Psi^{1}_{q}(k_{\text{in}}) \times \sum_{l',p',q'=1}^{4} \sum_{i'=1}^{2} \Psi^{j'}_{l'}(-k') \, \mathfrak{F}^{i}_{lp}(-k') \, \Psi^{1}_{q}(-k_{\text{in}}) \times \langle \mathfrak{D}^{i1}_{pq}(k \mid k_{\text{in}})_{+\epsilon} \mathfrak{D}^{i'1}_{p'q'}(-k' \mid -k_{\text{in}})_{-\epsilon} \rangle. \quad (92)$$

Applying the rules just derived to the lowest-order approximation for the incoherent part of $\langle \mathfrak{D}_{+\epsilon} \mathfrak{D}_{-\epsilon} \rangle$ which corresponds to the diagram shown in Fig. 5, gives

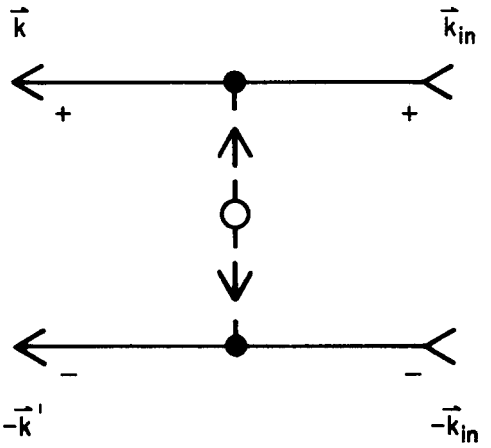

FIG. 5. The diagram corresponding to the lowest-order incoherently scattered intensity.

$$\langle \mathfrak{D}^{i1}_{pq}(k \mid k_{\text{in}})_{+\epsilon} \mathfrak{D}^{i'1}_{p'q'}(-k' \mid -k_{\text{in}})_{-\epsilon} \rangle = (2\pi)^2 \delta_2(k_\perp - k'_\perp) e^{-\Gamma(0)(k-k_{\text{in}})^2} \delta^{i1} \delta^{i'1} \times R_2(k - k_{\text{in}}, -k + k_{\text{in}}) \times V_{pq}(k - k_{\text{in}}) \, V_{p'q'}(-k' + k_{\text{in}}).$$

Substituting this into (91) and (92), then using (62)–(64) and (85) gives for the mutual coherence function

$$\langle \varphi^{\overset{1}{2}*}_{\text{out}}(x) \, \varphi^{\overset{1}{2}}_{\text{out}}(x') \rangle = (2\pi)^{-2} \int d^2k_\perp \, e^{ik\cdot(x-x')} k_z^{-2} \times e^{-\Gamma(0)(k_z - k_{(\text{in})z})^2} \left(R^{-1/4 \pm 1/4} \frac{k \cdot (k - k_{\text{in}})}{k_z - k_{(\text{in})z}} f(k_\perp) + R^{1/4 \mp 1/4} \frac{(k - k_{\text{in}}) \cdot k_{\text{in}}}{k_z - k_{(\text{in})z}} g(k_\perp) \right)^2 (2\pi)^{-2} \times \int d^2y_\perp \, e^{i(k_\perp - k_{(\text{in})\perp})\cdot y_\perp} (e^{\Gamma(y_\perp)(k_z - k_{(\text{in})z})^2} - 1).$$

The mutual coherence function is related to the intensity scattered in a particular direction by

$$\langle \varphi^{j*}_{\text{out}}(x) \, \varphi^{j}_{\text{out}}(x') \rangle = (2\pi)^{-2} \int d^2k_\perp \, e^{ik\cdot(x-x')} I^{j}(k).$$

Thus

$$I^{\overset{1}{2}}(k) = [k_z(k_z - k_{(\text{in})z})]^{-2} [R^{-1/4 \pm 1/4} k \cdot (k - k_{\text{in}}) f(k_\perp) + R^{1/4 \mp 1/4} (k - k_{\text{in}}) \cdot k_{\text{in}} g(k_\perp)]^2 \times (2\pi)^{-2} \int d^2y_\perp e^{i(k_\perp - k_{(\text{in})\perp})\cdot y_\perp} (e^{\Gamma(y_\perp)(k_z - k_{(\text{in})z})^2} - 1), \quad (93)$$

where $R = Z_+/Z_-$ and (49)–(51) give $f(k_\perp)$ and $g(k_\perp)$. In a similar way the lowest-order coherently scattered intensity, corresponding to the diagram shown in Fig. 6, can be shown to be given by

$$I^{\overset{1}{2}}(k) = [k_z(k_z - k_{(\text{in})z})]^{-2} [R^{-1/4 \pm 1/4} k \cdot (k - k_{\text{in}}) f(k_\perp) + R^{1/4 \mp 1/4} (k - k_{\text{in}}) \cdot k_{\text{in}} g(k_\perp)]^2 \, \delta_2(k_\perp - k_{(\text{in})\perp}). \quad (94)$$

Due to the fact that the diagrammatic rules have the same connectivity structure as those for the hard surface,[1] the classification procedure in terms of connected and disconnected diagrams and partial summation in terms of linear integral equations can be carried out in an identical manner here. A diagram is disconnected if it can be broken into two parts not connected to each other by any lines, by removing one propagator line $\mathcal{P}^{mn}_{lq}(k)$ from each of the solid lines that flow continuously

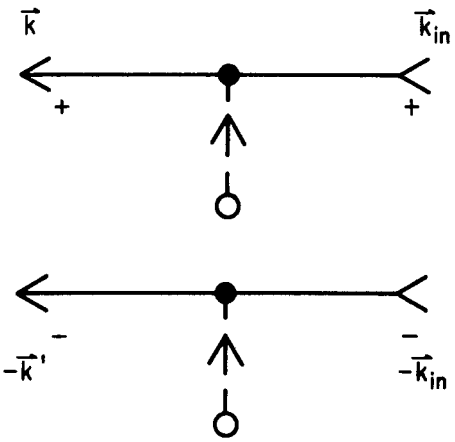

FIG. 6. The diagram corresponding to the lowest-order coherently scattered intensity.

through the diagram. A connected diagram is one that is not disconnected.

Examination of the series for $\langle \mathfrak{D} \rangle$ shows that one can partially sum the series to give the integral equation

$$\langle \mathfrak{D}^{ij}_{pq}(\mathbf{k}'|\mathbf{k}'')\rangle = \mathcal{C}^{ij}_{pq}(\mathbf{k}'|\mathbf{k}'') + \sum_{m,n=1}^{2}\sum_{l,s=1}^{4}\int d^3k\ \mathcal{C}^{in}_{pl}(\mathbf{k}'|\mathbf{k})\ \mathcal{P}^{nm}_{ls}(\mathbf{k})\langle \mathfrak{D}^{mj}_{sq}(\mathbf{k}|\mathbf{k}'')\rangle \tag{95}$$

where $\mathcal{C}^{ij}_{pq}(\mathbf{k}|\mathbf{k}')$ is the sum of the connected parts of $\langle \mathfrak{D}^{ij}_{pq}(\mathbf{k}|\mathbf{k}')\rangle$. Actually (95) is a one-dimensional integral equation because $\mathcal{C}^{ij}_{pq}(\mathbf{k}|\mathbf{k}')$ contains a factor of $\delta_2(\mathbf{k}_\perp - \mathbf{k}'_\perp)$. This can be seen directly from the rules for $\langle \mathfrak{D} \rangle$ and is due to the translational invariance of the surface statistics in the $\mathbf{x}_\perp$ plane. Letting

$$\mathcal{C}^{ij}_{pq}(\mathbf{k}|\mathbf{k}') = \delta_2(\mathbf{k}_\perp - \mathbf{k}'_\perp)\, C^{ij}_{pq}(k_z|k'_z),$$

$$\langle \mathfrak{D}^{ij}_{pq}(\mathbf{k}|\mathbf{k}')\rangle = \delta_2(\mathbf{k}_\perp - \mathbf{k}'_\perp) D^{ij}_{pq}(k_z|k'_z),$$

(95) then gives

$$D^{ij}_{pq}(k'_z|k''_z) = C^{ij}_{pq}(k'_z|k''_z) + \sum_{m,n=1}^{2}\sum_{l,s=1}^{4}\int dk_z C^{in}_{pl}(k'_z|k_z)\,\mathcal{P}^{nm}_{ls}(k_z)\,D^{mj}_{sq}(k_z|k''_z), \tag{96}$$

where the $\mathbf{k}_\perp$ dependence of the quantities appearing in (96) has not been shown since $\mathbf{k}_\perp$ is no more than a parameter in (96).

Similarly examination of the series for $\langle \mathfrak{D}\mathfrak{D} \rangle$ shows that one can partially sum the series to give the integral equation

$$\begin{aligned}\langle \mathfrak{D}^{lm}_{pq}(\mathbf{k}|\mathbf{k}_i)\mathfrak{D}^{l'm'}_{p'q'}(\mathbf{k}'|\mathbf{k}'_i)\rangle &= \langle \mathfrak{D}^{lm}_{pq}(\mathbf{k}|\mathbf{k}_i)\rangle\langle \mathfrak{D}^{l'm'}_{p'q'}(\mathbf{k}'|\mathbf{k}'_i)\rangle \\ &+ \sum_{n,n'=1}^{2}\sum_{s,s'=1}^{4}\int d^3k_2 d^3k'_2\ \mathcal{K}^{ll';nn'}_{pp';ss'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_2,\mathbf{k}'_2) \\ &\times\Big\{\delta^{nm}\delta_{sq}\delta_3(\mathbf{k}_2-\mathbf{k}_i)\delta^{n'm'}\delta_{s'q'}\delta_3(\mathbf{k}'_2-\mathbf{k}'_i) \\ &+ \sum_{j=1}^{2}\sum_{r=1}^{4}\big[\mathcal{P}^{nj}_{sr}(\mathbf{k}_2)\langle \mathfrak{D}^{jm}_{rq}(\mathbf{k}_2|\mathbf{k}_i)\rangle\delta^{n'm'}\delta_{s'q'}\delta_3(\mathbf{k}'_2-\mathbf{k}'_i) \\ &+ \delta^{nm}\delta_{sq}\delta_3(\mathbf{k}_2-\mathbf{k}_i)\mathcal{P}^{n'j}_{s'r}(\mathbf{k}'_2)\langle \mathfrak{D}^{jm'}_{rq'}(\mathbf{k}'_2|\mathbf{k}'_i)\rangle \\ &+ \sum_{j'=1}^{2}\sum_{r'=1}^{4}\mathcal{P}^{nj}_{sr}(\mathbf{k}_2)\mathcal{P}^{n'j'}_{s'r'}(\mathbf{k}'_2) \\ &\times\langle \mathfrak{D}^{jm}_{rq}(\mathbf{k}_2|\mathbf{k}_i)\mathfrak{D}^{j'm'}_{r'q'}(\mathbf{k}'_2|\mathbf{k}'_i)\rangle\big]\Big\},\end{aligned} \tag{97}$$

where

$$\begin{aligned}\mathcal{K}^{ll';mm'}_{pp';qq'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_2,\mathbf{k}'_2) &= \sum_{j,j'=1}^{2}\sum_{r,r'=1}^{4}\int d^3k_1 d^3k'_1 \\ &\times\Big(\sum_{n=1}^{2}\sum_{s=1}^{4}\langle \mathfrak{D}^{ln}_{ps}(\mathbf{k}|\mathbf{k}_1)\rangle\,\mathcal{P}^{nj}_{sr}(\mathbf{k}_1) + \delta^{lj}\delta_{pr}\delta_3(\mathbf{k}-\mathbf{k}_1)\Big) \\ &\times\Big(\sum_{n'=1}^{2}\sum_{s'=1}^{4}\langle \mathfrak{D}^{l'n'}_{p's'}(\mathbf{k}'|\mathbf{k}'_1)\rangle\,\mathcal{P}^{n'j'}_{s'r'}(\mathbf{k}'_1) \\ &+ \delta^{l'j'}\delta_{p'r'}\delta_3(\mathbf{k}'-\mathbf{k}'_1)\Big) \\ &\times\mathcal{C}^{jj';mm'}_{rr';qq'}(\mathbf{k}_1,\mathbf{k}'_1|\mathbf{k}_2,\mathbf{k}'_2),\end{aligned} \tag{98}$$

with $\mathcal{C}^{jj';mm'}_{rr';qq'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_1,\mathbf{k}'_1)$ being the connected part of $\langle \mathfrak{D}^{jm}_{rq}(\mathbf{k}|\mathbf{k}_1)\,\mathfrak{D}^{j'm'}_{r'q'}(\mathbf{k}'|\mathbf{k}'_1)\rangle$. Equation (97) can be put in a better form by letting

$$\begin{aligned}\mathcal{T}^{ll';mm'}_{pp';qq'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_i,\mathbf{k}'_i) &= \langle \mathfrak{D}^{lm}_{pq}(\mathbf{k}|\mathbf{k}_i)\,\mathfrak{D}^{l'm'}_{p'q'}(\mathbf{k}|\mathbf{k}'_i)\rangle \\ &- \langle \mathfrak{D}^{lm}_{pq}(\mathbf{k}|\mathbf{k}_i)\rangle\langle \mathfrak{D}^{l'm'}_{p'q'}(\mathbf{k}'|\mathbf{k}'_i)\rangle.\end{aligned} \tag{99}$$

$\mathcal{T}$ is the incoherent part of $\langle \mathfrak{D}\mathfrak{D} \rangle$ and is directly related to the incoherently scattered intensity. With this substitution, (97) becomes

$$\begin{aligned}\mathcal{T}^{ll';mm'}_{pp';qq'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_i,\mathbf{k}'_i) &= \sum_{n,n'=1}^{2}\sum_{s,s'=1}^{4}\int d^3k_2 d^3k'_2\ \mathcal{K}^{ll';nn'}_{pp';ss'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_2,\mathbf{k}'_2) \\ &\times\Big(\sum_{j=1}^{2}\sum_{r=1}^{4}\mathcal{P}^{nj}_{sr}(\mathbf{k}_2)\langle \mathfrak{D}^{jm}_{rq}(\mathbf{k}_2-\mathbf{k}_i)\rangle \\ &+ \delta^{nm}\delta_{sq}\delta_3(\mathbf{k}_2-\mathbf{k}_i)\Big) \\ &\times\Big(\sum_{j'=1}^{2}\sum_{r'=1}^{4}\mathcal{P}^{n'j'}_{s'r'}(\mathbf{k}'_2)\langle \mathfrak{D}^{j'm'}_{r'q'}(\mathbf{k}_2-\mathbf{k}_i)\rangle \\ &+ \delta^{n'm'}\delta_{s'q'}\delta_3(\mathbf{k}'_2-\mathbf{k}'_i)\Big) \\ &+ \sum_{n,n'=1}^{2}\sum_{s,s'=1}^{4}\int d^3k_2 d^3k'_2\ \mathcal{K}^{ll';nn'}_{pp';ss'}(\mathbf{k},\mathbf{k}'|\mathbf{k}_2,\mathbf{k}'_2) \\ &\times\mathcal{T}^{nn';mm'}_{ss';qq'}(\mathbf{k}_2,\mathbf{k}'_2|\mathbf{k}_i,\mathbf{k}'_i).\end{aligned} \tag{100}$$

Equations (98) and (100) could be further simplified by explicitly removing the two-dimensional δ function factors from $\langle \mathfrak{D} \rangle$, $\mathcal{C}$, $\mathcal{K}$ and $\mathcal{T}$. These factors mean that (98) is in reality a two-dimensional integral and (100) is a four-dimensional integral equation.

IV. CONCLUDING REMARKS

The problem of a scalar wave scattering from a random interface between two isotropic homogeneous media has been solved in terms of a series expansion for the moments of the scattered field. The lowest-order term for the scattered intensity as given in (93) and (94) differs from the Kirchhoff approximation, but it still retains the desirable feature that it becomes exact when the surface roughness approaches zero or the discontinuity of the interface is weak.

The series expansion has a diagrammatic representation which is useful for formal partial summation of the series to give linear or nonlinear integral equations whose solution can then be undertaken. Examples of partial summation in terms of linear integral equations are given in (96) and (100). One way to generate approximate solutions would be to approximate the kernel of an integral equation such as (96) and then to solve it either

analytically or numerically. Good experimental data using carefully controlled interfaces would be very useful in developing such approximation schemes.

APPENDIX

When the properties of the media are the same

$$R = 1, \quad f(\mathbf{k}_\perp) = g(\mathbf{k}_\perp) = \tfrac{1}{2}, \quad k_+ = k_- = k$$

and (34), (35), (37), and (39) give for $z' > h(\mathbf{x}'_\perp)$, $z'' > h(\mathbf{x}''_\perp)$

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_0(\mathbf{x}' - \mathbf{x}'') + i(2\pi)^{-6} M(\mathbf{x}' \mid \mathbf{x}'') \tag{A1}$$

and for $z' < h(\mathbf{x}'_\perp)$, $z'' > h(\mathbf{x}''_\perp)$,

$$G(\mathbf{x}' \mid \mathbf{x}'') = -\, i(2\pi)^{-6} M(\mathbf{x}' \mid \mathbf{x}''), \tag{A2}$$

where

$$M(\mathbf{x}' \mid \mathbf{x}'') = \int d^3k' d^3k''\, e^{i(\mathbf{k}'\cdot\mathbf{x}' - \mathbf{k}''\cdot\mathbf{x}'')} \times [(k^2 - \mathbf{k}'^2 + i\epsilon)^{-1} - (k^2 - \mathbf{k}''^2 + i\epsilon)^{-1}] \frac{A(\mathbf{k}' - \mathbf{k}'')}{k'_z - k''_z}. \tag{A3}$$

Using (21) it can be shown that

$$\int d^3k''\, e^{-i\mathbf{k}''\cdot\mathbf{x}''} \frac{A(\mathbf{k}' - \mathbf{k}'')}{k'_z - k''_z} = -\,(2\pi)^2 e^{-i\mathbf{k}'\cdot\mathbf{x}''} \times \int dk_z e^{-ik_z(z'' - h(\mathbf{x}''_\perp))} \left(\frac{a}{k_z - i\epsilon} + \frac{1-a}{k_z + i\epsilon}\right), \tag{A4}$$

$$\int d^3k'\, e^{i\mathbf{k}'\cdot\mathbf{x}'} \frac{A(\mathbf{k}' - \mathbf{k}'')}{k'_z - k''_z} = (2\pi)^2 e^{-i\mathbf{k}''\cdot\mathbf{x}'} \times \int dk_z\, e^{ik_z(z' - h(\mathbf{x}'_\perp))} \left(\frac{a}{k_z - i\epsilon} + \frac{1-a}{k_z + i\epsilon}\right), \tag{A5}$$

where the factors a and $1 - a$ indicate that the pole at $k'_z = k''_z$ has been partially split above and below the contour, i.e.,

$$\frac{1}{k'_z - k''_z} = \frac{a}{k'_z - k'' + i\epsilon} + \frac{1-a}{k'_z - k'_z - i\epsilon}.$$

This has been done to explicitly demonstrate independence of the contour choice at this pole. Performing the k_z integration in (A4) and (A5) using

$$\int dk_z \frac{e^{ik_z z}}{k_z \pm i\epsilon} = \pm\, 2\pi i \theta(\pm z)$$

gives for $z'' > h(\mathbf{x}''_\perp)$

$$M(\mathbf{x}' \mid \mathbf{x}'') = (2\pi)^6\, iG_0(\mathbf{x}' - \mathbf{x}'')\, \theta(-z' + h(\mathbf{x}'_\perp)).$$

Combining this result with (A1) and (A2) gives

$$G(\mathbf{x}' \mid \mathbf{x}'') = G_0(\mathbf{x}' - \mathbf{x}'').$$

ACKNOWLEDGMENTS

The author would like to thank Ken Flowers for his careful reading of the manuscript and to thank the National Research Council for its support of this work.

* Present address: Bell Telephone Laboratories, Whippany, N.J.

[1] G. G. Zipfel, Jr. and John A. DeSanto, J. Math. Phys. 13, 1903 (1972).
[2] U. Frisch, "Wave Propagation in a Random Medium," in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid (Academic, New York, 1968).
[3] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1968).
[4] P. Beckman and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergammon, New York, 1963).
[5] C. Eckart, J. Acoust. Soc. Am. 25, 566 (1953).
[6] O. D. Kellogg, *Foundations of Potential Theory* (Dover, New York, 1953).
[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[8] K. Huang, *Statistical Mechanics* (Wiley, New York, 1963), pp 305-306.

Natural basis expansion of functions of the distance between two points: Generalization of Laplace's and Neumann's expansions

Robert L. Matcha

Department of Chemistry, University of Houston, Houston, Texas 77004

Kenneth C. Daiker

Department of Chemistry, Simon Fraser University, Burnaby 2, British Columbia, Canada
(Received 18 April 1972)

A method of obtaining expansions of functions of r_{12} utilizing solutions of separable differential operators which annihilate the functions is presented. Expansions of r_{12}^n in spherical polar and confocal elliptic coordinates are obtained in terms of solutions of a generalized Laplace operator.

I. INTRODUCTION

In the following we present a somewhat novel method of obtaining expansions of functions of the distance between two points, $f(r_{12})$. The technique consists of constructing separable differential operators A which annihilate the function of interest,

$$Af(r_{12})=0. \tag{1}$$

The solutions of the general equation $A\psi=0$ obtained by separation of variables are then used as a basis for the expansions with coefficients obtained by applying the appropriate boundary conditions.

The technique is demonstrated by obtaining expansions of r_{12}^n in spherical polar and confocal elliptic coordinates by using the solutions of a generalized Laplacian operator.

II. OPERATOR CONSTRUCTION AND PROPERTIES

Consider the Laplace equation

$$\nabla^2(1/r_{12})=0. \tag{2}$$

As is well known, the solution to this equation can be written as a linear combination of solutions of the general equation $\nabla^2\psi=0$ obtained by separation of variables. Laplace's and Neumann's expansions of $1/r_{12}$ utilize this fact.[1] The operator ∇^2 can be viewed as an annihilator of the function $1/r_{12}$. Similarly, if we can find a separable second-order differential operator with the property that it annihilates r_{12}^n, then the separable solutions of the operator can serve as a basis for the expansion for the function.

