On Some Properties of Physical Symmetries*

Jan T. Lopuszanski†
Institute of Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790
(Received 8 February 1971)

It is shown that each symmetry of a field theory with interaction and with a mass gap, induced by a local current, is also induced by at least two other currents, local with respect to the asymptotic free fields and bilinear in these fields. The currents are asymptotic currents in the sense of Araki and Haag [H. Araki and R. Haag, Commun. Math. Phys. 4, 77 (1967)] of the original current. It is also shown that the charges, displaying spinorial transformation character with respect to the Lorentz group, vanish, while the tensorial charges are a linear combination of the scalar charges (internal symmetries) and the energy-momentum vector (translational symmetry).

1. INTRODUCTION

In some recent publications attention was paid to the problem of symmetries in the framework of axiomatic field theory (see, e.g., Refs. 1–9 and 10).

This paper is an extension of ideas previously put forward (see in particular Refs. 2 and 9).

We are concerned here with *physical* symmetries only, by which we mean symmetry groups whose generators are induced by currents, local, and local relative to the considered fields, locally conserved, with definite transformation properties under the Poincaré group.

Assuming that our field theory is constructed out of massive scalar, spinor, vector, spinor-vector, and tensor fields, we show that every physical symmetry is either an internal symmetry or a translation or a trivial combination of both, as soon as there is an interaction present among the fields under consideration.

The existence of "spinorial" charges as a counterpart to scalar (internal symmetry) and vector (translational symmetry) charges are ruled out by our results.

These results are not surprising, although not easy to prove. They state simply that in a relativistic model of field theory no mixing of space–time and internal symmetries can occur. We arrive at these conclusions without using group theoretical arguments.

Unfortunately, our arguments cannot be extended to the case of massless fields, since very little is known about their asymptotic behavior. Our conjecture would be, however, that our results hold true also in this case as far as strict symmetries are concerned, barring spontaneously broken symmetries; the latter—fortunately enough—affect only the case of particles of helicity zero.[11]

We would like to emphasize the crucial role that the interaction among the fields plays in our analysis. The role of interaction was exhibited first, in this context, by Orzalesi, Sucher, and Woo.[3]

For the free-field case, the aforementioned statement no longer need be true.

Since the essentially and spontaneously broken symmetries remain beyond the scope of our investigations, we consider our results as "no-go" theorems.

2. ASSUMPTIONS

We list the relevant assumptions:

(i) The Poincaré group is unitarily implemented in the Hilbert space $\mathcal{K}$ by the operators $U(A, a)$.

Here $A \in SL(2, c)$ and $\{\Lambda^{\nu}_{\mu}(A), a \equiv (a^0, a^1, a^2 a^3)\}$ represents the Poincaré group in the Minkowski space. In particular the translation group is given by $U(1, a) = \exp(iPa)$, where $Pa \equiv P_\mu a^\mu$. We require that the spectrum of P_μ lie in the forward light cone and that the spectrum of P^2 consist of two discrete points: 0, corresponding to the unique vacuum state Ω, and $m^2 > 0$, corresponding to the one-particle Hilbert subspace

$$\mathcal{K}_1 = E_1 \mathcal{K} \tag{2.1}$$

(E_1 is the projection operator on to the subspace $\mathcal{K}_1$), as well as a continuous part $\{\mu^2 > m^2, \infty\}$ corresponding to the scattering states.

(ii) We are interested in a theory of an arbitrary but finite number of quantal fields of spin $0, \frac{1}{2}, 1, \frac{3}{2}$, and 2 which are operator-valued distributions. We assume all the fields to be real.[12]

We denote scalar fields [the $(0, 0)$ representation for free fields] by $\phi^{(i)}(x)$, where $i = 1, \ldots, n_0$ enumerates the fields; the 4-spinor fields [$(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representation] by $\psi^{(i)}_\alpha(x)$, with $\alpha = 1, 2, 3, 4$, $i = 1, \ldots, n_{1/2}$, the vector fields [$(\frac{1}{2}, \frac{1}{2})$ representation] by $\phi^{(i)}_\mu(x)$, with

$$\partial^\mu \phi^{(i)}_\mu(x) = 0, \tag{2.2}$$

with $\mu = 0, 1, 2, 3$, $i = 1, \ldots, n_1$ (for the vector indices we adopt the summation convention); the mixed vector-spinor fields [$(1, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ representations] by $\psi^{(i)}_{\alpha;\mu}(x)$, with

$$\partial^\mu \psi^{(i)}_{\alpha;\mu} = 0 \tag{2.3a}$$

and

$$\sum_\beta \gamma^\mu_{\alpha\beta} \psi^{(i)}_{\beta;\mu} = 0, \tag{2.3b}$$

with $\alpha = 1, 2, 3, 4$, $\mu = 0, 1, 2, 3$, $i = 1, \ldots, n_{3/2}$; here γ^μ are the Dirac matrices in a Majorana representation, all being pure imaginary and unitary, γ^0 being antisymmetric, $\gamma^j, j = 1, 2, 3$, symmetric;[13] finally, tensor fields [$(1, 1)$ representation] by $\phi^{(i)}_{\mu\nu}(x)$, with

$$\phi^{(i)}_{\mu\nu} = \phi^{(i)}_{\nu\mu}, \tag{2.4a}$$

$$\partial^\mu \phi^{(i)}_{\mu\omega} = 0, \tag{2.4b}$$

$$\phi^\mu_\mu = 0, \tag{2.4c}$$

$\mu, \nu = 0, 1, 2, 3;\ i = 1, \ldots, n_2$.

Since the fields we are concerned with are interacting fields and the restrictions (2.2)–(2.4) are usually imposed on free fields, an explanation is in order. Is it possible to maintain these supplementary conditions also in the case of interacting fields? It turns out that this can be done without violating the locality conditions. In general one has only to replace the original interpolating interacting tensor or spinor field by a new local field, which is local with respect to the original one and has the same asymptotic (Lehmann, Symanzik, and Zimmermann) limits (in other words, belongs to the same Borchers class as the original field). The details are presented in Appendix A.

The list of fields given above does not encompass the fields corresponding to the representations $(3/2, 0)$ and $(0, 3/2)$ as well as the representations $(1, 0)$ and $(0, 1)$. These fields, however, can be expressed in terms of other fields already listed. This is also explained in detail in Appendix A on the example of the latter case.

To make precise what we understand by the transformation of a field, we assume that all the fields transform under the same Poincaré representation, *viz.*

$$U(A,a)\phi^{(i)}(x)U(A,a)^+ = \phi^{(i)}(\Lambda x + a), \qquad i = 1, \cdots, n_0, \tag{2.5a}$$

$$U(A,a)\psi^{(i)}_\alpha(x)U(A,a)^+ = \sum_\beta S^{-1}_{\alpha\beta}(A)\psi^{(i)}_\beta(\Lambda x + a), \qquad i = 1, \cdots, n_{1/2}, \tag{2.5b}$$

$$U(A,a)\phi^{(i)}_\mu(x)U(A,a)^+ = \Lambda^{-1}_\mu{}^\nu \phi^{(i)}_\nu(\Lambda x + a), \qquad i = 1, \cdots, n_1, \tag{2.5c}$$

etc., where $\Lambda^\nu_\mu = \Lambda^\nu_\mu(A)$ and $S_{\alpha\beta}(A)$ are four-dimensional representations of the Lorentz group, irreducible and reducible, respectively.

We assume further that the fields are (anti)local and relatively normally (anti)local [i.e., the field of spin $\frac{1}{2}(2k+1)$ anticommute with each other and commute with the fields of spin k, where k is an integer, for spacelike separations). For simplicity

$$[\Omega, \phi^{(j)}(f)\Omega] = 0, \quad \text{where } \phi^{(j)}(f) \equiv \int f(x)\phi^{(j)}(x)dx.$$

Since our results do not depend essentially on the kind of test functions used, we do not specify them, demanding only the locality, cluster decomposition, etc. maintain their meaning.

(iii) To promote aesthetic values as well as to get proper asymptotic conditions, we assume [see (2.1)]

$$(\Omega, \phi^{(i)}(x)E_1\phi^{(j)}(y)\Omega) = i\Delta^{(+)}(x-y; m^2)\delta^{ij}, \tag{2.6a}$$

$$(\Omega, \psi^{(i)}(x)E_1\psi^{(j)}(y)\Omega) = i(i\gamma^\mu\partial_\mu + m)\Delta^{(+)}(x-y; m^2)\delta^{ij}, \tag{2.6b}$$

etc. Since we assumed $m^2 \neq 0$, it follows from (2.6) that the free asymptotic fields exist.[14] We denote them by

$$\phi^{(i)}_{ex}(x), \psi^{(i)}_{ex}(x), \text{ etc.}, \tag{2.7}$$

respectively; "ex" stands for "in" or "out."

These free fields are properly normalized, also due to (2.6), and satisfy normal (anti) commutation relations.

(iv) We assume that one of the sets of asymptotic fields, say the incoming fields, is irreducible. This implies that the original set of interacting fields as well as the outgoing fields is also irreducible.

(v) Let us assume that in our theory the fields exist:

$$T_{\lambda,\omega_1,\cdots,\omega_b}(x) \equiv T_{\lambda,\omega}(x), \tag{2.8}$$

where $\lambda = 0, 1, 2, 3$ is a Minkowski vector index, $(\omega_1, \cdots, \omega_b) \equiv \omega$ can be either spinor or vector indices, such that they are normally (anti) local with respect to other fields, real, transform under the same representation of the Poincaré group as the original fields and are locally conserved with respect to the index "λ," *viz.*

$$\partial^\lambda T_{\lambda,\omega} = 0. \tag{2.9}$$

From the covariance of $T_{\lambda,\omega}$, from (2.9) and $m^2 \neq 0$ it then follows that a meaning can be given (see Ref. 15) to a three-dimensional integral over $T_{0,\omega}$. The quantities defined in this way, Q_ω—we are going to call them charges—are translationally invariant (see Ref. 15) and transform under the Lorentz transformation as geometrical objects characterized by the subindices ω. It can also be shown that[16]

$$Q_\omega\Omega = 0. \tag{2.10}$$

3. GENERAL RESULTS

Our starting point to present the results is an assertion stated in several papers on different occasions (see [1,2,6,7,9]). To avoid too many indices, we present it below for the case of scalar fields only (i.e., $n_{1/2} = n_1 = n_{3/2} = n_2 = 0$), although it is true—under the assumptions stated in Sec. 2—for arbitrarily (but finitely) many fields of different transformation character with respect to the Lorentz group. We have the following:

Statement 1: Under the assumptions stated before

(i) $$i[Q_\omega, \phi^{(j)}_{ex}(x)] = \sum_{l=1}^{n_0} k^{(jl)}_\omega \phi^{(l)}_{ex}(x); \qquad j = 1, \ldots, n_0, \tag{3.1}$$

where $k^{(jl)}_\omega$ is a numerical matrix in (j, l) and a polynomial in $-i\partial$;

(ii) Q_ω is uniquely defined by (3.1);

(iii) $[Q_\omega, S] = 0$; (3.2)

where S is the scattering operator.

The proof is given elsewhere (see, e.g., Ref. 1.). Of the following conclusions that can be drawn from this statement the first one is immediate.

Statement 2: Q_ω is always a bilinear expression in the incoming as well as in the outgoing asymptotic fields.

The other one follows in a slightly more involved way:

Statement 3: Assuming that the fields $T_{\lambda,\omega}(x)$ give rise to the charges Q_ω and that $S \neq 1$, we have at least two sets of fields $T^{(AH)}_{in;\lambda\omega}(x)$ and $T^{(AH)}_{out;\lambda\omega}(x)$, where AH stands for Araki–Haag, each of them is local, is bilinear in and local with respect to the corresponding asymptotic fields, is locally conserved and gives rise to the same charge Q_ω.

For simplicity and clarity of the lecture we shall present the outline of the proof for the special case of the field $T_\lambda(x)$ (no additional indices) and for the theory of scalar fields only.[9] The generalization for $T_{\lambda,\omega}(x)$ and for different kinds of fields will become obvious after this outline is given. It will also become clear from the later results of paragraph four that there is no need to consider other fields than the vector fields $T_\lambda(x)$ as the carriers of the physical symmetries.

According to Ref. 17, in addition to the LSZ limit[18] (see also Ref. 14 and 17) for $t \to \pm\infty$ (which behaves for large t as $t^{-3/2}$), there is an asymptotic limit which behaves like t^{-3}. We shall call it the AH limit. For $T_\lambda(x)$ the LSZ limit is zero since

$$E_1 T_\lambda(f)\Omega = 0, \tag{3.3}$$

so only the AH limit contributes at infinity.[19] This contribution is given by

$$\hat{T}^{(AH)}_{ex;\lambda}(x) = \iint dp dp' \delta(p^2 - m^2)\delta(q^2 - m^2)\theta(p_0)\theta(q_0) \times \sum_i \sum_j (\psi(p;i), T_\lambda(x)\psi(q;j))\tilde{\phi}^{(i)+}_{ex}(p)\tilde{\phi}^{(j)}_{ex}(q), \tag{3.4a}$$

where

$$(\psi(p;i), T_\lambda(x)\psi(q;j)) = e^{i(p-q)x}(p_\lambda + q_\lambda)F_{ij}[(p-q)^2]\theta(p_0)\theta(q_0). \tag{3.4b}$$

Here $\psi(p;j)$ are the scalar one-particle states of momentum $p(p^2 = m^2)$ and of kind j'; $F_{ij}(k^2)$ (defined here for $k^2 \leq 0$ only) is a form factor, analytic in the neighborhood of $k^2 = 0$ and for $k^2 < 0$,[20] with the following properties[21]:

$$F_{ij} = \overline{F_{ji}} = -\overline{F_{ij}}. \tag{3.4c}$$

The expression (3.4a), although locally conserved, is highly nonlocal, not only because it consists exclusively of spacelike field components, but also because of the presence of the form factor (which we assume not to be a constant). By proper manipulation it can be shown that (3.4) integrated over three-dimensional space makes sense and yields a proper charge:

$$Q = -i\sum_i \sum_j F_{ij}(0) \int \theta(p_0)\delta(p^2 - m^2) \times \tilde{\phi}^{(i)+}_{ex}(p)\tilde{\phi}^{(j)}_{ex}(p)dp. \tag{3.5}$$

In order to get (3.5) we made use of Statement 2, which ensures that the three-dimensional space integral of (3.4b) characterizes the charge completely.

We see that the only contribution to the charge comes from the point $p = q$ in (3.4). So we may construct a new field out of $\hat{T}^{(AH)}_{ex;\lambda}$ by replacing $F_{ij}(k^2)$ by $F_{ij}(0)$. This new field still consists only of spacelike components. It can, however, be completed to a local operator by adding terms with momenta lying in the forward and backward light cone. The completion is unique as soon as we demand that the field be local with respect to ϕ_{ex} and real. It yields

$$T^{(AH)}_{ex;\lambda}(x) \equiv -\frac{i}{\pi}\sum_i \sum_j F_{ij}(0) : \frac{\partial \phi^{(i)}_{ex}}{\partial x_\lambda} \phi^{(j)}_{ex} : (x). \tag{3.6}$$

This accomplishes the outline of the proof. We may call $T^{(AH)}_{ex;\lambda}(x)$ asymptotic fields to $T_\lambda(x)$ in the sense of AH in analogy to asymptotic fields in the sense of LSZ.

In general case (3.3) does not hold (see Ref. 19). This, however, does not affect our construction. In case (3.3) is violated both limits exist; the LSZ limit is then responsible for the particle properties of $T_\lambda(x)$, whilst the AH limit displays the symmetry properties induced by $T_\lambda(x)$.

In the general case when there are several kinds of underlying fields involved displaying different transformation character with respect to the Lorentz group the fields

$$T^{(AH)}_{ex,\lambda}(x)$$

are no longer uniquely defined. This is due to the fact that the structure of the one- one-particle amplitudes of $T_\lambda(x)$, as presented in (3.4b) and (3.4c), becomes more complicated. E.g., in the presence of the spinor field we have in addition to

$$T^{(AH)(1)}_{ex,\lambda}(x) \equiv -\frac{i}{2m} : (\partial_\mu \overline{\psi}\cdot\psi - \overline{\psi}\partial_\mu\psi) : (x)$$

[which is a counterpart of (3.6)] also

$$T^{(AH)(2)}_{ex,\lambda}(x) \equiv +\frac{i}{4m} : \partial^\nu\{\overline{\psi}(\gamma_\lambda\gamma_\nu - \gamma_\nu\gamma_\lambda)\psi\} : (x),$$

which does not contribute to the charge. Any linear combination of these two fields, real and properly normalized, yields the same charge.

A corollary follows.

Statement 4: Assume that the field $T_{\lambda\omega}(x)$ can be expanded in a functional series with respect to normal ordered products of the free asymptotic fields. Then the contribution to the charge comes only from the bilinear term in the free fields of the series expansion.

Thus if (3.3) does not hold and $T^{(\mathrm{LSZ})}_{\mathrm{ex};\lambda,\omega}(x)$ exists, we have always

$$\int d^3x T^{(\mathrm{LSZ})}_{\mathrm{ex};0,\omega}(x) = 0,$$

in spite of locality properties of $T_{\lambda,\omega}$ and $T^{(\mathrm{LSZ})}_{\mathrm{ex};\lambda,\omega}$ as well as

$$\partial^\mu T^{(\mathrm{LSZ})}_{\mathrm{ex};\mu,\omega} = 0.$$

This also follows easily from the observation that each Poincaré invariant operator that annihilates the vacuum and can be presented as a functional normal series expansion with respect to the free incoming fields, consists of terms 2-, 4-, 5-, etc., linear in these fields.

4. THE STRUCTURE OF THE CHARGES

We shall now turn to a more detailed analysis of the symmetries obtained on the basis of Statement 1 by examining the structure of the matrix $k^{(ij)}_\alpha(-i\partial)$.

Let us first consider the case when the underlying set of fields consists only of scalar fields (i.e., $n_{1/2} = n_1 = n_{3/2} = n_2 = 0$). We have

Statement 5: Let us assume that n_0 asymptotic scalar real fields form an irreducible set, the rest of the assumptions listed in Sec. 2 remaining unchanged. Let us make the additional hypothesis that each of the scalar fields interacts with the other fields.

(i) Then the charges, displaying a spinor character, vanish.

(ii) The set of charges, transforming as tensors (i.e., $\omega_1 \cdots \omega_b$ are vector indices), consists of scalar charges ($b = 0$), vector charges ($b = 1$), which are proportional to the energy-momentum vector P_μ, and of tensor charges of rank $b > 1$, which are linear combinations of the scalar charges, of P_μ and of the invariant tensors $\delta_{\mu\nu}$ and $\epsilon_{\kappa\lambda\mu\nu}$.

We would like to stress that for the proof of Statement 5, presented below, it is essential that the set of underlying fields consists of scalar fields, in contradistinction to previous statements. The extension of this statement to the case of several fields of different transformation character under the Lorentz group, listed in Sec. 2, is given later (see Statement 6 and Appendix C). However, we were not able to give a proof for the general case of arbitrary, but finite number of different fields.

We would also like to emphasize the importance of the additional assumption made at the end of Statement 5 about the role of the interaction, first exhibited in Ref. 3.

Proof: The first part of the Statement follows trivially. In order that the relation

$$i[Q_\alpha, \phi^{(i)}_{\mathrm{ex}}(x)] = \sum_{j=1}^{n_0} k^{(ij)}_\alpha \phi^{(j)}_{\mathrm{ex}}(x),$$

where $\alpha = 1, 2, 3, 4$ is a spinor index, be consistent, $k^{(ij)}_\alpha$ has to be a covariant bispinor function. However, no such function exists. Hence it follows that

$$k^{(ij)}_\alpha = 0.$$

Since $\{\phi_{\mathrm{ex}}\}$ form an irreducible set and Q_α annihilates the vacuum,

$$Q_\alpha = 0 \tag{4.1}$$

The proof can be immediately generalized to any case of spinorial charges.

The case of scalar charges does not require any additional comments (see, e.g., [1]); $k^{(ij)}$ is in this case a real, antisymmetric (anti-Hermitian) $(n_0 \times n_0)$ matrix.

The case of vector charges was also treated extensively[3,7] (see also Ref. 2). It is worthwhile to notice that here for the first time one makes use of the assumption about the interaction of the fields and that this assumption is crucial for the proof. All one needs is that there be elastic two-particle collisions between different species of particles. Then

$$k^{(ij)}_\nu(-i\partial) = c\delta^{ij}\partial_\nu, \tag{4.2}$$

where c is a real number.

Next we outline the proof for the case of tensor charges of rank 2. After this is done, it will become obvious how to extend the proof to the cases of tensors of higher rank.

We have

$$[Q_{\mu\nu}, \tilde{\phi}^{(i)}_{\mathrm{ex}}(p)] = \sum_{k=1}^{n_0} (a^{(ik)} p_\mu p_\nu + b^{(ik)} g_{\mu\nu}) \tilde{\phi}^{(k)}_{\mathrm{ex}}(p), \tag{4.3}$$

with $\mu, \nu = 0, 1, 2, 3$. $\tilde{\phi}(p)$ denotes the Fourier transform of $\phi(x)$. Both $a^{(ik)}$ and $b^{(ik)}$ are independent of p, Hermitian, pure imaginary (antisymmetric) $(n_0 \times n_0)$ matrices.[22] Notice that, although $ik^{(ik)}_{\mu\nu}(p)$ of (3.1) is here a Hermitian matrix and can be diagonalized, the matrices $a^{(ik)}$ and $b^{(ik)}$ do not need to commute, consequently the diagonalization of $ik^{(ik)}_{\mu\nu}(p)$ would depend crucially on the choice of the value of p. This is not only inconvenient but even makes the procedure useless since (4.3) has to be understood in the sense of distributions with respect to p.

Let us choose $\mu = 0$ and $\nu = 1$. We get from (4.3)

$$[Q_{01}, \tilde{\phi}^{(i)}_{\mathrm{ex}}(p)] = \sum_{k=0}^{n_0} a^{(ik)} p_0 p_1 \tilde{\phi}^{(k)}_{\mathrm{ex}}(p). \tag{4.4a}$$

Now we may diagonalize the matrix $a^{(ik)}$ by means of a complex unitary transformation (notice that this transformation does not depend on the index ex). Let us denote the new, in general, complex fields by

$$C_{ex}^{(i)}(x).$$

Then (4.4a) reads

$$[Q_{01}, C_{ex}^{(i)}(x)] = -\lambda^{(i)}\partial_0\partial_1 C_{ex}^{(i)}(x), \tag{4.4b}$$

where $\lambda^{(i)}$ are (real) eigenvalues of $a^{(i,k)}$.

The procedure outlined in Appendix B leads us to the conclusion that

$$\lambda^{(i)} \equiv 0.$$

Inserting this result back into (4.3) we get

$$\tfrac{1}{4}[Q_\mu^\mu, \tilde{\phi}_{ex}^{(i)}(p)] = \sum_{k=1}^{n_0} b^{(ik)}\tilde{\phi}_{ex}^{(k)}(p). \tag{4.4c}$$

In this way we reduce the problem to the case of a scalar charge, this charge being $\frac{1}{4}Q_\mu^\mu \equiv Q$. The ansatz

$$Q_{\mu\nu} = \tfrac{1}{4}g_{\mu\nu}Q_\mu^\mu = g_{\mu\nu}Q \tag{4.5}$$

solves (4.3). The uniqueness of the solution (4.5) follows from the assumed irreducibility of the set $\{\phi_{ex}\}$ as well as from

$$Q_{\mu\nu}\Omega = 0.$$

This completes the proof.

The extension of our proof to the case of tensor charges of higher rank follows more or less along the same lines, using slightly different but simple computational tricks and is rather straightforward.

Notice that the linearity of the tensor charges with respect to the scalar and vector charges also follows from Statement 2.

As soon as we give up the assumption of Statement 5 about the irreducibility of the set of scalar fields and admit other fields of a different transformation character under the Lorentz transformation, things become involved. We have the following statement.

Statement 6: Let us add to the assumptions listed in Sec. 2 the hypothesis that all the fields under consideration interact with each other.

Then (i) the commutator of the charge with a free asymptotic field is a linear combination of free asymptotic fields of the same transformation character under the Poincaré group. (ii) The assertion expressed in Statement 5 remains true.

Notice that we do not assume here that the fields $\phi_{\mu\nu}(x)$, giving rise to the vector charges Q_μ, are symmetric in the indices μ and ν (such an assumption was made in Refs. 3, 4, and 7).

The main idea of the proof is as follows: If $S \neq 1$, then there are infinitely many scalar products

$$(\Psi_{out}, \Phi_{in}) \neq 0$$

for which

$$(\Psi_{in}, \Phi_{in}) = 0.$$

If we restrict ourselves to Ψ and Φ of the dense set of vectors

$$P(\varphi_{ex}, \psi_{ex})\Omega,$$

where P is a polynomial in the smeared fields, then we may conclude from the equality

$$([Q_\omega, P_{out}^{(1)}]\Omega, P_{in}^{(2)}\Omega) = (P_{out}^{(1)}\Omega, [Q_\omega, P_{in}^{(2)}]\Omega) \tag{4.6}$$

and from (3.1) that (4.6) leads to a violation of the conservation law for energy and momentum, unless Q_ω possesses the nature pointed out in Statement 5 (ii) or the incoming and outgoing fields coincide (i.e., $S = 1$). The main reason for this is that the scattering amplitudes do not factorize into products of two-point functions except in a free field case.

In Appendix C we give the outline of the proof of the Statement 6 for the case when there is a nontrivial scattering of two particles. The extension of the proof to other cases of scattering becomes then obvious.

5. FINAL REMARK

Let us have a closer look at the results obtained so far.

It is clear from (i) of Statement 6 that the charges do not carry a spin, unless there is no interaction present.

Statement 6, which in our opinion can be extended to any finite number of fields of any transformation character under the Lorentz group, clearly exhibits the separation of the internal symmetries from the space-time symmetries for interacting systems with the smallest mass different from zero.[23] It seems to be of some interest that this separation no longer holds for the case of free fields (no interaction present). Let us elaborate on this point and consider the two following examples of free fields ϕ and ϕ_μ:

$$\begin{aligned} i[Q'_\mu, \tilde{\phi}(p)] &= b\tilde{\phi}_\mu(p), \\ i[Q'_\mu, \tilde{\phi}_\nu(p)] &= -b(g_{\mu\nu} - m^{-2}p_\mu p_\nu)\tilde{\phi}(p); \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} [Q'_\mu, \overset{(+)}{\tilde{\phi}}_\nu(p)] &= d\epsilon_{\mu\nu\lambda\kappa}p^\kappa \overset{(-)}{\tilde{\phi}}{}^\lambda(p), \\ [Q'_\mu, \overset{(-)}{\tilde{\phi}}_\nu(p)] &= -d\epsilon_{\mu\nu\lambda\kappa}p^\kappa \overset{(+)}{\tilde{\phi}}{}^\lambda(p), \end{aligned} \tag{5.2}$$

with b and d real (consult Eqs. (C.21a) and (C.21b) as well as (C.32b) and (C.32c) of which (5.1) and (5.2) are special cases). Both transformations,

(5.1) and (5.2), do not commute with the Lorentz group (they commute with the translation group) the first one mixing the vector and scalar; the other, the vector and skew-symmetric tensor fields. Both are unitarily implemented. To see it clearer, let us look at the finite transformations induced by the infinitesimal mapping (5.1). We have

$$\tilde{\phi} \to \tilde{\phi}' = C\tilde{\phi} + (S/V)a^\mu \tilde{\phi}_\mu,$$
$$(a^\mu \tilde{\phi}_\mu) \to (a^\mu \tilde{\phi}'_\mu) = -VS\tilde{\phi} + Ca^\mu \tilde{\phi}_\mu, \tag{5.3a}$$

$$C \equiv \cos(bV),$$
$$S \equiv \sin(bV), \tag{5.3b}$$
$$V \equiv +[a^2 - (a^\mu p\mu)^2 m^{-2}]^{1/2},$$

and $a_\mu = (a_0, a_1, a_2, a_3)$ are the parameters of the group.

Since the free-field Wightman functions are built out of two-point Wightman functions, it is enough to check whether the two-point Wightman functions do not change under the mapping (5.3). This is, indeed, the case, e.g.,

$$\begin{aligned} W_2(\phi') &\equiv (\Omega, \tilde{\phi}'^+(p)\tilde{\phi}'(p)\Omega) \\ &= \left(\Omega, \left(C\tilde{\phi}^+ + \frac{S}{V}a^\mu \tilde{\phi}^+_\mu\right)\left(C\tilde{\phi} + \frac{S}{V}a^\nu \tilde{\phi}_\nu\right)\Omega\right) \\ &= C^2(\Omega, \tilde{\phi}^+\tilde{\phi}\Omega) + \frac{S^2}{V^2}a^\mu a^\nu (\Omega, \tilde{\phi}^+_\mu \tilde{\phi}_\nu \Omega) \\ &= iC^2\tilde{\Delta}^{(+)}(p) + i\frac{S^2}{V^2}a^\mu a^\nu \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{m^2}\right)\tilde{\Delta}^{(+)}(p) \\ &= i\tilde{\Delta}^{(+)}(p). \end{aligned}$$

Thus the mapping (5.3) is unitarily implemented by the operators $W(a_0, a_1, a_2, a_3)$.

We can also find easily the explicit expression for the charges appearing in (5.1) and (5.2), e.g., in case (5.2) we find

$$Q'^\mu = -d\epsilon^{\mu\nu\kappa\lambda}\int p_\kappa \left(\overset{(+)}{\tilde{\phi}}{}^+_\lambda(p)\overset{(-)}{\tilde{\phi}}_\nu(p) + \overset{(-)}{\tilde{\phi}}{}^+_\nu(p)\overset{(+)}{\tilde{\phi}}_\lambda(p)\right) \times \delta(p^2 - m^2)\theta(p)dp. \tag{5.4}$$

We can also find the corresponding currents giving rise to the symmetries. These currents are local tensor fields, locally conserved and relatively local to the free fields underlying the theory.

It is obvious from (5.1) and (5.2) that these transformations are highly nonlocal. We see from this that locality of the current as well as bilinearity of the corresponding charge with respect to the free fields does not guarantee the induced symmetry to be local. This, however, seems no longer to be true as soon as interaction is switched on.

It is well known that local (internal) symmetries have to commute with the Poincaré group.[3,5,6] The converse statement is, of course, not true, to quote only two simple examples of a nontrivial S operator, *viz.*,

$$S\phi_{\text{in}}(x)S^+ = \phi_{\text{out}}(x), \tag{5.5}$$

or a symmetry generated by the number operator of the incoming particles, N_{in}, *viz.*,

$$e^{i\alpha N_{\text{in}}}\phi_{\text{in}}(x)\, e^{-i\alpha N_{\text{in}}} = e^{i\alpha}\phi^{(+)}_{\text{in}} + e^{-i\alpha}\phi^{(-)}_{\text{in}}, \tag{5.6}$$

where $\phi^{(+)}_{\text{in}}$ is the positive frequency part of ϕ_{in}; $0 \le \alpha < 2\pi$. In the latter case, (5.6), the fields ϕ_{in} and

$$e^{i\alpha}\phi^{(+)}_{\text{in}} + e^{-i\alpha}\phi^{(-)}_{\text{in}}$$

are at least relatively quasilocal, while in the case of (5.5) even that is not true. If we demand, in addition to the requirement of Poincaré invariance, that the symmetry transformation also be TCP invariant, both examples are ruled out.

We claim that in a theory of interacting fields every Poincaré and TCP invariant transformation group commuting with the S matrix either is strictly local or completely non-local (by the latter we mean that it is not even almost or quasi-local), the local transformations originating from generators induced by local currents. Unfortunately we are not able to prove this assertion. We may, however, give some arguments in favor of it.

Let us restrict ourselves again for simplicity to the case of scalar fields and vector currents. Consider an arbitrary Poincaré invariant selfadjoint operator Q which annihilates the vacuum, transforms under the TCP transformation according to

$$\theta Q \theta^{-1} = -Q$$

and can be represented as a functional power series normal expansion with respect to the free incoming fields.

Let us denote the term of Q, bilinear in the incoming fields, by Q'_{in}. The operator Q'_{in} has the structure of a physical charge giving rise to a local unitary transformation for the incoming fields and to a local current.

The operator

$$Q - Q'_{\text{in}} \equiv Q''_{\text{in}},$$

the lowest term of which in the power expansion is quadrilinear in the incoming fields (assuming that $Q''_{\text{in}} \neq 0$), can also be considered as a generator of a symmetry. However, this symmetry is not local with respect to the incoming fields. This follows from the observation that Q''_{in} annihilates the one particle states as well as the vacuum. Consequently

$$[Q''_{\text{in}}, \phi^{(j)}_{\text{in}}]\Omega = 0.$$

If

$$[Q''_{\text{in}}, \phi^{(j)}_{\text{in}}]$$

could be localized, then the application of Reeh–Schlieder theorem implies that

$$Q''_{\text{in}} = 0.$$

From that which was said so far we infer that any

operator Q, with $Q''_{\text{in}} \neq 0$, does give rise to a symmetry which is not local with respect to both the incoming and interacting fields. The latter can be shown as follows: Should

$$[Q, \phi^{(j)}]$$

be local with respect to ϕ's, this would imply a similar relation for the incoming fields, which is ruled out by our earlier considerations.

Although every $Q = Q'_{\text{in}}$ gives rise to a strictly local gauge symmetry for the incoming fields, this does not need to be true for the interacting fields. This is the part of the assertion we are not able to prove.

Assuming the assertion to be true we get the following

Conjecture Under the premises of Statement 6, every physical symmetry acts on the fields locally with a finite space–time translation.

A closer look at the transformations (5.1) and (5.2) reveals that the reason for their survival is rather trivial. The fields under consideration are free fields, thus independent of each other. The transformations induce a one-to-one mapping of two separate von Neumann algebras of these fields.[24] This mapping is disturbed by switching on the interaction among the fields.

We end with the following observation. The transformations (5.1) and (5.2) are prefectly legitimate for incoming as well as outgoing fields separately. Since the Wightman functions for the interacting fields can be constructed out of the Wightman functions, say, of the incoming fields, so the transformations (5.1) and (5.2) for the incoming fields give rise to a symmetry of the Wightman functionals of the interacting as well as outgoing fields. It is, however, obvious that this particular symmetry cannot be termed physical, since there are no local fields $T_{\mu\nu}(x)$ relatively local to the interacting or to the outgoing fields yielding the corresponding charges. In other words, the charges of (5.1) and (5.2) for the incoming fields do not commute with the S matrix.

ACKNOWLEDGMENTS

The author expresses his sincere gratitude to Professor C. N. Yang for the warm hospitality extended to him during his stay at S.U.N.Y. as well as to the Authorities of CERN where the earlier stage of this work was accomplished. The author is also indebted to Dr. Benjamin W. Lee, Dr. Michael Wise, and Dr. Ching Hung Woo for inspiring and fruitful discussions.

APPENDIX A

1. It is well known that each free vector field, let us call it $A^{(0)}_\mu$, can be uniquely decomposed into a spin 1 and a spin 0 part, *viz.*,

$$A^{(0)}_\mu = \phi^{(0)}_\mu + \partial_\mu \phi^{(0)}, \tag{A1a}$$

where

$$\phi^{(0)}_\mu \equiv (g_{\mu\nu} + m^{-2}\partial_\mu\partial_\nu)A^{(0)\nu}, \tag{A1b}$$

$$\phi^{(0)} \equiv -\, m^{-2}\partial^\nu A^{(0)}_\nu, \tag{A1c}$$

with

$$\partial^\mu \phi^{(0)}_\mu = 0. \tag{A1d}$$

Let us now examine the interacting vector field A which is supposed to tend, in the LSZ limit[14,18] to the free asymptotic fields $A_{\text{in},\mu}$ and $A_{\text{out},\mu}$. We have the identity

$$A_\mu = m^{-2}(\Box + m^2)A_\mu + \phi_\mu + \partial_\mu \phi, \tag{A2a}$$

with

$$\phi_\mu \equiv m^{-2}(-\Box g_{\mu\nu} + \partial_\mu\partial_\nu)A^\nu, \tag{A2b}$$

$$\phi \equiv -\, m^{-2}\partial^\nu A_\nu. \tag{A2c}$$

Obviously, in the LSZ limit one gets

$$\phi_\mu \xrightarrow[t \to \pm\infty]{\text{LSZ}} \phi_{\text{ex},\mu} \equiv (g_{\mu\nu} + m^{-2}\partial_\mu\partial_\nu)A^\nu_{\text{ex}}, \tag{A3a}$$

and

$$\phi \xrightarrow[t \to \pm\infty]{\text{LSZ}} \phi_{\text{ex}} \equiv -\, m^{-2}\partial^\nu A_{\text{ex},\nu}. \tag{A3b}$$

The field

$$m^{-2}(\Box + m^2)A_\mu,$$

appearing in (A2a), is local with respect to A and vanishes asymptotically. Thus

$$A'_\mu \equiv A_\mu - m^{-2}(\Box + m^2)A_\mu$$

is in the same Borchers class as A_μ and can be used as the interpolating field. Thus we have

$$A'_\mu = \phi_\mu + \partial_\mu \phi, \tag{A4a}$$

with [see (A3)]

$$A'_{\text{ex},\mu} = A_{\text{ex},\mu}, \tag{A4b}$$

and

$$\partial^\mu \phi_\mu = 0. \tag{A4c}$$

The last equation corresponds to (2.2).

Things go similarly for the symmetrical tensor field $A_{\mu\nu}$. The decomposition of the free field $A^{(0)}_{\mu\nu}$ into spin 2, 1, and 0 parts reads

$$A^{(0)}_{\mu\nu} = \phi^{(0)}_{\mu\nu} + \partial_\mu \phi^{(0)}_\nu + \partial_\nu \phi^{(0)}_\mu + \partial_\mu\partial_\nu \phi^{(0)}_I + g_{\mu\nu}\phi^{(0)}_{II}, \tag{A5a}$$

where

$$\begin{aligned}\phi^{(0)}_{\mu\nu} \equiv A^{(0)}_{\mu\nu} &+ \frac{1}{m^2}\left(g_{\rho\nu} + \frac{1}{m^2}\partial_\rho\partial_\nu\right)\partial_\mu\partial_\lambda A^{(0)\rho\lambda} \\ &+ \frac{1}{m^2}\left(g_{\rho\mu} + \frac{1}{m^2}\partial_\rho\partial_\mu\right)\partial_\nu\partial_\lambda A^{(0)\rho\lambda} - \frac{1}{3m^2}\partial_\mu\partial_\nu\left(A^{(0)\lambda}_{\lambda}\right. \\ &\left. + \frac{4}{m^2}\partial^\sigma\partial^\rho A^{(0)}_{\sigma\rho}\right) - \frac{g_{\mu\nu}}{3}\left(A^{(0)\lambda}_{\lambda} + \frac{1}{m^2}\partial^\sigma\partial^\rho A^{(0)}_{\sigma\rho}\right),\end{aligned} \tag{A5b}$$

$$\phi_\nu^{(0)} \equiv -\frac{1}{m^2}\left(g_{\rho\nu} + \frac{1}{m^2}\partial_\rho\partial_\nu\right)\partial_\mu A^{(0)}_{\rho\mu}, \tag{A5c}$$

$$\phi_I^{(0)} \equiv \frac{1}{3m^2}\left(A^{(0)\lambda}_\lambda + \frac{4}{m^2}\partial^\sigma\partial^\rho A^{(0)}_{\sigma\rho}\right), \tag{A5d}$$

$$\phi_{II}^{(0)} \equiv \tfrac{1}{3}\left(A^{(0)\lambda}_\lambda + \frac{1}{m^2}\partial^\sigma\partial^\rho A^{(0)}_{\sigma\rho}\right). \tag{A5e}$$

The interpolating interacting field $A'_{\mu\nu}$, equivalent to $A_{\mu\nu}$ and local with respect to it, is

$$A'_{\mu\nu} = \phi_{\mu\nu} + \partial_\mu\phi_\nu + \partial_\nu\phi_\mu + \partial_\mu\partial_\nu\phi_I + g_{\mu\nu}\phi_{II}, \tag{A6a}$$

with

$$\phi_{\mu\nu} \equiv \frac{1}{m^4}\{(\Box)^2 A_{\mu\nu} + (-\Box g_{\rho\nu} + \partial_\rho\partial_\nu)\partial_\mu\partial_\lambda A^{\rho\lambda} + (-\Box g_{\rho\mu} + \partial_\rho\partial_\mu)\partial_\nu\partial_\lambda A^{\rho\lambda} - \tfrac{1}{3}\partial_\mu\partial_\nu(-\Box A^\lambda_\lambda + 4\partial^\sigma\partial^\rho A_{\sigma\rho}) - \tfrac{1}{3}g_{\mu\nu}[(\Box)^2 A^\lambda_\lambda - \Box\partial^\sigma\partial^\rho A_{\sigma\rho}]\}, \tag{A6b}$$

$$\phi_\nu \equiv -\frac{1}{m^4}(-\Box g_{\rho\nu} + \partial_\rho\partial_\nu)\partial_\mu A^{\rho\mu}, \tag{A6c}$$

$$\phi_I \equiv \frac{1}{3m^4}(-\Box A^\lambda_\lambda + 4\partial^\sigma\partial^\rho A_{\sigma\rho}), \tag{A6d}$$

$$\phi_{II} \equiv \frac{1}{3m^2}(-\Box A^\lambda_\lambda + \partial^\sigma\partial^\rho A_{\sigma\rho}). \tag{A6e}$$

We have also

$$\partial^\nu\phi_\nu = 0, \tag{A6f}$$

as well as

$$\phi^\mu_\mu = 0 \tag{A6g}$$

and

$$\partial^\mu\phi_{\mu\nu} = 0. \tag{A6h}$$

Obviously $A_{\mu\nu}$ as well as $A'_{\mu\nu}$ converge in the LSZ limit to $A_{ex,\mu\nu}$.

So we succeeded in showing that (2.2–2.4) can also be imposed on the interacting fields without violation of the locality conditions.

2. To explain why we omitted in the list of fields the representation (1, 0) and (0, 1), let us recall that these representations correspond to antisymmetric tensor fields:

$$F_{\mu\nu}(x) = -F_{\nu\mu}(x). \tag{A7}$$

If the field (A7) is a free (real) field (with non-vanishing mass), it can be always presented as

$$F^{(0)}_{\mu\nu}(x) = \frac{1}{m^2}[\partial_\mu A^{(0)}_\nu(x) - \partial_\nu A^{(0)}_\mu(x)] + \frac{1}{m^2}\epsilon_{\mu\nu\kappa\lambda}\partial^\kappa B^{(0)\lambda}(x), \tag{A8a}$$

where $A^{(0)}_\mu$ and $B^{(0)}_\mu$ are two independent (real) free vector fields. The dual to $F^{(0)}_{\mu\nu}$ then reads[25,26]

$$\hat{F}^{(0)}_{\mu\nu}(x) = \frac{1}{m^2}[\partial_\mu B^{(0)}_\nu(x) - \partial_\nu B^{(0)}_\mu(x)] - \frac{1}{m^2}\epsilon_{\mu\nu\kappa\lambda}\partial^\kappa A^{(0)\lambda}(x). \tag{A8b}$$

In case $F_{\mu\nu}$ is no longer a free field, the field constructed according to the recipe (A8), let us call it $F'_{\mu\nu}$, no longer coincides in general with $F_{\mu\nu}$. According to (A8a) we define the (interacting) fields

$$A_\nu \equiv -\partial^\mu F_{\mu\nu},$$

$$B_\nu \equiv -\partial^\mu\hat{F}_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\kappa\lambda}\partial^\mu F^{\kappa\lambda}.$$

Consequently,

$$F'_{\mu\nu} = -\frac{1}{m^2}(\partial_\mu\partial^\lambda F_{\lambda\nu} - \partial_\nu\partial^\lambda F_{\lambda\mu}) + \frac{1}{2m^2}\epsilon_{\mu\nu\kappa\lambda}\epsilon^{\tau\lambda\sigma\rho}\partial^\kappa\partial_\tau F_{\sigma\rho} = -\frac{1}{m^2}\partial^\lambda\partial_\lambda F_{\mu\nu}. \tag{A9}$$

Of course, $F'_{\mu\nu}$ is local and local relative to $F_{\mu\nu}$ and gives rise to the same asymptotic fields. In the subsequent investigation, we are mainly concerned with the asymptotic free fields, so we may leave out the aforementioned cases since these can be covered by proper use of the fields $\phi^{(i)}_{\alpha,\mu}$ and $\phi^{(i)}_\mu$, respectively, in the asymptotic field.

APPENDIX B

According to (44b) we have

$$[Q_{o1}, C(x)] = -\lambda\partial_o\partial_1 C(x), \tag{B1}$$

where we have dropped the indices i and ex. Of course, we also have

$$[Q_{o1}, C^+(x)] = \lambda\partial_o\partial_1 C^+(x). \tag{B2}$$

If $C(x) = C^+(x)$, then $\lambda = 0$, so let us concentrate on the case where C is complex.

Assume that there is an elastic scattering between particles and antiparticles described by the fields C and C^+, different in the forward and backward directions; specifically, that the amplitude

$$[u^+_{out}(\overset{1}{p})\nu^+_{out}(\overset{2}{p})\Omega, u^+_{in}(\overset{3}{p})\nu^+_{in}(\overset{4}{p})\Omega] \neq 0,$$

where

$$\overset{1}{p} + \overset{2}{p} = \overset{3}{p} + \overset{4}{p}, \tag{B3}$$

and u, ν, and u^+, ν^+ are annihilation and creation operators for the particles and antiparticles, respectively. Taking into account (B1) and (B2) we have

$$\begin{aligned}&(u^+_{out}(\overset{1}{p})\nu^+_{out}(\overset{2}{p})\Omega, Q_{o1}u^+_{in}(\overset{3}{p})\nu^+_{in}(\overset{4}{p})\Omega)\\ &= -\lambda(\overset{1}{p}_o\overset{1}{p}_1 - \overset{2}{p}_o\overset{2}{p}_1)(u^+_{out}(\overset{1}{p})\nu^+_{out}(\overset{2}{p})\Omega, u^+_{in}(\overset{3}{p})\nu^+_{in}(\overset{4}{p})\Omega)\\ &= -\lambda(\overset{3}{p}_o\overset{3}{p}_1 - \overset{4}{p}_o\overset{4}{p}_1)\\ &\quad\times(u^+_{out}(\overset{1}{p})\nu^+_{out}(\overset{2}{p})\Omega, u^+_{in}(\overset{3}{p})\nu^+_{in}(\overset{4}{p})\Omega).\end{aligned} \tag{B4}$$

From (B4) it follows that either $\lambda = 0$ or

$$\begin{vmatrix} 1 & 2 \\ p_o p_1 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ p_o p_1 \end{vmatrix} \qquad \text{(B5)}$$

$$\begin{vmatrix} 2 & 1 \\ p_o p_1 \end{vmatrix} \qquad \begin{vmatrix} 4 & 3 \\ p_o p_1 \end{vmatrix}$$

or both.

Taking into account (B3) it follows from (B5) that

$$\frac{p_o^2 - p_o^4}{p_o^1 + p_o^2} = \frac{p_1^1 - p_1^3}{p_1^1 + p_1^2}. \qquad \text{(B6a)}$$

However, $p_o^1, p_o^2, p_o^4 > 0$, thus

$$\frac{p_o^2 - p_o^4}{p_o^1 + p_o^2} < 1. \qquad \text{(B6b)}$$

On the other hand, the choice of momenta p^1, p^2, p^3, and p^4 lying in the closest (open) neighborhood of

$$p_1^1 + p_1^2 = 0,$$
$$p_1^1 = p_1^4 \neq 0,$$
$$p_1^2 = p_1^3,$$

violates (B6b). Thus we have shown that $\lambda = 0$.

APPENDIX C

To prove Statement 6 we have to prove it for every kind of spinor and tensor charge separately.

1. We start with the scalar charge, for which the matter is easy. Since the commutator of a scalar charge with a field transforms in the same way as this field and since there are no numerical functions, transforming like a spinor, the commutator of scalar charge with a tensor field is a linear combination of tensor fields, and with a spinor field is a linear combination of spinor fields.

Let us consider

$$iQ\phi_{ex}^{(i)}(f)\Omega = \sum_{l=1}^{n_o} k^{(il)}\phi_{ex}^{(l)}(f)\Omega + \sum_{l=1}^{n_1} k_\mu^{(il)}\phi_{ex}^{(l)\mu}(f)\Omega + \text{etc.}$$

$$iQ\varphi_{ex,\mu}^{(i)}(f)\Omega = \sum_{l=1}^{n_o} q_\mu^{(il)}\varphi_{ex}^{(l)}(f)\Omega + \sum_{l=1}^{n_1} q_{\mu\nu}^{(il)}\varphi_{ex}^{(l)\nu}(f)\Omega + \text{etc.} \qquad \text{(C1)}$$

Q commutes with the Poincaré group, so it has to commute with the Casimir spin operator W^2. If we apply W^2 to both sides of (C1), take into account that we deal with a superposition of one-particle states of definite spin [eigenfunctions of W^2 with eigenvalues $J(J+1)$], it follows, by comparing the expression obtained with (C1), that

$$iQ\phi_{ex}^{(i)}(f)\Omega = \sum_{l=1}^{n_o} k^{(il)}\phi_{ex}^{(l)}(f)\Omega,$$
$$iQ\phi_{ex,\mu}^{(i)}(f)\Omega = \sum_{l=1}^{n_1} q^{(il)}\phi_{ex,\mu}^{(l)}(f)\Omega \text{ etc.} \qquad \text{(C2)}$$

The (ii) of Statement 6 does not need any comment. Similar reasoning can be applied in the case of spinorial fields.

2. Before we move on to examine the spinor charges we make the following obvious remark: the fields of integer and half-integer spins do not mix in the linear expression for a commutation of a charge with a field.

In the case of spinor charges the relations (3.1) read, e.g.,

$$i[Q_\alpha, \phi^{(i)}(x)] = \sum_{a=1}^{n_{1/2}} g_{\alpha\beta}^{(ia)}\psi_\beta^{(a)}(x),$$
$$i\{Q_\alpha, \overline{\psi}_\beta^{(a)}(x)\} = -\sum_{l=1}^{n_o}[g^{(la)}(i\gamma^\nu\partial_\nu + m)]_{\alpha\beta}\phi^{(l)}(x) + \sum_{t=1}^{n_1}[h^{(at)}(-i\gamma^\nu\partial_\nu + m)\gamma\mu]_{\alpha\beta}\phi_\mu^{(t)}(x), \qquad \text{(C3a)}$$

where

$$g^{(ia)} = \alpha^{(ia)} + \gamma^5\beta^{(ia)} = \overline{g^{(ia)}},$$
$$h^{(at)} = \alpha'^{(at)} + \gamma^5\beta'^{(at)} = \overline{h^{(at)}}, \qquad \text{(C3b)}$$

and where ϕ and ψ stands for ϕ_{ex} and ψ_{ex}. There are three more relations corresponding to ϕ_μ, ψ_μ, and $\phi_{\mu\nu}$.

To outline the proof and to avoid inessential complications, let us concentrate on a special case of (C3) when $\beta = \beta' = 0$. The matrix $\sum_a \alpha^{(ia)}\alpha^{(la)}$ is real and Hermitian and can be diagonalized to $g^{(i)2}\delta^{il}$[$g^{(i)}$ is real] by a real orthogonal transformation. We submit the fields ϕ as well as α to this transformation and denote the new quantities by $\hat{\phi}$ and $\hat{\alpha}$, respectively. We also introduce the auxiliary real free fields

$$\sum_a \hat{\alpha}^{(ia)}\psi^{(a)}(x) = g^{(i)}\hat{\psi}^{(i)}(x), \qquad \text{(C4)}$$

which satisfy the canonical anticommutation relations. In terms of $\hat{\Phi}$ and $\hat{\psi}$ (C3a) can be rewritten as[27]

$$i[Q, \hat{\phi}^{(i)}(x)] = g^{(i)}\hat{\psi}^{(i)}(x),$$
$$i\{Q, \overline{\hat{\psi}}^{(i)}(x)\} = -g^{(i)}(i\gamma^\nu\partial_\nu + m)\hat{\phi}^{(i)}(x) + \sum_t h^{(at)}(-i\gamma^\nu\partial_\nu + m)\gamma^\mu\phi_\mu^{(t)}(x). \qquad \text{(C5)}$$

Hereafter we shall drop the index i as well as the "*hat*" for notational convenience. Let us consider the matrix elements

$$[\overline{\tilde{\psi}}_{\alpha,\text{out}}^{+}(p^1)\overline{\tilde{\psi}}_{\beta,\text{out}}^{+}(p^2)\Omega, Q_\gamma\tilde{\phi}_{\text{in}}^{+}(p^3)\overline{\tilde{\psi}}_{\delta,\text{in}}^{+}(p^4)\Omega], \qquad \text{(C6a)}$$

and

$$(\tilde{\phi}_{\text{out}}^{+}(p^1)\tilde{\phi}_{\text{out}}^{+}(p^2)\Omega, Q_\gamma\tilde{\phi}_{\text{in}}^{+}(p^3)\overline{\tilde{\psi}}_{\delta,\text{in}}(p^4)\Omega). \qquad \text{(C6b)}$$

From (C6a) we get, assuming that g is different from zero,

$$k_{\gamma\alpha}^1 S_{o\beta;o\delta} - k_{\gamma\alpha}^2 S_{\alpha o;o\delta} = \gamma_{\gamma\rho}^o S_{\alpha\beta;\rho\delta} - (-\gamma^\nu p_\nu^4 + m)_{\gamma\delta} S_{\alpha\beta;oo}$$

$$+\sum_t \frac{h^{(t)}}{g} k^4_{\gamma\rho}\gamma^\mu_{\rho\delta} S'^{(t)}_{\alpha\beta;o\mu} + \text{(two further terms}$$

depending linearly on S′) $\equiv L^{(1)}_{\alpha\beta;\gamma\delta} \equiv L^{(1)}_{\alpha\beta}$, etc. (C7)

From (C6b) we get

$$S_{\gamma o;o\delta} + S_{o\gamma;o\delta} = S_{oo;\gamma\delta} - [\gamma^o(-\gamma^\nu \overset{4}{p}_\nu + m)]_{\gamma\delta} S_{oo;oo}$$
$$+ \text{(one term depending on } S') \equiv L^{(2)}_{\gamma;\delta} \equiv L^{(2)}_\gamma \text{ etc.}, \quad \text{(C8)}$$

with

$$S_{\alpha\beta;oo} \equiv [\tilde{\bar{\psi}}^+_{\alpha,\text{out}}(p^1)\tilde{\bar{\psi}}^+_{\beta,\text{out}}(\overset{2}{p})\Omega, \tilde{\phi}^+_{\text{in}}(\overset{3}{p})\tilde{\phi}^+_{\text{in}}(\overset{4}{p})\,\Omega],$$
$$S_{\alpha\beta;\gamma\delta} \equiv [\tilde{\bar{\psi}}^+_{\alpha,\text{out}}(\overset{1}{p})\tilde{\bar{\psi}}^+_{\beta,\text{out}}(\overset{2}{p})\,\Omega, \tilde{\bar{\psi}}^+_{\gamma,\text{in}}(\overset{3}{p})\tilde{\bar{\psi}}^+_{\delta,\text{in}}(\overset{4}{p})\,\Omega],$$
$$S'^{(t)}_{\alpha\beta;o\mu} \equiv [\tilde{\bar{\psi}}^+_{\alpha,\text{out}}(\overset{1}{p})\tilde{\bar{\psi}}^+_{\beta,\text{out}}(\overset{2}{p})\,\Omega, \tilde{\phi}^+(\overset{3}{p})\tilde{\phi}^{(t)+}_\mu(\overset{4}{p})\,\Omega], \quad \text{(C9a)}$$

$$S_{oo;oo} \equiv [\tilde{\phi}^+_{\text{out}}(\overset{1}{p})\tilde{\phi}^+_{\text{out}}(\overset{2}{p})\,\Omega, \tilde{\phi}^+_{\text{in}}(\overset{3}{p})\tilde{\phi}^+_{\text{in}}(\overset{4}{p})\,\Omega] \quad \text{etc.},$$
$$\overset{j}{k}_{\alpha\beta} \equiv (\gamma^\nu \overset{j}{p}_\nu + m)_{\alpha\beta} \equiv \overset{j}{k}_\beta. \quad \text{(C9b)}$$

Equations (C7) and (C8), viewed as a system of linear algebraic equations,

$$\begin{aligned} L^{(1)}_{\alpha\beta} &= \overset{1}{k}_{\alpha} S_{o\beta;o\delta} - \overset{2}{k}_\beta S_{\alpha o;o\delta}, \\ L^{(1)}_{\beta\alpha} &= \overset{1}{k}_\beta S_{o\alpha;o\delta} - \overset{2}{k}_\alpha S_{\beta o;o\delta}, \\ L^{(2)}_\alpha &= S_{o\alpha;o\delta} + S_{\alpha o;o\delta}, \\ L^{(2)}_\beta &= S_{o\beta;o\delta} + S_{\beta o;o\delta}, \end{aligned} \quad \text{(C10)}$$

yield the solution for $\overset{1}{p} \neq \overset{2}{p}$

$$D_{\alpha\beta} S_{\alpha o;o\delta} = L^{(1)}_{\alpha\beta}\overset{2}{k}_\alpha - L^{(1)}_{\beta\alpha}\overset{1}{k}_\alpha + [L^{(2)}_\alpha \overset{1}{k}_\beta - L^{(2)}_\beta \overset{2}{k}_\alpha]\overset{1}{k}_\alpha,$$
$$D_{\alpha\beta} S_{o\alpha;o\delta} = L^{(1)}_{\alpha\beta}\overset{2}{k}_\alpha + L^{(1)}_{\beta\alpha}\overset{1}{k}_\alpha - [L^{(2)}_\alpha \overset{2}{k}_\beta - L^{(2)}_\beta \overset{1}{k}_\alpha]\overset{2}{k}_\alpha,$$
$$D_{\alpha\beta} = \overset{1}{k}_\alpha \overset{1}{k}_\beta - \overset{2}{k}_\alpha \overset{2}{k}_\beta \text{(no summation over } \gamma). \quad \text{(C11)}$$

The solutions (C11) should not depend on the choice of β and γ, so we may choose $\alpha = \beta$, and we have

$$\begin{aligned} D_{\alpha\alpha} S_{\alpha o;o\delta} &= (\overset{2}{k}_\alpha - \overset{1}{k}_\alpha)(L^{(1)}_{\alpha\alpha} - L^{(2)}_\alpha \overset{1}{k}_\alpha), \\ D_{\alpha\alpha} S_{o\alpha;o\delta} &= (\overset{1}{k}_\alpha - \overset{2}{k}_\alpha)(L^{(1)}_{\alpha\alpha} + L^{(2)}_\alpha \overset{2}{k}_\alpha) \end{aligned} \quad \text{(C12)}$$

If we insert (C12) back into (C10) we get the relation

$$\frac{\overset{1}{k}_\alpha}{\overset{2}{k}_\beta + \overset{1}{k}_\beta}(L^{(1)}_{\beta\beta} + L^{(2)}_\beta \overset{2}{k}_\beta) + \frac{\overset{2}{k}_\beta}{\overset{2}{k}_\alpha + \overset{1}{k}_\alpha}(L^{(1)}_{\alpha\alpha} - L^{(2)}_\alpha \overset{1}{k}_\alpha) = L^{(1)}_{\alpha\beta}. \quad \text{(C13)}$$

Equation (B13) suggests that it should be possible to split $L^{(1)}_{\alpha\beta}/\overset{1}{k}_\alpha \overset{2}{k}_\beta$ into a sum of two terms each depending only on α or β, separately. For the case of free fields, when

$$S_{\alpha\beta;oo} = S'_{\alpha\beta;o\mu} = 0, \quad \text{etc.}, \quad \text{(C14)}$$

this can be accomplished because of the factorizability of the Wightman functions. In the case of interacting fields, however, even if (C14) holds, this is not possible since $S_{\alpha\beta;\gamma\delta}$ etc. do not factorize.

Thus $g = 0$.

In case (C14) does not hold we also have

$$h^{(it)} = 0 \quad \text{(C15)}$$

If, however, (C14) holds, we have to repeat our procedure for other choices of the S-matrix amplitude to get a contribution from the terms in (C5) involving $h^{(at)}$. One can show along the same lines, *mutatis mutandis*, that then (C15) follows.

This accomplishes the outline of the proof claiming that spinor charges do not exist in the case of interacting fields.

A similar proof can be presented in the case of spinor charges of higher rank.

3. We shall now explore the case of vector charges. In this case the relations (31) read:

$$i[Q_\mu, \phi^{(i)}(x)] = a^{(ik)}\partial_\mu \phi^{(k)}(x) + ib^{(im)}\phi^{(m)}_\mu(x), \quad \text{(C16a)}$$
$$i[Q_\mu, \phi^{(s)}_\nu(x)] = -ib^{(ks)}(g_{\mu\nu} + m^{-2}\partial_\mu\partial_\nu)\phi^{(k)}(x) + c^{(sm)}\partial_\mu \Phi^{(m)}_\nu + d^{(sm)}\epsilon_{\mu\nu\lambda\kappa}\partial^\kappa \phi^{(m)\lambda}(x) + \text{if}^{(sz)}\phi^{(z)}_{\mu\nu}, \quad \text{(C16b)}$$

$$i[Q_\mu, \phi^{(z)}_{\kappa\lambda}(x)] = -\text{if}^{(sz)}[(g_{\mu\kappa} + m^{-2}\partial_\mu\partial_\kappa)\phi^{(s)}_\lambda + (g_{\mu\lambda} + m^{-2}\partial_\mu\partial_\lambda)\phi^{(s)}_\kappa - \tfrac{2}{3}(g_{\kappa\lambda} + m^{-2}\partial_\kappa\partial_\lambda)\phi^{(s)}_\mu] + k^{(zy)}\partial_\mu \phi^{(y)}_{\kappa\lambda}(x). \quad \text{(C16c)}$$

To get (C16) we used Jacobi identities as well as properties of the fields. We also get the additional relations

$$i[Q_\mu, \partial^\mu \phi^{(s)}_\nu(x)] = -m^2 c^{(sm)}\phi^{(m)}_\nu(x), \quad \text{(C17a)}$$

$$i[Q_\mu, \partial^\mu \phi^{(z)}_{\kappa\lambda}(x)] = -m^2 k^{(zy)}\phi^{(y)}_{\kappa\lambda}(x), \quad \text{(C17b)}$$

as well as that $a^{(ik)}$, $c^{(sm)}$ and $k^{(zy)}$ are real Hermitian matrices, $d^{(sm)}$ is real but anti-Hermitian and $b^{(im)}$ and $f^{(sz)}$ are imaginary. We diagonalize the matrix $a^{(ik)}$ by a real orthogonal transformation. The standard examination of

$$[\tilde{\phi}^{(l_1)+}_{\text{out}}(p_1)\tilde{\phi}^{(l_2)+}_{\text{out}}(p_2)\Omega, Q_\mu \tilde{\phi}^{(l_3)+}_{\text{in}}(p_3)\tilde{\phi}^{(l_4)+}_{\text{in}}(p_4)\Omega],$$

taking into account the energy-momentum conservation,[2,3] yields

$$a^{(ik)} = \delta^{ik}a, \quad a = \bar{a}. \quad \text{(C18a)}$$

A similar procedure applied in the case of $c^{(sm)}$ and $k^{(zy)}$ yields

$$c^{(sm)} = \delta^{sm}c, \quad c = \bar{c}, \quad \text{(C18b)}$$

$$k^{(zy)} = \delta^{zy}k, \quad k = \bar{k}. \tag{C18c}$$

If the incoming and outgoing fields are properly normalized, it can be shown that

$$a = c = k. \tag{C18d}$$

We define

$$Q'_\mu = Q_\mu - aP_\mu, \tag{C19}$$

then (C16) reads as follows:

$$i[Q'_\mu, \phi^{(i)}(x)] = ib^{(im)}\phi_\mu^{(m)}(x), \tag{C20a}$$

$$i[Q'_\mu, \phi_\nu^{(s)}(x)] = -ib^{(ks)}(g_{\mu\nu} + m^{-2}\partial_\mu\partial_\nu)\phi^{(k)}(x) + d^{(sm)}\epsilon_{\mu\nu\lambda\kappa}\partial^\kappa\phi^{(m)\lambda}(x) + if^{(sz)}\phi_{\mu\nu}^{(z)}, \tag{C20b}$$

$$i[Q'_\mu, \phi_{\kappa\lambda}^{(z)}(x)] = -if^{(sz)}[(g_{\mu\kappa} + m^{-2}\partial_\mu\partial_\kappa)\phi_\lambda^{(s)} + (g_{\mu\lambda} + m^{-2}\partial_\mu\partial_\nu)\phi_\kappa^{(s)} - \tfrac{2}{3}(g_{\kappa\lambda} + m^{-2}\partial_\kappa\partial_\lambda)\phi_\mu^{(s)}]. \tag{C20c}$$

The terms containing $b^{(im)}$ interchange vector and scalar fields, the terms with $f^{(sz)}$ intertwine tensor and vector fields, while the correlation between the vector fields and the antisymmetric tensor fields

$$F_{\mu\nu}^{(m)} \equiv m^{-2}\epsilon_{\mu\nu\kappa\lambda}\partial^\kappa\phi^{(m)\lambda}, \quad \partial F_{\mu\nu}^{(m)} = 0$$

$$\hat{F}_{\mu\nu}^{(m)} = m^{-2}[\partial_\mu\phi_\nu^{(m)} - \partial_\nu\phi_\mu^{(m)}],$$

$$\partial\hat{F}_{\mu\nu}^{(m)} = -m^{-2}\phi_\nu^{(m)}$$

manifests itself in the terms containing $d^{(sm)}$. There is no interchange between scalar and tensor fields.

Our first task is to show that $b^{(im)}$ vanishes if there is an interaction present. For this purpose we cast the first two relations (C20) in the form

$$i[Q'_\mu, \hat{\phi}^{(m)}(x)] = b^{(m)}\hat{\phi}_\mu^{(m)}(x), \; m = 1, \cdots, n_1, \tag{C21a}$$

$$i[Q'_\mu, \hat{\phi}_\nu^{(m)}(x)] = -b^{(m)}(g_{\mu\nu} + \frac{1}{m^2}\partial_\mu\partial_\nu)\hat{\phi}^{(m)}(x) + \hat{d}^{(ml)}\epsilon_{\mu\nu\kappa\lambda}\partial^\lambda\hat{\phi}^{(l)k}(x) + i\hat{f}^{(mz)}\phi_{\mu\nu}^{(z)}(x), \tag{C21b}$$

with

$$-\sum_i \hat{b}^{(il)}\hat{b}^{(im)} = [b^{(m)}]^2\delta^{lm}, \quad b^{(m)} - real,$$
$$b^{(im)}S^{(ml)-1} \equiv \hat{b}^{(il)},$$
$$\phi_\mu^{(m)}S^{(ml)-1} \equiv \hat{\phi}_\mu^l,$$
$$\phi^{(i)}\hat{b}^{(im)} = \hat{\phi}^{(m)},$$
$$S^{(mr)}d^{(rs)}S^{(sl)-1} \equiv \hat{d}^{rs},$$
$$S^{(ms)}f^{(sz)} \equiv \hat{f}^{(mz)}. \tag{C21c}$$

Let us investigate more carefully the expression (we drop the index m as well as the "*hat*" in the following):

$$[\tilde{\phi}^+_{\kappa_{out}}(p_1)\tilde{\phi}^+_{\lambda_{out}}(p_2)\Omega, Q'_\nu\tilde{\phi}^+_{\mu_{in}}(p_3)\tilde{\phi}^+_{in}(p_4)\Omega], \tag{C22}$$

and

$$[\tilde{\phi}^+_{out}(p_1)\tilde{\phi}^+_{out}(p_2)\Omega, Q'_\nu\tilde{\phi}^+_{\mu_{in}}(p_3)\tilde{\phi}^+_{in}(p_4)\Omega]. \tag{C23}$$

We get from (C22) and (C23)

$$M^{(1)}_{\kappa\lambda\mu\nu} = \overset{1}{k}_{\nu\kappa}S_{o\lambda;\mu o} + \overset{2}{k}_{\nu\lambda}S_{\kappa o;\mu o}, \tag{C24}$$

$$M^{(2)}_{\mu\lambda} = S_{\lambda o;\mu o} + S_{o\lambda;\mu o}, \tag{C25}$$

with

$$ibM^{(1)}_{\kappa\lambda\mu\nu} = +ib\overset{3}{k}_{\mu\nu}S_{\kappa\lambda;oo} + d^{(l)}\epsilon^{\sigma\rho}_{\mu\nu}\overset{3}{p}_\rho S'^{(l)}_{\kappa\lambda;\sigma o} + ibS_{\kappa\lambda;\mu\nu} + f^{(z)}S''^{(z)}_{\kappa\lambda;\nu,\mu,o} + \text{further terms depending on } d \text{ and } f, \tag{C26a}$$

$$ibM^{(2)}_{\mu\nu} = ib\overset{3}{k}_{\mu\nu}S_{oo;oo} + ibS_{oo;\mu\nu} + d^{(l)}\epsilon^{\sigma\rho}_{\mu\nu}\overset{3}{p}_\rho S'^{(l)}_{oo;oo} + f^{(z)}S''^{(z)}_{\kappa\lambda;\mu\nu,o} \tag{C26b}$$

$$\overset{j}{k}_{\mu\nu} \equiv g_{\mu\nu} - m^{-2}\overset{j}{p}_\mu\overset{j}{p}_\nu \tag{C26c}$$

and

$$S_{\kappa\lambda;oo} \equiv [\tilde{\phi}^+_{\kappa_{out}}(p_1)\tilde{\phi}^+_{\lambda_{out}}(p_2)\Omega, \tilde{\phi}^+_{in}(p_3)\tilde{\phi}^+_{in}(p_4)\Omega],$$

$$S_{\nu o;\mu o} \equiv [\tilde{\phi}^+_{\nu_{out}}(p_1)\tilde{\phi}^+_{out}(p_2)\Omega, \tilde{\phi}^+_{\mu_{in}}(p_3)\tilde{\phi}^+_{in}(p_4)\Omega], \quad \text{etc.}$$

Now (C24) and (C25) for $\kappa = \lambda$ can be solved with respect to $S_{o\lambda;\mu o}$ and $S_{\lambda o;\mu o}$ for $\overset{1}{p}_\mu \neq \overset{2}{p}_\mu$. Since these quantities do not depend on the value of ν, we may put $\nu = \mu$. So we get

$$S_{o\lambda;\mu o} = \frac{1}{\overset{1}{k}_{\nu\lambda} - \overset{2}{k}_{\nu\lambda}}[M^{(1)}_{\lambda\lambda\mu\mu} - \overset{2}{k}_{\nu\lambda}M^{(2)}_{\mu\lambda}], \tag{C27a}$$

$$S_{\lambda o;\mu o} = \frac{1}{\overset{1}{k}_{\nu\lambda} - \overset{2}{k}_{\nu\lambda}}[\overset{1}{k}_{\nu\lambda}M^{(2)}_{\mu\lambda} - M^{(1)}_{\lambda\lambda\mu\mu}]. \tag{C27b}$$

(C25) is automatically satisfied by (C27); however, (C27) inserted into (C24) impose restrictions on the form of $M^{(1)}_{\kappa\lambda\mu\nu}$, *viz.*,

$$\frac{M^{(1)}_{\kappa\lambda\mu\nu}}{\overset{1}{k}_{\nu\kappa}\overset{2}{k}_{\nu\lambda}} = \frac{1}{\overset{2}{k}_{\nu\lambda}}S_{o\lambda;\mu o} + \frac{1}{\overset{1}{k}_{\nu\kappa}}S_{\kappa o;\mu o}. \tag{C28}$$

(C28) suggests that the lhs can be split into two terms each depending on κ and λ separately. The case $M^{(1)} = 0$ is obviously ruled out. A careful examination of $M^{(1)}$, taking into account is dependence on the momenta, leads to the conclusion that the term $S_{\kappa\lambda;\mu\nu}/\overset{1}{k}_{\nu\kappa}\overset{2}{k}_{\nu\lambda}$ should decompose into two κ- and λ-dependent parts, provided $b \neq 0$. But this is only true for the case of a free field. In fact, (C28) is satisfied in the absence of interaction. So in the case of interaction

$$b = 0. \tag{C29}$$

Similar considerations extended to (C20b) and (C20c) result in[29]

$$f^{(sz)} = 0. \tag{C30}$$

Eventually we are left with the relation

$$i[Q'_\mu, \phi^{(s)}(x)] = d^{(sm)}\epsilon_{\mu\nu\lambda\kappa}\partial^\kappa \phi^{(m)\lambda}(x). \tag{C31}$$

To show that $d^{(sm)}$ vanishes when interaction is present, we diagonalize $id^{(sm)}$ by a complex unitary transformation. After some manipulations we get

$$i[Q'_\mu, \overset{(0)}{\phi}{}^{(\alpha)}(x)] = 0, \tag{C32a}$$

$$i[Q'_\mu, \overset{(+)}{\phi}{}^{(\beta)}_\nu(x)] = d^{(\beta)}\epsilon_{\mu\nu\lambda\kappa}\partial^\kappa \overset{(-)}{\phi}{}^{(\beta)\lambda}, \tag{C32b}$$

$$i[Q'_\mu, \overset{(-)}{\phi}{}^{(\beta)}_\nu(x)] = -d^{(\beta)}\epsilon_{\mu\nu\lambda\kappa}\partial^\kappa \overset{(+)}{\phi}{}^{(\beta)\lambda}, \tag{C32c}$$

where

$$\alpha = 1 \cdots r_1, \quad \beta = 1 \cdots r_2,$$

$$r_1 + 2r_2 = n_1, \quad d \text{ is real},$$

and $\overset{(0)}{\phi}, \overset{(+)}{\phi}, \overset{(-)}{\phi}$, are real free vector fields.

To proceed further we assume that

$$[\tilde{\phi}^+_{out}(p_1)\tilde{\phi}^+(p_2)\Omega, \overset{(+)}{\tilde{\phi}}_\mu(p_3)\overset{(-)}{\tilde{\phi}}{}^+_\nu(p_4)\Omega] \neq 0.$$

Then the assertion follows easily by a proper choice of values of the momenta, (as pointed out in Appendix B).

Since the system of free fields, ϕ, ϕ_μ, and $\phi_{\mu\nu}$ forms an irreducible set as far as the boson fields are concerned, we find

$$Q_\mu = aP_\mu. \tag{C33}$$

A similar procedure applied to the case of fermion fields also leads to (C33). In this case the computations can be simplified considerably by making use of the fact that the free spinor fields satisfy the equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0.$$

The terms like $\gamma^5\psi$ are then separated immediately.

As soon as we know that there are no spinor charges as well as that (C33) holds, it is a relatively simple matter to show that Statement 6 holds generally, applying to the case of $Q_{\mu_1\cdots\mu_n}$, $n > 1$, methods similar to those presented above.

---

* Supported in part by A.E.C., under contract No. AT(30-1)-3668B.

† On leave of absence from the Institute of Theoretical Physics, University of Wroclaw, Wroclaw, Poland.

1 J. T. Lopuszanski, Commun. Math. Phys. 12, 80 (1969); 14, 158 (1969); *Springer Tracts in Modern Physics* (Springer-Verlag, Berlin, Heidelberg, 1970), Vol. 52, pp. 201-214.

2 J. T. Lopuszanski, Report No. E2-4832, 1969.

3 C. A. Orzalesi, J. Sucher, and C. H. Woo, Phys. Rev. Letters 21, 1550 (1968).

4 D. R. Divgi and C. H. Woo, University of Maryland Report No. 70-041, 1969.

5 L. J. Landau and F. H. Wichmann, J. Math. Phys. 11, 306 (1970).

6 L. J. Landau, Commun. Math. Phys. 17, 156 (1970).

7 C. A. Orzalesi, Rev. Mod. Phys. 42, 381 (1970).

8 C. H. Woo, University of Maryland Report No. 71-003, 1970.

9 L. Lopuszanski and J. Lukierski Rep. Math. Phys. (to be published).

10 A detailed report on the results achieved so far in this domain of interest is given in a review article of Orzalesi (Ref. 7).

11 D. Maison and H. Reeh, the Max-Planck Institute für Physik und Astrophysik, 1970.

12 Every complex scalar or vector field can be decomposed in a unique way into two real fields. For spinor fields, however, this can be achieved only if we use reducible representations of SL(2, C). E.g. for two-dimensional representation this can not be accomplished, for the four-dimensional reducible spinor representation we have the real Majorana 4-spinors.

13 Then $\gamma^5 = -(\gamma^5)^T$ is pure real, $\gamma^0\gamma^\mu\gamma^0 = +(\gamma^\mu)^+ = -(\gamma^\mu)^T$ and $(\gamma^0)^2 = 1$, $(\gamma^j)^2 = -1$, $(\gamma^5)^2 = -1$; we have also $\overline{\psi}(x) \equiv \gamma^0\psi(x)\ \theta\psi(x)\theta^{-1} = \pm\gamma^5\psi(-x)$ with θ denoting the TCP operator.

14 R. Haag, Phys. Rev. 112, 669 (1958); K. Hepp, Commun. Math. Phys. 1, 111 (1965).

15 D. Kastler, D. W. Robinson, and A. Swieca, Commun. Math. Phys. 2, 108 (1966); B. Schroer and J. Stichel, *ibid.*, 3, 258 (1966); A. Swieca, Nuovo Cimento 10 52, 242 (1967).

16 If $\omega_1, \cdots, \omega_{2k} (k = 0, 1, 2, \cdots)$ are vector indices, (2. 10) holds because of TCP invariance. Otherwise, it holds because of the transformation character of Q_ω under the Poincaré group.

17 H. Araki and R. Haag, Commun. Math. Phys. 4, 77 (1967).

18 H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento 1, 205 (1956); 6, 319 (1957); V. Glaser, H. Lehmann, and W. Zimmermann, *ibid.*, 6, 1122 (1957).

19 Notice that (3. 3) holds only because of our special choice of the underlying fields to be scalar. In case other kinds of fields are admitted, (3. 3) no longer need be true (for instance, if there are vector one-particle states in the theory).