Such an operator can be constructed in the following manner. Consider the differential operator expressed in spherical coordinates,

$$\mathscr{L}_N=[(r_1 r_2 \sin\theta_1 \sin\theta_2)^{-1}(d/d\cos\varphi)]^N, \tag{3}$$

where $\varphi=\varphi_1-\varphi_2$. This operator has the property that

$$\mathscr{L}_N(1/r_{12})=(2N-1)!!\, r_{12}^{-2N-1}, \tag{4}$$

where $r_{12}=[r_1^2+r_2^2-2r_1r_2(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2\cos\varphi)]^{1/2}$. Operating on both sides of Eq. (4) with the Laplacian operator gives the result

$$\nabla^2[\mathscr{L}_N(1/r_{12})]=\mathscr{L}_N\nabla^2(1/r_{12})+2\nabla\mathscr{L}_N\cdot\nabla(1/r_{12}) + (\nabla^2\mathscr{L}_N)(1/r_{12}). \tag{5}$$

Applying Laplace's equation and rearranging, we find

$$\nabla^2[\mathscr{L}_N(1/r_{12})]-2\nabla\mathscr{L}_N\cdot\nabla(1/r_{12})-(\nabla^2\mathscr{L}_N)(1/r_{12})=0. \tag{6}$$

Carrying out the indicated operations, rearranging into the form of a differential operator operating on $\mathscr{L}_N(1/r_{12})$, and introducing Eq. (4) gives

$$\left[\nabla^2+\frac{2N}{r_1^2}\left(r_1\frac{\partial}{\partial r_1}-x_1\frac{\partial}{\partial x_1}-\frac{y}{(1-x_1^2)}\frac{\partial}{\partial y}\right)\right]r_{12}^{-2N-1}=0, \tag{7}$$

where we have defined for convenience in notation $x_1=\cos\theta_1$, and $y=\cos\varphi$. Setting $n=-2N-1$, rearranging terms and defining the operator

$$\Delta^n=\frac{\partial^2}{\partial r_1^2}-\frac{(n-1)}{r_1}\frac{\partial}{\partial r_1}+\frac{1}{r_1^2}\left((1-x_1^2)\frac{\partial^2}{\partial x_1^2}+(n-1)x_1\frac{\partial}{\partial x_1}\right) + \frac{1}{r_1^2(1-x_1^2)}\left((1-y^2)\frac{\partial^2}{\partial y^2}+ny\frac{\partial}{\partial y}\right) \tag{8}$$

gives the result

$$\Delta^n r_{12}^n=0. \tag{9}$$

The operator Δ^n annihilates the nth power of r_{12}. Equation (9) holds for all positive and negative integer and noninteger values n, reducing to Laplace's equation for $n=-1$. It can thus be considered to be a generalized Laplace equation. By inspection it can be seen that Δ^n is separable in the variables r_1, x_1, and y. Eigenfunctions associated with Δ^n are products of solutions of the three separated equations,

$$\left\{(1-y^2)\frac{d^2}{dy^2}+ny\frac{d}{dy}-\lambda_y\right\}\Phi(y)=0, \tag{10}$$

$$\left\{(1-x_1^2)\frac{d^2}{dx_1^2}+(n-1)x_1\frac{d}{dx_1}+\frac{\lambda_y}{(1-x_1^2)}-\lambda_x\right\}P(x_1)=0, \tag{11}$$

and

$$\left\{\frac{d^2}{dr_1^2}-\frac{(n-1)}{r_1^2}\frac{d}{dr_1}+\frac{\lambda_x}{r_1^2}\right\}R(r_1)=0, \tag{12}$$

where λ_y and λ_x are separation constants and Φ, P, and R are eigenfunctions.

III. SOLUTIONS TO r_1, x_1, AND y EQUATIONS

By making the substitution $\lambda_y = -m(m-n-1)$, Eq. (10) is seen to reduce to Gegenbauer's equation[2] with solutions

$$\Phi_m^n(y) = \sum_{\mu=0}^{[m/2]} D_{m,\mu}^n \, y^{m-2\mu}, \tag{13}$$

where $[m/2]$ indicates the integer part of $m/2$. The coefficients $D_{m,\mu}^n$ have the definition

$$D_{m,\mu}^n = [(m-2\mu)!(2\mu)!\,!\,p(n-2m+2\mu,\mu)]^{-1}, \tag{14}$$

with the function p defined as

$$p(a,b) = \begin{cases} (a+1)(a-1)\ldots(a-2b+3), & b \text{ integer} > 0 \\ 1, & b = 0. \end{cases} \tag{15}$$

Equation (13) differs from the usual Gegenbauer polynomials $C_m^{-(n+1)/2}$ by a normalization factor. Explicitly

$$C_m^{-(n+1)/2} = p(n+1,m)\,(-1)^m \Phi_m^n. \tag{16}$$

This choice of normalization insures that the Φ_m^n are properly defined for $n = -1$.

A polynomial solution to Eq. (11) can be determined by substituting $\lambda_y = -m(m-n-1)$. However, it is first necessary to remove the singularity at the point $(1-x_1^2)^{1/2} = 0$ by making the substitution $P = (1-x_1^2)^{\alpha}\overline{P}$, where α is an undetermined variable. This leads to an indicial equation with solution $\alpha = m/2$.

Substituting $P = (1-x_1^2)^{m/2}\overline{P}$ into Eq. (11) we obtain the new equation

$$(1-x_1^2)\frac{d^2\overline{P}}{dx_1^2} + (n-2m-1)x_1\frac{d\overline{P}}{dx_1} - [m(m-n)+\lambda_x]\overline{P} = 0. \tag{17}$$

This is just Gegenbauer's equation. Thus, the general solutions may be written as

$$P \equiv P_l^{m,n} = (-1)^{l-m}(1-x_1^2)^{m/2}\,p(n-2m, l-m)\Phi_{l-m}^{n-2m-1}, \tag{18}$$

with $\lambda_x = -l(l-n)$. These functions reduce to associated Legendre functions when $n = -1$. Equation (11) is thus a generalization of the associated Legendre equation. In order to obtain proper solutions to the r_{12}^n expansion when n is an even integer, it is necessary to use the following normalized version of the functions given in Eq. (18):

$$\begin{aligned} \mathscr{P}_l^{m,n} &= (-1)^{l-m}[n(n-2)\ldots(n-[l+m]^{+}+2)]^{-1}\,P_l^{m,n} \\ &= (1-x_1^2)^{m/2}\sum_{\mu=0}^{[(l-m)/2]} d_{l,m}^{n,\mu}\, x_1^{l-m-2\mu}, \end{aligned} \tag{19}$$

where

$$d_{l,m}^{n,\mu} = \frac{(n-[l+m]^{+})(n-[l+m]^{+}-2)\ldots(n-2l+2\mu+2)}{(l-m-2\mu)!(2\mu)!\,!}. \tag{20}$$

The symbol $[l+m]^{+}$ is defined by the equation

$$[l+m]^{+} = l+m \pm \mathrm{OE}(l+m), \tag{21}$$

where the parity function $\mathrm{OE}(L)$ equals zero when L is an even integer and one when L is an odd integer.

A solution to the radial equation can be obtained by substituting $\lambda_x = -l(l-n)$ into Eq. (12). We find, $R = r_1^l$. A linearly independent second solution is r_1^{-l+n}.

IV. EXPANSION OF r_{12}^n IN SPHERICAL COORDINATES

We are now in a position to expand r_{12}^n in terms of the eigenfunctions associated with the Δ^n operator. These have the form

$$\psi_l^{m,n}(1) = (a_l r_1^l + b_l r_1^{-l+n})\mathscr{P}_l^{m,n}(x_1)\,\Phi_m^n(y) \tag{22}$$

and

$$\psi_l^{m,n}(2) = (c_l r_2^l + d_l r_2^{-l+n})\mathscr{P}_l^{m,n}(x_2)\,\Phi_m^n(y). \tag{23}$$

Assuming $r_1 < r_2$ and applying boundary conditions on r_1, we may thus write

$$r_{12}^n = \sum_{m=0}^{\infty}\sum_{l=m}^{\infty} c_l^m(2)\, r_1^l \mathscr{P}_l^{m,n}(x_1)\,\Phi_m^n(y), \tag{24}$$

where $c_l^m(2)$ is a function of the coordinates of r_2, x_2 only. Similarly

$$r_{12}^n = \sum_{m=0}^{\infty}\sum_{l'=m}^{\infty} d_{l'}^m(1)\, r_2^{-l'+n}\mathscr{P}_{l'}^{m,n}(x_2)\,\Phi_m^n(y), \tag{25}$$

where $d_{l'}^m(1)$ is a function of r_1, x_1 only. Comparing coefficients of $\Phi_m^n(y)$ yields the expansion

$$r_{12}^n = \sum_{m=0}^{\infty}\sum_{l,l'=m}^{\infty} g_{ll'}^m \frac{r_1^l}{r_2^{l'-n}}\mathscr{P}_l^{m,n}(x_1)\mathscr{P}_{l'}^{m,n}(x_2)\,\Phi_m^n(y), \tag{26}$$

where the $g_{ll'}^m$'s are undetermined constants. Since $r_{12}^n(x_1,x_2) = r_{12}^n(x_2,x_1)$,

$$g_{ll'}^m(r_1^l/r_2^{l'-n}) = g_{l'l}^m(r_1^{l'}/r_2^{l-n}), \tag{27}$$

which implies that $g_{ll'}^m = \delta_{ll'}\, g_{lm}^n$, where the n was suppressed in earlier coefficients for convenience in notation. Thus

$$r_{12}^n = \sum_{l=0}^{\infty}\sum_{m=0}^{l} g_{lm}^n \frac{r_1^l}{r_2^{l-n}}\mathscr{P}_l^{m,n}(x_1)\mathscr{P}_l^{m,n}(x_2)\,\Phi_m^n(y). \tag{28}$$

To determine g_{lm}^n, we rotate r_2 to fall along the z axis. Under these conditions x_2 equals one and x_1 equals $\cos\theta_{12}$ and the expansion is independent of y. Hence only the terms with $m = 0$ remain and

$$r_{12}^n = \sum_{l=0}^{\infty} g_{l0}^n \frac{r_1^l}{r_2^{l-n}}\mathscr{P}_l^{0,n}(\cos\theta_{12})\mathscr{P}_l^{0,n}(1). \tag{29}$$

The g_{l0}^n coefficients can now be obtained by setting $\cos\theta_{12} = 1$, expanding r_{12}^n in a power series in r_1/r_2, and comparing coefficients of r_1^l/r_2^{l-n}. The required expansion is

$$r_{12}^n = |r_1 - r_2|^n = \sum_{l=0}^{\infty}(-1)^l\binom{n}{l}\frac{r_1^l}{r_2^{l-n}}. \tag{30}$$

Comparing Eqs. (29) and (30) utilizing the fact that $\mathscr{P}_l^{0,n}(1) = [(n-1)(n-3)\ldots(n-[l]^{-}+1)]/l!$ obtained by induction from Eq. (19), we find

$$g_{l0}^n = \frac{(-1)^l\, l!\, n(n-2)\ldots(n+2-[l]^{+})}{[(n-1)(n-3)\ldots(n+1-[l]^{-})]} \tag{31}$$

and

$$g_{l0}^n\,\mathscr{P}_l^{0,n}(1)\mathscr{P}_l^{0,n}(\cos\theta_{12}) = P_l^{0,n}(\cos\theta_{12}). \tag{32}$$

Thus

$$r_{12}^n = \sum_{l=0}^{\infty}\frac{r_1^l}{r_2^{l-n}}P_l^{0,n}(\cos\theta_{12}). \tag{33}$$

Gegenbauer's expansion[2] of inverse powers of r_{12} may be obtained from Eq. (33) by setting $n=-2\nu$ and making the identification $P_l^{0,n}=C_l^{-n/2}$.

The remaining g_{lm}^n's in Eq. (28) may be obtained by first equating coefficients of Eqs. (33) and (28),

$$P_l^{0,n}=\sum_{m=0}^{l} g_{lm}^n \mathscr{P}_l^{m,n}\mathscr{P}_l^{m,n}\Phi_m^n. \tag{34}$$

We next note that for x_1, x_2 large

$$P_l^{0,n}(\cos\theta_{12})=c_{l,0}^n x_1^l x_2^l (1-y)^l, \tag{35}$$

where $c_{lm}^n=(-1)^l[n(n-2)\ldots(n-[l]^+ +2)]d_{l0}^{mn}$.

Also

$$\mathscr{P}_l^{m,n}(x_1)=(-1)^{m/2}x_1^l d_{lm}^{n,0} \tag{36}$$

and

$$\Phi_m^n(y)=\sum_{\mu=0}^{[m/2]} D_{m,\mu}^n y^{m-2\mu}. \tag{37}$$

Substituting Eqs. (35), (36), and (37) into (34) and equating coefficients of powers of y leads to the set of equations,

$$(-1)^{l-j}\binom{l}{j}=\sum_{k=0}^{[j/2]} a_{l,l-j+2k}^n D_{l-j+2k,k}^n, \tag{38}$$

with

$$a_{lm}^n=(-1)^m g_{lm}^n [d_{lm}^{n,0}]^2/c_{l,0}^n. \tag{39}$$

The g_{lm}^n coefficients can be obtained from Eq. (38) by induction. The result is

$$g_{lm}^n=\frac{(-1)^l(l-m)!\,n(n-2)(n-4)\ldots(n+2-[l+m]^+)}{(n-2m-1)(n-2m-3)\ldots(n+1-[l+m]^-)}. \tag{40}$$

Equation (28) holds for all n except positive odd integers. For these cases, due to the appearance of accidental degeneracies, the Φ_m^n's do not form a complete set. For positive even integers, the expansion truncates. This fact is evident from the g_{lm}^n definition. Note that Eq. (33) is valid for all n.

V. ONE CENTER EXPANSION OF r_{12}^n IN SPHERICAL HARMONICS

From Eq. (33), it is a relatively simple matter to obtain an expansion of r_{12}^n in spherical harmonics. To do this we first substitute the expansion in powers of $\cos\theta_{12}$ of $P_l^{0,n}$ into Eq. (33) giving

$$r_{12}^n=\sum_{l=0}^{\infty}\frac{r_1^l}{r_2^{l-n}}\sum_{m=0}^{[l/2]} c_{lm}^n \cos\theta_{12}^{l-2m}. \tag{41}$$

We now introduce the expansion[3]

$$\cos\theta_{12}^j=j!\sum_{k=0}^{[j/2]}\frac{(2j-4k+1)P_{j-2k}(\cos\theta_{12})}{(2k)!!(2j-2k+1)!!}, \tag{42}$$

giving

$$r_{12}^n=\sum_{l=0}^{\infty}\frac{r_1^l}{r_2^{l-n}}\sum_{m=0}^{[l/2]} c_{lm}^n \tag{43}$$

$$\times[l-2m]!\sum_{k=0}^{[(l-2m)/2]}\frac{(2l-4m-4k+1)}{(2k)!!(2l-4m-2k+1)!!}P_{l-2m-2k}^{(\cos\theta_{12})}.$$

Substituting $\lambda=l-2m-2k$ and reordering summation indices gives

$$r_{12}^n=\sum_{\lambda=0}^{\infty}(2\lambda+1)P_\lambda(\cos\theta_{12})\sum_{l(2)=\lambda}^{\infty} b_{\lambda l}^n (r_1^l/r_2^{l-n}), \tag{44}$$

where $l(2)$ indicates the summation proceeds in steps of two and

$$b_{\lambda l}^n=(-1)^l\frac{p(n,(l-\lambda)/2)\,p(n-1,(\lambda+l-1)/2)}{(l+\lambda+1)!!\,(l-\lambda)!!}. \tag{45}$$

with $b_{\lambda l}^0=\delta_{\lambda 0}\delta_{l0}$.

Under certain conditions, the l and λ summations truncate. For even positive n, $\lambda<n/2$ and $\lambda\leq l\leq n-\lambda$. For odd integer n greater than -3, $\lambda\leq l\leq n+1+\lambda$. These two cases can be stated explicitly as

$$r_{12}^{2N}=\sum_{\lambda=0}^{N}(2\lambda+1)P_\lambda(\cos\theta_{12})\sum_{l(2)=\lambda}^{2N-\lambda} b_{\lambda l}^{2N}\frac{r_1^l}{r_2^{l-2N}} \tag{46}$$

and

$$r_{12}^{2N-1}=\sum_{\lambda=0}^{\infty}(2\lambda+1)P_\lambda(\cos\theta_{12})\sum_{l(2)=\lambda}^{2N+\lambda} b_{\lambda l}^{2N-1}\frac{r_1^l}{r_2^{l-2N+1}}. \tag{47}$$

Introducing the addition theorem in spherical harmonics for $P_\lambda(\cos\theta_{12})$ leads to the expression

$$r_{12}^n=4\pi\sum_{\lambda=0}^{\infty}\sum_{l(2)=\lambda}^{\infty}\sum_{m=-\lambda}^{\lambda} b_{\lambda l}^n\frac{r_1^l}{r_2^{l-n}}Y_\lambda^m(\hat{r}_1)Y_\lambda^{m*}(\hat{r}_2). \tag{48}$$

This expression is equivalent to an expansion derived recently by Sack.[4]

VI. EXPANSION OF r_{12}^n IN CONFOCAL ELLIPTIC COORDINATES

In confocal elliptic coordinates, the annihilator Δ^n has the form,

$$\Delta^n=\frac{1}{R^2}\Big[(\xi_1^2-1)\frac{\partial^2}{\partial\xi_1^2}-(n-1)\xi_1\frac{\partial}{\partial\xi_1}+(1-\eta_1^2)\frac{\partial^2}{\partial\eta_1^2}$$

$$+(n-1)\eta_1\frac{\partial}{\partial\eta_1}+\left(\frac{1}{\xi_1^2-1}+\frac{1}{1-\eta_1^2}\right)\Big((1-y^2)\frac{\partial^2}{\partial y^2}$$

$$+ny\frac{\partial}{\partial y}\Big)\Big], \tag{49}$$

where R is the distance between the foci of the ellipse defining ξ_1 and η_1. This operator separates into y and η_1 equations identical to Eqs. (10) and (11) and the ξ_1 equation

$$\left((\xi_1^2-1)\frac{d^2}{d\xi_1^2}-(n-1)\xi_1\frac{d}{d\xi_1}+\frac{\lambda_\nu}{(\xi_1^2-1)}+\lambda_\eta\right)\hat{P}(\xi_1), \tag{50}$$

where $\lambda_\eta=-l(l-n)$. Equation (50) differs only by a sign from Eq. (11). The solution has the form

$$\hat{\mathscr{P}}_l^{m,n}(\xi_1)=(\xi_1^2-1)^{m/2}\sum_{\mu=0}^{[(l-2m)/2]} d_{lm}^{n,\mu}\,\xi_1^{l-m-2\mu}. \tag{51}$$

Since the ξ_1 coordinate ranges from one to infinity, in order for the eigenfunctions of Eq. (50) to serve as basis functions in the expansion of r_{12}^n, both the first solutions, $\hat{\mathscr{P}}_l^{m,n}$ and the second solutions, $\hat{Q}_l^{m,n}$ are required. The latter may be obtained by standard techniques.[3] In terms of Gegenbauer functions of the second kind,[2] D_λ^M,

$$\hat{Q}_l^{m,n}(\xi)=\frac{(2^{l+1})\,\Gamma(l-n/2+1)}{\Gamma(-n/2)\,\Gamma(l-n)}\,(\xi^2-1)^{m/2}\,D_{l-m}^{n-2m}(\xi). \tag{52}$$

As a power series in ξ,

$$\hat{Q}_l^{m,n}(\xi)=(\xi^2-1)^{m/2}\,\xi^{-l-m-n}$$

$$\times\sum_{\mu=0}^{\infty}(-1)^\mu\frac{(n-l-m)(n-l-m-1)\ldots(n+1-l-m-2\mu)}{(2\mu)!\,!(n-2l-2)(n-2l-4)\ldots(n-2l-2\mu)}$$

$$\cdot\,\xi^{-2\mu}. \tag{53}$$

The functions r_{12}^n can thus be expanded in terms of the basis functions

$$\chi_{lm}^n(1)=[A_l^m\,\hat{\mathscr{P}}_l^{mn}(\xi_1)+B_l^m\,\hat{Q}_l^{m,n}(\xi_1)]\mathscr{P}_l^{mn}(\eta_1)\Phi_m^n(y) \tag{54}$$

and

$$\chi_{lm}^n(2)=[C_l^m\,\hat{\mathscr{P}}_l^{mn}(\xi_2)+D_l^m\,\hat{Q}_l^{mn}(\xi_2)]\mathscr{P}_l^{mn}(\eta_2)\,\Phi_m^n(y). \tag{55}$$

Assuming $\xi_1<\xi_2$ and noting that $\hat{Q}_l^{mn}$ has a logarithmic singularity at $\xi_1=1$, we may write

$$r_{12}^n=\sum_{m=0}^{\infty}\sum_{l=m}^{\infty}\bar{a}_l^m(2)\,\hat{\mathscr{P}}_l^{mn}(\xi_1)\mathscr{P}_l^{mn}(\eta_1)\,\Phi_m^n(y), \tag{56}$$

where the $\bar{a}_l^m$ coefficients are functions of coordinates ξ_2,η_2 only. Noting that $\hat{\mathscr{P}}_l^{m,n}$ approaches infinity for large values of the argument, we write

$$r_{12}^n=\sum_{m=0}^{\infty}\sum_{l=m}^{\infty}\bar{d}_l^m(1)\,\hat{Q}_l^{mn}(\xi_2)\mathscr{P}_l^{mn}(\eta_2)\,\Phi_m^n(y), \tag{57}$$

where the $\bar{d}_l^m$ coefficients are functions of coordinates ξ_1,η_1 only. Applying arguments similar to those given in Sec. IV with respect to the interchange of η_1 and η_2 we find

$$r_{12}^n=\sum_{l=0}^{\infty}\sum_{m=0}^{l}A_{lm}^n\,\hat{Q}_l^{mn}(\xi_2)\hat{\mathscr{P}}_l^{mn}(\xi_1)\mathscr{P}_l^{mn}(\eta_1)\mathscr{P}_l^{mn}(\eta_2)\,\Phi_m^n(y), \tag{58}$$

where the A_{lm}^n are undetermined coefficients. To determine these, we take the limit $R\to 0$ and compare with Eq. (28). In this limit,

$$R\xi/2\to r,\quad \eta\to\cos\theta, \tag{59}$$

while y remains unchanged. Using Eqs. (51) and (53), we find

$$\lim_{R\to 0}\hat{Q}_l^{mn}(\xi_2)\hat{\mathscr{P}}_l^{mn}(\xi_1)\mathscr{P}_l^{mn}(\eta_1)\mathscr{P}_l^{mn}(\eta_2)\,\Phi_m^n(y)$$

$$=(2/R)^n d_{lm}^{n,0}(r_1^l/r_2^{l-n})\mathscr{P}_l^{mn}(x_1)\mathscr{P}_l^{mn}(x_2)\,\Phi_m^n(y). \tag{60}$$

Thus, in view of Eqs. (28), (58), and (60),

$$A_{lm}^n=(R/2)^n g_{lm}^n/d_{lm}^{n,0}. \tag{61}$$

Equation (58) is valid for all n except positive integers and can be shown to reduce to an expansion obtained by Wolniewicz[5] for negative n.

VII. CONFOCAL ELLIPTIC EXPANSION OF r_{12}^n IN SPHERICAL HARMONICS

Using the results of previous sections we can derive an expansion of r_{12}^n involving spherical harmonic functions of η_1 and η_2. To do this we introduce the expression [derived from Eq. (42) by applying $(1-x_1^2)^\mu\,d^\mu/dx_1^\mu$],

$$x_1^j=j!(1-x_1^2)^{-\mu}\sum_{k=0}^{[j/2]}\frac{(2j+2\mu-4k+1)}{(2k)!\,!(2j+2\mu-2k+1)!\,!}\,P_{j+\mu-2k}^\mu(x_1)$$

$$(\mu>0) \tag{62}$$

where

$$P_\lambda^\mu=(1-x_1^2)^{\mu/2}\frac{d^\mu}{dx_1^\mu}P_\lambda \tag{63}$$

into Eq. (19) and the expression

$$y^M=\frac{1}{2^M}\sum_{\mu=0}^{M}\binom{M}{\mu}\exp[i(M-2\mu)\varphi \tag{64}$$

into Eq. (37). These lead after some manipulation to the relationships

$$\mathscr{P}_l^{mn}(x_1)=\sum_{\lambda(2)=\mathrm{OE}(l-m)+\mu}^{l}I_{\lambda\mu}^{nlm}\,P_\lambda^\mu(x_1)\ (\mu>0) \tag{65}$$

and

$$\Phi_m^n(y)=\sum_{\mu(2)=-m}^{m}G_\mu^{mn}\exp(i\mu\varphi), \tag{66}$$

where

$$I_{\lambda\mu}^{nlm}=(2\lambda+1)\sum_\sigma(-1)^\sigma\frac{(n-[l+m]^+)(n-[l+m]^+-2)\ldots(n-2l+2\sigma+2)}{(2\sigma)!\,!(l-m-2\sigma)!}\sum_j\frac{(-1)^{(j-l+m)/2}\,j!}{(j+|\mu|-\lambda)!\,!(\lambda+j+|\mu|+1)!\,!}\binom{(m-|\mu|)/2}{(j-l+m+2\sigma)/2} \tag{67}$$

and

$$G_\mu^{mn}=\frac{(n+1)(n-1)\ldots(n+3-m+|\mu|)}{(m+\mu)!(m-\mu)!(n+1-m-|\mu|)(n-1-m-|\mu|)\ldots(n+3-2m)}. \tag{68}$$

The limits on σ and j in Eq. (67) are

$$0\le\sigma\le\min\{(l-\lambda)/2,\ [(l-m)/2]\}, \tag{69}$$

and

$$\max\{\lambda-|\mu|,l-m-2\sigma\}\le j(2)\le l-|\mu|-2\sigma.$$

A sum rule may exist for Eq. (67) but we have not found one. We now introduce Eqs. (65) and (66) into Eq. (58)

and obtain after some manipulation

$$r_{12}^n = \sum_{lm\lambda\lambda'\mu} T(nlm;\lambda\lambda'\mu)$$

$$\hat{Q}_l^{mn}(\xi_2)\hat{\mathscr{P}}_l^{mn}(\xi_1)\, Y_\lambda^\mu(\eta_1,\phi_1)\, Y_{\lambda'}^{\mu *}(\eta_2,\varphi_2), \tag{70}$$

where

$$T(nlm;\lambda\lambda'\mu) = 4\pi A_{lm}^n G_\mu^{lm}$$

$$\times\left(\frac{(\lambda+\mu)!(\lambda'+\mu)!}{(\lambda-\mu)!(\lambda'-\mu)!(2\lambda+1)(2\lambda'+1)}\right)^{1/2} I_{\lambda\mu}^{nlm}\, I_{\lambda'\mu}^{nlm}, \tag{71}$$

with $0 \leq l \leq \infty$, $0 \leq m \leq l$, $\mathrm{OE}(l-m) \leq \lambda(2) \leq l$, $\mathrm{OE}(l-m) \leq \lambda'(2) \leq l$, and

$$\max\{-m, -\lambda + \mathrm{OE}(l-m), -\lambda' + \mathrm{OE}(l-m)\}$$

$$\leq \mu \leq \min\{m, \lambda - \mathrm{OE}(l-m), \lambda' - \mathrm{OE}(l-m)\}. \tag{72}$$

Equation (70) holds for all n except positive even integers. Clearly, the same technique could be used to express $\hat{Q}_l^{mn}\hat{\mathscr{P}}_l^{mn}$ in terms of associated Legendre functions of the first and second kind. The resultant expansion would have all of the n dependence in the coefficients.

VIII. CONCLUSION

We have utilized an annihilator method to obtain expansions of r_{12}^n in spherical and confocal elliptic coordinates. This method is quite general and can be used to obtain expansions of other functions of r_{12} in various coordinate systems. We are currently working on further applications. We are also investigating in more detail the properties of the generalized Laplacian operator derived herein.

ACKNOWLEDGMENTS

One of the authors (RLM) gratefully acknowledges the support of the Robert A. Welch Foundation through Grant No. E-397. The other gratefully acknowledges the postdoctoral support of Dr. G. Malli through his National Research Council of Canada Grant No. 3510.

[1] J. Neumann, J. Reine Angew. Math. **36**, 21 (1848).
[2] L. Gegenbauer, Sitzungsber. Akad. Wiss. Wien Math.-Naturwiss. Kl. 2 A **70**, 433 (1874).
[3] E. D. Rainville, *Special Functions* (Macmillan, New York, 1960).
[4] R. A. Sack, J. Math. Phys. **5**, 245 (1964).
[5] L. Wolneiwicz, Acta Phys. Pol. **22**, 3 (1962).

Extra selection rules

N. B. Backhouse

Department of Applied Mathematics, University of Liverpool, Liverpool, England
(Received 31 January 1972; revised manuscript received 15 May 1972).

The recognition of antiunitary symmetry in a quantum mechanical system is considered in relation to selection rules. It is shown that the interaction of unitary symmetry and the anti-invariance of matrix elements under antiunitary operators can produce extra selection rules in magnetic systems. The criteria obtained for the occurrence of these additional selection rules are presented in the form of group character tests, and in a special case provide an application for the symmetrized and antisymmetrized squares of group representations.

1. INTRODUCTION

The theory of selection rules is concerned with the vanishing or nonvanishing of matrix elements purely on the basis of symmetry. In a quantum mechanical context, it is usual to consider matrix elements of the form

$$V_{rps} = \int \psi_r^* V_p \phi_s d\mu, \tag{1.1}$$

where the factors in the integrand are functions or operators, as appropriate, having known transformation properties under a unitary symmetry group **G**, and where μ is a **G**-invariant measure. In a typical application, the selection rule $V_{rps} = 0$, for all indices, indicates that a transition from the initial state ϕ to the final state ψ, under the perturbation **V**, is forbidden.

Group theoretically, if the quantities ψ_r, V_p, ϕ_s for all (r, p, s), transform according to the representations D_1, D_2, D_3, respectively, of **G**, then $V_{rps} = 0$ for all (r, p, s), if the Kronecker product $D^* \otimes D_2 \otimes D_3$ fails to contain the totally symmetric representation of **G**. Examples of the use of this criterion may be found in the standard works on the application of group theory to physics, for example see Refs. 1–5. However, it is a deep result of Wigner[1] that a quantum mechanical system may exhibit antiunitary symmetries in addition to its unitary symmetry group; so it is of interest to consider their effect on the existing selection rules. In the vast majority of cases the selection rule criterion given above is unaltered, except that, as Wigner[1] shows, the symmetries are realized by the use of corepresentations rather than representations. But not all vanishings of matrix elements are covered by the above rule; indeed it is well known, for example, see Hamermesh,[2] that further selection rules exist in the case of unitary symmetry when the functions ψ_r^* and ϕ_s belong to the same carrier space. Also, Lax[6] has demonstrated that there exist additional selection rules for matrix elements in which the functions ψ_r and ϕ_s are connected by time-reversal. In both of these special situations the analysis depends on the reduction of symmetrized and antisymmetrized squares of representations, see for example Sec. 2 of Bradley and Davies[7] for a simplified account of the second of these situations. It should be noted that in the second case it is unnecessary to employ corepresentations; this is so because the restriction to the unitary subgroup of the corepresentation, under which the matrix elements transform, contains more than the unitary representation on which the analysis of Lax, or Bradley and Davies, is based.

The antiunitary operators which spring immediately to mind are the time-reversal operators; more generally, and especially in the context of magnetic crystals, an antiunitary operator may appear as the product of a time-reversal operator together with a unitary operator. It is with magnetic crystals in mind that we have written this paper on the extra selection rules caused by the recognition of antiunitary symmetry in systems. More specifically we consider those matrix elements of the form (1.1) in which the functions ψ_r, ϕ_s, are related by an antiunitary operator, and where corepresentation theory does not yield a vanishing.

Recently, Aviran and Zak,[8] in an account of the Wigner–Eckart theorem for corepresentations, find sets of relationships between matrix elements caused by the presence of the antiunitary operators. The relationships are in general quite complicated, and their full implications may not be apparent. Indeed Aviran and Zak[8] did not set out to investigate these relations, except in special cases. Perhaps the aim of our paper, put in a slightly different way from that expressed above, is that of unravelling some of these relationships, in the cases when we might expect to be lead to vanishings. In this sense our work follows naturally on from that of Aviran and Zak.[8] However, because the restrictions we impose on certain quantities are not reflected in a definite assignment of corepresentation types, we do not work directly from Ref. 8. In support of this we note that the position operator and the momentum operator transform in the same way under rotation, yet differently under time-reversal.

The presentation of the paper is as follows. In Sec. 2 we define the matrix elements with which we are concerned, and by the imposition of reasonable conditions on the integrands we derive relationships between these matrix elements. Then in a manner analogous to the analysis of Eq. (2.14) of Bradley and Davies,[7] we find sufficient conditions for the occurrence of extra selection rules. These new criteria, embodied in Theorem 2.1, are translated into character tests in Sec. 3; as a special case Theorem 3.2 gives a direct generalization of the analysis of (2.14) of Ref. 7. It may be of some interest to note that even when the restrictive conditions of Theorem 2.1 fail to hold, the conclusions of Lemma 2.1 and its corollary force some, but not necessarily all, the matrix elements to vanish. The remainder of the paper is devoted to testing the conditions of Theorem 2.1 against the representations of crystallographic point groups and space groups. In the latter case the analysis is facilitated by the use of induced representations which enable the conditions of Theorem 2.1 to be rewritten in terms of little groups and their allowable representations.

The group representation theory involved in this paper

is elementary and may be found in the texts,[1-5] and the recent book of Bradley and Cracknell.[9] The literature in the field of magnetic symmetry is now so vast that it is difficult to do justice to it. Notable contributions have been made by Wigner,[1] Tavger and Zaitzev,[10] Opechowski and Guccione,[11] Dimmock and Wheeler,[12] Birss,[13] Bhagavantam,[14] and Bradley and Davies.[15] Also the extensive bibliography of Ref. 9 should be consulted for references to the numerous contributions from the Russian school and to the more recent papers of Cracknell and associates.

2. GENERAL EXTRA SELECTION RULES

Let **M** be a magnetic group having a unitary subgroup **G** and antiunitary coset generated by the antiunitary operator A. Then **M** may be written as

$$\mathbf{M}=\mathbf{G}+A\mathbf{G}. \tag{2.1}$$

In general A is of the form $R\theta$, where θ is a time-reversal operator commuting with all geometric operations and satisfying $\theta^2=wI$, with $w=\pm 1$ and R is a unitary operator (henceforth no distinction is made between group elements and their corresponding operators). Except in the case $R=$ the identity, R does not belong to **G**. The corepresentation theory of **M**, with respect to the unitary subgroup **G**, is sufficient to deal with the majority of selection rules in the underlying quantum mechanical system. However, it is the exceptional cases which are of interest here and they are to be found among matrix elements of the form

$$V_{rps}=\int (A\phi_r)^* V_p\phi_s\, d\mu, \tag{2.2}$$

$$=(A\phi_r,\ V_p\phi_s), \tag{2.3}$$

where the factors ϕ_r, $r=1,2,\ldots,d$; V_p, $p=1,2,\ldots,d'$, transform according to the unitary representations D, D', respectively, of **G**. It is shown in the text following (2.4) that R normalizes **G**, so it makes sense to define a new representation D_R of **G** by $D_R(g)=D(R^{-1}gR)$ for all $g\in\mathbf{G}$. Since it can be shown that the functions $(A\phi_r)^*$ transform according to D_R, then $V_{rps}=0$, for all (r,p,s), unless $D_R\otimes D\otimes D'$ contains the trivial representation of **G**. Although this approach is better than one based on corepresentations of **M**, it can be shown that there are cases in which a further sharpening is possible.

First, it is convenient to introduce the unitary group

$$\mathbf{G}'=\mathbf{G}+R\,\mathbf{G}. \tag{2.4}$$

Since **G** is a normal subgroup of **M** and θ commutes with each element of **G**, it follows that R normalizes **G**. Also because $A^2\in\mathbf{G}$ it follows that R^2 or $-R^2\in\mathbf{G}$. In the latter case **G** is a double group and contains R^2 for it necessarily contains $-I$. It has been shown that (2.4) does indeed define a unitary group.

Since $R^2\in\mathbf{G}$ the unitary matrix $D(R^2)$ exists and may be diagonalized. With respect to a suitable basis

$$D(R^2)_{jk}=\delta_{jk}\exp(i\theta_j), \tag{2.5}$$

for all j,k.

As it is necessary to consider the operators RV_pR^{-1}, for all p, in Lemma 2.1 below, it is convenient to assume that D' extends to a unitary representation of **G'**. Indeed D' could have been taken as a unitary representation of **G'** in the first place. Moreover, for definiteness, assume that D, D', are irreducible representations of their respective groups, and that

$$D'(R)_{jk}=\delta_{jk}\exp(i\alpha_j), \tag{2.6}$$

for all j,k. Since D and D' are defined on completely independent spaces the conditions (2.5) and (2.6) can always be made and in no way interfere with one another.

Now introduce a restriction on the operators V_p by requiring,

$$(\theta V_p\theta^{-1})^+=\beta V_p, \tag{2.7}$$

for all p, where $\beta=\pm 1$ is independent of p. An argument supporting this assumption is given in the Appendix.

By analogy with (2.16) of Ref. 7, and based on the antiinvariance of the scalar product with respect to A, the following relations hold between the matrix elements (2.3).

Lemma 2.1: Under the hypotheses expressed by (2.4) to (2.7),

$$V_{rps}=\omega\beta\exp[i(\theta_r+\alpha_p)]\,V_{spr}, \tag{2.8}$$

for all (r,p,s).

Proof: By definition (2.3),

$$V_{rps}=(A\phi_r,\ V_p\phi_s), \tag{2.9}$$

$$=(AV_p\phi_s, A^2\phi_r), \tag{2.10}$$

using the antiunitarity of A,

$$=\omega(AV_pA^{-1}A\phi_s, R^2\phi_r), \tag{2.11}$$

since $\theta^2=\omega I$,

$$=\omega\exp(i\theta_r)\,[A\phi_s,\ (AV_pA^{-1})^+\phi_r], \tag{2.12}$$

since $D(R^2)$ is diagonal,

$$=\omega\beta\exp(i\theta_r)\,[A\phi_s,\ (RV_pR^{-1})\,\phi_r], \tag{2.13}$$

using (2.7),

$$=\omega\beta\exp[i(\theta_r+\alpha_p)]\,V_{spr}, \tag{2.14}$$

since $D'(R)$ is diagonal.

Corollary: $V_{rps}=0$ unless $\exp[i(\theta_r+\theta_s+2\alpha_p)]=1$.

Proof: A double application of Lemma 2.1 gives

$$V_{rps}=\exp[i(\theta_r+\theta_s+2\alpha_p)]\,V_{rps}, \tag{2.15}$$

for all (r,p,s). The conclusion follows at once.

It should be noted that the relations obtained in Lemma 2.1 and in the corollary are not basis independent. This is not to be regarded as a disadvantage, because it is always possible to transform to any other basis set as required.

In analogy with the text following (2.16) of Ref. 7 we look for subspaces of Ω, the carrier space of $D_R\otimes D\otimes D'$, which bear some relationship to the equations (2.8). Let $F_{rps}=(A\phi_r)^*V_p\phi_s$, for all (r,p,s), and define

$$F^{\pm}_{rps}=F_{rps}\pm\omega\beta\exp[i(\theta_r+\alpha_p)]\,F_{spr}, \tag{2.16}$$

for all (r,p,s). These two sets of functions span the linear spaces $\Omega^{\pm}$, respectively. Although $\Omega=\Omega^+ + \Omega^-$ the sum is not necessarily a direct sum because the intersection of the subspaces $\Omega^{\pm}$ may be larger than the zero

space. Consider now the action of $g \in \mathbf{G}$ on the space Ω^+

$$g: F^+_{rps} \to \sum_{r',p',s'} D_R(g)_{r'r} D'(g)_{p'p} D(g)_{s's} F_{r'p's'} + \omega\beta \exp[i(\theta_r + \alpha_p)] \sum_{r',p',s'} D_R(g)_{s's} D'(g)_{p'p} D(g)_{r'r} F_{s'p'r'}. \tag{2.17}$$

If the assumption is made that

$$D_R(g)_{r'r} D'(g)_{p'p} D(g)_{s's} = \exp[i(\theta_r - \theta_{r'} + \alpha_p - \alpha_{p'})] \times D_R(g)_{s's} D'(g)_{p'p} D(g)_{r'r}, \tag{2.18}$$

for all (r,p,s), (r',p',s'), and for all $g \in \mathbf{G}$, then Ω^+ is stable under the action of $\mathbf{G}$. Equivalently, Ω^+ is stable if

$$D_{R^{-1}}(g)_{r'r} D'(g)_{p'p} D(g)_{s's} = D(g)_{r'r} D_R'(g)_{p'p} D_R(g)_{s's}, \tag{2.19}$$

for all (r,p,s), (r',p',s'), and $g \in \mathbf{G}$. A more convenient form of (2.19) is obtained in the following theorem.

Theorem 2.1: Under the hypotheses embodied in (2.4)–(2.7), the following conditions are sufficient to ensure the stability of Ω^+ and are equivalent to equations (2.19):

1. $D'_R = \chi_1 D'$, (2.20)
2. $D_R = \chi_2 D$, (2.21)
3. $D_{R^{-1}} = \chi_3 D$, (2.22)

where χ_1, χ_2, χ_3, are linear characters of $\mathbf{G}$ and satisfy $\chi_3 = \chi_1 \chi_2$.

Proof: It is straightforward to check that (2.19) holds, and hence that Ω^+ is stable, whenever conditions 1, 2, 3, obtain.

Suppose now that (2.19) holds for all sets of indices and for all $g \in \mathbf{G}$. For each $g \in \mathbf{G}$ both $D(g)$ and $D'(g)$ are nonzero matrices, so there exist pairs (s',s) and (p',p) such that $D'(g)_{p'p}$ and $D(g)_{s's}$ are nonzero. It follows that

$$D_{R^{-1}}(g)_{r'r} = D(g)_{r'r} \times [\text{a number independent of } (r',r)], \tag{2.23}$$

and hence that

$$D_{R^{-1}}(g) = \chi_3(g) D(g), \tag{2.24}$$

for all $g \in \mathbf{G}$, where χ_3 is a linear character of $\mathbf{G}$. Similarly

$$D_R(g) = \chi_2(g) D(g), \tag{2.25}$$

$$D_R'(g) = \chi_1(g) D'(g), \tag{2.26}$$

for all $g \in \mathbf{G}$, where χ_1 and χ_2 are linear characters of $\mathbf{G}$. Finally the relation between the linear characters follows from (2.19).

Now when Ω^+ is stable under the action of $\mathbf{G}$ a subrepresentation of $D_R \otimes D \otimes D'$ is defined. If this subrepresentation fails to contain the trivial representation of $\mathbf{G}$, then all members of Ω^+ have zero integral. In that case not only is (2.8) true but its negative also, and in consequence $V_{rps} = 0$ for all (r,p,s). Such a vanishing is an extra selection rule.

3. CHARACTER TESTS AND SPECIAL CASES

Group characters are far more accessible quantities than their associated representations, so it is natural to ask if Theorem 2.1 can be restated in terms of characters. That this is possible is due to the following result, whose proof is elementary.

Lemma 3.1: The unitary matrix T of dimension d is a scalar multiple of the identity if, and only if, Trace $T = dz$, where z is a complex number of unit modulus. Now Lemma 3.1 and Theorem 2.1 together give Theorem 3.1.

Theorem 3.1: If χ and χ' are the characters of D and D', respectively, the following conditions ensure the stability of Ω^+ under $\mathbf{G}$:

1. $\chi'(R^{-1} g R g^{-1}) = d' \chi_1(g)$, (3.1)
2. $\chi(R^{-1} g R g^{-1}) = d \chi_2(g)$, (3.2)
3. $\chi(R g R^{-1} g^{-1}) = d \chi_3(g)$, (3.3)

for all $g \in \mathbf{G}$, where χ_1, χ_2, χ_3, are linear characters of $\mathbf{G}$, and $\chi_3 = \chi_1 \chi_2$.

It should be noted that the equation $\chi(R^{-1} g R) = \chi_2(g) \chi(g)$, for all $g \in \mathbf{G}$, is not equivalent to (3.2), for the three-dimensional irreducible representation of the tetrahedral group T and R a mirror-reflection provides a counterexample.

Although Theorem 3.1 provides sufficient conditions for the stability of Ω^+ it does not give any hint as to the nature of the representation so defined. Indeed each example should be treated separately. There is, however, a frequently occurring situation which deserves special attention, and for which it is possible to explicitly perform the complete reduction of Ω. First note that (3.2) alone implies that $D_R \otimes D \otimes D' = \chi_2(D \otimes D \otimes D')$, which has two subrepresentations corresponding to the symmetrized and antisymmetrized parts of $D \otimes D$. In fact these subrepresentations have carrier spaces spanned by the functions $F_{rps} \pm F_{spr}$, but unfortunately they do not in general coincide with $\Omega^\pm$. There is coincidence, however, if the angles θ_r, α_p, are independent of their indices and satisfy $\exp[i(\theta_r + \alpha_p)] = \gamma$, where $\gamma^2 = 1$. Now restrict attention to this case, so that Ω^+ is the carrier space of $\chi_2[D \otimes D] \otimes D'$ or $\chi_2\{D \otimes D\} \otimes D'$ according as $z = \gamma\omega\beta$ is ± 1. Then the following is true.

Theorem 3.2: Suppose the following hold:

1. $\chi(R^2) = d \exp(i\theta)$, (3.4)
2. $\chi'(R) = d' \exp(i\alpha)$, (3.5)
3. $\chi(R^{-1} g R g^{-1}) = d \chi_2(g)$, (3.6)

for all $g \in \mathbf{G}$, where $\gamma = \exp[i(\theta + \alpha)]$ satisfies $\gamma^2 = 1$ and χ_2 is a linear character of $\mathbf{G}$. Then $V_{rps} = 0$ for all (r,p,s) if

(a) $\gamma\omega\beta = 1$ and $\chi_2[D \otimes D] \otimes D'$ does not contain the trivial representation of $\mathbf{G}$,

or

(b) $\gamma\omega\beta = -1$ and $\chi_2\{D \otimes D\} \otimes D'$ does not contain the trivial representation of $\mathbf{G}$.

Proof: The three conditions (3.4), (3.5), and (3.6), ensure that conditions (3.1), (3.2), and (3.3), of Theo-

rem 3.1 hold, and hence that Ω^+ is stable under the action of G. Also the value of $\gamma\omega\beta$ determines which subrepresentation of $D_R \otimes D \otimes D'$ corresponds to Ω^+, and hence how the vanishings of V_{rps} are distributed.

Note that $\gamma^2 = 1$ is a case not covered by the corollary to Lemma 2.1.

Focus is now directed towards the class of magnetic point groups, and the problem of checking in each case the application of Theorems 3.1 and 3.2.

4. MAGNETIC POINT GROUPS

Magnetic groups and their applications are well described in the literature; in particular the reader is referred to the review article of Bradley and Davies[15] which contains an extensive bibliography in addition to an account of the corepresentation theory of magnetic groups. An even more comprehensive source is the recent book of Bradley and Cracknell.[9] Also, Hamermesh[2] contains a derivation and a listing of all magnetic point groups as well as character tables of point groups and double point groups. The Schoenflies notation is used in the sequel.