20 J. Bros, H. Epstein, and V. Glaser, Nuovo Cimento 31, 1265 (1964); Commun. Math. Phys. 1, 240 (1965).

21 If we make use of the Wigner-Eckart theorem we may write $F_{ij}(k^2) = F(k^2)\epsilon_{ij}$. Notice also that for $k^2 \leq 0$, F_{ij} does not depend on "ex".

22 These properties can be easily derived from the hermiticity of $Q_{\mu\nu}$, as well as of $\phi^{(i)}(x)$, and from Jacobi identity $[\phi(Q, \phi)]$.

23 Notice that our statement 6 does not need to be true in the case of massless particles. Even if it holds true for exact symmetries of massless particles, it most likely breaks down in the case of a spontaneously broken symmetry. Of course, our statement does not apply to the case of broken symmetry either for massive or massless fields.

24 Two Hilbert spaces can always be mapped on each other.

25 Notice that the self-dual and anti-self-dual fields, corresponding to the (1, 0) and (0, 1) representations, constructed in a unique way from $F^{(0)}_{\mu\nu}$ and $\tilde{F}^{(0)}_{\mu\nu}$, cannot be real fields.

26 Notice that the free field creation and annihilation operators depend only on the absolute spin value $s \equiv j + j'$; what depends on the transformation characterized by (j, j') is the field.

27 If one insists on exploring the general expression (C3) with $\alpha, \alpha', \beta, \beta'$ different from zero, one should start with the relations

$$(1 + i\gamma^5)_{\alpha\beta} i[Q_\beta, \varphi^{(i)}(x)] = (1 + i\gamma^5)_{\alpha\beta} \sum_{a=1}^{n_{1/2}} z^{(ia)}\psi^{(a)}_\beta(x),$$

$$(1 + i\gamma^5)_{\alpha\beta} i\{Q_\beta, \overline{\psi}^{(a)}_\gamma(x)\} = -[(1 + i\gamma^5)(i\gamma^\nu\partial_\nu + m)]_{\alpha\gamma} \times \sum_{t=1}^{n_0} z^{(ta)}\varphi^{(t)}(x) + [(1 + i\gamma^5)(-i\gamma^\nu\partial_\nu + m)\gamma^\mu]_{\alpha\gamma} \sum_{t=1}^{n_1} z'^{(at)}\varphi^{(t)}_\mu(x),$$

where

$$z^{(ia)} = \alpha^{(ia)} - i\beta^{(ia)} \quad \text{etc.},$$

and the relations obtained from the former by replacing $1 + i\gamma^5$ by $1 - i\gamma^5$ and $z^{(ia)}$ by $z^{(ia)}$ etc. and transform them to a γ-representation in which γ^5 is diagonal. This leads, of course, to non-Hermitian charges as well as to non-Hermitian spinor fields.

28 Notice the quadratic dependence on momenta on the rhs of (C20c), which leads, in the case of interaction, to violation of the energy-momentum conservation law.

Computation of the Cluster Coefficients of the δ-Function Gases through the U Functions*

Sergio Servadio
Institute for Theoretical Physics, State University of New York, Stony Brook, New York 11790
(Received 26 June 1971)

It is shown that the wavefunctions of the δ-function l-body problem are simple enough to allow an explicit evaluation of the cluster operator U_l. Upon integration one can then obtain the general cluster coefficient b_l. The $l = 2$ case is explicitly solved for particles of any statistics.

INTRODUCTION

Since the first rigorous results by Lieb and Liniger,[1] considerable interest has been devoted to the study of the energy levels of N particles in one-dimension interacting via the two-body potential $2c\delta(x_i-x_j)$. The latest developments have led to the formulation of the thermodynamics both for Bose and Fermi statistics for any value of the coupling constant c. The basic assumption one makes is Bethe's ansatz for the wavefunction of the system; in the fermion case further assumptions about the counting of the states are needed. The ultimate justification of the whole procedure lies in the possibility of checking the results in the well-known (and trivial) limits of zero and infinite coupling constant and in the claim that by "continuity" the results should hold true also in between. Since such continuity arguments are never forcing, an independent and rigorous check would be quite welcome.

With this motivation in mind, we present a rigorous way of computing the second cluster coefficient of the δ-function systems for any value of the coupling constant c. Agreement is found with the results of Ref. 2. The method, making use of the U functions introduced by Lee and Yang,[3] can be applied to the evaluation of cluster integrals of any order.

1. REPULSIVE INTERACTION ($c > 0$)

A. Boltzmann Statistics

We recall[3] that the second virial coefficient b_2 is given by

$$b^2 = \tfrac{1}{2}\int_{-\infty}^{+\infty} d\xi\, \langle 0,\ \xi|U_2|0,\xi\rangle, \tag{1}$$

where

$$\langle 0,\xi|U_2|0,\xi\rangle \equiv \langle 0,\xi|W_2|0,\xi\rangle - \langle 0|W_1|0\rangle\langle\xi|W_1|\xi\rangle, \tag{2}$$

$$\langle 0,\xi|W_2|0,\xi\rangle \equiv \int_{-\infty}^{+\infty}\frac{dk_1}{2\pi}\int_{-\infty}^{+\infty}\frac{dk_2}{2\pi}\,|\Psi_{k_1,k_2}(0,\xi)|^2\exp[-\beta(k_1^2+k_2^2)], \tag{3}$$

$$\langle\xi|W_1|\xi\rangle \equiv \int_{-\infty}^{+\infty}\frac{dk_1}{2\pi}|\Psi_{k_1}(\xi)|^2 e^{-\beta k_1^2} = \frac{1}{\sqrt{4\pi\beta}} \equiv \frac{1}{\lambda}; \tag{4}$$

$\Psi_{k_1}(x)$ is the plane wave e^{ik_1x} and $\Psi_{k_1,k_2}(x_1,x_2)$ is the scattering solution of the two-body problem labelled by the incoming momenta.

One can easily prove that

$$\Psi_{k_1,k_2}(x_1,x_2) = \exp[i(k_1+k_2)(x_1+x_2)/2] \times \{e^{-i|k_2-k_1|(x_2-x_1)/2}[\theta(x_2-x_1) + T\theta(x_1-x_2)] + Re^{i|k_2-k_1|(x_2-x_1)/2}\theta(x_2-x_1)\},$$

with the reflection and transmission coefficients given by

$$R \equiv ic/(|k_2-k_1|-ic)$$

and

$$T \equiv |k_2-k_1|/(|k_2-k_1|-ic).$$

A straightforward computation shows that

$$\langle 0,\xi|U_2|0,\xi\rangle = \frac{1}{(2\pi)^2}\theta(\xi)\left[\frac{\pi}{\beta} + 2\sqrt{\frac{\pi}{\beta}}\int_0^\infty d\tau\, e^{-\beta\tau^2}\left(e^{-\sqrt{2}i\xi\tau}\frac{ic}{\sqrt{2}\tau - ic} + e^{\sqrt{2}i\epsilon\tau}\frac{-ic}{\sqrt{2}\tau+ic}\right) + 2\sqrt{\frac{\pi}{\beta}}\int_0^\infty d\tau e^{-\beta\tau^2}\frac{c^2}{2\tau^2+c^2}\right] + \frac{1}{(2\pi)^2}\theta(-\xi)2\sqrt{\frac{\pi}{\beta}}\int_0^\infty d\tau e^{-\beta\tau^2}\frac{2\tau^2}{2\tau^2+c^2} \tag{5}$$

and, upon integration (remembering $c > 0$),

$$b_2 = \frac{-1}{2\lambda}\left(\sqrt{\frac{2\beta}{\pi}}e^{\beta c^2/2}\int_0^{c/\sqrt{2}} dp\, e^{-\beta p^2} - \frac{1}{\sqrt{2}}(e^{\beta c^2/2}-1)\right).$$

B. Bose Statistics (Spin Zero)

By adding a superscript B to the various quantities, formulas (1)–(4) remain otherwise unchanged. However, now[1]

$$\Psi_{k_1,k_2}(x_2,x_1) = \Psi_{k_1,k_2}(x_1,x_2) = \alpha\exp[i(k_1x_1+k_2x_2)] + \beta\exp[i(k_2x_1+k_1x_2)], \qquad x_2 \geq x_1$$

where

$$\beta/\alpha = -[c - i(k_2-k_1)]/[c + i(k_2-k_1)]$$

and

$$|\alpha|^2 + |\beta|^2 = 1.$$

Indicating with the superscript (0) the ideal ($c \to 0$) limit, one finds

$$\langle 0,\xi|U_2^B|0,\xi\rangle - \langle 0,\xi|U_2^{B(0)}|0,\xi\rangle = \frac{-1}{(2\pi)^2}\frac{\pi}{\beta}2c\,e^{c|\xi|}\int_{|\xi|}^\infty d\tau\exp(-\tau^2/2\beta - c\tau) \tag{6}$$

from which follows

$$b_2^B - b_2^{B(0)} = \frac{-1}{\lambda}\left(\sqrt{\frac{2\beta}{\pi}}\,e^{\beta c^2/2}\int_0^{c/\sqrt{2}} dp\, e^{-\beta p^2} - \frac{1}{\sqrt{2}}(e^{\beta c^2/2}-1)\right)$$

with

$$b_2^{B(0)} = 1/2\sqrt{2}\lambda$$

C. Fermi Statistics

For fermions we should, depending on their spin, count the relative phase space of spacial symmetric and antisymmetric states and weight accordingly the sum over states of the two types. We shall instead simply compute the sum over antisymmetric states (labeled by the superscript A) or, equivalently, consider fermions of one species.

Using the formulas of Ref. 3, one has the relation

$$\langle 0,\xi|U_2^A|0,\xi\rangle = 2\langle 0,\xi|U_2|0,\xi\rangle - \langle 0,\xi|U_2^B|0,\xi\rangle,$$

which implies

$$b_2^A = 2b_2 - b_2^B;$$

remembering Eqs. (5) and (6), one sees that

$$b_2^A = -1/2\sqrt{2}\lambda = b_2^{A(0)}$$

[as expected since the $\delta(\xi)$ potential does not affect a wavefunction that vanishes at $\xi = 0$. Note also that $b_2^A = \lim_{c\to\infty} b_2^B$.]

D. Higher Coefficients

The calculations we have presented are, to our knowledge, the only nontrivial cases in which one has been able to compute explicitly the U_2 functions. It may be interesting to look into the higher U's. We present the results for Bose statistics.

By straightforward computation the U_3^B function is found to be

$$\begin{aligned}&\langle x_1,x_2,x_3|U_3^B|x_1,x_2,x_3\rangle \\ &\quad - \langle x_1,x_2,x_3|U_3^{B(0)}|x_1,x_2,x_3\rangle \\ &= \frac{1}{(2\pi)^3}\left(\frac{\pi}{\beta}\right)^{3/2} 8c^2 \exp[c(y_3-y_1)]\int_{y_3-y_2}^{\infty} d\xi_2 \\ &\quad \times \exp\left(-\frac{\xi_2^2}{4\beta} - c\xi_2\right)\int_{y_2-y_1}^{\infty} d\xi_1 \exp\left(\frac{\xi_1^2}{4\beta} - c\xi_1\right) \\ &\quad \times \exp\left(-\frac{(\xi_1+\xi_2)^2}{4\beta}\right), \end{aligned} \tag{7}$$

where (y_1, y_2, y_3) is the permutation of (x_1, x_2, x_3) such that $y_1 \le y_2 \le y_3$.

Looking at Eqs. (6) and (7), one can easily induce the functional form of the general U_l^B; less obvious is the pattern for the numerical coefficient for which one must go back to the details of the calculations. We then guess:

$$\begin{aligned}&\langle x_1,\cdots,x_l|U_l^B|x_1,\cdots,x_l\rangle - \langle x_1,\cdots,x_l|U_l^{B(0)}|x_1,\cdots,x_l\rangle \\ &= \frac{1}{(2\pi)^l}\left(\frac{\pi}{\beta}\right)^{l/2}(l-1)!(-2c)^{l-1}\exp[c(y_l-y_1)]\int_{y_l-y_{l-1}}^{\infty} d\xi_{l-1}\left(\exp\ -\frac{\xi_{l-1}^2}{4\beta} - c\xi_{l-1}\right) \\ &\quad \times \int_{y_{l-1}-y_{l-2}}^{\infty} d\xi_{l-2}\exp\left(-\frac{\xi_{l-2}^2}{4\beta} - c\xi_{l-2}\right)\int\cdots\int_{y_2-y_1}^{\infty} d\xi_1 \exp\left(-\frac{\xi_1^2}{4\beta} - c\xi_1\right)\exp\ -\left(\frac{(\xi_1+\cdots+\xi_l)^2}{4\beta}\right), \end{aligned}$$

where $(y_1,\cdots,y_l)$ is the permutation of $(x_1,\cdots,x_l)$ such that $y_1 \le y_2 \le \cdots \le y_l$. The general b_l^B can then be calculated by quadratures.

2. ATTRACTIVE INTERACTION ($c < 0$)

It is well known[4] that there exists only one N-body bound state with binding energy $E_0 = -c^2N(N^2-1)/12$ and that the corresponding wavefunction is symmetric. Besides, in N-body scattering, if one analyzes the asymptotic states according to channels of given number and type of bound states, the S matrix is known to be diagonal in the channel label. It is then quite natural to analyze the general U_l by identifying the contributions from the various channels.

If there are asymptotically no bound states, the evaluation of the sum over states runs completely parallel to that in the case of repulsive interaction. If, instead, there are bound states, modifications are needed[5] which amount to replacing some real and distinct pseudomomenta with clusters of complex pseudomomenta all with the same real part and distributed symmetrically around the real axis. The sum over states is then just an integration over all the possible real parts.

From this outline it is evident that one can, in principle, compute any U_l (with any statistics), although he would quite soon find it very tedious. We shall briefly present the results we have worked out.

A. Boltzmann Statistics

$$\langle 0,\xi|U_2|0,\xi\rangle = \frac{-c}{2\sqrt{2\pi\beta}}\, e^{\beta c^2/2}[e^{-c\xi}\theta(\xi) + e^{c\xi}\theta(-\xi)]$$

+ the whole right-hand side of (5),

$$b_2 = \frac{1}{\sqrt{2}\lambda}(e^{\beta c^2/2} - 1)$$

$$+ \frac{1}{2\lambda}\left(\sqrt{\frac{2\beta}{\pi}}\, e^{\beta c^2/2}\int_0^{-c/\sqrt{2}} dp\ e^{-\beta p^2} - \frac{1}{\sqrt{2}}(e^{\beta c^2/2} - 1)\right).$$

B. Bose Statistics (Spin Zero)

$$\langle 0,\xi|U_2^B|0,\xi\rangle - \langle 0,\xi|U_2^{B(0)}|0,\xi\rangle = \frac{1}{(2\pi)^2}\,\frac{\pi}{\beta}\,2c\, e^{c|\xi|}\int_{-\infty}^{|\xi|} d\tau\ \exp\left(-\frac{\tau^2}{2\beta} - c\tau\right),$$

$$b_2^B - b_2^{B(0)} = \frac{\sqrt{2}}{\lambda}(e^{\beta c^2/2} - 1) + \frac{1}{\lambda}\left(\sqrt{\frac{2\beta}{\pi}}\, e^{\beta c^2/2} \int_0^{-c/\sqrt{2}} dp\, e^{-\beta p^2} - \frac{1}{\sqrt{2}}(e^{\beta c^2/2} - 1)\right),$$

with

$$b_2^{B(0)} \equiv \lim_{c\to 0} b_2^B = 1/2\sqrt{2}\lambda.$$

C. Fermi Statistics (One-Species)

$$b_2^A = 2b_2 - b_2^S = -1/2\sqrt{2}\lambda.$$

D. Higher Coefficients

As in Sec. 1D, we only deal with Bose statistics:

$$\langle x_1,x_2,x_3|U_3^B|x_1,x_2,x_3\rangle - \langle x_1,x_2,x_3|U_3^{B(0)}|x_1,x_2,x_3\rangle = \frac{4c^2}{3!}\sqrt{\frac{3}{4\pi\beta}}\, e^{2\beta c^2} e^{2c(y_3-y_1)} + \frac{c^2}{\sqrt{2}\pi\beta}\, e^{\beta c^2/2} e^{c(y_3-y_1)}$$
$$\times \left\{\int_{-\infty}^{0} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{3c}{2}\xi\right) - \int_{-\infty}^{y_3-y_2} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right)\right.$$
$$- \int_{-\infty}^{y_2-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right) + \int_{-\infty}^{y_3-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right)$$
$$\left. - e^{c(y_3-y_1)} \int_{-\infty}^{y_3-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{3c}{2}\xi\right)\right\} + \frac{1}{(2\pi)^2}\left(\frac{\pi}{\beta}\right)^{3/2} 2!\,(2c)^2 e^{c(y_3-y_1)}$$
$$\times \int_{-\infty}^{y_3-y_2} d\xi_2 \left(\exp -\frac{\xi_2^2}{4\beta} - c\xi_2\right) \int_{-\infty}^{y_2-y_1} d\xi_1 \exp\left(-\frac{\xi_1^2}{4\beta} - c\xi_1\right) \exp\left(-\frac{(\xi_1+\xi_2)^2}{4\beta}\right),$$

where (y_1,y_2,y_3) is the permutation of (x_1,x_2,x_3) such that $y_1 \le y_2 \le y_3$.

The third virial coefficient b_3^B can then be computed by quadratures.

ACKNOWLEDGMENTS

The author would like to thank Professor C. N. Yang for suggesting this investigation, and his best friends of Piazza S. Caterina, in Pisa, for their warm hospitality and encouragement when the present work was undertaken.

* Work supported in part by U.S. Atomic Energy Commission Contract No. AT(30-1)3668B.

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); C. P. Yang, Phys. Rev. A 2, 154 (1970); C. K. Lai, to be published (1971).
3 T. D. Lee and C N. Yang, Phys. Rev. **113**, 1165 (1958).
4 J. B. McGuire, J. Math. Phys. 5, 622 (1964).
5 C. N. Yang, Phys. Rev. 168, 1920 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Classical Lorentz Invariant Particles

Richard Arens

Department of Mathematics, UCLA, Los Angeles, California 90024

A classical Lorentz invariant completely Hamiltonian elementary one-particle system is defined as having a state space K in which the Poincaré group acts transitively, its infinitesimal actions having generating functions relative to some Poisson bracket, such that there can be associated with each state k a world line $\Gamma(k)$ in Cartesian 4-space. It is determined that there are nine families of such particles. Two have their speed in the usual range, three travel at the speed of eight, and four always faster. In each family the members are distinguished by one or two parameters such as mass and spin.

1. INTRODUCTION

A one-particle system which is invariant (under the inhomogeneous Lorentz, or Poincaré, group $\mathcal{P}$) has a space of states K in which $\mathcal{P}$ acts. For an elementary particle we require this action to be transitive. The notion of one-particle system requires that it should be possible to infer from each state k a world line $\Gamma(k)$ in Cartesian 4-space in a coherent way.

We classify here all such systems which satisfy the condition of complete Hamiltonicity which means that there is on K a Poisson bracket $\{\ ,\ \}$ and ten functions $h_1,\ldots,h_{10}$ which are "generating" functions for the actions of the ten infinitesimal generators of $\mathcal{P}$.

The result is that there are nine families of such particles. Two types have their speed in the usual range, three travel at the speed 1, and four always

$$b_2^B - b_2^{B(0)} = \frac{\sqrt{2}}{\lambda}(e^{\beta c^2/2} - 1) + \frac{1}{\lambda}\left(\sqrt{\frac{2\beta}{\pi}}\, e^{\beta c^2/2} \int_0^{-c/\sqrt{2}} dp\, e^{-\beta p^2} - \frac{1}{\sqrt{2}}(e^{\beta c^2/2} - 1)\right),$$

with

$$b_2^{B(0)} \equiv \lim_{c\to 0} b_2^B = 1/2\sqrt{2}\lambda.$$

C. Fermi Statistics (One-Species)

$$b_2^A = 2b_2 - b_2^S = -1/2\sqrt{2}\lambda.$$

D. Higher Coefficients

As in Sec. 1D, we only deal with Bose statistics:

$$\langle x_1,x_2,x_3|U_3^B|x_1,x_2,x_3\rangle - \langle x_1,x_2,x_3|U_3^{B(0)}|x_1,x_2,x_3\rangle = \frac{4c^2}{3!}\sqrt{\frac{3}{4\pi\beta}}\, e^{2\beta c^2} e^{2c(y_3-y_1)} + \frac{c^2}{\sqrt{2}\pi\beta}\, e^{\beta c^2/2} e^{c(y_3-y_1)}$$
$$\times \left\{\int_{-\infty}^{0} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{3c}{2}\xi\right) - \int_{-\infty}^{y_3-y_2} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right)\right.$$
$$- \int_{-\infty}^{y_2-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right) + \int_{-\infty}^{y_3-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{c}{2}\xi\right)$$
$$\left. - e^{c(y_3-y_1)} \int_{-\infty}^{y_3-y_1} d\xi \exp\left(-\frac{3\xi^2}{8\beta} - \frac{3c}{2}\xi\right)\right\} + \frac{1}{(2\pi)^2}\left(\frac{\pi}{\beta}\right)^{3/2} 2!\,(2c)^2 e^{c(y_3-y_1)}$$
$$\times \int_{-\infty}^{y_3-y_2} d\xi_2 \left(\exp -\frac{\xi_2^2}{4\beta} - c\xi_2\right) \int_{-\infty}^{y_2-y_1} d\xi_1 \exp\left(-\frac{\xi_1^2}{4\beta} - c\xi_1\right) \exp\left(-\frac{(\xi_1+\xi_2)^2}{4\beta}\right),$$

where (y_1,y_2,y_3) is the permutation of (x_1,x_2,x_3) such that $y_1 \le y_2 \le y_3$.

The third virial coefficient b_3^B can then be computed by quadratures.

ACKNOWLEDGMENTS

The author would like to thank Professor C. N. Yang for suggesting this investigation, and his best friends of Piazza S. Caterina, in Pisa, for their warm hospitality and encouragement when the present work was undertaken.

* Work supported in part by U.S. Atomic Energy Commission Contract No. AT(30-1)3668B.

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); C. P. Yang, Phys. Rev. A 2, 154 (1970); C. K. Lai, to be published (1971).
3 T. D. Lee and C N. Yang, Phys. Rev. **113**, 1165 (1958).
4 J. B. McGuire, J. Math. Phys. 5, 622 (1964).
5 C. N. Yang, Phys. Rev. 168, 1920 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Classical Lorentz Invariant Particles

Richard Arens

Department of Mathematics, UCLA, Los Angeles, California 90024

A classical Lorentz invariant completely Hamiltonian elementary one-particle system is defined as having a state space K in which the Poincaré group acts transitively, its infinitesimal actions having generating functions relative to some Poisson bracket, such that there can be associated with each state k a world line $\Gamma(k)$ in Cartesian 4-space. It is determined that there are nine families of such particles. Two have their speed in the usual range, three travel at the speed of eight, and four always faster. In each family the members are distinguished by one or two parameters such as mass and spin.

1. INTRODUCTION

A one-particle system which is invariant (under the inhomogeneous Lorentz, or Poincaré, group $\mathcal{P}$) has a space of states K in which $\mathcal{P}$ acts. For an elementary particle we require this action to be transitive. The notion of one-particle system requires that it should be possible to infer from each state k a world line $\Gamma(k)$ in Cartesian 4-space in a coherent way.

We classify here all such systems which satisfy the condition of complete Hamiltonicity which means that there is on K a Poisson bracket $\{\ ,\ \}$ and ten functions $h_1,\ldots,h_{10}$ which are "generating" functions for the actions of the ten infinitesimal generators of $\mathcal{P}$.

The result is that there are nine families of such particles. Two types have their speed in the usual range, three travel at the speed 1, and four always

faster than that. In each family the members are distinguished by one or two real parameters such as mass and spin, or helicity.

The space of states K is in each case a coset space of $\mathcal{P}$ modulo a subgroup S which can have dimensions four or two but not zero.

For seven of the families, the subgroup S is the same for all members.

This study is suggested by (and indeed more elementary than) the corresponding wave-mechanical classification of Bargmann and Wigner.[1] It would be desirable to formulate a correspondence principle to relate the classifications.

An earlier paper of ours[2] dealt with the same problem as the present one with these differences: (1) The "one-particle" system is there (but not here) defined in a way forcing the speed to be less than 1; (2) the study was there limited to those systems for which K is *connected*.

When K is not connected, it can have two or four components which are permuted by space and time reflections.

For the familiar point particle, K is $\mathbb{R}^6$, the real six-dimensional space. The other particle with speeds less than 1 has $K = \mathbb{R}^6 \times S^2$, where S^2 is the familiar unit sphere. The quantum analog of this particle is the Dirac system.

Mathematically, in the terminology of Ref. 3, the content of the paper is as follows. A $\mathcal{P}$-invariant system is a $\mathcal{P}$-space. An elementary invariant system is a homogeneous $\mathcal{P}$-space. However, a completely Hamiltonian $\mathcal{P}$-invariant system need not be a $\mathcal{P}$-symplectic homogeneous space because in our concept the symplectic structure (or the Poisson Bracket) is not necessarily preserved by the elements of $\mathcal{P}$ not in the component G of the identity. If we drop down to G we lose the transitivity, but each of the connected components of the space are G-symplectic homogeneous spaces and hence by Ref. 3, 5.4.1, we have four (or less) covering spaces of orbits in $\mathfrak{g}'$ (the dual of the Lie algebra of G). These orbits have to be suitably related to each other.

The definition of our concept is now completed by giving a definition of *one-particle* system. For us this consists in requiring an equivariant map of the $\mathcal{P}$-symplectic homogeneous space into the $\mathcal{P}$-invariant system of straight lines in $\mathbb{R}^4$. The equivariance is not $\mathcal{P}$-equivariance but E-equivariance where E is the subgroup generated by the Euclidean motions and time translations. Our results consist essentially in an enumeration and description of the possible orbits in the dual of the Lie algebra which allow this equivariant map, and a determination of when it can be chosen $\mathcal{P}$-equivariant.

2. INVARIANT, ELEMENTARY, n-PARTICLE SYSTEMS

For dynamical systems we adopt the concepts and notation of Ref. 4, pp. 157–58. Thus a dynamical system has a state space K and for each two coordinators ("observers"?) x and y we have the map Δ_x^y of K onto K. Invariance is defined by asking that $\Delta_x^{\sigma \cdot x}$ depends only on the space–time group element σ. By an n-particle system we mean a system (K, Δ) where, for each coordinator x, there is a mapping

$$\Gamma_x : K \to \mathcal{C}^n, \tag{2.1}$$

where $\mathcal{C}^n$ is the class of n-tuples of curves in $\mathbb{R}^4$, and Γ_x has the property

$$\sigma^{-1} \circ \Gamma_x = \Gamma_x \circ \Delta^x_{\sigma \circ x} \tag{2.2}$$

for every σ in that subgroup E of the space time group in $\mathbb{R}^4$ which is generated by the time translations and the Euclidean motions. As our space–time group in this paper we take the Poincaré ("inhomogeneous Lorentz") group $\mathcal{P}$.

An example of an n-particle system is an *n-particle interaction* as presented in Ref. 4, p. 157. We use the notation employed there and Γ_x as $x \cdot (\Delta^x)^{-1}$. For k a state $(\Delta^x)^{-1}(k)$ is the set of curves in *space–time* for which k are the initial conditions relative to the coordinator x. We then use x to map those curves into $\mathbb{R}^4$. This is the meaning of $\Gamma_x(k)$. [In this example, $\Gamma_x(k)$ is the same for all x.] For the right side of (2.2), $k' = \Delta^x_{\sigma \circ x}(k)$ would be the transformed initial data and $\Gamma_x(k')$ would be the corresponding curves in $\mathbb{R}^4$. It is to be expected that they should be σ-related to $\Gamma_x(k)$. [A perfectly rigorous proof of (2.2) for this example begins by observing that $\Delta^{x \circ T} = \Delta^{x \circ T}$, replacing T by $x^{-1} \circ \sigma \circ x$, and using (2.42) of Ref. 4.]

In this example, for each x, the set of curves $\Gamma_x(k)$ completely determines k, that is to say, Γ_x is one-to-one. It is this one-to-one-ness which we wish to relax in the present paper. Furthermore, in this example, (2.2) holds for every σ in $\mathcal{P}$. We wish also to show that the weaker requirement [namely (2.2)] does allow some extra particles.

In an invariant system, the notation $U(\sigma)$ for $\Delta_x^{\sigma \cdot x}$ is justified, and $U(\sigma\tau) = U(\sigma)U(\tau)$. Thus $\mathcal{P}$ *acts* in K. If this action is transitive, the system is *elementary*.

For an invariant system, (2.2) says that $\sigma \cdot \Gamma_x = \Gamma_x \cdot U(\sigma)$ for all σ in E. In the case of an invariant system, K can be identified with the space $\mathcal{P}/S$ of cosets σS of some subgroup S of $\mathcal{P}$. The action of $\mathcal{P}$ in $\mathcal{P}/S$ is just such that τ sends σS into $\tau\sigma S$. There remains only the algebraic question of the existence of some map Γ satisfying (2.2). We note the following.

> Suppose S contains a one-parameter group of translations. Then $\Gamma(S)$ must be invariant under these translations, and, if Γ can be defined at all, the curves $\Gamma(\sigma S)$ are (straight) lines. If S contains a two-parameter group of translations, no Γ is possible. (2.3)

This is fairly obvious, for if τ is a translation, we must have $\tau\Gamma(S) = \Gamma(\tau S) = \Gamma(S)$ if τ is in S.

If S is contained in $E(1) \times O(3)$, the subgroup generated by time translations and reversal and the orthogonal group, then Γ can be defined. (2.4)

Indeed let $\Gamma(\sigma S)$ be the image under σ of the x^4 axis in $\mathbb{R}^4$ (the time axis). This Γ has property (2.2) for *all* σ, not merely in E. Such particles may be called *completely relativistic*. Their general theory is simple. Γ can be found precisely if there is a curve c in $\mathbb{R}^4$ left invariant by S.

Thus examples can be made in which the curves c are helices. [In the next section a condition will be imposed which implicitly, although not explicitly, forces the $\Gamma(\sigma S)$ to be straight lines.]

Some rather natural examples of totally relativistic invariant elementary one-particle systems can be obtained as follows. Let $\mathbf{T}, \mathbf{S}, \mathbf{L}$ be, respectively, the class of timelike, spacelike, lightlike lines in $\mathbb{R}^4$ relative to the Minkowski metric $(x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2$. By affixing a subscript 0 to any of these symbols we mean the lines of the class which go through the origin $(0,0,0,0)$. The group $\mathcal{P}$ acts in every one of these six spaces (when the 0 is affixed we use only the Lorentz part of the element of $\mathcal{P}$). Thus $\mathcal{P}$ also acts in $\mathbf{T} \times \mathbf{T}, \mathbf{T} \times \mathbf{S}_0$, etc. Thus we have here many invariant systems. It is not hard to see that it acts *transitively* in $\mathbf{T}_0 \times \mathbf{L}, \mathbf{T} \times \mathbf{L}_0, \mathbf{S}_0 \times \mathbf{L}, \mathbf{S} \times \mathbf{L}_0$. Hence these are elementary systems. Finally, in each of these four cases define $\Gamma(\alpha, \beta)$ to be either α or β according to whichever factor has not the suffix 0. This gives one-particle systems. The state spaces are eight dimensional.

The case $\mathbf{T} \times \mathbf{L}_0$ is the most interesting, and we mention the following properties:

$\mathbf{T} \times \mathbf{L}_0$ is equivalent to $T^1(\mathbb{R}^3)^0 \times S^2$. (2.5)

Here $T^1(\mathbb{R}^3)^0$ is the space of (bound) vectors in $\mathbb{R}^3$ of length less than 1 and S^2 is the 2-sphere (unit sphere in $\mathbb{R}^3$). Actually, $\mathbf{T}$ is equivalent to $T^1(\mathbb{R}^3)^0$ as follows. Take a line λ in $\mathbf{T}$. It hits the plane $\{x^4 = 0\}$ at $(a^1, a^2, a^3, 0)$. It has direction components $(b^1, b^2, b^3, 1)$, where $\sum |b^i|^2 < 1$. Also, $\mathbf{L}_0$ is equivalent to S^2. Take a line λ in L_0. It has direction components $(c^1, c^2, c^3, 1)$, where now $\sum |c^i|^2 = 1$.

Now there is a standard way in which the Euclidean group $E(3)$ acts in $T^1(\mathbb{R}^3)$ and in S^2.

The action of $E(3)$ in $T^1(\mathbb{R}^3)^0 \times S^2$ induced by the action of $\mathcal{P}$ in $\mathbf{T} \times \mathbf{L}_0$ via (2.5) is the standard one. (2.6)

There is a more or less standard way for $\mathcal{P}$ to act in $T^1(\mathbb{R}^3)^0$ (cf. Ref. 4, 2.7, with $A = 0$). In fact this action is derived by contemplating the preceding equivalence of T and $T^1(\mathbb{R}^3)^0$.

3. COMPLETELY HAMILTONIAN SYSTEMS

As in Ref. 5, we call an invariant system a *completely Hamiltonian* system if the infinitesimal dynamorphisms have generating functions relative to some Poisson bracket. This Poisson bracket is then necessarily preserved by the component of the identity of the space–time group, but not necessarily by all of that group. (For example, for a free particle, time inversion reverses the sign of all Poisson brackets.)

Differentiability is implicit in the concept of Hamiltonicity. Hence it is natural to require for elementary systems not only transitivity but also *local* transitivity (Ref. 5, 3.8).

We begin our investigations by first studying "systems" in which the group that acts is not $\mathcal{P}$ but the component G of the identity in $\mathcal{P}$. For reference, we call this concept a

G–invariant, completely Hamiltonian elementary one-particle system. (3.1)

Such a system is a

G–symplectic space (3.2)

in the sense of Ref. 3, 5.5 and satisfies the hypotheses of Ref. 3, Corollary, p. 192. Thus one could obtain at this point the conclusion of Ref. 3, 5.4.1 [which we do in fact establish in (3.7‴) below.] This would give us a Maurer–Cartan form that would have to be related to the Poisson bracket. The computations involved can easily be rearranged to give instead a proof of (3.7‴) based merely on Ref. 5.

A number of definitions and observations are needed before we can present even a partial version (Theorem 3.1 below) of the results. Let

$$M_{23}, M_{31}, M_{12}, M_{41}, M_{42}, M_{43}, M_1, M_2, M_3, M_4 \tag{3.3}$$

be the infinitesimal generators of the Poincaré group where the M_i are the infinitesimal *left* translations (usually called P_i; cf. Ref. 5, p. 134).

Suppose we had a (3.1). Then there would be generating functions

$$h_{23}, \ldots, h_{43}, h_1, \ldots, h_4 \tag{3.4}$$

for the action of the infinitesimal group elements (3.3).

These functions (3.4) are unique except for an additive constant of integration. However, in the circumstances of (3.1) they can be chosen to have the form (Ref. 5, proof of 6.2)

$$h_i = \langle \mu, M_i \rangle, \qquad h_{ij} = \langle \mu, M_{ij} \rangle \tag{3.4'}$$

where μ is a left-invariant 1-form on G [*N. B.*: the 3.3 are *right* invariant]. As a result, under left

translation by a Lorentz transformation on G,

the h_1, h_2, h_3, h_4 transform in the manner of the components of a vector, (3.4″)

the h_{ij} transform in the manner of the components of an alternating tensor, (3.4‴)

whence (Ref. 1, 46)

$w^i = \epsilon^{ijkl} h_j h_{kl}$ transform in the manner of a vector. (3.4⁗)

Moreover, (3.4″) and (3.4⁗) are invariant under translation so that (3.4″) and (3.4⁗) hold for any left translation on G. Thus (as in the quantum case) the two functions

$$h_4h_4 - h_1h_1 - h_2h_2 - h_3h_3 = g^{ij}h_ih_j = h^jh_j \quad (3.5)$$

and

$$w^4w^4 - w^1w^1 - w^2w^2 - w^3w^3 = g_{ij}w^iw^j = w_jw^j \quad (3.6)$$

are constant on $K' = G/S$.

We say that a vector (say 3.4″) is T, S, L, or 0 according to whether (3.5) is positive, negative, zero but that some h_i is not 0 or, finally, all the h_i are 0. Thus a given (3.2) gives rise to a pair of such letters. For example, a TS particle has (3.5) positive and (3.6) negative.

Theorem 3.1: A G-invariant, completely Hamiltonian elementary one-particle system is either TS, TO, ST, SS, SL, SO, LS, LL, or LO. The TS particles form a family indexed by two continuous parameters ("mass" and "spin"). The number of parameters for each of the other families is 1, 2, 2, 1, 1, 1, 1, 0, in the order of Theorem 3.1. The dimension of the state space is eight for TS, ST, SS, SL, LS, and six for TO, SO, LL, LO. Each TX particle can have any speed up to that of light. Each LX particle must have the speed of light. Each SX particle can have any speed exceeding that of light. Only LS and LL particles are not completely relativistic.

In stating this theorem we have presumed to make no distinction between any two *Hamiltonian-equivalent systems*, in the following sense. Let (K, Δ, p), $(\overline{K}, \overline{\Delta}, \overline{p})$ be completely Hamiltonian systems, p and $\overline{p}$ being the Poisson brackets. Then a map $T: K \to \overline{K}$ which makes $\overline{\Delta}^y_x \cdot T = \Delta^y_x$ and $\overline{p}(f, h) = p(f \cdot T, h \cdot T)$ is a Hamiltonian homomorphism, and, if T^{-1} is also a Hamiltonian homomorphism, then T is a Hamiltonian equivalence.

We have already remarked in Sec. 2 that for an invariant elementary system we can take K to be the coset space G/S, where S is a closed subgroup of G. We observed in Ref. 5, 3.7, how the 2-form β defining the alternating structure on K gives rise to closed (left) invariant 2-form α on G, which is of the form $d\mu$, where μ is a left-invariant 1-form (Ref. 5, 4.4).

We also observed that

the singular vectors at the identity, for $d\mu$, are precisely the vectors tangent to S there. (3.7)

We refer to this set of vectors as the Lie algebra of S. Another property of S is as follows:

The Maurer–Cartan form μ is invariant under right translation by elements of S. (3.7′)

Proof: From the way in which $d\mu$ is obtained on G from a 2-form on the space of cosets τS, it follows rather obviously that $d\mu$ is right invariant under S. Now we must prove that μ itself is thus invariant. Let μ' be a right translate of μ by a member of S, and let $\delta = \mu' - \mu$. Then $d\delta$ is certainly 0. Hence $\delta(Z) = 0$ for every commutator Z in $\mathfrak{g}$. But *the commutators span* $\mathfrak{g}$ (cf. Ref. 5, 4.6), so that $\delta = 0$, and $\mu' = \mu$. ($\mathfrak{g}$ is the Lie algebra of G.)

We now follow a suggestion of Simms to simplify the argument in Ref. 2 by explicitly observing the fact mentioned in Theorem 3.2 below. This theorem is precisely the conclusion of the relevant special case of Kostant's theorem (Ref. 3, 5.4.1). However, Simms[6] has kindly also suggested the following argument which establishes Theorem 3.2 on the basis of what we have done here up to this point.

S and $[\mu]$ have the same Lie algebra. (3.7″)

The proof requires recalling the adjoint representation ad of G in $\mathfrak{g}$. For τ in G, the transpose $\mathrm{ad}(\tau)^*$ acts in $\mathfrak{g}'$ and $\mathrm{ad}(\tau)^*\mu = \mu$ if and only if μ is preserved by the action of τ from the right. (For right-invariant μ, this *right* would be *left*.)

Therefore, consider an X in the Lie algebra of $[\mu]$. Then $\exp(tX) \in [\mu]$ and $\mathrm{ad}(\exp(tX))^*(\mu) = \mu$ and so (see Ref. 7, p. 118) $\exp(\mathrm{ad}tX)^*\mu = \mu$ for all t. Thus $(\mathrm{ad}X)^*\mu = 0$ and $(\mathrm{ad}X^*)(\mu)(Y) = 0$ for all Y. Therefore $\mu([X, Y]) = 0$ or $d\mu(X, Y) = 0$ for all Y. Using (3.7), we obtain (3.7″), and from this we obtain the result.

Theorem 3.2 (cf. Ref. 5, Theorem 5.4.1): G/S is a covering space of the orbit of μ in $\mathfrak{g}'$.

Here $\mathfrak{g}'$ is the dual linear space of $\mathfrak{g}$ and the "orbit" of μ is the set of images of μ under the action $\mu \to \mathrm{ad}(\tau^{-1})^*(\mu)$ in $\mathfrak{g}'$. Now S is, by (3.7″) and (3.7′), an open subgroup of $[\mu]$ whence we have the map from G/S to $G/[\mu]$ whence to the orbit $G\mu$ as asserted in Theorem 3.2.

One application of (3.7″) is that it enables one to assert that, for every μ in $\mathfrak{g}'$, the *orbit* of μ is a completely Hamiltonian system. To see this, one simply selects $S = [\mu]$ and forms $K = G/S$. As is well known, this is equivalent to the orbit. We must discover what the generating functions (3.4) are in this situation. To express the facts we must

observe that an element X of a linear space $\mathfrak{g}$ defines a (linear) function on $\mathfrak{g}'$ for which we use the same letter X.

Theorem 3.3: In the completely Hamiltonian system formed by the orbit of an arbitrary element μ of $\mathfrak{g}'$, the generating function for an infinitesimal Poincaré transformation X is precisely X itself regarded as a function on $\mathfrak{g}'$.

Proof: In Ref. 5, 5.3, the generating function was calculated as a function on the group G. We get from G to the orbit $G(\mu)$ by sending τ into $\mathrm{ad}(\tau^{-1})^*(\mu)$. Application of 5.3 of Ref. 5 to $\mu = A_i \mu^i$ and $X = \xi^j X_j$ (in the notation there used) tells us that the generating function for X on G is transformed precisely onto the function that X defines on g'.

4. A CRITERION FOR ONE-PARTICLE SYSTEMS

Theorem 4.1: Let (K, Δ) be a G-invariant, completely Hamiltonian system. Then it is a one-particle system if and only if the generating functions for translations (3.4″) are not all zero [see (3.1)].

Proof: If all four h_i are zero at some point, they are zero identically [(3.4″)]. Hence every point of K is invariant under all translations. Then (2.3) applies and shows that there is no Γ.

On the other hand, if they are never all zero, then the equations

$$h_4(k)x_i - h_i(k)x_4 = -h_{4i}(k), \qquad i = 1, 2, 3, \tag{4.1}$$

do define a line, which we will call $\Gamma(k)$. Here the x_i are not the Cartesian coordinates x^i of $\mathbb{R}^4$ but rather $x_i = g_{ij}x^j$ where the g_{ij} are defined *en passant* in (3.6).

Now perform an infinitesimal translation $\epsilon\partial/\partial x^k$. The perturbed form of (4.1) is then

$$h_4 x_i - h_i x_4 = -(h_{4i} - \epsilon g_{ki} h_4) \tag{4.2}$$

because[8] $\{h_k, h_{4i}\} = -g_{ki}h_4$. Suppose the point $(a^1, \ldots, a^4)$ lies on $\Gamma(k)$. The perturbed point has the contravariant coordinates $a^i + \epsilon\delta^i_k$ and the covariant coordinates $a_i + \epsilon g_{ik}$. These satisfy (4.2) if the a_i satisfy (4.1). Thus Γ commutes with translations.

We now consider rotations M_{jk}. Under such a map, the point with covariant coordinates a_i goes to $A_i = a_i + \epsilon(a_j g_{ki} - a_k g_{ji})$. The form of Eqs. (4.1) changes to

$$h_4 x_i - [h_i + \epsilon(g_{ik}h_j - g_{ij}h_k)]x_4 = -[h_{4i} + \epsilon(g_{ji}h_{k4} - g_{ki}h_{j4})].$$

We put $A_1, \ldots, A_4$ into this equation, noting that only $j, k < 4$ need be considered [see (2.2) again]. Thus $A_4 = a_4$, and the resulting equations are satisfied if the a_i satisfy (4.1). Thus Γ commutes with the rotations in the group E. Finally inspection shows that Γ commutes with space (and time) inversions. This establishes Theorem 4.1.

This mode of assigning lines in $\mathbb{R}^4$ to states of the system is generally not the only one possible. In fact, in some systems there is a more symmetric, in the space-time sense, formula possible. Let the value of (3.5) be denoted by m^2. When $m^2 \neq 0$, define the functions H^i by $H^i m^2 = h^i$ for $i = 1, 2, 3, 4$.

Theorem 4.2: Let (K, Δ) be as in Theorem 4.1 but assume also that $m^2 \neq 0$. Then this is a completely relativistic one-particle system.

Proof: The parametric equations $x_i = sh_i(k) - h_{ij}(k)H^j(k)$, $i = 1, 2, 3, 4$ and $-\infty < s < \infty$, define a line $\Gamma(k)$ for each state k. According to definition of "completely relativistic" we should show that Γ commutes with the action of the translations and also the entire Lorentz group. The method used in proving Theorem 4.1 does indeed work here too. However, the calculations are not very interesting and, since they are rather long, will be omitted.

It is a fact [proved in (4.9″) below] that for some one-particle systems, namely LS and LL, a completely relativistic world-line assignment is not possible. (To see this, we must continue our investigations until we find the corresponding subgroup S of G for which $K = G/S$.)

Since we have thus drawn attention to the non-uniqueness of the definition of $\Gamma(k)$, it seems appropriate to give a simple proposition showing just where the arbitrariness can, and cannot, lie.

> Assume the hypotheses of Theorem 4.1. When the h_i are not all 0, then $\Gamma(k)$ has to be a (straight) line and has to be parallel to (4.1). (4.3)

The reason is as follows. We know $\{g^{ij}h_i h_j, f\} = 0$ for all f [see (3.5)]. From this we see that $g^{ij}h_i \{h_j, f\} = 0$ and hence that $\{g^{ij}h_i(k)h_j, f\}$ is 0 at the point k of K. Thus k is unmoved by $g^{ij}h_i(k)M_j$ which generates a nonzero translation having the directions of the line (4.1). Thus $\Gamma(k)$ has to be invariant under such a translation.

We now return to the problem raised at the end of Sec. 3, namely to say which of the systems based on orbits in $\mathfrak{g}'$ give one-particle systems. To answer this [in (4.7), below], we need a little more notation. We have already selected as a basis for $\mathfrak{g}$ the vector fields (3.3) which are *right* invariant.

Now $\mathfrak{g}'$ is the dual linear space of $\mathfrak{g}$, but we will identify it with the *left*-invariant 1-forms on G. We take the *left* invariant rather than right invariant because the forms mentioned above, just before (3.7), are in fact left invariant [as they should be, because (3.3) are infinitesimal left multiplications.

Accordingly we choose as our basis for $\mathfrak{g}'$ the system

$$\mu^{23}, \mu^{31}, \mu^{12}, \mu^{41}, \mu^{42}, \mu^{43}, \mu^1, \mu^2, \mu^3, \mu^4 \quad (4.4)$$

such that the pairing of the i*th* element in (4.4) with the jth element of (3.3) gives δ_{ij} at the identity in G. By μ^{ij} not appearing in (4.4) we mean $-\mu^{ji}$

For the element

$$\mu = \sum_{i<j} A_{ij}\mu^{ij} + \sum_k A_k \mu^k, \quad (4.5)$$

we define $M_{ij}(\mu) = A_{ij}$ and $M_k(\mu) = A_k$. As already remarked in Sec. 3, these functions

$$M_{23}, M_{31}, \ldots, M_4 \quad (4.5')$$

are the generating functions for the similarly named infinitesimal actions (3.3). Thus we obtain the following result:

> The system associated with the orbit of (4.5) is a one-particle system if and only if the components $A_1, \ldots, A_4$ are not all 0. It is completely relativistic if $(A_4)^2 \neq (A_1)^2 + (A_3)^2$. (4.6)

Corollary 4.1: Systems (3.1) of the nine types listed in Theorem 3.1 and with the properties listed in Theorem 3.1 do exist.

To prove (4.6), we list nine Maurer–Cartan forms, expressed in the basis (4.4), where except in (4.7h) the coefficients are arbitrary real positive parameters. The two-letter symbols describe, in the notation of Theorem 3.1, the type of system which the orbit in g′ provides:

$$(TS) \quad \pm a\mu^4 + b\mu^{12}, \quad (4.7a)$$

$$(TO) \quad a\mu^4, \quad (4.7b)$$

$$(ST) \quad \pm a\mu^3 + b\mu^{12}, \quad (4.7c)$$

$$(SS) \quad a\mu^1 + b\mu^{43}, \quad (4.7d)$$

$$(SL) \quad a\mu^3 \pm \mu^{12} - \mu^{41}, \quad (4.7e)$$

$$(SO) \quad a\mu^3, \quad (4.7f)$$

$$(LS) \quad \mu^3 \pm \mu^4 + a\mu^{42}, \quad (4.7g)$$

$$(LL) \quad \mu^3 \pm \mu^4 + a\mu^{12} \quad (\text{where } a > 0 \text{ or } a < 0), \quad (4.7h)$$

$$(LO) \quad \mu^3 \pm \mu^4 \quad (4.7i)$$

Now these coefficients are the generating functions at one point of the orbit, namely (4.5) itself. Hence, for example, (4.7a) has m^2 [i.e., the value of (3.5)] equal to a^2, so that the T is justified. We leave it to the reader to calculate (3.4″″) and (3.6) in each case, and verify the propriety of each designation.

It is not relevant to the *existence*, but any (4.5), when the A_i are not all 0, lies in the orbit of one of (4.7a)–(4.7i) and, of course, in only one.

We now list bases for the Lie algebras of the subgroups of G which leave (4.7a)–(4.7i) (respectively) invariant. This will support the assertions about dimensions in Theorem 3.1:

$$(TS) \quad M_4, M_{12}, \quad (4.8a)$$

$$(TO) \quad M_4, M_{12}, M_{23}, M_{31}, \quad (4.8b)$$

$$(ST) \quad M_3, M_{12}, \quad (4.8c)$$

$$(SS) \quad M_1, M_{43}, \quad (4.8d)$$

$$(SL) \quad M_3, M_{12} \pm M_{41}, \quad (4.8e)$$

$$(SO) \quad M_3, M_{12}, M_{41}, M_{42}, \quad (4.8f)$$

$$(LS) \quad M_3 \mp M_4, M_{23} \mp M_{42} + 2aM_4, \quad (4.8g)$$

$$(LL) \quad M_3 \mp M_4, M_{12}, M_{23} \pm M_{42} - 2aM_1, \; M_{31} \mp M_{41} - 2aM_2, \quad (4.8h)$$

$$(LO) \quad M_3 \mp M_4, M_{12}, M_{23} \pm M_{42}, M_{31} \mp M_{41}. \quad (4.8i)$$

These can be obtained by the method of undetermined coefficients combined with these formula concerning g and g′:

$$\begin{aligned} (\mathrm{ad}\, M_i)^*(\mu^j) &= g_{ik}\mu^{kj}, \\ (\mathrm{ad}\, M_i)^*(\mu^{jk}) &= 0, \\ (\mathrm{ad}\, M_{ij})^*(\mu^k) &= (\delta_i^k g_{jm} - \delta_j^k g_{jm})\mu^m, \\ (\mathrm{ad}\, M_{ij})^*(\mu^{rs}) &= \delta_i^r \mu^{ks} g_{kj} - \delta_i^s \mu^{kr} g_{kj} \\ &\quad + \delta_j^s \mu^{kr} g_{ki} - \delta_j^r \mu^{ks} g_{ki}, \end{aligned} \quad (4.9)$$

which themselves follow from the elementary general fact that $(\mathrm{ad}\, X_i)^*(\mu^j) = c_{ik}^j \mu^k$.

Actually, we can assert more than is needed at this point.

> The group $[\mu]$ is generated by the Lie algebra given for the relevant case in (4.8). (4.9′)

The reason for this is that in each case the group $[\mu]$ is *connected*. The way to see this is to verify first that the subgroup S_0 of the proper orthochronous Lorentz group which leaves fixed the pair of vectors (3.4″) and (3.4″), is connected. (This requires some calculation for the last three.) Then we note that translations move an element μ of g′ all over the "fiber above" the corresponding pair of (3.4″), (3.4‴) values for that μ. The fiber is $\mathbb{R}^3$. Thus, if σ is in $[\mu]$, then $\sigma = \tau\lambda$, where $\lambda \in S_0$ and τ is a translation. We deform λ to 1 in S_0 and deform τ so that $\tau_t \lambda_t$ remains in $[\mu]$, showing the connectedness of the latter.

Finally, we are obliged to show that there is no G–invariant map of $G\mu$ into the lines in $\mathbf{R}^4$, in cases (4.7g), (4.7h). This is a consequence of the following, which is easily proved.

> Let $i = 1, 2, 3,$ or 4. There is no curve in $\mathbb{R}^4$ which is invariant under the group generated by $M_3 + M_4$ and $M_{23} + M_{24} + aM_i$ unless $a = 0$. (4.9″)

The case LO is completely relativistic, although

$m^2 = 0$, because the group (4.8i) is precisely the group leaving a certain line invariant.

5. THE ENUMERATION OF ELEMENTARY PARTICLE SYSTEMS

Theorem 5.1: Every [(3.1)] G-invariant, completely Hamiltonian one-particle system is of the form (4.6) with state space being an orbit of some element of $\mathfrak{g}$.

Proof: We have already mentioned [see (3.7)] how such a system gives rise to an element μ of $\mathfrak{g}'$. The problem is now just this: If μ is as in (4.5) and the A_i are not all 0, can one find a σ in G such that $(\mathrm{Ad}\,\sigma)^*(\mu)$ is in the list (4.7a)–(4.7i)? This is in fact true, and we sketch the proof. To begin with, we evaluate $\mathbf{P}(\mu)$ and $\mathbf{W}(\mu)$, the vectors (3.4″) and (3.4‴), and determine whether we have TS, $TO, \ldots, LO$. We then perform a *Lorentz* transformation, which reduces $\mathbf{P}$ and $\mathbf{W}$ to "normal form." We list these, for the nine cases, giving the contravariant components:

$$0, 0, 0, \pm a; 0, 0, \pm ab, 0, \tag{5.1a}$$

$$0, 0, 0, \pm a; 0, 0, 0, 0, \tag{5.1b}$$

$$0, 0, \pm ab, 0; 0, 0, 0, \pm a, \tag{5.1c}$$

$$a, 0, 0, 0; 0, b, 0, a, \tag{5.1d}$$

$$0, 0, -a, 0; 0, a, 0, \pm a, \tag{5.1e}$$

$$0, 0, -a, 0; 0, 0, 0, 0, \tag{5.1f}$$

$$0, 0, -1, \pm 1; a, 0, 0, 0, \tag{5.1g}$$

$$0, 0, -1; 0, 0, -a, \pm a, \tag{5.1h}$$

$$0, 0, -1, \pm 1; 0, 0, 0, 0. \tag{5.1i}$$

The knowledge of these components now fixes the coordinates A_i and A_{ij} of the Maurer–Cartan forms *except for three*. [These three degrees of freedom correspond to the fact that the fibers mentioned just below (4.9′) are like $\mathbb{R}^3$.] These three are in each case the ones that can be adjusted by *translations* (4.9). We assure the reader that (4.7a)–(4.7i) can be achieved.

The proof of (5.1) is completed by the final observation that S is in fact $[\mu]$ [see (3.7′), (3.7″), and (4.9′)]. Thus the specimens produced for proving (4.7) happen to give a complete list of all possible G-systems [(3.1)].

This result can be reformulated in an interesting way. We have already noted that $\mathfrak{g}'$ is a completely Hamiltonian system (Ref. 4, 5.1). By a *subsystem* (K', Δ') of a given system (K, Δ) we naturally mean that $K' \subset K$ and Δ' is the restriction of Δ to K'. *A completely Hamiltonian subsystem* (K', Δ') is one that is completely Hamiltonian, say with Poisson bracket $\{\ ,\ \}'$ such that if f and g are functions on K while f', g' are the restrictions to K', then $\{f', g'\}'$ is the restriction of $\{f, g\}$ to K'. In this sense, an open invariant submanifold of a completely Hamiltonian system is a completely Hamiltonian subsystem. In particular,

> The part $\mathfrak{m}$ of $\mathfrak{g}'$ where $(M_1)^2 + (M_2)^2 + (M_3)^2 + (M_4)^2$ is not zero is a completely Hamiltonian subsystem of $\mathfrak{g}'$. (5.2)

In fact, the uniform method 4.11 of defining $\Gamma(k)$ enables us to say that

> $\mathfrak{m}$ is an invariant one-particle system. (5.2′)

What it lacks is the property of "elementary." Now any invariant subset defines a subsystem, so that every invariant system decomposes into elementary systems. However, it is not clear that if the big system is completely Hamiltonian, then the subsystem is a completely Hamiltonian subsystem. But this is true for $\mathfrak{m}$. This is essentially what we have proved about G-systems.

Theorem 5.2: Every (3.1) is a completely Hamiltonian subsystem of the subsystem $\mathfrak{m}$ of the dual $\mathfrak{g}'$ of the Lie algebra of the Poincaré group.

We return now to (5.1) with the intention of describing the most general

> Poincaré invariant, completely Hamiltonian elementary one-particle system. (5.3)

Suppose K were the state space of a (5.3). Since G is a subgroup of $\mathcal{P}$, we obtain an action of G in K, but, of course, G need not act transitively. However, since $\mathcal{P}/G$ has only four elements, K will either be one, or two, or four open and closed subsets each of which is a single orbit under G.

Let us call these

> the G-constituents of the system (5.2). (5.4)

Each G-constituent will be of one of the types (4.7a)–(4.7i). However,

> if one G-constituent has a particular type, then the others have that same type. (5.5)

The reason for this is as follows. Let K_1 and K_2 be two constituents of K. We can pick k_1 in K_1 and k_2 in K_2 such that the subgroup S_1 (or S_2) of those elements of G that leave k_1 (or k_2, respectively) fixed is in the list (7.8a)–(4.8i). Now select σ in $\mathcal{P}$ such that $\sigma(k_1) = k_2$. Then $S_1 = \sigma^{-1} S_2 \sigma$. Reference to (4.8a)–(4.8i) shows that we must have $S_1 = S_2$, and so the G constituents have the same type. Of course, this does not mean that the parameter or parameters in the Maurer–Cartan form are necessarily (except in cases LS and LL) the same for K_1 and K_2. Thus

> the functions (3.5) and (3.6) are constant on each K_i but these constants may vary with i. However, the sign is constant on K, and if 3.43 is zero anywhere, it is O everywhere. Thus the (5.2) can be partially described by one of the nine symbols TS–LO. (5.6)

Thus a particle of type *TS* may have one, two, or four different masses, but two observers related by proper orthochronous Poincaré transformations will invariably see the same mass.

ACKNOWLEDGMENTS

This research was partially supported by the *NSF* Grant No. GP-18127. I wish to express my thanks to Professor J. D. Simms and Professor Robert Hermann for several valuable comments.

[1] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **34**, 211 (1948).
[2] R. Arens, Commun. Math. Phys. **21**, 139 (1971).
[3] B. Kostant, *Quantization and Unitary Representations*, Lecture Notes in Mathematics 170 (Springer, New York, 1970), pp. 87–208.
[4] R. Arens, Trans. Am. Math. Soc. **147**, 153 (1970).
[5] R. Arens, Commun. Math. Phys. **21**, 125 (1971).
[6] D. J. Simms, *Lie Groups and Quantum Mechanics IV*, Lecture Notes in Mathematics 52 (Springer, New York, 1968).
[7] S. Helgason, *Differential Geometry and Geometry Spaces* (Academic, New York, 1962).

[8] In greater detail, we assert that when the functions (3.4) are chosen in the manner of (3.4′), then their Poisson brackets are related in the same way as the Lie brackets of the corresponding infinitesimal transformations (3.3). This assertion is proved as follows. We examine the proof of 6.2 in Ref. 5 wherein vector fields Y_i have generating functions h_i. One sees there that $Y_m h_k = c^i_{km} h_i$. A little manipulation of 2.4 and 2.41 of Ref. 5 shows that (in the regular case at hand) $\{h_m, f\} = Y_m f$. Thus $\{h_m, h_k\} = c^i_{mk} h_i$. On the other hand, $[Y_m, Y_k] = -c^i_{mk} Y_i$. This proves our assertion.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Classical Moment Problem and the Calculation of Thermal Averages

A. Lonke
Department of Physics, University of the Negev, Beer-Sheva, Israel
(Received 22 March 1971)

The physical information contained in the first $2n$ moments of the single-particle spectral weight function of a fermionic many-body system is investigated. The approach is based on the mathematical theory of the classical moment problem. Under consideration are the thermal as well as dynamical properties of the system. Using this information, approximate n-pole single-particle thermal Green's functions and the corresponding spectral weight functions are constructed. It is shown that these approximations are not unique and depend on a real parameter. This dependence is used for the calculation of the rigorous error bounds of the approximate thermal averages.

1. INTRODUCTION

A time-dependent physical process is fully described by its spectral weight function (or spectral density, using a somewhat different terminology). Therefore, any possible information about this function is helpful in the understanding of the behavior of the physical system. An important source of such information are the moments of the spectral weight function (SWF). Moments of the SWF have been calculated for different systems in various areas of research as, for example, for the study of the NMR spectrum,[1] for neutron scattering,[2] for Raman light scattering,[3] for absorption of radiation,[4] etc. The advantage of these calculations dwells in the direct correspondence of their results to experimentally measurable quantities.

Suppose that a finite number of lowest-order SWF moments of a given system are known. It seems natural to inquire as to what information about the system can be derived from this knowledge. Consequently, since this information is, obviously, incomplete, a question arises concerning the error bounds of the approximate results. This question constitutes the object of the present paper.

We study here the physical information about a many-body system which might be obtained from a given number of lowest-order moments of its single-particle SWF. This concerns the dynamical properties of the system (states and energy spectrum) as well as statistical mechanics (thermal averages). Our approach is based on the theory of the classical (Hamburger) moment problem. This approach furnishes a method of establishing the rigorous upper and lower error bounds of the thermal averages in a given approximation.

The problem we are concerned with has been partially discussed by several authors. Harris and Lange[5] devised a moment technique for the study of single-particle excitations in metals with narrow energy bands. A method for the calculation of error bounds of high-temperature expansions in statistical mechanics has been given by Gordon,[6] who was first to apply the theoretical foundations of the moment problem to practical calculations. In Gordon's work use was made of the mathematical results of the theory of continued fractions, which is closely related to the classical moment problem. His calculation of the error bounds for the linear response of a system to a pulsed perturbation follows the same outlines.[7] These methods were applied by Wheeler and Gordon[8] for the calculation of the rigorous error bounds for the thermodynamical properties of harmonic solids.

Spectral weight functions of a variety of physical systems have one basic property in common: They are positive definite. We are referring to such systems in the present work. To be specific, we

Thus a particle of type *TS* may have one, two, or four different masses, but two observers related by proper orthochronous Poincaré transformations will invariably see the same mass.

ACKNOWLEDGMENTS

This research was partially supported by the *NSF* Grant No. GP-18127. I wish to express my thanks to Professor J. D. Simms and Professor Robert Hermann for several valuable comments.

[1] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **34**, 211 (1948).
[2] R. Arens, Commun. Math. Phys. **21**, 139 (1971).
[3] B. Kostant, *Quantization and Unitary Representations*, Lecture Notes in Mathematics 170 (Springer, New York, 1970), pp. 87–208.
[4] R. Arens, Trans. Am. Math. Soc. **147**, 153 (1970).
[5] R. Arens, Commun. Math. Phys. **21**, 125 (1971).
[6] D. J. Simms, *Lie Groups and Quantum Mechanics IV*, Lecture Notes in Mathematics 52 (Springer, New York, 1968).
[7] S. Helgason, *Differential Geometry and Geometry Spaces* (Academic, New York, 1962).

[8] In greater detail, we assert that when the functions (3.4) are chosen in the manner of (3.4′), then their Poisson brackets are related in the same way as the Lie brackets of the corresponding infinitesimal transformations (3.3). This assertion is proved as follows. We examine the proof of 6.2 in Ref. 5 wherein vector fields Y_i have generating functions h_i. One sees there that $Y_m h_k = c^i_{km} h_i$. A little manipulation of 2.4 and 2.41 of Ref. 5 shows that (in the regular case at hand) $\{h_m, f\} = Y_m f$. Thus $\{h_m, h_k\} = c^i_{mk} h_i$. On the other hand, $[Y_m, Y_k] = -c^i_{mk} Y_i$. This proves our assertion.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Classical Moment Problem and the Calculation of Thermal Averages

A. Lonke
Department of Physics, University of the Negev, Beer-Sheva, Israel
(Received 22 March 1971)

The physical information contained in the first $2n$ moments of the single-particle spectral weight function of a fermionic many-body system is investigated. The approach is based on the mathematical theory of the classical moment problem. Under consideration are the thermal as well as dynamical properties of the system. Using this information, approximate n-pole single-particle thermal Green's functions and the corresponding spectral weight functions are constructed. It is shown that these approximations are not unique and depend on a real parameter. This dependence is used for the calculation of the rigorous error bounds of the approximate thermal averages.

1. INTRODUCTION

A time-dependent physical process is fully described by its spectral weight function (or spectral density, using a somewhat different terminology). Therefore, any possible information about this function is helpful in the understanding of the behavior of the physical system. An important source of such information are the moments of the spectral weight function (SWF). Moments of the SWF have been calculated for different systems in various areas of research as, for example, for the study of the NMR spectrum,[1] for neutron scattering,[2] for Raman light scattering,[3] for absorption of radiation,[4] etc. The advantage of these calculations dwells in the direct correspondence of their results to experimentally measurable quantities.

Suppose that a finite number of lowest-order SWF moments of a given system are known. It seems natural to inquire as to what information about the system can be derived from this knowledge. Consequently, since this information is, obviously, incomplete, a question arises concerning the error bounds of the approximate results. This question constitutes the object of the present paper.

We study here the physical information about a many-body system which might be obtained from a given number of lowest-order moments of its single-particle SWF. This concerns the dynamical properties of the system (states and energy spectrum) as well as statistical mechanics (thermal averages). Our approach is based on the theory of the classical (Hamburger) moment problem. This approach furnishes a method of establishing the rigorous upper and lower error bounds of the thermal averages in a given approximation.

The problem we are concerned with has been partially discussed by several authors. Harris and Lange[5] devised a moment technique for the study of single-particle excitations in metals with narrow energy bands. A method for the calculation of error bounds of high-temperature expansions in statistical mechanics has been given by Gordon,[6] who was first to apply the theoretical foundations of the moment problem to practical calculations. In Gordon's work use was made of the mathematical results of the theory of continued fractions, which is closely related to the classical moment problem. His calculation of the error bounds for the linear response of a system to a pulsed perturbation follows the same outlines.[7] These methods were applied by Wheeler and Gordon[8] for the calculation of the rigorous error bounds for the thermodynamical properties of harmonic solids.

Spectral weight functions of a variety of physical systems have one basic property in common: They are positive definite. We are referring to such systems in the present work. To be specific, we

consider a system of fermions enclosed in a volume V, the dynamics of which is given by the Hamiltonian $\mathcal{H} = \mathcal{H}(c_\nu^+, c_\nu)$. The creation and annihilation operators c_ν^+ and c_ν satisfy the usual Fermi-Dirac anticommutation relations. The index ν may refer to a set of localized lattice sites (Wannier representation) or represent quantized moments, energy band indices, spin quantum numbers, etc.

The measurable statistical properties of the system can be conveniently described by means of the double-time Green's function formalism. Following Zubarev[9], we define for the system in thermodynamical equilibrium the retarded

$$G_\nu^{(r)}(t-t') = -i\theta(t-t')\langle[c_\nu(t), c_\nu^+(t')]_+\rangle \quad (1.1a)$$

and the advanced

$$G_\nu^{(a)}(t-t') = i\theta(t'-t)\langle[c_\nu(t), c_\nu^+(t')]_+\rangle \quad (1.1b)$$

single-particle Green's functions. Here $[\cdots]_+$ stands for the anticommutator, and the average is taken by means of the grand canonical ensemble

$$\langle\cdots\rangle = \mathrm{Tr}\{e^{\beta(\Omega - \mu N)}\cdots\}, \quad (1.2)$$

where Ω and μ are, respectively, the thermodynamical and chemical potentials and N is the number operator. The second quantization operators in (1.1a) and (1.1b) are taken in the Heisenberg representation

$$c_\nu^{(+)}(t) = e^{iHt}c_\nu^{(+)}e^{-iHt}, \quad (1.3)$$

with

$$H = \mathcal{H} - \mu N. \quad (1.4)$$

Finally, $\theta(t)$ is the usual step function

$$\theta(t) = \begin{cases} 1 \text{ for } t > 0, \\ 0 \text{ for } t < 0. \end{cases} \quad (1.5)$$

It is well known that the Fourier transform of the retarded or advanced Green's function

$$G_\nu^{(r,a)}(E) = (1/2\pi)\int_{-\infty}^{\infty} G_\nu^{(r,a)}(t)e^{iEt}dt \quad (1.6)$$

can be analytically continued into the complex energy plane E. As a result, we get an analytical function

$$G_\nu(E) = \begin{cases} G_\nu^{(r)}(E) & \text{for } \mathrm{Im}E > 0, \\ G_\nu^{(a)}(E) & \text{for } \mathrm{Im}E < 0, \end{cases} \quad (1.7)$$

consisting of two branches and defined in the whole complex E plane with a cut along the real axis. This function yields the spectral (Lehmann) representation

$$G_\nu(E) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\frac{\Lambda_\nu(\omega)}{E-\omega}, \quad \mathrm{Im}E \neq 0, \quad (1.8)$$

with the SWF

$$\Lambda_\nu(\omega) = i\{G_\nu(\omega + i\epsilon) - G_\nu(\omega - i\epsilon)\}$$

$$(\epsilon \to 0 \text{ and } \omega \text{ real}) \quad (1.9)$$

or

$$\Lambda_\nu(t) = \langle[c_\nu^+; c_\nu(t)]_+\rangle \quad (1.10)$$

in the time scale. (1.10) is usually referred to as the *characteristic function.*

In the following section we study in general terms the moments of the SWF, (1.9). We show that the determination of this function from a given sequence of moments represents a classical moment problem. In Sec. 3 we review some of the known mathematical results from the theory of the moment problem. These results are used for a geometrical interpretation of the thermal Green's function $G_\nu(E)$, (1.8). A detailed study of this geometrical picture of $G_\nu(E)$ on the complex G plane is given in Sec. 4. Using the information contained by a given number of SWF moments, we construct successive approximations to the SWF and $G_\nu(E)$ and give a geometrical interpretation of them. From this geometrical picture, general bounds on the function $G_\nu(E)$ may be established. The peculiar property of these approximations is their dependence on a real parameter. This implies that our procedure enables one to derive a continuum of functions, all of them equivalent in the framework of the given approximation.

Section 5 is devoted to the study of the dynamical aspects of our approximation scheme. We show that the consideration of physical systems leads to a restriction of the general mathematical problem by cases yielding unique solutions. Considering the dynamical information available from a finite number of lowest-order SWF moments, we introduce in this section a Hilbert space, the metrics of which is temperature dependent. In Sec. 6 the problem of the calculation of thermal averages is discussed. It is shown that the mentioned dependence of the approximate SWF on the numerical value of a real parameter enables one to establish the rigorous bounds of the error due to the incompleteness of the used physical information. A simple example is given in Sec. 7, where we discuss the problem of impurity scattering in metals. We close the paper by an appendix where it is shown that the Padé approximant of the thermal Green's function is a particular case of our approximation scheme.

2. THE MOMENTS OF THE SWF

In this section we will be concerned mainly with the general properties of the SWF defined by (1.9). Introducing a complete set of orthonormal eigenfunctions of the Hamiltonian (1.4), $H|n\rangle = E_n|n\rangle$, the SWF—the Fourier transform of the characteristic function (1.10)—is given by

$$\Lambda_\nu(\omega) = 2\pi f^{-1}(\omega)\sum_{m,n}\langle n|c_\nu^+|m\rangle\langle m|c_\nu|n\rangle \times e^{\beta(\Omega - E_n)}\delta(\omega + E_m - E_n), \quad (2.1)$$

where $f(\omega)$ stands for the Fermi function

$$f(\omega) = (e^{\beta\omega} + 1)^{-1}. \quad (2.2)$$

Thus, $\Lambda_\nu(\omega)$ is a real positive-definite function satisfying the normalization condition (the zeroth moment)

$$s_0^{(\nu)} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \Lambda_\nu(\omega) = 1. \tag{2.3}$$

Our posterior derivations, on the whole, are based on this small piece of information, namely, the positiveness of $\Lambda_\nu(\omega)$. Being absolutely general, it is found to be of great importance since it enables one to base the calculation of thermal Green's functions on the mathematical foundations of the classical moment problem.

Making use of the equation of motion of an operator in the Heisenberg representation, one easily derives the well-known formal expression for the nth moment of the SWF:

$$s_n^{(\nu)} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \omega^n \Lambda_\nu(\omega) = \langle [c_\nu^+; \mathbf{L}^n c_\nu]_+ \rangle, \qquad n = 0, 1, 2, \ldots, \tag{2.4}$$

where the operator $\mathbf{L}$ is given by

$$\mathbf{L} c_\nu = [c_\nu, H]_-. \tag{2.5}$$

The evaluation of the equal-time commutators involved in (2.4) is trivial (at least in principle). The result furnishes for $s_n^{(\nu)}$ an expression which is homogeneous (of order n) in the energy parameters of the particular system and which may involve thermal averages. In general, considering systems with binary interactions, these will be equal-time n-particle correlations. Thus, the sequence of Eqs. (2.4) represents a moment problem, the solution of which yields the SWF expressed in terms of given energy parameters and certain thermal averages. The latter, obviously, would have to be evaluated in a self-consistent manner.

Unfortunately, in practice this task can rarely be accomplished by a straightforward approach. It is possible, as it was proposed in a recent paper,[10] to take into account effectively the correlations up to the nth order by solving the first n-moment equations. To this end, it is sufficient to seek the approximate SWF in the linear manifold of n δ functions. For example, if n is even, then substituting

$$\Lambda_\nu(\omega) = \frac{1}{n} \sum_{i=1}^{n/2} [(1 - \alpha_i^{(\nu)})\delta(\omega + \gamma_i^{(\nu)}) + (1 + \alpha_i^{(\nu)})\delta(\omega - \gamma_i^{(\nu)})] \tag{2.6}$$

into the first n equations of the sequence (2.4), one derives the same number of algebraic relations for the parameters $\alpha_1^{(\nu)}, \ldots, \alpha_{n/2}^{(\nu)}, \gamma_1^{(\nu)}, \ldots, \gamma_{n/2}^{(\nu)}$. The thermal averages involved in these relations can be established self-consistently afterward by the consideration of higher-order SWF.[10]

This procedure, actually, implies the approximation of the Green function by a n-pole one. We will show, however, that this kind of solution is far from being unique. In other words, the same degree of approximation may be achieved by many other functions (in fact, a continuum) belonging to the same class. Moreover, among the functions of the same class, i.e., consisting of a weighted sum of n δ functions, there exist such that satisfy the sequence (2.4) up to the $(2n-2)$th moment equation, and under certain conditions even up to the order $2n-1$. Therefore, the fact that a self-consistent solution of the type (2.6), which was found in the described way, happens to satisfy the $(n+1)$th moment equation also, cannot be considered as a proof that the solution is exact and has to be taken *cum grano salis*.

Let us turn to the discussion of the general properties inherent in the moment problem (2.4). We start with the observation that the SWF may always be written in the form

$$\Lambda_\nu(\omega) = 2\pi \sum_{i=1}^{\infty} \mu_i^{(\nu)} \delta(\omega - \lambda_i^{(\nu)}), \tag{2.7}$$

where the coefficients $\mu_i^{(\nu)} \geq 0$ and satisfy the normalization condition

$$\sum_{i=1}^{\infty} \mu_i^{(\nu)} = 1. \tag{2.8}$$

From a pure mathematical point of view the representation of the SWF as an infinite weighted sum of δ functions follows from the fact that it has to satisfy a countable system of Eqs. (2.4). Indeed, we have seen that for any finite n the SWF may be chosen from the class of n-pole functions. Physically, (2.7) reflects the countability of the energy spectrum of a system enclosed in a finite volume. Finally, the nonnegativeness of the coefficients $\mu_i^{(\nu)}$ is obvious since $\Lambda_\nu(\omega)$ is a positive-definite function.

With (2.7) the moments of the SWF yield

$$s_n^{(\nu)} = \sum_{i=1}^{\infty} \mu_i^{(\nu)} (\lambda_i^{(\nu)})^n, \qquad n = 0, 1, 2, \ldots. \tag{2.9}$$

At this stage we would like to distinguish between the following two cases: (a) when the exact SWF of the particular problem under consideration belongs to a finite n-dimensional linear manifold of δ functions and (b) when this dimension is infinite. In the first case, only a finite number n of the coefficients $\mu_i^{(\nu)}$ in (2.9) are nonzero. This is the case of a *n-pole* SWF. In the second, there exists an infinite sequence of $\mu_i^{(\nu)} > 0$. Such a function will be referred to as a *continuous* SWF.