There are 90 crystallographic magnetic point groups distributed into two classes: first, the class of grey groups consists of the 32 crystallographic point groups each adjoined by the time-reversal operator θ; secondly, the class of black and white groups consists of the 58 groups of the form $\mathbf{M} = \mathbf{G} + A\,\mathbf{G}$, where the antiunitary element $A = R\theta$ and R does not belong to the unitary group G. In the latter class both G and $\mathbf{G}' = \mathbf{G} + R\,\mathbf{G}$ are crystallographic point groups; indeed the black and white groups can be enumerated by listing for each point group $\mathbf{G}'$ its subgroups G of index 2.

The extra selection rule theory appropriate to the grey groups is that given in Ref. 7; indeed there are many possible extra selection rules, for in almost every case the symmetrized and antisymmetrized squares of an irreducible point group representation have no representations in common.

On looking for possible extra selection rules among the black and white groups a large number of these groups can be dismissed. If G has an Abelian double group—for it is necessary to consider double-valued representations—then (3.2) and (3.3) hold automatically and (3.1), linked with (3.2) and (3.3) through the linear characters χ_1, χ_2, χ_3, only holds if D' is one-dimensional. In such a simple case there can be no extra selection rules because $D_R \otimes D \otimes D'$ is one-dimensional. For the remaining 26 pairs of groups it is always possible to choose R of order 2, and in these cases the conditions of Theorem 3.1 are equivalent to the following:

(A) $\chi'(R) = d' \exp(i\alpha)$,

(B) $\chi(R^{-1} g R g^{-1}) = d\,\chi_2(g)$,

for all $g \in \mathbf{G}$, where χ_2 is a linear character of G.

Proof: For if $R^2 = E$ and (A), (B), hold, then $d\chi_2(g) = \chi(R^2 R^{-1} g R g^{-1}) = \chi(R g R g^{-1}) = \chi(g^{-1} R g R) = \chi(g^{-1} R g R R^{-2}) = \chi(g^{-1} R g R^{-1}) = \chi(R g R^{-1} g^{-1})$, which means that (3.1) and (3.2) hold with $\chi_2 = \chi_3$. Also since $D'(R)$ is scalar diagonal, $\chi'(R^{-1} g R g^{-1}) = d'$, so that (3.1) holds with $\chi_1 =$ unity. Conversely if $R^2 = E$ and the conditions of Theorem 3.1 hold, then certainly (B) holds. Also since $D(R^2)$ is the identity, (3.2) and (3.2) imply that $\chi_2 = \chi_3$, and hence that χ_1 is unity. But then (3.1) means that $D'(R)$ commutes with the irreducible representation D', and that $D'(R)$ is scalar diagonal, as in (A). Of the 26 pairs of groups $(\mathbf{G}', \mathbf{G})$ there are 13 for which the choice $R =$ the space inversion is possible. Clearly for such groups both (A) and (B) hold trivially for all χ', χ, with $\exp(i\alpha) = \pm 1$ and χ_2 the trivial character. These cases therefore come within the scope of Theorem 3.2 in which $\gamma = \exp(i\alpha)$. Earlier remarks about the possible profusion of extra selection rules among the grey groups are also valid here, that is in the case of the following pairs: $(\mathbf{D}_{2h}, \mathbf{D}_2)$, $(\mathbf{D}_{2h}, \mathbf{c}_{2v})$, $(\mathbf{D}_{4h}, \mathbf{D}_4)$, $(\mathbf{D}_{4h}, \mathbf{c}_{4v})$, $(\mathbf{D}_{4h}, \mathbf{D}_{2d})$, $(\mathbf{D}_{3d}, \mathbf{c}_{3v})$, $(\mathbf{D}_{3d}, \mathbf{D}_3)$, $(\mathbf{D}_{6h}, \mathbf{D}_{3h})$, $(\mathbf{D}_{6h}, \mathbf{D}_6)$, $(\mathbf{D}_{6h}, \mathbf{c}_{6v})$, $(\mathbf{T}_h, \mathbf{T})$, $(\mathbf{O}_h, \mathbf{O})$, $(\mathbf{O}_h, \mathbf{T}_d)$.

Of the remaining 13 pairs of groups there are 8 for which only one-dimensional solutions of (A) and (B) can be found—these are dismissed in the context of extra selection rules. The 5 remaining pairs are: $(\mathbf{D}_{3h}, \mathbf{c}_{3v})$, $(\mathbf{D}_{3h}, \mathbf{D}_3)$, both with $R = \sigma_h$; $(\mathbf{D}_6, \mathbf{D}_3)$, $(\mathbf{c}_{6v}, \mathbf{c}_{3v})$, $(\mathbf{D}_{6h}, \mathbf{D}_{3d})$, each with $R = c_{2z}$. For these pairs (A) is only satisfied by the single-valued representations of $\mathbf{G}'$ with $\exp(i\alpha) = \pm 1$, while (B) holds trivially because for all these cases R commutes with each element of G. Again Theorem 3.2 is appropriate.

Before leaving magnetic point groups it is of interest to note the following potentially advantageous point. The conclusions of Theorem 3.1 are not independent of the choices of the element R in the coset $R\mathbf{G}$, so if an initial choice of R fails to lead to nontrivial results then there is freedom to replace it by Rg for any $g \in \mathbf{G}$. Since the matrix elements (2.2) based on R are linear combinations of the corresponding matrix elements based on Rg, there is no loss of selection rules. Indeed there may be some gain through the greater freedom now allowed in the hypotheses of Theorem 3.1. Unfortunately this extra freedom does not in fact result in any new extra selection rules associated with magnetic point groups.

5. MAGNETIC SPACE GROUPS

The problem of testing the hypotheses of Theorem 3.1 becomes more difficult when M is taken to be a magnetic crystallographic space group. Then G and $\mathbf{G}'$ are ordinary crystallographic space groups. The difficulties arise because space groups, being of infinite order, have representations which, although well known, are awkward to manipulate directly. There is a simplification however because, as is shown by Bradley and Cracknell,[9] an irreducible representation of a space group is obtained as an induced allowable representation of an appropriate little group. Since these allowable representations are much easier to handle than full group representations it is fortunate that the hypotheses of Theorem 3.1 can be translated into little group terms. This assertion is based on the following result.

Lemma 5.1: Let the representation D of the group G be induced from the representation Γ of the subgroup L. Let

$$\mathbf{G} = \sum_\sigma p_\sigma L$$

be a coset decomposition of $\mathbf{G}$ with respect to $\mathbf{L}$, where σ ranges over the finite factor space G/L. If $h \in \mathbf{G}$ then $D(h)$ is a scalar multiple of the identity if and only if

1. $h \in \bigcap_\sigma L_\sigma$, where $L_\sigma = p_\sigma L p_\sigma^{-1}$;
2. $\Gamma(p_\sigma^{-1} h p_\sigma)$ is the same scalar matrix for all σ.

Proof: Reference should be made to the definition of the induced representation $\Gamma \uparrow \mathbf{G}$ as given by Bradley.[16]

A similar result is true if the representation $\mathbf{D}'$ of $\mathbf{G}'$ is induced from the representation Γ' of $\mathbf{L}'$, and

$$\mathbf{G}' = \sum_\gamma r_\gamma \mathbf{L}'$$

is a suitable coset decomposition. Then Theorems 3.1, 3.2, can be restated as follows.

Theorem 5.1: Under the hypotheses of Lemma 5.1 and the subsequent text, and if χ, χ', are the characters of Γ, Γ', respectively, then the following conditions ensure the stability of Ω^+:

(A.1) $R^{-1} g R g^{-1} \in \bigcap_\gamma L'_\gamma$, where $L'_\gamma = r_\gamma L' r_\gamma^{-1}$,

(A.2) $R^{-1} g R g^{-1}$ and $R g R^{-1} g^{-1} \in \bigcap_\sigma L_\sigma$, where $L_\sigma = p_\sigma L p_\sigma^{-1}$;

(B.1) $\chi'(r_\gamma^{-1} R^{-1} g R g^{-1} r_\gamma) = f' \chi_1(g)$,

(B.2) $\chi(p_\sigma^{-1} R^{-1} g R g^{-1} p_\sigma) = f \chi_2(g)$,

(B.3) $\chi(p_\sigma^{-1} R g R^{-1} g^{-1} p_\sigma) = f \chi_3(g)$,

for all $g \in \mathbf{G}$, for all r_γ, p_σ, and where χ_1, χ_2, χ_3, are linear characters of $\mathbf{G}$ satisfying $\chi_3 = \chi_1 \chi_2$, and f, f', are the dimensions of Γ, Γ', respectively.

Theorem 5.2: Under the hypotheses of Theorem 5.1 and using the same notation, suppose

(A.1)′ $R \in \bigcap_\gamma L'_\gamma$,

(A.2)′ R^2 and $R^{-1} g R g^{-1} \in \bigcap_\sigma L_\sigma$;

(B.1)′ $\chi'(r_\gamma^{-1} R r_\gamma) = f' \exp(i\alpha)$,

(B.2)′ $\chi(p_\sigma^{-1} R^2 p_\sigma) = f \exp(i\sigma)$,

(B.3)′ $\chi(p_\sigma^{-1} R^{-1} g R g^{-1} p_\sigma) = f \chi_2(g)$,

for all $g \in \mathbf{G}$, where $\gamma = \exp[i(\theta + \alpha)]$ satisfies $\gamma^2 = 1$, and χ_2 is a linear character of $\mathbf{G}$. Then $V_{rps} = 0$ for all (r, p, s) if

(a) $\gamma\omega\beta = 1$ and $\chi_2[D \otimes D] \otimes D'$ does not contain the trivial representation of $\mathbf{G}$,

or

(b) $\gamma\omega\beta = -1$ and $\chi_2\{D \otimes D\} \otimes D'$ does not contain the trivial representation of $\mathbf{G}$.

Further analysis depends on an identification of the groups $\mathbf{L}$, $\mathbf{L}'$, as little groups. Since $\mathbf{G}$ is a space group its little groups $\mathbf{G}^{(\mathbf{k})}$ can be labeled by Brillouin zone vectors k (see for example Bradley[16] or Koster[17]); indeed $\mathbf{G}^{(\mathbf{k})}$ consists of those elements $\{S \mid \mathbf{v}(S) + \mathbf{t}\} \in \mathbf{G}$ which have the property $S\mathbf{k} \equiv \mathbf{k}$ modulo the reciprocal lattice of $\mathbf{G}$. But if $\mathbf{G}^{(\mathbf{k})}$ is the group of k, then $\mathbf{G}_\sigma^{(\mathbf{k})}$ is the group of $p_\sigma \mathbf{k}$. It follows that $\bigcap_\sigma \mathbf{G}_\sigma^{(\mathbf{k})}$ is the intersection of the groups of the prongs of the star of k. Here the star of k means the set of all vectors obtainable from k by the action of $\mathbf{G}$. For example if $\mathbf{k} \equiv 0$ then $\mathbf{G}^{(\mathbf{k})} = \mathbf{G}$, and the intersection group is $\mathbf{G}$ itself. At the other extreme if k is a general vector then $\mathbf{G}^{(\mathbf{k})}$ is the translation subgroup $\mathbf{T}$ of $\mathbf{G}$ and coincides with the intersection subgroup. As a practical aid it is worth noting that each factor group $\mathbf{G}_\sigma^{\mathbf{k}}/\mathbf{T}$ is a point group, and that

$$(\bigcap_\sigma \mathbf{G}_\sigma^{(\mathbf{k})})/\mathbf{T} = \bigcap_\sigma (\mathbf{G}_\sigma^{(\mathbf{k})}/\mathbf{T}), \tag{5.1}$$

so it is possible to compute $\bigcap_\sigma \mathbf{G}_\sigma^{(\mathbf{k})}$ at the point group level by examining the symmetrical unit cell of $\mathbf{G}$. Thus the (A)-conditions of Theorems 5.1 and 5.2 are straightforward to check. Also since the allowable representations of the little groups of a space group are always diagonal for the translational elements the (B)-conditions are essentially only conditions on point group representations and k vectors. For cases which satisfy the (A)-conditions, the (B)-conditions are straightforward, though tedious to check. As a final theoretical point the symmetrized and antisymmetrized representations which appear in Theorem 5.2 may be computed by a method due to Mackey,[18] which is explained and amplified in the context of space groups by Bradley and Davies.[7]

For a simple example consider the groups $\mathbf{G} = \mathbf{T}^2_d$ and $\mathbf{G}' = \mathbf{O}^5_h$, where R is the space inversion. Let $g = \{P \mid \mathbf{t}\}$ be the general element of $\mathbf{G}$, then $R^{-1} g R g^{-1}$ is $\{E \mid -2\mathbf{t}\}$, where E is the point group identity. It follows that (A.1) and (A.2) are automatically satisfied. Since the space inversion is self-inverse the (B)-conditions of Theorem 5.1 can only be satisfied if $\chi_1(g) = 1$ for all $g \in \mathbf{G}$.

Now if Γ, Γ', are associated with the vectors k, k′, respectively, it is easy to show that

$$\chi(p_\sigma^{-1} R^{-1} g R g^{-1} p_\sigma) = f \exp[i(p_0 \mathbf{k} \cdot 2\mathbf{t})], \tag{5.2}$$

and

$$\chi'(r_\gamma^{-1} R^{-1} g R g^{-1} r_\gamma) = f' \exp[i(r_\gamma \mathbf{k}' \cdot 2\mathbf{t})]. \tag{5.3}$$

for all $g \in \mathbf{G}$. It follows that the (B)-conditions can be satisfied whenever the prongs of the star of k differ from k by no more than a half reciprocal lattice vector, and similarly for k′. This restricts attention to those vectors k, k′, which are associated with the points Γ, X, L, of the Brillouin zone (see for example Fig. 4 of Bouckaert, Smoluchowski, and Wigner[19]). Under these circumstances Theorem 5.2 is applicable with $\exp(i\alpha) = \pm 1$, $\exp(i\sigma) = 1$, and χ_2 the trivial character of $\mathbf{G}$. In fact now D' is an irreducible representation of $\mathbf{G}$. Now reference to the tables in the appendix of Bradley and Davies[7] yields a wealth of extra selection rules.

For consider a transition between an initial state L_2 and $(I\theta)L_2$ which is effected by the electric dipole operator which transforms like a polar vector according to the representation Γ_5 of $T_d^2 (= \mathbf{G})$. Now $\beta = 1$, $\gamma = -1$, and $\omega = 1$, so $z = \beta\gamma\omega = -1$. Also

$$[(L_2 \uparrow \mathbf{G}) \otimes (L_2 \uparrow \mathbf{G})] = \Gamma_1 \uparrow \mathbf{G} + \Gamma_5 \uparrow \mathbf{G} + X_1 \uparrow \mathbf{G} + X_4 \uparrow \mathbf{G},$$

$$\{(L_2 \uparrow \mathbf{G}) \otimes (L_2 \uparrow \mathbf{G})\} = X_5 \uparrow \mathbf{G},$$

$$(X_5 \uparrow \mathbf{G}) \otimes (\Gamma_5 \uparrow \mathbf{G}) = \sum_{i=1}^{5} (X_i \uparrow \mathbf{G}).$$

Now $\beta\gamma\omega = -1$, χ_2 is the trivial character of $\mathbf{G}$, and $\chi_2(X_5 \uparrow \mathbf{G}) \otimes (\Gamma_5 \uparrow \mathbf{G})$ does not contain the trivial representation of $\mathbf{G}$, Theorem 5.2 implies that $V_{rps} = 0$ for all (r, p, s). Since $\Gamma_5 \uparrow \mathbf{G}$ is contained in $(L_2 \uparrow \mathbf{G}) \otimes (L_2 \uparrow \mathbf{G})$, the

above result is a true extra selection rule. Corepresentation theory does not give this result because the corepresentation associated with $\Gamma_5 \uparrow \mathbf{G}$ is contained in the square of the corepresentation of $L_2 \uparrow \mathbf{G}$.

APPENDIX

The following analysis is due to Dr. R. Shaw of the University of Hull (private communication).

The crucial assumption made in Sec. 2 was that the operators V_p satisfy $(\theta V_p \theta^{-1})^+ = \beta V_p$, where $\beta = \pm 1$ independently of p. Since this point is central to the argument it certainly deserves some explanation. In fact it is not an unreasonable assumption to make and is supported by the following group theoretical argument.

Let the operators V_p, $p = 1, 2, \ldots, d'$, which form a basis set for the vector space V, transform according to the irreducible representation D' of $\mathbf{G}'$. If $\overline{V}$ is the vector space generated by the operators $\overline{V}_p = (\theta V_p \theta^{-1})^+$, for all p, then $\overline{V}$ also carries the representation D'. For if $O(g)$ is the unitary operator corresponding to the element $g \in \mathbf{G}'$, the action of g on $\overline{V}_p$ is given by

$$O(g)\,\overline{V}_p\,O(g)^{-1} = [O(g)\,\theta V_p \theta^{-1} O(g)^{-1}]^+, \tag{A1}$$

using the unitarity of $O(g)$

$$= [\theta O(g)\, V_p O(g)^{-1}\, \theta^{-1}]^+, \tag{A2}$$

since θ commutes with all geometric operators,

$$= \sum_q [\theta D'(g)_{qp} V_q \theta^{-1}]^+, \tag{A3}$$

letting g act on V_p,

$$= \sum_q D'(g)_{qp} \overline{V}_q, \tag{A4}$$

as required. It follows that the intersection of V and $\overline{V}$ is an invariant subspace of D', and since D' is irreducible that either $V \cap \overline{V} = \{O\}$ or $V = \overline{V}$. In the latter case there exists a nonsingular matrix $c = (c_{pq})$, where

$$\overline{V}_p = \sum_q c_{qp} V_q, \tag{A5}$$

for all p. But the operators $\overline{V}_p$ transform under $\mathbf{G}'$ according to the representation D', from which it follows that c commutes with $D'(g)$ for all $g \in \mathbf{G}'$. This is only possible if c is a scalar multiple of the identity, so (A5) can be rewritten as $\overline{V}_p = \beta V_p$, for all p, for some complex number β. In fact $\beta^2 = 1$, for on the one hand $(\theta \overline{V}_p \theta^{-1})^+ = \beta(\theta V_p \theta^{-1})^+ = \beta^2 V_p$, and on the other hand $(\theta \overline{V}_p \theta^{-1})^+ = V_p$.

Group theoretically the other possibility $V \cap \overline{V} = \{O\}$ cannot be excluded. However it is an empirical fact that many physically relevant operators do conform to (2.7). For example, the position and momentum observables, the electric and magnetic dipole operators, all fall within its scope. What the above analysis shows is that it is sufficient to find one nonzero vector operator in the space $V \cap \overline{V}$ for (2.7) to hold.

ACKNOWLEDGMENTS

The author would like to express his thanks to C. J. Bradley for the suggestion which led to this paper, for his great help and encouragment towards its completion. Help in the preparation of the manuscript was gratefully received from B. L. Davies, D. H. Lewis, and Miss Patricia Gard.

[1] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959).
[2] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[3] M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
[4] J. S. Lomont, *Applications of Finite Groups* (Academic, New York, 1959).
[5] G. Ya. Lyubarskii, *The Application of Group Theory in Physics* (Pergamon, Oxford, 1960).
[6] M. Lax, *International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and The Physical Society, London, 1962).
[7] C. J. Bradley and B. L. Davies, J. Math. Phys. **11**, 1536 (1970).
[8] A. Aviran and J. Zak, J. Math. Phys. **9**, 2138 (1968).
[9] C. J. Bradley and A. P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon Press, Oxford, 1971).
[10] B. A. Tavger and C. M. Zaitzev, Sov. Phys.-JETP **3**, 430 (1956).
[11] W. Opechowski and R. Guccione, in *Magnetism*, edited by G. T. Rado and H. Suhl (Academic, New York, 1965), Vol. IIA.
[12] J. O. Dimmock and R. G. Wheeler, in *The Mathematics of Physics and Chemistry*, edited by H. Margenau and G. M. Murphy (Van Nostrand, New York, 1964), Vol. 2.
[13] R. R. Birss, *Symmetry and Magnetism* (North-Holland, Amsterdam, 1964).
[14] S. Bhagavantam, *Crystal Symmetry and Physical Properties* (Academic, New York, 1966).
[15] C. J. Bradley and B. L. Davies, Rev. Mod. Phys. **40**, 359 (1968).
[16] C. J. Bradley, J. Math. Phys. **7**, 1145 (1966).
[17] G. F. Koster, Solid State Phys. **5**, 173 (1957).
[18] G. W. Mackey, Am. J. Math. **75**, 387 (1953).
[19] L. P. Bouckaert, R. Smoluckowski, and E. P. Wigner, Phys. Rev. **50**, 58 (1936).

On the exponential superpropagator

C. G. Bollini*

Departamento de Física, Universidad Nacional de La Plata, La Plata, Argentina

J. J. Giambiagi†

International Centre for Theoretical Physics, Trieste, Italy
(Received 9 August 1971; revised manuscript received 28 February 1972)

Series developments for distributions of the type $R^{a}\exp(-\kappa^2/R)$ (with $R = r^2 - t^2$) are discussed in the light of the Gel'fand–Shilov method for the definition of causal distributions. It is shown that there is no arbitrariness in the series. In Sec. I the concept of causal distribution is reviewed, in Sec. II the series developments are given, fixing the values of the constants in the $\Box^n\delta$ terms. In Sec. III the Fourier transforms are computed. In Sec. IV we obtain the series development for the Fourier transform of $\exp(-\kappa^2/R)\partial_\mu\partial_\nu(R^{-1})$.

INTRODUCTION

Recently, the exponential superpropagator has come to the attention of physicists working with nonpolynomial Lagrangians in the form $e^{\kappa\phi}$.[1] The corresponding superpropagator has been discussed by Volkov[2] and in a different context by Lehmann and Pohlmeyer.[3]

We want to discuss this subject in the light of the Gel'fand–Shilov[4] method for the definition of causal distribution. Our method should be compared with that of Blomer and Constantinescu[5] where they use explicit regularization to construct the exponential superpropagator.

I. CAUSAL DISTRIBUTIONS

Let us first recall what is the meaning of "causal" distribution. We start from a well-defined distribution f which is a function of a positive definite quadratic form (see also Ref. 6), i.e.,

$$(f(r^2 + t^2),\ \varphi(xyzt)) = \psi(\varphi). \tag{1}$$

We now make a positive dilatation in the variable t (or, equivalently, we change the metric to $1, 1, 1, a^2$) obtaining

$$(f(r^2 + a^2t^2),\ \varphi(xyzt)) = \psi(\varphi, a). \tag{2}$$

ψ is now an analytic function of the parameter a, and the analytic continuation to $a = \pm i + \epsilon$ (i.e., $a^2 = -1 + 2i\epsilon$) defines, respectively, the causal and anticausal distributions $f(r^2 - t^2 \pm i\epsilon)$.[4]

It is easy to see, following this procedure, that the Fourier transform can be calculated as the Hankel transform of the Euclidean $f(r^2 + t^2)$ (see Ref. 7). We shall use this method to discuss the distribution

$$\exp(-\kappa^2/R \pm i0) \quad \text{where } R = r^2 - t^2 = x^2 + y^2 + z^2 - t^2. \tag{3}$$

According to the definition previously given, this distribution should be understood as

$$\psi = (\exp(-\kappa^2/R \pm i0)) = \lim_{a\to\pm i}(\exp[-\kappa^2/(r^2 + a^2t^2)],\ \varphi(xyzt))$$
$$= \lim_{a\to\pm i}(\exp[-\kappa^2/(r^2 + t^2)]\,\varphi[xyz(t/a)])$$

or, changing to polar variables in Euclidean four dimensions,

$$\psi = \lim_{a\to\pm i} a^{-1}\int \rho^3\,d\rho\,\exp(-\kappa^2/\rho^2)\int d\Omega\,\varphi[xyz(t/a)]$$
$$= \lim_{a\to\pm i} a^{-1}\int \rho^3\,d\rho\,\exp(-\kappa^2/\rho^2)\,\varphi(\rho, a), \tag{4}$$

where

$$\varphi(\rho, a) = \int d\Omega\,\varphi(xyz(t/a)).$$

This formula shows us explicitly that any causal distribution can be obtained from the corresponding Euclidean distribution by a suitable analytic continuation in the metric. As a consequence, it must be noted from (4) that the exponential superpropagator is well defined. Naively, one could think that $\exp(-\kappa^2/R \pm i0)$ has problems at the cone $R = 0$ when approaching this limit from $R < 0$, but it is clear from (4) that these negative values do not appear in the domain of integration. In particular, one can see that $\exp(-\kappa^2/R \pm i0)$ and all its derivatives are zero at the origin. As a matter of fact, in the Euclidean region, the exponential superpropagator is a smooth well-behaved function.

In the same way, one can see that $(R \pm i0)^\alpha \exp(-\kappa^2/R \pm i0)$ is also a well-defined distribution for any value of α, real or complex.

II. SERIES DEVELOPMENT

From now on we shall work with $R + i0$ which we will simply call R. For $(R - i0)$, the complex conjugate must be taken.

We want now to find the series development of the exponential superpropagator. The naive development is

$$\exp(-\kappa^2/R) = \sum_0^\infty \frac{1}{n!}\left(\frac{-\kappa^2}{R}\right)^n.$$

However, Gel'fand–Shilov[4] have proved that R^{-n} is not well-defined for $n \geq 2$. In these cases one defines a finite part

$$P_f R^{-n} = \frac{d}{d\alpha}(\alpha + n)R^\alpha\Big|_{\alpha=-n}, \tag{5}$$

but it remains an intrinsic indeterminacy proportional to the iterated d'Alembertian of the δ function. In other words,

$$R^{-n-2} \Rightarrow P_f R^{-n-2} + c_n\Box^n\delta,$$

where the first term of the rhs is well defined and the second term contains an arbitrary constant. So, we must expect that the series for the superpropagator will also contain a series in the iterated d'Alembertians with coefficients which we want to determine. In order to do that we take advantage of the given definition and consider the Euclidean development

$$\exp(-\kappa^2/r^2)=\sum_0^\infty \frac{(-\kappa^2)^n}{n!} P_f (r^2)^{-n}+\sum_0^\infty c_n \Delta^n \delta, \tag{6}$$

where

$$r^2=x^2+y^2+z^2+t^2, \quad \Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial t^2}+\frac{\partial^2}{\partial t^2}.$$

We now apply both members to the following family of trial functions:

$$\varphi(r)=r^{2\rho}\exp(-\beta^2 r^2).$$

The left-hand side of (6) gives[8]

$$[\exp(-\kappa^2/r^2), r^{2\rho}\exp(-\beta r^2)]=\int_0^\infty dr\, r^3 \exp(-\kappa^2/r^2)\, r^{2\rho} \times \exp(-\beta^2 r^2)\int d\Omega$$

$$=2\pi^2(\kappa/\beta)^{\rho+2} K_{\rho+2}(2\kappa\beta). \tag{7}$$

The first series in the right-hand side of (6) gives

$$2\pi^2 \sum_0^\infty \frac{(-\kappa^2)^n}{n!} P_f \int_0^\infty dr\, r^{2\alpha+3+2\rho}\exp(-\beta^2 r^2)\Big|_{\alpha=-n}$$

$$=2\pi^2 \sum_0^\infty \frac{(-\kappa^2)^n}{n!} P_f \frac{\Gamma(\rho+\alpha+2)}{2(\beta^2)^{\rho+\alpha+2}}=2\pi^2 \sum_0^{\rho+1} \frac{(-\kappa^2)^n}{n!}\frac{\Gamma(\rho+2-n)}{2(\beta^2)^{\rho+2-n}}$$

$$+2\pi^2 \sum_0^\infty \frac{(-\kappa^2)^{n+\rho+2}}{2\cdot n!(n+\rho+2)!}(-\beta^2)^n[\psi(n+1)-\ln\beta^2]. \tag{8}$$

The last series of (6) gives

$$\sum_{n=0}^\infty c_n \int d^4x\, \Delta^n\delta\, r^{2\rho}\exp(-\beta r^2)=\sum_n c_n \Delta^n[r^{2\rho}\exp(-\beta^2 r^2)]\Big|_{r=0}$$

$$=\sum_0^\infty \frac{c_n 4^n n!(n+1)!(-\beta^2)^{n-\rho}}{(n-\rho)!}, \tag{9}$$

where $\psi(n)$ is the Euler function and use has been made of

$$\lim_{\alpha\to -n}(\alpha+n)\,\Gamma(\rho+2+\alpha)=(-1)^{n-\rho}/\Gamma(n-\rho-1), \tag{10}$$

$$\lim_{\alpha\to -n}(\alpha+n)\,\Gamma(\rho+2+\alpha)=(-1)^{n-\rho}\psi(n-\rho-1)/\Gamma(n-\rho-1). \tag{11}$$

Comparing (8) and (9) with the series development of (7), namely,

$$2\pi^2\left(\frac{\kappa}{\beta}\right)^{\rho+2} K_{\rho+2}(2\beta\kappa)=2\pi^2\left(\sum_0^{\rho+1}\frac{(-1)^n(\rho-n+1)!(\kappa^2)^n}{2\cdot n!(\beta^2)^{\rho+2-n}}\right.$$

$$+(-1)^\rho\sum_0^\infty \frac{(\kappa^2)^{\rho+2+n}}{n!(\rho+2+n)!}(\beta^2)^n[-\ln(\beta\kappa)^2$$

$$\left.+\psi(n+1)+\psi(\rho+n+3)]\right),$$

we get

$$c_n=\frac{\pi^2(-\kappa^2)^{n+2}[\psi(n+3)-\ln\kappa^2]}{4^n n!(n+1)!(n+2)!}. \tag{12}$$

We finally obtain

$$\exp(-\kappa^2/r^2)=\sum_0^\infty \frac{(-\kappa^2)^n}{n!} P_f (r^2)^{-n}$$

$$+\pi^2\sum_0^\infty \frac{(-\kappa^2)^{n+2}[\psi(n+3)-\ln\kappa^2]}{4^n n!(n+1)!(n+2)!}$$

$$\Delta^n\delta, \tag{13}$$

which when written in the hyperbolic metric has the form (see note added in proof)

$$\exp(-\kappa^2/R)=\sum_0^\infty \frac{(-\kappa^2)^n}{n!} P_f R^{-n}-i\pi^2$$

$$\times\sum_0^\infty \frac{(-\kappa^2)^{n+2}[\psi(n+3)-\ln\kappa^2]}{4^n n!(n+1)!(n+2)!}\Box^n\delta, \tag{14}$$

where the i comes from the analytic continuation

$$\delta(t)\Rightarrow\delta(at) \quad (1/i)\delta(t).$$

We must point out that every term of Eq. (14) is perfectly well defined containing no arbitrariness whatsoever.

One can easily obtain the development of $R^{-s}\exp(-\kappa^2/R)$ by taking successive derivatives of (14) with respect to (κ^2) or s successive integrations if s is negative. However, we shall deduce them as particular cases of a more general expression.

We note first that

$$e^\chi=G_{01}^{10}(-\chi/0)=G_{10}^{01}[-1/\chi\,|\,1], \tag{15}$$

where G is the Meijer G function (Ref. 9). For our case

$$\exp(-\kappa^2/R)=G_{10}^{01}\left(\frac{R}{\kappa^2}\,\middle|\,1\right)=\frac{1}{2\pi i}\int_c \Gamma(s)\left(\frac{R}{\kappa^2}\right)^s ds, \tag{16}$$

where c runs from $-i\infty$ to $+i\infty$ leaving on the left the poles of $\Gamma(s)$ and those of R^s as R^α is analytic in α with poles for $\alpha=-2,\ -3,\ldots$ and residues

$$\mathrm{Res}\,R^\alpha\Big|_{\alpha=-2-k}=-i\pi^2\Box^k\delta/k!(k+1)!4^k. \tag{17}$$

It should be noted here that the usual definition of the Meijer G function[9] does not consider the poles of χ^s. Nevertheless, for distribution, these poles must be taken into account to obtain the correct results. In this way the evaluation by residues of the integral (16) gives back (14) when use is made of (17).

Now, multiplying (16) by R^α, we obtain

$$R^\alpha\exp(-\kappa^2/R)=\frac{(\kappa^2)^\alpha}{2\pi i}\int_c \Gamma(s)\left(\frac{R}{\kappa^2}\right)^{s+\alpha} ds$$

$$=\frac{(\kappa^2)^\alpha}{2\pi i}\int \Gamma(s-\alpha)\left(\frac{R}{\kappa^2}\right)^s ds$$

$$=(\kappa^2)^\alpha G_{10}^{01}\left(\frac{R}{\kappa^2}\,\middle|\,1+\alpha\right). \tag{18}$$

The evaluation by residues now gives

$$R^\alpha\exp(-\kappa^2/R)=\sum_0^\infty\frac{(-\kappa^2)^n}{n!}R^{\alpha-n}-i\pi^2\sum_0^\infty\frac{(\kappa^2)^{n+\alpha+2}}{4^n n!(n+1)!}$$

$$\Gamma(-n-\alpha-2)\Box^n\delta. \tag{19}$$

This equation can be checked with a trial function similar to the one already used.

It is easy to see that when α tends to an integer, the pole of each term of one of the series in the right-hand side of (19) is exactly compensated by another pole of the other series, leaving only the finite part of each term. In particular, for $\alpha \to 0$ we re-obtain (14). For $\alpha \to s$, s integer ≥ -2, it gives

$$R^s \exp(-\kappa^2/R) = \sum_0^{s+1} \frac{(-\kappa^2)^n}{n!} R^{s-n} + \sum_0^{\infty} \frac{(-\kappa^2)^{n+s+2}}{(n+s+2)!} P_f R^{-n-2} - i\pi^2 \sum_0^{\infty} \frac{(-\kappa^2)^{n+s+2}[\psi(n+s+3) - \ln\kappa^2]\Box^n\delta}{4^n n!(n+1)!(n+s+2)!}. \tag{20}$$

For $\alpha \to s \leq -2$

$$R^s \exp(-\kappa^2/R) = -i\pi^2 \sum_0^{-s-3} \frac{(\kappa^2)^{n+s+2}\,\Gamma(-s-n-2)\Box^n\delta}{4^n n!(n+1)!} + \sum_0^{\infty} \frac{(-\kappa^2)^n}{n!} P_f R^{s-n} - i\pi^2 \sum_0^{\infty} \frac{(-\kappa^2)^n[\psi(n+1) - \ln\kappa^2]\Box^{n-s-2}\delta}{4^{n-s-2}(n-s-2)!(n-s-1)!n!}. \tag{21}$$

III. FOURIER TRANSFORMS

The Fourier transform of all the expressions just obtained can be easily evaluated either by term-by-term transformation or by using the Hankel transform of Meijer's functions[10]

$$\mathfrak{F}\left[(\kappa^2)^\alpha G_{10}^{01}\left(\frac{R}{\kappa^2}\middle| 1+\alpha\right)\right] = -i\pi^2(\kappa^2)^{\alpha+2} G_{03}^{20}\left(\frac{\kappa^2 P}{4}\middle| -\alpha-2, 0, -1\right).$$

So

$$\mathfrak{F}[R^\alpha \exp(-\kappa^2/R)] = -i\pi^2 \sum_0^{\infty} \frac{(-\kappa^2)^n\,\Gamma(\alpha-n+2)}{n!\,\Gamma(n-\alpha)}\left(\frac{P}{4}\right)^{n-\alpha-2} - i\pi^2 \sum_0^{\infty} \frac{(\kappa^2)^{n+\alpha+2}\,\Gamma(-n-\alpha+2)}{n!(n+1)!}\left(-\frac{P}{4}\right)^n, \tag{22}$$

when $\alpha \to s$ integer with $s \geq -2$

$$\mathfrak{F}[R^s \exp(-\kappa^2/R)] = (2\pi)^4 \sum_0^{s} \frac{(-\kappa^2)^n(-\Box)^{s-n}\delta}{n!} - \frac{i\pi^2(-\kappa^2)^{s+1}}{(s+1)!}\frac{4}{P} - i\pi^2 \sum_0^{\infty} \frac{(-\kappa^2)^{n+s+2}(-P/4)^n[\psi(n+1)+\psi(n+2)+\psi(n+s+3) - \ln(\kappa^2 P/4)]}{n!(n+1)!(n+s+2)!} \tag{23}$$

and for $s \leq -2$

$$\mathfrak{F}[R^s \exp(-\kappa^2/R)] = -i\pi^2 \sum_0^{-s-3} \frac{(\kappa^2)^{n+s+2}\,\Gamma(-s-n+2)(-P)^n}{4^n n!(n+1)!} - i\pi^2 \sum_0^{\infty} \frac{(-\kappa^2)^n(-P/4)^{n-s-2}[\psi(n+1)+\psi(n-s)+\psi(n-s-1) - \ln(\kappa^2 P/4)]}{n!(n-s-2)!(n-s-1)!} \tag{24}$$

It should be noted here that the use of Volkov's formula given in Ref. 11 would give a wrong result when applied to this case as it reproduces the series in the rhs but it does not include the finite sum Σ^{-s-3} which is essential to get the result given in (24). See also Ref. 12.

IV. CALCULUS OF DERIVATIVES

The above-mentioned treatment allows us to solve the problem raised by Salam,[1] namely the evaluation of the Fourier transform of

$$\exp(-\kappa^2/R)\partial_\mu\partial_\nu(\kappa/R). \tag{25}$$

According to the given definition of causal distributions this product is well defined for any R and, in particular, is zero at the origin. One can easily verify the following identity:

$$\exp(-\kappa^2/R)\partial_\mu\partial_\nu(\kappa^2/R) = 2(\partial_\mu\partial_\nu - \tfrac{1}{4}g_{\mu\nu})(R/\kappa^2)\exp(-\kappa^2/R). \tag{26}$$

Note that in particular the trace of both members of (26) is zero. It must be noted that, according to the definition $\delta(xyzt)\exp(-\kappa^2/R)$ is zero.

The series development of the rhs can be written by means of Eq. (20) for $s=1$. To get the Fourier transform we use (23) and obtain

$$\mathfrak{F}[\exp(-\kappa^2/R)\partial_\mu\partial_\nu(\kappa^2/R)] = 4i\pi^2\kappa^2(P_\mu P_\nu - \tfrac{1}{4}g_{\mu\nu}) \times\left(1 - 2\sum_0^{\infty} \frac{(\kappa^2 P/4)^{n+1}[\psi(n+1)+\psi(n+2)+\psi(n+4) - \ln(\kappa^2 P/4)]}{n!(n+1)!(n+3)!}\right). \tag{27}$$

One can see that (27) goes to zero for $\kappa^2 \to 0$. This is also true for (26) in spite of the apparent pole in κ^2 in the r. h. s. which is eliminated by the differential operator.

V. CONCLUSIONS

By using the definition of causal distributions given at the beginning of the paper one can see that $\exp(-\kappa^2/R)$ is well defined and without problems at the origin. It has a uniquely determined series development without any arbitrary constants.

Note that the series development (14) contains an infinite series in $\Box^n\delta$, contrary to what one would expect from the principle of minimal singularity.[3] The same is valid for distributions of the form $R^\alpha \exp(-\kappa^2/R)$ (α, arbitrary). Further, no regularization is needed to write its series development (compare with Ref. 5). We also found the Fourier transforms of these distributions together with the corresponding series development. If one is interested in the series development of $\exp[\kappa^2/$

$(r^2 - t^2)]$ one can write it as $\exp[-\kappa^2/(t^2 - r^2)]$ and consider the Euclidean metric $t^2 + a^2r^2$. One then repeats the same argument that led to (14) by making at the end $a \to -i$. The result obtained in this way is equal to the one obtained by replacing $\kappa \to i\kappa$. (For anticausal $a \to +i$, $\kappa \to -i$.)

Equations (13) and (14) are a particular example of a general result which follows from the definition of causal distributions, namely, that any equation between them is the analytic continuation of a corresponding equation in Euclidean metric so that any theorem relating causal distributions can be deduced from a similar theorem in Euclidean metric. Series in $\Box^n \delta$ like those appearing in Eq. (14) have been extensively studied by Efimov. [13]

Finally, we want to remark that the logarithmic dependence of the coupling constant appears only when the number of dimensions is even. This follows from Gel'fand–Shilov's[14] result for $R^{-\lambda}$ which have poles for $\lambda = (d/2) + n$, where d equals the number of dimensions. For this reason the naive development of the exponential is valid for d odd, and the Fourier transform can be evaluated term by term without any problem[15] (compare with Ref. 5, p. 191).

Note: When this paper was finished we learned that G. Lazarides and A. A. Patani had obtained the same result [Eq. (14)] following a similar method.

ACKNOWLEDGMENTS

One of the authors (JJG) is indebted to Dr. J. Strathdee and Dr. C. J. Isham for useful discussions on the subject. He also wishes to thank Professor Abdus Salam and Professor P. Budini as well as the International Atomic Energy Agency and UNESCO for hospitality at the International Center for Theoretical Physics, Trieste.

*Also Consejo Nacional de Investigaciones Cientificas, Argentina.
†On leave of absence from Departamento de Física, Universidad Nacional de La Plata, Argentina.

[1] A. Salam, ICTP, Trieste, preprint IC/71/3; C. J. Isham, A. Salam, and J. Strathdee, ICTP, Trieste, preprint IC/71/13 (to appear in Phys. Letters).
[2] M. K. Volkov, JINR preprint E2-3472 (1967).
[3] H. Lehmann and K. Pohlmeyer, Commun. Math. Phys. **20**, 101 (1971).
[4] I. M. Gel'fand and G. E. Shilov, *Les Distributions* (Dunod, Paris, 1962).
[5] R. J. Blomer and F. Constantinescu, Nucl. Phys. B **27**, 173 (1971).
[6] H. H. Aly, W. Guttinger, and H. J. W. Müller, in *Lectures on Particles and Fields* (Gordon and Breach, New York, 1970), pp. 309–322.
[7] C. G. Bollini, J. J. Giambiagi, and A. González Domínguez, J. Math. Phys. **6**, 165 (1965).
[8] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965), p. 340.
[9] Bateman Project, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1, p. 206. See also: P. K. Mitter, Commun. Math. Phys. **20**, 251 (1971).
[10] Bateman Project, *Tables of Integral Transforms* (McGraw-Hill, New York, 1953), Vol. 2, p. 91.
[11] M. K. Volkov, Ann. Phys. (N.Y.) **49**, 202 (1968).
[12] C. G. Bollini and J. J. Giambiagi, Nuovo Cimento Lett. **4**, 711 (1970).
[13] G. V. Efimov, in Meeting on Renormalization Theory, ICTP, Trieste 1969 (IC/69/121).
[14] I. M. Gel'fand and G. E. Shilov, *Les Distributions* (Dunod, Paris, 1962), p. 269.
[15] I. M. Gel'fand and G. E. Shilov, *Les Distributions* (Dunod, Paris, 1962), p. 278.

KS-related *f-g* couples as exact vacuum solutions of Salam's two-tensor theory*

Helmuth K. Urbantke

Institut für Theoretische Physik, Universität Wien, Vienna, Austria
(Received 22 February 1972)

The exact solutions of the vacuum field equations of the two-tensor theory of gravitation proposed by Salam *et al.* are determined for the case that the two tensors differ only by the tensor product of a null vector field with itself. It is found that this null vector field must be geodesic with vanishing shear, expansion, and twist, and represents a field of multiple principal null directions for both Weyl tensors. As a consequence, the solutions can be derived from the results of a paper by Kundt.

I. INTRODUCTION

The field equations of the two-tensor theory of gravitation proposed by Salam *et al.*[1] are in general rather complicated, even *in vacuo*, so that it is hard to find exact solutions; but they simplify when special relations between the two tensors $f_{\mu\nu}$, $g_{\mu\nu}$ are assumed.

We will call a couple "$f_{\mu\nu}=g_{\mu\nu}=$ any solution of Einstein's vacuum equations" a trivial solution, and are interested in obtaining nontrivial couples $f_{\mu\nu}$, $g_{\mu\nu}$ of exact solutions. Pirani[2] has found that the assumption of a conformal relationship $f_{\mu\nu}=\Omega^2 g_{\mu\nu}$ leads to $\Omega^2\equiv 1$ when inserted into Salam's vacuum equations.

Here we want to consider a relationship of the form

$$f_{\mu\nu}=g_{\mu\nu}+2\mathfrak{K}k_\mu k_\nu, \tag{1}$$

where k_μ is a nonvanishing field of null vectors with respect to $g_{\mu\nu}$:

$$k^\mu k_\mu = k^\mu g_{\mu\nu}k^\nu = 0. \tag{2}$$

[At this point there is no gain of generality by writing a nonconstant $\mathfrak{K}$ in (1); but we will, later on, impose further conditions on k_μ, but do not want to restrict (1) by them, which is achieved by having $\mathfrak{K}$ explicitly at our disposal.]

A relationship of this kind between two symmetric tensors will be called a Kerr–Schild (KS) relationship: the class of metrics first considered by Kerr and Schild[3] is KS-related to the Minkowski metric, according to our terminology.

When this ansatz is introduced into the vacuum equations of Ref. 1, one obtains simply

$$G_{\mu\nu}(g)=-(\kappa_g^2/\kappa_f^2)M^2\mathfrak{K}k_\mu k_\nu,\quad G_{\mu\nu}(f)=M^2\mathfrak{K}k_\mu k_\nu, \tag{3}$$

where $G_{\mu\nu}(g)$ and $G_{\mu\nu}(f)$ are the Einstein tensors constructed from the "metrics" $g_{\mu\nu}$ and $f_{\mu\nu}$. It is the purpose of this note to derive the solutions of the system of equations (1), (2), (3). The result is that these solutions can almost be copied from a paper by Kundt,[4] after one has proven that the null congruence determinedby k_μ must be free of shear, twist, and expansion, and that the fields $g_{\mu\nu}$, $f_{\mu\nu}$ have algebraically special (or degenerate) Weyl tensors with k_μ a multiple (Debever–Penrose) principal null direction.

This proof is given in Sec. 3, after some auxiliary considerations in Sec. 2. In Sec. 4 we present the final "canonical" form of the solutions.

Finally, we mention that the same results hold in the two-tensor theory of massive gravitation which was obtained[5] by "covariantization" of the theory of Freund *et al.*[6]

II. AUXILIARY CONSIDERATIONS

The simplification that takes place when (1), (2) are inserted into Salam's equations comes from the fact that the "mixing terms" in these equations contain the difference $f_{\mu\nu}-g_{\mu\nu}$, and that we have the further relations between the inverses and determinants

$$f^{\mu\nu}=g^{\mu\nu}-2\mathfrak{K}k^\mu k^\nu \tag{4}$$

$$\det f_{\mu\nu}=\det g_{\mu\nu}. \tag{5}$$

k_μ is null with respect to $f_{\mu\nu}$ also, and the index on k can be raised and lowered by $f^{\mu\nu}$, $f_{\mu\nu}$ as well. Hence "KS-related by a vector field k_μ" is an equivalence relation, by which all metrics for which k_μ is null may be divided into equivalence classes ("KS classes"). Let us mention here some properties the proofs of which are simple and will not be carried out explicitly—some of them emerging in later calculations. First, if k_μ is geodesic for $g_{\mu\nu}$, it is geodesic for the whole KS-class containing $g_{\mu\nu}$. Second, if k_μ is geodesic for a class, its optical scalars[7] depend only on the class. Third, if $g_{\mu\nu}$ has an algebraically special Weyl tensor with k_μ as a multiple Debever–Penrose direction and a Ricci tensor proportional to $k_\mu k_\nu$, then all metrics KS-related to it by k_μ with a Ricci tensor of this kind[8] must have special Weyl tensor also (the type will, however, vary over the class!).