We consider now the sequence of the determinants

$$D_k = \begin{vmatrix} s_0 & s_1 \cdots s_k \\ s_1 & s_2 \cdots s_{k+1} \\ & \cdots \\ & \cdots \\ s_k & s_{k+1} \cdots s_{2k} \end{vmatrix}, \qquad k = 0, 1, 2, \ldots, \tag{2.10}$$

where for brevity we have neglected writing the quantum indices (ν). These determinants have the

following property: In the case of a n-pole SWF

$$D_k > 0 \quad \text{for} \quad k = 0, 1, 2, \ldots, n-1, \tag{2.11a}$$

$$D_k \equiv 0 \quad \text{for} \quad k \geq n. \tag{2.11b}$$

The validity of this statement can be easily proven. To this end we observe, first, that for the case under consideration we may write

$$D_{n-1} = \mu_{i_1}\mu_{i_2} \cdots \mu_{i_n}(d_n^{(i_1,i_2,\ldots,i_n)})^2, \tag{2.12}$$

where $d_n^{(i_1,i_2,\ldots,i_n)}$ stands for Vandermonde's determinant

$$d_n^{(i_1,i_2,\ldots,i_n)} = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_{i_1} & \lambda_{i_2} & \cdots & \lambda_{i_n} \\ \lambda_{i_1}^2 & \lambda_{i_2}^2 & \cdots & \lambda_{i_n}^2 \\ & \cdots & & \\ & \cdots & & \\ \lambda_{i_1}^{n-1} & \lambda_{i_2}^{n-1} & \cdots & \lambda_{i_n}^{n-1} \end{vmatrix} = \prod_{n \geq r > s \geq 1} (\lambda_{i_r} - \lambda_{i_s}). \tag{2.13}$$

Therefore, from the positiveness of the coefficients μ_i it follows that $D_{n-1} > 0$. Hence, since all the D_k for $k = 0, 1, \ldots, n-2$ are principal minors of a positive-definite determinant, they also have to be positive. This proves (2.11a). Finally, writing

$$D_k = \sum_{i_1, i_2, \ldots, i_{k+1}=1}^{n} \mu_{i_1}\mu_{i_2} \cdots \mu_{i_{k+1}} \lambda_{i_2} \times \lambda_{i_3}^2 \cdots \lambda_{i_{k+1}}^k d_{k+1}^{(i_1,i_2,\ldots,i_{k+1})}, \tag{2.14}$$

we note that for $k > n$ the determinants $d_{k+1}^{(i_1,i_2,\ldots,i_{k+1})}$ contain at least two identical columns, whence immediately follows (2.11b). This concludes the proof.

It becomes obvious now that the determinants (2.10) are all positive when a continuous SWF is considered. Thus, the moments of a continuous SWF, $\{s_n\}_0^\infty$, form a *positive sequence*.[11]

The problem (2.4) of the determination of a positive-definite function $\Lambda(\omega)$ from a given sequence $\{s_n\}_0^\infty$ is known as the classical moment problem. The positiveness of $\{s_n\}_0^\infty$, is a necessary and sufficient condition for the existence of a solution in the class of functions representing a continuous SWF.[12] Therefore, the fact that in this case all the determinants (2.10) are positive is an immediate result following from the assumption that the solution (2.4) does exist. This is plausible for physical reasons, taking into account the way in which our problem was posed. Thus, the proof presented above would appear to be irrelevant if only continuous SWF are considered. However, considering a given Hamiltonian the class of functions to which the SWF belongs is not known *a priori*. And then the relations (2.11) become important as a testing criterion. Suppose that for a particular physical model we have found that $D_n = 0$, whereas all D_k with $k < n$ are nonzero. This is sufficient to claim that the number of poles of the exact Green function is n.

3. THE GENERAL FORMALISM

The classical moment problem has been investigated in great detail. In this section we briefly review some of the well-known results. However, being interested mainly in the physical aspects of the theory, we will avoid reproducing lengthy proofs and mathematical derivations. For more detailed information concerning this subject the interested reader is referred to Refs. 12–14.

We start with the definition of a linear functional Φ which is based on the sequence $\{s_n\}_0^\infty$ and the domain of which is the space of all polynomials $\{R(E)\}$. To this end we relate the number

$$\Phi\{R(E)\} = p_0 s_0 + p_1 s_1 + \cdots + p_n s_n \tag{3.1}$$

to a given polynomial

$$R(E) = p_0 + p_1 E + \cdots + p_n E^n. \tag{3.2}$$

It can be shown that the positiveness of the sequence $\{s_n\}_0^\infty$ is a necessary and sufficient condition for the functional Φ to be positive. This implies that from $R(u) \geq 0$ $(-\infty < u < \infty)$ and $R(u) \not\equiv 0$ follows $\Phi(R) > 0$.[15]

In this section we consider the case of a continuous SWF exclusively. The appropriate generalizations for the n-pole case will be made below. Given a positive sequence $\{s_n\}_0^\infty$, it is possible to construct a sequence of polynomials $P_0(E), P_1(E), \ldots$, having the following properties:

(a) $P_n(E)$ is of order n and its highest coefficient is positive.

(b) The polynomials are orthonormal with regard to the sequence $\{s_n\}_0^\infty$, meaning by this

$$\Phi\{P_m(E) \cdot P_n(E)\} = \delta_{m,n}. \tag{3.3}$$

These polynomials, as can easily be checked, are given by

$$P_n(E) = (D_{n-1}D_n)^{-1/2} \begin{vmatrix} s_0 & s_1 \cdots & s_n \\ s_1 & s_2 \cdots & s_{n+1} \\ & \cdots & \\ s_{n-1} & s_n \cdots & s_{2n-1} \\ 1 & E \cdots & E^n \end{vmatrix}, \qquad n = 0, 1, 2, \ldots, \tag{3.4}$$

where $D_{-1} = 1$ by assumption.

The polynomials (3.4) represent one of the two linearly independent solutions of the finite difference equation

$$b_{k-1}y_{k-1} + a_k y_k + b_k y_{k+1} = E y_k, \qquad k = 1, 2, \ldots, \tag{3.5}$$

where

$$a_k = \Phi\{uP_k(u)P_k(u)\}, \quad b_k = (D_{k-1}D_{k+1})^{1/2}/D_k. \tag{3.6}$$

This solution is due to the initial conditions

$$P_0(E) = 1, \quad P_1(E) = (E - a_0)/b_0. \tag{3.7}$$

One may derive the second solution of Eq. (3.5) which is a sequence of polynomials $\{Q_n(E)\}_0^\infty$ by assuming the initial conditions

$$Q_0(E) = 0, \quad Q_1(E) = 1/b_0. \tag{3.8}$$

The polynomials $Q_n(E)$, being of order $n-1$, have the following representation:

$$Q_n(E) = \Phi_u\left\{\frac{P_n(E) - P_n(u)}{E - u}\right\}. \tag{3.9}$$

$Q_n(E)$ are usually referred to as the polynomials of of the *second kind*, whereas $P_n(E)$, as polynomials of the *first kind*.

Finally, it is convenient to introduce the quasi-orthogonal polynomials of order n, $P_n(E, \tau)$, which are given by

$$P_n(E, \tau) = P_n(E) - \tau P_{n-1}(E), \quad n = 1, 2, \ldots, \tag{3.10}$$

where τ is a parameter. It is easy to see that $P_n(E, \tau)$ satisfies the orthogonality conditions

$$\Phi_u\{P_n(u, \tau), u^k\} = 0 \quad \text{for } k = 0, 1, \ldots, n-2. \tag{3.11}$$

The polynomials which have been introduced above possess a number of peculiar properties. We mention here just these which are important to our present discussion:

(a) All the zeros of a real quasiorthogonal polynomial are real and simple.
(b) The zeros of the orthogonal polynomials $P_n(E)$ and $P_{n-1}(E)$ interlace.
(c) The zeros of the polynomial $Q_n(E)$ are real and simple and interlace with the zeros of $P_n(E)$.

Of particular interest to us will be the function

$$G_n(E, \tau) = Q_n(E, \tau)/P_n(E, \tau), \tag{3.12}$$

where the variable E is complex, τ is real ($-\infty < \tau < \infty$), and analogously with (3.10)

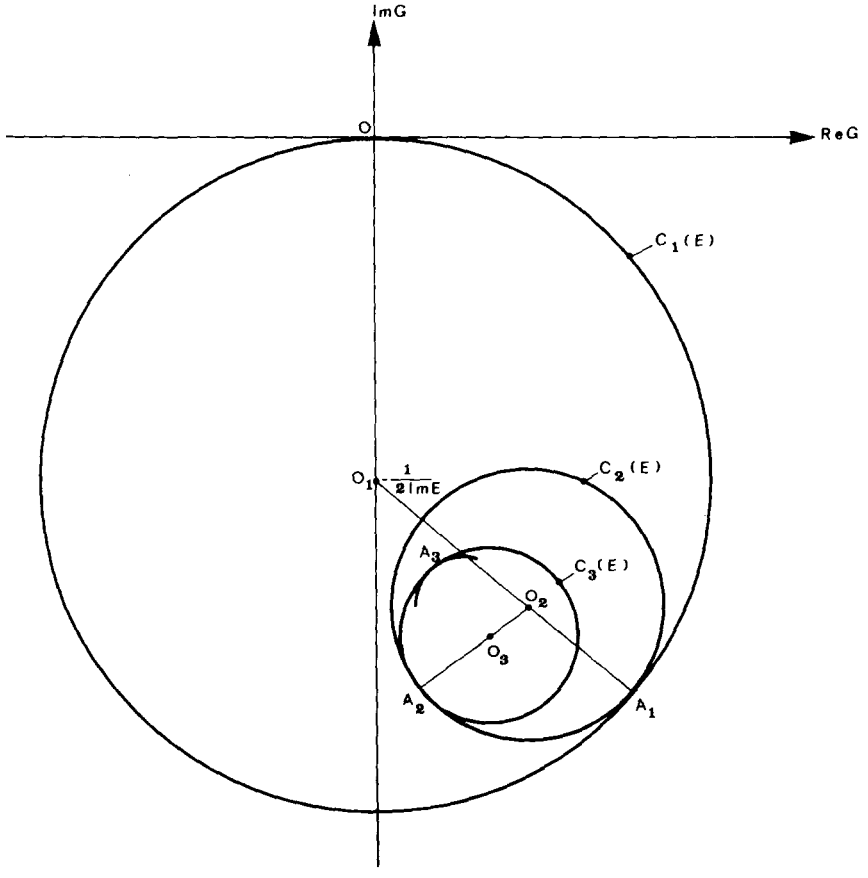

FIG. 1. An example of the first three circular contours $C_n(E)$ in the complex G plane. $C_n(E)$ represents the approximate n-pole thermal Green's functions $G_n(E, \tau)$ for a given E ($\mathrm{Im}E > 0$) and for all values of the real parameter τ ($-\infty < \tau < \infty$).

$$Q_n(E, \tau) = Q_n(E) - \tau Q_{n-1}(E). \quad (3.13)$$

There exists the following theorem:

Theorem[16]: Let E be fixed in the half-plane

$$\mathrm{Im}E > 0 \quad (\mathrm{Im}E < 0)$$

and let τ vary along the whole real axis. Then $g = G_n(E, \tau)$ describes a circular contour $\mathbf{C}_n(E)$ in the half-plane $\mathrm{Im}g < 0$ ($\mathrm{Im}g > 0$); the center of this circle is at the point

$$\frac{Q_n(E)\overline{P_{n-1}(E)} - Q_{n-1}(E)\overline{P_n(E)}}{P_n(E)\overline{P_{n-1}(E)} - P_{n-1}(E)\overline{P_n(E)}}, \quad (3.14)$$

and its radius is

$$\frac{1}{|E - \bar{E}|} \cdot \frac{1}{\sum_0^{n-1} |P_k(E)|^2}. \quad (3.15)$$

The equation of the circle $\mathbf{C}_n(E)$ may be written in the form

$$\frac{g - \bar{g}}{E - \bar{E}} + \sum_0^{n-1} |Q_k(E) - gP_k(E)|^2 = 0. \quad (3.16)$$

Furthermore, it can be shown that the circular area surrounded by the contour $\mathbf{C}_n(E)$ lies entirely within that of $\mathbf{C}_{n-1}(E)$ and the circumferences of the circles touch. The proof of the last statement is simple. We note first that the expression

$$\frac{g - \bar{g}}{E - \bar{E}} + \sum_0^{n-1} |Q_k(E) - gP_k(E)|^2 \quad (3.17)$$

is positive for lines which lie outside the circle $\mathbf{C}_n(E)$ and negative for lines lying inside. Indeed, since (3.17) is zero only on the contour $\mathbf{C}_n(E)$ it must, obviously, have different signs at both sides of it. Moreover,

$$\frac{g - \bar{g}}{E - \bar{E}} + \sum_0^{n-1} |Q_k(E) - gP_k(E)|^2 = A_n|g|^2 + Bg + \bar{B}\bar{g} + C, \quad (3.18)$$

where

$$A_n = \sum_0^{n-1} |P_k(E)|^2 > 0. \quad (3.19)$$

Therefore, (3.17) is certainly positive when $|g| \to \infty$. Thus, it is positive for all lines lying outside $\mathbf{C}_n(E)$ and negative for those inside. Now, dropping the last term from the sum in Eq. (3.16), we have for a point g lying on the contour $\mathbf{C}_n(E)$

$$\frac{g - \bar{g}}{E - \bar{E}} + \sum_0^{n-2} |Q_k(E) - gP_k(E)|^2 \leq 0, \quad (3.20)$$

which proves that the circle $\mathbf{C}_n(E)$ lies inside $\mathbf{C}_{n-1}(E)$. The equality sign in (3.20) is valid only for the common point of the two circles. Its existence follows straightforwardly from the definition (3.12), which yields

$$G_n(E, \infty) = G_{n-1}(E, 0). \quad (3.21)$$

As a result, we derive a geometrical picture which is schematically presented in Fig. 1 where the first three circles have been sketched. Equation (3.16) implies that the first circle $\mathbf{C}_1(E)$ is given by

$$(\mathrm{Re}g)^2 + \left(\mathrm{Im}g + \frac{1}{2\mathrm{Im}E}\right)^2 = \left(\frac{1}{2\mathrm{Im}E}\right)^2 \quad (3.22)$$

and does not depend on the moments of the SWF. Thus, $\mathbf{C}_1(E)$ is quite general and will be the same for an arbitrary physical model. The circles $\mathbf{C}_1(E)$ and $\mathbf{C}_2(E)$ touch at the point A_1,

$$G_1(E, 0) = 1/(E - s_1), \quad (3.23)$$

which is defined by the first moment of the SWF. The sequence of the circles, $\{\mathbf{C}_n(E)\}_0^\infty$, has to tend to some limit when n grows infinitely, since for the circular areas surrounded by these contours we have

$$\mathbf{C}_1(E) \subseteq \mathbf{C}_2(E) \subseteq \cdots \subseteq \mathbf{C}_n(E) \subseteq \cdots. \quad (3.24)$$

Let us adopt the notation

$$\mathbf{C}_\infty(E) = \lim_{n\to\infty} \mathbf{C}_n(E). \quad (3.25)$$

Generally speaking, $\mathbf{C}_\infty(E)$ may be either a *limiting point* or a *limiting circle*. In the former case we encounter the *determinate* moment problem (in the sense that its solution is unique); in the second, an infinite number of functions [all the points of the contour $\mathbf{C}_\infty(E)$] solve the problem. This makes it *indeterminate*.

From the physical point of view this poses an interesting question: Is it possible that the SWF moment problem for a certain dynamical system is indeterminate? This would happen when the radius of the limiting circle $\mathbf{C}_\infty(E)$, (3.15), would tend to a finite limit, or, in other words, when the series (3.19) converge for $n \to \infty$. We will return to this question in Sec. 5 where we show that the answer is unequivocally negative.

4. THE COMPLEX G PLANE

We proceed, concentrating our attention on the case of a continuous SWF. In the previous section, we introduced the functions $G_n(E, \tau)$ [see (3.12)] and described their geometrical behavior on the complex plane. We turn now to the question of what kind of physical information is carried by these functions.

Making use of Lagrange's interpolation formula, we write

$$G_n(E, \tau) = \frac{Q_n(E, \tau)}{P_n(E, \tau)} = \sum_{k=1}^{n} \frac{Q_n(\lambda_k^{(n)}, \tau)}{P_n'(\lambda_k^{(n)}, \tau)(E - \lambda_k^{(n)})}, \quad \mathrm{Im}\tau = 0, \quad (4.1)$$

where $P_n'(\lambda_k^{(n)}, \tau)$ stands for the derivative of the quasiorthonormal polynomial of order n at the point $\lambda_k^{(n)}$. By $\lambda_k^{(n)}$ $(k = 1, 2, \ldots, n)$, we denote the zeros of $P_n(E, \tau)$. As was mentioned, they are real, simple, and separated by the zeros of the polynomial $Q_n(E, \tau)$. We may, thus, write

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n, \quad [\lambda_k \equiv \lambda_k^{(n)}(\tau)]. \tag{4.2}$$

It can be easily shown that the coefficients

$$\mu_k \equiv \mu_k^{(n)}(\tau) = \frac{Q_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)} \tag{4.3}$$

have the following properties:

(i) They are real and positive.

(ii) An arbitrary polynomial $R(E)$ of degree $\leq 2n - 2$ satisfies the equality

$$\Phi_u\{R(u)\} = \sum_{k=1}^{n} \mu_k R(\lambda_k). \tag{4.4}$$

(iii) The relation holds:

$$\sum_{k=1}^{n} \mu_k^{(n)}(\tau) = s_0. \tag{4.5}$$

Indeed, property (i) follows from the fact that the coefficients (4.3) may be represented in the form[12]

$$\mu_k^{(n)}(\tau) = \left(\sum_{i=0}^{n-1} |P_i(\lambda_k^{(n)})|^2 \right)^{-1}. \tag{4.3'}$$

Hence, there exist polynomials $R_{n-2}(E)$ and $R_{n-1}(E)$ of degrees $\leq n - 2$ and $\leq n - 1$, respectively, such that

$$R_{2n-2}(E) = P_n(E, \tau) R_{n-2}(E) + R_{n-1}(E). \tag{4.6}$$

Now, by Lagrange's interpolation formula

$$R_{n-1}(E) = P_n(E, \tau) \sum_{k=1}^{n} \frac{R_{n-1}(\lambda_k)}{P_n'(\lambda_k, \tau)(E - \lambda_k)} \tag{4.7}$$

and using (4.6), we have

$$R_{n-1}(E) = P_n(E, \tau) \sum_{k=1}^{n} \frac{R_{2n-2}(\lambda_k)}{P_n'(\lambda_k, \tau)(E - \lambda_k)}, \tag{4.8}$$

since λ_k are zeros of the polynomial $P_n(E, \tau)$. Furthermore, from (4.6) and (3.11) we have

$$\Phi_u\{R_{2n-2}(u)\} = \Phi_u\{R_{n-1}(u)\}. \tag{4.9}$$

Thus,

$$\Phi_u\{R_{2n-2}(u)\} = \sum_{k=1}^{n} \frac{R_{2n-2}(\lambda_k)}{P'(\lambda_k, \tau)} \Phi_u\left\{\frac{P_n(u, \tau)}{u - \lambda_k}\right\}, \tag{4.10}$$

which completes the proof of (4.4) by virtue of (3.9). Finally, (4.5) follows straightforwardly from (4.4) under the assumption $R(E) \equiv 1$.

Returning to (4.1) let us first rewrite this relation by use of the notation (4.3):

$$G_n(E, \tau) = \sum_{k=1}^{n} \frac{\mu_k}{E - \lambda_k}. \tag{4.11}$$

A formal expansion of this expression yields the series

$$G_n(E, \tau) = \frac{1}{E} \sum_{k=1}^{n} \mu_k + \frac{1}{E^2} \sum_{k=1}^{n} \mu_k \lambda_k + \cdots + \frac{1}{E^{m+1}} \sum_{k=1}^{n} \mu_k \lambda_k^m + \cdots, \tag{4.12}$$

and using (4.4) we have

$$G_n(E, \tau) = \sum_{k=0}^{2n-2} \frac{s_k}{E^{k+1}} + \mathbf{O}\left(\frac{1}{E^{2n}}\right). \tag{4.13}$$

In the particular case of $\tau = 0$ the equality (4.4) holds for polynomials of degree $2n - 1$ also; therefore

$$G_n(E, 0) = \sum_{k=0}^{2n-1} \frac{s_k}{E^{k+1}} + \mathbf{O}\left(\frac{1}{E^{2n+1}}\right). \tag{4.14}$$

On the other hand, the analytical continuation of the Green's function, (1.8), is given in terms of the SWF moments (2.4) by the series

$$G(E) = \sum_{k=0}^{\infty} \frac{s_k}{E^{k+1}}. \tag{4.15}$$

The comparison of (4.14) and (4.15) illuminates the question concerning the information carried by the functions $G_n(E, \tau)$. It becomes clear that $G_n(E, \tau)$, with τ varying along the real axis $(-\infty < \tau < \infty)$, represent a continuum of functions which are the n-pole approximations to the exact Green's function $G(E)$. Being analytic, these functions have a Lehmann representation with the n-pole approximate SWF

$$\Lambda_n(\omega, \tau) = 2\pi \sum_{k=1}^{n} \mu_k^{(n)}(\tau) \delta(\omega - \lambda_k^{(n)}(\tau)), \tag{4.16}$$

which satisfies the first $2n - 2$ moment equations of the system (2.4). When $\tau = 0$, this is true for the $(2n - 1)$th moment equation also. Using (4.16), we may write (4.11) in the familiar form

$$G_n(E, \tau) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\Lambda_n(\omega, \tau)}{E - \omega}. \tag{4.17}$$

In the previous section we considered the geometrical picture of the successive n-pole approximations of the thermal Green's function. We recall that on the complex G plane, $G_n(E, \tau)$ is located on the circular contour $\mathbf{C}_n(E)$ which lies entirely in the lower (upper) half-plane if $\mathrm{Im}E > 0$ $(\mathrm{Im}E < 0)$. As was mentioned, the contour $\mathbf{C}_n(E)$ encircles all the contours of higher order while touching the circle $\mathbf{C}_{n-1}(E)$ (see Fig. 1).

Thus, the sequence $\{G_n(E, \tau)\}_0^{\infty}$ tends to a limit when $n \to \infty$. In case of a limiting point (a determinate moment problem) this limit will, obviously, be $G(E)$. The case of a limiting circle is irrelevant to our discussion. It will be shown in the following sec-

tion that the SWF moment problem for a well-defined physical system cannot be indeterminate.

Since the Green's function $G(E)$ will always be encircled by the contour $\mathbf{C}_1(E)$, we have

$$|\mathrm{Im}G(E)| \leq |1/\mathrm{Im}E|, \quad \mathrm{sgn}\ \mathrm{Im}G(E) = -\ \mathrm{sgn}\ \mathrm{Im}E \tag{4.18}$$

and

$$-\,|1/2\ \mathrm{Im}E| \leq \mathrm{Re}G(E) < |1/2\ \mathrm{Im}E|. \tag{4.19}$$

One may derive similar inequalities by considering circular contours of higher order. These increasingly will contain more information about $G(E)$. However, involving a successively growing number of moments s_n, they, obviously, depend on the dynamical properties of the system. Thus, (4.18) and (4.19) represent the most general restrictions on the real and imaginary parts of the analytically continued Green's function.

We turn now to the case of a finite-order multipole SWF. The problem we encounter here differs from the former one by the fact that the determinants D_k belonging to the given sequence $\{s_n\}_0^\infty$ vanish for $k \geq n$, n being the number of poles of $G(E)$. We recall that it is sufficient to show that $D_n \equiv 0$ to ensure the vanishing of all higher-order determinants. Intuitively, according to our geometrical picture, we expect here $\Lambda_n(E, 0)$ to solve the entire problem. In fact, the moments of the SWF depend on the energy parameters of the dynamical model. Varying these parameters (or introducing fictitious interactions), one might restore the positiveness of the sequence $\{s_n\}_0^\infty$. Let us assume that as a result of such a variation we obtain $D_n = \epsilon$, where ϵ is an arbitrary small positive quantity. Then due to the definition of the orthonormal polynomials (3.4), the radius of the circle $\mathbf{C}_{n+1}(E)$, (3.15), would also become arbitrarily small. Since the solution of the problem, $G(E)$, is enclosed by the contour $\mathbf{C}_{n+1}(E)$, it will tend with $\epsilon \to 0$ to the point where the circles $\mathbf{C}_{n+1}(E)$ and $\mathbf{C}_n(E)$ touch. And this is, obviously, $G_n(E, 0)$.

The problem, however, has to be treated in a more rigorous manner. Since $D_k > 0$ for $k = 1, 2, \ldots, n-1$, there exists a finite positive sequence $\{s_n\}_0^{2n-2}$. Thus, the functional (3.1) will remain positive if its domain has been restricted to the manifold of the polynomials of degree $\leq n-1$. Therefore, one may repeat all the previous derivations concerning the polynomials $P_k(E)$ and $Q_k(E)$ of degree $m \leq n-1$ and $m \leq n-2$, correspondingly. Hence, we introduce the polynomial

$$\tilde{P}_n(E) = \begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_n \\ s_1 & s_2 & s_3 & \cdots & s_{n+1} \\ & & \cdots & & \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-1} \\ 1 & E & E^2 & \cdots & E^n \end{vmatrix} . \tag{4.20}$$

The only difference between (4.20) and the orthonormal polynomial $P_n(E)$ given by (3.4) is the absence of the normalization factor $(D_{n-1}D_n)^{-1/2}$, which does not exist here since $D_n \equiv 0$. This polynomial, however, still satisfies the orthogonality conditions

$$\Phi\{\tilde{P}_n(E) \cdot E^k\} = 0, \quad k = 0, 1, \ldots, n-1. \tag{4.21}$$

The zeros of $\tilde{P}_n(E)$, $\lambda_k^{(n)}$ are, obviously, identical with those of $P_n(E)$. Using $\tilde{P}_n(E)$ instead of the non-existing orthonormal polynomial $P_n(E)$, we can construct the functions $G_n(E, \tau)$ which on the complex G plane will be represented by the circular contour $\mathbf{C}_n(E)$. The SWF belonging to $G_n(E, \tau)$ satisfies the first $2n-2$ moment equations (2.4) when τ is finite, and $2n-1$ equations when $\tau = 0$. Moreover, $\Lambda_n(E, 0)$ satisfies in this case the $2n$th moment equation as well, since the additional condition $D_n \equiv 0$ determines s_{2n} as a single-valued function of the moments $\{s_k\}_0^{2n-1}$.

It remains to show that $\Lambda_n(E, 0)$ solves the entire moment problem. To this end, let us recall first that the solution of our problem has to be sought in class of functions represented by (2.7) and consisting of n δ functions. This follows from the fact that $D_n \equiv 0$ and $D_k > 0$ for $k = 0, 1, \ldots, n-1$. Now, this class contains a subclass of functions leading to the same sequence $\{s_k\}_0^{2n-1}$. These will yield a continuum of n-pole functions $G_n(E, \tau)$, geometrically represented by the contour $\mathbf{C}_n(E)$. In other words, this contour (as well as all lower-order contours) will be in common for all physical systems, the differences in the dynamical behavior of which are not reflected in the values of the first $2n-1$ moments of the SWF. Thus, we have to seek the solution of our problem on the contour $\mathbf{C}_n(E)$.

Indeed, assume that this is not true. Then, obviously, the exact Green's function $G(E)$ had to be represented on the complex G plane by a point located inside $\mathbf{C}_n(E)$. In this case, one might construct on the basis of the given truncated sequence $\{s_k\}_0^{2n-1}$ a new moment problem where the contour $\mathbf{C}_{n+1}(E)$ would exist and would include the point $G(E)$. Since $\mathbf{C}_{n+1}(E)$ represents the class of functions consisting of $n+1$ poles, this would imply that for a certain $\tau(|\tau| < \infty)$ the coefficient $\mu_k^{(n+1)}(\tau)$ vanishes. This, however, is impossible as can easily be seen from the representation (4.3′).

Finally, since none of the functions represented by the contour $\mathbf{C}_n(E)$ but $G_n(E, 0)$ satisfies the moment equations for s_{2n-1} and s_{2n}, $\Lambda_n(E, 0)$ is the unique solution of the posed problem.

We summarize: *If for a given physical model the sequence $\{s_k\}_0^\infty$ yields $D_n \equiv 0$ and $D_k > 0$ for $k \leq n-1$, then the SWF (1.9) is given by*

$$\Lambda(\omega) = 2\pi \sum_{k=1}^{n} \mu_k^{(n)} \delta(\omega - \lambda_k^{(n)}), \tag{4.22}$$

where $\lambda_k^{(n)}$ are the zeros of the polynomial (4.20) and the coefficients $\mu_k^{(n)}$ are defined by (4.3′).

5. THE DYNAMICAL SPECIFICATIONS OF THE GENERAL PROBLEM

The fact that what is under consideration is a moment problem associated with the SWF of a dynamical system imposes certain specific features upon the problem itself. It can be shown, for instance, that in this case the problem is always a determinate one. We turn here to the discussion of those conclusions which follow from focusing our attention on physical systems, exclusively.

To this end let us introduce the Hilbert space **H** spanned by all various products of an odd number of one-particle fermionic creation and annihilation operators. The elements of **H** are, thus,

$$e_j = c^{(+)}_{\nu i_1} c^{(+)}_{\nu i_2} \cdots c^{(+)}_{\nu i_{2n+1}}, \qquad n = 0, 1, \ldots . \tag{5.1}$$

We define a scalar product by associating to every pair of elements e_i and e_j the number

$$(e_i, e_j) \equiv \langle [e_j^+, e_i]_+ \rangle. \tag{5.2}$$

The operator **L**, given by (2.5) is obviously defined on the whole space **H**. **L** is apparently symmetric:

$$(e_i, \mathbf{L}e_j) = (\mathbf{L}e_i, e_j), \tag{5.3}$$

however, not necessarily bounded.

Of particular interest to us will be the subspace $\mathbf{H}_\nu \subseteq \mathbf{H}$ generated by $c_\nu \in \mathbf{H}$ and represented by the elements

$$e_n^{(\nu)} \equiv \mathbf{L}^n c_\nu, \qquad n = 0, 1, \ldots . \tag{5.4}$$

The dimensionality of $\mathbf{H}_\nu$ *is given by the number of poles of the termal Green's function* $G_\nu(E)$, *(1.8).*

To prove this statement we observe first that

$$(e_n^{(\nu)}, e_m^{(\nu)}) \equiv (\mathbf{L}^n c_\nu, \mathbf{L}^m c_\nu) \equiv s_{m+n}^{(\nu)} \equiv \Phi^{(\nu)}(E^n \cdot E^m), \tag{5.5}$$

where $\Phi^{(\nu)}$ is the linear functional given by (3.1) and associated with the moment problem for the SWF (1.9). Using (5.5), it is easy to see that the elements

$$d_n^{(\nu)} \equiv P_n^{(\nu)}(\mathbf{L})c_\nu, \qquad n = 0, 1, \ldots, j \tag{5.6}$$

form an orthonormal set in $\mathbf{H}_\nu$. Here $P_n^{(\nu)}$ are the polynomials of the first kind given by (3.4) and their existence for $n > j$ has been assumed. As a matter of fact,

$$\begin{aligned}(d_n^{(\nu)}, d_m^{(\nu)}) &\equiv (P_n^{(\nu)}(\mathbf{L})c_\nu, P_m^{(\nu)}(\mathbf{L})c_\nu) \\ &\equiv \Phi^{(\nu)}(P_n^{(\nu)}(E) \cdot P_m^{(\nu)}(E)) \equiv \delta_{n,m}\end{aligned} \tag{5.7}$$

by virtue of (3.3) and the additiveness of the linear functionals.

Now, if the SWF under consideration is a continuous one, then the polynomials P_n exist for all n and (5.6) with $j \to \infty$ represent an infinite orthonormal basis in $\mathbf{H}_\nu$. Thus, in this case the dimensionality of $\mathbf{H}_\nu$ is infinite. On the other hand, in a case where the number n of poles of $G_\nu(E)$ is finite, then (5.6), with $j = n - 1$, represent an orthonormal set consisting of n elements. It remains to show that this set is complete in $\mathbf{H}_\nu$, i.e., that it is a basis in this subspace.

This becomes apparent by considering the polynomial $\tilde{P}_n$ which was introduced in the previous section by (4.20). Since in the n-pole case

$$(\tilde{P}_n^{(\nu)}(\mathbf{L})c_\nu, \tilde{P}_n^{(\nu)}(\mathbf{L})c_\nu) = D_n^{(\nu)} = 0, \tag{5.8}$$

it follows that

$$\tilde{P}_n^{(\nu)}(\mathbf{L})c_\nu = 0. \tag{5.9}$$

Therefore, $e_n^{(\nu)}$ may be represented in terms of the n orthonormal elements $d_i^{(\nu)} (i = 0, 1, \ldots, n-1)$. Consequently, the same will be true for every $e_m^{(\nu)}$ with $m \geq n$. Thus, the set (5.6) with $j = n - 1$ forms an orthonormal basis in $\mathbf{H}_\nu$, and its dimensionality is n. This completes the proof.

The dimensionality of $\mathbf{H}_\nu$ represents, obviously, the number of eigenstates of the dynamical system on which a single one-particle state might have been projected. The electronic state given by the set of quantum numbers $\{\nu\}$ is orthogonal to all the states contained by the subspace $\mathbf{H}'_\nu$, complementary to $\mathbf{H}_\nu$.

At this stage we would like to illuminate the question which was pointed out in Sec. 3: Does there exist a physical system for which the SWF moment equations would lead to an indeterminate moment problem? This question, obviously, arises exclusively in the case of a continuous SWF, since otherwise, as was formerly shown, the solution is unique.[17]

We consider the equation

$$\mathbf{L}v - Ev = 0, \qquad \mathrm{Im}E \neq 0. \tag{5.10}$$

If this equation has a nontrivial solution $v^{(\nu)} \in \mathbf{H}_\nu$, then expanding it in terms of the basis (5.6) one might write

$$v^{(\nu)} = \sum_{k=0}^{\infty} x_k d_k^{(\nu)}. \tag{5.11}$$

Here

$$x_k = (d_k^{(\nu)}, v^{(\nu)}) = (P_k^{(\nu)}(\mathbf{L})c_\nu, v^{(\nu)}) = (c_\nu, v^{(\nu)})P_k^{(\nu)}(E), \tag{5.12}$$

where in the derivation of the last equality we made use of (5.10). Thus,

$$v^{(\nu)} = x_0 \sum_{k=0}^{\infty} P_k^{(\nu)}(E) d_k^{(\nu)}. \tag{5.13}$$

Therefore, the condition for the existence of a nontrivial solution of Eq. (5.10) is the convergence of the series

$$A_\infty = \sum_{k=0}^{\infty} |P_k^{(\nu)}(E)|^2. \tag{5.14}$$

If (5.14) converges, then (5.13) represents the general solution of Eq. (5.10) and the indices of defect of the operator $\mathbf{L}$ are (1, 1). On the other hand, if (5.14) is a divergent series, then a non-trivial solution of Eq. (5.10) cannot be found and the defect indices of $\mathbf{L}$ are (0, 0), i.e., $\mathbf{L}$ is self-adjoint. We recall that the moment problem is indeterminate when the series (5.14) converges and determinate otherwise (see Sec. 3). Since the operator $\mathbf{L}$, (2.5), has to be by definition Hermitian in $\mathbf{H}_\nu$, we conclude: *A moment problem for the SWF of a dynamical system is always determinate.*

The subspace $\mathbf{H}_\nu$ is invariant with respect to the operator $\mathbf{L}$. Since all the zeros of a real quasi-orthogonal polynomial $P_n(E, \tau)$ are simple (see Sec. 3), the spectrum of the associated operator $\mathbf{L}$ has to be simple in $\mathbf{H}_\nu$. Thus: *The eigenelements of* $\mathbf{L}$ *in* $\mathbf{H}_\nu$ *are nondegenerate.*

In other words, representing a single-particle state in terms of the eigenstates of the dynamical system, one derives a linear combination of states referring to different total energies. Let us assume that the dimensionality of $\mathbf{H}_\nu$ is n. Then, beside (5.6) with $j = n - 1$, one might consider a different basis in $\mathbf{H}_\nu$ comprising the eigenelements $v_k^{(n)}$ of the self-adjoint operator $\mathbf{L}$. These, obviously, represent the subset of the eigenstates of the dynamical system contained by $\mathbf{H}_\nu$. Writing

$$c_\nu = \sum_{k=1}^{n} y_k^{(n)} v_k^{(n)}, \tag{5.15}$$

we have for the mth moment

$$s_m^{(\nu)} = (c_\nu, \mathbf{L}^m c_\nu) = \sum_{k=1}^{n} |y_k^{(n)}|^2 (\lambda_k^{(n)})^m, \tag{5.16}$$

where $\lambda_k^{(n)}$ is the eigenvalue corresponding to $v_k^{(n)}$:

$$\mathbf{L} v_k^{(n)} = \lambda_k^{(n)} v_k^{(n)}. \tag{5.17}$$

Recalling now the formal expression for the SWF, (4.22), the physical meaning of its constituents becomes clear. More specifically, the positive coefficient $\mu_k^{(n)}$, given by (4.3′), is nothing but the square absolute value of the projection of the single-particle state c_ν on the eigenstate $v_k^{(n)}$:

$$\mu_k^{(n)} = |y_k^{(n)}|^2 = |(v_k^{(n)}, c_k)|^2. \tag{5.18}$$

$\lambda_k^{(n)}$ is the energy corresponding to the state $v_k^{(n)}$. Thus, the spectrum of $\mathbf{L}$ in $\mathbf{H}_\nu$ is simple since the λ's are zeros of a quasiorthogonal polynomial. This remains true also for the case of a continuous SWF when $n \to \infty$.

The physical interpretation of the coefficients $\mu_k^{(m)}(\tau)$ and $\lambda_k^{(m)}(\tau)$ appearing in the expression (4.16) for an approximate SWF which satisfies the first $2m - 2$ SWF moment equations and does not satisfy the higher ones can be given in a similar way. To this end we consider the subspace $\mathbf{H}_m \subseteq \mathbf{H}_\nu \subseteq \mathbf{H}$ spanned by the first m elements of the basis (5.6). In this subspace we define an operator $\mathbf{L}_m(\tau)$, assuming

$$d_n^{(\nu)} \equiv P_n^{(\nu)}(\mathbf{L}) c_\nu = P_n^{(\nu)}[\mathbf{L}_m(\tau)] c_\nu, \qquad n = 0, 1, \ldots, m-1 \tag{5.19}$$

and satisfying the operator equation

$$P_m^{(\nu)}[\mathbf{L}_m(\tau)] c_\nu = \tau P_{m-1}^{(\nu)}[\mathbf{L}_m(\tau)] c_\nu, \tag{5.20}$$

where $-\infty < \tau \leq \infty$ is a real constant. Equations (5.19) and (5.20) define an operator on the linear manifold (5.4) with $n = 0, 1, \ldots, m-1$. The operator $\mathbf{L}_m(\tau)$ differs from $\mathbf{L}$ by the requirement that starting from the mth step, $\mathbf{L}_m(\tau)$ generates elements which are contained by $\mathbf{H}_m$, i.e.,

$$[\mathbf{L}_m(\tau)]^n c_\nu = e_n^{(\nu)} \quad \text{for } n = 0, 1, \ldots, m-1 \tag{5.21}$$

and

$$[\mathbf{L}_m(\tau)]^m c_\nu = \sum_{k=0}^{m-1} p_k(\tau) e_k^{(\nu)} \tag{5.22}$$

where $p_k(\tau)$ are certain τ-dependent coefficients. It is easy to see that (5.19) and (5.20) furnish a complete definition of $\mathbf{L}_m(\tau)$ and this operator is self-adjoint in $\mathbf{H}_m$. Denoting by $v_k^{(m)}(\tau)$ the eigenelements of $\mathbf{L}_m(\tau)$, we have

$$c_\nu = \sum_{k=0}^{m-1} z_k^{(m)} v_k^{(m)}(\tau), \qquad z_k^{(m)} = (v_k^{(m)}(\tau), c_\nu). \tag{5.23}$$

Hence, substituting this expansion into (5.20), we derive

$$\sum_{k=0}^{m-1} z_k^{(m)} P_m^{(\nu)}[\lambda_k^{(m)}(\tau), \tau] v_k^{(m)}(\tau) = 0, \tag{5.24}$$

where $\lambda_k^{(m)}(\tau)$ is the eigenvalue corresponding to $v_k^{(m)}(\tau)$:

$$\mathbf{L}_m(\tau) v_k^{(m)}(\tau) = \lambda_k^{(m)}(\tau) v_k^{(m)}(\tau).$$

It follows, therefore, that the $\lambda_k^{(m)}(\tau)$ are the zeros of the quasiorthogonal polynomial $P_m(E, \tau)$ given by (3.10). Similarly, the $\mu_k^{(m)}(\tau)$ represent the square absolute value of the projection of the single-particle state c_ν onto the eigenstates of the operator $\mathbf{L}_m(\tau)$. Thus, the calculation of the thermal Green function in terms of the approximate n-pole SWF implies the fact that the exact eigenstates of the physical system have been approximated by those of the operator $\mathbf{L}_m(\tau)$.

Significant to this approximation scheme is the speed of convergence of the eigenvalues and eigenelements of $\mathbf{L}_m(\tau)$ to those of the operator $\mathbf{L}$. Suppose that $\mathbf{L}$ is completely continuous and, therefore, bounded. Then the eigenvalues of $\mathbf{L}$ in $\mathbf{H}_\nu$ could be arranged in numerically decreasing order starting from $\lambda_0^{(\nu)}$, having the maximal absolute value, i.e.,

$$|\lambda_0^{(\nu)}| > |\lambda_1^{(\nu)}| > \cdots > |\lambda_n^{(\nu)}| > \cdots \to 0. \tag{5.25}$$

Consider the element $w_n^{(\nu)} \in \mathbf{H}_\nu$ given by

$$w_n^{(\nu)} = R_{n-1}^{(\nu)}(\mathbf{L}) c_\nu, \tag{5.26}$$

where

$$R^{(\nu)}_{n-1}(\lambda) = (\lambda - \lambda^{(\nu)}_0)(\lambda - \lambda^{(\nu)}_1) \cdots (\lambda - \lambda^{(\nu)}_{m-1}) \times (\lambda - \lambda^{(\nu)}_{m+1}) \cdots (\lambda - \lambda^{(\nu)}_{n-1}). \quad (5.27)$$

Substituting into (5.26) the expansion of c_ν in terms of the eigenelements of $\mathbf{L}$, $v^{(\nu)}_k \in \mathbf{H}_\nu$, we derive

$$w^{(\nu)}_n = y^{(\nu)}_m R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m) v^{(\nu)}_m + \sum_{k=n}^{\infty} y^{(\nu)}_k R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k) v^{(\nu)}_k, \qquad y^{(\nu)}_k = (v^{(\nu)}_k, c_\nu), \quad (5.28)$$

wherefrom it follows that $w^{(\nu)}_n$ belongs to the subspace $\mathbf{H}^{(m)}_\nu \subset \mathbf{H}_\nu$ orthogonal to the first m eigenelements $v^{(\nu)}_0, v^{(\nu)}_1, \ldots, v^{(\nu)}_{m-1}$. According to our notation (5.25), $\lambda^{(\nu)}_m$ is the eigenvalue having the maximal numerical value in this subspace. Further, let $\lambda^{(n+1)}_m(\tau)$ be the corresponding eigenvalue of the approximate operator $\mathbf{L}_{n+1}(\tau)$. For brevity of script we shall neglect, henceforth, the indication of the τ dependence. This will not affect the generality of our derivations. Now, since $w^{(\nu)}_n$ by definition belongs also to the subspace $\mathbf{H}_{n+1}$ spanned by the first $n+1$ elements of the basis (5.6) in $\mathbf{H}_\nu$, we have by virtue of the variational principle

$$\lambda^{(n+1)}_m \gtrless \frac{(L_{n+1} w^{(\nu)}_n, w^{(\nu)}_n)}{\|w^{(\nu)}_n\|^2}, \quad (5.29)$$

where the upper sign has to be taken for $\lambda^{(n+1)}_m > 0$ and the lower for $\lambda^{(n+1)}_m < 0$. We note, furthermore, that from the definition of the operator $\mathbf{L}_{n+1}$, (5.21), and due to the particular choice of the element $w^{(\nu)}_n$, (5.26), we have

$$(\mathbf{L}_{n+1} w^{(\nu)}_n, w^{(\nu)}_n) = (\mathbf{L} w^{(\nu)}_n, w^{(\nu)}_n). \quad (5.30)$$

The equality (5.28) implies

$$\|w^{(\nu)}_n\|^2 = \mu^{(\nu)}_m [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)]^2 + \sum_{k=n}^{\infty} \mu^{(\nu)}_k [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)]^2, \qquad \mu^{(\nu)} = |y^{(\nu)}|^2 \quad (5.31)$$

and

$$(\mathbf{L} w^{(\nu)}_n, w^{(\nu)}_n) = \lambda^{(\nu)}_m \mu^{(\nu)}_m [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)]^2 + \sum_{k=n}^{\infty} \lambda^{(\nu)}_k \mu^{(\nu)}_k [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)]^2. \quad (5.32)$$

The substitution of (5.30)–(5.32) into (5.29) yields

$$\lambda^{(n+1)}_m \gtrless \frac{\lambda^{(\nu)}_m + [1/\mu^{(\nu)}_m] \sum_{k=n}^{\infty} \lambda^{(\nu)}_k \mu^{(\nu)}_k [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)/R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)]^2}{1 + [1/\mu^{(\nu)}_m] \sum_{k=n}^{\infty} \mu^{(\nu)}_k [R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)/R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)]^2}. \quad (5.33)$$

Since $|\lambda^{(\nu)}_k| \leqslant |\lambda^{(\nu)}_n|$, $(k \geqslant n)$, it follows that

$$\frac{|R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)|}{|R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)|} \leqslant \frac{|\lambda^{(\nu)}_0| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_0| - |\lambda^{(\nu)}_m|} \cdots \frac{|\lambda^{(\nu)}_{m-1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_{m-1}| - |\lambda^{(\nu)}_m|} \frac{|\lambda^{(\nu)}_{m+1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_m| - |\lambda^{\nu}_{m+1}|} \cdots \frac{|\lambda^{(\nu)}_{n-1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_m| - |\lambda^{(\nu)}_{n-1}|}. \quad (5.34)$$

Thus, for a fixed m this quantity tends to zero when $n \to \infty$ [Cf. (5.25)]. Therefore, neglecting terms of second order, one might rewrite the inequality (5.33) as follows:

$$\lambda^{(n+1)}_m \gtrless \lambda^{(\nu)}_m - \frac{1}{\mu^{(\nu)}_m} \sum_{k=n}^{\infty} (\lambda^{(\nu)}_m - \lambda^{(\nu)}_n) \mu^{(\nu)}_k \left(\frac{R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)}{R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)} \right)^2 \quad (5.35)$$

or for any sign of $\lambda^{(n+1)}_m$

$$|\lambda^{(\nu)}_m - \lambda^{(n+1)}_m| \leqslant \frac{1}{\mu^{(\nu)}} \sum_{k=n}^{\infty} |\lambda^{(\nu)}_m - \lambda^{(\nu)}_k| \mu^{(\nu)}_k \left(\frac{R^{(\nu)}_{n-1}(\lambda^{(\nu)}_k)}{R^{(\nu)}_{n-1}(\lambda^{(\nu)}_m)} \right)^2 \leqslant \frac{|\lambda^{(\nu)}_m| + |\lambda^{(\nu)}_n|}{\mu^{(\nu)}_m} \left(\frac{|\lambda^{(\nu)}_0| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_0| - |\lambda^{(\nu)}_m|} \right)^2 \cdots \left(\frac{|\lambda^{(\nu)}_{m-1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_{m-1}| - |\lambda^{(\nu)}_m|} \right)^2 \cdot \left(\frac{|\lambda^{(\nu)}_{m+1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_m| - |\lambda^{(\nu)}_{m+1}|} \right)^2 \cdots \left(\frac{|\lambda^{(\nu)}_{n-1}| + |\lambda^{(\nu)}_n|}{|\lambda^{(\nu)}_m| - |\lambda^{(\nu)}_{n-1}|} \right)^2 \sum_{k=n}^{\infty} \mu^{(\nu)}_k. \quad (5.36)$$

This inequality indicates the speed of convergence of our approximation scheme. *The eigenvalues of the approximate operator* $\mathbf{L}_n(\tau)$ *[the zeros of the quasiorthogonal polynomial* $P_n(E, \tau)$*] tend to the corresponding values of the exact spectrum faster than any geometric progression.* The approximate eigenelements $v^{(m)}_k(\tau)$ have to behave, obviously, the same way.

6. THE APPROXIMATE QUADRATURES

The vast majority of physical systems of interest involve particle–particle interactions. These interactions, being represented in the Hamiltonian by products of at least four second quantization operators ($c^+_{\nu_1} c^+_{\nu_2} c_{\nu_3} c_{\nu_4}$ for binary collisions), give rise to the well-known hierarchy of equations of motion of the one-particle Green's function. In the language of our formalism, this means that here the SWF moments (2.4) are expressions involving thermal averages, correlations of successively growing order. Therefore, the only information which one may derive, in practice, from the calculation of the determinants D_n is confined to the possibility of proving that the Green's function $G(E)$ consists of more than n poles. This, when D_n is essentially positive. However, except for some trivial cases it can not be shown, usually, that a determinant D_n vanishes, since the thermal averages are unknown *a priori*. Bearing in mind our previous discussion, it is needless to say that the vanishing of the nth determinant D_n would indicate that the hierarchy of the equations of motion of the characteristic function (1.10) becomes closed after the nth step and *vice versa*.

Nevertheless, the moment approach enables one to construct formal expressions of the successive approximations of the SWF. In fact, an approximate

n-pole SWF can always be written as given by (4.16). The characteristic feature of such an approximation is its τ dependence. More specifically, calculating a certain statistical quantity $\bar{A}$ in the n-pole approximation, one derives an expression $\bar{A}_n(\{\tau_\nu\})$, the numerical value of which depends on the set of the real parameters $\{\tau_\nu\}$. It is clear that at the final stage the thermal averages involved in $\mu_k^{(n)}(\tau)$ and $\lambda_k^{(n)}(\tau)$ will have to be evaluated in some self-consistent manner. This implies that not just the, SWF but the moment problem itself (given by the positive sequence $\{s_n\}_0^\infty$), will become approximate. Let us, however, ignore this difficulty. For the purposes of the present paper it will be sufficient to discuss formal expressions for the approximate SWF or, in other words, to assume that the moments s_n are known explicitly. This, by the way, will always be the case when systems with one-particle interactions are considered (impurity scattering, electron–field interaction, etc.).

In the general case of a system with binary interactions we have

$$H = H_0 + H_I, \tag{6.1}$$

where H_0 stands for the unperturbed Hamiltonian

$$H_0 = \sum_{k,\sigma} \epsilon_{k,\sigma} c^+_{k,\sigma} c_{k,\sigma}. \tag{6.2}$$

The free-particle spectrum $\epsilon_{k,\sigma}$ is measured, as usual, with respect to the chemical potential μ. The interaction energy is given by

$$H_I = \frac{\gamma}{V} \sum_{\sigma_1, k_1, \cdots, k_4, \sigma_2} \delta_{k_1+k_2, k_3+k_4} \phi(k_2 - k_3) \times c^+_{k_1,\sigma_1} c^+_{k_2,\sigma_2} c_{k_3,\sigma_2} c_{k_4,\sigma_1}. \tag{6.3}$$

Here $\gamma > 0$ is a coupling constant and $\phi(k_2 - k_3)$ a real symmetric function.

We introduce the characteristic function

$$\Lambda^{(q,\sigma)}(t) = \langle [c^+_{q,\sigma}; c_{q,\sigma}(t)]_+ \rangle. \tag{6.4}$$

Now in terms of the corresponding SWF, the particle distribution function will be given by

$$\langle n_{q,\sigma} \rangle = \int_{-\infty}^{\infty} (d\omega/2\pi) f(\omega) \Lambda^{(q,\sigma)}(\omega). \tag{6.5}$$

Hence, using the equation of motion for (6.4),

$$\left(i \frac{\partial}{\partial t} - \epsilon_{q,\sigma} \right) \Lambda^{(q,\sigma)}(t) = \langle [[H_I, c^+_{q,\sigma}]_-; c_{q,\sigma}(t)]_+ \rangle, \tag{6.6}$$

one easily derives the formal expression for the mean interaction energy

$$\frac{\langle H_I \rangle}{V} = \frac{1}{2V} \sum_{q,\sigma} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} (\omega - \epsilon_{q,\sigma}) f(\omega) \Lambda^{(q,\sigma)}(\omega), \tag{6.7}$$

whence the total energy of the system is given by

$$\frac{\langle H \rangle}{V} = \frac{E - \mu N}{V} = \frac{1}{2V} \sum_{q,\sigma} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} (\omega + \epsilon_{q,\sigma}) f(\omega) \Lambda^{(q,\sigma)}(\omega). \tag{6.8}$$

A similar expression can be derived for the thermodynamical potential

$$\Omega = -\beta^{-1} \ln \operatorname{Tr} e^{-\beta H}. \tag{6.9}$$

It is easy to see that

$$\frac{\Omega - \Omega_0}{V} = \frac{1}{V} \int_0^\gamma \frac{d\gamma'}{\gamma'} \langle H_I \rangle_{(\gamma')} = \frac{1}{2V} \sum_{q,\sigma} \int_0^\gamma \frac{d\gamma'}{\gamma'} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} (\omega - \epsilon_{q,\sigma}) f(\omega) \Lambda^{(q,\sigma)}(\omega), \tag{6.10}$$

where Ω_0 stands for the free energy of the unperturbed system ($\gamma = 0$).

In what follows, we discuss the thermal average of an arbitrary dynamical operator, say A, under the assumption that it is given by

$$\bar{A} = \sum_\nu \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F(\omega) \Lambda^{(\nu)}(\omega), \tag{6.11}$$

where the function $F(\omega)$ is a product of a certain polynomial $R(\omega)$ and the Fermi function $f(\omega)$:

$$F(\omega) = R(\omega) f(\omega). \tag{6.12}$$

All the thermal averages mentioned above are particular cases of the generalized expression (6.11). Moreover, all equal-time correlations which one might derive from the hierarchy of equations of motion are expressible in the same form. Thus the expression (6.11) is quite general.

Since the exact SWF is unknown, we turn to the discussion of the approximate values of $\bar{A}$ which follow from (6.11) as a result of substitution of the approximate n-pole SWF instead of $\Lambda^{(\nu)}(\omega)$. We have

$$\bar{A}_n(\{\tau_\nu\}) = \sum_\nu \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F(\omega) \Lambda_n^{(\nu)}(\omega; \tau_\nu) = \sum_\nu \sum_{k=1}^{n} \mu_k^{(n)}(\lambda_k^{(\nu)}) F(\lambda_k^{(\nu)}), \tag{6.13}$$

where $\lambda_k^{(\nu)} \equiv \lambda_k^{(\nu)}(\tau_\nu)$ are the zeros of the quasi-orthogonal polynomial $P_n^{(\nu)}(E; \tau_\nu)$. We refer to $\bar{A}_n(\{\tau_\nu\})$ as the nth *approximate quadrature*. As was mentioned before, this is an approximation which depends, in general, on the set of real parameters $\{\tau_\nu\} (-\infty < \tau_\nu \leq \infty)$. In the framework of a given approximation, all the values of $\bar{A}_n(\{\tau_\nu\})$ for all possible sets of parameters $\{\tau_\nu\}$ have to be considered, in principle, as being equally reliable. There is no particular reason for preference.

The first question which arises here concerns the convergence of this approximation scheme. In a case where the numerical values of the parameters $\tau_\nu = 0$, this problem becomes settled by virtue of the following theorem.[18]

Theorem: Let $F_1(\omega) = \sum_{n=0}^{\infty} c_n \omega^{2n}$ be an integral transcendental function with nonnegative coefficients, such that $\sum_{i=0}^{\infty} \mu_{2i} c_i$ converges. Then if $F(\omega)$, when $|\omega| \to \infty$, is dominated by $F_1(\omega)$, then

$$\lim_{n \to \infty} \bar{A}_n^{(\nu)}(0) = \bar{A}^{(\nu)}. \tag{6.14}$$

Here $\bar{A}_n^{(\nu)}(0)$ is the one-particle nth approximate quadrature with $\tau_\nu = 0$:

$$\bar{A}_n^{(\nu)}(\tau_\nu) = \int_{-\infty}^{\infty} (d\omega/2\pi) F(\omega) \Lambda_n^{(\nu)}(\omega; \tau_\nu), \qquad (6.15)$$

and $\bar{A}^{(\nu)}$ is the exact quadrature given by the same integral with $\Lambda^{(\nu)}(\omega)$ instead of $\Lambda_n^{(\nu)}(\omega; \tau_\nu)$. The validity of this theorem in our case is obvious, recalling that $F(\omega)$ is given by (6.12).

Without loss of generality we will again neglect the quantum indices ν. Our intent is to show that *for every τ there exists such a τ_0' such that*

$$\Delta(\tau, \tau_0') = \bar{A}_n(\tau) - \bar{A}_{n-1}(\tau_0') = 0. \qquad (6.16)$$

To this end we construct a polynomial of the degree $2n-4$, $R_{2n-4}(\omega)$, such that

$$R_{2n-4}(\lambda_k^{(n-1)}(\tau')) = F(\lambda_k^{(n-1)}(\tau')), \quad k = 1, \ldots, n-1 \qquad (6.17a)$$

$$R_{2n-4}(\lambda_k^{(n)}(\tau)) = F(\lambda_k^{(n)}(\tau)), \quad k = 1, \ldots, r-1, r+2, \ldots, n. \qquad (6.17b)$$

Thus, $R_{2n-4}(\omega)$ is Lagrange's interpolation polynomial of the function $F(\omega)$. The interpolation is based on $2n-3$ points—all the zeros of the quasi-orthogonal polynomial $P_{n-1}(E; \tau')$ and $n-2$ zeros of $P_n(E; \tau)$.

Now, since an approximate n-pole SWF satisfies the first $2n-2$ moment equations, we have

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} R_{2n-4}(\omega) \Lambda_{n-1}(\omega; \tau') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} R_{2n-4}(\omega) \Lambda_n(\omega; \tau) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} R_{2n-4}(\omega) \Lambda(\omega), \qquad (6.18)$$

whence using (6.17a) and (6.17b), we derive

$$\Delta(\tau, \tau') = \mu_r^{(n)}[F(\lambda_r^{(n)}) - R_{2n-4}(\lambda_r^{(n)})] + \mu_{r+1}^{(n)}[F(\lambda_{r+1}^{(n)}) - R_{2n-4}(\lambda_{r+1}^{(n)})]. \qquad (6.19)$$

The polynomial $R_{2n-4}(\omega)$ depends on both τ and τ', since these parameters define the interpolation basis. We have to prove that for a fixed τ we can choose such a τ' that (6.19) vanishes.

Let us first investigate the behavior of $\lambda_k^{(n)}(\tau)$ (for an arbitrary n) as a function of τ. Since $\lambda_k^{(n)}(\tau)$ is a zero of $P_n(E; \tau)$,

$$P_n(\lambda_k^{(n)}; \tau) = P_n(\lambda_k^{(n)}) - \tau P_{n-1}(\lambda_k^{(n)}) = 0 \qquad (6.20)$$

and

$$\tau = P_n(\lambda_k^{(n)}) / P_{n-1}(\lambda_k^{(n)}). \qquad (6.21)$$

Differentiating (6.20) and substituting (6.21), we derive

$$\frac{d\lambda_k^{(n)}}{d\tau} = \frac{P_{n-1}^2(\lambda_k^{(n)})}{P_n'(\lambda_k^{(n)}) P_{n-1}(\lambda_k^{(n)}) - P_n(\lambda_k^{(n)}) P_{n-1}'(\lambda_k^{(n)})}. \qquad (6.22)$$

Making use of the Christoffel–Darboux formula,[19]

$$(\eta - \lambda) \sum_{k=0}^{n-1} P_k(\lambda) P_k(\eta) = b_{n-1}[P_n(\eta) P_{n-1}(\lambda) - P_n(\lambda) P_{n-1}(\eta)], \qquad (6.23)$$

where $b_{n-1} > 0$ is given by (3.6), we may rewrite the denominator of (6.22) in the form

$$P_n'(\lambda_k^{(n)}) P_{n-1}(\lambda_k^{(n)}) - P_n(\lambda_k^{(n)}) P_{n-1}'(\lambda_k^{(n)}) = \frac{1}{b_{n-1}} \sum_{i=0}^{n-1} |P_i(\lambda_k^{(n)})|^2. \qquad (6.24)$$

Finally, substituting (6.24) into (6.22) and using (4.3′), we have

$$\frac{d\lambda_k^{(n)}}{d\tau} = b_{n-1} \mu_k^{(n)} |P_{n-1}(\lambda_k^{(n)})|^2. \qquad (6.25)$$

Thus, when τ varies along the real axis, $\lambda_k^{(n)}(\tau)$ moves between two consequent zeros of the orthonormal polynomial $P_{n-1}(E)$. The speed of variation of $\lambda_k^{(n)}(\tau)$ is positive and reaches zero at both ends of this interval. Recalling that the zeros of the polynomial $P_n(E)$ are separated by that of $P_{n-1}(E)$, we draw a schematic picture of the intervals in which $\lambda_k^{(n-1)}(\tau')$ and $\lambda_k^{(n)}(\tau)$ vary (Fig. 2). Since $d\lambda/d\tau$ is positive, the growth of τ causes the $\lambda_k^{(n)}(\tau)$ to move in the positive direction between the boundaries of the corresponding interval.

Assume that τ is given. Whatever the value of τ is, there will exist two neighboring roots, $\lambda_r^{(n)}(\tau)$ and $\lambda_{r+1}^{(n)}(\tau)$, belonging to the same interval of variation of some $\lambda_k^{(n-1)}(\tau')$. These roots we choose as the missing ones in the interpolation basis of $R_{2n-4}(\omega)$ [Cf. (6.17b)].

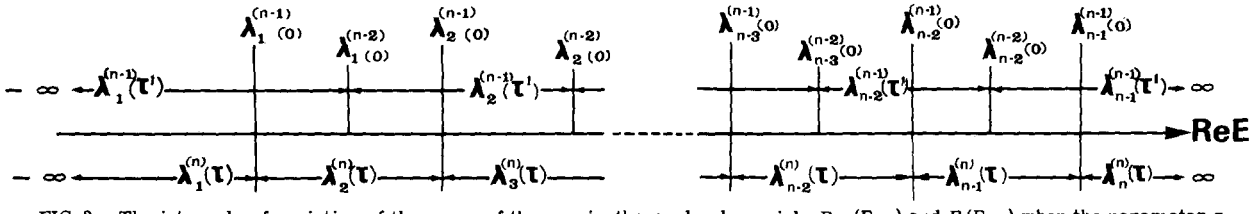

FIG. 2. The intervals of variation of the zeros of the quasiorthogonal polynomials $P_{n-1}(E, \tau)$ and $P_n(E, \tau)$ when the parameter τ varies along the real axis (schematic).

Now, consider such a τ' that

$$\lambda_r^{(n)}(\tau) < \lambda_k^{(n-1)}(\tau') < \lambda_{r+1}^{(n)}(\tau). \tag{6.26}$$

Then, obviously, the function

$$\phi(\omega) = F(\omega) - R_{2n-4}(\omega) \tag{6.27}$$

will have opposite signs at the points $\lambda_r^{(n)}(\tau)$ and $\lambda_{r+1}^{(n)}(\tau)$. Suppose it is positive for $\lambda_r^{(n)}$ and negative for $\lambda_{r+1}^{(n)}$ (Fig. 3). Then, since $\mu_r^{(n)}$ and $\mu_{r+1}^{(n)}$ depend only on τ which is fixed by assumption, $\Delta(\tau, \tau')$ will change the sign when $\lambda_k^{(n-1)}(\tau')$ varies in the interval $(\lambda_r^{(n)}, \lambda_{r+1}^{(n)})$. In fact, $\Delta(\tau, \tau') > 0$ when $\lambda_k^{(n-1)}(\tau')$ is arbitrarily close to $\lambda_{r+1}^{(n)}(\tau)$ and $\Delta(\tau, \tau') < 0$ when $\lambda_k^{(n-1)}(\tau')$ is sufficiently close to $\lambda_r^{(n)}(\tau)$.

Therefore, there will exist a $\tau' = \tau_0'$ for which $\Delta(\tau, \tau_0') = 0$, since $\Delta(\tau, \tau')$ is a continuous function of τ'. This completes the proof of (6.16).

Although (6.16) has been proven for the case of one-particle approximate quadratures (6.15), it will, obviously, remain true also for the general case (6.13). Thus, we may generalize our statement.

For every set of parameters $\{\tau_\nu\}$ *there exist such a set* $\{\tau_\nu'\}$ *such that*

$$\Delta(\{\tau_\nu\}, \{\tau_\nu'\}) = \bar{A}_n(\{\tau_\nu\}) - \bar{A}_{n-1}(\{\tau_\nu'\}) = 0, \tag{6.28}$$

where $\bar{A}_n(\{\tau_\nu\})$ is given by (6.13).

This is an important result. It shows, first, that the approximate quadratures are converging uniformly. In fact, taking a set of parameters $\{\tau_\nu\}$, where $|\tau_\nu| \to \infty$, we have

$$\lim_{|\tau_\nu| \to \infty} \bar{A}_n(\{\tau_\nu\}) = \bar{A}_{n-1}(\{0\}). \tag{6.29}$$

We recall that the values of $\lambda_k^{(n)}(\tau)$ tend to the roots of $P_{n-1}(E)$ when $|\tau| \to \infty$ [$\lambda_n^{(n)}(\tau) \to \infty$ when $\tau \to \infty$ and $\lambda_1^{(n)}(\tau) \to -\infty$ when $\tau \to -\infty$]. The convergence of the rhs of (6.29) is guaranteed by (6.14). Moreover, (6.28) indicates that the absolute maximum of the successive approximations decreases and, correspondingly, the absolute minimum increases.

This behavior of the approximate quadratures enables one to establish the exact upper and lower error bounds of a given approximation. These, obviously, will be the absolute maximum and minimum of $\bar{A}_n(\{\tau_\nu\})$. We note, finally, that for every approximation there always exists a set of parameters $\{\tau_\nu^{(0)}\}$ for which $\bar{A}_n(\{\tau_\nu^{(0)}\}) = \bar{A}$.

7. THE STATISTICS OF IMPURITY SCATTERED STATES

This section is devoted to an example. The specific model which we discuss below is a simple one. It is based on the physical picture of a metal where 'free' Bloch electrons are being scattered by impurities. The Hamiltonian of the system is given by

$$H = \sum_{k,\sigma} \epsilon_{k,\sigma} c_{k,\sigma}^+ c_{k,\sigma} + \gamma \sum_{k,\sigma;\, k',\sigma'} U_{k,\sigma;\, k',\sigma'}^{\{r\}} c_{k,\sigma}^+ c_{k',\sigma'}. \tag{7.1}$$

Here $\epsilon_{k,\sigma}$, as in the foregoing section, represents the unperturbed spectrum (the spin index σ becomes relevant in the presence of a magnetic field). $\gamma > 0$ is a coupling constant and $U_{k,\sigma;\, k',\sigma'}^{\{r\}}$ stands for the electron–impurity interaction. It depends, obviously, on the particular distribution of the impurities, i.e., on their coordinates $\{r\}$ in the lattice of the specimen. The interaction $U_{k,\sigma;\, k',\sigma'}^{\{r\}}$ as written in (6.1) is quite general and may represent ordinary potential scattering as well as spin-flip processes.

We confine ourselves to the calculation of two thermodynamic quantities, namely, the electron distribution function and the mean interaction energy. Considering the characteristic function (6.4), it is easy to see that the expressions (6.5) and (6.7) remain valid for the Hamiltonian (7.1), except for the factor $\frac{1}{2}$ in (6.7) which is absent in the present case. We have here

$$\frac{\langle H_I \rangle}{V} = \frac{1}{V} \sum_{q,\sigma} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} (\omega - \epsilon_{q,\sigma}) f(\omega) \Lambda^{(q,\sigma)}(\omega), \tag{7.2}$$

where H_I stands for the second term in (7.1).

The fact that the Hamiltonian under consideration does not contain electron–electron interactions considerably simplifies the calculations. In this case, the SWF moments are given explicitly by the energy parameters of the system. We have

$$s_n^{(q,\sigma)} = \langle [c_{q,\sigma}^+, \mathbf{L}^n c_{q,\sigma}]_+ \rangle = (\mathbf{W}^n)_{q,\sigma;q,\sigma}, \tag{7.3}$$

where W is the matrix whose elements are

$$W_{k,\sigma;\, k',\sigma'} = \epsilon_{k,\sigma} \delta_{k,k'} \delta_{\sigma,\sigma'} + \gamma U_{k,\sigma;\, k',\sigma'}^{\{r\}}. \tag{7.4}$$

This expression contains an unknown term—the chemical potential μ ($\epsilon_{k,\sigma} = \epsilon_{k,\sigma}^{(0)} - \mu$). In every approximation, it will have to be determined self-consistently from the given electron density

$$\bar{N} = \sum_{q,\sigma} \langle n_{q,\sigma} \rangle. \tag{7.5}$$

We turn now to our approximation scheme where we will denote the nth approximate quadrature of $\langle n_{q,\sigma} \rangle$ by $\bar{n}_n^{(q,\sigma)}(\tau)$, and that of $\langle H_I \rangle$ by $\bar{H}_n^{(I)}(\{\tau_{q,\sigma}\})$.

A. The Lowest Approximation

According to our procedure, we have to calculate first the determinant

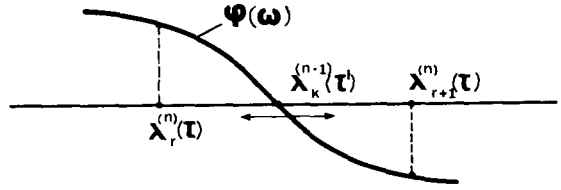

FIG. 3. See text for explanation.

$$D_1^{(q,\sigma)} = s_0^{(q,\sigma)} s_2^{(q,\sigma)} - (s_1^{(q,\sigma)})^2 = \gamma^2 \sum_{k \neq q,\sigma'} |U^{\{\mathbf{r}\}}_{q,\sigma;k,\sigma'}|^2 > 0. \quad (7.6)$$

It is, obviously, positive definite since our Hamiltonian is nondiagonal in the Bloch state representation. The approximate one-pole SWF is given by

$$\Lambda_1^{(q,\sigma)}(\omega, \tau_{q,\sigma}) = 2\pi\delta(\omega - \lambda_1^{(1)}(\tau_{q,\sigma})), \quad (7.7)$$

where $\lambda_1^{(1)}$ is the zero of the quasiorthogonal polynomial

$$P_1^{(q,\sigma)}(E, \tau_{q,\sigma}) = (D_1^{(q,\sigma)})^{-1/2}[E - s_1^{(q,\sigma)} - \tau_{q,\sigma}(D_1^{(q,\sigma)})^{1/2}]. \quad (7.8)$$

Thus,

$$\lambda_1^{(1)}(\tau_{q,\sigma}) = s_1^{(q,\sigma)} + \tau_{q,\sigma}(D_1^{(q,\sigma)})^{1/2}. \quad (7.9)$$

Using the approximation (7.7) and substituting (7.9) and (7.3), we derive for the first approximate quadrature of the electron distribution function (6.5):

$$\bar{n}_1^{(q,\sigma)}(\tau_{q,\sigma}) = f[W_{q,\sigma;q,\sigma} + \tau_{q,\sigma}(D_1^{(q,\sigma)})^{1/2}]. \quad (7.10)$$

When $\tau_{q,\sigma}$ varies along the real axis, (7.10) assumes all the numerical between zero ($\tau_{q,\sigma} \to \infty$) and 1 ($\tau_{q,\sigma} \to -\infty$) (see Fig. 4). This might be expected since approximation (7.7) with $\tau_{q,\sigma} \neq 0$ satisfies just the normalization condition (the zeroth moment) and, therefore, reflects a property which is in common for all systems in nature. If $\tau_{q,\sigma} = 0$, then (7.7) satisfies also the first SWF moment equation. This is equivalent to the first Born approximation (Hartree–Fock, in case of systems with binary interactions). Similarly, for the first approximate quadrature of the mean interaction energy (7.2), we have

$$\frac{1}{V}\bar{H}_1^{(1)}(\{\tau_{q,\sigma}\}) = \frac{1}{V}\sum_{q,\sigma}[\gamma U^{\{\mathbf{r}\}}_{q,\sigma;q,\sigma} + \tau_{q,\sigma}(D_1^{(q,\sigma)})^{1/2}]\bar{n}_1^{(q,\sigma)}(\tau_{q,\sigma}). \quad (7.11)$$

This quantity tends to $-\infty$ when $\tau_{q,\sigma} \to -\infty$. It vanishes at the point $\{\tau_{q,\sigma}^{(0)}\}$, where

$$\tau_{q,\sigma}^{(0)} = -\gamma U^{\{\mathbf{r}\}}_{q,\sigma;q,\sigma}/(D_1^{(q,\sigma)})^{1/2}. \quad (7.12)$$

It tends to zero when $\tau_{q,\sigma} \to \infty$.

B. The Second Approximation

The second determinant is given by

$$D_2^{(q,\sigma)} = \begin{vmatrix} s_0^{(q,\sigma)} & s_1^{(q,\sigma)} & s_2^{(q,\sigma)} \\ s_1^{(q,\sigma)} & s_2^{(q,\sigma)} & s_3^{(q,\sigma)} \\ s_2^{(q,\sigma)} & s_3^{(q,\sigma)} & s_4^{(q,\sigma)} \end{vmatrix} = (\mathbf{W}^{(4)})_{q,\sigma;q,\sigma} D_1^{(q,\sigma)} - (\mathbf{W}^3)_{q,\sigma;q,\sigma} d_1^{(q,\sigma)} + (\mathbf{W}^2)_{q,\sigma;q,\sigma} d_2^{(q,\sigma)} \geqslant 0, \quad (7.13)$$

where

$$d_1^{(q,\sigma)} = (\mathbf{W}^3)_{q,\sigma;q,\sigma} - W_{q,\sigma;q,\sigma}(\mathbf{W}^2)_{q,\sigma;q,\sigma} \quad (7.14)$$

and

$$d_2^{(q,\sigma)} = W_{q,\sigma;q,\sigma}(\mathbf{W}^3)_{q,\sigma;q,\sigma} - (\mathbf{W}^2)^2_{q,\sigma;q,\sigma}. \quad (7.15)$$

We assume that $D_2^{(q,\sigma)}$ does not vanish for the given scattering potential. Using the definitions (3.4) and (3.10), we have for the second quasiorthogonal polynomial

$$P_2^{(q,\sigma)}(E, \tau_{q,\sigma}) = (D_1^{(q,\sigma)} D_2^{(q,\sigma)})^{-1/2} \times \{E^2 D_1^{(q,\sigma)} - E[d_1^{(q,\sigma)} + \tau_{q,\sigma}(D_2^{(q,\sigma)})^{1/2}] + d_2^{(q,\sigma)} + \tau_{q,\sigma}(D_2^{(q,\sigma)})^{1/2} W_{q,\sigma;q,\sigma}\}. \quad (7.16)$$

The zeros of this polynomial are

$$\lambda_{1,2}^{(2)}(\tau_{q,\sigma}) = W_{q,\sigma;q,\sigma} + (1/2D_1^{(q,\sigma)})(d^{(q,\sigma)}(\tau_{q,\sigma}) \pm \{[d^{(q,\sigma)}(\tau_{q,\sigma})]^2 + 4(D_1^{(q,\sigma)})^3\}^{1/2}), \quad (7.17)$$

where

$$d^{(q,\sigma)}(\tau_{q,\sigma}) = d_1^{(q,\sigma)} + \tau_{q,\sigma}(D_2^{(q,\sigma)})^{1/2} - 2W_{q,\sigma;q,\sigma} D_1^{(q,\sigma)}. \quad (7.18)$$

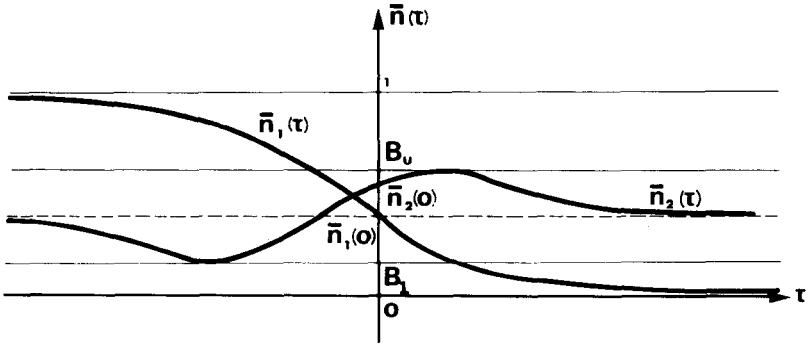

FIG. 4. The first two approximate quadratures for the electron distribution function (schematic). B_u and B_l are the rigorous upper and lower error bounds in the two-pole approximation.

We turn now to the approximate two-pole SWF

$$\Lambda_2^{(q,\sigma)}(\omega;\tau_{q,\sigma}) = 2\pi[\mu_1^{(2)}\delta(\omega-\lambda_1^{(2)}) + \mu_2^{(2)}\delta(\omega-\lambda_2^{(2)})]. \tag{7.19}$$

The coefficients $\mu_{1,2}^{(2)}$ are defined by (4.3). Using (3.13), (3.9), and (7.17), we derive

$$\mu_1^{(2)} = C^{(q,\sigma)}(\lambda_2^{(2)} - W_{q,\sigma;q,\sigma}) \tag{7.20}$$

and

$$\mu_2^{(2)} = -C^{(q,\sigma)}(\lambda_1^{(2)} - W_{q,\sigma;q,\sigma}), \tag{7.21}$$

where

$$C^{(q,\sigma)} = D_1^{(q,\sigma)}/\{[d^{(q,\sigma)}(\tau_{q,\sigma})]^2 + 4(D_1^{(q,\sigma)})^3\}^{1/2}. \tag{7.22}$$

Now, with (7.19) the second approximate quadrature of the electron distribution function (6.5) yields

$$\bar{n}_2^{(q,\sigma)}(\tau_{q,\sigma}) = C^{(q,\sigma)}\{(\lambda_2^{(2)} - W_{q,\sigma;q,\sigma})f(\lambda_1^{(2)}) - (\lambda_1^{(2)} - W_{q,\sigma;q,\sigma})f(\lambda_2^{(2)})\}. \tag{7.23}$$

This is a continuous function of $\tau_{q,\sigma}$. When $|\tau_{q,\sigma}| \to \infty$, it obviously tends to $\bar{n}_1^{(q,\sigma)}(0)$ (Fig. 4). Therefore, it has to have at least one maximum and one minimum. According to the derivations of the foregoing section, we find

$$B_u = \sup_{(\tau_{q,\sigma})} \bar{n}_2^{(q,\sigma)}(\tau_{q,\sigma}) \leq 1 \tag{7.24}$$

and

$$B_l = \inf_{(\tau_{q,\sigma})} \bar{n}_2^{(q,\sigma)}(\tau_{q,\sigma}) \geq 0. \tag{7.25}$$

B_u and B_l represent, correspondingly, the rigorous upper and lower error bounds of a calculation based on the information contained by the first three moments of the SWF. If the fourth moment is such that (7.13) vanishes, then $\bar{n}_2^{(q,\sigma)}(0)$ would yield the exact value $\bar{n}^{(q,\sigma)}$.