It is convenient for our purposes to extend the null tetrad formalism to the situation considered here, maintaining all basic conventions of a paper by Debney, Kerr, and Schild.[9] Then the optical scalars of k_μ will appear as certain rotation coefficients of the tetrad fields. Let $\{e^1, e^2, e^3=k_\mu dx^\mu, e^4\}$ be any null cotetrad for g, so that[10]

$$g_{\mu\nu}dx^\mu dx^\nu = 2(e^1e^2+e^3e^4)=g_{ab}e^a\otimes e^b. \tag{6}$$

With this we associate the following null cotetrad (h^a) for f:

$$h^1:=e^1,\quad h^2:=e^2,\quad h^3:=e^3,\quad h^4:=e^4+\mathfrak{K}e^3, \tag{7}$$

so that (1) becomes

$$f_{\mu\nu}dx^\mu dx^\nu = 2(h^1h^2+h^3h^4)=f_{ab}h^a\otimes h^b. \tag{6'}$$

The dual tetrads are the two quadruples of differential operators $\{e_1, e_2, e_3, e_4=k^\mu\partial_\mu\}$ and

$$h_1:=e_1,\quad h_2:=e_2,\quad h_3=e_2,\quad h_3=e_3-\mathfrak{K}e_4,\quad h_4=e_4. \tag{7'}$$

The anholonomy coefficients of these tetrads, defined by the LIE brackets

$$[e_a, e_b] = C^c_{ab} e_c, \quad [h_a, h_b] = D^c_{ab} h_c \tag{8}$$

(or, equivalently, by $d \wedge e^c = -\frac{1}{2} C^c_{ab} e^a \wedge e^b$, $d \wedge h^c = -\frac{1}{2} D^c_{ab} h^a \wedge h^b$), can be related to each other using (7′); here and in what follows we make the convention that latin indices on quantities derived from g or f refer them to the tetrad $\{e_a\}$ or $\{h_a\}$, respectively, and are raised and lowered by $g_{ab} = g^{ab} = f_{ab} = f^{ab}$. The Riemannian connection defined by $g(f)$ is given by the matrix $\Gamma_{ab}(\Phi_{ab})$ of connection forms; their decomposition

$$\Gamma_{ab} = \Gamma_{abc} e^c, \quad \phi_{ab} = \Phi_{abc} h^c \tag{9}$$

defines the rotation coefficients $\Gamma_{abc}(\Phi_{abc})$, which are expressed in terms of the $C^c_{ab}(D^c_{ab})$ as

$$2\Gamma_{abc} = -C_{abc} + C_{cab}, \quad 2\Phi_{abc} = -D_{abc} = -D_{abc} + D_{bac} + D_{cab}. \tag{10}$$

Using the relations mentioned above that exist between the C's and D's, we obtain (among others) the relations

$$\begin{aligned}
&\Phi_{ab4} = \Gamma_{ab4}, \quad \Phi_{422} = \Gamma_{422} =: \sigma, \quad \Phi_{421} = \Gamma_{421} =: -Z, \\
&\Phi_{423} = \Gamma_{423}, \\
&\Phi_{121} = \Gamma_{121}, \quad \Phi_{123} = \Gamma_{123} + \mathcal{K}(\overline{Z} - Z - \Gamma_{124}), \\
&\Phi_{342} = \Gamma_{342} - \mathcal{K}\Gamma_{424}, \quad \Phi_{343} = \Gamma_{343} - \mathcal{K}\Gamma_{434} + \mathcal{K}_{,4}, \\
&\Phi_{232} = \Gamma_{232} + \mathcal{K}\Gamma_{422}.
\end{aligned} \tag{11}$$

This gives relations between the connection forms:

$$\Phi_{42} = \Gamma_{42} + \mathcal{K}\Gamma_{424} e^3, \tag{12}$$

$$\begin{aligned}
\Phi_{12} + \Phi_{34} = \Gamma_{12} + \Gamma_{34} - \mathcal{K}(\Gamma_{414} e^1 + \Gamma_{424} e^2) \\
+ [\mathcal{K}_{,4} + \mathcal{K}(\overline{Z} - Z - 2\Gamma_{434})] e^3.
\end{aligned}$$

Here and in the following $F_{,4} = e_4(F) = h_4(F) = k^\mu \partial_\mu F$. A little care is necessary concerning the "comma" notation of Ref. 9 for directional derivatives in our case, as the derivatives $e_3(F)$ and $h_3(F)$ of an arbitrary function F differ from each other; but whenever $F_{,4} = 0$ is valid, they coincide and may then unambiguously be written $F_{,3}$.

The geodesic condition on k with respect to g is now $\Gamma_{424} = 0$, and from (11) it then follows $\Phi_{424} = 0$, the geodesic condition with respect to f. Affine parametrization is then expressed as $\Gamma_{434} = 0$, which would imply $\Phi_{434} = 0$. σ and Z are the complex shear and complex expansion of k; they are the same for g and f. For the divergence of k we get the relation

$$k^\mu{}_{;\mu} = Z + \overline{Z} - \Gamma_{434} = Z + \overline{Z} - \Phi_{434}. \tag{13}$$

If k is geodesic, the definition of the tetrad component $R^1{}_{441} = \frac{1}{2} R_{44}$ of the Riemann (or Ricci, resp.) tensor gives

$$Z_{,4} + Z(Z + \Gamma_{434}) + |\sigma|^2 = \tfrac{1}{2} R_{44}(g) = \tfrac{1}{2} R_{44}(f), \tag{14}$$

the usual "propagation law" of Z. Finally, under the geodesic condition on k, one readily verifies from (12) and Ref. 9 that the curvature forms $\mathfrak{R}_{42}$ of both metrics are related by

$$\mathfrak{R}_{42}(f) = \mathfrak{R}_{42}(g) + [\mathcal{K}_{,4} + \mathcal{K}(\overline{Z} - Z - 2\Gamma_{434})]\Gamma_{42} \wedge e^3. \tag{15}$$

This ends our collection of auxiliary formulae; note that we have not used any field equations here; only for (14) and (15) it was assumed that k is geodesic.

III. THE OPTICAL SCALARS OF k

Equipped with the formalism of Sec. 2, we come back to our field equations (3). It is our purpose here to determine their consequences on the optical scalars of k which is seen to be necessarily geodesic. This follows when the contracted Bianchi identities are applied to Eqs. (3) with $\mathcal{K} \not\equiv 0$:

$$\mathcal{K} k_{\mu;\nu} k^\nu + (k^\nu \partial_\nu \mathcal{K} + \mathcal{K} k^\nu{}_{;\nu}) k_\mu = 0. \tag{16}$$

Scalar multiplication of this equation with e_1, e_2, e_3 gives its translation into our tetrad formalism:

$$\Gamma_{414} = 0, \quad \Gamma_{424} = 0, \quad \mathcal{K}_{,4} + \mathcal{K}(Z + \overline{Z} - 2\Gamma_{434}) = 0. \tag{16′}$$

Contracting Eqs. (3) with $g^{\mu\nu}$ and $f^{\mu\nu}$ resp., we find that the curvature scalars must vanish, so that (3) is equivalent to

$$R_{\mu\nu}(g) = -(\kappa_g^2/\kappa_f^2) M^2 \mathcal{K} k_\mu k_\nu, \quad R_{\mu\nu}(f) = M^2 \mathcal{K} k_\mu k_\nu, \tag{17}$$

which reads in tetrad form

$$\begin{aligned}
&R_{33}(g) = -(\kappa_g^2/\kappa_f^2) M^2 \mathcal{K}, \quad R_{33}(f) = M^2 \mathcal{K}, \\
&R_{ab}(g) = 0 = R_{ab}(f) \quad \text{otherwise.}
\end{aligned} \tag{17′}$$

Now the field equations $R_{22}(f) = R_{22}(g) = 0$ together with $R_{22} = 2R_{4223}$ taken from (15) imply

$$\sigma[\mathcal{K}_{,4} + \mathcal{K}(\overline{Z} - Z - 2\Gamma_{434})] = 0. \tag{18}$$

Assume tentatively that $\sigma \neq 0$, then from (18) and (16′) $Z = 0$, and (14) gives $R_{44} = 2|\sigma|^2 \neq 0$, so that the field equations $R_{44} = 0$ cannot be satisfied. Hence we have shown that k must have *vanishing shear*.

It follows from the Goldberg–Sachs theorem[11] that both g and f must have algebraically special Weyl tensors, k being a multiple Debever–Penrose vector. This enables us in the case $Z \neq 0$ to follow the Kerr[12,9] procedure of constructing the algebraically special Einstein vacuum solutions as long as the field equation $R_{33} = 0$ is *not* used. In the case $Z = 0$ we can follow Ref. 4, where all metrics with $R_{\mu\nu} \propto k_\mu k_\nu$ that contain a $\sigma = Z = 0$ null congruence are given (Sec. 4).

We will show now that the assumption $Z \neq 0$ does not lead to nontrivial solutions either, using Kerr's procedure. Closely following Ref. 9 again, we adjust the tetrad $\{e_a\}$ such that $\Gamma_{423} = \Phi_{423} = 0$ and that $\Gamma_{42} = \Phi_{42}$ becomes a perfect differential $-dY$, thus introducing two complex conjugate coordinates Y, $\overline{Y}$. Also the construction of the third coordinate ρ goes through without change, according to our remark on comma derivatives, whereas, when the fourth coordinate r is introduced by the condition $r_{,4} = 1$, we must write

$$\begin{aligned}
e^4 &= dr + \beta dY + \overline{\beta} d\overline{Y} + \mathscr{G} e^3, \\
h^4 &= dr + \beta dY + \overline{\beta} d\overline{Y} + \mathfrak{F} e^3.
\end{aligned} \tag{19}$$

By (7), the real functions $\mathscr{G}$, $\mathfrak{F}$ satisfy

$$\mathfrak{F} - \mathscr{G} = \mathcal{K}, \tag{20}$$

and are determined, according to Ref. 9, as

$$\begin{aligned}
\mathfrak{F} &= |\dot{\Omega}|^2 + \mathrm{Re} LZ, \\
\mathscr{G} &= |\dot{\Omega}|^2 + \mathrm{Re} NZ,
\end{aligned} \tag{21}$$

where L, N are two complex functions (replacing the M of Ref. 9) which satisfy

$$\begin{aligned}
&L_{,4} = 0 = N_{,4}, \\
&\mathrm{Im} L = \mathrm{Im} \overline{D}\overline{D} D\Omega = \mathrm{Im} N,
\end{aligned} \tag{22}$$

$$\bar{D}L = 3L\dot{\bar{\Omega}},\quad \bar{D}N = 3N\dot{\bar{\Omega}}.$$

At this point we must deviate from Ref. 9, as we want to satisfy (17′) instead of $R_{33}=0$. But now we see that $\mathrm{Im}(L-N)=0$, so $L-N$ is real and satisfies

$$(L-N)_{,4}=0.,\quad (L-N)\mathrm{Re}Z=\mathcal{K}. \tag{23}$$

When this, together with $Z_{,4}$ from (14) (where now $\sigma=R_{44}=0$) and $\Gamma_{434}=0$ (a consequence of our tetrad adjustment, see Ref. 9) is inserted into (16′), the equation

$$(L-N)|Z|^2=0 \tag{24}$$

results, proving that for $Z\neq 0$ the solution can only be trivial.

The net result of this section therefore is that for a nontrivial solution we must restrict k to have *vanishing shear, expansion and twist*; g must have an *algebraically special Weyl tensor* with k as a multiple Debever–Penrose vector.

IV. CANONICAL FORM OF SOLUTIONS

In Ref. 4, a study was made of the metrics that admit a null congruence k_μ with vanishing shear, twist, and expansion and whose Ricci tensor is of the form

$$R_{\mu\nu}=-\mu k_\mu k_\nu, \tag{25}$$

where μ is an unspecified function. It was shown that such a metric may be brought into the form

$$ds^2=p^2|dz+Bdu|^2+2dudv+Hdu^2, \tag{26}$$

where $z=x+iy$ and B are complex, $e^3=k_\mu dx^\mu=du$ is affinely parametrized by $v, p>0$, $\partial_v p=0$. If we use the cotetrad

$$e^1=\tfrac{1}{2}p(dz+Bdu),\quad e^2=\tfrac{1}{2}p(d\bar{z}+\bar{B}du),\quad e^3=du,$$
$$e^4=dv+\tfrac{1}{2}Hdu,$$

then the field equations[13] $R_{44}=0$ are identically satisfied, $R_{41}=R_{42}=0$ imply $\partial_v^2 B=0$, $R_{11}=R_{22}=R_{12}=0$ then determine B as far as possible. H must be of the form

$$H=v^2H_2(u,z,\bar{z})+vH_1(u,z,\bar{z})+A(u,z,\bar{z}), \tag{27}$$

where H_2 and H_1 are completely determined by the equations $R_{31}=R_{32}=R_{34}=0$ once p, B have been chosen [an undetermined additive contribution $F(u)$ to H_1 is canceled by a coordinate transformation]. The remaining equation $R_{33}=-\mu$ becomes a *linear* second-order differential equation of elliptic type for the function, A, $\mathcal{D}A=-\mu$, say.

From this, the solutions of our problem are obtained as follows. First, write g in the form (26), i.e., choose p, B according to the restrictions above, and calculate H_2, H_1. Then put

$$g_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2\mathcal{G}du^2,\quad \partial_v\mathcal{G}=0 \tag{28}$$

where ds_0^2 is (26) with $H\equiv v^2H_2+vH_1$. From (1) we then get

$$f_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2\mathcal{F}du^2,\quad \partial_v\mathcal{F}=0 \tag{28′}$$

where $\mathcal{F}=\mathcal{G}+\mathcal{K}$.[14] This satisfies all equations, if $\mathcal{F}$, $\mathcal{G}$ satisfy the linear system

$$\mathcal{D}\mathcal{G}=-(\kappa_g^2/\kappa_f^2)M^2(\mathcal{F}-\mathcal{G}),$$
$$\mathcal{D}\mathcal{F}=M^2(\mathcal{F}-\mathcal{G}), \tag{29}$$

which is easily decoupled by going to the linear combinations $\mathcal{F}-\mathcal{G}$, $\mathcal{G}+(\kappa_g^2/\kappa_f^2)\mathcal{F}$.

We do not want to give more details in this general case, but point out how certain special solutions that have been given elsewhere[15,16] fit into this scheme. In the solutions indicated above both g and f will in general be of Petrov type II, but one can single out the types III and N. It is convenient to do this by making one further invariant distinction: in the $\sigma=Z=0$ case another optical scalar exists, the *rotation*[4,7] (in the formalism of Sec. 2, it is equal to $|\Gamma_{423}|$). Thus, one can classify[4] the solutions according to vanishing or nonvanishing rotation: (a) In the case of vanishing rotation ("recurrent rays") all solutions *are* of type III or more special. They can be written in Kundt's canonical form as

$$g_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2\mathcal{G}du^2,\quad f_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2\mathcal{F}du^2, \tag{30a}$$

where

$$ds_0^2=|dz+bdu|^2+2dudv-v\partial_x bdu^2, \tag{31a}$$
$$\partial_v b=\Delta_2 b=0(\Delta_2\equiv\partial_x^2+\partial_y^2),\ b \text{ real},$$

and where $\mathcal{F},\mathcal{G}$ satisfy

$$\partial_v\mathcal{F}=\partial_v\mathcal{G}=0,$$
$$\Delta_2(\mathcal{F}-\mathcal{G})=[1+(\kappa_g^2/\kappa_f^2)]M^2(\mathcal{F}-\mathcal{G}), \tag{32a}$$
$$\Delta_2\{[\mathcal{G}+(\kappa_g^2/\kappa_f^2)\mathcal{F}]/[1+(\kappa_g^2/\kappa_f^2)]\}$$
$$=\partial_u\partial_x b-\tfrac{1}{2}(\partial_y b)^2-b\partial_x^2 b-\tfrac{3}{2}(\partial_x b)^2.$$

The type N subcase can be brought to this form with $b\equiv 0$; these are the *pp*- wave solutions described in Ref. 15. When b is not linear in x, y, we have type III and no plane-frontedness in Kundt's sense. (b) In the case of nonvanishing rotation the canonical form is rather simple when we *require* type III or a more special one. The solutions can be written as (note the replacement $\mathcal{G}\to x\mathcal{G}$, $\mathcal{F}\to x\mathcal{F}$ made for convenience)

$$g_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2x\mathcal{G}du^2,\quad f_{\mu\nu}dx^\mu dx^\nu=ds_0^2+2x\mathcal{F}du^2, \tag{30b}$$

where

$$ds_0^2=|dz-(2vx^{-1}+c)du|^2+2dudv+[\partial_x-2x^{-1})c-3v^2x^{-2}]du^2, \tag{31b}$$
$$\partial_v c=\Delta_2(xc)=0,\quad c \text{ real},$$

and where $\mathcal{F},\mathcal{G}$ satisfy

$$\partial_v\mathcal{F}=\partial_v\mathcal{G}=0,$$
$$\Delta_2(\mathcal{F}-\mathcal{G})=[1+(\kappa_g^2/\kappa_f^2)]M^2(\mathcal{F}-\mathcal{G}), \tag{32b}$$
$$\Delta_2\{[\mathcal{G}+(\kappa_g^2/\kappa_f^2)\mathcal{F}]/[1+(\kappa_g^2/\kappa_f^2)]\}$$
$$=c(x^{-1}-\partial_x)\partial_x c-\partial_u\partial_x c-\tfrac{1}{2}(\partial_y c)^2-\tfrac{3}{2}(\partial_x c)^2.$$

The type N subcase can be brought tho this form with $c\equiv 0$; these are the plane-fronted wave solutions with rotating rays described in Ref. 16; ds_0^2 is then flat but written in some noninertial curvilinear coordinates, so that g, f are KS-related to the Minkowski metric (as they are in the N subcase of (a), but in a very different way). When xc is not a linear function of x, y, we have type III and no plane-frontedness.

Finally, we see, as a by-product, that, for all possible solutions, f and g have the same Petrov type.

Note added in proof: It has been pointed out to the author that a relation of the kind defined by Eqs. (1), (2) appears to have been first considered by Trautman.[17] It is algebraically characterized uniquely by the condition that an ansatz $f_{\mu\nu} = g_{\mu\nu} + 2\mathcal{K} h_{\mu\nu}$ imply $f^{\mu\nu}$ = linear function of $\mathcal{K}$. Whereas Trautman considered the more restrictive case where one demands that the (Einstein) field equations permit $\mathcal{K}$ to be an arbitrary function of some scalar function $\sigma(x)$, Kerr and Schild[3,9] and Robinson and Robinson[8] considered the general case (with $Z \neq 0$, by which assumption the two cases in fact become disjoint). The relationship between f and g should therefore be called TKS relation. It may also be characterized geometrically by the two null cones *hypersculate* each other along a common generator k. This remark clarifies some questions concerning causality in our solutions.

ACKNOWLEDGMENTS

The author wishes to thank Dr. P. C. Aichelburg and R. Mansouri for helpful discussions.

*Supported by "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich", Nr. 1534.

[1]C. J. Isham, A. Salam, and J. Strathdee, Phys. Rev. D **3**, 867 (1971).
[2]F. A. E. Pirani, Contribution at the Topical Meeting on Gravitation and Field Theory, Trieste, 1971.
[3]R. P. Kerr and A. Schild, *Proceedings of the Meeting of General Relativity* (Anniversary Volume, Fourth Centenary of Galileo's Birth), edited by G. Barèra (Firenze, 1965).
[4]W. Kundt, Z. Phys. **163**, 77 (1961).
[5]P. C. Aichelburg and R. Mansouri, Proceedings of the Topical Meeting on Gravitation and Field Theory, Trieste, 1971; Nuovo Cimento **10B**, 483 (1972).
[6]P. G. O. Freund, A. Maheshwari, and E. Schonberg, Astrophys. J. **157**, 857 (1969).
[7]P. Jordan, J. Ehlers, and R. K. Sachs, Abh. Math.-Naturwiss. Kl. Akad. Wiss. Lit. Mainz **1**, (1961).
[8]KS-related Einstein vacuum solutions of this sort have, as a by-product, been considered by I. Robinson and J. R. Robinson, Int. J. Theor. Phys. **2**, 231 (1969).
[9]G. C. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969). Note that these authors call rotation what in Refs. 4 and 7 is called twist. We stick to the latter terminology which reserves "rotation" to another property.
[10]This is the usual way of writing $g_{\mu\nu} dx^\mu \otimes dx^\nu = e^1 \otimes e^2 + e^2 \otimes e^1 + e^3 \otimes e^4 + e^4 \otimes e^3$.
[11]See J. N. Goldberg and R. K. Sachs, Acta Phys. Pol. **22**, 13 (1962). Lemma 1; observe that the tetrad indices 3,4 in this paper are used in the opposite was as done here.
[12]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[13]Explicit formulae for all relevant quantities have been published by K. Bichteler, Z. Phys. **178**, 488 (1964), with only slightly different conventions.
[14]$\partial_\nu \mathcal{F} = 0$ follows from $\partial_\nu \mathcal{G} = 0$ and $\partial_\nu \mathcal{K} = 0$; the latter is just Eq. (16'), since affine parametrization means $\Gamma_{434} = 0$, and e_4 equals ∂_ν here.
[15]P. C. Aichelburg, R. Mansouri, and H. K. Urbantke, Phys. Rev. Lett. **27**, 1533 (1971); in (5), (6) of this reference, $\mathcal{F}$ and $\mathcal{G}$ appear erroneously interchanged.
[16]R. Mansouri and H. K. Urbantke, "On Exact Kerr-Schild Type Vacuum Solutions to Salam's Two-tensor Theory of Gravity" Commun. Math. Phys. **26**, 301 (1972).
[17]A. Trautman, in *Recent Developments in General Relativity* (Polish Scientific Publishers, Pergamon, 1962).

One-dimensional model of the rearrangement process and the Faddeev equations

K. Lipszyc

Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland
(Received 16 May 1972)

In the paper we study the Faddeev–Lovelace equations for the one-dimensional system of three particles interacting via zero-range potentials. The half-off-shell rearrangement and elastic amplitudes are constructed explicitly for the particular case of the model, namely for the system consisting of three particles with equal masses and with two-body potentials of equal strength. These exact amplitudes are shown to satisfy the modified form of the Faddeev-Lovelace equations.

I. INTRODUCTION

In Ref. 1 we have studied the one-dimensional model of three particles interacting via zero-range potentials. The two-body potentials of the system under consideration were described by the boundary conditions requiring that the logarithmic derivative of the wave function with respect to the two-particle relative distance variable takes a constant value $-\alpha$ at the point of zero interparticle distance. The exact solution of the three-body Schrödinger equation was written using the similarity of our quantum mechanical problem and a certain scattering problem in acoustics.[2] For the general case of arbitrary masses and arbitrary two-body constants α the solution in coordinate space was given in the form of Sommerfeld's contour integral, the integrand being the combination of the trigonometric functions and of certain special functions defined and applied by Maluzhinetz when studying the problem of diffraction of an acoustic wave by a wedge with given face impedances. The possible scattering processes in the three-body system under consideration are elastic scattering, rearrangement, and break-up. The amplitudes of those processes were written explicitly in Ref. 1, in terms of these special functions.

A somewhat similar three-body problem was studied in detail by Nussenzweig.[3] In his model one of the particle masses is assumed to be infinitely large, the remaining two being equal to each other. One of the two-body potentials was assumed to be of the type described above, and the other one to be of the hard-core type. As mentioned in Ref. 1, the model discussed by Nussenzweig may be considered as a special case of our model.

In this paper we calculate the two-body t matrix for the two-body subsystems of our model and use it to construct the Faddeev equations for the elastic and rearrangement transition amplitudes. As the simplest case of our model we discuss the system of three particles with equal masses interacting via two-body potentials of equal strength. For this symmetrical model we have calculated the half-off-the-energy-shell amplitudes from the three-body solution in coordinate space. We have undertaken to verify the Faddeev—Lovelace equations using these exact transition amplitudes. This procedure lead us to expressions with divergent integrals, this fact being the consequence of the singularity of the potentials. In order to remove these divergences we subtracted some "on-shell" terms from both sides of the Faddeev equations and checked that our exact amplitudes satisfy the resulting equations.

Dodd[4] discussed the Faddeev—Lovelace equations for the one-dimensional system of three particles interacting via delta-function potentials. The exact amplitudes were written in such form that they contained the factor which exactly cancelled the propagator in the kernel of the integral equations. The integrals resulting from the substitution of these exact amplitudes into the integral equations were convergent and the equations could be verified directly. In our case the amplitudes contain a factor which cancels the propagator only partly. The resulting integrals are divergent. These difficulties are apparently connected with the fact that the potentials we use contain a hard-core part, i.e., the particles are impenetrable.

II. THE TWO-BODY T MATRIX

As in Ref. 1 we discuss the Schrödinger equation which is of the following form:

$$-\frac{d^2}{dx^2}\Psi(x)=\hat{E}\Psi(x), \tag{1}$$

with a boundary condition

$$\frac{d}{dx}\Psi(x)+\alpha\Psi(x)=0, \quad \text{for } x=0, \tag{2}$$

where α is a real number.

The boundary condition (2) describes the action of the zero-range potential, which is impenetrable, i.e., the wave coming from one end of the x axis cannot be transmitted through the scattering center at $x=0$. We shall also assume that

$$\alpha<0. \tag{3}$$

Under this assumption Eq. (1) with the condition (2) has one bound state solution different from zero on the semiaxis $x<0$:

$$\Psi_b(x)=H(-x)(-2\alpha)^{1/2}\exp(-\alpha x), \tag{4}$$

where $H(x)$ is the Heaviside step function.

The solutions corresponding to $\hat{E}>0$, different from zero for negative values of x, read as follows:

$$\Psi(x)=H(-x)\left(\frac{1}{2\pi}\right)^{1/2}\left[\exp(i\sqrt{\hat{E}}x)-\left(\frac{\alpha+i\sqrt{\hat{E}}}{\alpha-i\sqrt{\hat{E}}}\right)\exp(-i\sqrt{\hat{E}}x)\right]. \tag{5}$$

Equation (1) with the boundary condition (2) can be written in the form of a single equation:

$$\left(-\frac{d^2}{dx^2}+v(x)\right)\Psi(x)=\hat{E}\Psi(x), \tag{6}$$

where $v(x)\Psi(x)$ is the distribution of the following type:

$$v(x)\Psi(x)=\Psi'(0)\cdot\delta(x)+\Psi(0)\cdot\delta'(x)$$
$$=\Psi(0)(-\alpha\delta(x)+\delta'(x)), \quad (7)$$
$$\Psi(0)\stackrel{\text{df}}{=}\Psi(x+0)-\Psi(x-0). \quad (7')$$

The t matrix for the potential under consideration can be calculated applying the method of the Møller wave operator.[5] Proceeding as in Ref. 5 we write the expression for the t matrix elements off the energy shell in the following form:

$$\langle k|t(z)|k'\rangle=(1/2\pi)^{1/2}\int_{-\infty}^{\infty}dx\exp(-ikx)v(x)\Psi_{k',z}(x), \quad (8)$$
$$-\infty<k,k'<\infty,$$

where k is the momentum coordinate conjugated with the coordinate x, and $\Psi_{k',z}(x)$ is the solution of the following equation:

$$\left(-\frac{d^2}{dx^2}-z\right)\Psi_{k',z}(x)=\left(\frac{1}{2\pi}\right)^{1/2}(k'^2-z)\exp(ik'x), \quad (9)$$

with the boundary condition

$$\frac{d}{dx}\Psi_{k',z}(x)+\alpha\Psi_{k',z}(x)=0, \quad \text{for } x=0. \quad (9')$$

The meaning of the expression $v(x)\Psi_{k',z}(x)$ in (8) is the same as that given by (7), i.e.,

$$v(x)\Psi_{k',z}(x)=\Psi_{k',z}(0)(-\alpha\delta(x)+\delta'(x)). \quad (10)$$

At $|x|\to\infty$, the difference $[\Psi_{k',z}(x)-(1/2\pi)^{1/2}\exp(ik'x)]$ should tend to zero for all complex values of z not lying on the positive real semiaxis.

The solution of Eq. (9) satisfying these conditions is

$$\Psi_{k',z}(x)=\begin{cases}(1/2\pi)^{1/2}\left(\exp(ik'x)-\dfrac{\alpha+ik'}{\alpha-i\sqrt{z}}\exp(i\sqrt{z}x)\right), & \text{for } x>0,\\ (1/2\pi)^{1/2}\left(\exp(ik'x+\dfrac{\alpha+ik'}{-\alpha+i\sqrt{z}}\exp(-i\sqrt{z}x)\right), & \text{for } x<0,\end{cases}$$
$$\operatorname{Im}\sqrt{z}\geq 0. \quad (11)$$

From (8), (10), and (11) we obtain

$$\langle k|t(z)|k'\rangle=\left(\frac{1}{2\pi}\right)^{1/2}\int_{-\infty}^{\infty}dx\exp(-ikx)v(x)\Psi_{k',z}(x),$$
$$=\frac{1}{2\pi}(k+i\alpha)\frac{2i\sqrt{z}}{z+\alpha^2}(-k'+i\alpha). \quad (12)$$

In the limit $|\alpha|\to\infty$ the right-hand side of (12) tends to $-2i\sqrt{z}(2\pi)^{-1}$ which is the expression for the hard-core t matrix.[5]

The t matrix element given by Eq. (12) is separable in the coordinates k, k' and can be written in the form

$$\langle k|t(z)|k'\rangle=g^*(k)\tau(z)g(k')=\langle k|i\rangle\tau(z)\langle i|k'\rangle, \quad (13)$$

where $g(k)$ is the form factor of the bound state

$$g^*(k)=\langle k|i\rangle=-(k^2+\alpha^2)\langle k|\Psi_b\rangle=-i(-\alpha/\pi)^{1/2}(k+i\alpha) \quad (14)$$

and the propagator $\tau(z)$ reads as follows:

$$\tau(z)=\frac{i\sqrt{z}}{\alpha(z+\alpha^2)}. \quad (15)$$

III. THE FADDEEV-LOVELACE EQUATIONS FOR THE ELASTIC AND REARRANGEMENT TRANSITION AMPLITUDES

In the following we shall discuss the problem of scattering of three particles moving in one dimension. The potentials V_1 and V_3 between the particles 2 and 3 and between the particles 1 and 2, respectively, is assumed to be of the type described in Sec. I and the potential V_2 of the pair 1, 3 is put equal to zero.

The position and momentum coordinates used in the three-body system are defined as follows[1]:

$$s_1=(r_2-r_3)[2m_2m_3/(m_2+m_3)]^{1/2},$$
$$t_1=\{[(m_2r_2+m_3r_3)/(m_2+m_3)]-r_1\}\cdot[2m_1(m_2+m_3)/(m_1+m_2+m_3)]^{1/2}, \quad (16)$$
$$k_1=(m_3p_2-m_2p_3)/[2m_2m_3(m_2+m_3)]^{1/2},$$
$$q_1=[m_1(p_2+p_3)-(m_2+m_3)p_1]/[2m_1(m_2+m_3)(m_1+m_2+m_3)]^{1/2}, \quad (17)$$

where r_i and p_i are the position and momentum coordinates, respectively, in the three-body center of mass system. We shall also use the coordinates (s_3, t_3) and (k_3, q_3), the definitions of which can be obtained from the above ones by cyclic permutations of the indices 1, 2, 3. The system (s_1, t_1) [or (k_1, q_1)] is related to that of (s_3, t_3) [or (k_3, q_3)] by the transformation

$$\begin{bmatrix}s_1\\ t_1\end{bmatrix}=\begin{bmatrix}-b & a\\ -a & -b\end{bmatrix}\begin{bmatrix}s_3\\ t_3\end{bmatrix}, \quad (18)$$

where

$$a=[m_2(m_1+m_2+m_3)/(m_1+m_2)(m_2+m_3)]^{1/2}\equiv\sin2\Phi$$
$$b=[m_1m_3/(m_1+m_2)(m_2+m_3)]^{1/2}\equiv\cos2\Phi. \quad (19)$$

The Schrödinger equation for the problem under consideration takes the following form:

$$\left(-\frac{\partial^2}{\partial s^2}-\frac{\partial^2}{\partial t^2}\right)\Psi(s,t)=E\Psi(s,t), \quad (20)$$
$$\frac{\partial\psi}{\partial s_3}+\alpha_3\Psi(s_3,t_3)=0, \quad \text{for } s_3=0, \quad (20')$$
$$\frac{\partial\psi}{\partial s_1}+\alpha_1\Psi(s_1,t_1)=0, \quad \text{for } s_1=0, \quad (20'')$$

where (s, t) is one of the systems (s_1, t_1) or (s_3, t_3) and the negative numbers α_1 and α_3 describe the strength of interaction between the particles 2 and 3 and between the particles 1 and 2, respectively.

In the following we shall concentrate on the problem of scattering of particle 3 on the bound state of particles 1 and 2. The possible processes are the elastic scattering, the dissociation of the pair 1, 2, and the rearrangement process in which the bound state of particles 2, 3 is formed and particle 1 leaves freely.

The exact solution of Eq. (20) was given in Ref. 1 where the expressions for the scattering probabilities were also written in analytic form.

Here we use the Faddeev–Lovelace formalism[6,7] to write the integral equations for the elastic and rearrangement transition amplitudes off the energy shell. These equations take a relatively simple form due to

the fact that the two-body t matrices as given by expression (12) are of the one-component separable form.

We define the elastic scattering amplitude $h_{33}(q_3, p)$ and the rearrangement amplitude $h_{13}(q_1, p)$ as follows:

$$h_{i3}(q_i, p) = \langle q_i, i | G_0(E+i0)U_{i3}G_0(E+i0)|3, p\rangle + (1-\delta_{i3})\cdot[\langle q_i, i|G_0(E+i0)|3,p\rangle - \langle q_i, i|G_0(E+i0)V_iG_0(E+i0)|3,p\rangle], \quad (21)$$
$$i = 1, 3,$$

where for simplicity we have denoted the relative momentum of particle 3 in the initial state by p, and

$$G_0(E+i0) = [E+i0-H_0]^{-1}, \quad (22)$$

$$U_{ij}(E+i0) = (V_i+V_k)+(V_j+V_k)G(E+i0)(V_i+V_k), \quad i,j,k=1,2,3, \quad (23)$$

$$G(E+i0) = (E+i0-H)^{-1}, \quad (24)$$

$|i\rangle$ is the form factor of the bound state of the subsystem i [see Eq. (14)].

The term in the square brackets in Eq. (21) vanishes when h_{i3} are calculated on the energy shell. The reason for including it in the definitions (21) is that it leads to a more symmetrical form of the integral equations.[6]

The Faddeev–Lovelace equations for the amplitudes h_{13} and h_{33} are as follows:

$$h_{13}(q,p;E) = b_{13}(q,p;E) + \int_{-\infty}^{\infty} dq' b_{13}(q,q';E)\tau_3(E-q'^2) \times h_{33}(q',p;E), \quad (25)$$

$$h_{33}(q,p;E) = \int_{-\infty}^{\infty} dq' b_{31}(q,q';E)\tau_1(E-q'^2)h_{13}(q',p;E), \quad (26)$$

where $\tau_1(E)$ and $\tau_3(E)$ are the propagators of the two-body t matrices $t_1(E)$ and $t_3(E)$, given by (15). In Eqs. (25) and (26) we have dropped the infinitesimal imaginary part $i0$ of the energy variable E. The Born terms b_{13} and b_{31} are defined by the expressions

$$b_{13}(q,q';E) = \langle 1, q|G_0(E)|3,q'\rangle, \quad (27)$$

$$b_{31}(q,q';E) = \langle 3, q|G_0(E)|1,q'\rangle. \quad (28)$$

From (14), (27), (28) and using the relations (18) [written for the (k,q) variables] we obtain

$$b_{13}(q,q';E) = -(\pi a)^{-1}(\alpha_1\alpha_3)^{1/2}[E-(a^2)^{-1}(q^2+2bqq'+q'^2)]^{-1} \times[-(b/a)q - a^{-1}q' + i\alpha_1][-a^{-1}q-(b/a)q'+i\alpha_3], \quad (29)$$

$$b_{31}(q,q';E) = -(\pi a)^{-1}(\alpha_1\alpha_3)^{1/2}[E-(a^2)^{-1}(q^2+2bqq+q^2)]^{-1} \times[(b/a)q + a^{-1}q' + i\alpha_3][a^{-1}q+(b/a)q'+i\alpha_1)]. \quad (30)$$

If we put

$$E = p^2 - \alpha_3^2 \quad (31)$$

in the definitions (21) and in the integral equations (26) we obtain the equations for the half-on-the-energy-shell amplitudes. The energy of the initial state, i.e., the sum of the kinetic energy of particle 3 and of the bound state energy of the pair 1, 2 is now related to the energy variable E by the energy conservation rule.

In this case the amplitudes h_{13} and h_{33} can be expressed in terms of the exact three-body solution[4] $|\Psi_p\rangle$ corresponding to the incident state $|\varphi_3, p\rangle$, where φ_3 is the bound state function of the pair 1, 2.

Applying the definitions (21), (23) and noting that

$$G_0(p^2-\alpha_3^2)|3,p\rangle = |\varphi_3, p\rangle, \quad (32)$$

we obtain

$$h_{33}(q,p) = \langle q,3|G_0(E)[V_1+V_1G(E)V_1]|\varphi_3,p\rangle = \langle q,3|G_0(E)V_1|\Psi_p\rangle, \quad (33)$$

$$h_{13}(q,p) = \langle q,1|G_0(E)[V_1+V_3G(E)V_1]|\varphi_3,p\rangle + \langle q,1|G_0(E)V_3|\varphi_3,p\rangle - \langle q,1|G_0(E)V_1|\varphi_3,p\rangle = \langle q,1|G_0(E)V_3[1+G(E)V_1]|\varphi_3,p\rangle = \langle q,1|G_0(E)V_3|\Psi_p\rangle, \quad (34)$$

where we have used the fact that Ψ_p is the solution of the Lippmann–Schwinger equation

$$|\Psi_p\rangle = |\varphi_3,p\rangle + G(E+i0)V_1|\varphi_3,p\rangle. \quad (35)$$

The dissociation amplitude (on-the-energy-shell) can be expressed in terms of the amplitudes (33) and (34). Using the Lovelace formalism we obtain the following result:

$$T_{03}(k_1,q_1;p) \equiv \langle k_1,q_1|V_1+V_3|\Psi_p\rangle = \langle k_1,q_1|T_1(E)G_0(E)V_3|\Psi_p\rangle + \langle k_1,q_1|T_3(E)G_0(E)V_1|\Psi_p\rangle, \quad (36)$$

where we put $k_1^2+q_1^2 = E = p^2-\alpha_3^2$.

$T_1(E)$ and $T_3(E)$ are the two-body t-operators acting in the three-particle space.

Using the separable forms of T_1 and T_3, and the expressions (33) and (34), we obtain

$$T_{03}(k_1,q_1;p) = g_1^*(k_1)\tau_1(E-q_1^2)h_{13}(q_1,p) + g_3^*(k_3)\tau_3(E-q_3^2)h_{33}(q_3,p), \quad (37)$$

where (k_3, q_3) are related to (k_1, q_1) by the transformation (18), and $g_i^*(k_i)$ $(i=1,3)$ are given by (14).

IV. THE MODEL OF EQUAL MASSES AND EQUAL POTENTIALS

In Ref. 1 we noted that the solution of the scattering problem under consideration takes a very simple form if we put:

$$m_1 = m_2 = m_3, \quad \alpha_1 = \alpha_3 \equiv \alpha, \quad \text{i.e.,} \quad V_1 = V_3. \quad (38)$$

In this section we calculate the explicit forms of the half-on-the-energy-shell amplitudes for that symmetrical problem and discuss the Faddeev equations in that case.

Using (30) we find from (18) that

$$2\Phi = \pi/3, \quad a = \sin 2\Phi = \tfrac{1}{2}\sqrt{3}; \quad b = \cos 2\Phi = \tfrac{1}{2}. \quad (39)$$

The position coordinate system is shown in Fig. 1. We consider the solution different from zero in region I where the order of particles is 1, 2, 3 and both co-

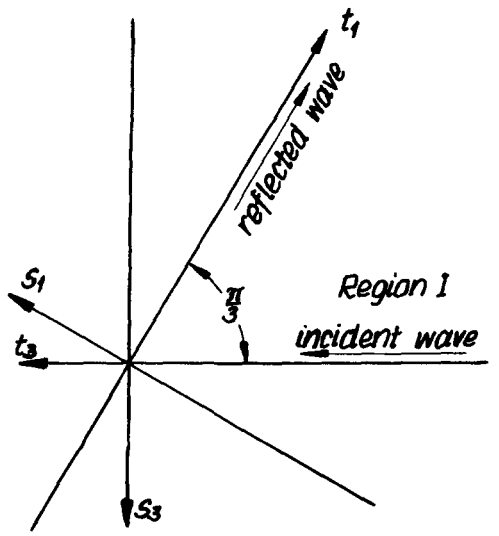

FIG. 1. The three-body coordinate system.

ordinates s_1 and s_3 are negative. In that region the wave coming from infinity has a positive value of the momentum coordinate q_3, and the wave leaving the reaction center has a positive value of q_1. As mentioned above, we assume that in the initial state particles 1 and 2 are bound which corresponds to the situation of the "surface" wave coming from infinity along the boundary $s_3=0$ in the region I.

The exact solution can be derived in this case from the general solution of Eq. (20) as shown in Ref. 1 or it can be readily calculated directly by the method of images. One has to add to the incident wave the two images resulting from the reflection of the incident wave from the "mirrors" $s_3=0$ and $s_1=0$ in such a way that the corresponding boundary conditions will be fulfilled. The resulting solution takes the following form:

$$\Psi_p(s,t)=\frac{-\alpha}{\pi}^{1/2}\Big[\exp(-\alpha s_3+ipt_3)-\Big(\frac{3\alpha+i\sqrt{3}p}{\alpha-i\sqrt{3}p}\Big)\times\exp s_3\Big(-\tfrac{1}{2}\alpha+\frac{\sqrt{3}}{2}ip\Big)\ \exp t_3\Big(-\frac{\sqrt{3}}{2}\alpha-\tfrac{1}{2}ip\Big)+\Big(\frac{3\alpha+i\sqrt{3}p}{\alpha-i\sqrt{3}p}\Big)\Big(\frac{\alpha+i\sqrt{3}p}{3\alpha-i\sqrt{3}p}\Big)\exp(-\alpha s_1+it_1p)\Big] \tag{40}$$

in region I, and $\Psi_p(s,t)=0$ in the remaining part of the (s,t) plane. The last term in (40) was written using the coordinates (s_1,t_1) in order to show that it represents the wave leaving the reaction center along the boundary $s_1=0$. It represents the final state in which particles 2 and 3 are bound. We see also from (40) that the amplitude of dissociation is equal to zero.

In order to calculate the half-off-shell transition amplitudes we have to study the behavior of the functions $V_1\Psi_p$ and $V_3\Psi_p$ which are involved in expressions (33) and (34). Inserting into the Schrödinger equation the formal expressions $V_1(s_1)$ and $V_3(s_3)$ for the potentials described by the boundary conditions (20′) and (20″) and treating the discontinuous functions $\Psi(s,t)$, $\partial\Psi/\partial s_3$, and $\partial\Psi/\partial s_1$ as distributions,[8] we obtain the prescription how to treat the expression $(V_1+V_3)\Psi$ under the integral sign:

$$\begin{aligned}\int_{-\infty}^{\infty}ds\int_{-\infty}^{\infty}dt\,g(s,t)(V_1+V_3)\Psi(s,t)&=\Big(-\int_0^{\infty}dt_1\frac{\partial\Psi}{\partial s_1}(s_1=0,t_1)\cdot g(s_1=0,t_1)\\&+\int_0^{\infty}dt_1\frac{\partial g}{\partial s_1}(s_1=0,t_1)\cdot\Psi(s_1=0,t_1)\Big)\\&+\Big(-\int_{-\infty}^{0}dt_3\frac{\partial\Psi}{\partial s_3}(s_3=0,t_3)\cdot g(s_3=0,t_3)\\&+\int_{-\infty}^{0}dt_3\frac{\partial g}{\partial s_3}(s_3=0,t_3)\cdot\Psi(s_3=0,t_3)\Big),\end{aligned} \tag{41}$$

where $g(s,t)$ is a test function continuously differentiable and the symbol $\lrcorner\!\!\lceil$ is defined as follows:

$$\lrcorner\!\!\lceil f(s=0,t)\stackrel{\mathrm{df}}{=}f(s+0,t)-f(s-0,t). \tag{41′}$$

From the zero-rangeness of the potentials we can deduce that the first term in the square brackets in (41) describes the action of $V_1\Psi$ and the second term in the square brackets describes the action of $V_3\Psi$.