Before turning to the calculation of the second approximate quadrature of the mean interaction energy, let us rewrite (7.2) as follows:

$$\frac{1}{V}\langle H_I\rangle = \frac{1}{V}\sum_{q,\sigma}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}(\omega - W_{q,\sigma;q,\sigma})f(\omega)\Lambda^{(q,\sigma)}(\omega) + \frac{\gamma}{V}\sum_{q,\sigma}U^{(r)}_{q,\sigma;q,\sigma}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}f(\omega)\Lambda^{(q,\sigma)}(\omega). \tag{7.26}$$

The second term is nothing but the diagonal part of the mean interaction energy of the system (7.1):

$$\frac{1}{V}\langle H_I\rangle_{\text{diag}} = \frac{\gamma}{V}\sum_{q,\sigma}U^{(r)}_{q,\sigma;q,\sigma}\langle n_{q,\sigma}\rangle. \tag{7.27}$$

This part takes care of the forward scattering processes. Substituting the approximate two-pole SWF into the first term of (7.26), we derive

$$\frac{1}{V}\bar{H}_2^{(I)}(\{\tau_{q,\sigma}\}) - \frac{1}{V}\langle H_I\rangle_{\text{diag}} = -\sum_{q,\sigma}C^{(q,\sigma)}D_1^{(q,\sigma)}[f(\lambda_1^{(2)}) - f(\lambda_2^{(2)})]. \tag{7.28}$$

Since by definition $\lambda_1^{(2)} < \lambda_2^{(2)}$, it follows that $f(\lambda_1^{(2)}) > f(\lambda_2^{(2)})$. $C^{(q,\sigma)}$ and $D_1^{(q,\sigma)}$ are positive. Therefore, (7.28) is negative for all possible sets $\{\tau_{q,\sigma}\}$. Among these will be such a set of the parameters $\tau_{q,\sigma}$ for which the second approximate quadrature coincides with the exact value of the mean interaction energy. We derive, thus, the following conclusion:

The thermal average of the diagonal part of a Hamiltonian which represents a system with single-particle interactions overestimates algebraically the mean total energy of it.

This result

$$\langle H\rangle \leq \langle H\rangle_{\text{diag}} \tag{7.29}$$

is valid at all temperatures. It might be considered as a generalization of the modified Peierls inequality[20]

$$\Omega(\{H_0\}) + \lambda\frac{\mathrm{Tr}H_I e^{-\beta H_0}}{\mathrm{Tr}e^{-\beta H_0}} \geq \Omega(\{H_0 + \lambda H_I\}), \tag{7.30}$$

where $\Omega(\{H\})$ is the free energy as a functional of the Hamiltonian, which implies that the thermal average of the interaction energy taken with the unperturbed density matrix overestimates the true value of it.

Higher-order approximations can be derived in the same manner.

ACKNOWLEDGMENT

The author would like to thank Dr. A. Solomon for helpful discussions.

APPENDIX: THE PADE APPROXIMANT OF THE THERMAL GREEN'S FUNCTION

A Padé approximant[21] of a certain function, say $G(E)$, is given as a fraction of two polynomials. If the numerator $q(E)$ is a polynomial of the degree M and the denominator $p(E)$ has the degree N, then the fraction defines the $[N, M]$ Padé approximant. The coefficients of these polynomials are determined by equating the terms of equal power in E in the equation

$$G(E)p(E) - q(E) = AE^{M+N+1} + BE^{M+N+2} + \cdots, \quad p(0) = 1. \tag{A1}$$

Following the usual notation, we write

$$[N, M]G(E) = q(E)/p(E). \tag{A2}$$

Let us consider the continued fraction

$$\cfrac{1}{E - a_0 - \cfrac{b_0^2}{E - a_1 - \cfrac{b_1^2}{E - a_2 - \ddots}}}, \quad \mathrm{Im}E \neq 0, \tag{A3}$$

where a_k and b_k are the coefficients (3.6). It is well known that in case of a determinate moment problem, this fraction is converging and represents the Green function (1.8).[22] (A3) is known as the real J fraction. Consider, hence, the successive approximants of (A3), namely,

$$\frac{q_1(E)}{p_1(E)} = \frac{1}{E - a_0}, \quad \frac{q_2(E)}{p_2(E)} = \frac{1}{E - a_0 - b_0^2/(E - a_1)}, \cdots. \tag{A4}$$

It is easy to see that the numerators and denominators of these fractions satisfy the finite-difference equation

$$Y_{k+1} = (E - a_k)Y_k - b_{k-1}^2 Y_{k-1}, \quad (k = 1, 2, \ldots), \tag{A5}$$

with the initial conditions

$$q_0(E) = 0, \quad q_1(E) = 1, \quad p_0(E) = 1, \quad p_1(E) = E - a_0. \tag{A6}$$

Furthermore, the substitution

$$y_0 = y_0, \quad y_k = \frac{1}{b_0 b_1 \cdots b_{k-1}} Y_k, \quad k = 1, 2, \ldots \tag{A7}$$

transforms (A5) and (A6) into the finite-difference equation (3.5) with the initial conditions (3.7) and (3.8), the solution of which are, correspondingly, the polynomials of the first and second kind $P_k(E)$ and $Q_k(E)$. Since $q_N(E)/p_N(E)$ in (A4) represents the $[N, N-1]$ Padé approximant of (A3), we have

$$[N, N-1]G(E) = G_N(E) = Q_N(E)/P_N(E). \tag{A8}$$

Thus, on the complex G plane (see Fig. 1) the successive Padé approximants will be represented by the points of touch of the successive circular contours $A_1, A_2, A_3, \cdots$.

1 J. H. van Vleck, Phys. Rev. **74**, 1168 (1948).
2 G. Placzek, Phys. Rev. **86**, 377 (1952).
3 R. G. Gordon, J. Chem. Phys. **40**, 1973 (1964).
4 R. G. Gordon, J. Chem. Phys. **41**, 1819 (1964).
5 A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967).
6 R. G. Gordon, J. Math. Phys. **9**, 655 (1968).
7 R. G. Gordon, J. Math. Phys. **9**, 1087 (1968).
8 J. C. Wheeler and R. G. Gordon, J. Chem. Phys. **51**, 5566 (1969).
9 D. N. Zubarev, Usp. Fiz. Nauk **71**, 71 (1960) [Sov. Phys. Usp. 3, 320 (1960)].
10 D. K. Kalashnikov and E. S. Fradkin, Zh. Eksp. Teor. Fiz. **55**, 607 (1968) [Sov. Phys. JETP **28**, 317 (1969)].
11 In the sense that all Hankel forms $\sum_{i,k=0}^{m} s_{i+k} x_i x_k$ are positive definite.
12 N. I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1965).
13 J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, New York, 1943).
14 Y. V. Vorobyev, *Methods of Moments in Applied Mathematics* (Gordon and Breach, New York, 1965).
15 The functional Φ has the integral representation
$$\Phi\{R(u)\} = \int_{-\infty}^{\infty} (du/2\pi)\, R(u)\Lambda(u).$$
16 This theorem is due to E. Hellinger [*Math. Ann.* **86** (1922)]. The proof is reproduced in Ref. 12, p. 11. For reasons which will become clear below, we consider a function of the opposite sign than this in Ref. 12.
17 Speaking about a unique solution, we have in mind a solution for a given set of moments $\{s_n^{(v)}\}_0^\infty$. However, when the dynamical system under consideration contains two- (or many-) particle interactions, the moments $s_n^{(v)}$ will be given in terms of many-particle equal-time correlations. Thus, they by themselves are functionals of the SWF. As a result, one might obtain a number of different SWF's satisfying the same system of equations (2.4). This would furnish, for example, different free-energy functions, the intersection of which on the temperature scale would indicate a possible phase transition. However, since different values of the thermal averages lead to different sets of moments, we regard them as different moment problems.
18 Reference 13, p. 123.
19 Reference 12, p. 9.
20 The author owes this remark to the referee.
21 See, e.g., G. A. Baker, Jr., in *Advances in Theoretical Physics*, K. E. Brueckner, Ed. (Academic, New York, 1965), Vol. 1, p. 1.
22 Reference 12, Sec. 4. Also see D. Masson, *Hilbert Space and the Padé Approximant* (to be published).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

A Technique for Solving Recurrence Relations Approximately and Its Application to the 3-*J* and 6-*J* Symbols

Donald Neville
Department of Physics, Temple University, Philadelphia, Pennsylvania 19122
(Received 17 June 1971)

It is shown that not only some second-order differential equations, but also some second-order difference equations (three-term recurrence relations) possess solutions of WKBJ type. WKBJ-like solutions are derived and used to study the "classical" (large J) limits of both the 6-J symbol and (in an appendix) the 3-J symbol.

1. INTRODUCTION

As every physicist knows, in quantum mechanics the classical limit of the wavefunction is computed by solving Schrödinger's differential equation in WKBJ approximation.[1] Ordinarily one thinks of the WKBJ approach as applicable to functions depending on a continuous variable and obeying a differential equation. This is not the whole story, however, for in the present paper we shall apply the WKBJ approach to a function (the 6-J symbol) depending on a discrete variable and obeying a difference equation.

where a_k and b_k are the coefficients (3.6). It is well known that in case of a determinate moment problem, this fraction is converging and represents the Green function (1.8).[22] (A3) is known as the real J fraction. Consider, hence, the successive approximants of (A3), namely,

$$\frac{q_1(E)}{p_1(E)} = \frac{1}{E - a_0}, \quad \frac{q_2(E)}{p_2(E)} = \frac{1}{E - a_0 - b_0^2/(E - a_1)}, \cdots. \tag{A4}$$

It is easy to see that the numerators and denominators of these fractions satisfy the finite-difference equation

$$Y_{k+1} = (E - a_k)Y_k - b_{k-1}^2 Y_{k-1}, \quad (k = 1, 2, \ldots), \tag{A5}$$

with the initial conditions

$$q_0(E) = 0, \quad q_1(E) = 1, \quad p_0(E) = 1, \quad p_1(E) = E - a_0. \tag{A6}$$

Furthermore, the substitution

$$y_0 = y_0, \quad y_k = \frac{1}{b_0 b_1 \cdots b_{k-1}} Y_k, \quad k = 1, 2, \ldots \tag{A7}$$

transforms (A5) and (A6) into the finite-difference equation (3.5) with the initial conditions (3.7) and (3.8), the solution of which are, correspondingly, the polynomials of the first and second kind $P_k(E)$ and $Q_k(E)$. Since $q_N(E)/p_N(E)$ in (A4) represents the $[N, N-1]$ Padé approximant of (A3), we have

$$[N, N-1]G(E) = G_N(E) = Q_N(E)/P_N(E). \tag{A8}$$

Thus, on the complex G plane (see Fig. 1) the successive Padé approximants will be represented by the points of touch of the successive circular contours $A_1, A_2, A_3, \cdots$.

1 J. H. van Vleck, Phys. Rev. **74**, 1168 (1948).
2 G. Placzek, Phys. Rev. **86**, 377 (1952).
3 R. G. Gordon, J. Chem. Phys. **40**, 1973 (1964).
4 R. G. Gordon, J. Chem. Phys. **41**, 1819 (1964).
5 A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967).
6 R. G. Gordon, J. Math. Phys. **9**, 655 (1968).
7 R. G. Gordon, J. Math. Phys. **9**, 1087 (1968).
8 J. C. Wheeler and R. G. Gordon, J. Chem. Phys. **51**, 5566 (1969).
9 D. N. Zubarev, Usp. Fiz. Nauk **71**, 71 (1960) [Sov. Phys. Usp. 3, 320 (1960)].
10 D. K. Kalashnikov and E. S. Fradkin, Zh. Eksp. Teor. Fiz. **55**, 607 (1968) [Sov. Phys. JETP **28**, 317 (1969)].
11 In the sense that all Hankel forms $\sum_{i,k=0}^{m} s_{i+k} x_i x_k$ are positive definite.
12 N. I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1965).
13 J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, New York, 1943).
14 Y. V. Vorobyev, *Methods of Moments in Applied Mathematics* (Gordon and Breach, New York, 1965).
15 The functional Φ has the integral representation
$$\Phi\{R(u)\} = \int_{-\infty}^{\infty} (du/2\pi)\, R(u)\Lambda(u).$$
16 This theorem is due to E. Hellinger [*Math. Ann.* **86** (1922)]. The proof is reproduced in Ref. 12, p. 11. For reasons which will become clear below, we consider a function of the opposite sign than this in Ref. 12.
17 Speaking about a unique solution, we have in mind a solution for a given set of moments $\{s_n^{(v)}\}_0^\infty$. However, when the dynamical system under consideration contains two- (or many-) particle interactions, the moments $s_n^{(v)}$ will be given in terms of many-particle equal-time correlations. Thus, they by themselves are functionals of the SWF. As a result, one might obtain a number of different SWF's satisfying the same system of equations (2.4). This would furnish, for example, different free-energy functions, the intersection of which on the temperature scale would indicate a possible phase transition. However, since different values of the thermal averages lead to different sets of moments, we regard them as different moment problems.
18 Reference 13, p. 123.
19 Reference 12, p. 9.
20 The author owes this remark to the referee.
21 See, e.g., G. A. Baker, Jr., in *Advances in Theoretical Physics*, K. E. Brueckner, Ed. (Academic, New York, 1965), Vol. 1, p. 1.
22 Reference 12, Sec. 4. Also see D. Masson, *Hilbert Space and the Padé Approximant* (to be published).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

A Technique for Solving Recurrence Relations Approximately and Its Application to the 3-*J* and 6-*J* Symbols

Donald Neville
Department of Physics, Temple University, Philadelphia, Pennsylvania 19122
(Received 17 June 1971)

It is shown that not only some second-order differential equations, but also some second-order difference equations (three-term recurrence relations) possess solutions of WKBJ type. WKBJ-like solutions are derived and used to study the "classical" (large J) limits of both the 6-J symbol and (in an appendix) the 3-J symbol.

1. INTRODUCTION

As every physicist knows, in quantum mechanics the classical limit of the wavefunction is computed by solving Schrödinger's differential equation in WKBJ approximation.[1] Ordinarily one thinks of the WKBJ approach as applicable to functions depending on a continuous variable and obeying a differential equation. This is not the whole story, however, for in the present paper we shall apply the WKBJ approach to a function (the 6-J symbol) depending on a discrete variable and obeying a difference equation.

This difference equation is of the form

$$f_2\delta^{(2)}W + f_1\delta^{(1)}W + f_0W = 0, \tag{1.1}$$

where $W = W(J)$ is the 6-J symbol (symmetrized Racah coefficient)[2]

$$W(J) = \begin{Bmatrix} J_1 & J_2 & J \\ J_4 & J_3 & J_5 \end{Bmatrix}. \tag{1.2}$$

The f's are known functions of the J's, and the $\delta^{(n)}$'s are "finite difference operators." The latter are to the difference calculus what the d^n/dJ^n operators are to ordinary calculus:

$$\begin{aligned} \delta^{(1)}W &\equiv W(J+1) - W(J), \\ \delta^{(2)}W &\equiv [W(J+1) - W(J)] - [W(J) - W(J-1)], \\ &= W(J+1) - 2W(J) + W(J-1), \text{etc.} \end{aligned} \tag{1.3}$$

Equation (1.1) is strongly reminiscent of a second-order differential equation, the type that is usually solved in WKBJ approximation, and this fact, plus the quantum-mechanical nature of the 6-J symbol, is what got us started thinking of WKBJ-type solutions in the first place.

We say that a solution is of WKBJ type if it can be written in the form

$$W = Ae^{iS}, \tag{1.4}$$

where A and S can be expanded in the series of terms of successively lower order in J: $S = S_0 + S_1 + S_2 + \cdots$, with $S_{n+1}/S_n = \text{order } (1/J)$; and similarly for A. (In the usual WKBJ approximation, the small expansion parameter is $\hbar$; here it is $1/J$.) Just as in the Schrödinger case, W is "classically forbidden" for some values of J (that is, classical mechanics predicts that W is identically zero); and for these values S is complex and W is a decaying exponential. For other "clasically allowed" J values, S is real and W is oscillatory. Also, there are "turning points" or J values where solution (1.4) changes from oscillatory to exponential. Just as in the usual case, the region near turning points requires special treatment.

Also as in the usual case, the accuracy of the WKBJ solution is limited. In order to compute W with an error of order $1/J$, it is necessary to compute both S_0 and S_1, because S_2 is the first term in the expansion which is order $1/J$. (In contrast, A need be computed only to lowest order because it does not occur in an exponential.) S_0 is easily computed, but S_1 and succeeding terms are another matter.

Near a turning point the solution is not so phase-sensitive; however, the formula used to locate the turning point may introduce an error, by giving a value for the turning point off by a small integer.

In practice these inaccuracies may not be a serious drawback. In the classical limit, what is needed usually is an average of W over a small range of J values, rather than the value of W at a single J value, because the former is what is measured by a classical "coarse grained" apparatus:

$$\begin{aligned} \overline{W}^2 &= \sum_{J-\Delta}^{J+\Delta} W(J)^2/(2\Delta + 1), \\ &\cong A^2/2. \end{aligned}$$

The bar over $\overline{W}$ indicates a root-mean-square average, and the 2 in $A^2/2$ comes from the mean-square average of $\cos S$, supposedly a rapidly varying sinusoid. Thus, in the classical limit, the exact value of S may not be important.

Furthermore, the WKBJ solution has the advantage of compactness and simplicity. If one wants accuracy, one can calculate W exactly on a computer using (say) the power series for the 6-J symbol (Racah's sum).[3,4] But such an approach tends to bury the general features of the solution under an acre of computer printout.

So far as we can see, the present paper is important to 6-J theory especially for the light which it sheds on the behavior of W near turning points. Previously, Wigner had calculated A by a geometrical method (and as we have just seen, A may be all that is needed in the classical limit), but his approach does not work near turning points.[5] We remark that W has all its relative maxima at turning points.

Equation (1.1) can be rewritten as a three-term recurrence relation, rather than a second-order difference equation. If we eliminate the $\delta^{(n)}W$ using Eqs. (1.3), we get

$$g_+W(J+1) + g_0W(J) + g_-W(J-1) = 0, \tag{1.5}$$

where

$$\begin{aligned} g_+ &= f_2 + f_1, \\ g_0 &= f_0 - f_1 - 2f_2, \\ g_- &= f_2. \end{aligned} \tag{1.6}$$

Of course, the present method can be applied to difference equations (or recurrence relations) other than the one for the 6-J symbol. Sufficient conditions on the g's for this to be possible are

$$\lim_{J\to\infty} [(g_+ - g_-)/g_-] \leq \text{order } (1/J), \tag{1.7a}$$

$$\lim_{J\to\infty} (g_0/g_\pm) \leq \text{order unity}, \tag{1.7b}$$

$$\lim_{J\to\infty} \delta^{(n)}g/g = \text{order } J^{-n}. \tag{1.7c}$$

The last condition means that the g's behave like polynomials under differencing rather than like (say) sinusoids. Differencing lowers the order of a polynomial by one but leaves the order of a sinusoid ($\sin J$, say) unaffected, just as differentiating does. Often the g's are polynomials (or simple roots of polynomials) so that condition (1.7c) is usually satisfied in practice. Condition (1.7b) is the one that might be difficult to satisfy. If the recurrence relation does not satisfy condition (1.7a),

usually a change of dependent variable helps. For example, suppose $g_+ = (J-a)b$ and $g_- = 1$, where a and b are constants. Make the change of dependent variable $W = [b^J \Gamma(J-a)]^{-1/2} W'$, where $\Gamma(z)$ is the gamma function, satisfying $z\Gamma(z) = \Gamma(z+1)$. The recurrence relation for W' then satisfies condition (1.7a). More complicated cases, where the $g_\pm$ are polynomials or roots of polynomials, may be treated by induction on this simple example.

The present paper can be read both by those interested only in the general WKBJ method as well as by those interested only in the 6-J symbol. The former should concentrate on Secs. 2 and 4, and perhaps glance at Appendices B and C for hints on how to treat recurrence relations near zeros of $g_\pm$, where conditions (1.7) are often violated. The latter can skim Secs. 2–3 and perhaps skip Sec. 4 altogether, since the main results of Sec. 4 are summarized at the beginning of Sec. 5. Section 3 discusses two geometrical concepts, Wigner's tetrahedron and the (J, J_5) plane. We have found these quite useful for visualizing and organizing the algebraic and somewhat abstract results of Secs. 2 and 5. Appendix A derives the g's for the 6-J symbol from the Biedenharn–Elliott identity. Appendices B and C discuss the 6-J symbol near the maximum and minimum values of J allowed quantum-mechanically. At these extrema, conditions (1.7) are not satisfied and a separate treatment is necessary. The extrema are the only points where neither the turning point nor the WKBJ oscillating or exponential solutions work; therefore, the present paper, including Appendices B and C, provides an approximation scheme for every range of J values in the interval $\min J \le J \le \max J$ (so long as $\max J - \min J$ is large, of course). Appendix D is a mathematical footnote to Sec. 4. Appendix E applies the techniques of this paper to the 3-J symbol (symmetrized Clebsch–Gordan coefficient). We have concentrated on the 6-J symbol in the body of the text entirely for historical reasons: The techniques of the present paper were developed incidental to a study of isospin crossing matrices, which are essentially 6-J (or rather 6-I) symbols.[6] Appendix E shows that the present techniques work just as well when applied to a 3-J symbol.

2. THE WKBJ SOLUTION AWAY FROM TURNING POINTS

In this section for the most part we shall not specialize to the 6-J case, but shall assume merely that the g's obey conditions (1.7). However, it will not hurt to quote immediately the result we get when the formalism of this section is applied to the 6-J case:

$$A_0^2 = \text{const}\ [f_0(4f_2 - f_0)]^{-1/2}. \tag{2.1}$$

The formalism also gives us a difference equation for S_0 [Eq. (2.10)], which we interpret geometrically in Sec. 3.

We remark that conditions (1.7) on the g's imply conditions on the f's defined at Eq. (1.1). If we express the f's in terms of the g's by inverting Eqs. (1.6),

$$\begin{aligned} f_2 &= g_-, \\ f_1 &= g_+ - g_-, \\ f_0 &= g_+ + g_- + g_0, \end{aligned} \tag{2.2}$$

we find that the f's, like the g's, must be polynomial like under differencing; furthermore, conditions (1.7a) and (1.7b) imply

$$\lim_{J\to\infty} (f_1/f_2) \le \text{order } (1/J), \tag{2.3a}$$

$$\lim_{J\to\infty} (f_0/f_2) \le \text{order unity}. \tag{2.3b}$$

Now let us insert ansatz (1.4) into Eq. (1.1) and use the following identities for the difference of the product of two functions:

$$\delta^{(1)}(AB) = A\delta^{(1)}B + B\delta^{(1)}A + \delta^{(1)}A\delta^{(1)}B, \tag{2.4}$$

$$\begin{aligned} \delta^{(2)}(AB) = {} & A\delta^{(2)}B + B\delta^{(2)}B + 2\delta^{(1)}A\delta^{(1)}B \\ & - (\delta^{(2)}A\delta^{(1)}B + \delta^{(2)}B\delta^{(1)}A) \\ & + \delta^{(2)}A\delta^{(2)}B. \end{aligned} \tag{2.5}$$

These identities follow readily from the definitions (1.3). Setting $W = AB$ and using Eqs. (2.4) and (2.5), we get

$$\begin{aligned} & f_2[\delta^{(2)}A - \delta^{(1)}A + A]\delta^{(2)}B \\ & + [f_2(2\delta^{(1)}A - \delta^{(2)}A) + f_1(\delta^{(1)}A + A)]\delta^{(1)}B \\ & + [f_2\delta^{(2)}A + f_1\delta^{(1)}A + f_0A]B = 0. \end{aligned} \tag{2.6}$$

We now separate out the terms of highest-order in J. We may make the ansatz that A and S behave like polynomials under differencing, hence that A and B ($\equiv \exp iS$) in Eq. (2.6) behave like a polynomial and a sinusoid, respectively, under differencing. Therefore, $\delta^{(n)}A$ is n orders smaller than A, and $\delta^{(n)}B$ is the same order as B. Furthermore, f_1 is one order lower than f_2, from Eq. (2.3a). Therefore, the highest-order terms in Eq. (2.6) are the $f_2A\delta^{(2)}B$ and f_0AB terms. We get

$$f_2[\delta^{(2)}B]_{HO} + f_0[B]_{HO} = 0, \tag{2.7}$$

where $[\delta^{(2)}B]_{HO}$ means we are to difference $\mathrm{Re}B$ and throw away all but the highest-order term which results. The first and second differences of $\mathrm{Re}B \equiv \cos S$ follow from definitions (1.3) and some trigonometry [use the identities $\sin(\theta_1 \pm \theta_2) = \sin\theta_1\cos\theta_2 \pm \cos\theta_1\sin\theta_2$, $\cos(\theta_1 \pm \theta_2) = \cos\theta_1\cos\theta_2 \mp \sin\theta_1\sin\theta_2$]:

$$\delta^{(1)}\cos S = \cos S(\cos\delta^{(1)}S - 1) - \sin S\sin\delta^{(1)}S, \tag{2.8}$$

$$\begin{aligned} \delta^{(2)}\cos S = {} & 2\cos S(\cos\delta^{(1)}S - 1) + \cos(S - \delta^{(1)}S) \\ & \times (\cos\delta^{(2)}S - 1) - \sin(S - \delta^{(1)}S)\sin\delta^{(2)}S. \end{aligned} \tag{2.9}$$

We have to proceed cautiously in deciding which terms in Eq. (2.9) are of highest-order. If $S \geq$ order J^2, then $\delta^{(2)}S \geq$ order unity, and all terms in Eq. (2.9) are order unity. If all terms in Eq. (2.9) are order unity, then Eq. (2.7) becomes a complicated equation connecting $\delta^{(2)}S$ and $\delta^{(1)}S$ to tan$S$. Very likely all the S's satisfying such an equation would be sinusoidal rather than polynomial-like under differencing. We therefore try the ansatz $S \leq$ order J for large J. S must be at least order J, since if $S \leq$ order unity, $\exp(iS)$ would be approximately equal to $(1 + iS)$, which is polynomial-like under differencing. By process of elimination, therefore, S must be order J for large J, and this in turn implies that $[\delta^{(2)}B]_{HO}$ is the 2 cos$S$ term in Eq. (2.9) (with S replaced by $[S]_{HO} = S_0$ everywhere). Equation (2.7) becomes

$$\cos S[2f_2(\cos\delta^{(1)}S_0 - 1) + f_0] = 0$$

or

$$\delta^{(1)}S_0 = 2\arcsin(f_0/4f_2)^{1/2}. \tag{2.10}$$

We can check that this result is consistent with our ansatz that S_0 is order J and polynomial-like under differencing. Equation (2.10) and (2.3b) imply that $\sin\frac{1}{2}\delta^{(1)}S_0$ is order unity and polynomial-like under differencing. This does not look plausible at first glance, because $\sin\frac{1}{2}\delta^{(1)}S_0$ is a sinusoid, not a polynomial. However, a sinusoid will become polynomial-like when its argument is order unity and polynomial-like. [Proof: Use identity (2.8). If S on the left is order unity and polynomial-like, then $\delta^{(1)}S$ on the right is small, and $\sin\delta^{(1)}S \cong \delta^{(1)}S$, etc. Hence cos$S$ differences like a polynomial. QED] Hence Eq. (2.10) allows $\delta^{(1)}S_0$ to be order unity and polynomial-like. Therefore, S_0 itself can be order J and polynomial-like, consistent with our ansatz.

In order to obtain an equation for A_0, we must consider the next-to-highest order terms in Eqs. (2.6) and (2.7). These are (with $\sin\delta^{(2)}S \cong \delta^{(2)}S_0$)

$$-f_2A[2\cos S\sin\delta^{(1)}S_0\delta^{(1)}S_1 + \delta^{(2)}S_0\sin(S - \delta^{(1)}S)] - f_2\delta^{(1)}A[2\cos S(\cos\delta^{(1)}S - 1)]$$
$$+ (2f_2\delta^{(1)}A + f_1A)[\cos S(\cos\delta^{(1)}S - 1) - \sin S\sin\delta^{(1)}S] = 0. \tag{2.11}$$

The last square bracket comes from the $\delta^{(1)}B$ term in Eq. (2.6), the middle square bracket comes from the $\delta^{(2)}B$ term in Eq. (2.6), and the first square bracket comes from the lower-order corrections dropped on going from Eq. (2.7) to Eq. (2.10). Equation (2.11) can be separated into a cos$S$ and a sin$S$ term:

$$A\cos S[(-2\sin\delta^{(1)}S_0\delta^{(1)}S_1 + \delta^{(2)}S_0\sin\delta^{(1)}S)f_2$$
$$+ (\cos\delta^{(1)}S - 1)f_1] + \sin S[-(\delta^{(2)}S_0\cos\delta^{(1)}S)f_2A$$
$$-(2\delta^{(1)}Af_2 + Af_1)\sin\delta^{(1)}S] = 0. \tag{2.12}$$

Now suppose that the second square bracket is not identically zero, and imagine (2.12) divided through by $A\cos S$ times this bracket. The equation would then say that a function with sinusoidal behavior under differencing (tan$S$) equals a function with polynomial behavior. This is a contradiction; hence both square brackets in Eq. (2.12) must vanish. This gives us two equations, one for higher-order corrections to S_0 (we ignore this equation) and one for A_0:

$$2\delta^{(1)}A_0/A_0 = -\delta^{(2)}S_0\cot\delta^{(1)}S_0 - f_1/f_2. \tag{2.13}$$

This A_0 is polynomial-like under differencing, consistent with our original ansatz.

Equations (2.10) and (2.13) are two first-order difference equations which replace the single second-order difference equation (1.1). They are the basic equations of the WKBJ approach.

We can obtain an approximate solution to these equations as follows. Make a Taylor series expansion of $\delta^{(n)}A_0$ and $\delta^{(n)}S_0$,

$$\delta^{(1)}A_0 = A_0(J+1) - A_0(J)$$
$$= \frac{dA_0(J)}{dJ} + \left(\frac{1}{2!}\right)\frac{d^2A_0(J)}{dJ^2} + \cdots$$

and similarly for the other differences of A_0 and S_0. Since A_0 and S_0 are polynomial-like, successive derivatives will be down by factors of $(1/J)$; hence, only the highest derivatives need be kept:

$$\delta^{(1)}A_0 \cong \frac{dA_0}{dJ}, \quad \text{etc;} \tag{2.14}$$

i.e., we can replace differences by derivatives everywhere in Eqs. (2.10) and (2.13). In simple cases, these equations can then be solved immediately by integration. In the 6-J case, for example, where $f_1 = \delta^{(1)}f_2$ (cf. Appendix A, Eq. (A9), Eq. (2.13) gives

$$2\ln A_0 = -\ln[\sin\delta^{(1)}S_0] - \ln f_2 + \ln(C/2),$$

where C is an integration constant. Substituting for $\sin\delta^{(1)}S_0$ from Eq. (2.10), we get

$$A_0^2 \cong C[f_0(4f_2 - f_0)]^{-1/2}. \tag{2.15}$$

Even when $f_1 \neq \delta^{(1)}f_2$, we will get a solution of type (2.15) if $f_1 \cong (p+1)df_2/dJ$ to order J, where p is some constant:

$$A_0^2 \cong C[f_0(4f_2 - f_0)]^{-1/2}f_2^{-p}. \tag{2.16}$$

Evidently the solution breaks down when f_0 equals either zero or f_2 [or when $f_2 \equiv g-$ vanishes; but this possibility is excluded by conditions (1.7a) and (1.7b)]. In other words, the solution breaks

down at turning points, since, at $f_0 = 0$ or $4f_2$, W changes from oscillatory to exponential [cf. Eq. (2.10)]. After a digression into the geometry of the 6-J symbol (Sec. 3), we shall return to study W near turning points in Sec. 4.

3. GEOMETRICAL INTERPRETATION OF THE 6-J SYMBOL

In this section we restrict W to be the 6-J symbol and Eq. (1.5) to be the recurrence relation for the 6-J symbol derived in Appendix A. We shall sketch Wigner's geometrical interpretation of this symbol and then derive a geometrical interpretation of the g's, f's, and S_0. This interpretation is useful for locating "turning points" of W quickly. In addition, Wigner's result for A_0, which follows readily from the geometry, can be used as a "boundary condition" to fix the constant in Eq. (2.1) and the normalization at turning points in Sec. 4.

We start from the interpretation of W as a probability amplitude for adding four angular momentum vectors $\mathbf{J}_1$–$\mathbf{J}_4$ to form a resultant angular momentum of zero. Let $|J_{13}, J_{24}\rangle^{(0)}$ be the ket for the state in which these four vectors have been summed to zero in the order

$$\mathbf{J}_1 + \mathbf{J}_3 = \mathbf{J}_{13}, \quad \mathbf{J}_2 + \mathbf{J}_4 = \mathbf{J}_{24}, \quad \mathbf{J}_{13} + \mathbf{J}_{24} = 0.$$

From the last equations, necessarily $J_{13} = J_{24}$, and, if we denote $\mathbf{J}_{13} = \mathbf{J}_{24}$ by $\mathbf{J}_5$, we have the situation diagrammed in Fig. 1. Then, again referring to Fig. 1, we have that

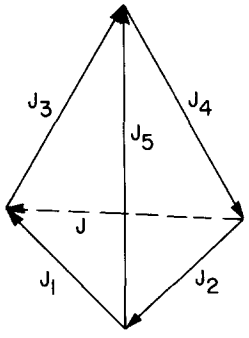

FIG. 1. Wigner's 6-J tetrahedron.

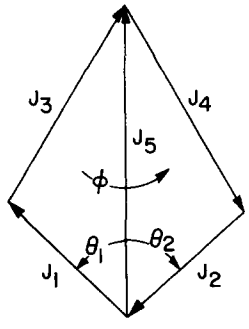

FIG. 2. $\mathbf{J}_5$ as the z axis of a spherical coordinate system. ϕ is azimuth and $\mathbf{J}_1$ lies in the xz plane. For clarity J is not shown.

$$P(J) = |\langle J_{12}J_{34}|J_{13}J_{24}\rangle^{(0)}|^2 \tag{3.1}$$

is the quantum-mechanical probability that the vector $\mathbf{J} = \mathbf{J}_{12} = \mathbf{J}_{34}$ will have length J, given that vectors $\mathbf{J}_1$–$\mathbf{J}_5$ have lengths J_1–J_5 and add to zero as shown in Fig. 1. Evidently this probability is a function of the 6 J's $J_1, \ldots, J_5, J$, and the function W is defined and normalized so that

$$(2J + 1)(2J_5 + 1)|W|^2 = P(J). \tag{3.2}$$

Now if we imagine that we are adding together very large vectors J, we can set $P(J)$ equal to the *classical* value for this probability, and thereby obtain an equation for $|W|^2$, [5]

$$P(J) \to 2\Delta\phi/2\pi \quad \text{or zero.} \tag{3.3}$$

$\phi = \phi(J)$ is the angle between the planes $J_1J_3J_5$ and $J_2J_4J_5$ in Fig. 1 or Fig. 2, and it is not fixed even after all lengths J_1–J_5 are fixed. Every value $0 \le \phi \le 2\pi$ is equally probable; hence $P(J)$, the probability that J will have length J, is proportional to $\Delta\phi$, the increment in ϕ as $\mathbf{J}$ changes in length from J to $J + 1$. There is a 2 in the numerator of Eq. (3.3) because ϕ passes through a given J twice between 0 and 2π.

$P(J)$ is zero, classically, if J is so large or so small that the tetrahedron of Fig. 1 cannot be formed. To see how this might happen, imagine that in Fig. 1 the (dotted) vector $\mathbf{J}$ is increased in length indefinitely, the other vectors remaining fixed in length. The sides $J_1J_3J_5$ and $J_2J_4J_5$ of the tetrahedron will pivot about J_5 as common axis until eventually the tetrahedron flattens to a quadrilateral. At this point, J still has not reached the largest value allowed quantum-mechanically [$\max J = \min(J_1 + J_2, J_3 + J_4)$]; but, classically, larger values of J are not allowed, as further increases in J would break the figure. The classical $P(J)$ is zero for those J's which break the figure. The quantum mechanical $P(J)$ is not zero, but we may expect that *any J-value at which the tetrahedron flattens to a figure of lower dimension corresponds to a "turning point" of the WKBJ solution*, where W changes from oscillatory to damped exponential in character. In the previous section we found that turning points correspond to zeros of f_0 or $4f_2 - f_0$, and later on in this section we shall find that such zeros indeed do correspond to zeros of V, the volume of the 6-J tetrahedron.

But first let us complete the calculation of the classical $P(J)$, Eq. (3.3). After several lines of vector algebra, one finds $\cos\phi$ in terms of the angles given in Fig. 2:

$$\cos\phi = \frac{J_1^2 + J_2^2 - 2J_1J_2\cos\theta_1\cos\theta_2 - J^2}{2J_1J_2\sin\theta_1\sin\theta_2}. \tag{3.4}$$

(Strictly speaking $\mathbf{J}^2$ is $J(J + 1)$, not J^2; we are approximating $J(J + 1)$ by J^2 throughout this section.) Equation (3.4) can be differentiated to give

an implicit expression for $d\phi/dJ$. Then one can calculate $\Delta\phi$,

$$\begin{aligned}\Delta\phi &= \int_J^{J+1} \frac{d\phi}{dx}\, dx \\ &= \int [2J_1J_2 \sin\theta_1 \sin\theta_2 \sin\phi(x)]^{-1}\, 2x dx \\ &\cong [2J_1J_2 \sin\theta_1 \sin\theta_2 \sin\phi(J)]^{-1} \int 2x dx \\ &\cong [2J_1J_2 \sin\theta_1 \sin\theta_2 \sin\phi]^{-1} J. \end{aligned} \tag{3.5}$$

In the second from last step we have assumed that $\phi(x)$ varies slowly between $x = J$ and $x = J + 1$, so that $\sin\phi(x) \cong \sin\phi(J)$. [This is not true for $\phi \cong 0$ or π, where $d\phi/dx \propto (\sin\phi)^{-1}$ which blows up: The present calculation breaks down at turning points.] In the last step we have taken $2J + 1 \cong 2J$. The square bracket in Eq. (3.5) is essentially the volume V of the 6-J tetrahedron:

$$6V = |\mathbf{J}_1 \times \mathbf{J}_2 \cdot \mathbf{J}_5| = J_1J_2J_5 \sin\theta_1 \sin\theta_2 \sin\phi. \tag{3.6}$$

Hence $\Delta\phi \cong JJ_5/6V$. Combining this with Eqs. (3.2) and (3.3) and taking a square root, we get

$$\overline{W} = (24\pi V)^{-1/2} \tag{3.7}$$

or zero, for ϕ not too close to 0 or π where the $V^{-1/2}$ blows up. The bar over W is a reminder that $\overline{W}$ is a root-mean-square average over a small range of values of J, as in the Introduction:

$$\overline{W}(J) = \Big[(2\Delta + 1)^{-1} \sum_{x=J-\Delta}^{J+\Delta} |W(x)|^2\Big]^{1/2}. \tag{3.7'}$$

Now let us interpret the g's and f's geometrically, and relate V to the f's. For the remainder of the section, we shall switch from $\mathbf{J}_5$ to $\mathbf{J}$ as z axis, and call our new azimuthal and polar angles ξ and ψ to avoid confusing them with the old ϕ and θ (see Fig. 3).

By repeated application of the law of cosines, one can calculate the cosine and sine of ψ_2 and ψ_4 in terms of the lengths of the edges; for example,

$$2JJ_2 \cos\psi_2 = J^2 + J_2^2 - J_1^2, \tag{3.8}$$

$$\begin{aligned}(2JJ_2 \sin\psi_2)^2 = {} & (J + J_2 - J_1)(J + J_2 + J_1) \\ & \times (-J + J_2 + J_1)(J - J_2 + J_1).\end{aligned} \tag{3.9}$$

Therefore, when the J's are large,

$$g_+ \cong g_- \cong 4JJ_2J_4 \sin\psi_2 \sin\psi_4 \tag{3.10}$$

to order J^2. For g_0 we get a slightly more complicated expression,

$$g_0 \cong 8JJ_2J_4(-\cos\psi_2 \cos\psi_4 + \cos\psi_{24}), \tag{3.11}$$

where $\cos\psi_{24}$ is the angle between $\mathbf{J}_2$ and $\mathbf{J}_4$. We need to express this last cosine in terms of the ψ's and ξ. Let $\hat{J}$ be a unit vector along $\mathbf{J}$, etc.; then

$$\begin{aligned}\cos\psi_{24} &= \hat{J}_2 \cdot \hat{J}_4 \\ &= (\sin\psi_2, 0, \cos\psi_2) \\ &\quad \cdot (\sin\psi_4 \cos\xi, \sin\psi_4 \sin\xi, \cos\psi_4) \\ &= \sin\psi_2 \sin\psi_4 \cos\xi + \cos\psi_2 \cos\psi_4.\end{aligned}$$

Using this expression to eliminate $\cos\psi_{24}$, we get

$$g_0 \cong 8JJ_2J_4 \sin\psi_2 \sin\psi_4 \cos\xi. \tag{3.12}$$

A geometrical interpretation for the f's now follows from relations (2.2):

$$\begin{aligned}f_2 &\cong 4JJ_2J_4 \sin\psi_2 \sin\psi_4, \\ f_1 &\cong 0, \\ f_0 &\cong 8JJ_2J_4 \sin\psi_2 \sin\psi_4(1 + \cos\xi).\end{aligned} \tag{3.13}$$

Now we can express V in terms of the f's. At Eq. (3.6) we computed V from $\mathbf{J}_1 \times \mathbf{J}_2 \cdot \mathbf{J}_5$. We could equally well have used $\mathbf{J}_2 \times \mathbf{J}_4 \cdot \mathbf{J}$ which would have given us

$$\begin{aligned}6V &= JJ_2J_4 \sin\psi_2 \sin\psi_4 \sin\xi \\ &= JJ_2J_4 \sin\psi_2 \sin\psi_4 (1 + \cos\xi)^{1/2} (1 - \cos\xi)^{1/2}\end{aligned}$$

or

$$6V \cong [f_0(4f_2 - f_0)]^{1/2}/8. \tag{3.14}$$

Equation (3.14) establishes our contention that turning points of the WKBJ solution ($f_0 = 0$ or $f_0 = 4f_2$) correspond to points where the 6-J tetrahedron collapses to a figure of lower dimension ($V \to 0$). Inserting Eq. (3.14) into Eq. (3.7), we get

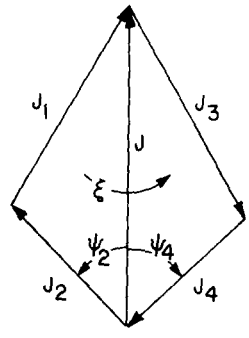

FIG. 3. J as the z axis of a spherical coordinate system. ξ is azimuth and $\mathbf{J}_1$ lies in the xz plane. For clarity $\mathbf{J}_5$ is not shown.

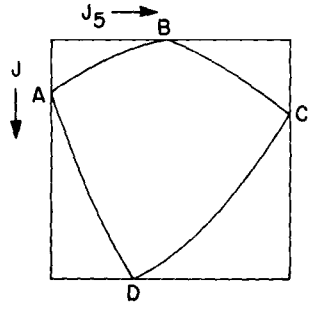

FIG. 4. Boundaries of the classically allowed region in the (J, J_5) plane.

a formula for the rms value of W in the classically allowed region,

$$\overline{W} = (2/\pi)^{1/2}[f_0(4f_2 - f_0)]^{-1/4}. \tag{3.15}$$

If we recall that $\overline{W^2}$, the mean-square value of W, is $A_0^2/2$, it follows that the constant in Eq. (2.1) is $(4/\pi)$.

Formulas (3.13) for f_0 and f_2 give us a geometrical interpretation for $\delta^{(1)}S_0$, Eq. (2.10):

$$\begin{aligned}\delta^{(1)}S_0 &= 2\arcsin[(1+\cos\xi)/2]^{1/2} \\ &= 2\arcsin[\cos(\xi/2)] \\ &= 2\arcsin[\sin(\xi/2+\pi/2)]\end{aligned} \tag{3.16}$$

or

$$\delta^{(1)}S_0 = \xi + \pi \cong dS_0/dJ.$$

A useful way to visualize the classically allowed region is to think of W as a square matrix array depending on *two* variables, a row index J and a column index J_5, and then plot the boundaries of the classically allowed region in (J, J_5) space. Thinking of W as a function of two variables is natural in the many applications where J_1–J_4 are held fixed and W is essentially a matrix transformation from states labeled by J_5 to states labeled by J. In Fig. 4, we show the part of the (J, J_5) plane bounded by the four straight lines $J = \max J$, $J = \min J$, $J_5 = \max J_5$, $J_5 = \min J_5$. In addition, we show (schematically) the boundaries of the classically allowed region for a case in which no two of the vectors J_1–J_4 are equal. The classically allowed region is bounded by the four curves AB, BC, CD, DA. These are the curves where f_0 or $4f_2 - f_0 = 0$ [from Eq. (3.15)] or, equivalently, where $\cos\xi = \pm 1$ [from Eqs. (3.13)]. As one moves from the center out across any one of these boundaries toward the edges in Fig. 4, one enters a classically forbidden region and the elements begin to decrease exponentially.

In the most general case (no two vectors J_1–J_4 equal) the classically allowed region touches each edge at only one point, as shown in Fig. 4. In special cases where two or more of J_1–J_4 are equal, however, two or more of the points $ABCD$ may move to corners, with the result that an entire edge $J_5 = \min J_5$ or $J = \min J$ becomes classically allowed. (The boundary curves never intersect the edges $J = \max J$ or $J_5 = \max J_5$ at more than one point, however.) We leave it to the interested reader to verify the foregoing statements by experimenting with the geometry of the 6-J tetrahedron, Fig. 3.

4. THE SOLUTION NEAR TURNING POINTS

After this digression into geometry for the special case W = 6-J symbol, we now return to the general case and solve Eq. (1.1) near turning points ($f_0 = 0$ or $f_0 = 4f_2$). We continue to assume that conditions (1.7) or (2.3) are satisfied, even near turning points, and in particular we assume that f_2 ($\equiv g$ –) is nonzero. (For a discussion of the case $f_2 \cong 0$, see Appendices B and C.) In addition, we assume that the functions f_0 and $f_0 - 4f_2$ have zeros of *integer* order at their turning points J_0 [that is, the functions vanish as $J - J_0)^n$, n = integer], though we do make a few remarks about how one might handle zeros of noninteger order. (In the 6-J case f_0 and $f_0 - 4f_2$ vanish linearly at their turning points: $n = 1$.)

If Eq. (1.1) were a differential equation, one would obtain solutions near turning points by changing the dependent variable so as to get a wave equation $d^2W/dJ^2 + k^2(J)W = 0$. Then one would approximate $k^2 \cong \text{const} \times (J - J_0)^n$ near the turning-point J_0. Finally, one would solve the resulting equation exactly and get a Bessel function. Something quite similar works for the difference equation. The solution in the difference case also turns out to be a Bessel function, although not always the same Bessel function as in the differential case.

In order to get Eq. (1.1) into "wave equation" form (i.e., eliminate the first difference term), we substitute

$$W \equiv RV, \tag{4.1}$$

so that Eq. (1.1) becomes Eq. (2.6) with $A, B \to R, V$. We then demand that the coefficient of $\delta^{(1)}V$ in Eq. (2.6) be identically zero. This determines R:

$$f_2(2\delta^{(1)}R - \delta^{(2)}R) + f_1(\delta^{(1)}R + R) = 0, \tag{4.2}$$

and Eq. (2.6) becomes a "wave equation" for V:

$$\begin{aligned}f_2[\delta^{(2)}R - \delta^{(1)}R + R]\delta^{(2)}V \\ + [f_2\delta^{(2)}R + f_1\delta^{(1)}R + f_0R]V = 0.\end{aligned} \tag{4.3}$$

Of course, the coefficients in Eq. (4.3) depend on R, so that we must find a solution to Eq. (4.2) before we can solve Eq. (4.3). We shall try to find an approximate solution, valid to order $1/J$. We try the ansatz that at least one solution to Eq. (4.2) is polynomial-like under differencing. Since f_1 is one order smaller than f_2, according to condition (2.3a), the $f_2\delta^{(2)}R$ and $f_1\delta^{(1)}R$ terms are negligible in Eq. (4.2). This equation then collapses to a first-order equation

$$\delta^{(1)}R/R \cong -f_1/2f_2. \tag{4.4}$$

Clearly this equation is consistent with our ansatz that R is polynomial-like, since f_1/f_2 = order $(1/J)$ from condition (2.3a). Further, since the difference $\delta^{(1)}R$ is small compared to R, $\delta^{(1)}$ can be replaced by d/dJ as was done in Eq. (2.14), and Eq. (4.4) can be integrated to obtain the desired approximate solution. In the 6-J case, where $f_1 = \delta^{(1)}f_2 \cong df_2/dJ$, this solution is

$$R \cong (f_2)^{-1/2}, \tag{4.5}$$

where we have set a constant of integration equal to one, without loss of generality.

An exact solution to Eq. (4.2) can also be obtained in the 6-J case, or in any case where $[f_2/(f_2+f_1)]^n$ is a ratio of polynomials for some choice of n. For the details see Appendix D.

Now we return to Eq. (4.3) and eliminate R from it. We write the second square bracket of this equation as

$$f_2 R[\delta^{(2)}R/R + (f_1/f_2)(\delta^{(1)}R/R) + (f_0/f_2)].$$

Evidently, from Eq. (4.4) and the polynomial-like nature of R_1, the first two terms in the bracket are order $(1/J^2)$ and can be neglected. The third could be small also, though perhaps not negligibly small like the first two, because we are supposing ourselves near turning points, and one way to get a turning point is to have f_0 vanish. (The other way is to have $f_0 = 4f_2$; we shall consider this case in the second from last paragraph of this section.) We assume $f_0 \cong 0$ and expand (f_0/f_2) in a Taylor series about its zero:

$$(f_0/f_2) \cong z^n/d,$$

where z is the distance from turning point J_0, n is the order of the zero of f_0 at J_0, and d is a constant:

$$z \equiv J - J_0, \tag{4.6}$$

$$f_0(J_0) = 0, \tag{4.7}$$

$$1/d = \lim_{z\to 0} (f_0/f_2 z^n). \tag{4.8}$$

Since $f_0/f_2 = (g_+ + g_- + g_0)/g_- \cong 2 + g_0/g_-$ from Eq. (2.2) and condition (1.7a), g_0/g_- and therefore f_0/f_2 must be dimensionless near the zero, which means $(f_0/f_2) =$ order $(z/J)^n$, or

$$d = \text{order}\,(J^n). \tag{4.9}$$

Hence the (f_0/f_2) term is order J^{-n}, small but not necessarily negligible. Dividing through Eq. (4.3) by the coefficient of $\delta^{(2)}V$ $(\approx f_2 R)$ and neglecting all but highest-order terms, we get

$$\delta^{(2)}V + (z^n/d)V = \text{order}\,(V/J^{n+1}, V/J^2), \tag{4.10}$$

where the $(z^n/d)V$ term is order V/J^n, larger than those neglected. [Of course, if $n \geq 2$, the $(z^n/d)V$ term is as small as the (V/J^2) terms we have neglected, and Eq. (4.10) reduces to $\delta^{(2)}V \cong 0$, $V \cong$ linear function of z. Therefore, $n \geq 2$ is a trivial case, and from now on we need consider only the nontrivial case $0 < n < 2$.] From now on we drop the terms on the right-hand side of Eq. (4.10), replacing them by zero.

If n is not an integer (if $n \neq 1$), the only way we know of to solve Eq. (4.10) is to approximate $\delta^{(2)}V$ by d^2V/dJ^2; the solutions to the resultant differential equation are essentially Bessel functions

$$V = c_\pm z^{1/2} J_{\pm p}[2p(-z^{1/p}/d)^{1/2}],$$
$$P = 1/(n+2). \tag{4.11}$$

Equation (4.11) is just the usual WKBJ solution to the Schrödinger equation, discussed in many quantum mechanics text books.[1] The approximation $\delta^{(2)}V \cong d^2V/dJ^2$ has to be scrutinized carefully, however, because conceivably it could lead to errors of the same order as the $(z^n/d)V$ term in Eq. (4.10). By an argument similar to the one used in Eq. (2.14), $\delta^{(2)}V - d^2V/dJ^2$ is order V/J^2, provided V is polynomial-like under differencing. In the present case, V is *not* polynomial-like, however, since $\delta^{(2)}V/V$ would then be order $1/J^2$ rather than order $z^n/d =$ order $1/J^n$. Hence a careful discussion, which we do not give, is needed to establish the limits of validity of solution (4.11) when $n \neq$ integer.

If $n = 1$, Eq. (4.10) can be solved as it stands without replacing $\delta^{(2)}$ by d^2/dJ^2. The solutions are Bessel functions, though not the Bessel functions of Eq. (4.11). To show that V is a Bessel function for $n = 1$, we start from the recurrence relation

$$Z_{p+1}(\omega) + Z_{p-1}(\omega) = (2p/\omega) Z_p(\omega), \tag{4.12}$$

which is satisfied for Z, a Bessel, Hankel, or Neumann function. Subtracting $2Z_p(\omega)$ from both sides of this equation, we get a second difference on the left:

$$\delta_p^{(2)} Z_p(\omega) = 2(p-\omega) Z_p(\omega)/\omega. \tag{4.13}$$

The subscript on $\delta_p^{(2)}$ indicates we are differencing with respect to the index, not the argument. Now taking

$$\omega = -2d,$$
$$p = z - 2d, \tag{4.14}$$
$$\delta_p^{(n)} = \delta_z^{(n)} \equiv \delta^{(n)},$$

we find

$$\delta^{(2)} Z_{z-2d}(-2d) = -(z/d) Z_{z-2d}(-2d), \tag{4.15}$$

which is just Eq. (4.10) for $n = 1$. Therefore, near $z = 0$,

$$V(z) = c_1 J_{z-2d}(-2d) + c_2 N_{z-2d}(-2d), \quad d < 0. \tag{4.16}$$

We should explain why we impose the condition $d < 0$ at Eq. (4.16). The next step in determining the solution near turning points is to expand Eq. (4.11) or (4.16) asymptotically in both directions from the turning point $(z \to \pm\infty)$ and join it up to the WKBJ solutions found in Sec. II. The functions occuring in Eq. (4.11) can be expanded readily enough using Hankel's asymptotic series for functions of small order but large argument. The functions occuring in Eq. (4.16), however, have both large order $(z - 2d)$ and large argument $(-2d)$. In such a case it is necessary to use Debye's series rather than Hankel's.[7] In the literature, Debye's series are often quoted only for the case in which the order $(z - 2d)$ and the argument $(-2d)$

are both positive. If d as defined in Eq. (4.8) turns out to be positive, it is therefore more convenient to use

$$V(z) = c_1 J_{-z+2d}(+2d) + c_2 N_{-z+2d}(+2d), \quad d > 0, \tag{4.17}$$

as our solution. Equation (4.17) is equivalent to Eq. (4.16) because $J_p(x)$ and $N_p(x)$ can be expressed as linear combinations of $J_{-p}(-x)$ and $N_{-p}(-x)$.

So far we have considered only turning points $f_0 = 0$. If we are near a turning point where $f_0 = 4f_2$, we change the dependent variable from W to W' defined by

$$W(J) = (-1)^J W'(J). \tag{4.18}$$

By direct substitution into Eq. (1.5) it is easy to see that $W'(J)$ obeys the same recurrence relation as W, except for $g_0 \to -g_0$ everywhere. Hence W' obeys the same difference equation as W, except that in Eqs. (2.2)

$$f_0 \to f_0' \equiv g_+ + g_- - g_0. \tag{4.19}$$

Using definitions (2.2) and condition (1.7a), we verify that f_0' vanishes at turning points $f_0 - 4f_2 = 0$:

$$\begin{aligned} f_0' &= f_0 - 2g_- - 2g_+ \\ &= f_0 - 2g_- - 2g_- + 2(g_- - g_+) \\ &= f_0 - 4g_-[1 + \text{order } (1/J)] \\ &\cong f_0 - 4f_2. \end{aligned} \tag{4.20}$$

Hence near turning points $f_0 - 4f_2 = 0$ we can repeat the whole of the preceding discussion with f_0 replaced by f_0', W replaced by W', and d replaced by a quantity d' defined by

$$1/d' = \lim_{z \to 0} (f_0'/f_2 z^n). \tag{4.21}$$

To summarize, the solution near turning points $f_0 = (J - J_0)^n/d$ is $W = RV$ with V determined by Eq. (4.16) for $n = 1$ and d negative, or Eq. (4.17) for $n = 1$ and d positive, and R determined by Eq. (4.4). For $n = 2, 3, 4, \ldots$, V is a linear function of z. The solution near turning points $f_0 = 4f_2$ is the same as that near turning points $f_0 = 0$, except $W = (-1)^J RV$; and, in V, d is replaced by d', the quantity defined at Eqs. (4.19) and (4.21).

5. THE 6-J SYMBOL NEAR TURNING POINTS

Now let us apply the theory of Sec. IV to the 6-J symbol. So far we know that W is of the form $W = RV$ near $f_0 = 0$ turning points, and $W = (-1)^J RV$ near $f_0 = 4f_2$ turning points. R for the 6-J case has been determined already at Eq. (4.5). V has been determined up to the constants c_i at Eq. (4.16) or (4.17). What remains is to determine these constants. As discussed in the previous section, Eqs. (4.16) and (4.17) apply only at $f_0 = 0$ turning points. At $f_0 = 4f_2$ turning points, the d in the equations is to be replaced by d', the quantity defined at Eqs. (4.19) and (4.21). At the end of this section we use the geometrical theory of Sec. 3 to distinguish the J-values where $f_0 = 0$ from those where $f_0 = 4f_2$, and the J-values where d or d' is positive from those where d or d' is negative.

To determine c_2, we evaluate Eq. (4.16) for $z(-2d)^{-1/3}$ large and positive [or Eq. (4.17) for $-z(+2d)^{1/3}$ large and positive].[7] This region of z-values is "classically forbidden" for the 6-J symbol. The Neumann function is a diverging exponential in this region; hence $c_2 = 0$.

To determine $|c_1|$, we evaluate Eq. (4.16) for $z(-2d)^{-1/3}$ large and negative (a "classically allowed" region for the 6-J symbol), join it up with the WKBJ solution there, and impose Wigner's requirement that the rms value of W must be Eq. (3.15) in a classically allowed region. According to Debye, for $z < 0$ and $-z > 3(z - 2d)^{1/3}$ ($|z|$ large but not necessarily order J),[7]

$$\begin{aligned} &J_{z-2d}(-2d) \\ &\quad = \sin\{-2d[\cos\alpha + (\alpha - \pi/2)\sin\alpha] + \pi/4\} \\ &\qquad \times[(\pi d \cos\alpha)^{-1/2} + \text{order } (d\cos\alpha)^{-3/2}], \end{aligned} \tag{5.1}$$

where

$$\begin{aligned} \sin\alpha &\equiv 1 - z/2d, \\ \cos\alpha &= -[z/d - (z/2d)^2]^{1/2}. \end{aligned} \tag{5.2}$$

The region of validity of Eq. (5.1) begins at $|z| =$ order $(-2d)^{1/3}$; hence [from Eq. (4.9) with $n = 1$] $-z/2d =$ order $J^{-2/3}$, still small, and $\alpha \cong \pi/2$. Therefore,

$$\begin{aligned} \cos\alpha &\cong -(-z/-d)^{1/2} \\ &= \sin(\pi/2 - \alpha) \\ &\cong (\pi/2 - \alpha) - (\pi/2 - \alpha)^3/6, \\ \sin\alpha &= \cos(\pi/2 - \alpha) \cong 1 - (\pi/2 - \alpha)^2/2, \\ \cos\alpha &- (\pi/2 - \alpha)\sin\alpha \cong (\pi/2 - \alpha)^3/3. \end{aligned} \tag{5.3}$$

Relations (5.3) can be used to eliminate α from expression (5.1)

$$J_{z-2d}(-2d) \cong \frac{\sin[-\frac{2}{3}(-z)^{3/2}(-d)^{-1/2} + \pi/4]}{[\pi^{1/2}(-d)^{1/4}(-z)^{1/4}]}. \tag{5.4}$$

[The $(-z)^{3/2}$ behavior of this phase checks with that expected from the WKBJ solution. Near the zero of f_0, the phase S_0 of the WKBJ solution is determined by Eq. (2.10) to be

$$\begin{aligned} \tfrac{1}{2}\delta^{(1)} S_0 &= \arcsin(f_0/4f_2)^{1/2} \\ &\cong (f_0/4f_2)^{1/2} \cong \tfrac{1}{2}(-z/-d)^{1/2} \end{aligned}$$

or

$$\begin{aligned} S_0(z) &= \textstyle\sum_z \delta^{(1)} S^{(1)} S_0(z) \cong \int^z \delta^{(1)} S_0 dz \\ &= \int (-z/-d)^{1/2} \\ &= -\tfrac{2}{3}(-z)^{3/2}(-d)^{-1/2} + \text{const.}] \end{aligned}$$

Now we get an equation for $|c_1|$ by putting together Eqs. (3.15), (4.1), (4.5), (4.16) with $c_2 = 0$, and (5.4) with the sine replaced by its rms average $(1/2)^{1/2}$:

$$(2/\pi)^{1/2}[f_0(4f_2 - f_0)]^{-1/4} = [|c_1|(f_2)^{-1/2}(1/2)^{1/2}]/[\pi^{1/2}(-d)^{1/4}(-z)^{1/4}].$$

On the left we remember that $(f_0/f_2) \cong -z/-d \ll 1$ [Eq. (4.8), and Eq. (5.2)ff.] The square bracket then simplifies to $(2f_2)^{-1/2}(-z/-d)^{-1/4}$ and we find $|c_1| = (-2d)^{1/2}$.

Therefore, near turning points $f_0 = 0, d < 0$,

$$W = e^{i\delta}(-2d/f_2)^{1/2} J_{z-2d}(-2d), \quad d < 0, \qquad (5.5)$$

where δ, the phase of c_1, is as yet undetermined.

For a turning point with $f_0 = 0$ but d positive, replace z and d in Eq. (5.5) by $-z$ and $-d$.

For a turning point with $f_0 = 4f_2$ and $d' \lessgtr 0$, replace z and d in Eq. (5.5) by $\pm z$ and $\pm d'$, and multiply Eq. (5.5) by $(-1)^J$.

We emphasize that *one* formula, Eq. (5.5), works on *both* sides of the turning point $z = 0$: There is no need to use Bessel functions of real argument on one side, Bessel functions of imaginary argument on the other, and then match them at a boundary, the way one must in solving a differential equation by the WKBJ method. In this respect, the solution to the difference equation is actually simpler than that to the differential equation.

Next we determine δ. This could be done easily enough if we could continue Eq. (5.5) to either of the values $J = \min J$ or $J = \max J$. At these extreme values of J, Racah's sum for the 6-J symbol reduces to a single term, and the phase of W is readily evaluated:[3]

$$\text{phase of } W(J = \max J) = \pi(J_1 + J_2 + J_3 + J_4)$$
$$\text{phase of } W(J = \min J = J_i - J_j) = \pi(J_i + J_k + J_5), \qquad (5.6)$$

where, on the second line, J_k is that angular momentum which forms a vector coupling triangle with J_i and J_5. (In other words, there are four Clebsch–Gordan coefficients in the expression defining the 6-J symbol, and one of them couples J_i plus J_k to J_5.) For example, if $\min J = J_1 - J_2$, then the phase of $W(\min J)$ is $\pi(J_1 + J_3 + J_5)$. Unfortunately, formula (5.5) cannot be continued to either extremum: It breaks down near $J = J_e$ (e for extremum) because the formula was derived assuming f_2 had no zeros, and f_2 does have zeros near $J = J_e$. Furthermore, the WKBJ formulas of Sec. 2 do not work near $J = J_e$ for the same reason. Therefore, what we must do first is investigate the behavior of $W(J)$ near extrema. We have done this in Appendices C and D. We find that, if W is behaving like $e^{i\delta}$ times a monotonically decreasing exponential in the forbidden region near the turning point, then W is still behaving this way at the extremum; hence δ is the phase of $W(J_e)$. Similarly, if W is behaving like $(-1)^J e^{i\delta}$ times a monotonically decreasing exponential in the forbidden region near the turning point, then W is still behaving this way at the extremum; hence δ is the phase of $W(J_e)$ minus πJ_e. At $f_0 = 0$ turning points, therefore, δ is given by Eq. (5.6); while at $f_0 = 4f_2$ turning points, δ is given by eq. (5.6) minus πJ_e.

Now let us interpret the $f_0 = 0$ and $f_0 = 4f_2$ turning points geometrically. From Eqs. (3.13), points where $f_0 = 0$ and $f_0 = 4f_2$ correspond geometrically to points where $\cos\xi = -1$ and $\cos\xi = +1$, respectively, where ξ is the azimuthal angle introduced in Fig. 3. At $f_0 = 0$ the 6-J tetrahedron is "flattened out," while at $f_0 = 4f_2$ it is "folded over." In the (J, J_5) plane introduced in Fig. 4 and at the end of Sec. 3, therefore, $f_0 = 0$ turning points lie on the BC and CD boundaries of the classically forbidden region (see Fig. 4), the ones closest to max J_5, while the $f_0 = 4f_2$ turning points lie on the AB and DA boundaries, close to min J_5.

We can also interpret $d \lessgtr 0$ and $d' \lessgtr 0$ geometrically. The sign of d (say) is the sign of (f_0/zf_2) at $z = 0$. From Eqs. (3.13), f_0 is positive within the classically allowed region only, while f_2 is positive in both classical and quantum-mechanically allowed regions. Furthermore, f_0 must have two zeros, one at the BC boundary and one at the CD boundary (Fig. 4). Therefore within and on the BC-CD boundaries,

$$f_0/f_2 = (J_{02} - J)(J - J_{01})p(J), \qquad (5.7)$$

where J_0 is the BC turning point, J_{02} is the CD turning point, and $p(J)$ is a positive function of J because f_0/f_2, $(J_{02} - J)$, and $(J - J_{01})$ are positive functions of J in the classically allowed region. Now dividing Eq. (5.7) by $z(= J - J_{02}$, depending on which turning point we are at) and evaluating the quotient at $z = 0$, we get

$$\begin{aligned} &d > 0 \text{ at } J = J_{01}, \quad BC \text{ boundary}, \\ &d < 0 \text{ at } J = J_{02}, \quad CD \text{ boundary}. \end{aligned} \qquad (5.8)$$

Similar arguments applied to

$$(4f_2 - f_0)/f_2 \equiv (J'_{02} - J)(J - J'_{01})p'(J), \quad p' \text{ positive},$$

give

$$\begin{aligned} &d' > 0 \quad \text{at} \quad J = J'_{01}, \quad (AB \text{ boundary}), \\ &d' < 0 \quad \text{at} \quad J = J'_{02}, \quad (AD \text{ boundary}). \end{aligned} \qquad (5.9)$$

As remarked in Sec. 3, for special values of $J_1J_2J_3J_4$, one or more of the points $ABCD$ can move into a corner of the plot in Fig. 4, (never the lower right-hand corner), so that one or more of the boundaries DA, AB, BC coincides with an edge $J = \min J$ or $J_5 = \min J_5$. The turning point solution (5.5) then breaks down [for the reasons given after Eq. (5.6)]; but one can use the formulas of

Appendix C if the turning point is too near min J and the formulas of Appendix C again (after a little relabeling of the variables) if the turning point is too near min J_5.

6. SPECULATIVE EXTENSIONS OF THE THEORY

Conceivably, recurrence techniques could be used to obtain approximate formulas for the 6μ symbols of the higher groups ($SU(3)$, etc; μ labels the representation]. Certainly a Biedenharn–Elliott identity exists for the higher groups as well as for $SU(2)$. There is a complication, however; μ for $SU(3)$ (say) is two numbers, not one, because $SU(3)$ has two Casimir operators. Probably the most natural representation labels to use would be the number of quark and antiquark (3 and $\bar{3}$) indices on the irreducible tensor corresponding to representation μ. This labeling would be a natural generalization from $SU(2)$, where $2J$ is the number of quark indices. No matter how one chooses the labels, presumably one would have to deal with a partial difference equation in two variables.

Something similar seems to happen when we stay within $SU(2)$ and attempt to extend the theory to the 9-J symbol. So far we have not been able to find an equation of type (1.1) for the 9-J symbol, although we have found a system of three coupled second-order partial difference equations in three variables. We have not worked with these equations enough to know how tractable they are, but the coefficients in the equations have a simple geometrical interpretation, just as in the 3-J and 6-J cases. Hence one could bring considerable geometrical and physical intuition to bear, and solutions might turn out to be reasonably simple to find.

APPENDIX A: A RECURRENCE RELATION FOR THE 6-J SYMBOL

To derive a recurrence relation of type (1.1) for the 6-J symbol, we start from the Biedenharn–Elliott identity:[8]

$$W(J)\begin{Bmatrix}\nu & J_2 & \mu\\ J_4 & \rho & J_5\end{Bmatrix} = \sum_\lambda (2\lambda+1)(-1)^{\lambda+s}\begin{Bmatrix}J_1 & J_2 & J\\ \mu & \lambda & \nu\end{Bmatrix}\times\begin{Bmatrix}\mu & \lambda & J\\ J_3 & J_4 & \rho\end{Bmatrix}\begin{Bmatrix}J_1 & \nu & \lambda\\ \rho & J_3 & J_5\end{Bmatrix}, \tag{A1}$$

$$s \equiv J_1+J_2+J_3+J_4+J_5+J+\mu+\nu+\rho. \tag{A2}$$

$W(J)$ is defined at Eq. (1.2). If we choose $\nu = J_2$, $\rho = J_4$, the last 6-J bracket in Eq. (A1) becomes $W(\lambda)$; furthermore, if we choose $\mu = 1$, the sum over λ in Eq. (A1) reduces to $J-1 \le \lambda \le J+1$. Equation (A1) has now become a three-term recurrence relation of type (1.5), as desired, with

$$Gg_\pm = (-1)^{J+S\pm1}(2J+1\pm2)\times\begin{Bmatrix}J_1 & J_2 & J\\ 1 & J\pm1 & J_2\end{Bmatrix}\begin{Bmatrix}1 & J\pm1 & J\\ J_3 & J_4 & J_4\end{Bmatrix}, \tag{A3}$$

$$Gg_0 = (-1)^{J+S}(2J+1)\times\begin{Bmatrix}J_1 & J_2 & J\\ 1 & J & J_2\end{Bmatrix}\begin{Bmatrix}1 & J & J\\ J_3 & J_4 & J_4\end{Bmatrix} - \begin{Bmatrix}J_2 & J_2 & 1\\ J_4 & J_4 & J_5\end{Bmatrix}. \tag{A4}$$

We are anticipating that the g's will contain a common factor G to be specified later. Since Eq. (1.5) is homogeneous in the g's, such a factor will cancel out eventually.

All the 6-J symbols in Eqs. (A3), (A4) have at least one of their arguments equal to unity. Such 6-J symbols can be calculated readily from Racah's sum, and general formulas for them are available.[9] After substituting these formulas into Eqs. (A3) and (A4), we get

$$\begin{aligned} g_+(J) = {} & [(J_1+J_2+J+2)(J_1+J_2-J)\\ & \times(J_1-J_2+J+1)(-J_1+J_2+J+1)\\ & \times(J_3+J_4+J+2)(J_3+J_4-J)\\ & \times(J_3-J_4+J+1)(-J_3+J_4+J+1)]^{1/2}\\ & \times(J+1)^{-1} \end{aligned} \tag{A5}$$

$$g_-(J) = g_+(J-1), \tag{A6}$$

$$\begin{aligned} g_0(J) = {} & 2(2J+1)[J_4(J_4+1)+J_2(J_2+1)\\ & -J_5(J_5+1)]-(2J+1)[J_2(J_2+1)\\ & -J_1(J_1+1)+J(J+1)]\,[J_4(J_4+1)\\ & -J_3(J_3+1)+J(J+1)]/J(J+1). \end{aligned} \tag{A7}$$

The common factor G, appearing in Eqs. (A3), (A4) but canceling out thereafter, is

$$G^{-1} = (-1)^{J_2+J_4+J_5}\,4(2J+1)[J_2(J_2+1)(2J_2+1)\times(J_4(J_4+1)(2J_4+1)]^{1/2}. \tag{A8}$$

From Eq. (A6) we have derived a useful result for $\delta^{(1)}g_-$, needed in Sec. 2:

$$\begin{aligned}\delta^{(1)}g_- &= g_-(J+1)-g_-(J)\\ &= g_+(J)-g_-(J),\end{aligned}$$

or, in terms of the f's, Eqs. (1.6):

$$\delta^{(1)}f_2 = f_1. \tag{A9}$$

APPENDIX B: THE 6-J SYMBOL NEAR ZEROS OF $g^\pm$

This appendix and the next discuss the solution to Eq. (1.1), when $W(J)$ is a 6-J symbol and J is near max J (Appendix B) or min J (Appendix C). Near these points the functions $g^\pm$ have zeros, as does the function f_-, and the WKBJ-like solutions of Secs. 4–5 are not valid.

Without loss of generality we can take $\max J = J_1 + J_2$ and define a quantity r by

$$\max J = J_1 + J_2 = J_3 + J_4 - r, \tag{B1}$$

r an integer ≥ 0. In addition we introduce the symbol x for the distance from J to the extremum:

$$x \equiv J_1 + J_2 - J. \tag{B2}$$

From the results of Appendix A, the g's behave as follows near $x \cong 0$:

$$\begin{aligned} g_0(x) &\cong EH, \\ g_+(x) &\cong [x(r+x)]^{1/2} E, \\ g_-(x) &\cong [(x+1)(r+x+1)]^{1/2} E, \end{aligned} \tag{B3}$$

where H and E are constants independent of x and we have neglected terms of order $x/J_i, i = 1$ to 4. Inserting Eqs. (B3) into Eq. (1.1), we get

$$W(x-1)[x(r+x)]^{1/2} + W(x)H + W(x+1) \times [(x+1)(r+x+1)]^{1/2} \cong 0. \tag{B4}$$

(Here and for the remainder of this appendix, for simplicity we have changed the argument of W from $J = \max J - x$ to x.) By changing the dependent variable in Eq. B4, we can transform it to the recurrence relation for the hypergeometric function. Take

$$W(x) = \text{const}[\Gamma(x+1)\Gamma(r+x+1)]^{-1/2} \times a^x U(x)/\Gamma(-x+b), \tag{B5}$$

where a and b are constant parameters to be determined and $U(x)$ is the new dependent variable. Then $U(x)$ obeys

$$U(x-1)[x(r+x)/(-x+b)a] + U(x)H + U(x+1)(-x+b-1)a = 0. \tag{B6}$$

This should be compared to the following recurrence relation for the hypergeometric function $F(\alpha, \beta; \gamma; k) \equiv F(\alpha, \gamma)$[10]:

$$F(\alpha+1, \gamma+1)[(\beta-\gamma)\alpha k/\gamma] + [(\alpha-1) - (\beta-\gamma)k - (2-\gamma)(1-k)]F(\alpha,\gamma) - (\gamma-1)F(\alpha-1, \gamma-1) = 0. \tag{B7}$$

Equations (B6) and (B7) coincide provided we take

$$\begin{aligned} \alpha &= -x, \\ \beta &= [r+1+iH]/2, \\ \gamma &= \beta - (r+x), \\ a &= i, \\ b &= B - r, \\ k &= -1. \end{aligned}$$

Therefore

$$U(x) = F(-x, \beta; \beta - r - x; -1), \tag{B8}$$

and, from Eqs. (B5) and (B8), we have

$$W(x) \cong W(0)[\Gamma(r+1)/\Gamma(x+1)\Gamma(r+x+1)]^{1/2}(i)^x \times [\Gamma(\beta-r)/\Gamma(-x+\beta-r)] \times F(-x, \beta; \beta-r-x; -1), \tag{B9}$$

where $\beta = (r+1+iH)/2$. There is a second solution to Eq. (B7), but it is not of physical interest because it contains a factor of $1/\Gamma(\alpha) = 1/\Gamma(-x)$ which vanishes at physical values of x. Call this second solution $G(\alpha, \gamma)$:

$$G(\alpha,\gamma) \equiv [\Gamma(\gamma)\Gamma(\beta-\gamma+1)/\Gamma(\beta-\alpha+1)\Gamma(\alpha)] \times F(\beta-\gamma+1, \beta; \beta-\alpha+1; k^{-1}).$$

In order to verify that $G(\alpha, \gamma)$ satisfies the same equation as $F(\alpha, \gamma)$, substitute G for F into the left-hand side of Eq. (B7); then relabel $\beta - \gamma + 1 \to \alpha, \beta \to \beta, \beta - \alpha + 1 \to \gamma, k^{-1} \to k$. The expression which results is just the left-hand side of Eq. (B7), times a factor, hence must vanish. QED

Expression (B9) for W looks complex because it contains a factor $(i)^x$ and it depends on β, which is complex. In fact, expression (B9) *is* complex for general x, but not for physical values of x ($x =$ integer ≥ 0. There is a connection formula for the hypergeometric function which states that[11]

$$\begin{aligned} F(\alpha, \beta; \gamma; k) &= [\Gamma(\gamma)\Gamma(\beta-\alpha)/\Gamma(\beta)\Gamma(\gamma-\alpha)](-k)^{-\alpha} \\ &\quad \times F(\alpha, 1-\gamma+\alpha; 1-\beta+\alpha; k^{-1}) \\ &\quad + [\Gamma(\gamma)\Gamma(\alpha-\beta)/\Gamma(\alpha)\Gamma(\gamma-\beta)](-k)^{-\beta} \\ &\quad \times F(\beta, 1-\gamma+\beta; 1-\alpha+\beta; k^{-1}). \end{aligned} \tag{B10}$$

At physical values of x the second term on the right vanishes due to the $\Gamma(\alpha) = \Gamma(-x)$ in the denominator. If one then eliminates $F(\alpha, \beta; \gamma; k)$ from Eq. (B9) using Eq. (B10), one obtains the complex conjugate of Eq. (B9). QED

For determining the phase of W near turning points (in Sec. 5), we need to know the phase of W at small values of x. To know the phase, we must expand F, and, to expand F, we must know the order of magnitude of the parameter H in Eqs. (B3–B4). [W depends on H via $\beta \equiv (r+1+iH)/2$.] From Eqs. (B3), $H = g_0$ ($J = \max J$)$/E$, where E is order $(J_i)^2$ [from Eq. (A5)] and g_0 is usually order $(J_1)^3$ [from Eqs. (A7), (3.11), or (3.12)] unless $\cos\xi$ or $(\cos\psi_{24} - \cos\psi_2 \cos\psi_4)$ vanishes at $J = \max J$. [The angles are defined at Eq. (3.11) and Fig. 3.] Hence H is usually order J_i—i.e., large compared to x.