From (40) we obtain

$$\begin{aligned}-\lrcorner\!\!\lceil\Psi_p(s_1=0,t_1)&=\Big(-\frac{\alpha}{\pi}\Big)^{1/2}\Big[-2\Big(\frac{\alpha+i\sqrt{3}p}{\alpha-i\sqrt{3}p}\Big)\\&\times\exp\Big[t_1\Big(\frac{\sqrt{3}}{2}\alpha-\frac{1}{2}ip\Big)\Big]+\Big(\frac{\sqrt{3}\alpha+ip}{\alpha-i\sqrt{3}p}\Big)\\&\times\Big(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\Big)\exp(it_1p)\Big],\\&=\frac{1}{\alpha}\frac{\partial\Psi_p}{\partial s_1}(s_1=0,t_1)\end{aligned} \tag{42}$$

and

$$\begin{aligned}-\lrcorner\!\!\lceil\Psi_p(s_3=0,t_3)&=\Big(-\frac{\alpha}{\pi}\Big)^{1/2}\Big[\exp(ipt_3)-2\Big(\frac{\sqrt{3}\alpha+ip}{\sqrt{3}\alpha-ip}\Big)\\&\times\exp\Big[t_3\Big(-\frac{\sqrt{3}}{2}\alpha-\frac{1}{2}ip\Big)\Big]\Big]\\&=\frac{1}{\alpha}\lrcorner\!\!\lceil\frac{\partial\Psi_p}{\partial s_3}(s_3=0,t_3).\end{aligned} \tag{43}$$

Using (41), (42), (43) and the definitions (33) and (34) we obtain the integral expressions for the function h_{13} and h_{33}. For the function h_{33} we obtain

$$\begin{aligned}h_{33}(q,p)&=\frac{(-\alpha)^{1/2}}{\sqrt{2}\,\pi i}\int_{-\infty}^{\infty}dk_3(-k_3+i\alpha)\frac{1}{E-k_3^2-q^2+i0}\\&\times\int_{-\infty}^{\infty}ds_1\int_{-\infty}^{\infty}dt_1\exp\Big[-is_1\Big(-\tfrac{1}{2}k_3+\frac{\sqrt{3}}{2}q\Big)\\&-it_1\Big(-\frac{\sqrt{3}}{2}k_3-\tfrac{1}{2}q\Big)\Big]\cdot V_1\Psi_p(s_1,t_1),\\&=+i\frac{(-2\alpha)^{1/2}}{2\pi}\int_{-\infty}^{\infty}dk_3(-k_3+i\alpha)\\&\times\Big(-\alpha-\frac{i}{2}k_3+\frac{i}{2}\sqrt{3}q\Big)\frac{1}{k_3^2-(p^2-q^2-\alpha^2+i0)}\\&\times\int_0^{\infty}dt_1\exp\Big[it_1\Big(\frac{\sqrt{3}}{2}k_3+\tfrac{1}{2}q\Big)\Big](-)\lrcorner\!\!\lceil\Psi_p(s=0,t_1),\end{aligned} \tag{44}$$

with $E=p^2-\alpha^2$.

We perform the integration over k_3 by closing the contour in the upper half-plane and by calculating the residue at

$$t=\sqrt{p^2-q^2-\alpha^2+i0}. \tag{45}$$

This procedure is justified because of the exponentially decreasing component for $\mathrm{Im}k_3>0$ under the second integral sign. Performing subsequently the integration over t_1, we obtain after algebraic transformations

$$h_{33}(q,p)=-\frac{\sqrt{3}\alpha}{2\pi}\frac{1}{t}(p+q)\left(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\right)\left(\frac{p-2q-i\sqrt{3}\alpha}{p+2q-i\sqrt{3}\alpha}\right). \tag{46}$$

Proceeding in the same way we obtain the following expression for the rearrangement amplitude:

$$h_{13}(q,p)=+\frac{\sqrt{3}}{2\pi}\frac{\alpha}{t}(p+q)\left(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\right)\left(\frac{p-2q+i\sqrt{3}\alpha}{p+2q+i\sqrt{3}\alpha}\right), \tag{47}$$

with t given by (45).

The amplitudes on the energy shell can by calculated from (46) and (47) by putting

$$t=\sqrt{p^2-q^2-\alpha^2}=-i\alpha, \quad \text{for } p^2=q^2. \tag{48}$$

Thus, we have

$$\begin{aligned} &h_{33}(-p,p)=h_{13}(-p,p)=0,\\ &h_{33}(p,p)=-(2\pi)^{-1}2ip,\\ &h_{13}(p,p)=+\frac{1}{2\pi}2ip\left(\frac{\alpha+i\sqrt{3}p}{3\alpha-i\sqrt{3}p}\right)\left(\frac{3\alpha+i\sqrt{3}p}{\alpha-i\sqrt{3}p}\right). \end{aligned} \tag{49}$$

The relation between the on-shell values of the amplitudes h_{i3} and the transmission and reflection probabilities can be deduced by studying the asymptotic behavior of the formal Lippmann—Schwinger equations. The function $\Psi_p(s,t)$ may be written as a solution of the following integral equations:

$$\begin{aligned} \Psi_p(s_3,t_3)=&(-\alpha/\pi)^{1/2}H(-s_3)\exp(-\alpha s_3)\exp(ipt_3)\\ &+\int_{-\infty}^{\infty}dt_3'\int_{-\infty}^{0}ds_3'G_3(s_3,t_3;s_3',t_3';p^2+\alpha^2)\\ &\times V_1(s_3')\Psi_p(s_3',t_3'), \end{aligned} \tag{50}$$

or

$$\begin{aligned} \Psi_p(s_1,t_1)=&\int_{-\infty}^{\infty}dt_1'\int_{-\infty}^{0}ds_1'G_1(s_1,t_1;s_1',t_1';p^2+\alpha^2)\\ &\times V_3(s_1')\Psi_p(s_1',t_1'), \end{aligned} \tag{51}$$

where

$$\begin{aligned} G_i(s_i,t_i;s_i',t_i';p^2-\alpha^2)=&H(-s_i)H(-s_i')(-2\alpha)\exp[-\alpha(s_i+s_i')]\\ &\times\frac{+1}{2ip}\exp(ip|t_i'-t_i|)+\int_0^{\infty}dp'\\ &\times\frac{\exp(i\sqrt{p^2-p'^2-\alpha^2+i0}\,|t_i'-t_i|)}{2i\sqrt{p^2-p'^2-\alpha^2+i0}}\\ &\times\chi_{p'}(s_i)\chi_{p'}(s_i'), \quad i=1,3 \end{aligned} \tag{52}$$

with

$$\begin{aligned} \chi_p(s_i)=(2\pi)^{-1/2}H(-s_i)\{&\exp(ips_i)-[(\alpha+ip)/(\alpha-ip)]\\ &\times\exp(-ips_i)\}. \end{aligned}$$

Studying the asymptotic behavior of the right-hand sides of Eqs. (50) and (51) for $t_3\to\pm\infty$ and $t_1\to\pm\infty$, respectively, we conclude that:

(1) the amplitude of reflection in channel 3 is

$$R_{33}=+(2\pi/2ip)\langle\varphi_3,-p|V_1|\Psi_p\rangle=(2\pi/2ip)h_{33}(-p,p)=0, \tag{53}$$

(b) the amplitude transmission in channel 3 is

$$T_{33}=1+(2\pi/2ip)\langle\varphi_3,p|V_1|\Psi_p\rangle=1+(2\pi/2ip)h_{33}(p,p)=0, \tag{54}$$

(c) the amplitude of the rearrangement process, with the final state momentum q_1 equal to p, is

$$\begin{aligned} R_{13}=&+(2\pi/2ip)\langle\varphi_1,p|V_3|\Psi_p\rangle=+(2\pi/2ip)h_{13}(p,p)\\ =&\left(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\right)\left(\frac{\sqrt{3}\alpha+ip}{\alpha-i\sqrt{3}p}\right), \end{aligned} \tag{55}$$

(d) the amplitude of the rearrangement process with $q_1=-p$ is

$$T_{13}=+(2\pi/2ip)\langle\varphi_1,-p|V_3|\Psi_p\rangle=+(2\pi/2ip)h_{13}(-p,p)=0. \tag{56}$$

The amplitude of the dissociation process can be expressed in terms of the functions h_{13} and h_{33}. Employing (36) we get

$$\begin{aligned} T_{03}(k_1,q_1;p)=&+(-\alpha\pi)^{-1/2}\Bigg[(k_1+i\alpha)\frac{\sqrt{p^2-q_1^2-\alpha^2+i0}}{p^2-q_1^2+i0}h_{13}(q_1,p)\\ &+(k_3+i\alpha)\frac{\sqrt{p^2-q_3^2-\alpha^2+i0}}{p^2-q_3^2+i0}h_{33}(q_3,p)\Bigg], \end{aligned} \tag{57}$$

where $k_1^2+q_1^2=k_3^2+q_3^2=p^2-\alpha^2$ and

$$k_3=-\tfrac{1}{2}k_1-(\sqrt{3}/2)q_1; \quad q_3=(\sqrt{3}/2)k_1-\tfrac{1}{2}q_1.$$

Substituting the expressions (46) and (47) into (57) and performing the calculations, we obtain

$$T_{03}(k_1,q_1;p)\equiv\langle k_1,q_1|V_1+V_3|\Psi_p\rangle=0, \tag{58}$$

in agreement with the result which can be obtained directly by studying the solution (40). The existence of nonzero dissociation probability should manifest itself in the expression of the wave function as a term describing a circular wave and such a term is absent in (40).

No we shall study the Faddeev equations (25) and (26) for the problem of equal masses and equal potentials. In the case, when discussing the half-off-shell form of the equations, we get

$$\begin{aligned} h_{13}(q,p)=&+\left(\frac{2\alpha}{\sqrt{3}\pi}\right)\frac{-q-2p+i\sqrt{3}\alpha}{2q+p+i\sqrt{3}\alpha}+\frac{2i}{\pi}\\ &\times\int_{-\infty}^{\infty}dq'(-q-2q'+i\sqrt{3}\alpha)\\ &\times\frac{1}{3p^2-3\alpha^2-4q^2-4qq'-4q'^2+i0}\\ &\times(-2q-q'+i\sqrt{3}\alpha)\frac{\sqrt{p^2-\alpha^2-q'^2+i0}}{p^2-q'^2+i0}h_{33}(q',p), \end{aligned} \tag{59}$$

$$\begin{aligned} h_{33}(q,p)=&\frac{2i}{\pi}\int_{-\infty}^{\infty}dq'(q+2q'+i\sqrt{3}\alpha)\\ &\times\frac{1}{3p^2-3\alpha^2-4q^2-4qq'-4q'^2+i0} \end{aligned}$$

$$\times(2q+q'+i\sqrt{3}\alpha)\frac{\sqrt{p^2-\alpha^2-q'^2+i0}}{p^2-q'^2+i0}\,h_{13}(q',p). \tag{60}$$

We substitute the exact expressions (46) and (47) for h_{13} and h_{33} into (59) and (60) and try to check the equality of both sides in each of the resulting expressions. The calculations are simplified by the fact that the factor $t=\sqrt{p^2-q'^2-\alpha^2+i0}$ of the integral kernels is cancelled by the factor t^{-1} of the amplitudes. However, differently from the problem of three collinear particles interacting via delta function potentials discussed by Dodd in Ref. 4, the resulting integrals on the right-hand sides of Eqs. (59) and (60) are not convergent:

$$h_{33}(q,p)\overset{?}{=}\left(\frac{+i\alpha}{\pi^2}\right)\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\int_{-\infty}^{\infty}dq' \times\left(\frac{(q+2q'+i\sqrt{3}\alpha)(2q+q'+i\sqrt{3}\alpha)}{4q'^2+4qq'+4q^2+3\alpha^2+3p^2-i0} \times\frac{1}{q'-p-i0}\,\frac{p-2q'+i\sqrt{3}\alpha}{p+2q+i\sqrt{3}\alpha}\right), \tag{61}$$

$$h_{13}(q,p)\overset{?}{=}\left(+\frac{2\alpha}{\sqrt{3}\pi}\right)\frac{-q-2p+i\sqrt{3}\alpha}{2q+p+i\sqrt{3}\alpha}-\left(\frac{i\alpha}{\pi^2}\right)\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip} \times\int_{-\infty}^{\infty}dq'\left(\frac{(-q-2q'+i\sqrt{3}\alpha)(-2q-q'+i\sqrt{3}\alpha)}{4q'^2+4qq'+4q^2-3p^2+3\alpha^2-i0} \times\frac{1}{q'-p-i0}\,\frac{p-2q'-i\sqrt{3}\alpha}{p+2q'-i\sqrt{3}\alpha}\right). \tag{62}$$

The integrals in (61) and (62) are divergent because of the too slow decrease of the integrands for $|q'|\to\infty$. The integrals exist in the principal value meaning, but the calculations performed with the p.v. definitions[9] of the integrals *do not* lead to the equality of the right- and left-hand sides in the Eqs. (61) and (62).

The divergent integrals in (61) and (62) may be formally written as

$$\int_{-\infty}^{\infty}K_3(q,p,q')dq'$$

and

$$\int_{-\infty}^{\infty}K_1(q,p,q')dq', \text{ respectively.} \tag{63}$$

We find that the integrals

$$\int_{-\infty}^{\infty}[K_3(q,p,q')-K_3(\pm p,p,q')]\,dq', \quad \int_{-\infty}^{\infty}[K_1(q,p,q')-K_1(\pm p,p,q')]\,dq' \tag{64}$$

are convergent in the ordinary sense. We can calculate them by summing up the residua at the poles in the upper half-plane:

$$q_{\mathrm{I}}'=p+i0, \quad q_{\mathrm{II}}'=-\tfrac{1}{2}p-i\frac{\sqrt{3}}{2}\alpha, \quad q_{\mathrm{III}}'=-\tfrac{1}{2}q+\frac{\sqrt{3}}{2}t.$$

Proceeding in this way we have checked the following equalities:

$$h_{33}(q,p)-h_{33}(\pm p,p)=\left(\frac{-i\alpha}{\pi^2}\right)\left(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\right) \times\int_{-\infty}^{\infty}dq'[K_3(q,p,q')-K_3(\pm p,p,q')], \tag{65}$$

$$h_{13}(q,p)-h_{13}(\pm p,p)=b_{13}(q,p)-b_{13}(\pm p,p)+\left(\frac{-i\alpha}{\pi^2}\right) \times\left(\frac{\alpha+i\sqrt{3}p}{\sqrt{3}\alpha-ip}\right)\int_{-\infty}^{\infty}dq'[K_1(q,p,q') -K_1(\pm p,p,q')], \tag{66}$$

where $h_{33}(q,p)$, $h_{13}(q,p)$, $h_{33}(\pm p,p)$, $h_{13}(\pm p,p)$ are given by (46), (47), and (49), and we have denoted the inhomogeneous term in the right-hand side of (62) by $b_{13}(q,p)$.

In conclusion, the Faddeev–Lovelace equations (25) and (26) could not be verifed directly. On the other hand, their modified forms which deal with the differences of the amplitudes h_{13} and h_{33} off the energy shell can be directly verified as follows:

$$h_{33}(q,p)-h_{33}(\pm p,p)=\int_{-\infty}^{\infty}dq'[b_{31}(q,q')-b_{31}(\pm p,q')] \times\tau(E-q'^2)h_{13}(q',p), \tag{67}$$

$$h_{13}(q,p)-h_{13}(\pm p,p)=b_{13}(q,p)-b_{13}(\pm p,p) +\int_{-\infty}^{\infty}dq'[b_{13}(q,q')-b_{13}(\pm p,q')] \times\tau(E-q'^2)h_{33}(q',p), \tag{68}$$

It remains an open question, what meaning should be attributed, in our case, to the Lovelace–Faddeev equations in their general form.

Note added in proof: During the process of publication of this paper, the author succeeded in finding the expressions for the cross sections for arbitrary masses and arbitrary constants α.[1] The same result was obtained independently by McGuire and Hurst [J.B. McGuire, C.A. Hurst, J. Math. Phys. **13**, 1595 (1972)]. The energy of the three-body bound state has also been found and for the case $m_1=m_2=m_3$, $\alpha_1=\alpha_2\equiv\alpha$, considered here, is equal to $-4\alpha^2$.

ACKNOWLEDGMENTS

The author is indebted to Professor J. Stecki for his encouragement and interest in this work, and to Dr. J. Wrzecionko for several discussions.

[1]K. Lipszyc, Acta Phys. Pol. **A42**, 571 (1972); **A44**, 115 (1973).
[2]G. D. Maluzhinetz, Dokl. Akad. Nauk SSSR **121**, 436 (1958).
[3]H. M. Nussenzweig, Proc. R. Soc. Lond. **264A**, 408 (1961).
[4]L. R. Dodd, J. Math. Phys. **11**, 207 (1970).
[5]K. Lipszyc and J. Stecki, Bull. Acad. Pol. Sci. Ser. Sci. Chim. **18**, 541 (1970).
[6]C. Lovelace, Phys. Rev. **135**, B1225 (1964).
[7]I. Duck, in *Advances in Nuclear Physics* (Plenum, New York, 1968), Vol. I.
[8]L. Schwartz, *Mathematics for the Physical Sciences* (Hermann, Paris, 1966), English ed., Chap. II, Sec. 2.3.
[9]The definition of the principal value of the integral with infinite limits of integration is as follows: p.v. $\int_{-\infty}^{\infty}f(x)dx=\lim_{M\to\infty}\int_{-M}^{M}f(x)dx$.

A comment on the paper "Lower Bounds on the Energy Eigenvalues of Systems Containing Identical Particles"

F. Calogero

Istituto di Fisica, Università di Roma, 00185 Rome, Italy

Istituto Nazionale di Fisica Nucleare—Sezione di Roma, Rome, Italy

C. Marchioro

Istituto di Matematica, Università di Roma, Rome, Italy

(Received 17 January 1973)

It is pointed out that the extension by Yeh [J. Math. Phys. **13**, 227 (1972)] of the result by Calogero and Marchioro [J. Math. Phys. **10**, 562 (1969)] is invalid. Counterexamples are indicated; the fault hidden in the proof of the extended result is spotted.

Lower bounds to the ground-state energy of systems containing many identical particles interacting by two-body forces (and possibly also with an external potential) can be obtained from a powerful lemma, whose general formulation and proof has been given recently.[1] They are expressed in terms of the ground-state energies of (simpler) systems composed of a smaller number of particles having masses and/or interactions obtained from the original ones by appropriate scaling. This reduction process can be continued all the way down to systems involving one and two particles only. As an example of the results that can be obtained in this manner, we report the formula[1]

$$E_N \geq \tfrac{1}{2}NE_2(1;N-1;1), \tag{1}$$

where E_N is the ground-state energy of the system composed of N identical particles interacting pairwise via a given potential V and with a given external (single-particle) potential W, while $E_2(1;N-1;1)$ is the ground-state energy of the system composed of only two such particles, interacting with the same external potential and, among themselves, by the potential $(N-1)V$, namely by a two-body potential that is $N-1$ times stronger than the original one.

Very recently a paper has been published, in which these results are extended to other energy levels besides the ground state.[2] Thus, for instance, in place of (1), the inequality

$$E_N^i \geq \tfrac{1}{2}NE_2^i(1;N-1;1) \tag{2}$$

has been obtained, corresponding to the statement that "the ith eigenvalue of H_N is not less than $N/2$ times the ith energy eigenvalue of the system composed of two such particles interacting with the same external potential and among themselves through a two-body interparticle potential, which is $N-1$ times stronger than the original interparticle potential".[2] Here of course H_N is the N-particle Hamiltonian, and the energy levels are ordered so that $E_N^i \geq E_N^j$ if $i>j$. For $i=1$, the inequality (2) coincides with (1), and it is certainly true[1]; for $i>1$, the inequality (2) does not generally hold. The more general statement,[2] that extends to excited levels the lemma of Calogero and Marchioro,[1] and from which the inequality (2) follows as a special case, is also invalid.

Indeed, the inequality (2) could not be right, since in the case without external potentials and with two-body regular potentials vanishing fast at infinity there is generally only a finite number of negative-energy states in the two-body problem, but an infinite number of negative states in the N-body problem with $N \geq 3$ (corresponding, for instance, to configurations containing a 2-body cluster and $N-2$ free particles far apart).[3] It is also easy to construct explicit counterexamples to (2), using exactly solvable N-body problems.[4]

The fault in the proof by Yeh[2] need not be discussed in detail, since it is immediately apparent once the following trivial, but subtle, remark is made. Consider the Hamiltonian

$$H_2 = T_1 + T_2 + V_{12} + W_1 + W_2, \tag{3}$$

where T_i is the kinetic energy of the ith particle, W_i the external potential with which the ith particle interacts, and V_{12} the interparticle potential between particles 1 and 2. Let $\psi(1,2)$ be the ground-state wavefunction of the 2-body problem with this Hamiltonian, and E_2 the corresponding energy:

$$H_2\psi(1,2) = E_2\psi(1,2). \tag{4}$$

Consider now a 3-body problem *with this same Hamiltonian*. Clearly the ground state of this problem is characterized by the wavefunction $\psi(1,2)\varphi(3)$, where $\varphi(3)$ is arbitrary (if you like, normalizable), since H_2 neither acts nor depends in any way on the coordinates of the third particle; and the corresponding energy coincides with the ground-state energy of the two-body problem:

$$H_2\psi(1,2)\varphi(3) = E_2\psi(1,2)\varphi(3). \tag{5}$$

But now this energy level is infinitely degenerate (due to the arbitrariness in the choice of φ), in contrast to the situation in the two-body case. This degeneracy, that is ignored in the proof by Yeh,[2] explains why the paradoxical result of Eq. (2) was obtained. If this degeneracy is taken into account, it prevents the extension of the Calogero and Marchioro result[1] to other energy levels besides the ground state, at least for all cases in which the comparison systems contain less particles than the original system.

Only if the comparison system contains as many particles as the original system, and moreover at least its first few eigenvalues are discrete and not infinitely degenerate, the extension of the lower bound to excited states yields nontrivial results.[5] It is possible to invent examples of this kind where the comparison system is

still considerably simpler than the original problem, this being of course a necessary condition for the usefulness of this kind of approach.[5] None of these examples seems, however, to be particularly interesting.

[1]F. Calogero and C. Marchioro, J. Math. Phys. 10, 562 (1969). For previous, but less general, versions of this result see the other papers quoted in this reference.
[2]R.H.T. Yeh, J. Math. Phys. 13, 227 (1972).
[3]If the lack of normalizability of the wavefunctions describing the free particles is considered objectionable, the systems can be enclosed into a large, but finite, box; then the number of negative-energy states in the many-body case is not infinite, but it is still arbitrarily large (since the size of the box can be chosen arbitrarily large), and this is clearly inconsistent with Eq. (2).
[4]F. Calogero, J. Math. Phys. 12, 419 (1971).
[5]The lower bound result can be extended to excited states, using comparison systems involving N particles not all of which interact among themselves. For instance the JMP referee has pointed out that, "for an even number N of particles, $E_N^{(i)} \geq \sum_{s=1}^{N/2} E_2^{(\nu_s^{(i)})}(1;N-1;1)$ $(\nu_1^{(i)} \geq \nu_2^{(i)} \geq \nu_3^{(i)} \geq \cdots \nu_N^{(i)})$, where the notation is that used in (this paper), and the $\nu_s^{(i)}$ are to be chosen so that the sum on the right-hand side achieves the i' th-lowest value possible; e.g., for the first excited state, the inequality becomes $E_N^{(1)} \geq (\frac{1}{2}N-1)E_2^{(0)}(1;N-1;1) + E_2^{(1)}(1;N-1;1)$. This result follows at once from the observation that the exact eigenfunctions for the full N-particle system are legitimate trial eigenfunctions for a system of $\frac{1}{2}N$ independent particle pairs, each having the Hamiltonian $T_1+T_2+W_1+W_2+(N-1)V_{12}$." It should be noted that this result, while certainly correct, is nontrivial [namely, it yields a lower bound for the first excited state that is higher than the bound for the ground state given by Eq. (1)] only if the two-body problem characterized by the Hamiltonian $T_1+T_2+W_1+W_2+(N-1)V_{12}$ has a nondegenerate ground state. A necessary condition for that is the presence of the external potential W, for otherwise translation invariance implies the well known (and, in this context, highly relevant) degeneracy associated with the localization of the center of mass. Thus this lower bound can be stringent only for systems where the external potential plays a rôle, in binding the system, at least comparable to that of the interparticle potential. (This footnote has been added to take into account a point raised by the referee).