Furthermore, H is small only when the turning point is quite near $\max J$ and the formulas of Sec. 5 break down anyway. We can restate this geometrically: $\cos\xi \cong 0$ or $(\cos\psi_{24} - \cos\psi_{24} \cos\psi_4) \cong 0$ only near point D in the (J, J_5) plane (Fig. 4), where the classically allowed region touches $\max J$.

Proof: from the geometry of the 6-J tetrahedron, $\cos\psi_{24} - \cos\psi_2 \cos\psi_4 = \cos\xi = 0$ exactly at D. At points near D on the edge $J = \max, J, \cos\xi$ is no longer defined because the 6-J tetrahedron cannot be formed; but $\cos\psi_{24} - \cos\psi_2 \cos\psi_4$ is still meaningful, and the magnitude of this quantity must increase as we move away from point D in either direction along $J = \max J$, because $\cos\psi_{24}$ changes monotonically while $\cos\psi_2 \cos\psi_4$ remains fixed. Therefore, H is small only near point D where max J is very close to a turning point. QED

In determining the phase of W for the purposes of Sec. 5, therefore, we may assume that H—and therefore β—is large compared to x. [No assumption will be made about the size of the parameter r in Eq. (B1).] We now expand the F in Eq. (B9), using $\beta + \beta^* = r + 1$:

$$
\begin{aligned}
F(-x, \beta; \beta - r - x; -1) &\cong F(-x, \beta; -\beta^*; -1) \\
&= 1 + \frac{(-x)\beta}{(-\beta^*)} \frac{(-1)}{1!} \\
&\quad + \frac{(-x)(-x+1)\beta(\beta+1)(-1)^2}{(-\beta^*)(-\beta^*+1)} \frac{}{2!} + \cdots \\
&\cong 1 + (-x) \frac{(\beta/\beta^*)}{1!} \\
&\quad + (-x)(-x+1) \frac{(\beta/\beta^*)^2}{2!} + \cdots \\
&= F(-x, 1; 1; \beta/\beta^*) = (1 - \beta/\beta^*)^x \\
&= [(\beta^* - \beta)/\beta^*]^x \cong [iH/(\beta - r)]^x .
\end{aligned}
$$

On the last line we used $\beta^* - \beta = -iH$.

Inserting this result into Eq. (B9) and ignoring for the moment some factors which do not contribute to the phase, we get

$$W(x) \propto W(0)(i)^x \Gamma(\beta - r)[iH/(\beta - r)]^x / \Gamma(\beta - r - x)$$

or finally, since $\Gamma(\beta - r)/\Gamma(\beta - r - x) \cong (\beta - r)^x$,

$$W(x) \cong W(0)[\Gamma(r+1)/\Gamma(x+1)\,\Gamma(x+r+1)]^{1/2} \times (-H)^x . \quad \text{(B11)}$$

Evidently the phase behavior of $W(x)/W(0)$ depends on the sign of H: $W(x)/W(0)$ is monotonic (respectively, alternates in sign) for H negative (positive). The sign of H follows from the same geometrical arguments used to establish the magnitude of H. Considered as a function of max J, H has a zero at the single classically allowed value of max J (point D in Fig. 4); when H is finite, the sign of H equals the sign of $\cos\psi_{24} - \cos\psi_2 \cos\psi_4$ and is positive (respectively, negative) for max J smaller (larger) than the classically allowed value. Therefore, $W(x)/W(0)$ is monotonic (respectively, alternates in sign) near max J for (J, J_5) in the forbidden region to the right (left) of point D in Fig. 4. Comparing this result for $J \cong \max J$ to the result of Sec. 5 for $J \cong J_0$, we find that $W(x)/W(0)$ has the same phase behavior (monotonic, or alternating in sign) near J_0 as it has near max J. This is the point which we needed to establish for the purposes of Sec. 5.

APPENDIX C: THE 6-J SYMBOL NEAR ZEROS OF $g^{\pm}$ (CONT.)

We continue the discussion of Appendix B, now considering zeros of $g^{\pm}$ near min J rather than max J.

We introduce the notation

$$m \equiv J_1 - J_2, \qquad n \equiv J_4 - J_3, \quad \text{(C1)}$$

so that min $J = \max(|m|, |n|)$. Then near $J =$ min J,

$$
\begin{aligned}
g_+ &\propto [(J+1)^2 - m^2]^{1/2}[(J+1)^2 - n^2]^{1/2}, \\
g_- &\propto [J^2 - m^2]^{1/2}[J^2 - n^2]^{1/2} .
\end{aligned} \quad \text{(C2)}
$$

When min J is large, Eqs. (C2) reduce to equations of the form (B3), and the theory of the previous section becomes applicable.

Proof: Without loss of generality, we can take min $J = m = J_1 - J_2$. We define quantities x', r' by

$$
\begin{aligned}
x' &\equiv J - \min J = J - m, \\
r' &\equiv m - |n| \geq 0;
\end{aligned} \quad \text{(C3)}
$$

then

$$
\begin{aligned}
g_- &\propto [x'(x' + 2m)(x' + r')(x' + m + |n|)]^{1/2}, \\
&\cong [x'(2m)(x' + r')(m + |n|)]^{1/2} ,
\end{aligned}
$$

where we have neglected terms of order $x'/(m + |n|) \cong x'/\min J$, and similarly for g_+. Therefore, whenever $x'/\min J$ is negligible,

$$
\begin{aligned}
g_0(x') &\cong E'H', \\
g_+(x') &\cong [(x'+1)(r'+x'+1)]^{1/2} E', \\
g_-(x') &\cong [x'(r'+x')]^{1/2} E',
\end{aligned} \quad \text{(C4)}
$$

just as in Eqs. (B3) except for the primes and an interchange of g_+ and g_-. This interchange is compensated for at a later point when we replace $W(x \pm 1)$ by $W(x' \mp 1)$ in Eq. (B4). Since x is linear in $-J$ and x' is linear in $+J$, $x \pm 1$ corresponds to $x' \mp 1$. Near $J =$ min J, therefore, $W(x')$ obeys Eq. (B4) with x, r, and H replaced by x', r', and H'. QED

Applying the theory of Appendix B, we verify the statements of Sec. 5. In the forbidden region to the right (respectively, left) of point B in Fig. 4, W alternates in sign (is monotonic).

Now let us consider the case of small min J. Here it turns out that the turning points J_0 are always near min J (or, in geometrical terms, the classical boundaries AB and BC in Fig. 4 are always near the upper edge $J = \max J$) so that the formulas of Sec. 5 break down. Instead of these formulas we use an approximation valid for $J <$ order $J_i^{1/2}$ $(1 \leq i \leq 4)$, and first derived by Racah and Edmonds.[12] As far as we can determine, they obtained the approximation by manipulating the power

series expansion for W. It will be useful to rederive the approximation using the difference equation techniques of the present paper to illustrate a technique which may work in cases where there is no power series.

As a first step, we rewrite the g's of Appendix A in terms of the parameters m and n defined at Eq. (C1) and the angle ψ opposite side J_5 in triangle $J_4J_2J_5$:

$$\cos\psi \equiv (|J_5|^2 - |J_4|^2 - |J_2|^2)/2|J_4\|J_2|, \quad \text{(C5)}$$

$$|J_i|^2 \equiv J_i(J_i+1), \quad \text{(C6)}$$

$$\begin{aligned} g_+(J) &= \{[(J_1+J_2+1)^2-(J+1)^2] \\ &\quad \times [(J+1)^2-m^2][(J_3+J_4+1)^2-(J+1)^2] \\ &\quad \times [(J+1)^2-n^2]\}^{1/2}/(J+1), \end{aligned}$$

$$g_-(J) = g_+(J+1),$$

$$\begin{aligned} g_0(J) &= (2J+1)\{-4|J_2\|J_4|\cos\psi \\ &\quad -[|J_2|^2-|J_1|^2+|J|^2] \\ &\quad \times [|J_4|^2-|J_3|^2+|J|^2]/|J|^2\}. \end{aligned} \quad \text{(C7)}$$

Next, we drop the $(J+1)^2$ in the first and third square brackets of g_+, and neglect the $|J|^2$'s in the numerator of g_0:

$$\begin{aligned} &(J_1+J_2+1)^2-(J+1)^2 \\ &\quad \cong (J_1+J_2+1)^2, \quad \text{etc.}, \\ &|J_2|^2-|J_1|^2+|J|^2 \\ &\quad \cong (J_2-J_1)(J_2+J_1+1)+|J|^2 \\ &\quad \cong -m(J_2+J_1+1), \quad \text{etc.} \end{aligned}$$

In the last line, where we neglected a term of order $|J|^2$ compared to a term of order J_i, we used our assumption that $J <$ order $J_i^{1/2}$. At this point, most of the terms in the g's contain a factor $(J_1+J_2+1)(J_3+J_4+1)$, which suggests that we should divide through by this constant:

$$\begin{aligned} &\frac{g_+(J)}{(J_1+J_2+1)(J_3+J_4+1)} \\ &\quad \cong \frac{\{[(J+1)^2-m^2][(J+1)^2-n^2]\}^{1/2}}{(J+1)}, \\ &\quad g_-(J) = g_+(J-1), \\ &\frac{g_0(J)}{(J_1+J_2+1)(J_3+J_4+1)} \\ &\quad \cong (2J+1)\left(-\cos\psi + \frac{mn}{|J|^2}\right) \end{aligned} \quad \text{(C8)}$$

In simplifying the coefficient of $\cos\psi$, we have used $(J_1+J_2+1)(J_3+J_4+1) = (2J_2+m+1)(2J_4-n+1) \cong 4|J_2\|J_4|$. Next, since most recurrence relations found in the mathematical literature do not contain square roots, we get rid of the roots in Eq. (C8) by an appropriate change of dependent variable. Let

$$W(J) \equiv [(J+m)!\,(J-m)!/(J+n)!\,(J-n)!]^{1/2}P(J). \quad \text{(C9)}$$

Then the recurrence relation for $P(J)$ is

$$\begin{aligned} &P(J+1)[(J+1)^2-m^2]/(J+1) \\ &\quad +(2J+1)(-\cos\psi + mn/|J|^2)P(J) \\ &\quad +P(J-1)[J^2-n^2]/J \cong 0, \end{aligned} \quad \text{(C10)}$$

without any square roots. Comparing Eq. (C10) to a standard recurrence relation for the Jacobi polynomial,[13] we find that

$$P(J) = (\text{const})P_{J-m}^{(m-n,\,m+n)} \quad \text{(C11)}$$

where the constant may be determined from boundary conditions. At $J-m=0$, the Jacobi polynomial reduces to unity. Putting Eq. (C11) into Eq. (C9), setting $J-m=0$ on both sides, and evaluating $W(J=\min J)$ using Racah's sum,[3] we find that

$$\begin{aligned} (\text{const})^2 &= \left(\frac{(2J_2)!\,(2J_4-n-m)!}{(2J_2+2m+1)!\,(2J_4-n+m+1)!}\right)\left(\frac{(J_2+J_4+m-n+J_5+1)!}{(J_2+J_4+J_5+1)!}\,\frac{(J_2+J_4+m-n-J_5)!}{(J_2+J_4-J_5)!}\right) \\ &\quad \times \left(\frac{(J_2-J_4+m+n+J_5)!\,(-J_2+J_4+J_5)!}{(J_2-J_4+J_5)!\,(-J_2+J_4-m-n+J_5)!}\right) \end{aligned} \quad \text{(C11′)}$$

$$\begin{aligned} &\cong (2J_2)^{-2m-1}(2J_4)^{-2m-1}(J_2+J_4+J_5+1)^{m-n}(J_2+J_4-J_5)^{m-n}(J_2-J_4+J_5)^{m+n}(-J_2+J_4+J_5)^{m+n} \\ &\cong (4J_2J_4)^{-1}2^{-2m}(1-\cos\psi)^{m-n}(1+\cos\psi)^{m+n}. \end{aligned} \quad \text{(C12)}$$

The phase of const, from Eq. (5.6) with $J_i - J_j = J_1 - J_2$, is $\pi(J_1+J_3+J_5)$. Then

$$\begin{aligned} W(J) &\cong (-1)^{J_1+J_3+J_5}(4J_2J_4)^{-1/2} \\ &\quad \times [(J+m)!(J-m)!/(J+n)!(J-n)!]^{1/2} \\ &\quad \times 2^{-m}(1-\cos\psi)^{(m-n)/2}(1+\cos\psi)^{(m+n)/2} \\ &\quad \times P_{J-m}^{(m-n,\,m+n)}(\cos\psi) \end{aligned}$$

or

$$W(J) \cong (-1)^{J_1+J_3+J_5}(4J_2J_4)^{-1/2}\,d_{mn}^{J}(\psi), \quad J \ll J_i^{1/2}. \quad \text{(C13)}$$

d^J is the usual $SU(2)$ rotation matrix element.[14] Equation (C13) is the Racah–Edmonds result.

APPENDIX D: AN EXACT SOLUTION TO EQ. (4.2)

If $[f_2/(f_2 + f_1)]^n = (g_-/g_+)^n$ is a ratio of polynomials for some choice of n, then there exists an exact solution to Eq. (4.2). To obtain this solution, we eliminate the difference operators from Eq. (4.2), using definitions (1.3). We get

$$(g_+/g_-) = (f_2 + f_1)/f_2 = R(J-1)/R(J+1). \quad \text{(D1)}$$

Next we raise both sides of this equation to an appropriate power n, in order to get rid of any roots present on the left. In the 6-J case, for example, we square both sides ($n = 2$). Also, we make the changes of variable $J - 1 \equiv 2z$, $[R(2z)]^n \equiv Y(z)$. Equation (D1) becomes

$$(g_-/g_+)^n\, Y(z) = Y(z+1). \quad \text{(D2)}$$

Equation (D2) is a first-order difference equation in z with rational coefficients, a standard form solved in Sec. 11.2 of Milne-Thompson.[15] Suppose, for example, that

$$(g_+/g_-)^n = a(J-r_1)(r_2 - J)/(J-r_3)(r_4 - J), \quad \text{(D3)}$$

where a and the r_i are constants and where for later convenience each parenthesis has been arranged so as to be positive for J values of physical interest. Then the solution to Eq. (D2) is determined up to a constant:

$$[R(J-1)]^n = Y(z) \propto \frac{a^z \Gamma[(J-r_1)/2]\Gamma[(r_4 - J + 2)/2]}{\Gamma[(J-r_3)/2]\Gamma[(r_2 - J + 2)/2]}. \quad \text{(D4)}$$

The argument of each gamma function in Eq. (D4) is positive because each parenthesis in Eq. (D3) is positive. Polynomials more complicated than Eq. (D3) can be handled by induction on this simple example.

In the 6-J case the $g_\pm(J)$ are of the form

$$g_-(J) = \left(\prod_{i=1}^{q}(J-r_i)\prod_{j=q+1}^{p}(r_j - J)\right)^{1/2}/J,$$
$$g_+(J) = g_-(J+1). \quad \text{(D5)}$$

The exact solution for this case is

$$[R(J-1)]^2 = [\Gamma(\tfrac{1}{2}J + \tfrac{1}{2})/\Gamma(\tfrac{1}{2}J)]^2 \times \prod_{ij}\Gamma[(J-r_i)/2]\,\Gamma[(r_j - J + 1)/2]/\Gamma[(J-r_i+1)/2]\,\Gamma[(r_j - J + 2)/2] \quad \text{(D6)}$$

$$\cong 2^{p/2-1}/f_2(J-1). \quad \text{(D7)}$$

The approximate result (D7) was obtained from Eq. (D6) using Stirlings' approximation, valid for J not too close to an r_i or r_j, and is the same as the approximate result obtained using another method at Eq. (4.5) (up to a constant of proportionality which can be ignored). p is the number of factors in the numerator of Eq. (D5).

APPENDIX E: THE 3-J SYMBOL

In this appendix let $W(J)$ be a symmetrized Clebsch-Gordan coefficient, or 3-J symbol[16]:

$$W(J) = (-1)^{J_1+J_2+m}\langle J_1 m_1 J_2 m_2 | Jm\rangle/(2J+1)^{1/2} \equiv \begin{pmatrix} J_1 & J_2 & J \\ m_1 & m_2 & -m \end{pmatrix}. \quad \text{(E1)}$$

Then W obeys a recurrence relation of type (1.5) with[17]

$$g_-(J) = \frac{\{(J^2 - m^2)[J^2 - (J_1 - J_2)^2][J_1 + J_2 + 1)^2 - J^2]\}^{1/2}}{2J},$$
$$g_+(J) = g_-(J+1), \quad g_0(J) = (2J+1)\left(\frac{-m[J(J+1) + J_2(J_2+1) - J_1(J_1+1)]}{2J(J+1)} + m_2\right). \quad \text{(E2)}$$

The geometric interpretation of these g's is remarkably similar to the geometrical interpretation of the 6-J g's. In Fig. 3 relabel

$$JJ_1J_2J_4 \to JJ_1J_2m. \quad \text{(E3)}$$

The replacement of J_4 by a vector of length m means that the old J_4 axis becomes the new z axis. The old J_3 and J_5 in Fig. 3 have no particular physical significance after the relabeling; however, the length of J_3 must be such that the projection of J upon the new z axis (old J_4) is m (i.e., J_3 and old J_4 must form a right angle), and the length of J_5 must be such that the projection of J_2 upon the new z axis is m_2.

Hence there is a tetrahedron in the 3-J case as well as in the 6-J case! In terms of the geometry of this tetrahedron, we have

$$g_+ \cong g_- \cong JJ_2 \sin\psi_2 \sin\psi_4 \quad \text{(E4)}$$

$$g_0 \cong 2JJ_2[\cos\psi_{24} - \cos\psi_2\cos\psi_4] \quad \text{(E5)}$$

$$= 2JJ_2 \sin\psi_2 \sin\psi_4 \cos\xi. \quad \text{(E6)}$$

These formulas are remarkably similar to the 6-J ones, Eqs. (3.10) and (3.11). The only significant difference is in the direction of simplicity: Since the old "4" direction is the new z direction, the formulas for $\cos\psi_{24}$ and $\cos\psi_4$ are simpler:

$$\cos\psi_{24} = m_2/J_2, \tag{E7}$$

$$\cos\psi_4 = m/J. \tag{E8}$$

The mean-square value of W in the classically allowed region is again proportional to V^{-1}, although with a different constant of proportionality[5]:

$$\overline{W}^2 = |m|/12\pi V, \quad \text{or zero.} \tag{E9}$$

A small-J analysis carried out according to the methods of Appendix C again gives a W proportional to a rotation matrix:

$$W \cong (-1)^{J+J_1+m_2} d^J_{J_2-J_1,\,m}(\psi_{24})/(2J+1)^{1/2}, \quad J < \text{order } J_1^{1/2}. \tag{E10}$$

This result was first derived by Brussard and Tolhoek.[18]

Likewise, the discussion of 6-J turning points in Secs. 4 and 5 may be repeated nearly verbatim for the 3-J symbol. There are only two minor changes, both connected with normalization. Firstly, the magnitude of the 3-J W, Eq. (E9), is larger than the magnitude of the 6-J $\overline{W}$, Eq. (3.7), by a factor of $(2|m|)^{1/2}$; hence Eq. (5.5) must be multiplied by $(2|m|)^{1/2}$ before it is correct for the 3-J symbol. Secondly, in place of Eqs. (5.6) we must use

$$\text{phase of } W\,(J = \max J) = \pi(J_1 - J_2 + m_1 + m_2) \tag{E11}$$

and, if $\min J = J_2 - J_1$,

$$\text{phase of } W\,(J = J_2 - J_1) = \pi(2J_1 - m_2 + J_2) \tag{E12}$$

or, if $\min J = J_1 - J_2$,

$$\text{phase of } W\,(J = J_1 - J_2) = \pi(2J_2 - m_1 - J_1). \tag{E13}$$

If we then repeat the discussion following Eqs. (5.6) but use Eqs. (E11)–(E13) in place of Eqs. (5.6), we get a different phase δ in Eq. (5.5).

---

1 For a WKBJ bibliography see Leonard I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955, 1968), 2nd ed., p. 184, 3rd ed., p. 268.

2 Eugene P. Wigner, *Group Theory* (Academic, New York, 1959); A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1960).

3 For Racah's sum for the 6-J symbol see Edmonds, Ref. 2, p. 99.

4 Two extensive computer complications of 6-J symbols are M. Rotenberg, R. Bivens, N. Metropolis, and J. K. Wooten, Jr., *3-J and 6-J Symbols* (Technology, Cambridge, Mass., 1959), and *Tables of Racah Coefficients*, edited by Ishidzu Takehiko (Pan-Pacific, Tokyo, 1960).

5 For a geometrical calculation of A in both the 3-J and the 6-J cases, see Wigner, Ref. 2, Eqs. (27.6) and (27.12), pp. 353 ff.

6 D. E. Neville, Phys. Rev. **160**, 1375 (1967); **163**, 1582 (1967).

7 W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea, New York, 1954). Debye's series are on p. 23.

8 For a derivation of the Biedenharn–Elliott identity, see Edmonds, Ref. 2, p. 96–97.

9 6-J symbols with one argument unity are quoted by Edmonds, Ref. 2, Table 5, p. 130.

10 Chester Snow, *Hypergeometric and Legendre Functions*, N.B.S. Applied Math. Series # 19 (U.S. Govt. Printing Office, Washington, D.C., 1952), p. 32. Equation (B7) follows from Eq. (12) (with α replaced by $\alpha - 1$) plus Eq. (12) (with γ replaced by $\gamma + 1$) plus Eq. (15) (with α and β interchanged).

11 Magnus and Oberhettinger, Ref. 7, p. 9, third line.

12 Edmonds, Ref. 2, Eq. (A2.2), p. 122. Racah derived Eq. (C13) for the special case $m = n = 0$.

13 Gabor Szego, *Orthogonal Polynomials*, AMS Colloquium Publications, Vol. 23 (Am. Math. Soc., Providence, R.I., 1939), p. 71.

14 Edmonds, Ref. 2, p. 58.

15 L. M. Milne-Thompson, *Calculus of Finite Differences* (Macmillan, London, 1933), Sec. 11.2.

16 Edmonds, Ref. 2, Eq. (3.7.3), p. 46.

17 E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge U.P., New York, 1959), p. 74, Eq. (4).

18 P. Brussaard and J. H. Tolhoek, *Physica*, **23**, 955 (1957). See also Edmonds, Ref. 2, Eq. (A2.1), p. 122.

Irreducible Vector and Ray Representations of Some Cubic Crystal Point Groups in Four Dimensions*

Li-Ching Chen† and Joseph L. Birman
Physics Department, New York University, New York, N.Y. 10012
(Received 3 June 1971)

A new, and computationally simple, method is given for obtaining the characters of all the inequivalent irreducible vector representations of a finite group. The method was applied to determine the characters of the irreducible vector and ray representations of the four-dimensional cubic crystal point groups: group 47 and group 45. These groups are of order 384 and 1152, respectively, and contain the cubic point group in three dimensions $O_h[3]$ as a subgroup. Tables are given of the irreducible representations of O_h subduced by the irreducible representations of group 47 and group 45. These tables may be useful in testing the conjecture that accidental degeneracy in problems in solid state physics in three dimensions reflects a higher symmetry in four dimensions.

1. INTRODUCTION

Character tables of the inequivalent irreducible vector representations, and of the p-inequivalent irreducible ray representations, of finite groups have been widely used in physics. For example, such tables for the crystallographic point groups in three dimensions are indispensable in problems in solid state physics.

The present paper is concerned with such character tables for certain crystallographic cubic point groups in four dimensions. This study was motivated by the conjecture by one of us (J.B.) that accidental degeneracy in problems in solid state physics (in three dimensions) is related to an underlying four-dimensional symmetry group: namely, a crystallographic point subgroup of the group $O(4)$. We will not discuss this conjecture extensively in this paper but restrict ourselves to the mathematical, and computational aspects of obtaining the character tables needed for further study of this problem.

In principle, given the defining generators and relations for any finite group, the character tables can be found straightforwardly by solving a set of bilinear equations.[1] But when the order of the group is large with correspondingly large number of generators and relations, one must also be concerned with the practicality of a computational method used to obtain the characters. For this reason a systematic and practical method of obtaining characters is given in Sec. 2. To our knowledge, this method is new.

In Sec. 3 we apply this method to two cubic crystal point groups in four dimensions. These groups are known as the groups 45 and 47 in [4].[2] The group 45 is a subgroup of order 1152 of the orthogonal group $O(4)$. The groups 47 is a subgroup of index 3 of 45. The familiar cubic group in 3 dimensions: $O_h(3)$ is a subgroup of index 8 in 47. Thus, we give in Sec. 3 the complete set of character tables for groups 45 and 47 and also in Sec. 4 the subduction tables which give the connection with the irreducible representations of $O_h[3]$. It is the latter table which will ultimately be of value in testing the conjecture on the relation to accidental degeneracy. The concluding Sec. 5 gives a brief discussion of our results.

2. METHOD OF OBTAINING CHARACTERS OF VECTOR IRREDUCIBLE REPRESENTATIONS

All methods of obtaining the full set of characters of a finite group start from the equations which relate the characters to certain structure constants of the group.

Let G be a finite group of order g. Let the class C_i of G be of order r_i, and the total number of classes be r. Let χ_i^μ be the character of class C_i, in the μth irreducible vector representation Γ_μ of G; let d_μ be the dimension of Γ_μ.

There are then four basic equations to be used.[1,3] The positive integers d_μ satisfy

$$\sum_{\mu=1}^{r} d_\mu^2 = g. \tag{2.1}$$

The characters χ_i^μ satisfy the orthonormality relations

$$\sum_{\mu=1}^{r} r_i \chi_i^\mu \chi_j^{\mu *} = g\delta_{ij} \tag{2.2}$$

and

$$\sum_{i=1}^{r} r_i \chi_i^\mu \chi_i^{\nu *} = g\delta_{\mu\nu}. \tag{2.3}$$

Finally, we have the bilinear equations

$$r_i r_j \chi_i^\mu \chi_j^\mu = d_\mu \sum_{k=1}^{r} h_{ij,k} r_k \chi_k^\mu, \qquad i,j, = 1,\ldots,r. \tag{2.4}$$

In (2.4) the $h_{ij,k}$ are the class multiplication coefficients, which have the symmetry properties

$$h_{ij,k} = h_{ji,k} \tag{2.5}$$

and also obey the sum rule

$$\sum_{k=1}^{r} h_{ij,k} h_{kl,m} = \sum_{k=1}^{r} h_{jl,k} h_{ik,m} \tag{2.6}$$

and

$$r_k h_{ij,k} = r_i h_{jk',i'} = r_j h_{ik',j'}. \tag{2.7}$$

In (2.7) we have that the primed letters (e.g., i', j', ...) correspond to the index of the inverse class which contains the inverse elements to those in $(i,j,\ldots)$. Evidently $r_j = r_{j'}$.

For finite groups of low order, such as the crystal point groups in [3], one can proceed in an elementary fashion: first finding the positive integer

d_μ from (2.1), then the χ_i^μ from (2.2)–(2.4). Another method, introduced by Burnside,[4] involves finding eigenvalues and eigenvectors of r different matrices.

But we have used a method which seems simpler and more facile computationally and proceeds directly from (2.2)–(2.4), thereby automatically ensuring that the orthonormality relations are satisfied. Define an $r \times r$ matrix U by

$$U_{\mu i} \equiv (r_i/g)^{1/2}\chi_i^\mu. \tag{2.8}$$

From (2.2) and (2.3) we see that U is unitary. Then (2.4) becomes

$$\sqrt{gr_ir_j}\; U_{\mu i}U_{\mu j} = d_\mu \sum_{k=1}^{r} h_{ij,k}\sqrt{r_k}\, U_{\mu k}. \tag{2.9}$$

Multiply both sides of (2.9) by $U_{\nu j}^*/(d_\mu\sqrt{r_j})$ and sum on j from 1 to r. We have

$$\sqrt{\frac{gr_i}{d_\mu}}\; U_{\mu i}\delta_{\mu\nu} = \sum_{j=1}^{r}\sum_{k=1}^{r} h_{ij,k}\sqrt{\frac{r_k}{r_j}}\; U_{\mu k}U_{\nu j}^*. \tag{2.10}$$

A. Dimensionalities of Representations

We first show the use of (2.10) to determine the dimensionalities d_μ of the irreducible vector representations. Let y_i with $i = 1, \ldots, r$ be a set of indeterminates, and multiply both sides of (2.10) by y_i and sum over i from 1 to r:

$$\left(\sum_i (\sqrt{gr_i}/d_\mu)\, U_{\mu i}y_i\right)\delta_{\mu\nu} = \sum_j\sum_k U_{\mu k}\left(\sum_i h_{ij,k}y_i\sqrt{r_k/r_j}\right)U_{\nu j}^*. \tag{2.11}$$

Define the r numbers λ_μ and the matrix Y by

$$\lambda_\mu \equiv \sum_{i=1}^{r}(\sqrt{gr_i}/d_\mu)U_{\mu i}y_i \tag{2.12}$$

and

$$Y_{kj} \equiv \sqrt{r_k/r_j}\sum_i h_{ij,k}y_i. \tag{2.13}$$

Then (2.11) becomes

$$\lambda_\mu\delta_{\mu\nu} = \sum_{j,k=1}^{r} U_{\mu k}Y_{kj}U_{j\nu}^{-1}, \tag{2.14}$$

where we used the unitary property of U. Also from (2.14),

$$\sum_k Y_{jk}U_{k\mu}^{-1} = \lambda_\mu U_{j\mu}^{-1}. \tag{2.15}$$

Hence the unitary matrix U diagonalizes the matrix Y, and the column vector $(U^{-1})_{j\mu}$ (with μ fixed and $j = 1, \ldots, r$) is an eigenvector of Y associated with eigenvalue λ_μ. With a particular choice of the y_i, the problem can be completely solved.

Now in (2.10), let $\mu = \nu$ and sum over μ from $1, \ldots, r$. Then

$$\sum_\mu \sqrt{\frac{gr_i}{d_\mu}}\; U_{\mu i} = \sum_{j,k=1}^{r} h_{ij,k}\sqrt{\frac{r_k}{r_j}}\left(\sum_{\mu=1}^{r} U_{\mu k}U_{\mu j}^*\right) = \sum_{k=1}^{r} h_{ik,k}. \tag{2.16}$$

Then multiply both sides of (2.16) by $(1/\sqrt{r_i})U_{\omega i}^*$ and sum on i. Then we have

$$\sum_{\mu=1}^{r}(\sqrt{g}/d_\mu)\delta_{\mu\omega} = \sqrt{g}/d_\omega = \sum_{i,k=1}^{r}(1/\sqrt{r_i})\, h_{ik,k}U_{\omega i}^* \tag{2.17}$$

or

$$g/d_\omega^2 = \sum_{i=1}^{r}(\sqrt{gr_i}/d_\omega)\, U_{\omega i}\left(\sum_{k=1}^{r}(1/r_i)\, h_{ik,k}\right). \tag{2.18}$$

But now compare (2.18) and (2.12). If we choose

$$y_i \equiv \sum_{k=1}^{r}(1/r_i)\, h_{ik,k}, \tag{2.19}$$

then

$$\lambda_\mu = g/d_\mu^2 \tag{2.20}$$

and

$$Y_{kj} \equiv \sqrt{\frac{r_k}{r_j}}\sum_{i=1}^{r} h_{ij,k}\left(\frac{1}{r_i}\sum_{l=1}^{r} h_{il,l}\right). \tag{2.21}$$

From (2.21), using (2.6) and (2.7), we easily prove that

$$Y_{kj} = Y_{jk}. \tag{2.22}$$

Then Y is a real symmetric matrix. Further, it follows from (2.20) that the eigenvalues of Y are trivially related to the dimensionalities d_μ of the irreducible representations of the group.

Thus the first stage of the program is easily concluded. From the basic structure constants of the group (the r_i and $h_{ij,k}$), the matrix Y can be constructed. The matrix Y is $r \times r$ and its real eigenvalues λ_μ, and its eigenvectors can be found by using a standard computer routine.[5]

B. Characters of Representations

A standard eigenvector-eigenvalue routine can be used to diagonalize the matrix Y. Such a routine will produce an orthogonal matrix V, which diagonalizes Y,

$$VYV^{-1} = \lambda, \tag{2.23}$$

or, using the symmetry and reality of Y,

$$\sum_{k=1}^{r} Y_{jk}V_{\mu k} = (g/d_\mu^2)V_{\mu j}, \tag{2.24}$$

where we use (2.20). But we need to examine now the relation between the matrix V given by the computer and the desired matrix U of (2.8), whose elements are the characters.

If a given eigenvalue $\lambda_\nu = g/d_\nu^2$ of Y occurs once, the corresponding eigenvector $V_{\nu k}$ is unique up to a sign. Then this eigenvector $V_{\nu k}(k = 1, \ldots, r)$ is related to the corresponding $U_{\nu k}$ by

$$(U_{\nu 1}, U_{\nu 2}, \ldots, U_{\nu r}) = \pm\,(V_{\nu 1}, \ldots, V_{\nu r}), \tag{2.25}$$

and the sign of (2.25) can be found by taking

$$U_{\nu 1} \equiv \sqrt{g}\ \chi_1^{\nu} = \sqrt{g}\, d_{\nu}. \tag{2.26}$$

This concludes the analysis for the cases where only one irreducible vector representation exists of a given dimension d_ν.

Now assume that, for certain μ, the eigenvalue (g/d_μ^2) is multiple:

$$d_1 = d_2 = \cdots = d_\mu = \cdots d_n. \tag{2.27}$$

That is, there are n irreducible vector representations of the same dimension d_μ. In this case, the eigenvectors $V_{\mu k}$ are not automatically the desired $U_{\mu k}$. To find the $U_{\mu k}$, we must make a unitary transformation of the $V_{\mu k}$:

$$U_{\mu k} = \sum_{\nu=1}^{n} a_{\mu\nu} V_{\nu k}, \quad k = 1, \ldots, r, \ \mu = 1, \ldots, n. \tag{2.28}$$

To find the elements $a_{\mu\nu}$, we need to substitute (2.28) into (2.9). We have

$$\sqrt{g r_i r_j}\left(\sum_{\nu=1}^{n} a_{\mu\nu} V_{\nu i}\right)\left(\sum_{\omega=1}^{n} a_{\mu\omega} V_{\omega j}\right) = d_\mu \sum_{k=1}^{r} h_{ij,k} \sqrt{r_k}\left(\sum_{\gamma=1}^{n} a_{\mu\gamma} V_{\gamma k}\right) \tag{2.29}$$

or, if we multiply both sides of (2.29) by

$$(1/\sqrt{g r_i r_j}) V_{\alpha i}^* V_{\beta j}^* \tag{2.30}$$

and sum over i and j, we have

$$a_{\mu\alpha} a_{\mu\beta} = d_\mu \sum_{\gamma=1}^{n} f_{\alpha\beta\gamma} a_{\mu\gamma} \tag{2.31}$$

with

$$f_{\alpha\beta\gamma} \equiv \sum_{i,j,k=1}^{r} \sqrt{\frac{r_k}{g r_i r_j}}\, h_{ij,k} V_{\alpha i}^* V_{\beta j}^* V_{\gamma k}. \tag{2.32}$$

The coefficients $f_{\alpha\beta\gamma}$ are real and are given, since the diagonalization of Y has produced the real eigenvectors $V_{\alpha i}$. In the work reported here, Eqs. (2.31) were solved directly by inspection. Since the subspace involved is relatively small (n, the multiplicity of eigenvalue λ_μ much smaller than r, the number of classes of the group), this is a straightforward task. After this work was completed, we realized that a simple application of the same kind of analysis used earlier could be applied also to Eqs. (2.31). This is described in the Appendix.

In summary, the method given here reduces the problem of determining the characters of the irreducible vector representations of a finite group to that of a diagonalization of a matrix, which is easily constructed from known structure constants of the group. Since the p-inequivalent irreducible ray representations of a given group are obtainable from the irreducible vector representations of a certain minimum covering group (or "representation group"), our procedure suffices to find ray representations once we find the representation group for the given group. But in principle this is a solved problem.

3. IRREDUCIBLE REPRESENTATIONS OF GROUPS 45 AND 47 IN [4]

Hurley[2,6] has derived the 227 crystallographic point groups in four-dimensional Euclidean space [4], basing his work on Goursat's[7] analysis of the orthogonal groups in [4]. The largest of these groups is of order 1152, and is labeled Group 45, following Hurley. It is the symmetry group of the regular 24-cell ($\delta 4$). The symmetry group of the hypercubic in [4] is of order 384, and is called Group 47; it is isomorphic to a subgroup of Group 45.

In this section, we give the characters of the inequivalent irreducible vector representations of these two groups as obtained by the methods discussed in Sec. 2. Also we give all the p-inequivalent irreducible ray representations of these groups, found from the irreducible vector representations of the representation (covering) group.

A. Group 47: Hypercubic

A set of generators and relations for the group 47 is given in Table I, and correspondence with Goursat's substitutions are given in Table II. Column 3 of the latter table gives the 4×4 matrix associated with each substitution, and column 4 the invariants $\chi(A)$, $\sigma(A)$, and $d(A)$ for every orthogonal matrix A in [4]. The invariants are defined by the characteristic equation of A:

$$\det(\lambda I - A) = \lambda^4 - \chi(A)\lambda^3 + \sigma(A)\lambda^2 - d(A)\chi(A)\lambda + d(A) = 0 \tag{3.1}$$

The last column gives Hurley's letters which correspond to the 24 symmetry operations distinguished by different values of χ, σ, or d.

Mackay and Pawley[8] have given the Bravais lattices in [4]. For Group 47, the invariant metric tensor g_{ij} is

$$g_{ij} = \begin{pmatrix} a_1^2 & 0 & 0 & 0 \\ 0 & a_1^2 & 0 & 0 \\ 0 & 0 & a_1^2 & 0 \\ 0 & 0 & 0 & a_1^2 \end{pmatrix}, \tag{3.2}$$

where the unit cell edges in terms of orthogonal unit vectors $(\hat{\imath}, \hat{\jmath}, \hat{k}, \hat{l})$ are $a_1, \hat{\imath}, a_1, \hat{\jmath}, a_1, \hat{k}, a_1, \hat{l}$.

The group 47 has 20 conjugacy classes. All are ambivalent so that the vector characters of this group are all real. In Table III, we give the character table for inequivalent irreducible vector representations of 47. In the table, we give a typical element of each class, the order of the class, and the full set of characters for the 20

irreducible representations $\Gamma^{(j)}, j = 1, \ldots, 20$. All are obtained by the method of Sec. 2.

To obtain all the p-inequivalent, irreducible ray representations of a group G, we need to construct the representation group G^* and find the irreducible vector representations of the latter. Here G^* has the property

$$G^*/M = G, \tag{3.3}$$

where M is in the center of G^*. We modified the procedures described by Döring[9] and Hamermesh[3] to find the minimum group G^* or 47*. Details are given in the thesis of Chen.[10] Defining relations of the group 47* are given in Table IV. The constants α, β, γ in that table can have values ± 1; different choices give p-inequivalent factor systems for the irreducible ray representations of 47. Finally, in Tables V–XI, the complete set of p-inequivalent irreducible ray representations of 47 are given.

B. Group 45

Generators and defining relations of Group 45 are given in Table I. Compared with Group 47, this group has an additional generator K; the corresponding substitution and the 4×4 matrix are given in Table II. The metric for this group is

$$g_{ij} = \begin{pmatrix} a_1^2 & -a_1^2/2 & 0 & a_1^2/2 \\ & a_1^2 & -a_1^2/2 & 0 \\ & & a_1^2 & -a_1^2/2 \\ & & & a_1^2 \end{pmatrix}$$

with cell edges $\frac{1}{4} a_1 (-\hat{\imath} + \hat{\jmath} + \hat{k} + \hat{l})$, $\frac{1}{4} a_1 (\hat{\imath} - \hat{\jmath} - \hat{k} + \hat{l})$, $\frac{1}{4} a_1 (\hat{\imath} + \hat{\jmath} + \hat{k} - \hat{l})$, and $\frac{1}{4} a_1 (-\hat{\imath} - \hat{\jmath} + \hat{k} + \hat{l})$.

All 25 classes of Group 45 are ambivalent. The vector characters of Group 45 are listed in Table XII.

The defining relations for the representation Group 45* are given in Table XIII. The constants α, γ can take values $\pm$. Using the method of Sec. 2, we found the characters of the irreducible vector representations of Group 45*, and thence the characters of the p-inequivalent irreducible ray representations of Group 45. These are given in Tables XIV–XVI.

4. O_h AS A SUBGROUP OF GROUP 47 AND GROUP 45

From Table I, we know that the five generators T, R, S, W, and Y form a subgroup H of both Group 47 and Group 45. The group H is defined by the following:

$$T^2 = R^2 = S^3 = W^4 = Y^2 = I,$$

$$RT = TR, \quad ST = TRS, \quad SR = TS, \quad W^2 = T,$$

$$WR = TRW, \quad WS = TS^2W, \quad YT = TY, \quad YR = RY,$$

$$YS = SY, \quad YW = WY.$$

If we associate W with A_1, T with A_1^2, S with A_2, R with $(A_2A_1)^2$, and Y with I', then it can easily

TABLE I. Defining relations of Group 47 and Group 45.

Group 47: $T^2 = R^2 = S^3 = I$,

$RT = TR$,	$ST = TRS$.	$SR = TS$,
$W^2 = T$,	$WR = TRW$,	$WS = TS^2W$,
$I'^2 = I$,	$A^2 = B^2 = I'$,	
$I'T = TI'$,	$I'R = RI'$,	$I'S = SI'$,
$I'W = WI'$,	$I'A = AI'$,	$I'B = BI'$,
$AT = TA$,	$AR = I'RA$,	$AS = I'SA$,
$AW = WA$,	$BT = I'TB$,	$BR = RB$,
$BS = SAB$,	$BW = I'WAB$,	$BA = I'AB$,
$Y^2 = I$,	$YT = TY$,	$YR = RY$,
$YS = SY$,	$YW = WY$,	$YI' = I'Y$.

Group 45 = Group 47 + (Group 47)K + (Group 47)K^2

$K^3 = I$,	$KT = I'TBK$,	$KR = I'RABK$,
$KS = SK$,	$KW = WBK^2$,	$KI' = I'K$,
$KA = I'ABK$,	$KB = I'AK$,	$KY = SYK^2$.

TABLE II. Generators of Group 47 and Group 45.

Generator	Goursat's substitution	Matrix	(χ, σ, d)	Hurley's letter
I	$[\eta, \xi; \eta, \xi]$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	(4, 6, 1)	I
I'	$[\eta, \xi; \eta, \xi]$	$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	(−4, 6, 1)	I'
T	$[\eta, \xi; -\eta, -\xi]$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	(0, −2, 1)	E
R	$\left[\eta, \xi; \frac{1}{\eta}, \frac{1}{\xi}\right]$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	(0, −2, 1)	E
S	$\left[\eta, \xi; i\frac{\eta+1}{\eta-1}, i\frac{\xi+1}{\xi-1}\right]$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	(1, 0, 1)	K
W	$[\eta, \xi; i\eta, i\xi]$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	(2, 2, 1)	R
A	$[\eta, \xi; \eta, -\xi]$	$\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	(0, 2, 1)	D
B	$\left[\eta, \xi; \eta, \frac{1}{\xi}\right]$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$	(0, 2, 1)	D
Y	$[\eta, \xi; \xi, \eta]$	$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	(2, 0, −1)	T
K	$\left[\eta, \xi; \eta, i\frac{\xi+1}{\xi-1}\right]$	$\frac{1}{2}\begin{bmatrix} -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix}$	(−2, 3, 1)	S'

TABLE III. Vector Characters of Group 47.

Class	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
Typical Element	I	S	SY	W	Y	RWY	R	RW	A	RA	SWA	WY	YA	$SWYA$	I'	$I'S$	$I'SY$	WA	$I'Y$	$I'RWY$
r_i	1	32	32	12	4	12	6	24	12	12	48	24	24	48	1	32	32	12	4	12
$\Gamma^{(1)}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$\Gamma^{(2)}$	1	1	1	−1	1	−1	1	−1	1	1	−1	−1	1	−1	1	1	1	−1	1	−1
$\Gamma^{(3)}$	1	1	−1	1	−1	−1	1	1	1	1	1	−1	−1	−1	1	1	−1	1	−1	−1
$\Gamma^{(4)}$	1	1	−1	−1	−1	1	1	−1	1	1	−1	1	−1	1	1	1	−1	−1	−1	1
$\Gamma^{(5)}$	2	−1	1	0	−2	0	2	0	2	2	0	0	−2	0	2	−1	1	0	−2	0
$\Gamma^{(6)}$	2	−1	−1	0	2	0	2	0	2	2	0	0	2	0	2	−1	−1	0	2	0
$\Gamma^{(7)}$	3	0	0	1	3	1	3	1	−1	−1	−1	1	−1	−1	3	0	0	1	3	1
$\Gamma^{(8)}$	3	0	0	−1	3	−1	3	−1	−1	−1	1	−1	−1	1	3	0	0	−1	3	−1
$\Gamma^{(9)}$	3	0	0	−1	−3	1	3	−1	−1	−1	1	1	1	−1	3	0	0	−1	−3	1
$\Gamma^{(10)}$	3	0	0	1	−3	−1	3	1	−1	−1	−1	−1	1	1	3	0	0	1	−3	−1
$\Gamma^{(11)}$	4	1	1	−2	−2	−2	0	0	0	0	0	0	0	0	−4	−1	−1	2	2	2
$\Gamma^{(12)}$	4	1	−1	−2	2	2	0	0	0	0	0	0	0	0	−4	−1	1	2	−2	−2
$\Gamma^{(13)}$	4	1	−1	2	2	−2	0	0	0	0	0	0	0	0	−4	−1	1	−2	−2	2
$\Gamma^{(14)}$	4	1	1	2	−2	2	0	0	0	0	0	0	0	0	−4	−1	−1	−2	2	−2
$\Gamma^{(15)}$	6	0	0	2	0	0	−2	−2	2	−2	0	0	0	0	6	0	0	2	0	0
$\Gamma^{(16)}$	6	0	0	−2	0	0	−2	2	2	−2	0	0	0	0	6	0	0	−2	0	0
$\Gamma^{(17)}$	6	0	0	0	0	2	−2	0	−2	2	0	−2	0	0	6	0	0	0	0	2
$\Gamma^{(18)}$	6	0	0	0	0	−2	−2	0	−2	2	0	2	0	0	6	0	0	0	0	−2
$\Gamma^{(19)}$	8	−1	1	0	4	0	0	0	0	0	0	0	0	0	−8	1	−1	0	−4	0
$\Gamma^{(20)}$	8	−1	−1	0	−4	0	0	0	0	0	0	0	0	0	−8	1	1	0	4	0

TABLE IV. Group 47*: Generators and defining relations.

$\alpha^2 = \beta^2 = \gamma^2 = 1$		
$T^2 = \alpha I$		
$R^2 = \alpha I$	$RT = \alpha TR$	
$S^3 = I$	$ST = TRS$	$SR = TS$
$W^2 = T$	$WR = TRW$	$WS = \alpha TS^2W$
$A^2 = \beta I'$	$AT = TA$	$AR = I'RA$
	$AS = I'SB$	$AW = \beta WA$
$B^2 = \beta I'$	$BA = \beta I'AB$	$BT = I'TB$
	$BR = RB$	$BS = SAB$
	$BW = \beta I'WAB$	
$I'^2 = I$	$I'T = TI'$	$I'R = RI'$
	$I'S = SI'$	$I'W = WI'$
	$I'A = AI'$	$I'B = BI'$
$Y^2 = I$	$YT = TY$	$YR = RY$
	$YS = SY$	$YW = \gamma WY$
	$YI' = \alpha I'Y$	$YA = I'TAY$
	$YB = \alpha I'RBY$	

TABLE V. Ray characters of Group 47 belonging to the factor system $\alpha = \beta = 1, \gamma = -1$. For classes 4, 5, 6, 8, 12, 13, 18, 19, 20, the ray characters = 0.

Class	1.	2.	3.	7.	9.	10.	11.	14.	15.	16.	17.
	2	2	0	2	2	2	0	0	2	2	0
	2	−1	$i\sqrt{3}$	2	2	2	0	0	2	−1	$i\sqrt{3}$
	2	−1	$-i\sqrt{3}$	2	2	2	0	0	2	−1	$-i\sqrt{3}$
	6	0	0	6	−2	−2	0	0	6	0	0
	6	0	0	−2	2	−2	2	0	6	0	0
	6	0	0	−2	2	−2	−2	0	6	0	0
	6	0	0	−2	−2	2	0	$i2$	6	0	0
	6	0	0	−2	−2	2	0	$-i2$	6	0	0
	8	2	0	0	0	0	0	0	−8	−2	0
	8	−1	$i\sqrt{3}$	0	0	0	0	0	−8	1	$-i\sqrt{3}$
	8	−1	$-i\sqrt{3}$	0	0	0	0	0	−8	1	$i\sqrt{3}$

TABLE VI. Ray characters of Group 47 belonging to the factor system $\alpha = \gamma = 1, \beta = -1$. For classes 4, 6, 9, 10, 13, 18, 20, the ray characters = 0.

Class	1.	2.	3.	5.	7.	8.	11.	12.	14.	15.	16.	17.	19.
	2	−1	1	−2	2	0	$i\sqrt{2}$	0	$-i\sqrt{2}$	2	−1	1	−2
	2	−1	1	−2	2	0	$-i\sqrt{2}$	0	$i\sqrt{2}$	2	−1	1	−2
	2	−1	−1	2	2	0	$i\sqrt{2}$	0	$i\sqrt{2}$	2	−1	−1	2
	2	−1	−1	2	2	0	$-i\sqrt{2}$	0	$-i\sqrt{2}$	2	−1	−1	2
	4	1	1	−2	0	2	0	2	0	−4	−1	−1	2
	4	1	−1	2	0	−2	0	2	0	−4	−1	1	−2
	4	1	−1	2	0	2	0	−2	0	−4	−1	1	−2
	4	1	1	−2	0	−2	0	−2	0	−4	−1	−1	2
	4	1	1	4	4	0	0	0	0	4	1	1	4
	4	1	−1	−4	4	0	0	0	0	4	1	−1	−4
	8	−1	1	4	0	0	0	0	0	−8	1	−1	−4
	8	−1	−1	−4	0	0	0	0	0	−8	1	1	4
	12	0	0	0	−4	0	0	0	0	12	0	0	0

TABLE VII. Ray characters of Group 47 belonging to the factor system $\alpha = 1, \beta = \gamma = -1$. For classes 4, 5, 6, 8, 9, 10, 12, 18, 19, 20, the ray characters = 0.

Class	1.	2.	3.	7.	11.	13.	14.	15.	16.	17.
	4	1	$i\sqrt{3}$	4	0	0	0	4	1	$i\sqrt{3}$
	4	1	$-i\sqrt{3}$	4	0	0	0	4	1	$-i\sqrt{3}$
	4	−2	0	4	0	0	0	4	−2	0
	6	0	0	−2	$i\sqrt{2}$	$-i2$	$\sqrt{2}$	6	0	0
	6	0	0	−2	$i\sqrt{2}$	$i2$	$-\sqrt{2}$	6	0	0
	6	0	0	−2	$-i\sqrt{2}$	$i2$	$\sqrt{2}$	6	0	0
	6	0	0	−2	$-i\sqrt{2}$	$-i2$	$-\sqrt{2}$	6	0	0
	8	2	0	0	0	0	0	−8	−2	0
	8	−1	$-i\sqrt{3}$	0	0	0	0	−8	1	$i\sqrt{3}$
	8	−1	$i\sqrt{3}$	0	0	0	0	−8	1	$-i\sqrt{3}$

be shown that H is isomorphic to the cubic symmetry group O_h in [3]. The group O_h can be defined by the following relations:

$$A_1^4 = A_2^3 = (A_1A_2^2)^2 = I'^2 = I,$$
$$A_1I' = I'A_1, \quad A_2I' = I'A_2.$$

The order of O_h is 48, its number of classes is 10.

We can write Group 47 and Group 45 in terms of O_h:

$$\text{Group 47} = O_h + O_hA + O_hB + O_hAB + O_hI' + O_hI'A + O_hI'B + O_hI'AB,$$

$$\text{Group 45} = \text{Group 47} + (\text{Group 47})K + (\text{Group 47})K^2,$$

TABLE VIII. Ray characters of Group 47 belonging to the factor system $\alpha = -1, \beta = \gamma = 1$. Ray characters = 0, for all other classes.

Class	1.	2.	4.	9.	11.
	4	−2	$-2\sqrt{2}$	2	$-\sqrt{2}$
	4	−2	$2\sqrt{2}$	2	$\sqrt{2}$
	8	2	0	4	0
	12	0	$2\sqrt{2}$	−2	$-\sqrt{2}$
	12	0	$-2\sqrt{2}$	−2	$\sqrt{2}$

TABLE IX. Ray characters of Group 47 belonging to the factor system $\alpha = \gamma = -1, \beta = 1$. Ray characters = 0, for all other classes.

Class	1.	2.	9.	11.	18.
	4	−2	2	$-\sqrt{2}$	$-2\sqrt{2}$
	4	−2	2	$\sqrt{2}$	$2\sqrt{2}$
	8	2	4	0	0
	12	0	−2	$-\sqrt{2}$	$2\sqrt{2}$
	12	0	−2	$\sqrt{2}$	$-2\sqrt{2}$

TABLE X. Ray characters of Group 47 belonging to the factor system $\alpha = \beta = -1, \gamma = 1$. Ray characters = 0, for all other classes.

Class	1.	2.	11.
	8	2	$i2$
	8	2	$-i2$
	16	−2	0

TABLE XI. Ray characters of Group 47 belonging to the factor system $\alpha = \beta = \gamma = -1$. Ray characters = 0, for all other classes.

Class	1.	2.	11.
	8	2	$i2$
	8	2	$-i2$
	16	−2	0

TABLE XIII. Defining relations of Group 45*. $g = 1152$, $r = 25$, $g^* = 4608$, $r^* = 59$.

$\alpha^2 I = I$,	$T^2 = R^2 = \alpha I$,	$S^3 = W^8 = I$,
$\alpha = \pm 1$	$RT = \alpha TR$,	$ST = TRS$,
	$SR = TS$,	$W^2 = T$,
	$WR = TRW$,	$WS = \alpha TS^2 W$,
	$A^2 = B^2 = I'$,	$I'^2 = Y^2 = I$,
	$BA = I'AB$,	$AT = TA$,
	$AR = I'RA$,	$AS = I'SB$,
	$AW = WA$.	$BT = I'TB$,
	$BR = RB$,	$BS = SAB$,
	$BW = I'WAB$,	
	$I'T = TI'$,	$I'R = RI'$,
	$I'S = SI'$;	$I'W = WI'$,
	$I'A = AI'$,	$I'B = BI'$,
$\gamma^2 I = I$	$YT = TY$,	$YR = RY$,
$\gamma = \pm 1$	$YS = SY$,	$YW = \gamma WY$,
	$YA = I'TAY$,	$YB = I'RBY$,
	$YI' = \alpha I'Y$,	
	$K^3 = I$,	
	$KT = I'TBK$,	$KR = I'RABK$,
	$KS = SK$,	$KW = WBK^2$,
	$KA = I'ABK$,	$KB = I'AK$,
	$KY = SYK^2$,	$KI' = I'K$.

where the definitions of the generators A, B, I', and K are given in Table I.

The irreducible vector representations of Group 47 and Group 45 will subduce representations of O_h, which are listed in Tables XVII and XVIII.

5. SUMMARY AND OBSERVATIONS

The method given here has been found to be straightforward and easily used with available computer routines for the diagonalization of real symmetric matrices. It has the advantage of requiring only a single diagonalization of an $r \times r$ matrix: The remainder of the work can then be done by solving (2.31) by inspection or by a fur-

TABLE XII. Vector Characters of Group 45.

Class	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.
Typical Element	I	K	S	SK	SY	WYK	W	Y	RWY	A	SWA	RK	$\acute{R}$	RW	WY	YA	I'	$I'K$	$I'S$	$I'SK$	$I'SY$	$I'WYK$	$I'W$	$I'Y$	$I'RWY$
r_i	1	16	32	32	96	96	36	12	12	12	144	96	18	72	72	72	1	16	32	32	96	96	36	12	12
$\Gamma^{(1)}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$\Gamma^{(2)}$	1	1	1	1	−1	−1	1	−1	−1	1	1	1	1	1	−1	−1	1	1	1	1	−1	−1	1	−1	−1
$\Gamma^{(3)}$	1	1	1	1	−1	1	−1	−1	1	1	−1	1	1	−1	1	−1	1	1	1	1	−1	1	−1	−1	1
$\Gamma^{(4)}$	1	1	1	1	1	−1	−1	1	−1	1	−1	1	1	−1	−1	1	1	1	1	1	1	−1	−1	1	−1
$\Gamma^{(5)}$	2	−1	2	−1	0	−1	0	0	2	2	0	−1	2	0	2	0	2	−1	2	−1	0	−1	0	0	2
$\Gamma^{(6)}$	2	−1	2	−1	0	1	0	0	−2	2	0	−1	2	0	−2	0	2	−1	2	−1	0	1	0	0	−2
$\Gamma^{(7)}$	2	−1	−1	2	−1	0	0	2	0	2	0	−1	2	0	0	2	2	−1	−1	2	−1	0	0	2	0
$\Gamma^{(8)}$	2	−1	−1	2	1	0	0	−2	0	2	0	−1	2	0	0	−2	2	−1	−1	2	1	0	0	−2	0
$\Gamma^{(9)}$	4	1	−2	−2	0	0	0	0	0	4	0	1	4	0	0	0	4	1	−2	−2	0	0	0	0	0
$\Gamma^{(10)}$	4	−2	1	1	−1	1	−2	2	2	0	0	0	0	0	0	0	−4	2	−1	−1	1	−1	2	−2	−2
$\Gamma^{(11)}$	4	−2	1	1	1	−1	−2	−2	−2	0	0	0	0	0	0	0	−4	2	−1	−1	−1	1	2	2	2
$\Gamma^{(12)}$	4	−2	1	1	−1	−1	2	2	−2	0	0	0	0	0	0	0	−4	2	−1	−1	1	1	−2	−2	2
$\Gamma^{(13)}$	4	−2	1	1	1	1	2	−2	2	0	0	0	0	0	0	0	−4	2	−1	−1	−1	−1	−2	2	−2
$\Gamma^{(14)}$	6	3	0	0	0	0	2	0	0	2	0	−1	−2	−2	0	0	6	3	0	0	0	0	2	0	0
$\Gamma^{(15)}$	6	3	0	0	0	0	−2	0	0	2	0	−1	−2	2	0	0	6	3	0	0	0	0	−2	0	0
$\Gamma^{(16)}$	8	2	2	−1	0	1	0	0	−4	0	0	0	0	0	0	0	−8	−2	−2	1	0	−1	0	0	4
$\Gamma^{(17)}$	8	2	2	−1	0	−1	0	0	4	0	0	0	0	0	0	0	−8	−2	−2	1	0	1	0	0	−4
$\Gamma^{(18)}$	8	2	−1	2	1	0	0	4	0	0	0	0	0	0	0	0	−8	−2	1	−2	−1	0	0	−4	0
$\Gamma^{(19)}$	8	2	−1	2	−1	0	0	−4	0	0	0	0	0	0	0	0	−8	−2	1	−2	1	0	0	4	0
$\Gamma^{(20)}$	9	0	0	0	0	0	1	3	3	−3	−1	0	1	1	−1	−1	9	0	0	0	0	0	1	3	3
$\Gamma^{(21)}$	9	0	0	0	0	0	1	−3	−3	−3	−1	0	1	1	1	1	9	0	0	0	0	0	1	−3	−3
$\Gamma^{(22)}$	9	0	0	0	0	0	−1	3	−3	−3	1	0	1	−1	1	−1	9	0	0	0	0	0	−1	3	−3
$\Gamma^{(23)}$	9	0	0	0	0	0	−1	−3	3	−3	1	0	1	−1	−1	1	9	0	0	0	0	0	−1	−3	3
$\Gamma^{(24)}$	12	−3	0	0	0	0	0	0	0	4	0	1	−4	0	0	0	12	−3	0	0	0	0	0	0	0
$\Gamma^{(25)}$	16	−2	−2	−2	0	0	0	0	0	0	0	0	0	0	0	0	−16	2	2	2	0	0	0	0	0

TABLE XIV. Ray characters of Group 45 belonging to the factor system $\alpha = 1, \gamma = -1$. Classes 7, 8, 9, 14, 15, 16, 23, 24, 25 provide zero ray characters only.

Class	1.	2.	3.	4.	5.	6.	10.	11.	12.	13.	17.	18.	19.	20.	21.	22.
	2	2	2	2	0	0	2	0	2	2	2	2	2	2	0	0
	2	−1	−1	2	$i\sqrt{3}$	0	2	0	−1	2	2	−1	−1	2	$i\sqrt{3}$	0
	2	−1	−1	2	$-i\sqrt{3}$	0	2	0	−1	2	2	−1	−1	2	$-i\sqrt{3}$	0
	2	−1	2	−1	0	$\sqrt{3}$	2	0	−1	2	2	−1	2	−1	0	$\sqrt{3}$
	2	−1	2	−1	0	$-\sqrt{3}$	2	0	−1	2	2	−1	2	−1	0	$-\sqrt{3}$
	4	1	−2	−2	0	0	4	0	1	4	4	1	−2	−2	0	0
	6	3	0	0	0	0	2	2	−1	−2	6	3	0	0	0	0
	6	3	0	0	0	0	2	−2	−1	−2	6	3	0	0	0	0
	8	−4	2	2	0	0	0	0	0	0	−8	4	−2	−2	0	0
	8	2	−1	2	$-i\sqrt{3}$	0	0	0	0	0	−8	−2	1	−2	$i\sqrt{3}$	0
	8	2	−1	2	$i\sqrt{3}$	0	0	0	0	0	−8	−2	1	−2	$-i\sqrt{3}$	0
	8	2	2	−1	0	$-\sqrt{3}$	0	0	0	0	−8	−2	−2	1	0	$\sqrt{3}$
	8	2	2	−1	0	$\sqrt{3}$	0	0	0	0	−8	−2	−2	1	0	$-\sqrt{3}$
	12	−3	0	0	0	0	4	0	1	−4	12	−3	0	0	0	0
	16	−2	−2	−2	0	0	0	0	0	0	−16	2	2	2	0	0
	18	0	0	0	0	0	−6	0	0	2	18	0	0	0	0	0

TABLE XV. Ray characters of Group 45 belonging to the factor system $\alpha = -1$, $\gamma = 1$. Other classes provide only zero ray characters.

Class	1.	2.	3.	4.	7.	10.	11.	12.	18.
	4	1	−2	−2	$-2\sqrt{2}$	2	$-\sqrt{2}$	−1	3
	4	1	−2	−2	$2\sqrt{2}$	2	$\sqrt{2}$	−1	3
	8	−1	2	−4	0	4	0	1	−3
	8	−1	−4	2	0	4	0	1	−3
	8	−4	2	2	0	4	0	−2	0
	8	5	2	2	0	4	0	1	3
	12	−3	0	0	$2\sqrt{2}$	−2	$-\sqrt{2}$	1	3
	12	−3	0	0	$-2\sqrt{2}$	−2	$\sqrt{2}$	1	3
	24	3	0	0	0	−4	0	−1	−3

TABLE XVI. Ray characters of Group 45 belonging to the factor system $\alpha = \gamma = -1$. Other classes provide only zero ray characters.

Class	1.	2.	3.	4.	10.	11.	12.	18.	23.
	4	1	−2	−2	2	$-\sqrt{2}$	−1	3	$-2\sqrt{2}$
	4	1	−2	−2	2	$\sqrt{2}$	−1	3	$2\sqrt{2}$
	8	−1	2	−4	4	0	1	−3	0
	8	−1	−4	2	4	0	1	−3	0
	8	−4	2	2	4	0	−2	0	0
	8	5	2	2	4	0	1	3	0
	12	−3	0	0	−2	$-\sqrt{2}$	1	3	$2\sqrt{2}$
	12	−3	0	0	−2	$\sqrt{2}$	1	3	$-2\sqrt{2}$
	24	3	0	0	−4	0	−1	−3	0

TABLE XVII. Representations of O_h subduced by the irreducible vector representations of Group 47. $\Gamma_{47}^{(\mu)}$, with $\mu = 1$–20, are representations of O_h subducted by the irreducible vector representations of Group 47, the dimensionalities of which are $d_1 = d_2 = d_3 = d_4 = 1$, $d_5 = d_6 = 2$, $d_7 = d_8 = d_9 = d_{10} = 3$, $d_{11} = d_{12} = d_{13} = d_{14} = 4$, $d_{15} = d_{16} = d_{17} = d_{18} = 6$, $d_{19} = d_{20} = 8$. $\Gamma_{O_h}^{(\nu)}$, with $\nu = 1$–10, are the irreducible vector representations of O_h, the dimensionalities of which are $d_1 = d_2 = d_{1-} = d_{2-} = 1$, $d_{12} = d_{12-} = 2$, $d_{15} = d_{25} = d_{15-} = d_{25-} = 3$.

$\Gamma_{47}^{(\mu)} = \sum_\nu a_{\mu\nu} \Gamma_{O_h}^{(\nu)}$		$\Gamma_{47}^{(\mu)} = \sum_\nu a_{\mu\nu} \Gamma_{O_h}^{(\nu)}$	
(1)	(1)	(11)	(2) ⊕ (25−)
(2)	(2)	(12)	(25) ⊕ (2−)
(3)	(1−)	(13)	(15) ⊕ (1−)
(4)	(2−)	(14)	(1) ⊕ (15−)
(5)	(12−)	(15)	(15) ⊕ (15−)
(6)	(12)	(16)	(25) ⊕ (25−)
(7)	(1) ⊕ (12)	(17)	(25) ⊕ (15−)
(8)	(2) ⊕ (12)	(18)	(15) ⊕ (25−)
(9)	(2−) ⊕ (12−)	(19)	(25) ⊕ (15) ⊕ (12−)
(10)	(1−) ⊕ (12−)	(20)	(12) ⊕ (25−) ⊕ (15−)

TABLE XVIII. Representations of O_h subduced by the irreducible vector representations of Group 45. $\Gamma_{45}^{(\mu)}$, with $\mu = 1$–25, are representations of O_h subducted by the irreducible vector representations of Group 45, the dimensionalities of which are $d_1 = d_2 = d_3 = d_4 = 1$, $d_5 = d_6 = d_7 = d_8 = 2$, $d_9 = d_{10} = d_{11} = d_{12} = d_{13} = 4$, $d_{14} = d_{15} = 6$, $d_{16} = d_{17} = d_{18} = d_{19} = 8$, $d_{20} = d_{21} = d_{22} = d_{23} = 9$, $d_{24} = 12$, $d_{25} = 16$.

$\Gamma_{45}^{(\mu)}$	$= \sum_\nu a_{\mu\nu} \Gamma_{47}^{(\nu)}$	$= \sum_\omega b_{\mu\omega} \Gamma_{O_h}^{(\omega)}$
(1)	(1)	(1)
(2)	(3)	(1−)
(3)	(4)	(2−)
(4)	(2)	(2)
(5)	(1) ⊕ (4)	(1) ⊕ (2−)
(6)	(2) ⊕ (3)	(2) ⊕ (1−)
(7)	(6)	(12)
(8)	(5)	(12−)
(9)	(5) ⊕ (6)	(12−) ⊕ (12)
(10)	(12)	(25) ⊕ (2−)
(11)	(11)	(2) ⊕ (25−)
(12)	(13)	(15) ⊕ (1−)
(13)	(14)	(1) ⊕ (15−)
(14)	(15)	(15) ⊕ (15−)
(15)	(16)	(25) ⊕ (15−)
(16)	(11) ⊕ (13)	(2) ⊕ (25−) ⊕ (15) ⊕ (1−)
(17)	(12) ⊕ (14)	(25) ⊕ (2−) ⊕ (1) ⊕ (15−)
(18)	(19)	(25) ⊕ (15) ⊕ (12−)
(19)	(20)	(12) ⊕ (25−) ⊕ (15−)
(20)	(7) ⊕ (16)	(1) ⊕ (12) ⊕ (25) ⊕ (25−)
(21)	(10) ⊕ (18)	(1−) ⊕ (12−) ⊕ (15) ⊕ (25−)
(22)	(8) ⊕ (18)	(2) ⊕ (12) ⊕ (15) ⊕ (25−)
(23)	(9) ⊕ (17)	(2−) ⊕ (12−) ⊕ (25) ⊕ (15−)
(24)	(15) ⊕ (16)	(15) ⊕ (15−) ⊕ (25) ⊕ (25−)
(25)	(19) ⊕ (20)	(25) ⊕ (15) ⊕ (12−) ⊕ (12) ⊕ (25−) ⊕ (15−)

ther sequence of diagonalizations of much smaller matrices ($n \times n$), as in the Appendix. We believe it should be very useful even for groups of still larger order and even more complicated structure. Of course, in any event, one requires the structure constants of the group as input data. Since one is using Eqs. (2.2)–(2.4), the orthonormality relations for the group characters are automatically satisfied.

Turning to the results for Group 47, Group 45, and O_h[3] we have given in the subduction tables, we can make two observations. These may be related to the original conjecture on a relation between

accidental degeneracy in [3] (for example, in O_h) and a higher symmetry in [4] (for example, in the Group 47 and Group 45.) First note that, with few exceptions, the representations of O_h which "stick" to give a representation of 47 or 45 are of different parity. Thus, from Table XVII, (15—) + (15+) of O_h combine to give (15) of 47, and similarly for 45. Thus a sticking of states of different parity in O_h may be related to the higher group symmetry.

Secondly, observe that the representations of O_h which "stick" are never two of the same species. Equivalently stated: In the subduction from $\Gamma^{(j)}$ of 47, or 45, one never finds a $\Gamma^{(k)}$ of O_h more than once. This seems related to the "no crossing" rules so well known already: Two states of the same symmetry (e.g., the symmetry $\Gamma^{(k)}$ of O_h) cannot touch.

APPENDIX

A method will be given for solving the bilinear equations (2.31) obtained in Sec. 2B. This completes our method of obtaining characters of irreducible vector representations of finite groups.

We have, for $d_1 = d_2 = \cdots = d_\mu = \cdots = d_n$,

$$a_{\mu\alpha}a_{\mu\beta} = d_\mu \sum_{\gamma=1}^{n} f_{\alpha\beta\gamma}a_{\mu\gamma}, \qquad \mu, \alpha, \beta = 1, 2, \ldots, n, \tag{2.31}$$

with

$$f_{\alpha\beta\gamma} \equiv \sum_{i,j,k=1}^{r} \sqrt{\frac{\gamma_k}{g\gamma_i\gamma_j}}\, h_{ij,k} V_{\alpha i} V_{\beta j} V_{\gamma k}. \tag{2.32}$$

The coefficients $f_{\alpha\beta\gamma}$ are real and given. Rewrite (2.31) in the form

$$\sum_{\gamma=1}^{n} (d_\mu f_{\alpha\beta\gamma} - a_{\mu\alpha}\delta_{\beta\gamma})a_{\mu\gamma} = 0 \tag{A1}$$

For a given α, we can define a matrix M^α with

$$M^\alpha_{\beta\gamma} \equiv d_\mu f_{\alpha\beta\gamma}. \tag{A2}$$

Then (A1) implies that $a_{\mu\alpha}$, $\mu = 1, 2, \ldots, n$, are the eigenvalues of M^α, associated with eigenvectors $a_{\mu\gamma}$, $\gamma = 1, 2, \ldots, n$, which are independent of the index α. It can easily be shown that M^α is a normal matrix; thus there exists in general a unitary matrix V' which diagonalizes

$$M^\alpha \quad \text{or} \quad \sum_{\gamma=1}^{n} M^\alpha_{\beta\gamma} V'_{\mu\gamma} = a_{\mu\alpha} V'_{\mu\beta}. \tag{A3}$$

As analysed in Sec. 2B, if an eigenvalue $a_{\mu\alpha}$ is distinct, its associated eigenvector is uniquely determined up to a $\pm$ sign, i.e.,

$$a_{\mu\gamma} = \pm\, V'_{\mu\gamma}. \tag{A4}$$

The sign will be determined by using (2.26) and (2.28).

If we have multiple eigenvalues $a_{1\alpha} = a_{2\alpha} = \cdots = a_{n'\alpha'}$,[11] then we have to make a unitary transformation on the $V'_{\mu\gamma}$. Similar to (2.28), we have

$$a_{\mu\gamma} = \sum_{\nu}^{n'} b_{\mu\nu} V'_{\nu\gamma}, \quad \mu = 1, 2, \ldots, n', \quad \gamma = 1, 2, \ldots, n. \tag{A5}$$

Substituting (A5) into (2.31), we have

$$b_{\mu\alpha}b_{\mu\beta} = d_\mu \sum_{\gamma=1}^{n'} f'_{\alpha\beta\gamma} b_{\mu\gamma}, \quad \mu, \alpha, \beta = 1, 2, \ldots, n', \tag{A6}$$

with

$$f'_{\alpha\beta\gamma} = \sum_{i,j,k=1}^{n'} f_{ijk} V'_{\gamma k} V'^{*}_{\alpha i} V'^{*}_{\beta j}. \tag{A7}$$

The coefficients $f'_{\alpha\beta\gamma}$ are complex and given. Equations (A6) and (A7) are similar to (2.31) and (2.32), except that we have a smaller matrix $b_{\mu\alpha}$. We can apply the same method as described above to solve for the n'^2 $b_{\mu\alpha}$. Continuing in this way, each time the dimension of the unknown matrix becomes smaller, we finally have all the $a_{\mu\gamma}$ from a sequence of equations of the form (A5).

Once the $a_{\mu\gamma}$ are known, characters can be obtained from Eqs. (2.28) and (2.8).

The difference between (2.31) and (2.4) or (2.9) is that

(i) In (2.31) the d_μ, $\mu = 1, 2, \ldots, n$, are known and equal.

(ii) Here we always deal with a normal matrix; thus there exists a complete orthonormal set of eigenvectors which diagonalize the matrix.

---

* Some of the work in this paper was contained in a thesis for the Ph.D. degree presented to the Physics Department, New York University, June, 1970 by Li-Ching Chen. The work was supported in part by the Army Research Office Durham and the Aerospace Research Laboratories WPAFB, Dayton, Ohio. Publication supported by the National Science Foundation.

† Present Address: Instituut voor Theoretische Fysica, Catholic University, Nijmegen, Netherlands.

1 J. S. Lomont, *Applications of Finite Groups* (Academic, New York, 1959).

2 A. C. Hurley, Proc. Cambridge Phil. Soc. **47**, 650 (1951).

3 M. Hamermesh, *Group Theory* (Addison-Wesley, New York, 1964).

4 W. Burnside, *Theory of Groups of Finite Order* (Dover, New York, 1955).

5 The computer routine we used is called *Subroutine Eigen*, which is based on the diagonalization method originated by Jacobi and adapted by von Neumann for large computers as found in *Mathematical Methods for Digital Computers*, edited by A. Ralston and H. S. Wilf (Wiley, New York, 1962), Chap. 7.

6 A. C. Hurley, *Quantum Theory of Atoms, Molecules, and the Solid State*, edited by P. O. Lowdin (Academic, New York, 1967).

7 E. Goursat, Ann. Sci. Ecole Norm. Sup. Paris **3**, 6 (1889).

8 A. L. Mackay and G. S. Pawley, Acta Cryst. **16**, 11 (1963).

9 W. Z. Döring, Naturforschr. **14**, 343 (1959).

10 Li-Ching Chen, Ph.D thesis (New York University, 1970), available from University Microfilms Inc., Ann Arbor., Michigan.

11 From (2.28) we can show that the $n \times n$ matrix $(a_{\mu\alpha})$ is unitary. We cannot have two columns of $(a_{\mu\alpha})$ such that all elements are equal in each column. Thus there exist $(n-1)$ columns such that $n' < n$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

A Note on Asymptotically Flat Spaces*

B. Aronson, R. Lind, J. Messmer, and E. Newman
(Faculty of Arts and Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15213
(Received 26 May 1971)

It is shown that there is a simple alternative to the Bondi type of coordinate condition in asymptotically flat spaces, which leads to a different asymptotic coordinate group whose structure is simpler than that of the BMS group. As a concomitant to this work, there was discovered a *scalar* function on null infinity. From an analysis of this function one can obtain, starting from the asymptotic shear of one family of hypersurface orthogonal rays, (1) all null asymptotically shear-free congruences, (2) the asymptotic shear of all hypersurface orthogonal rays.

1. INTRODUCTION

In the early 1960's Bondi[1] *et al.* and then Sachs,[2,3] while studying asymptotically flat solutions to the Einstein field equations, discovered and studied a group now known as the Bondi–Metzner–Sachs (BMS) group, which briefly can be looked upon as the semidirect product of the homogeneous Lorentz group and an infinite-dimensional Abelian group[4] of which the first four elements can be identified with the translations of the Poincaré group, the remaining elements being known as supertranslations. The BMS group first arose as the coordinate group that preserved certain *coordinate conditions* (we will call them Bondi type) which were imposed at infinity. Later Sachs[3] and Penrose[5,6] realized that there was an intrinsic geometric significance to the BMS group; namely, that it was generated by the conformal Killing vectors of the null surface at infinity. Though a great deal of effort has gone into studying the geometric meaning, the group structure and representations[3,7] of the BMS group, little of physical relevance seems to have emerged. Almost to the contrary it has proved to be a hindrance in understanding the physical variables in gravitational radiation theory. Though energy and linear momentum are well defined via the generators of the translation subgroup, angular momentum is not.[8,9] Another difficulty is the vast proliferation of generators known as supermomenta, arising from the supertranslation part of the BMS group. Owing to these difficulties, considerable thought has gone into the question of whether further conditions (either coordinate or boundary conditions) could be imposed to restrict the asymptotic coordinate group to the Poincaré group. This effort has been either unsuccessful[10] or artificial.[11]

The first purpose of the present note is to show that it is possible by means of asymptotic coordinate conditions different from the Bondi type to obtain a different group. This new group, though still large, is considerably smaller than the BMS group. Its structure is that of the direct product of the homogeneous Lorentz group with an Abelian group, the Abelian group being closely associated with the transformations leading from one arbitrary timelike world line to another in Minkowski space. It thus appears as if this group has a more intuitive meaning than the BMS. It should be emphasized that we are not suggesting that this group replace the BMS group as the asymptotic symmetry group, but as the asymptotic coordinate group.

Of possibly equal or even greater interest was the closely related discovery of a *scalar* function on null infinity (on J^+ in Penrose's conformal language) from which can be obtained by differentiation

(a) null rays which are asymptotically shear-free (but twisting) and

(b) the asymptotic shear or news function.

In Sec. 2, we review some pertinent properties of asymptotically flat spaces. In Sec. 3 the scalar field on null infinity is introduced which in turn is used in Sec. 4 to impose new coordinate conditions. The remaining coordinate freedom is discussed. In Sec. 5, by computing the commutator of two of the allowed coordinate transformations, the fact that they form an Abelian group is ascertained. In Sec. 6 the geometric meaning of one of the basic variables is discussed.

2. NULL INFINITY

In this section where we discuss some properties of null infinity, we will use the descriptive language of the conformal infinity of Penrose, though all of the analysis will be done in physical space. We thus look upon infinity as a three-dimensional null surface with a degenerate line element, g being the null rays or generators of the surface with tangent vectors n^μ.

The topology is $S^2 \times R$. (See Fig. 1). The surface referred to as J^+ is coordinated in the following fashion: (a) Since the generators g can be mapped one-to-one on S^2, they can be labeled by the complex stereographic coordinates of a sphere, ζ and $\bar{\zeta}$; and (b) arbitrary spacelike, nonintersecting cuts can be taken in J^+ and labeled by $u = \text{const}$ in a monotonic fashion. The permitted coordinate transformations are (a) the one-to-one conformal transformations of the sphere onto itself, i.e., the relabeling of the generators by

$$\zeta' = f(\zeta) = \frac{a\zeta + b}{c\zeta + d}, \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 1, \tag{2.1}$$

and (b) taking different families of cuts through J^+,

$$u' = G(u, \zeta, \bar{\zeta}). \tag{2.2}$$

Choosing a particular coordinate system on J^+ permits us to coordinatize uniquely a neighborhood of J^+ in the four-dimensional space–time by the following procedure: from each point of each $u = \text{const}$ cut, take the null geodesic with tangent vector l^μ, which strikes the cut orthogonally and comes

from the interior. (See Fig. 1.) This leads to a family of null surfaces each being labeled by its cut of J^+, the generators of these surfaces being labeled by their intersection with the generators of J^+. If we choose the "distance" along each generator as the affine length r (properly normalized[12]), we get a unique null coordinate system. It was this system that was used in NU[12] to describe the asymptotically flat solutions to Einstein's equations.

In NU it was shown that there were two functions defined on J^+ that constituted the free data to determine a solution of Einstein's equations. (There are other data, not given on J^+, which do not concern us.) The first of these is the real function $P(u, \zeta, \bar{\zeta})$ defined from the metric of the limiting two-surfaces u and r const by

$$\lim_{r\to\infty} r^{-2}\, ds^2 = \frac{d\zeta d\bar{\zeta}}{P^2}.$$

In other words knowing P is equivalent to knowing the metric of the two-dimensional cuts of J^+.

The second function, $\sigma^0(u, \zeta, \bar{\zeta})$, is the (complex) asymptotic value of the shear σ of the null rays l^μ.

It is defined from σ by

$$\sigma^0 = \lim_{r\to\infty} r^2\sigma\ .$$

It is important for us to know how both σ^0 and P transform under (2.1) and (2.2). This can be inferred from NU (after a very long calculation) with the result that

(a) from (2.1),

$$\sigma'^0 = e^{2i\lambda}\sigma^0,$$
$$P'^2 = f_{,\zeta}\bar{f}_{\bar{\zeta}}P^2, \tag{2.3}$$

where $e^{2i\lambda} = (\bar{f}_{\bar{\zeta}}/f_{,\zeta}) = [(c\zeta + d)/(\overline{c\zeta} + \bar{d})]^2$,

(b) from (2.2),

$$\sigma'^0 = \dot{G}^{-1}\sigma^0 + \dot{G}^{-4}\ddot{G}\eth G\cdot\eth G - 2\dot{G}^{-3}\eth G\cdot\eth\dot{G} + \dot{G}^{-2}\eth^2 G - \frac{\dot{P}}{P}\dot{G}^{-3}\eth G\cdot\eth G, \tag{2.4}$$

$$P' = \dot{G}^{-1}P, \tag{2.5}$$

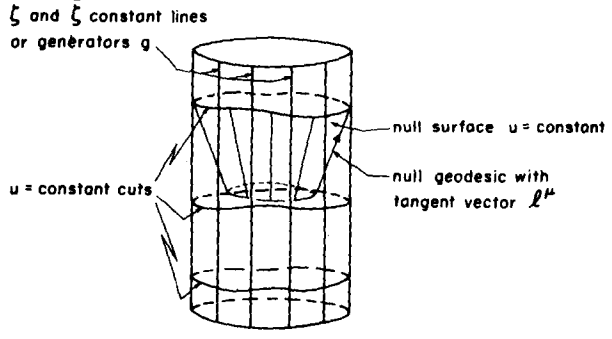

FIG. 1. Future null infinity, J^+.

where $\dot{G} \equiv G_{,u}$ and $\eth$ is an angular differential operator.[13,11]

It is seen that though we have three real functions (P and the real and imaginary parts of σ^0) to specify freely, only two are really essential because we still have the arbitrary coordinate transformation $u' = G(u, \zeta, \bar{\zeta})$ to impose a restriction.

For example, $\dot{G}$ could be chosen to make $P' = \frac{1}{2}(1 + \zeta\bar{\zeta})$, in which case only σ^0 can be chosen freely. This condition is equivalent to the Bondi-type coordinate conditions. In fact it is easily seen from (2.3) and (2.5) that the remaining coordinate freedom would be just the BMS group.

In the above paragraph we used (2.5) to impose conditions on P. The question arises whether or not we can use (2.4) to impose conditions on σ^0 to eliminate one of its two degrees of freedom and, if so, what the remaining coordinate freedom will be. The remainder of the paper is devoted to answering these questions.

3. THE SCALAR FUNCTION ϕ on J^+

We will now introduce a "potential" for σ^0 by the nonlinear equation

$$\sigma^0 = \eth L + LL_{,u} \tag{3.1}$$

and a potential for L by

$$0 = \eth\phi + L\phi_{,u}. \tag{3.2}$$

We claim that, though σ^0 has the very complicated transformation law (2.4) under $u' = G(u, \zeta, \bar{\zeta})$, L will transform as

$$L' = L + \dot{G}^{-1}\eth G \tag{3.3}$$

and ϕ will transform as a scalar

$$\phi'(u', \zeta, \bar{\zeta}) = \phi(u, \zeta, \bar{\zeta}), \quad u' = G(u, \zeta, \bar{\zeta}). \tag{3.4}$$

The easiest way to prove these assertions is to work backwards starting from (3.4). First note that from the transformations

$$u' = G(u, \zeta, \bar{\zeta}), \quad \zeta' = \zeta,$$

we immediately get

$$1 = \dot{G}\frac{\partial u}{\partial u'} \to \frac{\partial u}{\partial u'} = \dot{G}^{-1},$$
$$0 = \dot{G}\frac{\partial u}{\partial\zeta'} + G_{,\zeta} \to \frac{\partial u}{\partial\zeta'} = -\dot{G}^{-1}G_{,\zeta}. \tag{3.5}$$

Then by differentiating (3.4) with respect to u' and ζ', we obtain

$$\phi'_{,u'} = \phi_{,u}\frac{\partial u}{\partial u'} = \phi_{,u}\dot{G}^{-1},$$
$$\phi'_{,\zeta'} = \phi_{,\zeta} + \phi_{,u}\frac{\partial u}{\partial\zeta'} = \phi_{,\zeta} - \phi_{,u}\dot{G}^{-1}G_{,\zeta}. \tag{3.6}$$

Taking the ratio, yields

$$\phi'_{,\zeta'}/\phi'_{,u'} = \dot{G}\phi_{,\zeta}/\phi_{,u} - G_{,\zeta} \tag{3.7}$$

If we multiply by $P' = \dot{G}^{-1}P$ and use the definition of $\eth$ acting on a spin-weight zero quantity, we have

$$\frac{\eth'\phi'}{\phi'_{,u'}} = \frac{\eth\phi}{\phi_{,u}} - \dot{G}^{-1}\eth G. \tag{3.8}$$

This is precisely (3.3) if we take $L = -\eth\phi/\phi_{,u}$ from (3.2). By continuing this process, the transformation law of σ^0 obtained from (3.1) and (3.3) can be shown to be the same as (2.4).

In addition it can be shown that the σ^0 derived from ϕ by (3.1) and (3.2) transforms correctly under (2.1) if ϕ is assumed to be of spin and conformal weight zero.[14]

It should be emphasized that the L and ϕ derived from a σ^0 are not unique. This lack of uniqueness will be mirrored in the coordinate freedom after we impose our new coordinate conditions.

In addition, though we do not know the geometric significance of ϕ, L has a clear and significant meaning, the discussion of which will be postponed till later.

4. COORDINATE CONDITIONS

Assume that we are given J^+ with a slicing, as well as the σ^0 and the metric, $P(u,\zeta,\bar{\zeta})$ associated with the slicing. If we transform to a second slicing, the new σ'^0 and P' could be calculated directly from their transformation laws. It would, however, be much easier if a potential $\phi(u,\zeta,\bar{\zeta})$ for the original σ^0 were known, for by just differentiating it in the new coordinate system the new σ'^0 would be obtained. In other words the difficult transformation law for σ^0 can be shifted to the simple law for ϕ. It is this observation that permits us to choose a canonical family of slicings.

From a given σ^0 and P, assume that we have a ϕ written in the form

$$\phi_1 = \alpha_1 + i\beta_1. \tag{4.1}$$

If we now perform the transformation

$$u' = G(u,\zeta,\bar{\zeta}) \equiv \alpha_1(u,\zeta,\bar{\zeta}) \tag{4.2}$$

the ϕ becomes

$$\phi'_1 = u' + i\beta'_1(u',\zeta,\bar{\zeta}). \tag{4.3}$$

If a second ϕ, from the same σ^0 and P were found, i.e.,

$$\phi_2 = \alpha_2 + i\beta_2,$$

the transformation $u'' = \alpha_2$ would put it into the form

$$\phi''_2 = u'' + i\beta''_2.$$

Whenever a slicing or coordinate system is found such that ϕ is in the form (4.3) we will refer to it as a canonical slicing and the ϕ and the associated σ^0 and L will be referred to, respectively, as a canonical ϕ, σ^0, and L. It is clear that the number of different canonical slicings are equal to the number of different (real parts of) solutions to the equations $\sigma^0 = \eth L + LL_{,u}$, $0 = \eth\phi + L\phi_{,u}$.

The coordinate conditions we then choose are the canonical slicings, with the remaining freedom being the transformations between canonical slicings.

In order to obtain a feel for the amount of the freedom left, we shall study the equations

$$\sigma^0 = \eth L + LL_{,u}, \quad 0 = \eth\phi + L\phi_{,u}. \tag{4.4}$$

We first write P in the form

$$P = P_0 V \text{ with } P_0 = \tfrac{1}{2}(1 + \zeta\bar{\zeta})$$

and rewrite (4.4) as

$$\sigma^0 = \eth_0(VL) + LL_{,u}, \quad 0 = V\eth_0\phi + L\phi_{,u} \tag{4.5}$$

with $\eth_0$ being taken with respect to P_0. There is a *unique integral* operator[15] denoted by $\eth_0^\dagger$ which is the generalized inverse to $\eth_0$. With its aid, (4.5) can be converted to the integral equations

$$L = V^{-1}\left(\eth_0^\dagger(\sigma^0 - LL_{,u}) + \sum_{m=-1}^{1} A_m(u)\,{}_1Y_{1m}\right) \tag{4.6}$$

$$\phi = \eth_0^\dagger \frac{L}{V}\phi_{,u} + A_0(u)Y_{00}, \tag{4.7}$$

with the $A_m(u)$ and $A_0(u)$ being four arbitrary functions of u. It is clear from this that the freedom of solutions depends only on these four functions, which, therefore, must also determine the coordinate freedom.

This is not to say that we know or can explicitly obtain expressions for the allowed coordinate transformations. Nevertheless, we will be able to show that these transformations form an Abelian group.

5. THE COMMUTATOR

The main idea is to first choose a canonical slicing, i.e., $\phi_c = u + i\beta$, and then to use the associated canonical σ_c^0 to find other neighboring solutions $\phi = \phi_c + \epsilon\psi_1$ from which the coordinate transformation to a new canonical slicing may be found. From two of these transformations the infinitesimal commutator can be formed. Unfortunately, it is very difficult to carry out this program. Use had to be made of the following *artifice*. No place in Secs. 3 and 4 was essential use made of the facts that u is a real variable and that $\bar{\zeta}$ is the complex conjugate of ζ. In fact all the arguments for the introduction of the scalar ϕ remain unchanged even if $u' = G(u,\zeta,\bar{\zeta})$ is a complex transformation. (Though we are here treating this complexification as a pure artifice with no significance,

it should be noted that a special case of it is the complex coordinate transformation which led from the Schwarzschild solution to the Kerr metric.[16] We use this complex G to impose stronger "coordinate conditions" than were used in Sec. 4, namely we choose $u' = G(u,\zeta,\bar{\zeta}) = \phi(u,\zeta,\bar{\zeta})$, so that the new "canonical" form is $\phi_c = u$ with $L_c = 0$ and $\sigma_c^0 = 0$. We thus wish to solve the equations $\sigma_c^0 = 0 = \eth L + L L_{,u}$ and $0 = \eth\phi + L\phi_{,u}$ with L and ϕ close to their "canonical" forms, i.e., with $L = L_c + \epsilon L_1 = \epsilon L_1$ and $\phi = \phi_c + \epsilon\psi_1 = u + \epsilon\psi_1$. It is clear that, to first order, ψ_1 satisfies

$$\eth^2\psi_1 \equiv 4\frac{\partial}{\partial\zeta}P^2\frac{\partial\psi_1}{\partial\zeta} = 0 \tag{5.1}$$

and has solutions of the form

$$\psi_1 = y(u,\bar{\zeta})F(u,\zeta,\bar{\zeta}) + z(u,\bar{\zeta}), \tag{5.2}$$

with $F = \int_\infty^\zeta d\zeta/P^2$ and y and z arbitrary functions of $u,\bar{\zeta}$. Note also that ψ_1 is the infinitesimal transformation, $u' = u + \epsilon\psi_1$ leading from one "canonical" form to another. It is thus from two separate solutions of (5.1) that the commutator will be constructed.

Labeling the two solutions $\psi_1(u,\zeta,\bar{\zeta})$ and $\psi_2(u,\zeta,\bar{\zeta})$, the commutator is the difference $u'' - u^{**} = \epsilon^2\psi_c$, where

$$\begin{aligned} u' &= u + \epsilon\psi_1(u,\zeta,\bar{\zeta}),\\ u'' &= u' + \epsilon\psi_2'(u',\zeta,\bar{\zeta}) \\ &= u + \epsilon\psi_1(u,\zeta,\bar{\zeta}) + \epsilon\psi_2'(u+\epsilon\psi_1,\zeta,\bar{\zeta}), \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} u^* &= u + \epsilon\psi_2(u,\zeta,\bar{\zeta}),\\ u^{**} &= u^* + \epsilon\psi'_1(u^*,\zeta,\bar{\zeta}) = u + \epsilon\psi_2(u,\zeta,\bar{\zeta}) \\ &\quad + \epsilon\psi'_1(u+\epsilon\psi_2,\zeta,\bar{\zeta}). \end{aligned} \tag{5.4}$$

Before the commutator can be evaluated, using the properties of (5.2), a subtlety must first be discussed. We began by considering two transformations, ψ_1 and ψ_2 which change the coordinate u. The question then arises as to what we mean by the *same* transformation acting on the coordinate u' or u^*. [This ambiguity was symbolized by the use of a prime on ψ in (5.3) and (5.4).] We will avoid the issue for the time being and just write that when $\psi_2(u,\zeta,\bar{\zeta}) = y_2(u,\zeta,\bar{\zeta})F + z_2(u,\zeta,\bar{\zeta})$, then

$$\psi_2'(u',\zeta,\bar{\zeta}) = [y_2(u',\bar{\zeta}) + \epsilon\delta_1 y_2(u',\bar{\zeta})]F' + [z_2(u',\bar{\zeta}) + \epsilon\delta_1 z_2(u',\bar{\zeta})], \tag{5.5}$$

where

$$F' = \int_\infty^\zeta \frac{d\zeta}{P'^2}, \quad P' = (1 - \epsilon\dot{\psi}_1)P$$

and the $\delta_1 y_2$ and $\delta_1 z_2$ are for the moment undefined changes in y_2 and z_2 brought about by the transformation $u' = u + \epsilon\psi_1$. Exactly the same thing would be done to ψ'_1 with the 1 and 2 interchanged.

By substituting $u' = u + \epsilon\psi_1$ into (5.5) and writing

$$F' = F + \epsilon\delta_1 F, \delta_1 F = 2\int_\infty^\zeta \frac{d\zeta\dot{\psi}_1}{P^2},$$

we have

$$\psi_2'(u',\zeta,\bar{\zeta}) = \psi_2(u,\zeta,\bar{\zeta}) + \epsilon[(\dot{y}_2\psi_1 + \delta_1 y_2)F + y_2\delta_1 F + \dot{z}_2\psi_1 + \delta_1 z_2].$$

Thus, from (5.3) and (5.4) the commutator ψ_c becomes

$$\begin{aligned} \epsilon^{-2}(u'' - u^{**}) = \psi_c &= (\dot{y}_2\psi_1 + \delta_1 y_2)F + y_2\delta_1 F \\ &+ \dot{z}_2\psi_1 + \delta_1 z_2 - \{(\dot{y}_1\psi_2 + \delta_2 y_1)F + y_1\delta_2 F \\ &+ \dot{z}_1\psi_2 + \delta_2 z_1\} = [\dot{y}_2\psi_1 - \dot{y}_1\psi_2 + \delta_1 y_2 \\ &- \delta_2 y_1]F + [\dot{z}_2\psi_1 - \dot{z}_1\psi_2 + \delta_1 z_2 - \delta_2 z_1] \\ &+ y_2\delta_1 F - y_1\delta_2 F. \end{aligned}$$

Using $\psi_1 = y_1 F + z_1$, $\psi_2 = y_2 F + z_2$, we have after rearrangement

$$\begin{aligned} \psi_c = (\dot{y}_2 y_1 - \dot{y}_1 y_2)F^2 &+ [(z_1\dot{y}_2 - z_2\dot{y}_1) + (\dot{z}_2 y_1 \\ &- \dot{z}_1 y_2) + (\delta_1 y_2 - \delta_2 y_1)]F + (\dot{z}_2 z_1 - \dot{z}_1 z_2) \\ &+ (\delta_1 z_2 - \delta_2 z_1) + y_2\delta_1 F - y_1\delta_2 F. \end{aligned}$$

The last two terms can be rewritten as

$$\begin{aligned} y_2\delta_1 F - y_1\delta_2 F &= 2\Big((y_2\dot{y}_1 - y_1\dot{y}_2)\int_\infty^\zeta \frac{d\zeta}{P^2}F \\ &+ (y_2 y_1 - y_1 y_2)\int_\infty^\zeta \frac{d\zeta}{P^2}\dot{F} \cdot \\ &+ (y_2\dot{z}_1 - y_1\dot{z}_2)\int_\infty^\zeta \frac{d\zeta}{P^2}\Big) \\ &= (y_2\dot{y}_1 - y_1\dot{y}_2)F^2 + 2(y_2\dot{z}_1 - y_1\dot{z}_2)F, \end{aligned}$$

where the first term has been simplified by integration by parts. We thus obtain

$$\begin{aligned} \psi_c = [(y_2\dot{z}_1 - y_1\dot{z}_2) &- (z_2\dot{y}_1 - z_1\dot{y}_2) \\ &+ (\delta_1 y_2 - \delta_2 y_1)]F \\ &+ (\dot{z}_2 z_1 - \dot{z}_1 z_2) \\ &+ (\delta_1 z_2 - \delta_2 z_1), \end{aligned} \tag{5.6}$$

which is of the form (5.2), from which it follows that the coordinate transformations form a group. Note also that the commutator is independent of the form of F (or P). We will exploit this by showing that for a special P the commutator vanishes and thus we should in general choose the δy and δz to make (5.6) vanish. This is accomplished by

$$\delta_1 y_2 = -y_2\dot{z}_1 + z_2\dot{y}_1 + S_y, \quad \delta_1 z_2 = z_2\dot{z}_1 + S_z,$$

where S_y and S_z are symmetric in 1 and 2.

For the special P we choose

$$P = vP_0 = \tfrac{1}{2}\, v(1 + \zeta, \bar{\zeta}), \tag{5.7}$$

where $v(u, \zeta, \bar{\zeta}) = \sum_{l=0}^{1} A_{lm}(u) Y_{lm}(\zeta, \bar{\zeta})$ such that it has no zeros. Then Eq. (5.1) takes the form

$$\eth^2 \psi = \eth_0(v^2 \eth_0 \psi) = 0, \tag{5.8}$$

with the general regular solution being

$$\psi(u, \zeta, \bar{\zeta}) = v^{-1} \sum_{l=0}^{1} b_{lm}(u) Y_{lm}(\zeta, \bar{\zeta}). \tag{5.9}$$

If we now say, by *definition*, that the *same* transformation starting from a primed system, i.e., $P' = \frac{1}{2}\, v'(1 + \zeta\bar{\zeta})$ is given by

$$\psi'(u', \zeta, \bar{\zeta}) = v'^{-1} \sum_{l=0}^{1} b_{lm}(u') Y_{lm}(\zeta, \bar{\zeta}),$$

it is *not* difficult to show that the commutator of two transformations (5.9) is zero. [Use, of course, must be made of the transformation properties of P, i.e.,

$$P' = (1 - \epsilon\dot{\psi})P, \quad \cdot\ v' = (1 - \epsilon\dot{\psi})v,$$

which guarantees that $v' = \sum_{l=0}^{1} A'_{lm}(u') Y_{lm}$.]

If we now return to the real canonical slicing by the inverse of the complex coordinate transformation $u' = G$, the commutator remains zero and thus our original contention is proven.

As an inessential point, we wish to remark that the transformations obtained from (5.9) have a very simple geometric meaning in Minkowski space. If one takes an arbitrary timelike world line, the light cone from every point on the line gives a slicing of J^+. A neighboring world line yields a neighboring slicing with the transformation connecting the slicings being given by $u' = u + \epsilon\psi$. The four $b_{lm}(u)$ are related to the arbitrary time-dependent 4-vector connecting the two world lines.

6. GEOMETRIC SIGNIFICANCE OF L

In Sec. 2, we introduced on null infinity two-dimensional cuts and generators g with tangent vectors n^μ. By considering the second set of rays which are orthogonal to the cuts (with tangent vectors l^μ), we were able to construct a family of null hypersurfaces in a four-dimensional neighborhood of null infinity. In general the tangent vectors (or gradients of the null surfaces) have an asymptotic shear σ^0, arising from

$$\sigma = \frac{\sigma^0}{r^2} + O(r^{-3}). \tag{6.1}$$

The question can be raised—is it possible to introduce other null rays striking the cuts, but now *not* orthogonally, such that their asymptotic shear vanishes. The new tangent vectors denoted by $l^{*\mu}$, can be obtained from null rotations around $\tilde{n}^\mu$, namely by

$$l^{*\mu} = l^\mu + A\bar{m}^\mu + \bar{A}m^\mu + A\bar{A}n^\mu, \tag{6.2}$$

where m^μ and $\bar{m}^\mu$ are the (complex) tangent vectors to the cut, n^μ is the inward pointing null vector orthogonal to the cut, and the A and $\bar{A}$ are functions on the neighborhood of J^+, that are to go to zero on J^+. (All rays striking a point of J^+ are parallel to l^μ at J^+.) Thus we may write (6.2) as

$$l^{*\mu} = l^\mu + \frac{L}{r}\,\bar{m}^\mu + \frac{\bar{L}}{r}\, m^\mu + O(r^{-2}), \tag{6.3}$$

with $L(u, \zeta, \bar{\zeta})$ still arbitrary. Note that the meaning of L is effectively the scalar product of $l^{*\mu}$ with the tangent vector $\bar{m}^\mu$ or

$$L = \lim_{r\to\infty}(r l^{*\mu} \bar{m}_\mu). \tag{6.4}$$

Computing the shear σ^* of $l^{*\mu}$, we obtain

$$\sigma^* = \frac{\sigma^{0*}}{r^2} + O(r^{-3}) \tag{6.5}$$

with

$$\sigma^{0*} = \sigma^0 - \eth L - L\dot{L}. \tag{6.6}$$

If we thus demand the vanishing of the asymptotic shear σ^{0*}, we recover (3.1), the starting point of our investigation.

7. DISCUSSION

We have shown in this paper how to introduce coordinate conditions in asymptotically flat space which are different from those of Bondi. These conditions lead to a coordinate group not only of greater simplicity than the Bondi–Metzner–Sachs group, but one which also has a simple straightforward geometric interpretation in Minkowski space.

Concomitant with these results was the discovery of the differential equation (3.1), leading to asymptotically shear-free null congruences. Already this equation has been exploited by Penrose[17] to define twistors in asymptotically flat space. In a future paper we will show how to introduce a coordinate system associated with these twisting shear-free rays and the transformation to the hypersurface orthogonal ray coordinate system. This will be used to analyze the physical meaning of the geometric quantities which appear in, for example, the algebraically special twisting metrics.

One of the major applications we envisage for the work presented here, is to use up the remaining coordinate freedom, by, in the sense to be explained, transforming to the center of mass frame. Physical quantities such as energy or linear momentum can be defined[9] at infinity by integrals over cuts. If the center of mass can be so defined relative to one of our canonical cuts, then it is hoped that a particular canonical cut will yield a zero center of mass and thereby yield a unique coordinate system associated with a particular solution. As a reward we would presumably obtain

the equations of motion of this center of mass and the "internal" angular momentum, i.e., angular momentum relative to the center of mass. Preliminary investigations bear this out.

ACKNOWLEDGMENTS

We would like to than our colleagues E. Couch, J. Goldberg, A. Janis, R. Penrose, and J. Winicour for many useful and stimulating discussions.

* This research has been supported by the NSF, Grant No. GP-22789.

[1] H. Bondi, M. van der Burg, and A. Metzner, Proc. Roy. Soc. (London) A269, 21 (1962).
[2] R. Sachs, Proc. Roy. Soc. (London) A270, 103 (1962).
[3] R. Sachs, Phys. Rev. 128, 2851 (1962).
[4] R. Geroch and E. T. Newman, J. Math. Phys. 12, 314 (1971).
[5] R. Penrose, Phys. Rev. Letters 10, 66 (1963).
[6] R. Penrose, in *Relativity, Groups, and Topology*, edited by C. Dewitt and B. Dewitt (Gordon and Breach, New York, 1964).
[7] V. Cantoni, J. Math. Phys. 8, 1700 (1967).
[8] R. Sachs, in *Relativity, Groups and Topology*, edited by C. Dewitt and B. Dewitt (Gordon and Breach, New York, 1964).
[9] L. Tamburino and J. Winicour, Phys. Rev. 150, 1039 (1966).
[10] R. Sachs (private communication).
[11] E. T. Newman and R. Penrose, J. Math. Phys. 7, 863 (1966).
[12] E. T. Newman and T. Unti, J. Math. Phys. 3, 891 (1962)
[13] For properties of the operator ð see J. Goldberg, A. MacFarlane, E. T. Newman, F. Rohlich, and G. Sudarshan, J. Math. Phys. 8, 2155 (1967).
[14] A. Held, E. T. Newman, and R. Posadas, J. Math. Phys. 11, 3145 (1970).
[15] A. Exton, E. T. Newman, and R. Penrose, J. Math. Phys. 10, 1570 (1969).
[16] A. Janis and E. T. Newman, J. Math. Phys. 6, 915 (1965).
[17] R. Penrose (private communication).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Many-Body Point Transforms. II. An Exact Noncluster Approach to the Hard-Core Many-Body Problem

Norman M. Witriol
Solid State Physics Branch, Physical Sciences Laboratory, Redstone Arsenal, Alabama 35809

The method of many-body point transforms previously developed by the author is modified to eliminate the interacting particle cluster formalism. The cutoff terms in the transformation, and thus in the Hamiltonian, are therefore removed. For the hard-core many-body problem, as in the previous work, a Hamiltonian is obtained which is Fourier analyzable, Hermitian and amenable to ordinary perturbation and variational techniques while still being equivalent to the original Hamiltonian. No approximations are made. It is demonstrated that the use of the standard zeroth-order approximation of the ground state Bose system, i.e., a constant, for the transformed wavefunction gives a negative expectation value of the energy for the hard core system. This discrepancy arises from the extended range of the new potential generated by the transformation, which necessitates explicit consideration of boundary conditions not satisfied by the above wavefunction.

1. INTRODUCTION

In a previous paper,[1] referred to herein as I, we developed an interacting particle cluster method for transforming the many-body hard-core Hamiltonian into a Fourier analyzable, Hermitian Hamiltonian which was completely equivalent to the original Hamiltonian while being amenable to ordinary perturbation and variational techniques. The method was based on the point transformation method developed by Eger and Gross.[2-4] Using this method with the assumption of pairwise additivity, the ground-state energy and the low-level excitations of the hard-core Bose liquid have been calculated[3] and found to agree in the dilute gas limit with the low-density results of Lee, Huang, and Yang.[5] Similarly, the ground-state energy and Landau parameters of the hard-core Fermi liquid have been calculated[6] and found to agree in the dilute gas limit with the series expansion in $c\rho^{1/3}$ (c being the hard-core radius, ρ being the density of the system) of Huang and Yang[7] and Lee and Yang[8] for the ground-state energy and with the series expansions in $c\rho^{1/3}$ of Abrikosov and Khalatnikov[9] for the Landau parameters. In addition, Cooper[10] has used this method to calculate the virial coefficients for the classical hard-core gas. In I the point transformation method was extended to the consideration of the many-body terms of many-body systems. The point transformation was taken to be a continuous cluster expansion (one body isolated, two bodies interacting, three bodies interacting, ..., N bodies interacting) in the $3N$ dimensional transformed space. The transformed Hamiltonian was found to be Fourier analyzable and Hermitian, while having the same continuous cluster expansion form as the transformation. For large clusters this expansion technique gave rise to mathematical computations of prohibitive complexity, resulting from the mixing of the individual cluster transformations, caused by the continuity requirement. In this article we shall derive results amenable to the calculation of many-body effects. We note that the mathematical computations in this paper, while being similar to, are completely independent from those of I. In Sec. 2 we shall review the point transformation method, and the results of I, for use in the derivation of and comparison with the results obtained in this article. In Sec. 3 we choose general forms for the N-body transformation, and derive the transformed Hamiltonians. In Sec. 4 we apply this method to the N-body hard-core problem. Choosing a specific transformation which removes the hard-core potential, we calculate the transformed Hamiltonian.

the equations of motion of this center of mass and the "internal" angular momentum, i.e., angular momentum relative to the center of mass. Preliminary investigations bear this out.

ACKNOWLEDGMENTS

We would like to than our colleagues E. Couch, J. Goldberg, A. Janis, R. Penrose, and J. Winicour for many useful and stimulating discussions.

* This research has been supported by the NSF, Grant No. GP-22789.

[1] H. Bondi, M. van der Burg, and A. Metzner, Proc. Roy. Soc. (London) A269, 21 (1962).
[2] R. Sachs, Proc. Roy. Soc. (London) A270, 103 (1962).
[3] R. Sachs, Phys. Rev. 128, 2851 (1962).
[4] R. Geroch and E. T. Newman, J. Math. Phys. 12, 314 (1971).
[5] R. Penrose, Phys. Rev. Letters 10, 66 (1963).
[6] R. Penrose, in *Relativity, Groups, and Topology*, edited by C. Dewitt and B. Dewitt (Gordon and Breach, New York, 1964).
[7] V. Cantoni, J. Math. Phys. 8, 1700 (1967).
[8] R. Sachs, in *Relativity, Groups and Topology*, edited by C. Dewitt and B. Dewitt (Gordon and Breach, New York, 1964).
[9] L. Tamburino and J. Winicour, Phys. Rev. 150, 1039 (1966).
[10] R. Sachs (private communication).
[11] E. T. Newman and R. Penrose, J. Math. Phys. 7, 863 (1966).
[12] E. T. Newman and T. Unti, J. Math. Phys. 3, 891 (1962)
[13] For properties of the operator ð see J. Goldberg, A. MacFarlane, E. T. Newman, F. Rohlich, and G. Sudarshan, J. Math. Phys. 8, 2155 (1967).
[14] A. Held, E. T. Newman, and R. Posadas, J. Math. Phys. 11, 3145 (1970).
[15] A. Exton, E. T. Newman, and R. Penrose, J. Math. Phys. 10, 1570 (1969).
[16] A. Janis and E. T. Newman, J. Math. Phys. 6, 915 (1965).
[17] R. Penrose (private communication).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Many-Body Point Transforms. II. An Exact Noncluster Approach to the Hard-Core Many-Body Problem

Norman M. Witriol
Solid State Physics Branch, Physical Sciences Laboratory, Redstone Arsenal, Alabama 35809

The method of many-body point transforms previously developed by the author is modified to eliminate the interacting particle cluster formalism. The cutoff terms in the transformation, and thus in the Hamiltonian, are therefore removed. For the hard-core many-body problem, as in the previous work, a Hamiltonian is obtained which is Fourier analyzable, Hermitian and amenable to ordinary perturbation and variational techniques while still being equivalent to the original Hamiltonian. No approximations are made. It is demonstrated that the use of the standard zeroth-order approximation of the ground state Bose system, i.e., a constant, for the transformed wavefunction gives a negative expectation value of the energy for the hard core system. This discrepancy arises from the extended range of the new potential generated by the transformation, which necessitates explicit consideration of boundary conditions not satisfied by the above wavefunction.

1. INTRODUCTION

In a previous paper,[1] referred to herein as I, we developed an interacting particle cluster method for transforming the many-body hard-core Hamiltonian into a Fourier analyzable, Hermitian Hamiltonian which was completely equivalent to the original Hamiltonian while being amenable to ordinary perturbation and variational techniques. The method was based on the point transformation method developed by Eger and Gross.[2-4] Using this method with the assumption of pairwise additivity, the ground-state energy and the low-level excitations of the hard-core Bose liquid have been calculated[3] and found to agree in the dilute gas limit with the low-density results of Lee, Huang, and Yang.[5] Similarly, the ground-state energy and Landau parameters of the hard-core Fermi liquid have been calculated[6] and found to agree in the dilute gas limit with the series expansion in $c\rho^{1/3}$ (c being the hard-core radius, ρ being the density of the system) of Huang and Yang[7] and Lee and Yang[8] for the ground-state energy and with the series expansions in $c\rho^{1/3}$ of Abrikosov and Khalatnikov[9] for the Landau parameters. In addition, Cooper[10] has used this method to calculate the virial coefficients for the classical hard-core gas. In I the point transformation method was extended to the consideration of the many-body terms of many-body systems. The point transformation was taken to be a continuous cluster expansion (one body isolated, two bodies interacting, three bodies interacting, ..., N bodies interacting) in the $3N$ dimensional transformed space. The transformed Hamiltonian was found to be Fourier analyzable and Hermitian, while having the same continuous cluster expansion form as the transformation. For large clusters this expansion technique gave rise to mathematical computations of prohibitive complexity, resulting from the mixing of the individual cluster transformations, caused by the continuity requirement. In this article we shall derive results amenable to the calculation of many-body effects. We note that the mathematical computations in this paper, while being similar to, are completely independent from those of I. In Sec. 2 we shall review the point transformation method, and the results of I, for use in the derivation of and comparison with the results obtained in this article. In Sec. 3 we choose general forms for the N-body transformation, and derive the transformed Hamiltonians. In Sec. 4 we apply this method to the N-body hard-core problem. Choosing a specific transformation which removes the hard-core potential, we calculate the transformed Hamiltonian.

This Hamiltonian is Fourier analyzable, Hermitian and equivalent to the original Hamiltonian, while being amenable to ordinary perturbation and variational techniques. In Sec. 5 we consider the application of this transformed Hamiltonian to the calculation of the ground-state energy. We first show that we obtain the standard results for the two-body isolated hard-core system. We then consider the calculation of the ground-state energy of the N-body hard-core Bose system. Using the variational principle and choosing the trial transformed wavefunction to be a constant, we obtain a negative result. This discrepancy is shown to be due to the neglect of boundary conditions. Sec. 6 contains our conclusions.

2. RELEVANCE TO PREVIOUS WORK

In I the point transformation method[2-4] was extended to the consideration of the many-body terms of many-body systems. The point transformation was taken to be regional (one body isolated, two bodies interacting, three bodies interacting, ..., N bodies interacting) and continuous in the $3N$-dimensional transformed space. Notationally, Latin indices represent the particle indices 1 to N; Greek indices represent the Cartesian coordinates 1 to 3; the $X_{i\alpha}$ (the $\mathbf{R}_i$) are the original coordinates; the $x_{i\alpha}$ (the $\mathbf{r}_i$) are the transformed coordinates; the $P_{i\alpha}$ are the original momenta; the $p_{i\alpha}$ are the transformed momenta; $X_{ij\alpha} = X_{i\alpha} - X_{j\alpha}$; $x_{ij\alpha} = x_{i\alpha} - x_{j\alpha}$; $r_{ij} = |\mathbf{r}_{ij}|$; and $\mathbf{R}, \mathbf{P}, \mathbf{r}, \mathbf{p}$ represent their respective set of $3N$ variables, $\mathbf{R}_i, \mathbf{P}_i, \mathbf{r}_i$ and $\mathbf{p}_i$. The regional many-body point transformation was thereby written in the form

$$
\begin{aligned}
X_{i\alpha} = &\left(\prod_{\substack{a=1\\a\neq i}}^{N}\theta(r_{ia}-b)\right)(x_{i\alpha})\\
&+\sum_{\substack{j=1\\j\neq i}}^{N}\left(\prod_{\substack{a=1\\a\neq i,j}}^{N}\theta(r_{ia}-b)\theta(r_{ja}-b)\right)[\theta(b-r_{ij})]\\
&\times[A_{ija}(1-h_{ij})+A_{i\alpha}h_{ij}]\\
&+\sum_{\substack{j,k=1\\k<j\\k,j\neq i}}^{N}\left(\prod_{\substack{a=1\\a\neq i,j,k}}^{N}\theta(r_{ia}-b)\theta(r_{ja}-b)\theta(r_{ka}-b)\right)\\
&\times[\theta(b-r_{ik})\theta(b-r_{jk})\theta(r_{ij}-b)+\theta(b-r_{jk})\\
&\times\theta(b-r_{ij})\theta(r_{ik}-b)+\theta(b-r_{ij})\theta(b-r_{ik})\\
&\times\theta(r_{jk}-b)+\theta(b-r_{ij})\\
&\times\theta(b-r_{ik})\theta(b-r_{jk})][A_{ijk\alpha}(1-h_{ik}h_{jk}(1-h_{ij})\\
&-h_{ij}h_{kj}(1-h_{ik})-h_{ik}h_{ij})+A_{ij\alpha}h_{ik}h_{jk}(1-h_{ij})\\
&+A_{ik\alpha}h_{ij}h_{kj}(1-h_{ik})+A_{i\alpha}h_{ik}h_{ij}]+\cdots,
\end{aligned}
\tag{2.1}
$$

$$
P_{i\alpha}=\tfrac{1}{2}\sum_{j=1}^{N}\sum_{\beta=1}^{3}\left(\frac{\partial x_{j\beta}}{\partial X_{i\alpha}}p_{j\beta}+p_{j\beta}\frac{\partial x_{j\beta}}{\partial X_{i\alpha}}\right), \tag{2.2}
$$

where the θ functions, specified by

$$
\begin{aligned}
\theta(r_{ia}-b)=0, &\quad \theta(b-r_{ia})=1, &&r_{ia}\le b\\
=1, &\quad =0, &&r_{ia}>b,
\end{aligned}
\tag{2.3}
$$

define the regions for the cluster formalism, b being the arbitrary range of the interaction in both the transformed and original spaces: the $A_{i_1\cdots i_n\alpha}$ are the desired cluster $(i_1\cdots i_n)$ transformations of $X_{i_1\alpha}$ and the h_{ij} are arbitrary functions subject to the conditions

$$
\begin{aligned}
&h_{ij}=0, \quad r_{ij}=0,\\
&\quad =1, \quad r_{ij}=1,\\
&dh_{ij}/dr_{ij}=d^2h_{ij}/dr_{ij}^2=0, \quad r_{ij}=b,
\end{aligned}
\tag{2.4}
$$

and required for the continuity of the wavefunction; e.g., for the first two cluster terms in Eq. (2.1), as $r_{ij}\to b$ the (ij) cluster transformation of $X_{i\alpha}, X_{i\alpha}=A_{ij\alpha}(1-h_{ij})+x_{i\alpha}h_{ij}$, must approach the isolated one-body identify transformation $X_{i\alpha}=x_{i\alpha}$. Under the point-transformation method the original Hamiltonian for equal mass particles of mass m,

$$
\tilde{H}(\mathbf{R},\mathbf{P})=\frac{1}{2m}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}P_{i\alpha}^2+V(\mathbf{R},\mathbf{P}) \tag{2.5}
$$

is transformed into itself expressed in terms of the transformed coordinates,[2]

$$
H(\mathbf{r},\mathbf{p})=\tilde{H}(\mathbf{R}(\mathbf{r}),\mathbf{P}(\mathbf{r},\mathbf{p})) \tag{2.6}
$$

$$
=\frac{1}{2m}\sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}(p_{i\alpha}g_{i\alpha j\beta}p_{j\beta})+V(\mathbf{r},\mathbf{p})+W(\mathbf{r}), \tag{2.7}
$$

where

$$
g_{i\alpha j\beta}=\sum_{l=1}^{N}\sum_{\tau=1}^{3}\frac{\partial x_{i\alpha}}{\partial X_{l\tau}}\frac{\partial x_{j\beta}}{\partial X_{l\tau}}, \tag{2.8}
$$

$$
W(r)=\frac{\hbar^2}{2m}\sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\sqrt{B}\,\frac{\partial}{\partial x_{i\alpha}}\left[g_{i\alpha j\beta}\frac{\partial}{\partial x_{j\beta}}\left(\frac{1}{\sqrt{B}}\right)\right], \tag{2.9}
$$

B is the Jacobian of the inverse transformation $(\mathbf{r}\to\mathbf{R})$, and $V(\mathbf{r},\mathbf{p})$ is $V(\mathbf{R},\mathbf{P})$ written in terms of $\mathbf{r},\mathbf{p}$. In addition, the original normalized wavefunction $\tilde{\psi}(\mathbf{R})$ is related to the transformed normalized wavefunction $\psi(\mathbf{r})$ by

$$
\tilde{\psi}(\mathbf{R}(\mathbf{r}))=\sqrt{B(\mathbf{r})}\,\psi(\mathbf{r}). \tag{2.10}
$$

For a regional point transformation of the form Eq. (2.1), we found[1]

$$
\partial x_{k\lambda}/\partial X_{l\tau}=\theta((k))\delta_{kl}\delta_{\lambda\tau}+\sum_{\substack{(.k.l.)\\(.k.l.)\neq(k)}}\theta((.k.l.))\left(\partial x_{k\lambda}/\partial X_{l\tau}\right)_{(.k.l.)}, \tag{2.11}
$$

$$
g_{i\alpha j\beta}=\theta((i))\delta_{ij}\delta_{\alpha\beta}+\sum_{\substack{(.i.j.)\\(.i.j.)\neq(i)}}\theta((.i.j.))\,g_{i\alpha j\beta(.i.j.)}, \tag{2.12}
$$

$$
B=\sum_{\{n(N)\}}\theta(\{n(N)\})\,\Pi_i B_{n_i}, \tag{2.13}
$$

$$W(\mathbf{r}) = \sum_{n=1}^{N} \sum_{(i_1 \cdots i_n)} \theta((i_1 \cdots i_n))\, W_{(i_1 \cdots i_n)} \quad \text{with} \quad W_{(i_1)} = 0, \tag{2.14}$$

and

$$H = \Bigg\{ \sum_{n=1}^{N} \sum_{(i_1 \cdots i_n)} \theta((i_1 \cdots i_n)) \times \Bigg[\Bigg(\sum_{a,b=1}^{N} \sum_{\alpha,\beta=1}^{3} p_{i_a \alpha}\, g_{i_a \alpha i_b \beta (i_1 \cdots i_n)}\, p_{i_b \beta} \Bigg) + W_{(i_1 \cdots i_n)} \Bigg] \Bigg\} + V(\mathbf{r}, \mathbf{p}), \tag{2.15}$$

where $\sum_{(.k.l.)}$ is the sum over all clusters including the particles k and l, $\theta((.k.l.))$ represents the product (and sums) of the θ functions defining the $(.k.l.)$ cluster, $(\partial x_{k\lambda}/\partial X_{l\tau})_{(.k.l.)}$ is $\partial x_{k\lambda}/\partial X_{l\tau}$ for the isolated $(.k.l.)$ cluster, $g_{i\alpha j\beta\,(i.j.)}$ is $g_{i\alpha j\beta}$ for the isolated $(.i.j.)$ cluster, $\{n(N)\}$ represents a specific combination of clusters of the N particles, $\sum_{\{n(N)\}}$ is the sum over all possible combinations of clusters, $B_{(n)_i}$ is the Jacobian of the inverse transformation $(\mathbf{r} \rightarrow \mathbf{R})$ for the isolated $(n)_i$ cluster, and $W_{(i_1 \cdots i_n)}$ is the new coordinate only dependent potential W for the isolated $(i_1 \cdots i_n)$ cluster. If the transformed potential has a range less than b, then Eq. (2.15) can be written in the form

$$H = \sum_{n=1}^{N} \sum_{(i_1 \cdots i_n)} \theta((i_1 \cdots i_n)) \times \Bigg[\Bigg(\frac{1}{2m} \sum_{a,b=1}^{n} \sum_{\alpha,\beta=1}^{3} p_{i_a \alpha} g_{i_a \alpha i_b \beta (i_1 \cdots i_n)} p_{i_b \beta} \Bigg) + W_{(i_1 \cdots i_n)} + V_{(i_1 \cdots i_n)} \Bigg] \tag{2.16}$$

$$= \sum_{n=1}^{N} \sum_{(i_1 \cdots i_n)} \theta((i_1 \cdots i_n))\, H_{(i_1 \cdots i_n)}, \tag{2.17}$$

where $V_{(i_1 \cdots i_n)}$ represents the transformed potentials acting between the particles $i_1 \cdots i_n$ in the isolated $(i_1 \cdots i_n)$ cluster, and $H_{(i_1 \cdots i_n)}$ is the transformed n-body Hamiltonian of the isolated $(i_1 \cdots i_n)$ cluster.

In I we considered the case of the original potential $V(\mathbf{R}, \mathbf{P})$ being a short-range, strongly repulsive or hard-core potential,

$$V(\mathbf{R}, \mathbf{P}) = V(R) = V_0, \quad R \le c, \qquad = 0, \quad R > c. \tag{2.18}$$

Now the effects of a potential are proportional to both its strength and its range. We thereby showed that we could reduce the effects of this potential by reducing its range. Diagrammatically the desired effect of the transformation is shown in Fig. 1. To achieve this goal the desired cluster transformations $A_{i_1 \ldots i_n \alpha}$, were defined in a manner such that each n-particle cluster $(i_1 \ldots i_n)$ shrunk about its center of mass $\mathbf{R}_{cm(i_1 \ldots i_n)}$, the shrinkage factor depending on the smallest interparticle distance. This shrinkage factor was so chosen that when the smallest interparticle separation R_{ij} was equal to the strong interaction radius c, the smallest transformed interparticle separation r_{ij} was equal to the transformed strong interaction radius ϵ, where ϵ is an arbitrarily chosen small distance. This concept is illustrated in Fig. 2 for the three-body cluster. Mathematically the desired cluster transformations $A_{i_1 \ldots i_n \alpha}$ have the form[11]

$$A_{i\alpha} = x_{i\alpha}, \tag{2.19}$$

$$A_{i_1 \cdots i_n \alpha} = u_{i_1 \cdots i_n}\, x_{i_1 \alpha} + (1 - u_{i_1 \cdots i_n}) \times \frac{(x_{i_1 \alpha} + \cdots + x_{i_n \alpha})}{n}, \tag{2.20}$$

where

$$u_{i_1 \cdots i_n} = \frac{1}{\tau_n} \sum_{\substack{c,d=1 \\ c<d}}^{n} \prod_{\substack{a,b=1 \\ a<b \\ i_a i_b \neq i_c i_d}}^{n} \theta(r_{i_a i_b} - r_{i_c i_d})\, e(\mathbf{r}_{i_c i_d}), \tag{2.21}$$

in which τ_n is the normalization function

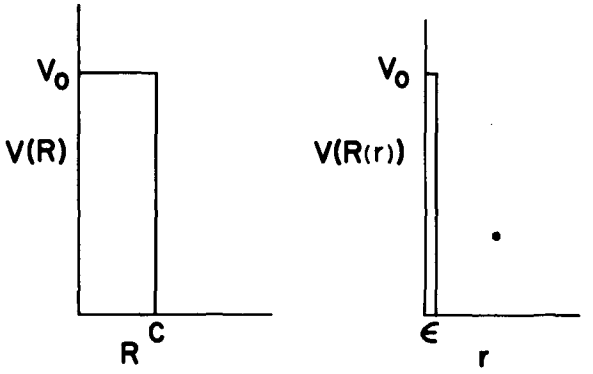

FIG. 1. Reduction (by the transformation) of the range of the strongly repulsive potential V_0.

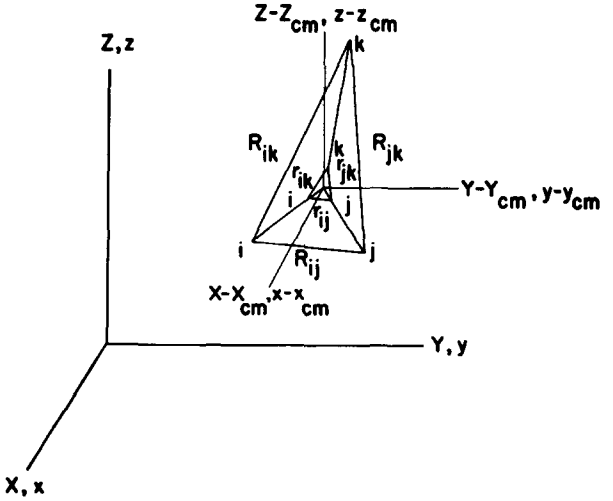

FIG. 2. Three-body (ijk) cluster transformation; cm = center of mass of (ijk) cluster.

$$\tau_n = \sum_{\substack{c,d=1 \\ c<d}}^{n} \prod_{\substack{a,b=1 \\ i<b \\ i_a i_b \neq i_c i_d}}^{n} \theta(r_{i_a i_b} - r_{i_c i_d}), \qquad (2.22)$$

and $i_a i_b \neq i_c i_d$ means that $r_{i_a i_b}$ and $r_{i_c i_d}$ are not the same interparticle separation. The product over the θ is zero unless $r_{i_c i_d}$ is less than or equal to each one of the other interparticle distances $r_{i_a i_b}$ in the n-body cluster $(i_1 \cdots i_n)$. The function $e(\mathbf{r}_{i_c i_d})$ is the actual shrinkage factor for $c \to \epsilon$ in each cluster transformation, and is taken to be of the same form for all $i_c i_d$. One possible expression given for $e(\mathbf{r}_{i_c i_d})$ was the pairwise transformation form

$$e(\mathbf{r}_{i_c i_d}) = 1 + \frac{c}{(r_{i_c i_d} + \epsilon'')}, \qquad (2.23)$$

where ϵ'' is a small number related to ϵ. For hard cores we showed that we can let ϵ and thus $\epsilon'' \to 0$, thereby completely removing the region $0 \leq R < c$, and thus the hard-core potential, Eq. (2.18) with $V_0 = \infty$. To guarantee the one to oneness of the transformation, we required $b > c$. Therefore if our original potential were only the hard-core potential, we could obtain a transformed Hamiltonian of the form Eq. (2.17), which is Hermitian, Fourier analyzable and amenable to ordinary perturbation and variational techniques, while still being equivalent to the original Hamiltonian. The new momentum-dependent potentials $\vec{p} \cdot (\overleftrightarrow{g} - \overleftrightarrow{\delta}) \cdot \vec{p}$ and the new coordinate only dependent potentials W have range b in the transformed space. Due to the regional nature of Eq. (2.16) [caused by the regional nature of the transformation, Eq. (2.1)] the transformed Hamiltonian H can easily be approximated by assuming n-body additivity $(1 \leq n \leq N)$, i.e., by assuming the probability for $n+1$ and more body clusters to be negligible where n is arbitrarily chosen as a compromise between accuracy and laborius mathematical computation. Now the choice of the functional form of h (and thus b) is relatively arbitrary. If we solved the problem exactly (i.e., kept all the terms), the ground state energy and the original wavefunction should be independent of this choice. However, when making approximations by specifying the maximum size of the clusters considered, we expect our results to be dependent upon the h. Our results are therefore accurate to the extent that they do not involve these parameters. As the size of the clusters considered increases, we expect the dependence of our results upon the choice of h to decrease.

In this article we shall remove this cluster formalism. Essentially our method will be equivalent to that of I with $b = \infty$, $h = 1$ for all r space; we shall therefore be dealing only with the N-body cluster. In I we derived a continuous cluster transformed Hamiltonian, Eq. (2.16). For large clusters, this expansion technique gave rise to prohibitively complex mathematical computations, resulting from the mixing of the desired cluster transformations $A_{i_1 \cdots i_n \alpha}$ caused by the h. We shall now use a noncluster transformation which will eliminate the hard-core potential while giving results amenable to the calculation of many-body effects.

3. GENERAL TRANSFORMATION

We start with a general form for the transformation with the center of mass taken as the reference point (i.e., the center of mass is not changed by the transformation),

$$X_{i\alpha} = u x_{i\alpha} + (1-u)\bar{x}_\alpha, \qquad (3.1)$$

where u is the transformation function and $\bar{x}_\alpha$ represents the center of mass coordinate of the N-body system,

$$\bar{x}_\alpha = (x_{i\alpha} + \cdots + x_{N\alpha})/N. \qquad (3.2)$$

The center of mass is chosen as the reference point since the purpose of this method is to simplify the interparticle potentials; we do not wish to alter the center of mass motion. To find the transformed Hamiltonian, we first need to calculate the derivatives $\partial x_{k\lambda}/\partial X_{l\tau}$, the effective metrics $g_{i\alpha j\beta}$, the Jacobian B, and the new coordinate dependent potential W. The transformed Hamiltonian is then obtained by substituting these expressions into Eq. (2.7). For strongly repulsive, or hard-core potentials, we shall, as in I, choose u to be of the form of Eq. (2.21) with Eq. (2.23). However, for the purpose of future generalizations, we shall proceed to this final form of u by stages. We start by taking u to be an arbitrary function of the $3N$ transformed coordinates $x_{k\lambda}$.

Consider the $\partial x_{k\lambda}/\partial X_{l\tau}$. These terms are found by taking the derivative with respect to $X_{l\tau}$ of Eq. (3.1) for all $i = 1 \cdots N$, $\alpha = 1.3$. Applying the chain rule, we obtain the resulting set of simultaneous linear equations in the $3N$ variables $\partial x_{k\lambda}/\partial X_{l\tau} (k = 1 \cdots N, \quad \lambda = 1.3; l\tau$ fixed);

$$\delta_{il}\delta_{\alpha\tau} = \sum_{k=1}^{N} \sum_{\lambda=1}^{3} \frac{\partial u}{\partial x_{k\lambda}} \frac{\partial x_{k\lambda}}{\partial X_{l\tau}} (x_{i\alpha} - \bar{x}_\alpha) + u \frac{\partial x_{i\alpha}}{\partial X_{l\tau}} + \frac{(1-u)}{N} \sum_{k=1}^{N} \frac{\partial x_{k\alpha}}{\partial X_{l\tau}}. \qquad (3.3)$$

Using Cramer's rule,[12] we find the set of equations for the derivatives $\partial x_{k\lambda}/\partial X_{l\tau}$ to be of the form

$$\frac{\partial x_{k\lambda}}{\partial X_{l\tau}} = \frac{E_{k\lambda,l\tau}}{D}, \qquad (3.4)$$

where D is the $3N$th order determinant of the coefficients of the $\partial x_{k\lambda}/\partial X_{l\tau}$ in Eq. (3.3) and $E_{k\lambda,l\tau}$ is D with the $k\lambda$th column replaced by the coefficients on the left-hand side of Eq. (3.3) the $\delta_{il}\delta_{\alpha\tau}$. We write D in the form

$$D = |D_{k\lambda,i\alpha}| = \left| C_{i\alpha}\left(\frac{\partial x_{k\lambda}}{\partial X_{l\tau}}\right) \right|, \qquad (3.5)$$

where dictionary ordering is used (for example, the ordering of the rows $k\lambda$ is $11, 12, 13, 21, \ldots, N3$) and $C_{i\alpha}(\partial x_{k\lambda}/\partial X_{l\tau})$ is the coefficient of $(\partial x_{k\lambda}/\partial X_{l\tau})$ in the $(i\alpha)$ equation of the set of equations given by Eq. (3.3). Notationally we set

$$\partial u/\partial x_{k\lambda} \equiv u'_{k\lambda}, \qquad (x_{i\alpha} - \bar{x}_\alpha) = \bar{x}_{i\alpha}, \qquad [(1-u)/N] = \bar{u}. \tag{3.6}$$

The $3N$th order determinant D is calculated in Appendix A. We find, Eq. (A24),

$$D = u^{3N-4}\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right). \tag{3.7}$$

The $3N$th order determinant $E_{k\lambda,l\tau}$ is calculated in Appendix B. We find Eq. (B8),

$$E_{k\lambda,l\tau} = u^{3N-5}\left[\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right)(\delta_{kl} - \bar{u})\delta_{\lambda\tau} - \bar{x}_{k\lambda}\left(u'_{l\tau} - \bar{u}\sum_{j=1}^{N}u'_{j\tau}\right)\right]. \tag{3.8}$$

Substituting Eqs. (3.7) and (3.8) into Eq. (3.4), we obtain

$$\frac{\partial x_{k\lambda}}{\partial X_{l\tau}} = \frac{1}{u}\left[(\delta_{kl} - \bar{u})\delta_{\lambda\tau} - \frac{\bar{x}_{k\lambda}\left(u'_{l\tau} - \bar{u}\sum_{j=1}^{N}u'_{j\tau}\right)}{u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}}\right]. \tag{3.9}$$

We now calculate the effective metrics $g_{i\alpha j\beta}$. Substituting Eq. (3.9) into Eq. (2.8), we find

$$g_{i\alpha j\beta} = \frac{1}{u^2}\left\{\delta_{\alpha\beta}\left[\delta_{ij} + \frac{(u^2-1)}{N}\right] - \frac{\bar{x}_{i\alpha}\left[u'_{j\beta} + [(u^2-1)/N]\sum_{l=1}^{N}u'_{l\beta}\right] + \bar{x}_{j\beta}\left[u'_{i\alpha} + [(u^2-1)/N]\sum_{l=1}^{N}u'_{l\alpha}\right]}{\left(u + \sum_{l=1}^{N}\sum_{\tau=1}^{3}\bar{x}_{l\tau}u'_{l\tau}\right)} + \frac{\bar{x}_{i\alpha}\bar{x}_{j\beta}\left[\sum_{k=1}^{N}\sum_{\lambda=1}^{3}(u'_{k\lambda})^2 - 2\bar{u}\sum_{k=1}^{N}\sum_{\lambda=1}^{3}\left(u'_{k\lambda}\sum_{l=1}^{N}u'_{l\lambda}\right) + N\bar{u}^2\sum_{\lambda=1}^{3}\left(\sum_{l=1}^{N}u'_{l\lambda}\right)^2\right]}{\left(u + \sum_{l=1}^{N}\sum_{\tau=1}^{3}\bar{x}_{l\tau}u'_{l\tau}\right)^2}\right\}. \tag{3.10}$$

We next consider the Jacobian $B = |\partial x_{k\lambda}/\partial X_{l\tau}|$. This $3N$th order determinant is calculated in Appendix C. We find, Eq. (C9),

$$B = \left[u^{3N-4}\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right)\right]^{-1}. \tag{3.11}$$

The new coordinate-dependent potential $W(\mathbf{r})$ is found by substituting Eqs. (3.10) and (3.11) into Eq. (2.9); and the transformed Hamiltonian $H(\mathbf{r}, \mathbf{p})$ is found by substituting this result and Eq. (3.10) into Eq. (2.7).

For the calculation of the hard-core many-body problem we have found in I that a useful technique was to have the transformation function u be regional, with regions determined by the smallest interparticle separation, namely

$$u(\mathbf{r}) = \frac{1}{\tau}\sum_{\substack{a,b=1\\a<b}}^{N}\prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N}\theta(r_{cd} - r_{ab})\,u_{ab}(\mathbf{r}_{ij}), \tag{3.12}$$

where

$$\tau = \sum_{\substack{a,b=1\\a<b}}^{N}\prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N}\theta(r_{cd} - r_{ab}), \tag{3.13}$$

$$u_{ab}(\mathbf{r}_{ij}) = u_{ef}(\mathbf{r}_{ij}) \text{ when } r_{ab} = r_{ef}; \tag{3.14}$$

Eq. (3.14) was required for the continuity of the transformation, and the regional transformation function u_{ab} was taken to be only a function of the interparticle coordinates $\mathbf{r}_{ij}$, since we were only concerned with transforming the intermolecular potentials. This last requirement simplifies our expressions [Eqs. (3.7)–(3.11)] since

$$\sum_{j=1}^{N}u'_{j\beta} = \sum_{j=1}^{N}\frac{\partial u(\mathbf{r}_{21}\cdots\mathbf{r}_{N(N-1)})}{\partial x_{j\beta}} \tag{3.15}$$

$$= 0. \tag{3.16}$$

Let us now consider the derivatives of the θ functions explicitly displayed in Eq. (3.12). These θ functions are used to specify regions in the $3N$-dimensional space. The derivatives of the θ functions are nonzero only at the boundary between regions, i.e., at some $r_{ef} = r_{gh}$ in Eq. (3.12). In the sum over a, b we have $\theta(r_{ef} - r_{gh})u_{gh}(\mathbf{r}_{ij}) + \theta(r_{gh} - r_{ef})u_{ef}(\mathbf{r}_{ij})$. Now $\partial\theta(r_{ef} - r_{gh})/\partial x_{k\lambda} = -\partial\theta(r_{gh} - r_{ef})/\partial x_{k\lambda}$. Therefore, since $u_{gh}(\mathbf{r}_{ij}) = u_{ef}(\mathbf{r}_{ij})$, when $r_{gh} = r_{ef}$, terms containing derivatives of θ functions will cancel one another. We note that Eq. (3.16) was required just to effect this cancellation. By considering the definition of the θ function [Eq. (2.3)], we can easily see that in the operations of addition, subtraction, multiplication, or division of functions of the regional form [Eq. (3.12)], the re-

sult will just be of that same regional form. Thus in all our expressions the θ functions will only define regions. Equations (3.10) and (3.11) can thereby be written as

$$g_{i\alpha j\beta} = \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) \times \left\{ \frac{1}{(u_{ab})^2} \left[\delta_{\alpha\beta} \left(\delta_{ij} + \frac{(u_{ab}^2 - 1)}{N} \right) - \frac{(\bar{x}_{i\alpha}(u_{ab})'_{j\beta} + \bar{x}_{j\beta}(u_{ab})'_{i\alpha})}{u_{ab} + \sum_{l=1}^{N} \sum_{\tau=1}^{3} x_{l\tau}(u_{ab})'_{l\tau}} + \frac{\bar{x}_{i\alpha}\bar{x}_{j\beta} \sum_{k=1}^{N} \sum_{\lambda=1}^{3} ((u_{ab})'_{k\lambda})^2}{(u_{ab} + \sum_{l=1}^{N} \sum_{\tau=1}^{3} x_{l\tau}(u_{ab})'_{l\tau})^2} \right] \right\} \tag{3.17}$$

$$\equiv \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab})(g_{ab})_{i\alpha j\beta}, \tag{3.18}$$

$$B = \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) \left[(u_{ab})^{(3N-4)} \left(u_{ab} + \sum_{j=1}^{N} \sum_{\beta=1}^{3} x_{j\beta}(u_{ab})'_{j\beta} \right) \right]^{-1} \tag{3.19}$$

$$\equiv \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) B_{ab}, \tag{3.20}$$

where $(u_{mn})'_{i\alpha} = \partial u_{mn}/\partial x_{i\alpha}$. Similarly, from Eqs. (2.7), (2.9), (3.18), and (3.20), the coordinate-dependent potential W can be written as

$$W(\mathbf{r}) = \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) \frac{\hbar^2}{2m} \times \sum_{i,j=1}^{N} \sum_{\alpha,\beta=1}^{3} \sqrt{B_{ab}} \frac{\partial}{\partial x_{i\alpha}} \left((g_{ab})_{i\alpha j\beta} \frac{\partial}{\partial x_{j\beta}} \frac{1}{\sqrt{B_{ab}}} \right) \tag{3.21}$$

$$\equiv \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) W_{ab}, \tag{3.22}$$

and the transformed Hamiltonian H can be written as

$$H = \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) \frac{1}{2m} \times \sum_{i,j=1}^{N} \left(\sum_{\alpha,\beta=1}^{3} p_{i\alpha}(g_{ab})_{i\alpha j\beta} p_{j\beta} + V(\mathbf{r},\mathbf{p}) + W_{ab} \right) \tag{3.23}$$

$$\equiv \frac{1}{\tau} \sum_{\substack{a,b=1\\a<b}}^{N} \prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N} \theta(r_{cd} - r_{ab}) H_{ab}. \tag{3.24}$$

In order to obtain more explicit results, while still not restricting $u_{ab}(\mathbf{r}_{ij})$ to the form of the right-hand side of Eq. (2.23), the expression used in I for the hard-core many-body problem, we shall now require

$$u_{ab}(\mathbf{r}_{ij}) = e(r_{ab}) \equiv e. \tag{3.25}$$

The calculation for the effective metric $(g_{ab})_{i\alpha j\beta}$ is performed in Appendix D. We find, Eq. (D5),

$$(g_{ab})_{i\alpha j\beta} = \frac{1}{e^2} \left[\delta_{\alpha\beta} \left(\delta_{ij} + \frac{(e^2 - 1)}{N} \right) - \left(\frac{\bar{x}_{i\alpha} e'(x_{ab\beta}/r_{ab})(\delta_{ja} - \delta_{jb}) + \bar{x}_{j\beta} e'(x_{ab\alpha}/r_{ab})(\delta_{ia} - \delta_{ib})}{(r_{ab}e)'} + \frac{2\bar{x}_{i\alpha}\bar{x}_{j\beta} e'^2}{(r_{ab}e)'^2} \right) \right], \tag{3.26}$$

where prime denotes the derivative with respect to r_{ab}. The Jacobian B_{ab} [Eqs. (3.19) and (3.20)] becomes

$$B_{ab} = [e^{3N-4}(r_{ab}e)']^{-1}. \tag{3.27}$$

The calculation for the new coordinate-dependent potential W_{ab} is performed in Appendix E. We find, Eqs. (E6) and (E7),

$$W_{ab} = \frac{\hbar^2}{m} \sqrt{B_{ab}} \sum_{\alpha=1}^{3} \frac{\partial}{\partial x_{ab\alpha}} \frac{x_{ab\alpha}}{(r_{ab}e)(r_{ab}e)'} \left(\frac{1}{\sqrt{B_{ab}}} \right)' - \frac{\hbar^2}{m} \sqrt{B_{ab}} \sum_{i=1}^{N} \sum_{\alpha=1}^{3} \frac{\partial}{\partial x_{i\alpha}} \frac{\bar{x}_{i\alpha} e'}{e(r_{ab}e)'^2} \left(\frac{1}{\sqrt{B_{ab}}} \right)', \tag{3.28}$$

$$= 3\frac{\hbar^2}{m}\frac{\sqrt{B_{ab}}}{(r_{ab}e)(r_{ab}e)'}\left(\frac{1}{\sqrt{B_{ab}}}\right)' - 3(N-1)\frac{\hbar^2}{m}\frac{\sqrt{B_{ab}}e'}{e(r_{ab}e)'^2}$$
$$\times\left(\frac{1}{\sqrt{B_{ab}}}\right)' + \frac{\hbar^2}{m}\sqrt{B_{ab}}\,r_{ab}\left[\frac{1}{r_{ab}(r_{ab}e)'^2}\left(\frac{1}{\sqrt{B_{ab}}}\right)'\right]'. \tag{3.29}$$

The transformed Hamiltonian H is obtained by substituting Eq. (3.26) and Eq. (3.28) or (3.29) into Eq. (3.23).

4. THE HARD-CORE HAMILTONIAN

In this section we shall apply our results to the hard-core problem. The hard-core Hamiltonian in the original space is

$$\tilde{H} = \frac{1}{2m}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}P_{i\alpha}^2 + \sum_{\substack{i,j=1\\ i<j}}^{N}V_{ij}, \tag{4.1}$$

where

$$\begin{aligned} V_{ij} &= \infty, \quad R_{ij} \le c, \\ &= 0, \quad R_{ij} > c. \end{aligned} \tag{4.2}$$

As discussed in Sec. 2, and in I, the transformation

$$e \equiv e(r_{ab}) = 1 + (c/r_{ab}) \tag{4.3}$$

will eliminate the regions of $3N$-dimensional space in which any interparticle separation $\boldsymbol{R}_{ij}$ is less than the hard-core diameter c. Therefore from Eq. (4.2) the transformed potential $V_{ij}(r)$ will be zero. The transformation Eq. (3.1) with Eqs. (3.12), (3.25), and (4.3) is continuous and one to one[13] in the regions in which no interparticle separation R_{ij} is less than c. We shall now calculate the transformed hard-core Hamiltonian with the transformation specified by Eq. (4.3).

Substituting Eq. (4.3) into Eqs. (3.26)–(3.29), we obtain

$$(g_{ab})_{i\alpha j\beta} = \frac{1}{[1+(c/r_{ab})]^2}\left[\delta_{a\beta}\left(\delta_{ij} + \frac{2c/r_{ab} + c^2/r_{ab}^2}{N}\right)\right.$$
$$+ (c/r_{ab}^3)[x_{i\alpha}x_{ab\beta}(\delta_{ja} - \delta_{jb})$$
$$+ \bar{x}_{j\beta}x_{ab\alpha}(\delta_{ia} - \delta_{ib})]$$
$$\left. + 2c^2 x_{i\alpha}\bar{x}_{j\beta}/r_{ab}^4\right], \tag{4.4}$$

$$B_{ab} = [1 + (c/r_{ab})]^{-(3N-4)}, \tag{4.5}$$

$$W_{ab} = -\frac{\hbar^2}{2m}\frac{(3N-4)c}{[1+(c/r_{ab})]^{(3N-4)/2}}$$
$$\times \sum_{\alpha=1}^{3}\frac{\partial}{\partial x_{ab\alpha}}\left(\frac{x_{ab\alpha}[1+(c/r_{ab})]^{(3N-6)/2}}{r_{ab}^2(r_{ab}+c)}\right) - \frac{\hbar^2}{2m}\frac{(3N-4)c^2}{[1+(c/r_{ab})]^{(3N-4)/2}}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_{i\alpha}}\left(\frac{\bar{x}_{i\alpha}(1+(c/r_{ab}))^{(3N-6)/2}}{r_{ab}^3(r_{ab}+c)}\right) \tag{4.6}$$

$$= -3\frac{\hbar^2}{2m}\frac{(3N-4)c}{r_{ab}(r_{ab}+c)^2} - 3(N-1)\frac{\hbar^2}{2m}\frac{(3N-4)c^2}{r_{ab}^2(r_{ab}+c)^2}$$
$$-\frac{\hbar^2}{2m}\frac{(3N-4)c\,r_{ab}}{(1+c/r_{ab})^{(3N-4)/2}}\frac{1}{r_{ab}^3}\left[(1+c/r_{ab})^{(3N-6)/2}\right]' \tag{4.7}$$

$$= -3\frac{\hbar^2}{2m}\frac{(3N-4)(N-2)c^2}{2r_{ab}^2(r_{ab}+c)^2}. \tag{4.8}$$

Substituting Eqs. (4.4) and (4.8) into Eq. (3.23) we obtain the transformed hard-core Hamiltonian.

$$H = \frac{1}{\tau}\sum_{\substack{a,b=1\\ a<b}}^{N}\prod_{\substack{c,d=1\\ c<d\\ cd\neq ab}}^{N}\theta(r_{cd}-r_{ab})\left\{\frac{1}{2m}\sum_{k,l=1}^{N}\sum_{\alpha,\beta=1}^{3}p_{k\alpha}\frac{1}{[1+(c/r_{ab})]^2}\right.$$
$$\times\left[\delta_{\alpha\beta}\left(\delta_{kl} + \frac{2c/r_{ab}+c^2/r_{ab}^2}{N}\right) + (c/r_{ab}^3)[\bar{x}_{k\alpha}x_{ab\beta}(\delta_{la}-\delta_{lb}) + \bar{x}_{l\beta}x_{ab\alpha}(\delta_{ka}-\delta_{kb})] + 2c^2\bar{x}_{k\alpha}\bar{x}_{l\beta}/r_{ab}^4\right]p_{l\beta}$$
$$\left. -3\frac{\hbar^2}{2m}\frac{(3N-4)(N-2)c^2}{2r_{ab}^2(r_{ab}+c)^2}\right\} \tag{4.9}$$

$$= \frac{1}{\tau} \sum_{\substack{a,b=1 \\ a<b}}^{N} \prod_{\substack{c,d=1 \\ c<d \\ cd \neq ab}}^{N} \theta(r_{cd} - r_{ab}) \left[\frac{1}{2m} \sum_{k=1}^{N} \sum_{\alpha=1}^{3} p_{k\alpha} \frac{1}{[1 + (c/r_{ab})]^2} p_{k\alpha} \right.$$

$$+ \frac{1}{2m} \sum_{k,l=1}^{N} \sum_{\alpha,\beta=1}^{3} p_{k\alpha} \frac{[(2c/r_{ab} + c^2/r_{ab}^2)/N]}{[1 + (c/r_{ab})]^2} p_{l\beta} + \frac{c}{m} \sum_{k=1}^{N} \sum_{\alpha,\beta=1}^{3} \left(p_{k\alpha} \frac{\bar{x}_{k\alpha} x_{ab\beta}}{r_{ab}(r_{ab} + c)^2} p_{ab\beta} \right.$$

$$\left. \left. + p_{ab\alpha} \frac{\bar{x}_{k\beta} x_{ab\alpha}}{r_{ab}(r_{ab} + c)^2} p_{k\beta} \right) + \frac{c^2}{m} \sum_{k,l=1}^{N} \sum_{\alpha,\beta=1}^{3} p_{k\alpha} \frac{\bar{x}_{k\alpha} \bar{x}_{l\beta}}{r_{ab}^2(r_{ab} + c)^2} p_{l\beta} - 3 \frac{\hbar^2}{2m} \frac{(3N-4)(N-2)c^2}{2r_{ab}^2(r_{ab} + c)^2} \right] , \quad (4.10)$$

where $p_{ab\alpha}$ is the relative momentum, $p_{ab\alpha} = \frac{1}{2}(p_{a\alpha} - p_{b\alpha})$. The transformed Hamiltonian, Eq. (4.10), is Hermitian, Fourier analyzable, and amenable to ordinary perturbation and variational techniques, while being completely equivalent to the original hard-core Hamiltonian [Eq. (4.1) with Eq. (4.2)]. We can now use this Hamiltonian to calculate the energy spectrum of the system.

5. GROUND STATE ENERGY

The energy of a system is given by

$$E = \int \frac{\psi^\dagger(\mathbf{r}) H \psi(\mathbf{r}) d^{3N}r}{\psi^\dagger(\mathbf{r})\psi(\mathbf{r}) d^{3N}r} , \quad (5.1)$$

where $\psi(\mathbf{r})$ is the exact wavefunction of the system in the transformed space. If $\psi(\mathbf{r})$ is not the correct ground state wavefunction then $E > E_G$, where E_G is the correct ground energy. In this section we shall consider the application of the method developed to the calculation of the ground state energy of hard-core systems.

Let us first consider the isolated two-body system. For the ground state energy calculation, we can choose, as our trial wavefunction, the zero momentum wavefunction in the transformed space

$$\psi(\mathbf{r}) = \text{const.} \quad (5.2)$$

The momentum-dependent terms in Eq. (5.1) operating on $\psi(\mathbf{r})$ given by Eq. (5.2) are zero. From Eq. (4.8), $W = 0$. Therefore we obtain $E = 0$, the exact minimum energy for this isolated two-body system.

Let us now consider the many-body system. For the ground state energy of the hard-core Bose system problems arise if one naively chooses Eq. (5.2), the zero momentum wavefunction in the transformed space, as a trial wavefunction. The momentum-dependent terms in Eq. (4.10) operating on $\psi(\mathbf{r})$ given by Eq. (5.2) are zero. Therefore substituting Eqs. (4.10) and (5.2) into Eq. (5.1) one would obtain a negative energy solution for the ground state energy for $N > 2$. This solution is obviously incorrect since the original potential is purely repulsive. Let us now consider the cause of this discrepancy.

If our system is confined in a volume Ω, then

$$\tilde{\psi}(\mathbf{R}) = 0 \text{ at the walls of the container.} \quad (5.3)$$

From Eqs. (2.10) and (4.5) we therefore must require

$$\psi(\mathbf{r}) = 0 \text{ at the walls of the container.} \quad (5.4)$$

The choice of $\tilde{\psi}(\mathbf{r}) = \text{const}$, Eq. (5.2), with the normalization condition on $\psi(\mathbf{r})$, $\int \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) d^{3N}r = 1$, does not satisfy this condition. Therefore Eq. (5.2) cannot be used as a trial wavefunction for $\psi(\mathbf{r})$.

To demonstrate the importance of this point, consider the contribution to the ground state energy of W,

$$E_W = \int \psi^\dagger(\mathbf{r}) W(\mathbf{r}) \psi(\mathbf{r}) d^{3N}r. \quad (5.5)$$

From Eqs. (3.22) and (4.8)

$$E_W = \frac{-3(3N-4)(N-2)c^2\hbar^2}{4m} \int \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \times \frac{1}{\tau} \sum_{\substack{a,b=1 \\ a<b}}^{N} \prod_{\substack{c,d=1 \\ c<d \\ cd \neq ab}}^{N} \theta(r_{cd} - r_{ab}) \frac{d^{3N}r}{r_{ab}^2(r_{ab} + c)^2} . \quad (5.6)$$

Therefore E_W is always less than zero. However, let us return to W given in the form Eq. (2.9). In that case we have

$$E_W = \frac{\hbar^2}{2m} \int \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \sum_{i,j=1}^{N} \sum_{\alpha,\beta=1}^{3} \sqrt{B} \times \frac{\partial}{\partial x_{i\alpha}} \left[g_{i\alpha j\beta} \left(\frac{\partial}{\partial x_{j\beta}} \frac{1}{\sqrt{B}} \right) \right] d^{3N}r . \quad (5.7)$$

Integrating by parts we find

$$\int \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \sqrt{B} \frac{\partial}{\partial x_{i\alpha}} \left(g_{i\alpha j\beta} \frac{\partial}{\partial x_{j\beta}} \frac{1}{\sqrt{B}} \right) dx_{i\alpha}$$

$$= \left\{ \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \sqrt{B} \frac{\partial}{\partial x_{i\alpha}} \left[g_{i\alpha j\beta} \left(\frac{\partial}{\partial x_{j\beta}} \frac{1}{\sqrt{B}} \right) \right] \right\}_{\text{boundary}}$$

$$- \int g_{i\alpha j\beta} \left(\frac{\partial}{\partial x_{j\beta}} \frac{1}{\sqrt{B}} \right) \left(\frac{\partial}{\partial x_{i\alpha}} \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \sqrt{B} \right) dx_{i\alpha} . \quad (5.8)$$

Using the condition $\psi(\mathbf{r}) = 0$ at the boundary and substituting Eq. (5.8) into Eq. (5.7), we obtain

$$E_W = \frac{\hbar^2}{2m}\Bigg[-\int \sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\sqrt{B}\; g_{i\alpha j\beta}\left(\frac{\partial}{\partial x_{j\beta}}\;\frac{1}{\sqrt{B}}\right)\times\left(\psi^\dagger(\mathbf{r})\frac{\partial\psi(\mathbf{r})}{\partial x_{i\alpha}}+\psi(\mathbf{r})\frac{\partial\psi^\dagger(\mathbf{r})}{\partial x_{i\alpha}}\right)d^{3N}r - \int\psi^\dagger(\mathbf{r})\psi(\mathbf{r})\sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3} g_{i\alpha j\beta}\left(\frac{\partial}{\partial x_{j\beta}}\;\frac{1}{\sqrt{B}}\right)\times\left(\frac{\partial\sqrt{B}}{\partial x_{i\alpha}}\right)d^{3N}r\Bigg]. \tag{5.9}$$

The integrand in the second term in the brackets is calculated in Appendix F. Substituting Eq. (F5) into Eq. (5.9), we have

$$E_W = \frac{\hbar^2}{2m}\Bigg[-\int \sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\sqrt{B}\; g_{i\alpha i\beta}\left(\frac{\partial}{\partial x_{j\beta}}\;\frac{1}{\sqrt{B}}\right)\times\left(\psi^\dagger(\mathbf{r})\frac{\partial\psi(\mathbf{r})}{\partial x_{i\alpha}}+\psi(\mathbf{r})\frac{\partial\psi^\dagger(\mathbf{r})}{\partial x_{i\alpha}}\right)d^{3N}r + 2(\tfrac{3}{2}N-2)^2c^2\int\psi^\dagger(\mathbf{r})\psi(\mathbf{r})\frac{1}{\tau}\times\sum_{\substack{a,b=1\\a<b}}^{N}\prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N}\theta(r_{cd}-r_{ab})\frac{1}{r_{ab}^2(r_{ab}+c)^2}\,d^{3N}r\Bigg]. \tag{5.10}$$

If we now choose $\psi(\mathbf{r}) = \text{const}$ [Eq. (5.2)] as our wavefunction, the first integral in Eq. (5.10) would be zero. As before, the momentum-dependent terms in Eq. (4.10) operating on $\psi(\mathbf{r})$ would be zero. Therefore, assuming $\psi(\mathbf{r})$ to be normalized, we would be left with

$$E = (\tfrac{3}{2}N-2)^2\;\frac{\hbar^2}{m}\;c^2\int\psi^\dagger(\mathbf{r})\psi(\mathbf{r})\frac{1}{\tau}\times\sum_{\substack{a,b\\a<b}}^{N}\prod_{\substack{c,d=1\\c<d\\cd\neq ab}}^{N}\theta(r_{cd}-r_{ab})\;\frac{1}{r_{ab}^2(r_{ab}+c)^2}\;d^{3N}r, \tag{5.11}$$

a purely positive quantity. Therefore, the boundary conditions on $\tilde{\psi}(\mathbf{R})$, and thus on $\psi(\mathbf{r})$, have "changed" the purely negative E_W into a positive quantity. The contribution at the walls of the container, which the boundary conditions cause to be neglected, would thus be highly negative if $\psi(\mathbf{r})$ were taken to be a constant. This contribution is caused by the extended range, in the $3N$-dimensional transformed space, of the new potential. We note that for the case $N = 2$, $\psi(\mathbf{r}) = \text{const}$ determined by the normalization condition we would obtain from Eq. (5.11),

$$E = \frac{\hbar^2}{m}\;\frac{c^2}{K}\int\frac{d^3r_{21}}{r_{21}^2(r_{21}+c)^2} \tag{5.12}$$

$$= 4\pi\, c\hbar^2/mK, \tag{5.13}$$

where K is a constant. In the dilute system limit $K \simeq \Omega$ and assuming pair-wise additivity, we would have for $N(N-1)/2$ pairs, $E \to 2\pi\, c\hbar^2N^2/m\Omega$ the proper result.[5] In addition, we note that before performing the integration by parts, the answer for $N = 2$ was $E = 0$, the proper isolated two-body scattering result. The effect of a different formulation of the boundary conditions [Eq. (5.3)] has previously been considered for the $N = 2$ problem.[14]

From the foregoing, it appears that boundary conditions are necessary and require further investigation. We note that in I, boundary conditions of the form Eq. (5.3) were not used; rather the Hamiltonian contained cutoff functions. In the case of the Hamiltonian [Eq. (4.10)] a procedure equivalent to that of I would be to expand the trial wavefunction in a regional form similar to that of Eq. (2.1). The problem of regionality would still remain, but the procedure appears to be computationally easier than that of I and could be useful for calculations of the properties of medium density systems.

6. CONCLUSIONS

We have derived an explicit Hamiltonian which is completely equivalent to the original hard-core many-body Hamiltonian, while being Fourier analyzable, Hermitian, and amenable to perturbational and variational techniques. In addition, the interacting particle cluster expansion used in a previous paper[1] has been eliminated. By no longer needing to use a cluster expansion, we are no longer restricted to the consideration of dense gases, but can now consider true liquids.

The disadvantages of this Hamiltonian are (i) it is regional in that it depends upon the smallest interparticle separation and (ii) it does not implicitly contain a cutoff function, i.e., for particle i, the Hamiltonian does not approach the free-body Hamiltonian for large values of r_i. This latter difficulty results in restrictions on the set of trial wavefunctions as discussed in Sec. 5.

We note the following:

(1) This method, applied to the hard-core problem in Sec. 4 and 5, could also be used for strongly repulsive but nonhard-core potentials if we used $e(\mathbf{r}_{ab}) = 1 + c/(r_{ab} + \epsilon'')$ instead of $e(\mathbf{r}_{ab}) = 1 + c/r_{ab}$ [Eq. (4.3)]. The parameter ϵ'' would be determined, as discussed in I, by the shrinkage $c \to \epsilon$ required to make the effects of the transformed strong, repulsive potential, sufficiently small.

(2) An alternate method of eliminating the transformed hard-core potential would be to have the transformation shrink the system about an arbitrary origin instead of about the center of mass of the system. The transformation would then be given by Eq. (3.1) without the $(1-u)\bar{x}_\alpha$ term, i.e., by

$$X_{i\alpha} = ux_{i\alpha}. \tag{6.1}$$

The functional form of u would be unchanged from that given in this paper. The calculation of the Jacobian of the transformation, and of the transformed Hamiltonian is straightforward.

ACKNOWLEDGMENT

The author would like to thank Dr. John D. Stettler for his helpful suggestions in the preparation of this manuscript.

APPENDIX A: CALCULATION OF D

By definition, from Eqs. (3.3), (3.4),

$$D = \begin{vmatrix} (u+\bar{u}+\bar{x}_{11}u'_{11})(\bar{u}+\bar{x}_{11}u'_{21}) & \dots (\bar{u}+\bar{x}_{11}u'_{N1}) & \bar{x}_{11}u'_{12} & \dots \bar{x}_{11}u'_{N3} \\ (\bar{u}+\bar{x}_{21}u'_{11}) & (u+\bar{u}+\bar{x}_{21}u'_{21}) \dots (\bar{u}+\bar{x}_{21}u'_{N1}) & \bar{x}_{21}u'_{12} & \dots \bar{x}_{21}u'_{N3} \\ \dots & \dots & \dots & \\ (\bar{u}+\bar{x}_{N1}u'_{11}) & (\bar{u}+\bar{x}_{N1}u'_{21}) \dots (u+\bar{u}+\bar{x}_{N1}u'_{N1}) & \bar{x}_{N1}u'_{12} & \dots \bar{x}_{N1}u'_{N3} \\ \bar{x}_{12}u'_{11} & \bar{x}_{12}u'_{21} \dots \bar{x}_{12}u'_{N1} & (u+\bar{u}+\bar{x}_{12}u'_{12}) \dots \bar{x}_{12}u'_{N3} & \\ \dots & \dots & & \\ \bar{x}_{N3}u'_{11} & \bar{x}_{N3}u'_{21} \dots \bar{x}_{N3}u'_{N1} & \bar{x}_{N3}u'_{12} & \dots (u+\bar{u}+\bar{x}_{N3}u'_{N3}) \end{vmatrix} \tag{A1}$$

We shall use the following theorem[15]:

$$|a_{ik}+b_{ik}| = |a_{ik}| + |b_{ik}| + \sum \Delta a_{3r-1-\alpha} b_{\alpha+1}, \tag{A2}$$

where $3r-2 \geq \alpha \geq 0$; $\Delta\, a_{3r-1-\alpha}b_{\alpha+1}$ denotes the determinant formed as follows: The first $3r-1-\alpha$ columns are taken from $|a_{ik}|$ and the remaining $\alpha+1$ columns are taken from $|b_{ik}|$ with the proviso that no two columns thus taken have the same column number; and $\sum$ denotes the sum over all the possible combinations. Identify $(u\delta_{ik}\delta_{\alpha\lambda} + \bar{u}\delta_{\alpha\lambda})$ with $a_{[i+N(\alpha-1)][k+N(\lambda-1)]}$ and $x_{i\alpha}u'_{k\lambda}$ with $b_{[i+N(\alpha-1)][k+N(\lambda-1)]}$, $1 \leq i, k \leq N$, $1 \leq \alpha, \lambda \leq 3$.

Set

$$D^{(1)} = \left| a_{[i+N(\alpha-1)][k+N(\lambda-1)]} \right| \tag{A3}$$

$$= \left[\begin{vmatrix} (u+\bar{u}) & \bar{u} & \cdot & \bar{u} \\ \bar{u} & (u+\bar{u}) & \cdot & \bar{u} \\ \cdot & \cdot & \cdot & \cdot \\ \bar{u} & \bar{u} & \cdot & (u+\bar{u}) \end{vmatrix} \right]^3 \tag{A4}$$

$$\equiv [D^{(1a)}]^3. \tag{A5}$$

Identifying $u\delta_{ik}$ with $a^{(1a)}_{ik}$ and $\bar{u}$ with $b^{(1a)}_{ik}$, $[1 \leq i, k \leq N]$, we find

$$\left| a^{(1a)}_{ik} \right| = u^N, \tag{A6}$$

$$\left| b^{(1a)}_{ik} \right| = \bar{u}^N \begin{vmatrix} 1 & 1 & \cdot & 1 \\ 1 & 1 & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & 1 \end{vmatrix} \tag{A7}$$

$$= 0. \tag{A8}$$

From Eq. (A7) all terms in $\sum \Delta a^{(1a)}_{3r-1-\alpha} b^{(1a)}_{\alpha+1}$ are zero except those containing only one column of $b^{(1a)}$. Therefore

$$\sum \Delta a^{(1a)}_{3r-1-\alpha} b^{(1a)}_{\alpha+1} = N u^{3N-1} \bar{u}. \tag{A9}$$

Thus combining Eqs. (A3)–(A9), we obtain

$$D^{(1)} = u^{3(N-1)}(u+N\bar{u})^3 \tag{A10}$$

$$= u^{3(N-1)}. \tag{A11}$$

Set

$$D^{(2)} = \left| b_{[i+N(\alpha-1)][k+N(\lambda-1)]} \right| \tag{A12}$$

$$= \prod_{m,n=1}^{N} \prod_{\alpha,\xi=1}^{3} u'_{n\xi} \begin{vmatrix} \bar{x}_{11} & \bar{x}_{11} & \cdot & \bar{x}_{11} \\ \bar{x}_{21} & \bar{x}_{21} & \cdot & \bar{x}_{21} \\ \cdot & \cdot & \cdot & \cdot \\ \bar{x}_{N3} & \bar{x}_{N3} & \cdot & \bar{x}_{N3} \end{vmatrix} \tag{A13}$$

$$= 0. \tag{A14}$$

Set

$$D^{(3)} = \sum \Delta a_{3r-1-\alpha} b_{\alpha+1}. \tag{A15}$$

From Eq. (A13) all terms in $\sum \Delta a_{3r-1-\alpha} b_{\alpha+1}$ are zero except those containing only one column of b. Consider the case where the $j$bth column is taken from b. Simplifying the determinant we obtain

$$D^{(3)}_{j\beta} = u^{2N-2}u'_{j\beta}\begin{vmatrix} (u+\bar{u}) & \bar{u} & . & \bar{x}_{1\beta} & . & \bar{u} \\ \bar{u} & (u+\bar{u}) & . & \bar{x}_{2\beta} & . & \bar{u} \\ . & . & . & . & . & . \\ \bar{u} & \bar{u} & . & \bar{x}_{j\beta} & . & \bar{u} \\ . & . & . & . & . & . \\ \bar{u} & \bar{u} & . & \bar{x}_{N\beta} & . & (u+\bar{u}) \end{vmatrix} . \tag{A16}$$

where for illustrational convenience we have set $j \neq 1, 2, N$. Identifying $u\delta_{ik}(1-\delta_{kj}) + x_{i\beta}\delta_{kj}$ with $a^{(3)}_{ik}$ and $\bar{u}(1-\delta_{kj})$ with $b^{(3)}_{ik}$, $1 \leq i, k \leq N$, we find

$$\left|a^{(3)}_{ik}\right| = u^{N-1}\bar{x}_{j\beta}\,, \tag{A17}$$

$$\left|b^{(3)}_{ik}\right| = 0. \tag{A18}$$

All terms in $\sum \Delta a^{(3)}_{3\gamma-1-\alpha} b^{(3)}_{\alpha+1}$ are zero except those containing only one column of b, in particular, one column of $\bar{u}$. Considering the case where this column is the m^{th} column, $m \neq j$, we obtain, after rearranging terms,

$$\Delta a^{(3)}_{3\gamma-1-\alpha} b^{(3)}_{\alpha+1}\Big|_m = \begin{vmatrix} u & 0 & . & 0 & \bar{u} & \bar{x}_{1\beta} \\ 0 & u & . & 0 & \bar{u} & \bar{x}_{2\beta} \\ . & . & . & . & . & . \\ 0 & 0 & . & u & \bar{u} & \bar{x}_{N\beta} \\ 0 & 0 & . & 0 & \bar{u} & \bar{x}_{m\beta} \\ 0 & 0 & . & 0 & \bar{u} & \bar{x}_{j\beta} \end{vmatrix} \tag{A19}$$

$$= u^{N-2}\bar{u}(\bar{x}_{j\beta} - \bar{x}_{m\beta}), \tag{A20}$$

where for illustrational convenience we have set $j, m \neq 1, 2, N$. For $m = j$,

$$\Delta a^{(3)}_{3\gamma-1-\alpha} b^{(3)}_{\alpha+1}\Big|_{m=j} = 0. \tag{A21}$$

Summing over m and combining Eqs. (A16)–(A21), we find

$$D^{(3)}_{j\beta} = u^{3N-4}u'_{j\beta}\left([u+(N-1)\bar{u}]\bar{x}_{j\beta} - \bar{u}\sum_{\substack{m=1 \\ m\neq j}}^{N}\bar{x}_{m\beta}\right) \tag{A22}$$

$$= u^{3N-4}\bar{x}_{j\beta}u'_{j\beta}\,. \tag{A23}$$

Therefore summing over j, β and combining Eqs. (A11), (A14), and (A23), we obtain

$$D = u^{3N-4}\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right). \tag{A24}$$

APPENDIX B: CALCULATION OF $E_{k\lambda,l\tau}$

By definition, from Eqs. (3.3), (3.4),

$$E_{k\lambda,l\tau} = \begin{vmatrix} & & & 0 & & & \\ & & & . & & & \\ & & & . & & & \\ & & & 0 & & & \\ & & & 1 & & & \\ & & & 0 & & & \\ & & & . & & & \\ & & & . & & & \\ & & & 0 & & & \end{vmatrix}, \tag{B1}$$

where the constant 1 is in the $[k + N(\gamma-1)]$ column, $[l + N(\tau-1)]$ row, and the matrices represented by large parentheses are the same terms as those in the same position in the determinant of the coefficients D, given by Eq. (A1).

We separate $E_{k\lambda,l\tau}$ into three possibilities:

(1) $k = l$, $\lambda = \tau$, (2) $k \neq l$, $\lambda = \tau$, (3) $\lambda \neq \tau$.

The calculations for each of these possibilities is performed using the method illustrated in Appendix A. For possibility (1) we find

$$E_{k\lambda,k\lambda} = \{u^{3N-4}[u+(N-1)\bar{u}]\} + \{0\} + \left\{u^{3N-5}\left[\sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}u'_{k\lambda} - \bar{u}\left(\sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\lambda}\right)\right]\right\} \tag{B2}$$

$$= u^{3N-5}\left[u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}u'_{k\lambda} - \bar{u}\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\lambda}\right)\right]. \tag{B3}$$

For possibility (2) we find

$$E_{k\lambda,l\lambda}\big|_{k\neq l} = \{-\alpha^{3N-4}\bar{u}\} + \{0\} + \left\{-\alpha^{3N-5}\left[\bar{x}_{k\lambda}u'_{l\lambda} + \bar{u}\left(\sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\lambda}\right)\right]\right\} \tag{B4}$$

$$= -u^{3N-5}\left[\bar{x}_{k\lambda}u'_{l\lambda} + \bar{u}\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta} - \bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\lambda}\right)\right]. \tag{B5}$$

For possibility (3) we find

$$E_{k\lambda,l\tau}|_{\lambda\neq\tau} = \{0\} + \{0\} + \left\{u^{3N-5}\left(\bar{u}\bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\tau} - \bar{x}_{k\lambda}u'_{l\tau}\right)\right\} \tag{B6}$$

$$= -u^{3N-5}\left(\bar{x}_{k\lambda}u'_{l\tau} - \bar{u}\bar{x}_{k\lambda}\sum_{j=1}^{N}u'_{j\tau}\right). \tag{B7}$$

Combining Eqs. (B3), (B5), and (B7) we obtain the general result,

$$E_{k\lambda,l\tau} = u^{3N-5}\left[\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right)(\delta_{kl} - \bar{u})\,\delta_{\lambda\tau} - \bar{x}_{k\lambda}\left(u'_{l\tau} - \bar{u}\sum_{j=1}^{N}u'_{j\tau}\right)\right]. \tag{B8}$$

APPENDIX C: CALCULATION OF $B \equiv |\partial x_{i\alpha}/\partial X_{k\gamma}|$

For convenience in notation we set

$$G = u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}, \tag{C1}$$

$$J'_{l\tau} = u'_{l\tau} - \bar{u}\sum_{j=1}^{N}u'_{j\tau}. \tag{C2}$$

By definition, from Eq. (3.9), we can thereby write B as

$$B = \left(\frac{1}{u^{3N}}\right)\begin{vmatrix} \left(1 - \bar{u} - \frac{J'_{11}x_{11}}{G}\right) & \left(-\bar{u} - \frac{J'_{21}x_{11}}{G}\right) & \cdots & \left(-\frac{J'_{N3}x_{11}}{G}\right) \\ \left(-\bar{u} - \frac{J'_{11}x_{21}}{G}\right) & 1 - \bar{u} - \frac{J'_{21}x_{21}}{G} & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \left(-\frac{J'_{11}x_{N3}}{G}\right) & \cdot & \cdots & \left(1 - \bar{u} - \frac{J'_{N3}x_{N3}}{G}\right) \end{vmatrix}. \tag{C3}$$

This determinant is of the same form as that of Eq. (A1). Thus from Eqs. (A10), (A14), and (A22), we obtain

$$B^{(1)} = (1 - N\bar{u})^3/u^{3N}, \tag{C4}$$

$$B^{(2)} = 0, \tag{C5}$$

$$B^{(3)}_{j\beta} = -(1 - N\bar{u})^2 J'_{j\beta}\left([1 - (N-1)\bar{u}]\bar{x}_{j\beta} + \bar{u}\sum_{\substack{m=1\\ m\neq j}}^{N}\bar{x}_{m\beta}\right)/Gu^{3N} \tag{C6}$$

$$= -(1 - N\bar{u})^3\bar{x}_{j\beta}J'_{j\beta}/Gu^{3N}. \tag{C7}$$

Therefore, summing over j, β and combining terms, we obtain

$$B = u^{-(3N-3)}\left[1 - \frac{\sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}\left(u'_{j\beta} - \bar{u}\sum_{k=1}^{N}u'_{k\beta}\right)}{\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right)}\right] \tag{C8}$$

$$= u^{-(3N-4)}\left[\left(u + \sum_{j=1}^{N}\sum_{\beta=1}^{3}\bar{x}_{j\beta}u'_{j\beta}\right)\right]^{-1}. \tag{C9}$$

APPENDIX D: CALCULATION OF $(g_{ab})_{i\alpha j\beta}$ WITH $u_{ab}(\mathbf{r}_{ij}) = e(r_{ab})$

Notationally we set $x_\alpha = x_{ab\alpha}$, $r = r_{ab}$, $e = e(r_{ab})$, $A'_{j\beta} = \partial A/\partial x_{j\beta}$, and $A' = \partial A/\partial r$, where A is an arbitrary function. From Eqs. (3.17) and (3.18),

$$(g_{ab})_{i\alpha j\beta} = \frac{1}{e^2}\left\{\delta_{\alpha\beta}\left[\delta_{ij} + \frac{(e^2-1)}{N}\right] - \left(\frac{x_{i\alpha}e'_{j\beta} + x_{j\beta}e'_{i\alpha}}{e + \sum_{l=1}^{N}\sum_{\tau=1}^{3}x_{l\tau}e'_{l\tau}}\right) + \frac{\bar{x}_{i\alpha}\bar{x}_{j\beta}\sum_{k=1}^{N}\sum_{\lambda=1}^{3}e'^2_{k\lambda}}{\left(e + \sum_{l=1}^{N}\sum_{\tau=1}^{3}x_{l\tau}e'_{l\tau}\right)^2}\right\}. \tag{D1}$$

The following relations hold:

$$\sum_{j=1}^{N}\sum_{\beta=1}^{3}x_{j\beta}f'_{j\beta}(r) = rf'(r), \tag{D2}$$

$$\sum_{j=1}^{N}\sum_{\beta=1}^{3}f'_{j\beta}(r)g'_{j\beta}(r) = 2f'g', \tag{D3}$$

where f and g are arbitrary functions. Therefore,

$$e + \sum_{l=1}^{N}\sum_{\tau=1}^{3}x_{l\tau}e'_{l\tau} = (re)', \tag{D4}$$

and Eq. (D1) becomes

$$(g_{ab})_{i\alpha j\beta} = \frac{1}{e^2}\left[\delta_{\alpha\beta}\left(\delta_{ij} + \frac{(e^2-1)}{N}\right) - \frac{\bar{x}_{i\alpha}e'(x_\beta/r)(\delta_{ja}-\delta_{jb}) + \bar{x}_{j\beta}e'(x_\alpha/r)\delta_{ia}-\delta_{ib})}{(re)'} + \frac{2\bar{x}_{i\alpha}\bar{x}_{j\beta}e'^2}{(re)'^2}\right]. \quad \text{(D5)}$$

APPENDIX E: CALCULATION OF W_{ab} WITH $u_{ab}(\mathbf{r}_{ij}) = e(r_{ab})$

From Eqs. (3.21) and (3.22),

$$W_{ab} = \frac{\hbar^2}{2m}\sqrt{B_{ab}} \times \sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\frac{\partial}{\partial x_{i\alpha}}\left(g_{ab})_{i\alpha j\beta}\frac{\partial}{\partial x_{j\beta}}\frac{1}{\sqrt{B_{ab}}}\right). \quad \text{(E1)}$$

Using Eq. (3.26) and the notation of Appendix D with $\mathfrak{B} = B_{ab}$, Eq. (E1) becomes

$$W_{ab} = \frac{\hbar^2}{2m}\sqrt{\mathfrak{B}} \times \sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\left\{\frac{\partial}{\partial x_{i\alpha}}\frac{1}{e^2}\left[\delta_{\alpha\beta}\left(\delta_{ij}+\frac{(e^2-1)}{N}\right) - \frac{\bar{x}_{i\alpha}e'(x_\beta/r)(\delta_{ja}-\delta_{jb}) + \bar{x}_{j\beta}e'(x_\alpha/r)(\delta_{ia}-\delta_{ib})}{(re)'} + \frac{2\bar{x}_{i\alpha}\bar{x}_{j\beta}e'^2}{(re)'^2}\right]\frac{\partial}{\partial x_{j\beta}}\frac{1}{\sqrt{\mathfrak{B}}}\right\}, \quad \text{(E2)}$$

where B is given by Eq. (3.27). The following relations hold:

$$\sum_{i=1}^{N}\frac{\partial f(r)}{\partial x_{i\alpha}} = 0, \quad \text{(E3)}$$

$$\frac{\partial}{\partial x_{a\alpha}} - \frac{\partial}{\partial x_{b\alpha}} = 2\frac{\partial}{\partial x_\alpha}, \quad \text{(E4)}$$

where f is an arbitrary function. Equation (E4) is obtained from the transformation $x_{a\alpha}, x_{b\alpha} \to x_\alpha = x_{a\alpha} - x_{b\alpha}$, $\mathfrak{X}_\alpha = (x_{a\alpha} + x_{b\alpha})/2$. Using Eqs. (E3), (E4), and (E2) we find

$$W_{ab} = \hbar^2/m\sqrt{\mathfrak{B}}\left\{\left[\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_\alpha}\frac{1}{e^2}\frac{x_\alpha}{r}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'\right] - \left[\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_{i\alpha}}\frac{1}{e^2}\frac{\bar{x}_{i\alpha}e'}{(re)'}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'\right] - \left[\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_\alpha}\frac{1}{e^2}\frac{x_\alpha e'}{(re)'}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'\right] + \left[\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_{i\alpha}}\frac{1}{e^2}\frac{\bar{x}_{i\alpha}e'^2r}{(re)'^2}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'\right]\right\} \quad \text{(E5)}$$

$$= \frac{\hbar^2}{m}\sqrt{\mathfrak{B}}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_\alpha}\frac{x_\alpha}{(re)(re)'}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)' - \frac{\hbar^2}{m}\sqrt{\mathfrak{B}}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x_{i\alpha}}\frac{\bar{x}_{i\alpha}e'}{e(re)'^2}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'. \quad \text{(E6)}$$

Taking the derivatives with respect to x_α and $x_{i\alpha}$, we obtain

$$W_{ab} = 3\frac{\hbar^2}{m}\frac{\sqrt{\mathfrak{B}}}{(re)(re)'}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)' - 3(N-1)\frac{\hbar^2}{m}\frac{\sqrt{\mathfrak{B}}e'}{e(re)'^2}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)' + \frac{\hbar^2}{m}\sqrt{\mathfrak{B}}\,r\left[\frac{1}{r(re)'^2}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'\right]'. \quad \text{(E7)}$$

APPENDIX F: CALCULATION FOR EQ. (5.9)

Set

$$I = -\frac{\hbar^2}{m}\sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}g_{i\alpha j\beta}\left(\frac{\partial}{\partial x_{j\beta}}\frac{1}{\sqrt{\mathfrak{B}}}\right)\left(\frac{\partial\sqrt{\mathfrak{B}}}{\partial x_{i\alpha}}\right) \quad \text{(F1)}$$

$$= \frac{1}{\tau}\sum_{\substack{c,d=1\\c<d}}^{N}\prod_{\substack{a,b=1\\a<b\\i_ai_b\neq i_ci_d}}^{N}\theta(r_{i_ai_b}-r_{i_ci_d})I_{i_ci_d}. \quad \text{(F2)}$$

From Eqs. (E1)–(E6),

$$I_{ab} = -\frac{\hbar^2}{m}\sum_{\alpha=1}^{3}\left(\frac{\partial}{\partial x_\alpha}\sqrt{\mathfrak{B}}\right)\frac{x_\alpha}{(re)(re)'}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)' + \frac{\hbar}{m}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left(\frac{\partial\sqrt{\mathfrak{B}}}{\partial x_{i\alpha}}\right)\frac{\bar{x}_{i\alpha}e'}{e(re)'^2}\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)' \quad \text{(F3)}$$

$$= -\frac{\hbar^2}{m}\frac{1}{(re)'^2}\left(\sqrt{\mathfrak{B}}\right)'\left(\frac{1}{\sqrt{\mathfrak{B}}}\right)'. \quad \text{(F4)}$$

For e given by Eq. (4.3), we obtain

$$I_{ab} = (\tfrac{3}{2}N-2)^2\frac{\hbar^2}{m}\frac{c^2}{r^2(r+c)^2}. \quad \text{(F5)}$$

[1] N. M. Witriol, J. Math. Phys. 11, 669 (1970).
[2] F. M. Eger and E. P. Gross, Ann. Phys. (N.Y.) 24, 63 (1963).
[3] F. M. Eger and E. P. Gross, Nuovo Cimento 34, 1225 (1964).
[4] F. M. Eger and E. P. Gross, J. Math. Phys. 7, 578 (1966).
[5] T. D. Lee, K. Huang, and C. N. Yang, Phys. Rev. 106, 1135 (1957).
[6] N. M. Witriol, Phys. Rev. A 1, 1775 (1970).
[7] K. Huang and C. N. Yang, Phys. Rev. 105, 767 (1957).
[8] T. D. Lee and C. N. Yang, Phys. Rev. 105, 1119 (1957).
[9] A. A. Abrikosov and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. 33, 1154 (1957) [Sov. Phys. JETP 6, 888 (1958)].
[10] M. J. Cooper, J. Math. Phys. 9, 571 (1968).
[11] Notationally we use u for the shrinkage functions instead of α as was used in I.

[12] See, for example, G. B. Thomas, *Calculus and Analytic Geometry* (Addison-Wesley, Reading, Mass., 1953), Chap. 12.

[13] We note that the transformation Eq. (3.1) is not one to one over all R, r space with the function u given by Eq. (3.12) with Eqs. (3, 25) and (4, 3). To obtain a one-to-one transformation we would use the transformation with e given by Eq. (2.23) instead of Eq. (4, 3). To eliminate the transformed hard-core potential we would let $\epsilon'' \to 0$ at the end of the calculations. No inconsistencies arise if we let $\epsilon'' = 0$ at the beginning of the calculations.

[14] N. M. Witriol, J. Math. Phys. 12, 177 (1971).

[15] Thomas Muir, A Treatise on the Theory of Determinants (Dover, New York, 1960), Theorem 104, p. 89.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

One-Particle Operators and Local Internal Symmetries*†

Ira W. Herbst
Lawrence Radiation Laboratory, University of California, Berkeley, California 94720
(Received 18 May 1971)

The structure of local operators linear in the creation and destruction operators of a finite number of particles is investigated. It is shown that these operators determine a set of local and relatively local free fields relative to which the original operators are local. This result is used to study local internal symmetries in a theory of interactions and to show that these symmetries commute with the Lorentz group. The assumptions of Haag and Ruelle which lead to a complete particle interpretation are also discussed. Asymptotic many-particle states are constructed under less restrictive assumptions.

I. INTRODUCTION

In this article we will first investigate the structure of a particularly simple local theory. In a Fock space of N free particles of mass $m > 0$ and arbitrary spins, we consider local Hermitian operators linear in the creation and destruction operators of these particles. These "one-particle operators" are associated with bounded regions of space–time in a way familiar from the local algebra theories of Haag and Araki, [1,2] with one important difference: The set of space–time regions to which the one-particle operators are localized need not include regions of arbitrarily small size. (For a precise description of this set of regions, see Theorem 1 of Sec. III.) We show that a set of one-particle operators determines a standard set of relatively local free fields relative to which the original one-particle operators are local. This is the subject of Sec. III.

In Sec. IV we use this result to analyze local internal symmetries of a type considered by Landau and Wichmann. [3,4] These authors have demonstrated the translational invariance of such symmetries, and Landau has further shown that, in a *field* theory with an interpolating field[5] for each particle, these symmetries must commute with the Lorentz group. In this section we extend this result to the apparently more general local algebra theories and field theories for which a complete set of interpolating fields is not assumed. We also discuss the relationship of the Haag–Ruelle theory of asymptotic states to selection rules. Asymptotic many-particle states are constructed under less restrictive assumptions.

To establish notation, we first review the Poincaré transformation properties of one-particle states.

II. ONE-PARTICLE REPRESENTATIONS OF THE POINCARÉ GROUP

The representation $\Gamma_{m,s}$ of the Poincaré group (denoted $\bar{\mathcal{G}}$) appropriate to a particle of mass $m > 0$ and spin s can be realized by unitary operators $U(x,g) = U(x)\,U(g)$ [x a 4-vector, g in $\mathcal{G} \equiv SL(2, C)$] on a Hilbert space of $(2s+1)$-component functions $f(\mathbf{p})$ in the following way:

$$(U(x,g)f)(\mathbf{p}) = e^{ix\cdot p} D^s(g) f(\Lambda^{-1}(g)\mathbf{p})$$
$$x\cdot p = t\omega - \mathbf{x}\cdot\mathbf{p},$$
$$\omega(\mathbf{p}) = (\mathbf{p}^2 + m^2)^{1/2}. \tag{1}$$

Here $D^s(g)$ is the standard $(2s+1)$-dimensional representation $(0, s)$ of $\mathcal{G}$ and the notation $\Lambda(g)\mathbf{p}$ means the spatial part of $\Lambda(g)p$, where $\Lambda(g)$ is the 4×4 Lorentz matrix. The invariant scalar product is

$$\langle f|h\rangle \equiv \int \frac{d^3\mathbf{p}}{2\omega} B(f,h;\mathbf{p}),$$

where $B(f,h;\mathbf{p}) = f^\dagger(\mathbf{p}) D^s(\tilde{p}) h(\mathbf{p})$, $\tilde{p} = [\omega(\mathbf{p}) - \mathbf{p}\cdot\sigma]/m$, and the σ_i are Pauli matrices. Note that since $D^s(\tilde{p})$ is a positive-definite Hermitian matrix, for fixed $\mathbf{p}$ $B(f,h;\mathbf{p})$ is a scalar product of the $(2s+1)$-component vectors $f(\mathbf{p})$ and $h(\mathbf{p})$. We also mention here that $D^s(g)$ is a homogeneous polynomial of degree $2s$ in the matrix elements of g and that therefore $D^s(\tilde{p})$ is a polynomial in the components of $p = (\mathbf{p}, \omega)$.

We will, in addition, make use of the infinitesimal operators of the representation $U(x,g)$. A simple calculation gives the following results:

(a) The generators of rotations and velocity transformations are given by $\mathbf{J} = \mathbf{S} + \mathbf{L}$, $\mathbf{K} = i(\mathbf{S} - \omega\nabla)$ respectively. Here $\mathbf{S}$ is the standard $(2s+1)$-dimensional angular momentum matrix and $\mathbf{L} = -i\mathbf{p} \times \nabla$. Note that $\mathbf{S}$ is not Hermitian relative to the scalar product $\langle f|g\rangle$ and hence is not the spin observable.

(b) The operators $\mathbf{S}$ and $\mathbf{L}$ are expressible in terms of the generators of Poincare transformations. The formula is

$$m^2\mathbf{L} = \mathbf{p} \times [\mathbf{p} \times (\mathbf{J} + i\mathbf{K})] - i\omega\mathbf{p} \times (\mathbf{J} + i\mathbf{K}). \tag{2}$$

[12] See, for example, G. B. Thomas, *Calculus and Analytic Geometry* (Addison-Wesley, Reading, Mass., 1953), Chap. 12.

[13] We note that the transformation Eq. (3.1) is not one to one over all R, r space with the function u given by Eq. (3.12) with Eqs. (3, 25) and (4, 3). To obtain a one-to-one transformation we would use the transformation with e given by Eq. (2.23) instead of Eq. (4, 3). To eliminate the transformed hard-core potential we would let $\epsilon'' \to 0$ at the end of the calculations. No inconsistencies arise if we let $\epsilon'' = 0$ at the beginning of the calculations.

[14] N. M. Witriol, J. Math. Phys. 12, 177 (1971).

[15] Thomas Muir, A Treatise on the Theory of Determinants (Dover, New York, 1960), Theorem 104, p. 89.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

One-Particle Operators and Local Internal Symmetries*†

Ira W. Herbst
Lawrence Radiation Laboratory, University of California, Berkeley, California 94720
(Received 18 May 1971)

The structure of local operators linear in the creation and destruction operators of a finite number of particles is investigated. It is shown that these operators determine a set of local and relatively local free fields relative to which the original operators are local. This result is used to study local internal symmetries in a theory of interactions and to show that these symmetries commute with the Lorentz group. The assumptions of Haag and Ruelle which lead to a complete particle interpretation are also discussed. Asymptotic many-particle states are constructed under less restrictive assumptions.

I. INTRODUCTION

In this article we will first investigate the structure of a particularly simple local theory. In a Fock space of N free particles of mass $m > 0$ and arbitrary spins, we consider local Hermitian operators linear in the creation and destruction operators of these particles. These "one-particle operators" are associated with bounded regions of space–time in a way familiar from the local algebra theories of Haag and Araki, [1,2] with one important difference: The set of space–time regions to which the one-particle operators are localized need not include regions of arbitrarily small size. (For a precise description of this set of regions, see Theorem 1 of Sec. III.) We show that a set of one-particle operators determines a standard set of relatively local free fields relative to which the original one-particle operators are local. This is the subject of Sec. III.

In Sec. IV we use this result to analyze local internal symmetries of a type considered by Landau and Wichmann. [3,4] These authors have demonstrated the translational invariance of such symmetries, and Landau has further shown that, in a *field* theory with an interpolating field[5] for each particle, these symmetries must commute with the Lorentz group. In this section we extend this result to the apparently more general local algebra theories and field theories for which a complete set of interpolating fields is not assumed. We also discuss the relationship of the Haag–Ruelle theory of asymptotic states to selection rules. Asymptotic many-particle states are constructed under less restrictive assumptions.

To establish notation, we first review the Poincaré transformation properties of one-particle states.

II. ONE-PARTICLE REPRESENTATIONS OF THE POINCARÉ GROUP

The representation $\Gamma_{m,s}$ of the Poincaré group (denoted $\bar{\mathcal{G}}$) appropriate to a particle of mass $m > 0$ and spin s can be realized by unitary operators $U(x,g) = U(x)\,U(g)$ [x a 4-vector, g in $\mathcal{G} \equiv SL(2, C)$] on a Hilbert space of $(2s+1)$-component functions $f(\mathbf{p})$ in the following way:

$$(U(x,g)f)(\mathbf{p}) = e^{ix\cdot p} D^s(g) f(\Lambda^{-1}(g)\mathbf{p})$$
$$x\cdot p = t\omega - \mathbf{x}\cdot\mathbf{p},$$
$$\omega(\mathbf{p}) = (\mathbf{p}^2 + m^2)^{1/2}. \tag{1}$$

Here $D^s(g)$ is the standard $(2s+1)$-dimensional representation $(0, s)$ of $\mathcal{G}$ and the notation $\Lambda(g)\mathbf{p}$ means the spatial part of $\Lambda(g)p$, where $\Lambda(g)$ is the 4×4 Lorentz matrix. The invariant scalar product is

$$\langle f|h\rangle \equiv \int \frac{d^3\mathbf{p}}{2\omega} B(f,h;\mathbf{p}),$$

where $B(f,h;\mathbf{p}) = f^\dagger(\mathbf{p}) D^s(\tilde{p}) h(\mathbf{p})$, $\tilde{p} = [\omega(\mathbf{p}) - \mathbf{p}\cdot\sigma]/m$, and the σ_i are Pauli matrices. Note that since $D^s(\tilde{p})$ is a positive-definite Hermitian matrix, for fixed $\mathbf{p}$ $B(f,h;\mathbf{p})$ is a scalar product of the $(2s+1)$-component vectors $f(\mathbf{p})$ and $h(\mathbf{p})$. We also mention here that $D^s(g)$ is a homogeneous polynomial of degree $2s$ in the matrix elements of g and that therefore $D^s(\tilde{p})$ is a polynomial in the components of $p = (\mathbf{p}, \omega)$.

We will, in addition, make use of the infinitesimal operators of the representation $U(x,g)$. A simple calculation gives the following results:

(a) The generators of rotations and velocity transformations are given by $\mathbf{J} = \mathbf{S} + \mathbf{L}$, $\mathbf{K} = i(\mathbf{S} - \omega\nabla)$ respectively. Here $\mathbf{S}$ is the standard $(2s+1)$-dimensional angular momentum matrix and $\mathbf{L} = -i\mathbf{p} \times \nabla$. Note that $\mathbf{S}$ is not Hermitian relative to the scalar product $\langle f|g\rangle$ and hence is not the spin observable.

(b) The operators $\mathbf{S}$ and $\mathbf{L}$ are expressible in terms of the generators of Poincare transformations. The formula is

$$m^2\mathbf{L} = \mathbf{p} \times [\mathbf{p} \times (\mathbf{J} + i\mathbf{K})] - i\omega\mathbf{p} \times (\mathbf{J} + i\mathbf{K}). \tag{2}$$

Finally we will make use of the creation and destruction operators appropriate to a Fock space description of this particle. Hence we introduce the operators $a_\mu^\dagger(\mathbf{p})$ such that if $f(\mathbf{p})$ is the $(2s+1)$-component function transforming according to Eq. 1, then

$$|f\rangle = \sum_{\mu=-s}^{s} \int \frac{d^3\mathbf{p}}{2\omega} f_\mu(\mathbf{p}) a_\mu^\dagger(\mathbf{p})|0\rangle$$

is the one-particle Fock space state it represents. Here $|0\rangle$ is the vacuum vector. It is then easy to verify the transformation law

$$U(x,g)a_\mu^\dagger(\mathbf{p})U^{-1}(x,g) = \sum_{\sigma=-s}^{s} e^{ix\cdot\Lambda(g)p} a_\sigma^\dagger(\Lambda(g)\mathbf{p}) D^s_{\sigma\mu}(g). \tag{3}$$

If the particle satisfies either Fermi or Bose statistics, the creation and destruction operators must satisfy the usual anticommutation or commutation relations

$$[a_\mu(\mathbf{p}), a_\nu^\dagger(\mathbf{q})]_\pm = 2\omega\delta(\mathbf{p}-\mathbf{q})D^s_{\mu\nu}(\tilde{p}),$$

respectively.

We introduce the free field[6] $\phi_\mu(x)$ transforming according to $D^s(g)$ as follows:

$$\phi_\mu(x) = \int \frac{d^3\mathbf{p}}{2\omega}[a_\mu^\dagger(\mathbf{p})e^{ix\cdot p} + \lambda\theta a_\mu(\mathbf{p})\theta^{-1}e^{-ix\cdot p}]. \tag{4}$$

Here θ is the TCP operator given by

$$\theta a_\mu(\mathbf{p})\theta^{-1} = \sum_\sigma a_\sigma(\mathbf{p})[C_s D^s(\tilde{p})]_{\sigma\mu}\lambda^*,$$

where $C_s = \exp(-i\pi S_2)$ and λ is a complex number of modulus unity. If θ is to act locally on $\phi_\mu(x)$, we must choose $\lambda^2 = (-1)^{2s}$. Note that θ has the following action on the one-particle wavefunctions,

$$(\theta f)(\mathbf{p}) = [C_s D^s(\tilde{p})f(\mathbf{p})]^*,$$

and that

$$B(\theta f, \theta h; \mathbf{p}) = B(h, f; \mathbf{p}), \tag{5a}$$

$$[U(g), \theta] = 0. \tag{5b}$$

Since $C_s D^s(g)^* C_s^{-1} = D^s(g^{-1})\dagger$, $D^{s\dagger}(g)D^s(\tilde{p})D^s(g) = D^s(\tilde{p}')$, with $p' = \Lambda(g^{-1})p$, we find that

$$U(g)\phi_\mu(x)U^{-1}(g) = \sum_\lambda \phi_\lambda(\Lambda(g)x)D_{\lambda\mu}{}^s(g)$$

and

$$[\phi_\mu(x), \phi_\lambda(y)]_\pm = i(C_s^{-1})_{\mu\lambda}\Delta_0(x-y;m).$$

III. RELATIVELY LOCAL SETS OF ONE-PARTICLE OPERATORS

We consider a Fock space of N particles with a common mass $m > 0$ and arbitrary spins. Each particle is assumed to be either a fermion or a boson, and normal commutation relations are assumed for the creation and destruction operators of different particles.

Suppose $\mathfrak{D}$ is a Poincare invariant class of bounded open space–time regions. That is, if $D \in \mathfrak{D}$ then $[\Lambda(g)D + x_0] \in \mathfrak{D}$, where $\Lambda(g)D + x_0$ is the region $\{x | x = \Lambda(g)x' + x_0, x' \in D\}$. With each $D \in \mathfrak{D}$ we associate a set $S(D)$ of Hermitian one-particle operators each of which is linear in a subset of the N creation and destruction operators. Each operator in $S(D)$ is assumed to involve either fermions alone or bosons alone defining the sets $S_+(D)$ and $S_-(D)$, respectively; $S(D) = S_+(D) \cup S_-(D)$. To insure the completeness of the description of the theory in terms of the $S(D)$ we assume that the linear span of $\{\cup_{D\in\mathfrak{D}} S(D)\}|0\rangle$ is dense in the one-particle Hilbert space. The operators of $S(D)$ are "localized" to the region D in the following sense:

1. $U(x,g)S(D)U^{-1}(x,g) = S(\Lambda(g)D + x)$;

2. If $D_1 \subseteq D_2'$, then $[S_\pm(D_1), S_\pm(D_2)]_\pm = 0$.

(Here D' is the region spacelike to the closure of D.) To simplify the geometrical aspects of our study, we will assume in what follows that all regions $D \in \mathfrak{D}$ are of a particular kind called "diamonds." A diamond with vertices x_1 and x_2 (where $x_2 - x_1$ is forward timelike) is the intersection of the open backward light cone from x_2 with the open forward cone from x_1. Note that $\mathfrak{D}$ may consist of the Poincaré transforms of only a single diamond.

The assumptions stated above will allow us to prove the following:

Theorem 1: There exists a set of N free fields $\phi^k(x)$, $k = 1, \ldots, N$, of the type discussed in Sec. II [ϕ^k transforms according to $D^s(g)$, where s depends on k] which satisfy

$$[\phi_\mu^k(x), \phi_\nu^l(y)]_\pm = i\delta_{kl}(C_s^{-1})_{\mu\nu}\Delta_0(x-y;m)$$

and which are local relative to the $S(D)$. That is, if $Q \in S_\pm(D)$ and $x \in D'$, then $[Q, \phi_\mu^k(x)]_- = 0$ if at least one of Q and ϕ^k are Bose operators and $[Q, \phi_\mu^k(x)]_+ = 0$ if both are Fermi operators.

As we shall see later the above set of fields is essentially unique. Furthermore, each operator of $S(D)$ is essentially a sum of fields "smeared" with functions with support in D. Hence the structure of the operators of $S(D)$ is determined by that of a free field theory.

In proving Theorem 1 we will want to integrate and differentiate quantities such as $U(x,g)QU^{-1}(x,g)$ with $Q \in S(D)$. While only the commutation properties of these operators which follow from assumptions 1 and 2 above will be used, for notational convenience we will say that the integrated or differentiated operators are members of the appropriate $S(D)$. A series of three lemmas will now be given which result in Theorem 1.

We first show that only one spin at a time need be considered.

Lemma 1: Suppose $Q \in S(D)$ with $U(x,g)Q|0\rangle$ C^∞ in (x,g). Then each Hermitian operator Q_k which occurs in the decomposition of Q into a sum of operators involving particles of spin $j(k)$ is also in $S(D)$.

Proof: We introduce the polarization 4-vector operator $W^k = (1/2m)\sum \epsilon^{klpq} P_l J_{pq}$. (Here P_l and J_{pq} are the usual generators of translations and Lorentz transformation.) The operator $W^2 \equiv \sum W_k W^k$ has the property that for a one-particle state ψ of spin j and mass m, $[W^2 + j(j+1)]\psi = 0$. The action of W^2 on Q is defined by the appropriate repeated differentiations of $U(x,g)QU^{-1}(x,g)$. For example,

$$(P_l J_{pq})(Q) = \left(i\frac{d}{d\lambda}\; i\frac{d}{d\lambda'} U(x,g)QU^{-1}(x,g)\right)_{\lambda'=\lambda=0},$$

where $U(x,g) = \exp(-i\lambda P_l)\exp(-i\lambda' J_{pq})$. Since this operation is local, i.e., for $R \in S(D_1)$ and $D_1 \subseteq D'$ $[W^2(Q), R]_\pm = 0$, we have $W^2(Q) \in S(D)$. With $Q = \sum_{k=1}^m Q_k$ and $\alpha_k = j(k)[j(k)+1]$, we note that

$$Q_i = \left(\prod_{k=1}^{m}{}' \frac{W^2 + \alpha_k}{\alpha_k - \alpha_i}\right)(Q),$$

where the prime sign indicates that the factor containing the index i is to be omitted. Hence $Q_i \in S(D)$.

We now consider one of the Hermitian operators Q_k in $S(D)$ (hereafter called Q) involving only particles of spin s. Poincaré invariance allows us to transform D to the special form $D_r = \{x \mid \|\mathbf{x}\| + |t| < r\}$ for some r. The operator Q can be written

$$Q = \sum_{k=1}^{n} \sum_{\mu=-s}^{s} \int \frac{d^3\mathbf{p}}{2\omega} [a_\mu^\dagger(k,\mathbf{p}) f_\mu(k,\mathbf{p}) + a_\mu(k,\mathbf{p}) f_\mu^*(k,\mathbf{p})].$$

Here $a_\mu^\dagger(k,\mathbf{p})$ is a creation operator for particle k as discussed in Sec. II and $[a_\mu(k,\mathbf{p}), a_\nu^\dagger(l,\mathbf{q})]_\pm = 0$ for $k \neq l$. We use the notation $f_i(\mathbf{p}) = \{f_{-s}(i,\mathbf{p}), \cdots, f_s(i,\mathbf{p})\}$, $f(\mathbf{p}) = \{f_1(\mathbf{p}), \cdots f_n(\mathbf{p})\}$, and $D(g)$ for the $[(2s+1)\times n]$-dimensional matrix whose elements consist of n matrices $D^s(g)$ on the diagonal. Thus $f_i(\mathbf{p})$ and $f(\mathbf{p})$ are respectively $(2s+1)$- and $[(2s+1)\times n]$-component column vectors. Without loss of generality, we will assume that the $f_i(\mathbf{p})$ are linearly independent vector-valued functions of $\mathbf{p}$. [If there are $m < n$ linearly independent $f_i(\mathbf{p})$, we "rename" the particles by choosing $b_\mu^\dagger(k,\mathbf{p}) = \sum_{l=1}^n a_\mu^\dagger(l,\mathbf{p})\beta_{lk}$ with unitary β. For a suitable choice of β the new wavefunctions $f_i'(\mathbf{p}) = \sum_{l=1}^n \beta_{il}^{-1} f_l(\mathbf{p})$ will satisfy $f_i'(\mathbf{p}) \equiv 0$ for $i > m$, the remainder being linearly independent.] Note that by virtue of the differentiability assumption in Lemma 1, the $f_\mu(k,\mathbf{p})$ are in the function class $\mathcal{S}$.

We want to establish the result that $f(\mathbf{p})$ is the restriction of an entire function $\hat{f}(p)$ of the 4-vector p to $p_0 = \omega(\mathbf{p})$ such that the $\mathbf{x}$-space Fourier transforms of $\hat{f}(\mathbf{p},\omega) + \hat{f}(\mathbf{p},-\omega)$ and of $[\hat{f}(\mathbf{p},\omega) - \hat{f}(\mathbf{p},-\omega)]/\omega$ have support in $\|\mathbf{x}\| \leq r$. These facts and certain other results of an algebraic nature will be needed to establish Theorem 1.

Hence consider the function

$$\mathfrak{F}(x;g_1,g_2) \equiv [U(x,g_1)QU^{-1}(x,g_1), U(g_2)QU^{-1}(g_2)]_\pm,$$

which by assumption has compact support in $\mathbf{x}$ at $t = 0$ along with all its derivatives with repect to x and g_1. With F_1 and F_2, respectively, the Fourier transforms of $[i\partial_t \mathfrak{F}(x;g_1,g_2)]_{t=0}$ and $[\mathfrak{F}(x;g_1,g_2)]_{t=0}$ we calculate

$$B(U(g_1)f, U(g_2)f;\mathbf{p}) = F_1(\mathbf{p};g_1,g_2) + \omega F_2(\mathbf{p};g_1,g_2) \equiv A(\mathbf{p};g_1,g_2), \quad (6)$$

where

$$B(U(g_1)f, U(g_2)f;\mathbf{p}) \equiv \sum_{k=1}^{n} B(U(g_1)f_k, U(g_2)f_k;\mathbf{p}).$$

Note that $A(\mathbf{p};g_1,g_2)$ has the following two properties:

1. It is the restriction to $p_0 = \omega(\mathbf{p})$ of an entire function $\hat{A}(p;g_1,g_2)$ of the 4-vector p; for example, $\hat{A}(p;g_1,g_2) = F_1(\mathbf{p};g_1,g_2) + p_0 F_2(\mathbf{p};g_1,g_2)$. This is also true of all derivatives of $A(\mathbf{p};g_1,g_2)$ with respect to g_1. [For brevity, functions of the 3-vector $\mathbf{p}$ which are restrictions to $p_0 = \omega(\mathbf{p})$ of entire functions of the 4-vector p will sometimes be called entire functions of p.]

2. For real p and $p_0 = \omega$, $\hat{A}$ satisfies the reality condition $\hat{A}^*(p;g_1,g_2) = \mp \hat{A}(-p;g_1,g_2)$. These properties of A can be used to prove the following:

Lemma 2: (a) The function $f(\mathbf{p}) = \hat{f}(\mathbf{p},\omega)$, where $\hat{f}(p)$ is an entire function of p.

(b) A suitable renaming of the n particles can be carried out to give for the n new wavefunctions [again denoted $f_i(\mathbf{p})$] and for real p: $f^\theta(\mathbf{p}) = e^{i\pi s}\hat{f}(-\mathbf{p},-\omega)$. Here the abbreviation $f^\theta(\mathbf{p}) \equiv e^{-i\pi s}[CD(\tilde{p})f(\mathbf{p})]^*$ has been introduced.

The proof of this lemma has essentially been carried out by Epstein.[7] We present a somewhat different proof:

Let $\mathfrak{A}$ be the algebra of all polynomials in the representatives of the generators of the Poincaré group. For each $\mathbf{p}$ let $\mathfrak{V}(\mathbf{p})$ be that subspace of $[(2s+1)\times n]$-dimensional complex space equal to $\{\mathfrak{A}f(\mathbf{p})\}$ [i.e., $\mathfrak{V}(\mathbf{p}) = \{w \mid w = (Rf)(\mathbf{p}), R \in \mathfrak{A}\}$]. Suppose $M = \max_{\mathbf{p}}$ [dimension of $\mathfrak{V}(\mathbf{p})$]. We note the following important facts:

1. If $R_1, R_2 \in \mathfrak{A}$ then $B(R_1 f, R_2 f;\mathbf{p})$ is an entire function of $\tilde{p}$.

2. We can choose $R_1, \ldots, R_M \in \mathfrak{A}$ with $R_1 = 1$ such that the Gram determinant $G(\mathbf{p}) = \det[B(R_i f, R_j f;\mathbf{p})]$ does not vanish identically. [Note also that $G(\mathbf{p}) \geq 0$ for real $\mathbf{p}$.]

3. Equation (2) implies that $\omega\nabla \in \mathfrak{A}$. Statements 1, 2, and 3 then imply that, for $\mathbf{p}$ real and $G(\mathbf{p}) \neq 0$,

$$\frac{\partial (R_i f)(\mathbf{p})}{\partial p_k} = \sum_{j=1}^{M} (R_j f)(\mathbf{p}) \alpha_{ji}^k(\mathbf{p}). \quad (7)$$

Here $\alpha^k_{jl}(\mathbf{p})$ can be calculated by Cramer's rule in terms of the inner products $B(\omega\nabla_k R_i f, R_j f; \mathbf{p})$ and $B(R_i f, R_j f; \mathbf{p})$ and is seen to equal $[\omega G(\mathbf{p})]^{-1} \times$ (an entire function of p).

Substituting $\partial^2(R_i f)(\mathbf{p})/\partial p_l \partial p_k = \partial^2(R_i f)(\mathbf{p})/\partial p_k \partial p_l$ and noting the linear independence of the column vectors $(R_i f)(\mathbf{p})$ at real points $\mathbf{p}$ for which $G(\mathbf{p}) > 0$, we derive the integrability conditions

$$\frac{\partial \alpha^k(\mathbf{p})}{\partial p_l} - \frac{\partial \alpha^l(\mathbf{p})}{\partial p_k} = [\alpha^k(\mathbf{p}), \alpha^l(\mathbf{p})] \tag{8}$$

for all points $\mathbf{p} \in C^3$ for which $\omega(\mathbf{p})G(\mathbf{p}) \neq 0$. The matrices $\alpha^k(\mathbf{p})$ are analytic at least at these points.

The analyticity properties of the solutions of differential equations such as Eq. (7) (total differential equations) with coefficients satisfying (8) are well known[8,9]: In any simply connected neighborhood (real or complex) of a point $\mathbf{z}_0$ where the α^k are analytic, the equation has a unique solution (with given initial conditions). This solution is analytic in this neighborhood and has an analytic continuation along any path of analyticity of the coefficients.

To determine additional analyticity properties of $f(\mathbf{p})$, we will make use of finite Lorentz transformations. Choose a polycylinder C_0, centered at $\mathbf{p}_0 = \mathbf{p}_0^*$, in which $\omega(\mathbf{z})G(\mathbf{z}) \neq 0$. We solve the DE there to get a function $F(\mathbf{z})$ analytic in C_0 and equal to $f(\mathbf{p})$ in R_0, the intersection of C_0 with the real space. Let us assume that the particle wavefunctions $f_1(\mathbf{p}), \ldots, f_n(\mathbf{p})$ are linearly independent (as vector-valued functions of $\mathbf{p}$) in R_0. [If this is not so, we arrange a renaming of the particles so that there are n', $n' < n$, linearly independent $f_i(\mathbf{p})$ in R_0, the remainder being zero in R_0.] Since $F(\mathbf{p}) = f(\mathbf{p})$ for $\mathbf{p} \in R_0$, the equation

$$B(U(g)F, F; \mathbf{p}) = A(\mathbf{p}; g, 1) \equiv A(\mathbf{p}; g) \tag{6'}$$

holds for all g in some real neighborhood $N(1)$ of the identity and all $\mathbf{p}$ in some real neighborhood R_1 of $\mathbf{p}_0$.

We will now use Eq. (6′) to show that in fact $F(\mathbf{z})$ can be analytically continued along any path $\mathbf{z}(t)$ which begins at $\mathbf{p}_0$ and does not pass through the surface $\omega(\mathbf{z}) = 0$. Hence suppose $F(\mathbf{z})$ is analytic at all points $\mathbf{z}(t)$, $0 \leq t < t_1$. We will show that $F(\mathbf{z})$ is also analytic at $\mathbf{z}(t_1)$. We choose $g_0 \in N(1)$ and a neighborhood $N(g_0)$ of g_0 with $N(g_0) \subseteq N(1)$ such that

(a) $|\omega(\Lambda^{-1}(g)\mathbf{z}(t))| > 0$ for $0 \leq t \leq t_1, g \in N(g_0)$,

(b) $|G(\Lambda^{-1}(g)\mathbf{z}(t))| > 0$ for $t_1 - \epsilon \leq t \leq t_1$,

$g \in N(g_0)$ with an $\epsilon > 0$ which depends on g_0 and $N(g_0)$.

Such a choice is possible because the condition $G(\Lambda^{-1}(g)\mathbf{z}(t_1)) = 0$ for all g in a real neighborhood of the identity implies that G vanishes in a complex neighborhood of the identity and therefore that $G(\mathbf{z})$ vanishes in a complex neighborhood of $\mathbf{z}(t_1)$. [Note that $\Lambda^{-1}\mathbf{z}$ is the spatial component of $\Lambda^{-1}(\mathbf{z}, \omega(\mathbf{z}))$, where the value of $\omega(\mathbf{z})$ depends on the path taken to the point $\mathbf{z}$.] Let $J(\mathbf{z}; g_1, \ldots, g_\beta)$, $\beta = (2s+1) \times n$, be the matrix whose rows are the vectors $v_i \equiv [D(g_i)F(\Lambda^{-1}(g_i)\mathbf{z}^*)]^*$. Let us now choose a set of β elements $g_i \in N(g_0)$ such that $\det J(\mathbf{z}; g_1, \ldots, g_\beta) \neq 0$ for $\mathbf{z}$ in a neighborhood of $\mathbf{z}(t_1)$. This choice is possible for the following reason: The vanishing of $\det J(\mathbf{z}(t_1); g_1, \ldots, g_\beta)$ for all $g_i \in N(g_0)$ is equivalent to the statement that the v_i do not span the β-dimensional space no matter how the g_i are chosen in $N(g_0)$; i.e., there is a nonzero vector v orthogonal to $v(g) \equiv [D(g)F([\Lambda^{-1}(g)\mathbf{z}(t_1)]^*)]^*$ for all $g \in N(g_0)$. But $v^\dagger v(g)$ is therefore zero in a complex neighborhood N_0 of g_0. We choose $u \in SU(2)$ and g_1 such that $\Lambda^{-1}(g_1)\mathbf{z}(t_1) = \mathbf{0}$. Then, if $g^{-1} = h^{-1}g_1 u g_1^{-1}$, $h \in N_0$, we have $v^\dagger v(g) = 0$ for all $u \in SU(2)$. A short calculation then shows the $F_i^*(\mathbf{z}^*)$ to be linearly dependent in a neighborhood of $\Lambda^{-1}(g_0)\mathbf{z}(t_1)$. Continuing back to R_0 gives the linear dependence of the $f_i(\mathbf{p})$ in R_0, contradicting our assumption that they are linearly *independent* in R_0.

We will now use Eq. (6′) to show that $F(\mathbf{z})$ is analytic in a neighborhood of $\mathbf{z}(t_1)$. We choose a path $\mathbf{w}(t)$ such that (a) $\mathbf{w}(t)$ coincides with $\mathbf{p}$ and $\mathbf{z}(t_1 - \epsilon)$ at its end points, (b) $\mathbf{w}(t)$ is close enough to $\mathbf{z}(t)$ for $0 \leq t \leq t_1 - \epsilon$ so that it can be continuously distorted to $\mathbf{z}(t)$ without crossing any singularities of $F(\mathbf{z})$, and (c) for $0 \leq t \leq t_1 - \epsilon$, $|G(\Lambda^{-1}(g_i)\mathbf{w}(t))| > 0$, $i = 1, \ldots, \beta$. With these three conditions met, we can continue the equations

$$[D(g_i)F([\Lambda^{-1}(g_i)\mathbf{z}]^*)]^\dagger D(\tilde{z})F(\mathbf{z}) = A(\mathbf{z}; g_i) \tag{6''}$$

along $\mathbf{w}(t)$ to $\mathbf{z}(t_1 - \epsilon)$, resulting in the same $F(\mathbf{z})$ as would have been obtained by continuation along $\mathbf{z}(t)$. [Note that conditions (b) and (c) can easily be met, for example, in the following way: First choose a polygonal path which approximates $\mathbf{z}(t)$ well enough to satisfy (a) and (b) and whose component line segments neither begin nor end at zeroes of $G(\Lambda^{-1}(g_1)\mathbf{z})$ for any i. If for each straight path segment $\mathbf{w}_k(t)$ we choose a linear parametrization, $G(\Lambda^{-1}(g_i)\mathbf{w}_k(t))$ is analytic in t and hence $\mathbf{w}_k(t)$ can be distorted infinitesimally in the t complex plane to miss the zeroes of $G(\Lambda^{-1}(g_i)\mathbf{z})$, $i = 1, \ldots, \beta$. The new path will satisfy (a), (b), and (c).] Then the Eqs. (6″) can be continued further from $\mathbf{z}(t_1 - \epsilon)$ along $\mathbf{z}(t)$. If $A(\mathbf{z})$ is the column vector whose components are $A(\mathbf{z}; g_i)$, we thus obtain the representation

$$F(\mathbf{z}) = [J(\mathbf{z}; g_1, \ldots, g_\beta)D(\tilde{z})]^{-1}A(\mathbf{z}) \tag{9}$$

in a neighborhood of $\mathbf{z}(t_1)$. This shows explicitly that $F(\mathbf{z})$ is analytic in this neighborhood. By repeated application of this argument we see that no singularities can appear on the path $\mathbf{z}(t)$. The monodromy theorem then shows $F(\mathbf{z})$ to be analytic in any simply connected domain not containing points of the surface $\omega(\mathbf{z}) = 0$.

We now go on to examine $F(\mathbf{z})$ in a neighborhood of

a point $\mathbf{z}_0$ on this surface. We can easily choose $g_i \in N(1)$, $i = 1, \ldots, \beta$, so that $\omega(\Lambda^{-1}(g_i)\mathbf{z}_0) \neq 0$, $\det J(\mathbf{z}_0; g_1, \cdots, g_\beta) \neq 0$, and so that the $\Lambda^{-1}(g_i)\mathbf{z}_0$ lie in a neighborhood of a single point $\mathbf{z}_1$ containing no points of the surface $\omega(\mathbf{z}) = 0$. Note that $J(\mathbf{z}; g_1, \ldots, g_\beta)$ can be written as $J_1(\mathbf{z}; g_1, \ldots, g_\beta) + \omega(\mathbf{z})J_2(\mathbf{z}; g_1, \ldots, g_\beta)$, where the $J_i(\mathbf{z})$ are analytic in a neighborhood of $\mathbf{z}_0$. Again Eq. (9) gives an explicit representation for $F(\mathbf{z})$ in a neighborhood of $\mathbf{z}_0$. This representation shows two things. Firstly, since all curves $\mathbf{z}(t)$ which encircle $-m^2$ twice in the $\mathbf{z}^2$ plane can be continuously distorted to lie arbitrarily close to $\mathbf{z}_0$ without crossing $\omega(\mathbf{z}) = 0$, we have $F(\mathbf{z}) = F_1(\mathbf{z}) + \omega(\mathbf{z})F_2(\mathbf{z})$, where $F_i(\mathbf{z})$ is single valued and analytic for $\omega(\mathbf{z}) \neq 0$. Secondly, Eq. (9) gives an explicit representation for the $F_i(\mathbf{z})$ in a neighborhood of $\mathbf{z}_0$, showing that they are analytic there. Since $\mathbf{z}_0$ is arbitrary, the $F_i(\mathbf{z})$ are in fact entire functions.

To complete part (a) of the lemma, it remains to show that the branch of $F(\mathbf{z})$ resulting from a continuation along a real path from $\mathbf{p}_0$ is in fact equal to $f(\mathbf{p})$ for all real $\mathbf{p}$. To accomplish this, we will demonstrate the equality on all real straight lines emanating from $\mathbf{p}_0$. Suppose $\mathbf{p}(t) = \mathbf{p}_0 + \mathbf{p}_1 t$ is such a line. Then along $\mathbf{p}(t)$ Eq. (7) reduces to an ordinary linear differential equation in t. The coefficients are analytic functions of t in a neighborhood of the real axis except at the zeroes of $G(\mathbf{p}(t))$, and hence $F(\mathbf{p}(t)) = f(\mathbf{p}(t))$ for $t \geq 0$ up to the nearest zero of $G(\mathbf{p}(t))$. We now consider the solution to the original DE [Eq. (7)] in a neighborhood of a point on the line between this and the next zero of $G(\mathbf{p}(t))$. This solution, $H(\mathbf{z})$, has the same analyticity properties as $F(\mathbf{z})$ and $H(\mathbf{p}(t)) = f(\mathbf{p}(t))$ between the first and second zeroes of $G(\mathbf{p}(t))$. But since $H(\mathbf{p}(t))$ and $F(\mathbf{p}(t))$ are analytic at the first zero and, since $f(\mathbf{p}(t))$ is infinitely differentiable, we have equality of all derivatives with respect to t of H and F at the first zero. Hence $H(\mathbf{p}(t)) = F(\mathbf{p}(t)) = f(\mathbf{p}(t))$ for $t \geq 0$ up to the second zero of $G(\mathbf{p}(t))$. Repeating this argument, we reach any point on the line in a finite number of steps. This completes the proof of part (a) of the lemma.

We now derive the connection between the two branches of the function $f(\mathbf{p}) = \hat{f}(\mathbf{p}, \omega)$ which is implied by $\hat{A}^*(p; g_1, g_2) = \mp \hat{A}(-p; g_1, g_2)$ for real $\mathbf{p}$ and $p_0 = \omega$. Let $\bar{f}(\mathbf{p}) = \hat{f}(-\mathbf{p}, -\omega)$ and $h(\mathbf{p}) = [CD(\tilde{p})\bar{f}(\mathbf{p})]^*$ for $\mathbf{p}$ real. Note that since $D(\tilde{p})$ is a homogeneous polynomial of degree $2s$ in the matrix elements of $\tilde{p} = [\omega(\mathbf{p}) - \mathbf{p}\cdot\sigma]/m$, we have $D(-\tilde{p}) = (-1)^{2s}D(\tilde{p})$. We now consider Eq. (6). We continue both sides around the branch point $\mathbf{z}^2 = -m^2$ in the $\mathbf{z}^2$ plane and back to the real domain, and then we replace the argument $\mathbf{p}$ by $-\mathbf{p}$. (These operations correspond to the variable transformation $p \to -p$.) We thus obtain

$$B(U(g_1)\bar{f}, U(g_2)\bar{f}; \mathbf{p}) = (-1)^{2s}\hat{A}(-\mathbf{p}, -\omega; g_1, g_2).$$

If we now take the complex conjugate of this equation and make use of the above property of A and the relations (5), we obtain the result

$$B(U(g_1)h, U(g_2)h; \mathbf{p}) = \mp(-1)^{2s}B(U(g_1)f, U(g_2)f; \mathbf{p}). \tag{10a}$$

Setting $g_i = 1$, we find that for f to be nonzero we must have the usual connection between spin and statistics.[10] Setting $g_2^{-1} = g^{-1}\nu(\mathbf{p})u\nu^{-1}(\mathbf{p})$ [$u \in SU(2)$ and $\nu(\mathbf{p})$ the boost to momentum $\mathbf{p}$] and noting the linear independence of the functions $D^s_{\mu\nu}(u)$, we conclude that

$$\sum_{i=1}^{n} h_i(\mathbf{p})h_i\dagger(\mathbf{q}) = \sum_{i=1}^{n} f_i(\mathbf{p})f_i\dagger(\mathbf{q}). \tag{10b}$$

Since the $f_i(\mathbf{p})$ are linearly independent, (10b) implies that in fact $f_i(\mathbf{p}) = \sum_{j=1}^{n}\gamma_{ij}h_j(\mathbf{p})$ with γ a unitary matrix. From this and from the definition of $h(\mathbf{p})$ we find that $\gamma^*\gamma = 1$ and hence that γ has the representation $\gamma = \beta\beta^T$ with β unitary. A renaming of the particles with the unitary transformation β leads to part (b) of the lemma for the new wavefunctions $f'(\mathbf{p}) = \beta^{-1}f(\mathbf{p})$.

In terms of these new wavefunctions [which we again call $f_i(\mathbf{p})$] and new creation operators [again called $a_\mu\dagger(k, \mathbf{p})$], Q can be written

$$\begin{aligned} Q &= \sum_{k=1}^{n}\sum_{\mu=-s}^{s}\int\frac{d^3\mathbf{p}}{2\omega}\{a_\mu\dagger(k,\mathbf{p})\hat{f}_\mu(k,p) \\ &\quad + a_\mu(k,\mathbf{p})[C_s D^s(\tilde{p})\hat{f}_k(-p)]_\mu\}_{p_0=\omega} \\ &= \sum_{k=1}^{n}\sum_{\mu=-s}^{s}\int d^3\mathbf{x}[\phi^k_\mu(x)\tilde{f}^1_\mu(k,\mathbf{x}) - \partial_t\phi^k_\mu(x)\tilde{f}^2_\mu(k,\mathbf{x})]_{t=0}, \end{aligned} \tag{11}$$

where here $f(\mathbf{p})$ has been expressed as $f^1(\mathbf{p}) - i\omega(\mathbf{p})f^2(\mathbf{p})$, with the $f^i(\mathbf{p})$ entire functions of $\mathbf{p}$, and where $\tilde{f}^i$ is the Fourier transform of f^i. The field $\phi^k(x)$ is given by Eq. (4) with $a^\dagger_\mu(\mathbf{p})$ replaced by $a^\dagger_\mu(k, \mathbf{p})$. Note that, by virtue of part (b) of Lemma 2, the $f^i(\mathbf{p})$ are in the function class $\mathcal{S}$ when considered as functions of the real variable $\mathbf{p}$. We shall next prove the following:

Lemma 3[11]: The functions $\tilde{f}^i_\mu(k, \mathbf{x})$ have support in $\|\mathbf{x}\| \leq r$, where r is the radius of the base of the diamond D_r.

This means that the operator Q commutes (or anticommutes, depending on s) with the fields $\phi^k(x)$, $k = 1, \ldots, n$, for x spacelike relative to the region D_r. Equation (11) shows that Q depends only on the field variables inside D_r. Note that because the fields satisfy the Klein–Gordon equation, Q is expressible in terms of the field variables at $t = 0$.

To prove Lemma 3, we first estimate

$$c_k(g_1, g_2) = \sup_x |D^k\mathfrak{F}(x; g_1, g_2)|$$

for certain (g_1, g_2). (Here D^k is a kth derivative with respect to x.) Let ν be a pure velocity transformation in $\mathcal{G}$. Suppose $\epsilon > 0$ and let $g \in \mathcal{G}$ be such that $\|\Lambda(g) - I\| < \epsilon/\sqrt{2}$. (Here $\|M\|^2 = \mathrm{Tr}\, M\dagger M$.) We set $g_1 = \nu g$, $g_2 = \nu^{-1}$, and $\gamma = \Lambda_{00}(\nu)$.

From the fact that $U(x)Q|0\rangle$ is C^∞ in x, we easily derive the inequality $c_k(g_1,g_2) < c_{Nk}(\epsilon)\gamma^{-N}$ for every N and every k. Via locality, a simple calculation shows that $[D^k\mathfrak{F}(x;g_1,g_2)]_{t=0}$ vanishes for $\|\mathbf{x}\| > 2\gamma r_\epsilon$, with $r_\epsilon = r(1+\epsilon)$, leading to estimates on the Fourier transforms $F_i(\mathbf{p};g_1,g_2)$:

$$\|\mathbf{z}\|^k |F_i(\mathbf{z};g_1,g_2)| < c'_{Nk}(\epsilon)\gamma^{-N}\exp(2\gamma r_\epsilon \|\mathrm{Im}\mathbf{z}\|) \tag{12}$$

uniformly in g whenever $\|\Lambda(g) - I\| < \epsilon/\sqrt{2}$.

Letting $\mathbf{e}$ be a real unit vector and $\mathbf{p}_0$ a real 3-vector perpendicular to $\mathbf{e}$, we define $f(z) \equiv f(z\mathbf{e}+\mathbf{p}_0)$. Denoting $\lambda = (\mathbf{p}_0^2 + m^2)^{1/2}$, we see that $f(z)$ is analytic on a two-sheeted Riemann surface with branch points at $\omega(z) \equiv (z^2+\lambda^2)^{1/2} = 0$. We cut the z plane along the imaginary axis from $i\lambda$ to $i\infty$ and from $-i\lambda$ to $-i\infty$. We first want to show that, for any nonnegative integer k, $\|z^k f(z)\| < c_k \exp(r_\epsilon|\mathrm{Im}z|)$. Hence we choose $\beta = (2s+1)\times n$ elements $g_i \in \mathcal{G}$ with $\|\Lambda(g_i) - I\| < \epsilon/\sqrt{2}$ so that $J(z) \equiv J(z\mathbf{e}+\mathbf{p}_0;g_1,\ldots,g_\beta)$ has a determinant which does not vanish identically as a function of z. We next choose two circles (one on each sheet) centered at $z=0$ with equal radii greater than λ and such that on these curves $\det J(z)$ has no zeroes. Let $\mathcal{C}$ denote the union of the two circles. We now consider velocity transformations ν in the $\mathbf{e}$ direction, and, since these do not change $\mathbf{p}_0$, we write $z(\nu)\mathbf{e}+\mathbf{p}_0 = \Lambda(\nu)(z\mathbf{e}+\mathbf{p}_0)$. Since $A(\Lambda(g_0)\mathbf{p};g_0g_1, g_0g_2) = A(\mathbf{p};g_1,g_2)$, Eq. (6) implies that

$$B(U(g_i)f, U^{-1}(\nu^2)f;\mathbf{p}) = A(\Lambda(\nu)\mathbf{p};\nu g_i,\nu^{-1}),$$

or, when written in terms of the matrix J,

$$f(z(\nu^2)) = D(\nu^2)[J(z)D(z)]^{-1}A(z(\nu)), \tag{13}$$

where $D(z) \equiv D(\tilde{q})$, $q = (z\mathbf{e}+\mathbf{p}_0, \omega(z))$, and $A(z(\nu))$ is the column vector with components $A(z(\nu);\nu g_i,\nu^{-1})$. With $w = z(\nu^2)$ we easily calculate $2\gamma z(\nu) = w + z$, and therefore the estimate (12) results in

$$\left\|\left(\frac{w+z}{2\gamma}\right)^k f(w)\right\| < c''_{Nk}(\epsilon)\gamma^{-N}\exp[r_\epsilon|\mathrm{Im}(w+z)|] \tag{14}$$

for z on the "double circle" $\mathcal{C}$. Here we have used the fact that $\|D(\nu^2)\|\gamma^{-2s}$ is bounded. This gives the result

$$\|w^k f(w)\| < c_k \exp(r_\epsilon|\mathrm{Im}w|) \tag{15}$$

for all w which can be reached from some z on the double circle $\mathcal{C}$ with a real velocity transformation. That these w in fact make up the whole two-sheeted Riemann surface can be seen by direct calculation or by the following argument: We consider the real 2-vectors $p = (\mathrm{Re}z, \mathrm{Re}(z^2+\lambda^2)^{1/2})$ and $q = (\mathrm{Im}z, \mathrm{Im}(z^2+\lambda^2)^{1/2})$. Since the invariant $I(z) = p\cdot p + q\cdot q = |z^2+\lambda^2| - |z|^2$ varies in the interval $[-\lambda^2,\lambda^2]$ and since $p\cdot p = (I+\lambda^2)/2$ and $q\cdot q = (I-\lambda^2)/2$, p and q are respectively timelike or null and spacelike or null. Hence the real velocity transformations preserve sgn $[\mathrm{Re}(z^2+\lambda^2)^{1/2}]$ (and therefore the sheet structure) and sgn $(\mathrm{Im}z)$. It is not difficult to see that all points z with the same value of the invariants are related by a real velocity transformation. The demonstration is then completed by noting that each curve described by a fixed value of the invariants has an intersection with one of the circles. {These curves are the hyperbolas $(1+t)^{1/2}\mathrm{Im}z = \pm[(1-t)(\mathrm{Re}z)^2 + (1-t^2)(\lambda^2/2)]^{1/2}$ for $t \equiv I/\lambda^2$ in the interval $(-1,1]$ and the lines $\pm\mathrm{Im}z \geq \lambda, \mathrm{Re}z = 0$.} We then have Eq. (15) for all w and hence the $f^i(w\mathbf{e}+\mathbf{p}_0)$ also satisfy Eq. (15). This is just the condition which insures that

$$\int_{-\infty}^{\infty} dp\, f^i(p\mathbf{e}+\mathbf{p}_0)\, e^{ipx} = 0 \quad \text{for } |x| > r_\epsilon.$$

Since this holds for arbitrary $\mathbf{e}$ and $\mathbf{p}_0$, integration over all $\mathbf{p}_0$ perpendicular to $\mathbf{e}$ gives $\tilde{f}^i(x\mathbf{e}) = 0$ for $|x| > r_\epsilon$ and all $\mathbf{e}$. Since $\epsilon > 0$ is arbitrary, $\tilde{f}^i(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > r$, which establishes Lemma 3.

Thus the operator Q of Eq. (10) determines a set of fields relative to which Q is local. We would now like to show that two different operators, Q_1 and Q_2, of the same type as Q, determine relatively local free fields. Thus suppose

$$Q_1 = \sum_{k=1}^{n_1}\sum_{\mu=-s}^{s}\int\frac{d^3\mathbf{p}}{2\omega}\,[a^\dagger_\mu(k,\mathbf{p})f_\mu(k,\mathbf{p}) + a_\mu(k,\mathbf{p})f^*_\mu(k,\mathbf{p})],$$

$$Q_2 = \sum_{k=1}^{n_2}\sum_{\mu=-s}^{s}\int\frac{d^3\mathbf{p}}{2\omega}\,[b^\dagger_\mu(k,\mathbf{p})h_\mu(k,\mathbf{p}) + b_\mu(k,\mathbf{p})h^*_\mu(k,\mathbf{p})],$$

with $[a_\mu(k,\mathbf{p}), b^\dagger_\nu(l,\mathbf{q})]_\pm = \epsilon_{kl}D^s_{\mu\nu}(\tilde{p})2\omega\delta(\mathbf{p}-\mathbf{q})$. Here ϵ_{kl} is an $n_1\times n_2$ matrix. As was shown in Lemma 2, the particle creation operators can be selected so that the wavefunctions $f_k(\mathbf{p})$ are linearly independent and satisfy $f^\theta(\mathbf{p}) = e^{i\pi s}\hat{f}(-\mathbf{p},-\omega)$, $f(\mathbf{p}) = \hat{f}(\mathbf{p},\omega)$. The same applies to the $h_k(\mathbf{p})$. The analog of Eq. (6) for the function

$$\mathfrak{F}_{12}(x;g_1,g_2) = [U(x,g_1)Q_1U^{-1}(x,g_1), U(g_2)Q_2U^{-1}(g_2)]_\pm$$

is

$$B(U(g_1)f, U(g_2)\epsilon h;\mathbf{p}) = A_{12}(\mathbf{p};g_1,g_2), \tag{6'''}$$

where $(\epsilon h)_k(\mathbf{p}) = \sum_{l=1}^{n_2}\epsilon_{kl}h_l(p)$ and where A_{12} has the same analyticity and reality properties as the function $A(\mathbf{p};g_1,g_2)$. The analog of Eq. (10a) is therefore

$$B(U(g_1)f, (\epsilon-\epsilon^*)U(g_2)h;\mathbf{p}) = 0, \tag{10a'}$$

leading to

$$\sum_{k=1}^{n_1}((\epsilon-\epsilon^*)h)_k(\mathbf{p})f^\dagger_k(\mathbf{q}) = 0, \tag{10b'}$$

which implies $\epsilon = \epsilon^*$. This is just the condition that the fields constructed from the "a" creation operators are local relative to those constructed from the "b" operators. It also guarantees that the "b" fields anticommute or commute with Q_1 when these fields are at a point x spacelike separated from the region of localization of Q_1.

We have thus demonstrated that the set of all $Q \in \{\bigcup_{D\in\mathfrak{D}} S(D)\}$ which also have the property that $U(x,g)Q|0\rangle$ is C^∞ in (x,g) determines a set $\mathcal{L}$ of relatively local fields which are also local relative to those Q. It is now a simple matter to construct the standard set of fields referred to in the theorem.[12] Choosing N_s linearly independent creation operators $b^\dagger_\mu(k,\mathbf{p})$, $k = 1, \ldots, N_s$, of spin s from those associated with fields in $\mathcal{L}$, we note that $\langle 0|b_\mu(k,\mathbf{p})b^\dagger_\nu(l,\mathbf{q})|0\rangle = 2\omega\delta(\mathbf{p}-\mathbf{q})D^s_{\mu\nu}(\tilde{p})M_{kl}$, where M is *real*, symmetric, and positive definite. If we choose $c^\dagger_\mu(k,\mathbf{p}) = \sum_l b^\dagger_\mu(l,\mathbf{p})\alpha_{lk}$ and construct fields $\phi^k(x)$ from these new creation operators, we obtain fields that are local relative to this set of Q's if and only if α is real. Since under this change of creation operators $M \to M' = \alpha^\dagger M\alpha$, we can achieve $M' = I$ with a suitable real α. This gives the standard set of fields referred to in the theorem. We finally show that these fields are also local relative to $Q \in S(D)$ for which $U(x,g)Q|0\rangle$ is not C^∞ in (x,g). The proof goes via a limiting process: Let $Q_n = \int d^4x dg\, f_n(x,g)U(x,g)QU^{-1}(x,g)$, where $f_n(x,g)$ is a sequence of real, non-negative C^∞ functions with compact support Δ_n converging to the point $(0,1)$. With $\int d^4x dg f_n(x,g) = 1$, it is easy to see that $Q_n|0\rangle$ converges strongly to $Q|0\rangle$. If $h(x) \in \mathcal{S}$ has support in D', then for large enough n, $\langle 0|[Q_n, \phi^k_\mu(h)]_\pm|0\rangle = 0$. [Here $\phi^k_\mu(h) = \int d^4x\ \phi^k_\mu(x)h(x)$.] Therefore, $[Q, \phi^k_\mu(h)]_\pm = 0$ and the proof of Theorem 1 is complete.

We conclude this section with a few remarks. We first note that the creation operators used to construct the N fields are unique up to a real orthogonal transformation. That is, if the equation $VC^\dagger_\mu(k,\mathbf{p})V^{-1} = \sum_l c^\dagger_\mu(l,\mathbf{p})V_{lk}$ defines a unitary operator V which commutes with the Poincaré group and which satisfies $V|0\rangle = |0\rangle$, then the fields $V\phi^k(x)V^{-1}$ are local relative to the $S(D)$ if and only if the matrix V_{lk} is real. Another way of stating this is also of interest: The abbreviation $f^0(\mathbf{p})$ used in Theorem 1 actually defines an antiunitary TCP operator when the $f_i(\mathbf{p})$ refer to the standard creation operators $c^\dagger_\mu(k,\mathbf{p})$. We see that the set of operators $\theta S(D)\theta^{-1}$ is local relative to the fields in the region $[-D]'$. With this definition of θ, $V\phi^k(x)V^{-1}$ is local relative to the $S(D)$ if and only if $[\theta, V] = 0$.[13,14]

Secondly, we remark on the simple representation (11) for the operators $Q \in S(D)$. This representation, with $\tilde{f}^i(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > r$, has been shown to hold for those $Q \in S(D_r)$, $D_r = \{x \mid \|\mathbf{x}\| + |t| < r\}$, for which $U(x,g)\,Q|0\rangle$ is C^∞ in (x,g). For a general operator $Q \in S(D_r)$ the differentiability condition need not hold; we only require that $Q|0\rangle$ exist. In this case the vanishing of the tempered distributions $[Q, \phi^k_\mu(x)]_\pm$ for $x \in D'_r$ still implies Lemma 2 for wavefunctions $f_i(\mathbf{p})$ which may now grow like a power of $\mathbf{p}$ in the real domain; hence the representation (11) no longer applies. Lemma 3 can be replaced by the statement that the Fourier transforms of $f^1(\mathbf{p})$ and $f^2(\mathbf{p})$ are tempered distributions with support in $\|\mathbf{x}\| \leq r$.

IV. APPLICATION TO FIELD THEORIES AND LOCAL ALGEBRA THEORIES

A. Asymptotic Fields

We consider algebras of local operators $\mathfrak{F}(D)$ associated with diamonds D. These might be the polynomial algebras of a field theory associated with bounded regions of space–time or Haag–Araki algebras of bounded operators.[1,2] Considered as a vector space, $\mathfrak{F}(D)$ is spanned by the two subspaces $\mathfrak{F}_\pm(D)$ containing Fermi and Bose elements, respectively. The observables of the theory are, of course, contained in the $\mathfrak{F}_-(D)$. Locality is introduced by assuming that, for $D_1 \subseteq D'_2$,

$$\begin{aligned}[\mathfrak{F}_\pm(D_1), \mathfrak{F}_\pm(D_2)]_\pm &= 0,\\ [\mathfrak{F}_\mp(D_1), \mathfrak{F}_\pm(D_2)]_- &= 0.\end{aligned} \tag{16}$$

Note that, in field theory, Eq. (16) contains the statement of normal statistics which in this case involves no loss of generality.[15] We now assume that the theory is asymptotically complete so that there are Fock spaces of "incoming" and "outgoing" particles each of which is in fact equal to the entire Hilbert space. The creation operators are denoted $a^\dagger_{\mathrm{ex}}(i,\mathbf{p})$, where i collectively denotes spin and particle type and "ex" means "in" or "out." If the $a^\dagger_{\mathrm{ex}}(i,\mathbf{p})$ are defined so that

$$\langle 0|a_{\mathrm{ex}}(i,\mathbf{p})a^\dagger_{\mathrm{ex}}(j,\mathbf{q})|0\rangle = \delta_{ij}\delta(\mathbf{p}-\mathbf{q}),$$

then for each operator Q in $\mathfrak{F}_\pm(D)$ and each particle mass the formula[2]

$$Q_{\mathrm{ex}} = \sum_i \int d^3\mathbf{p}[\langle 0|a_{\mathrm{ex}}(i,\mathbf{p})Q|0\rangle a^\dagger_{\mathrm{ex}}(i,\mathbf{p}) + a_{\mathrm{ex}}(i,\mathbf{p})\langle 0|Qa^\dagger_{\mathrm{ex}}(i,\mathbf{p})|0\rangle] \tag{17}$$

defines a certain one-particle operator. For a fixed D, the set of all such Q_{ex} is denoted $\mathfrak{F}^{\mathrm{ex}}_\pm(D)$. The sum is over all particles of a given mass m and hence $\mathfrak{F}^{\mathrm{ex}}(D)$ depends implicitly on the mass of the particles involved. Landau has shown[4] that the $\mathfrak{F}^{\mathrm{ex}}_\pm(D)$ satisfy the locality properties assumed for the $S_\pm(D)$ in Sec. III. Hence, restricting our further attention to theories with a minimum mass > 0 and at most a finite number of particles at each mass, Theorem 1 implies the existence of a relatively local set of asymptotic fields $\phi^k_{\mathrm{ex}}(x)$ which are local relative to the $\mathfrak{F}^{\mathrm{ex}}_\pm(D)$.

B. Local Internal Symmetries

Following Landau and Wichmann,[3,4] we call a unitary operator V a local internal symmetry if $V|0\rangle = |0\rangle$ and if for $D_1 \subseteq D'_2$ we have, in analogy to (16),

$$\begin{aligned}[V\mathfrak{F}_\pm(D_1)V^{-1}, \mathfrak{F}_\pm(D_2)]_\pm &= 0,\\ [V\mathfrak{F}_\mp(D_1)V^{-1}, \mathfrak{F}_\pm(D_2)]_- &= 0.\end{aligned} \tag{18}$$

Under these assumptions $[U(x), V] = 0$, as was shown in Ref. 3. Furthermore, Landau has shown,[4] using the Haag–Ruelle construction,[16,17] that V

commutes with the S matrix and that

$$Va^{\dagger}_{ex}(i,\mathbf{p})V^{-1} = \sum_j a^{\dagger}_{ex}(j,\mathbf{p})\mathcal{U}_{ji}(\mathbf{p}). \tag{19}$$

In a field theory, with the additional assumption of the existence of an interpolating field for each particle, Landau was able to show that

$$[U(x,g), V] = 0 \tag{20}$$

and

$$V\psi^k_{ex}(x)V^{-1} = \sum_l \psi^l_{ex}(x)\mathcal{K}_{lk}. \tag{21}$$

Here the $\psi^k_{ex}(x)$ are suitably chosen asymptotic fields associated with the interpolating fields and $\mathcal{K}$ is a *real* orthogonal matrix. We will now derive (20) and (21) without assuming the existence of interpolating fields. Equation (21) will then be true for the asymptotic fields $\phi^k_{ex}(x)$ discussed in Sec. IVA above.

We first note that the assumption (18) coupled with the result of Landau mentioned in Sec. IVA shows that for $D_1 \subseteq D_2'$ the vacuum expectation values of the expressions

$$[V\mathfrak{F}^{ex}_{\pm}(D_1)V^{-1}, \mathfrak{F}^{ex}_{\pm}(D_2)]_{\pm}$$

and

$$[V\mathfrak{F}^{ex}_{\mp}(D_1)V^{-1}, \mathfrak{F}^{ex}_{\pm}(D_2)]_{-}$$

both vanish. The problem is thus reduced to one involving one-particle operators alone. Let us choose a set of creation operators $c^{\dagger ex}_{\mu}(k,\mathbf{p})$ transforming according to Eq. (3) and such that $\phi^k_{\mu ex}(x)$ is linear in $c^{\dagger ex}_{\mu}(k,\mathbf{p})$ and its Hermitian conjugate. If Eq. (19) is rewritten in terms of these creation operators, then the matrix $\mathcal{U}(\mathbf{p})$ is replaced by a matrix which we denote $K(\mathbf{p})$. In order to examine $K(\mathbf{p})$, we choose $Q_1, Q_2 \in \mathfrak{F}^{ex}_{\pm}(D)$ such that for each spin the particle wavefunctions $f_i(\mathbf{p})$ [relative to the $c^{\dagger ex}_{\mu}(k,\mathbf{p})$] associated with $Q_1|0\rangle$ are linearly independent, and similarly for the $h_i(\mathbf{p})$ associated with $Q_2|0\rangle$. Furthermore, we require that $Q_i|0\rangle$ contain all the different types of particles of mass m and half-integer (or integer) spin and that both $U(x)Q_1|0\rangle$ and $U(x)Q_1^{\dagger}|0\rangle$ are C^{∞} in the variable x. We then consider the function

$$\mathfrak{F}(x;g_1,g_2,g) \equiv \langle 0|[U(x,g_1)Q_1^{\dagger}U^{-1}(x,g_1), V(g)U(g_2)Q_2U^{-1}(g_2)V^{-1}(g)]_{\pm}|0\rangle, \tag{22}$$

where $V(g) \equiv U(g)VU^{-1}(g)$ is also a local internal symmetry. If we fix $\epsilon > 0$ and require $\| \Lambda(g_i) - I\| < \epsilon$ for $i = 1, 2$, then $\mathfrak{F}(x;g_1,g_2,g) = 0$ for x spacelike relative to a fixed diamond D^{ϵ}, uniformly in g. In p-space this leads to the statement that

$$B(U(g_1)f, V(g)U(g_2)h;\mathbf{p}) = A(\mathbf{p};g_1,g_2,g), \tag{23}$$

where $A(\mathbf{p};g_1,g_2,g)$ is an entire function of p satisfying

$$\|\mathbf{p}\|^k |A(\mathbf{p};g_1,g_2,g)| < c_k(\epsilon)\exp(\alpha_{\epsilon}\|\mathrm{Im}\mathbf{p}\|) \tag{24}$$

with α_{ϵ} and $c_k(\epsilon)$ independent of g. The Hermitian form $B(f,h;\mathbf{p})$ has the same meaning as in Eq. (6) except that a sum over spins is also understood. Considering along with the expressions (22) the corresponding expressions obtained when $Q_1 \in \mathfrak{F}^{ex}_{\mp}$ and $Q_2 \in \mathfrak{F}^{ex}_{\pm}$, we see that both integer and half-integer spins can be assumed to be represented in f and h. Note also that $V(g)$ in (23) corresponds to the matrix $K(\mathbf{p};g) = D(g)K(\Lambda^{-1}(g)\mathbf{p})D^{-1}(g)$. The construction of matrices of the type $J(\mathbf{p})$ as in Eq. (9) from Lorentz-transformed f's and h's allows us to invert Eq. (23) to get

$$K(\mathbf{p};g) = [J_1(\mathbf{p})D(\tilde{p})]^{-1}A(\mathbf{p};g)[J_2^T(\mathbf{p})]^{-1}, \tag{25}$$

where $J_1^*(\mathbf{p})$ and $J_2(\mathbf{p})$ are constructed respectively from $(U(g_1)f)(\mathbf{p})$ and $(U(g_i')h)(\mathbf{p})$ and where $A_{ij}(\mathbf{p};g) = A(\mathbf{p};g_i,g_j',g)$. Since $f(\mathbf{p})$ and $h(\mathbf{p})$ are entire functions of p, Eq. (25) shows explicitly that $K(\mathbf{p})$ is a meromorphic function of the 4-vector p. A suitable choice of the g_i and g_i' can remove any of the zeroes of the denominator, which means that in fact $K(\mathbf{p})$ is an entire function of p. We now consider the growth of $K(\mathbf{p})$ for complex $\mathbf{p}$. As in Lemma 3, we examine $K(\Lambda^{-1}\mathbf{p})$ for $\mathbf{p} = z\mathbf{e} + \mathbf{p}_0$. We choose a double circle $\mathcal{C}$ in the z complex space of radius greater than $(\mathbf{p}_0^2 + m^2)^{1/2}$ on which $\det J_1(z)$ and $\det J_2(z)$ have no zeroes. For $z \in \mathcal{C}$, the inequality (24) gives

$$\|K(z\mathbf{e} + \mathbf{p}_0;g)\| < c. \tag{26}$$

We choose g to be a velocity transformation ν of velocity v along $\mathbf{e}$ and define $\gamma = \Lambda_{00}(\nu)$. Since $\|D(\nu)\|\gamma^{-s}$ is bounded, we have

$$\|K(w\mathbf{e} + \mathbf{p}_0)\| < c'\gamma^{2s}, \tag{27}$$

where $w = z(\nu^{-1}) = \gamma[z - vw(z)]$. Since, for large w, v must be near 1 and $z - vw(z)$ cannot be near zero, we have

$$\|K(w\mathbf{e} + \mathbf{p}_0)\| < c''(1 + |w|^{2s}). \tag{28}$$

This means that $K(w\mathbf{e} + \mathbf{p}_0)$ is in fact a polynomial in w and $\omega(w)$.[18] An argument given by Landau[4] then suffices to show that V commutes with the Lorentz group. A variant of this argument will be outlined here for the reader's convenience: We first note that the unitarity of V is reflected in the equation

$$K^{\dagger}(\mathbf{p})D(\tilde{p})K(\mathbf{p}) = D(\tilde{p}) \tag{29}$$

and hence, since $D(\tilde{p}) = D^2(\nu(-\mathbf{p}))$, where $\nu(\mathbf{p}) \in g$ is the boost to momentum $\mathbf{p}$, that $\Gamma(\mathbf{p}) \equiv D^{-1}(\nu(\mathbf{p}))K(\mathbf{p})D(\nu(\mathbf{p}))$ is a unitary matrix. Let $p = \Lambda(g)(0,m)$, where g is the velocity transformation $\cosh x + \sinh x\mathbf{e}\cdot\sigma$. Since $\Lambda(g)$ and $D(g)$ are polynomials in the matrix elements of g and g^*, for fixed $\mathbf{e}$ the matrix Γ is a polynomial in $\lambda = e^x$ and $\lambda^{-1} = e^{-x}$. But since Γ must remain bounded as λ varies in $0 < \lambda < \infty$, it is in fact independent of λ. Setting $\lambda = 1$ gives $\Gamma(\mathbf{p}) = K(0)$. We now apply the same argument to $K(\mathbf{p};g)$ with the result that $D^{-1}(\nu(\mathbf{p}))K(\mathbf{p};g)D(\nu(\mathbf{p}))$ is also independent of $\mathbf{p}$. If $g = \nu(\mathbf{q})$,

we can choose $\mathbf{p} = \mathbf{q}$ in which case the latter expression becomes $K(\mathbf{0})$. This shows that for g a velocity transformation $K(\mathbf{p}; g) = K(\mathbf{p})$. But since the velocity transformations generate $\mathcal{G}$, we have $K(\mathbf{p}; g) = K(\mathbf{p})$ for all $g \in \mathcal{G}$; i.e., $[U(g), V] = 0$.

Finally, we take $u \in SU(2)$ and note that $K(\mathbf{0}; u) = D(u)K(\mathbf{0})D^{-1}(u) = K(\mathbf{0})$ implies that $K(\mathbf{p}) = D(\nu(\mathbf{p}))\Gamma(\mathbf{p})D^{-1}(\nu(\mathbf{p})) = K(\mathbf{0}) \equiv \mathcal{K}$. Hence $K(\mathbf{p}) = \mathcal{K}$ is a unitary matrix which acts only on the particle indices and does not mix particles of different spin. The closing remarks of Sec. III then suffice to show that $\mathcal{K}$ is a real matrix and thus to complete the demonstration of the Eqs. (20) and (21).

C. Haag–Ruelle Theory

In the discussion of local internal symmetries the Haag–Ruelle construction of asymptotic states was needed to relate the action of V on the one-particle states to its action on the many-particle states. [See the derivation of Eq. (19) in Ref. 4.] We first want to point out that the method of Haag and Ruelle[16,17] rests on an assumption which is not entirely appropriate in a study of the possible symmetries of a local theory: It is assumed[16,17] that pure one-particle states can be constructed by letting quasi-local operators[19] Q act on the vacuum. This means that quasilocal operators Q can be found such that the states $Q|0\rangle$ are eigenstates of the mass operator. In certain situations this supposition follows directly from the assumed mass spectrum: Suppose that the particle in question has mass $m > 0$ and that E_m and $E(\mu^2)$ project onto the subspaces with $P^2 = m^2$ and $P^2 \leq \mu^2$, respectively. If m^2 is an isolated point in the spectrum of P^2, then it is easy to find the required quasilocal Q.[20] For $R \in \mathfrak{F}_+(D)$ satisfying $E_m R|0\rangle \neq 0$, the operator $R(x) = U(x)RU(-x)$ can be suitably smeared with a function $f(x) \in \mathcal{S}$ to give a quasilocal $Q = \int d^4x\, R(x)f(x)$ satisfying $Q|0\rangle = E_m Q|0\rangle = R|0\rangle$. However, it may be the case that m is in the mass continuum of some many-particle states. In this case one might try to justify the Haag-Ruelle assumption in the following manner: Any continuum contribution to the support of $dE(\mu^2)$ in a neighborhood of m^2 should correspond to states with "quantum numbers" different from those of the particle in question; otherwise the particle would not be stable. At this point, one might argue that in a "reasonable theory" there ought to exist local operators carrying the appropriate quantum numbers. For example, if these quantum numbers were associated with a compact local internal symmetry group, then local operators carrying definite values of these quantum numbers could be constructed from arbitrary local operators (by means of the Haar integral, for example). We could then choose a local operator Q which carried the quantum numbers of the particle in question. This operator would satisfy $dE(\mu^2)(1 - E_m)Q|0\rangle = 0$ in an interval around m^2, while $E_m Q|0\rangle \neq 0$. By smearing $Q(x)$ as before, one would arrive at a quasilocal operator satisfying the assumption of Haag and Ruelle.

In view of this close connection of the Haag–Ruelle assumption with the symmetries of the theory, we feel that a better understanding of these symmetries could result if asymptotic states could be constructed without making this assumption.

In a theory of strong interactions we are no doubt justified in assuming that $E(\mu^2)(1 - E_m)$ is continuous in a neighborhood of m^2 (for all particle masses m). This is just the assumption that particle masses have no point of accumulation. If one makes a slightly stronger assumption concerning the smoothness of $E(\mu^2)(1 - E_m)Q|0\rangle$, then, as we shall show, it is possible to carry out the Haag–Ruelle construction of asymptotic states. Specifically we assume that, for each particle mass m, operators $Q_i \in \mathfrak{F}_+(D_i)$ can be found such that

(a) $E_m Q_i|0\rangle \neq 0$,

(b) $E(\mu^2)(1 - E_m)Q_i|0\rangle$ is Hölder continuous at m^2 [i.e., there exist ϵ_i, δ_i, and c_i all greater than zero such that, for $|\mu^2 - m^2| < \delta_i$, $\psi_i = (1 - E_m)Q_i|0\rangle$ satisfies

$$\|[E(\mu^2) - E(m^2)]\,\psi_i\| < c_i\,|\mu^2 - m^2|^{\epsilon_i}\}.$$

If, furthermore, for each mass m, the Poincaré transforms of $\{E_m Q_i|0\rangle\}$ span the one-particle sub-Hilbert space, asymptotic Fock spaces can be constructed. We note that although the Hölder continuity condition is a considerable relaxation of the Haag–Ruelle assumption, it may not be physically motivated and hence our result should be considered provisional.

The idea of the construction is to use quasilocal "creation operators" $Q_i(f_i, t)$ at finite times t, which, as $t \to \pm\infty$, have a mass spectrum "converging" to the point m. This enables us to construct a dense set of collision states. Details are given in the Appendix.

ACKNOWLEDGMENTS

I would like to thank Professor E. H. Wichmann for suggesting that a result such as the theorem of Sec. III should be true and for many valuable discussions. It is also a pleasure to thank Mr. Peter Mohr for helpful discussions. I am very grateful to my wife Judith for her encouragement during the course of this work.

APPENDIX

Suppose that, for $i = 1, \ldots, n$, $Q_i \in \mathfrak{F}_+(D_i)$ satisfies the conditions (a) and (b) of Sec. IVC with m replaced by m_i. Let $h(s)$ be a real C^∞ function of the real variable s with compact support and with $h(0) = 1$. We choose n C^∞ functions $f_i(p)$ of the 4-vector p with compact support inside the forward light cone and define

$$Q_i(f_i, t) = \int d^4p\, \tilde{Q}_i(p) f_i(p) h(s(p^2 - m_i^2)^2) \times \exp\{i[p_0 - \omega_i(\mathbf{p})]\,t\},$$

where $\tilde{Q}_i(p)$ is the Fourier transform of $Q_i(x)$,

$\omega_i(\mathbf{p}) = (\mathbf{p}^2 + m_i^2)^{1/2}$, and $s = s(t)$ is a function (to be specified later) which increases with t.

The main tool used in the construction of asymptotic states is a generalization of a spacelike cluster property proved by Ruelle[17]: Let $\langle 0 | R_1(x_1) \cdots R_n(x_n) | 0 \rangle_T$ be the truncated vacuum expectation value of the product of operators $R_i(x_i)$, with $R_i(0) \in \mathfrak{F}_\pm(D_i)$. Let, furthermore, $f(s; y_1, \cdots, y_n)$ be the Fourier transform of a function in $\mathcal{S}$ multiplied by $\prod_{i=1}^{n} h(s(p^2 - m_i^2)^2)$. Then for some $N_0 \geq 0$ the function

$$\mathfrak{F}(s; \mathbf{x}_1, \ldots, \mathbf{x}_n) = \int \prod_{i=1}^{n} d^4y_i \, f(s; y_1, \ldots, y_n) \times \langle 0 | R_1(y_1 + \mathbf{x}_1) \cdots R_n(y_n + \mathbf{x}_n) | 0 \rangle_T$$

satisfies the inequality

$$| D^k \mathfrak{F}(s; \mathbf{x}_1, \ldots, \mathbf{x}_n) | < c_{Nk} \, d^{-N} (| s |^{N/2+N_0} + 1) \tag{A1}$$

with $d = \max_{i,j} \| \mathbf{x}_i - \mathbf{x}_j \|$. Hence D^k is any monomial in the derivatives with respect to the $\mathbf{x}_i$ and N is an arbitrary nonnegative integer. The proof will not be given since it involves only a straightforward generalization of Ruelle's estimates. Note that the inequality (A1) might be expected since the function $h(s(p^2 - m^2)^2)$ contributes an $(s)^{1/2}$ "spreading" in coordinate space.

We now investigate the convergence of the vector $\psi(t) = Q_1(f_1, t) \cdots Q_n(f_n, t) | 0 \rangle$, where the $f_i(p)$ are chosen "nonoverlapping in velocity space"[21]; that is, for all $i \neq j$, if $p_i \in \operatorname{supp} f_i$ and $p_j \in \operatorname{supp} f_j$, then $\mathbf{p}_i/\omega_i(\mathbf{p}_i) \neq \mathbf{p}_j/\omega_j(\mathbf{p}_j)$. Expanding $\| d\psi(t)/dt \|^2$ in truncated vacuum expectation values gives a sum of products of N-point functions. We first consider an N-point function where $N \geq 3$ and which involves no operator Q_i which has been differentiated with respect to t. After a renumbering of the operators, the N-point function can be written

$$\int \prod_{i=2}^{N} d^3p_i \, \tilde{F}(s; \mathbf{p}_2, \cdots, \mathbf{p}_N) \exp\{ i[\omega_1(\mathbf{p}_1) + \cdots + \omega_m(\mathbf{p}_m) - \omega_{m+1}(\mathbf{p}_{m+1}) - \cdots - \omega_N(\mathbf{p}_N)] t \}, \tag{A2}$$

where $F(s; \mathbf{x}_2, \cdots, \mathbf{x}_N) \equiv \langle 0 | Q_1^\dagger(s; \mathbf{0}) Q_2^\dagger(s; -\mathbf{x}_2) \cdots Q_m^\dagger(s; -\mathbf{x}_m) Q_{m+1}(s; \mathbf{x}_{m+1}) \cdots Q_N(s; \mathbf{x}_N) | 0 \rangle_T$ and $Q_i(s; \mathbf{x}) \equiv \int d^4p \, \tilde{Q}_i(p) f_i(p) h(s(p^2 - m_i^2)^2) e^{-i\mathbf{p}\cdot\mathbf{x}}$. Note that, in (A2), $\mathbf{p}_1 + \cdots + \mathbf{p}_m = \mathbf{p}_{m+1} + \cdots + \mathbf{p}_N$ and that $1 \leq m < N$. The estimate (A1) then shows that if D^l is an lth derivative with respect to $\mathbf{p} = (\mathbf{p}_2, \cdots, \mathbf{p}_N)$, then for some N_0'

$$\| \mathbf{p} \|^k | D^l \mathbf{F}(s; \mathbf{p}_2, \ldots, \mathbf{p}_N) | < c_{kl}' (| s |^{l/2+N_0'} + 1). \tag{A3}$$

If we now take $s = | t |^{2-\epsilon}$ with $2 \geq \epsilon > 0$, repeated integration by parts in the expression (A2) shows (see Ref. 21) that it is bounded by $c_k | t |^{-k}$ for any $k \geq 0$. Note that this bound remains valid when one or more $Q_i(f_i, t)$ are replaced by their time derivatives; the differentiation only results in the replacement of $h(s(p^2 - m_i^2)^2)$ by $i(p_0 - \omega_i) \times h(s(p^2 - m_i^2)^2) + (d/dt) h(s(p^2 - m_i^2)^2)$.

We now consider the two-point function. The assumption of Hölder continuity is easily seen to imply that if $\epsilon < 2\epsilon_i/(1 + \epsilon_i)$, $s = | t |^{2-\epsilon}$, then for some $\beta_i > 0$

$$\left\| \frac{dQ_i}{dt} (f_i, t) | 0 \rangle \right\| < \frac{\alpha_i}{(1 + | t |)^{1+\beta_i}} .$$

Hence, for appropriate ϵ, each two-point function is bounded, and the terms in the expansion of $\| d\psi(t)/dt \|^2$ containing products of two-point functions only is bounded by $c | t |^{-2-\delta}$ with $\delta > 0$. Thus $\psi(t)$ converges strongly. An argument similar to that given by Hepp[20] establishes the Lorentz frame independence of the construction. The existence of asymptotic Fock spaces is thus demonstrated. Note that it is sufficient to consider only sets of functions $f_i(p)$ with nonoverlapping supports in velocity space because the linear span of the vectors constructed using such functions is already dense in Fock space.[2,21]

We remark finally that once the convergence of $\psi(t)$ to ψ has been demonstrated as above, it is no longer necessary to choose $s(t)$ as we did. In fact suppose $\mu(t)$ is any function of t satisfying

(a) $\mu(t) \to +\infty$ as $t \to \pm\infty$,

(b) for some $\sigma > 0$, $| t |^\sigma (\mu(t)/t^2) \to 0$ as $t \to \pm\infty$.

Then it is easy to see that if $\phi(t)$ is constructed exactly as $\psi(t)$ above except $s(t)$ is replaced by $\mu(t)$, we have $\| \phi(t) - \psi(t) \| \to 0$ as $t \to \pm\infty$ and hence $\phi(t)$ converges to the same vector as $\psi(t)$. Although we will not demonstrate it, this fact can be used to prove strong convergence without the restriction to $f_i(p)$ which are nonoverlapping in velocity space.

---

* This work was supported in part by the U.S. Atomic Energy Commission.

† Research supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Grant No. AF-AFOSR-68-1471.

1 R. Haag and B. Schroer, J. Math. Phys. 3, 248 (1962).

2 H. Araki and R. Haag, Commun. Math. Phys. 4, 77 (1967).

3 L. J. Landau and E. Wichmann, J. Math. Phys. 11, 306 (1970).

4 L. J. Landau, Commun. Math. Phys. 17, 156 (1970).

5 The field $\phi_\mu(x)$ is an interpolating field for particle i if the projection E_i onto the one-particle states of this particle satisfies $E_i \phi_\mu(x) | 0 \rangle \neq 0$.

6 S. Weinberg, Phys. Rev. 133, B1318 (1964).

7 H. Epstein, J. Math. Phys. 8, 750 (1967). Epstein proved Lemma 2 in the case $n = 1$. His method can easily be generalized to arbitrary n.

8 E. Goursat, *Differential Equations* (Dover, New York, 1959).

9 P. Hartman, *Ordinary Differential Equations* (Wiley, New York, 1964).

10 This observation is also made in Ref. 7.

11 H. Epstein, V. Glaser, and A. Martin, Commun. Math. Phys. 13, 257 (1969). In the case $n = 1$, $s = 0$, it is shown here that the functions $\tilde{f}^i(\mathbf{x})$ have compact support. The precise region of support is not determined.

[12] J. Schwartz, J. Math. Phys. 2, 271 (1961).
[13] P. Carruthers, Phys. Rev. Letters 18, 353 (1967).
[14] B. Zumino and D. Zwanziger, Phys. Rev. 164, 1959 (1967).
[15] H. Araki, J. Math. Phys. 2, 267 (1961).
[16] R. Haag, Phys. Rev. 112, 669 (1958).
[17] D. Ruelle, Helv. Phys. Acta 35, 147 (1962).
[18] Another perhaps more direct way of proving this fact is to use Eq. (25) and Theorem A3. 2 of Ref. 11.
[19] Quasilocal operators can be defined as those operators which satisfy a spacelike cluster property given by Haag in Ref. 16. See also Eq. (A1) in the Appendix with $s = 0$.
[20] K. Hepp, in 1965 Brandeis University Summer Institute in Theoretical Physics, *Axiomatic Field Theory*, edited by M. Chretien and S. Deser (Gordon and Breach, New York, 1966), Vol. I.
[21] K. Hepp, Commun. Math. Phys. 1, 95 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Electromagnetic Radiation in Curved Spaces*

Richard Sigal
Belfer Graduate School of Science, Yeshiva University, New York, N.Y., 10033
(Received April 20 1971)

The classes of trajectories of charged particles, for which the radiation reaction force vanishes, in curved spaces are examined. The trajectories in the Robertson–Walker cosmologies are given in detail and it is shown that the only reaction-free geodesic trajectories are those at rest relative to the local matter density.

I. INTRODUCTION

The question of which classes of trajectories of charged particles in a curved space are free of a radiation reaction force is an old one in general relativity. We examine this question through the use of the conformal symmetries of Maxwell's equations in cases where the metric is taken as given.

A reaction-free trajectory shall be defined as one for which the radiation reaction force, as originally defined by Dirac, vanishes. We have chosen to study radiation from this aspect to avoid global problems such as the structure of infinity. The meaning of this condition in flat space has been widely discussed in relation to the problem of a uniformly accelerated charged particle.[1] From these discussions it is clear that this radiation reaction formalism and definitions of radiation as a nonvanishing Poynting vector at infinity are not unambiguously equivalent, and also that the definitions at infinity themselves suffer from problems.

We have done detailed analysis of the structure of the reaction-free trajectories for the cases of Robertson–Walker cosmologies, where the analysis is somewhat simplified due to their conformal flatness.[2] These models, although simple, are of considerable interest as they represent a close approximation to the large scale structure of the universe as observed.[3]

We have shown in the Robertson–Walker cosmologies that the only reaction-free geodesics are those followed by particles at rest relative to the local matter density.

II. CONFORMAL STRUCTURE

Maxwell's equations[4] for the field $F_{\mu\nu}$, current j_μ, and metric $g_{\mu\nu}$ are given by

$$g^{\mu\alpha}F_{\mu\nu;\alpha} = 4\pi j_\nu, \tag{1}$$

$$F_{[\mu\nu;\lambda]} = F_{[\mu\nu,\lambda]} = 0, \tag{2}$$

where square brackets mean antisymmetrization, ";" covariant derivative with respect to $g_{\mu\nu}$, and "," ordinary derivative. The metric is taken to have signature -2.

Under a conformal transformation

$$g_{\mu\nu} \to \bar{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}, \quad g^{\mu\nu} \to \bar{g}^{\mu\nu} = \Omega^{-2} g^{\mu\nu}$$

Maxwell equations are invariant if we take[5]

$$\bar{F}_{\mu\nu} = F_{\mu\nu}, \quad \bar{j}_\mu = \Omega^{-2} j_\mu.$$

We can then define the densities

$$\mathfrak{F}^{\mu\nu} = \sqrt{-g}F^{\mu\nu}, \quad J^\nu = \sqrt{-g}j^\nu,$$

where g = determinant $g_{\mu\nu}$ and

$$\sqrt{-\bar{g}} = \Omega^4\sqrt{-g}.$$

We then have

$$\bar{\mathfrak{F}}^{\mu\nu} = \mathfrak{F}^{\mu\nu}, \quad \bar{J}^\nu = J^\nu$$

and Eq. (1) reads

$$\bar{\mathfrak{F}}^{\mu\nu}{}_{,\mu} = \mathfrak{F}^{\mu\nu}{}_{,\mu} = \bar{J}^\nu = J^\nu,$$

which gives the conservation law

$$\bar{J}^\nu{}_{,\nu} = J^\nu{}_{,\nu} = 0$$

and thus gives the total conserved electric charge Q as

$$Q = \int J^\mu dS_\mu,$$

which is seen to be a conformal invariant.

This means that under a conformal mapping currents that represent point charges with charge q_i

[12] J. Schwartz, J. Math. Phys. 2, 271 (1961).
[13] P. Carruthers, Phys. Rev. Letters 18, 353 (1967).
[14] B. Zumino and D. Zwanziger, Phys. Rev. 164, 1959 (1967).
[15] H. Araki, J. Math. Phys. 2, 267 (1961).
[16] R. Haag, Phys. Rev. 112, 669 (1958).
[17] D. Ruelle, Helv. Phys. Acta 35, 147 (1962).
[18] Another perhaps more direct way of proving this fact is to use Eq. (25) and Theorem A3. 2 of Ref. 11.
[19] Quasilocal operators can be defined as those operators which satisfy a spacelike cluster property given by Haag in Ref. 16. See also Eq. (A1) in the Appendix with $s = 0$.
[20] K. Hepp, in 1965 Brandeis University Summer Institute in Theoretical Physics, *Axiomatic Field Theory*, edited by M. Chretien and S. Deser (Gordon and Breach, New York, 1966), Vol. I.
[21] K. Hepp, Commun. Math. Phys. 1, 95 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Electromagnetic Radiation in Curved Spaces*

Richard Sigal
Belfer Graduate School of Science, Yeshiva University, New York, N.Y., 10033
(Received April 20 1971)

The classes of trajectories of charged particles, for which the radiation reaction force vanishes, in curved spaces are examined. The trajectories in the Robertson–Walker cosmologies are given in detail and it is shown that the only reaction-free geodesic trajectories are those at rest relative to the local matter density.

I. INTRODUCTION

The question of which classes of trajectories of charged particles in a curved space are free of a radiation reaction force is an old one in general relativity. We examine this question through the use of the conformal symmetries of Maxwell's equations in cases where the metric is taken as given.

A reaction-free trajectory shall be defined as one for which the radiation reaction force, as originally defined by Dirac, vanishes. We have chosen to study radiation from this aspect to avoid global problems such as the structure of infinity. The meaning of this condition in flat space has been widely discussed in relation to the problem of a uniformly accelerated charged particle.[1] From these discussions it is clear that this radiation reaction formalism and definitions of radiation as a nonvanishing Poynting vector at infinity are not unambiguously equivalent, and also that the definitions at infinity themselves suffer from problems.

We have done detailed analysis of the structure of the reaction-free trajectories for the cases of Robertson–Walker cosmologies, where the analysis is somewhat simplified due to their conformal flatness.[2] These models, although simple, are of considerable interest as they represent a close approximation to the large scale structure of the universe as observed.[3]

We have shown in the Robertson–Walker cosmologies that the only reaction-free geodesics are those followed by particles at rest relative to the local matter density.

II. CONFORMAL STRUCTURE

Maxwell's equations[4] for the field $F_{\mu\nu}$, current j_μ, and metric $g_{\mu\nu}$ are given by

$$g^{\mu\alpha}F_{\mu\nu;\alpha} = 4\pi j_\nu, \tag{1}$$

$$F_{[\mu\nu;\lambda]} = F_{[\mu\nu,\lambda]} = 0, \tag{2}$$

where square brackets mean antisymmetrization, ";" covariant derivative with respect to $g_{\mu\nu}$, and "," ordinary derivative. The metric is taken to have signature -2.

Under a conformal transformation

$$g_{\mu\nu} \to \bar{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}, \quad g^{\mu\nu} \to \bar{g}^{\mu\nu} = \Omega^{-2} g^{\mu\nu}$$

Maxwell equations are invariant if we take[5]

$$\bar{F}_{\mu\nu} = F_{\mu\nu}, \quad \bar{j}_\mu = \Omega^{-2} j_\mu.$$

We can then define the densities

$$\mathfrak{F}^{\mu\nu} = \sqrt{-g}F^{\mu\nu}, \quad J^\nu = \sqrt{-g}j^\nu,$$

where g = determinant $g_{\mu\nu}$ and

$$\sqrt{-\bar{g}} = \Omega^4\sqrt{-g}.$$

We then have

$$\bar{\mathfrak{F}}^{\mu\nu} = \mathfrak{F}^{\mu\nu}, \quad \bar{J}^\nu = J^\nu$$

and Eq. (1) reads

$$\bar{\mathfrak{F}}^{\mu\nu}{}_{,\mu} = \mathfrak{F}^{\mu\nu}{}_{,\mu} = \bar{J}^\nu = J^\nu,$$

which gives the conservation law

$$\bar{J}^\nu{}_{,\nu} = J^\nu{}_{,\nu} = 0$$

and thus gives the total conserved electric charge Q as

$$Q = \int J^\mu dS_\mu,$$

which is seen to be a conformal invariant.

This means that under a conformal mapping currents that represent point charges with charge q_i

and trajectories Z^{μ}_{i} transform into currents that also represent charges q_i and trajectories Z^{μ}_{i}. It is important to note that in general the trajectories so associated do not represent the same physical situation, as many physical parameters used to describe a trajectory are changed.

An important property of conformal transformations of the metric is that, as they locally leave the direction of time unchanged and leave the light cone invariant, retarded or advanced solutions of Maxwell's equations will remain such under them.

III. CONFORMAL MOTIONS

The covariance of Maxwell's equations under conformal transformations of the metric is closely related to their invariance, like all zero rest mass fields, under the group of conformal motions of a space.[6] The conformal motions[7] of a space V_4 are the group of automorphisms of V_4 onto V_4 that do not change the angle between two directions at a point. For an infinitesimal transformation

$$X'^{\alpha} = X^{\alpha} + \xi^{\alpha}(x),$$

this is equivalent to

$$\underset{\xi^{\alpha}}{\mathcal{L}} g_{\mu\nu} = \phi(x) g_{\mu\nu}, \tag{3}$$

where $\underset{\xi^{\alpha}}{\mathcal{L}}$ is the Lie derivative with respect to ξ^{α}. A vector field that satisfies (3) is a conformal Killing vector and (3) is the conformal Killing equation. Expanded out, (3) reads

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} - \phi g_{\mu\nu} = 0 \tag{4}$$

and contracting with $g^{\mu\nu}$ on (4) gives

$$\phi = \tfrac{1}{2}\xi^{\sigma}{}_{;\sigma}$$

If we define the quantity

$$\mathcal{G}_{\mu\nu} = g_{\mu\nu}/(-g)^{1/4},$$

then (4) can be written as

$$\underset{\xi^{\alpha}}{\mathcal{L}} \mathcal{G}_{\mu\nu} = 0.$$

As $\overline{\mathcal{G}}_{\mu\nu} = \mathcal{G}_{\mu\nu}$, we can see that conformal Killing vectors are invariant under conformal transformations of the metric.

Conformal motions can be characterized by a class of curves, conformal circles. If for the curve $Z^{\mu}(s)$, where s is an affine parameter, and its "unit tangent vector"

$$v^{\alpha} \equiv \frac{dZ^{\alpha}}{ds} \equiv \dot{Z}^{\alpha}, \quad v^{\alpha}v_{\alpha} = \epsilon = \begin{cases} 1 \\ -1 \end{cases}$$

we define

$$u^{\alpha} \equiv \epsilon \ddot{v}^{\alpha} + \dot{v}^{\rho}\dot{v}_{\rho}v^{\alpha} + P^{\alpha}_{\sigma}v^{\sigma} - \epsilon P^{\rho}_{\sigma}v_{\rho}v^{\sigma}v^{\alpha}, \tag{5}$$

then the equation $u^{\alpha} = 0$ defines a conformal circle. In (5)

$$P_{\rho\sigma} \equiv \tfrac{1}{2}R_{\rho\sigma} - \tfrac{1}{12}g_{\rho\sigma}R, \quad R_{\rho\sigma} \equiv R^{\alpha}_{\cdot\rho\alpha\sigma},$$

where $R^{\alpha}_{\rho\omega\sigma}$ is the Riemann Christoffel tensor and our sign convention is given by

$$2K^{\rho}_{[\mu;\nu]} = R^{\rho}_{\cdot\sigma\mu\nu}K^{\sigma}.$$

It can be shown[8] that for ξ^{α} to generate a conformal motion, it is necessary and sufficient that it take conformal circles into conformal circles.

If we perform a conformal transformation of the metric $g_{\mu\nu} \to \Omega^2 g_{\mu\nu}$, then for a timelike curve $Z^{\mu}(s)$ we have a new proper time $\bar{s}$ and[9]

$$\bar{v}^{\mu} \equiv dZ^{\mu}(\bar{s})/d\bar{s} = \Omega^{-1}v^{\mu}.$$

We also have[10]

$$\bar{P}_{\mu\alpha} = P_{\mu\alpha} + \Omega^{-1}\Omega_{;\mu\alpha} - 2\Omega^{-2}\Omega_{;\mu}\Omega_{;\alpha} + \tfrac{1}{2}\Omega^{-2}g_{\mu\alpha}g^{\rho\sigma}\Omega_{;\rho}\Omega_{;\sigma}.$$

From these a straightforward calculation gives

$$\bar{u}^{\alpha} = \Omega^{-3}u^{\alpha}.$$

In flat space the timelike conformal circles are the curves of uniform acceleration. In any empty space ($R_{\mu\nu} = 0$) they are the curves with constant first curvature and vanishing second curvature,[11] and are generally called geodesic circles. This definition of geodesic circles holds also in spaces where $R_{\mu\nu} \neq 0$ but then the two types of curves are distinct. Geodesic circles are the trajectories that Rindler[12] has identified as the uniformally accelerated trajectories in a curved space.

Congruences of conformal circles are generated by the orbits of some conformal transformations. For timelike conformal circles, if W^{μ} is our conformal Killing field, then define v^{μ} such that

$$W^{\mu} = \alpha v^{\mu}, \quad v^{\mu}v_{\mu} = 1.$$

Then Eq. (4) for W^{μ} gives, using $v_{\alpha;\beta}v^{\alpha} = 0$,

$$\dot{\alpha} \equiv \alpha_{,\mu}v^{\mu} = \phi/2,$$
$$v^{\mu}{}_{;\nu}v^{\nu} \equiv \dot{v}_{\mu} = -1/\alpha(\alpha_{;\mu} - \dot{\alpha}v_{\mu}),$$
$$\ddot{v}_{\mu} = -1/\alpha(\alpha_{;\mu\nu}v^{\nu} - \ddot{\alpha}v_{\mu})$$

Putting these conditions into Eq. (5) and setting it equal to zero, we get that *the necessary and sufficient condition for a conformal Killing field to generate a conformal circle is*

$$\frac{1}{\alpha}\alpha_{;\mu\nu}v^{\nu} - \tfrac{1}{2}R_{\mu\nu}v^{\nu} = v_{\mu}\left(\frac{1}{\alpha}\alpha_{;\rho\sigma}v^{\sigma}v^{\rho} - \tfrac{1}{2}R_{\rho\sigma}v^{\sigma}v^{\rho}\right).$$

For flat space the group of conformal motion has 15 parameters and in the standard representation the generators are divided up into the group of translations (four parameters), the homogeneous

Lorentz group (six parameters), the group of scale transformations (one parameter), and the group of uniform accelerations (four parameters).[13]

The translations and the uniform accelerations both contain timelike infinitesimal generators and it is easy to show that these generators satisfy the previous conditions.

IV. THE RADIATION REACTION

The electromagnetic radiation reaction terms in an arbitrary given curved space were first calculated by DeWitt and Brehme[14] in 1960 using bitensors and were recalculated by Hobbs[15] in 1968 using a vierbein formalism. The Hobbs result differs from DeWitt and Brehme due to a mistake discovered by Hobbs in their work. We will use Hobbs result, which is consistent with the conformal structure of Maxwell's equations while the result of DeWitt and Brehme is not.

DeWitt and Brehme and Hobbs both calculate the radiation reaction force for a single point particle by generalizing Dirac's classic paper[16] to curved space–times. As in Dirac's work, the total electromagnetic field $F_{\alpha\beta}$ at a point near the source is broken up into several terms. $F^{\rm ret}_{\alpha\beta}$ is a formal retarded solution for the given source distribution and $F^{\rm adv}_{\alpha\beta}$ is the corresponding advanced solution. These terms are then used to define

$$F_{\alpha\beta} \equiv F^{\rm in}_{\alpha\beta} + F^{\rm ret}_{\alpha\beta} \equiv F^{\rm out}_{\alpha\beta} + F^{\rm adv}_{\alpha\beta},$$

$$\tilde{F}_{\alpha\beta} \equiv \tfrac{1}{2}(F^{\rm ret}_{\alpha\beta} + F^{\rm adv}_{\alpha\beta}),$$

$$F^{\rm rad}_{\alpha\beta} \equiv F^{\rm ret}_{\alpha\beta} - F^{\rm adv}_{\alpha\beta},$$

$$F^{\rm free}_{\alpha\beta} \equiv \tfrac{1}{2}(F^{\rm in}_{\alpha\beta} + F^{\rm out}_{\alpha\beta}) = F^{\rm in}_{\alpha\beta} + \tfrac{1}{2}F^{\rm rad}_{\alpha\beta}$$
$$= F^{\rm out}_{\alpha\beta} - \tfrac{1}{2}F^{\rm rad}_{\alpha\beta},$$

$$F_{\alpha B} = \tilde{F}^{\rm free}_{\alpha\beta} + \tilde{F}_{\alpha\beta}.$$

$F^{\rm in}$, $F^{\rm out}$, $F^{\rm free}$, and $F^{\rm rad}$ are singularity free at the source while $\tilde{F}_{\alpha\beta}$, $F^{\rm ret}_{\alpha\beta}$, and $F^{\rm adv}_{\alpha\beta}$ satisfy Maxwell's equations for the given source.

These fields are expanded around the singular source to get their expressions on a world tube constructed about it. They expand $T_{\mu\nu}$, the Maxwell stress energy tensor, in terms of $\tilde{F}_{\alpha\beta}$ and $F^{\rm free}_{\alpha\beta}$ and use $T^{\mu\nu}{}_{;\nu} = 0$ to obtain equations of motion. This procedure involves at the end a renormalization to kill the divergent terms.

The results of these calculations of Hobbs[17] for a particle with trajectory $Z^\mu(s)$ and charge e with $c = 1$ are

$$F^{\rm rad}_{\alpha\beta} = \tfrac{4}{3}e(\dot{Z}_\beta \dddot{Z}_\alpha - \ddot{Z}_\beta \ddot{Z}_\alpha) + \tfrac{2}{3}eR_{\alpha\gamma}\dot{Z}^\gamma \dot{Z}_\beta$$
$$- \tfrac{2}{3}eR_{\beta\gamma}\dot{Z}^\gamma \dot{Z}_\alpha + e\int_{-\infty}^{\infty} \epsilon(s-s') f_{\alpha\beta\gamma}\dot{Z}^{\gamma\prime} ds'$$

$$M\ddot{Z}_\alpha = eF^{\rm free}_{\alpha\beta}\dot{Z}^\beta + \tfrac{1}{2}e^2\dot{Z}^\beta \int_{-\infty}^{\infty} f_{\alpha\beta\gamma}\dot{Z}^\gamma(s')ds' \quad (6)$$

$$= eF^{\rm in}_{\alpha\beta}\dot{Z}^\beta + \tfrac{2}{3}e^2(\ddot{Z}_\alpha + \dot{Z}_\alpha \ddot{Z}^\beta \ddot{Z}_\beta) + \tfrac{1}{3}e^2 R_{\alpha\gamma}\dot{Z}^\gamma$$
$$- \tfrac{1}{3}e^2 \dot{Z}_\alpha R_{\beta\gamma}\dot{Z}^\beta \dot{Z}^\gamma + e^2\dot{Z}^\beta \int_{-\infty}^{s} f_{\alpha\beta\gamma'}\dot{Z}^{\gamma\prime}(s')ds'$$
$$= eF^{\rm in}_{\alpha\beta}\dot{Z}^\beta + \tfrac{2}{3}e^2 u_\alpha + e^2\dot{Z}^\beta \int_{-\infty}^{s} f_{\alpha\beta\gamma'}\dot{Z}^{\gamma\prime}ds',$$

where $\epsilon(\tau) = \Theta(\tau) - \Theta(\tau')$,

$$\Theta(\tau) = \begin{cases} 0, & \tau < 0 \\ 1, & \tau > 0 \end{cases},$$

and $f_{\alpha\beta\gamma'}$ is a bitensor that is zero in flat space but occurs in curved spaces as in general the Green's function for the vector wave equation is nonzero not only on the light cone but also inside it. The integral term is generally called the tail term. Some comment is necessary on these results.

In the calculations of DeWitt and Brehme and Hobbs, formal retarded and advanced fields are written down for an arbitrary space. These fields at a point depend only on the source distribution inside the past or future light cone of the point. But it has been shown by Penrose[18] that if there exists a particle horizon in space, purely retarded solutions to Maxwell's equations do not exist for general source distributions, and, if there exists an event horizon, the same is true for advanced solutions. Thus unlike the case in flat space the field $F^{\rm in}_{\alpha\beta}$ is not completely at our disposal.

Also unlike in flat space the radiation reaction term is not given completely by $F^{\rm rad}_{\alpha\beta}$, but has a contribution in the tail from a term that in flat space is totally divergent. It is perhaps remarkable that it is just this additional term that gives us an equation of motion that depends at least formally only on the past behavior of the particle and not also on its future.

Rohrlich has suggested the introduction of an already renormalized Lagrangian for Maxwell's equations so as to avoid divergent terms from the beginning.[19] If this is done, the radiation reaction term is just given by $\dot{Z}^\beta F^{\rm rad}_{\alpha\beta}$. Thus the equation of motion obtained is somewhat unsatisfactory, being completely noncausal, not only in a "small region" as is true even in flat space, but over all space–time.

We can use conformal covariance to obtain information about the tail term. If the advanced and retarded potentials are unique, then $\overline{F}^{{\rm rad}\,\alpha\beta} = \Omega^{-4}F^{{\rm rad}\,\alpha\beta}$ and it is easy to calculate that

$$U^{\alpha\beta} \equiv \tfrac{4}{3}e[(\dot{Z}^\beta \dddot{Z}^\alpha - \ddot{Z}^\beta \ddot{Z}^\alpha) + \tfrac{1}{2}R^\alpha_\gamma \dot{Z}^\gamma \dot{Z}^\beta,$$
$$- \tfrac{1}{2}R^\beta_\gamma \dot{Z}^\gamma \dot{Z}^\alpha] = \Omega^4 \overline{U}^{\alpha\beta}.$$

Thus

$$e\int_{-\infty}^{\infty} \epsilon(s-s') f^{\alpha\beta}_{\gamma'} \dot{Z}^{\gamma\prime}(s')ds'$$
$$= \Omega^{+4} e\int_{-\infty}^{\infty} \epsilon(s-s') f^{\alpha\beta}_{\gamma'} \overline{\dot{Z}}^{\gamma\prime}(\bar{s}')d\bar{s}'$$

and is only a function of the conformal curv-

ature. Since we can take $\Omega = 1$ for $s' > s$, the $\int_{-\infty}^{s} f^{\alpha}_{\beta\gamma'} \dot{Z}^{\gamma'}(s')ds'$ also depends only on the conformal curvature.

We therefore have that *the tail is zero in conformally flat spaces*, where there is, of course, no problem about the uniqueness of retarded and advanced fields. Hobbs has also obtained this result by direct calculation.[20]

Equation (6) shows that it is only the tail term that prevents conformal circles from being the curves of vanishing radiation reaction. For some conformal circles the tail term may vanish separately, but its general complexity prevents its evaluation except in the most elementary cases. It has been shown for example, by DeWitt and DeWitt,[21] that the tail term vanishes for a particle at rest in a weak static field.

V. ROBERTSON–WALKER COSMOLOGIES

We will take the Robertson–Walker line element in the form

$$ds^2 = dt^2 - S^2(t)du^2,$$

where

$$du^2 = \frac{\delta_{ij}dx^i dx^j}{[1 + \frac{1}{4}k\delta_{ij}x^i x^j]^2}, \qquad k = \begin{cases} 0 \\ \pm 1 \end{cases}.$$

In this metric form letting $r^2 = \delta_{ij}x^i x^j$, we get[22]

$$R_{00} = 3S''/S,$$

$$R_{11} = R_{22} = R_{33} = \frac{-1}{[1 + \frac{1}{4}kr^2]^2}(SS'' + 2S'^2 + 2k),$$

where

$$S' \equiv \frac{dS}{dt},$$

$$R_{\mu\nu} = 0, \quad \mu \neq \nu,$$

$$R = \frac{6S''}{S} + \frac{6S'^2}{S^2} + \frac{6k}{S^2},$$

$$P_{00} = P^0_0 = \frac{S''}{S} - \frac{S'^2}{2S^2} - \frac{k}{2S^2},$$

$$P_{11} = P_{22} = P_{33} = -\frac{[S'^2 + k]}{2[1 + \frac{1}{4}kr^2]^2},$$

$$P^1_1 = P^2_2 = P^3_3 = \frac{[S'^2 + k]}{2S^2}.$$

Therefore, for timelike trajectories, we get for u^{α}

$$u^0 = \ddot{v}^0 + \dot{v}^{\gamma}\dot{v}_{\gamma}v^0 + v^0(1 - v^0v_0)(P^0_0 - P^1_1),$$

$$u^i = \ddot{v}^i + \dot{v}^{\gamma}\dot{v}_{\gamma}v^0 + v^iv^0v_0(P^1_1 - P^0_0),$$

where

$$(P^1_1 - P^0_0) = \frac{S''}{S} - \frac{S'^2}{S^2} - \frac{k}{S^2}.$$

We first look at the conditions that

$$\frac{S''}{S} - \frac{S'^2}{S^2} - \frac{k}{S^2} = 0. \tag{7}$$

This implies $d/dt(S''/S) = 0$ and thus

$$S'' = -CS \tag{8}$$

and

$$CS^2 + S'^2 + k = 0.$$

On the other hand this set of equations implies (7). In the Robertson–Walker metric, Eqs. (7) are just the condition that a space be of constant curvature C.[23]

That is,

$$R_{\mu\nu\rho\sigma} = -C(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}).$$

As all spaces of constant curvature can be put in the Robertson–Walker form at least locally, we have the theorem that *all conformal circles are geodesic circles in spaces of constant curvature and it is only in these cosmologies for which this is true.*

We have therefore proven that *all geodesics in spaces of constant curvature are reaction-free.*

For cosmologies where $P^0_0 - P^1_1 \neq 0$ then the radiation reaction force can vanish for geodesics only if

$$v^i = 0,$$

$$1 - v^0v_0 = 0.$$

That is, $v^{\mu} = \delta^{\mu}_0$. These trajectories are just the streamlines of the fluid which is the source for the Robertson–Walker metric.

We thus have the result that *for Robertson–Walker cosmologies of nonconstant curvature the only reaction-free geodesic trajectories are those that are at rest relative to the local matter density.*

This result can be understood as a consequence of the existence of an absolute time and an absolute rest frame in these Robertson–Walker cosmologies, unlike the cases of spaces of constant curvature, where there exists the full ten-parameter isometry group, and thus a full generalized Lorentz group.

Besides geodesics, another class of curves of interest are those generated by timelike Killing fields if they exist. There are three nonflat Robertson–Walker cosmologies of this type. Two are the static cosmologies ($S' = 0$ and $k = \pm 1$) and the third is the de Sitter universe ($S' \neq 0$) which is also of constant curvature.[23]

In the first two cases the Killing field is given by $W^{\mu} = \delta^{\mu}_0$ so that the curves are the already discussed reaction-free geodesics.

In the de Sitter case, $k = 0$ and $S^2 = e^{2at/b}$. The Killing field is not everywhere timelike, so it is

not strictly stationary but is classified as such by many. The field is given by

$$W^0 = a, \quad W^i = ax^i/b$$

and its associated curves have "unit tangent vectors"

$$v^0 = b/(b^2 - e^{2t/b}r^2)^{1/2},$$
$$v^i = -x^i/(b^2 - e^{2t/b}r^2)^{1/2}.$$

From these we can easily calculate that the v^μ *are conformal circles and thus a particle following them will not experience a radiation reaction force.*

ACKNOWLEDGMENTS

I would like to thank Professor Emil Grgin and Professor Roger Penrose for valuable discussions on this work.

* This work was in part supported by NSF grant GP 8850 and NASA grant NGR 33-023-018.
1 J. L. Anderson, *Relativity Physics* (Academic, New York, 1967), p. 245; H. Bondi and T. Gold, Proc. Roy. Soc. (London) **A229**, 416 (1955); T. Fulton and F. Rohrlich, Ann. Phys. (N.Y.) **9**, 499 (1960); C. Leibovitz and A. Peres, Ann. Phys. **25**, 400 (1963); F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965).
2 R. Penrose, *Relativity Groups and Topology* (Gordon and Breach, New York, 1964), p. 580.
3 H. P. Robertson and T. Noonan, *Relativity and Cosmology* (Saunders, Philadelphia, 1968).
4 J. L. Synge, *Relativity. The General Theory* (North-Holland, Amsterdam, 1966), p. 356.
5 R. Penrose, Ref. 2, p. 566; P. C. Peters, J. Math. Phys. **10**, 1216 (1969).
6 F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965), p. 104.
7 K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland, Amsterdam, 1955), p. 32.
8 K. Yano, Ref. 7, p. 159.
9 J. A. Schouten, *Ricci-Calculus* (Springer-Verlang, Berlin, 1954), p. 309.
10 R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
11 J. A. Schouten, Ref. 9, p. 331.
12 W. Rindler, Phys. Rev. **119**, 2082 (1960).
13 E. Grgin, J. Math. Phys. **9**, 1595 (1968).
14 B. DeWitt and R. Brehme, Ann. Phys. (N.Y.) **9**, 220 (1960).
15 J. M. Hobbs, Ann. Phys. **47**, 141 (1968).
16 P. A. M. Dirac, Proc. Roy Soc. (London) **A167**, 148 (1938).
17 See Ref. 15. (My results differ in sign from those of Hobbs due to different sign conventions.)
18 R. Penrose, Ref. 2, p. 574.
19 F. Rohrlich, Ref. 6, p. 220.
20 J. M. Hobbs, Ann. Phys. **47**, 166 (1968).
21 C. M. DeWitt and B. S. DeWitt, Physics **1**, No. 1, 3 (1964).
22 Ref. 3, p. 343.
23 Ref. 3, p. 346.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Combined Neutrino-Gravitational Fields in General Relativity*

D. Trim† and J. Wainwright
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada.
(Received 24 May 1971)

The combined gravitational-neutrino field equations are solved, subject only to the restriction that the energy-flow vector of the neutrino field be timelike or null. The principal null congruence (pnc) of the neutrino field is necessarily geodesic and shear-free, and coincides with a repeated pnc of the gravitational field, which is thus algebraically special. The twist of the neutrino pnc plays an important role, and is zero if and only if the neutrino energy-flow vector is null, in which case the neutrino field represents pure radiation.

1. INTRODUCTION

In this paper we consider gravitational fields which have as source a neutrino field.[1] We impose the restriction that the energy flow vector of the neutrino field with respect to an arbitrary observer u^a, defined to be $T_{ab}u^b$ in terms of the neutrino energy tensor[1] T_{ab}, be a timelike or null vector; i.e.,

$$T^c_a T_{cb} u^a u^b \geq 0, \quad T_{ab}u^b \neq 0 \tag{1.1}$$

for all timelike (unit) vectors u^a. The condition (1.1) immediately implies[2] that

$$T_{ab}u^a u^b \neq 0, \tag{1.2}$$

so that the field has nonzero energy density with respect to all observers.

The geometrical effect of assumption (1.1) is to impose strong restrictions on the principal null congruence[1] (pnc) k^a of the neutrino field, namely that it be geodesic and shear-free:

$$k_{a;b}k^b = 0, \quad [k^a_{\ ;a}]^2 = 2k_{(a;b)}k^{a;b}, \tag{1.3}$$

as was shown in an earlier paper.[3] In constructing the general line element it is natural, in view of the work of Robinson and others[4] to choose a coordinate system $(u, r, x, y) = (x^1, x^2, x^3, x^4)$ based on this particular congruence, with respect to which

$$k^a = \delta^a_2. \tag{1.4}$$

By virtue of (1.3) we can further specialize the coordinates to that the line element assumes the form

$$ds^2 = -\tfrac{1}{2}G\bar{G}dzd\bar{z} + 2(k_a dx^a)[dr - \tfrac{1}{2}\bar{W}dz - \tfrac{1}{2}W d\bar{z} - U(k_b dx^b)], \tag{1.5}$$

not strictly stationary but is classified as such by many. The field is given by

$$W^0 = a, \quad W^i = ax^i/b$$

and its associated curves have "unit tangent vectors"

$$v^0 = b/(b^2 - e^{2t/b}r^2)^{1/2},$$
$$v^i = -x^i/(b^2 - e^{2t/b}r^2)^{1/2}.$$

From these we can easily calculate that the v^μ *are conformal circles and thus a particle following them will not experience a radiation reaction force.*

ACKNOWLEDGMENTS

I would like to thank Professor Emil Grgin and Professor Roger Penrose for valuable discussions on this work.

* This work was in part supported by NSF grant GP 8850 and NASA grant NGR 33-023-018.
1 J. L. Anderson, *Relativity Physics* (Academic, New York, 1967), p. 245; H. Bondi and T. Gold, Proc. Roy. Soc. (London) **A229**, 416 (1955); T. Fulton and F. Rohrlich, Ann. Phys. (N.Y.) **9**, 499 (1960); C. Leibovitz and A. Peres, Ann. Phys. **25**, 400 (1963); F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965).
2 R. Penrose, *Relativity Groups and Topology* (Gordon and Breach, New York, 1964), p. 580.
3 H. P. Robertson and T. Noonan, *Relativity and Cosmology* (Saunders, Philadelphia, 1968).
4 J. L. Synge, *Relativity. The General Theory* (North-Holland, Amsterdam, 1966), p. 356.
5 R. Penrose, Ref. 2, p. 566; P. C. Peters, J. Math. Phys. **10**, 1216 (1969).
6 F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965), p. 104.
7 K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland, Amsterdam, 1955), p. 32.
8 K. Yano, Ref. 7, p. 159.
9 J. A. Schouten, *Ricci-Calculus* (Springer-Verlang, Berlin, 1954), p. 309.
10 R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
11 J. A. Schouten, Ref. 9, p. 331.
12 W. Rindler, Phys. Rev. **119**, 2082 (1960).
13 E. Grgin, J. Math. Phys. **9**, 1595 (1968).
14 B. DeWitt and R. Brehme, Ann. Phys. (N.Y.) **9**, 220 (1960).
15 J. M. Hobbs, Ann. Phys. **47**, 141 (1968).
16 P. A. M. Dirac, Proc. Roy Soc. (London) **A167**, 148 (1938).
17 See Ref. 15. (My results differ in sign from those of Hobbs due to different sign conventions.)
18 R. Penrose, Ref. 2, p. 574.
19 F. Rohrlich, Ref. 6, p. 220.
20 J. M. Hobbs, Ann. Phys. **47**, 166 (1968).
21 C. M. DeWitt and B. S. DeWitt, Physics **1**, No. 1, 3 (1964).
22 Ref. 3, p. 343.
23 Ref. 3, p. 346.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Combined Neutrino-Gravitational Fields in General Relativity*

D. Trim† and J. Wainwright
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada.
(Received 24 May 1971)

The combined gravitational-neutrino field equations are solved, subject only to the restriction that the energy-flow vector of the neutrino field be timelike or null. The principal null congruence (pnc) of the neutrino field is necessarily geodesic and shear-free, and coincides with a repeated pnc of the gravitational field, which is thus algebraically special. The twist of the neutrino pnc plays an important role, and is zero if and only if the neutrino energy-flow vector is null, in which case the neutrino field represents pure radiation.

1. INTRODUCTION

In this paper we consider gravitational fields which have as source a neutrino field.[1] We impose the restriction that the energy flow vector of the neutrino field with respect to an arbitrary observer u^a, defined to be $T_{ab}u^b$ in terms of the neutrino energy tensor[1] T_{ab}, be a timelike or null vector; i.e.,

$$T^c_a T_{cb} u^a u^b \geq 0, \quad T_{ab}u^b \neq 0 \tag{1.1}$$

for all timelike (unit) vectors u^a. The condition (1.1) immediately implies[2] that

$$T_{ab}u^a u^b \neq 0, \tag{1.2}$$

so that the field has nonzero energy density with respect to all observers.

The geometrical effect of assumption (1.1) is to impose strong restrictions on the principal null congruence[1] (pnc) k^a of the neutrino field, namely that it be geodesic and shear-free:

$$k_{a;b}k^b = 0, \quad [k^a_{\ ;a}]^2 = 2k_{(a;b)}k^{a;b}, \tag{1.3}$$

as was shown in an earlier paper.[3] In constructing the general line element it is natural, in view of the work of Robinson and others[4] to choose a coordinate system $(u, r, x, y) = (x^1, x^2, x^3, x^4)$ based on this particular congruence, with respect to which

$$k^a = \delta^a_2. \tag{1.4}$$

By virtue of (1.3) we can further specialize the coordinates to that the line element assumes the form

$$ds^2 = -\tfrac{1}{2}G\bar{G}dzd\bar{z} + 2(k_a dx^a)[dr - \tfrac{1}{2}\bar{W}dz - \tfrac{1}{2}W d\bar{z} - U(k_b dx^b)], \tag{1.5}$$

where

$$k_a dx^a = du - \tfrac{1}{2}\bar{Q}dz - \tfrac{1}{2}Q d\bar{z}, \quad z = x + iy,$$

as was shown by Robinson.[4] The function U which is real, and G, W which are complex, depend on all the coordinates, while the complex function Q is independent of r. The coordinate r here is an affine parameter along the pnc.

If, in addition, the congruence k^a is twist-free, one can choose the coordinate u to be constant on the corresponding family of null hypersurfaces so that $k_a = \delta_a^1$, which entails

$$Q = 0. \tag{1.6}$$

In Sec. 2 the restrictions which are imposed on the Ricci tensor due to the presence of the neutrino field are derived. The further simplification of the line element (1.5) depends on whether or not the pnc of the neutrino field has zero expansion. In the expansion-free case, which is discussed in Sec. 3, the pnc is necessarily twist-free and the neutrino field represents pure radiation.[5] Particular solutions in this class have been described by a number of authors.[5-7] The general line element in the case of nonzero expansion is described in Sec. 4. The solutions of this class in which the pnc is twist-free have been given previously.[7] The twisting solutions, however, are new to the best of the authors' knowledge, and establish the existence of combined neutrino-gravitational fields in which the neutrino field does not represent pure radiation. In the Appendix, a summary of the calculations leading to the twisting line element is given, using the Newman–Penrose[8] formalism. In order to increase their applicability, these calculations are given in a form which includes as special cases, the general algebraically special vacuum gravitational field whose repeated pnc (which is necessarily geodesic and shear-free) is twisting,[4] as well as a wide class of combined electromagnetic–gravitational fields,[9] although these are not dealt with explicitly in this paper.

2. NEUTRINO FIELDS AND THE RICCI TENSOR

A neutrino field in curved space-time can be described by a 2-spinor field $\phi^A(x^a)$ which satisfies the field equations[10]

$$\sigma^{aAA'}\phi_{A;a} = 0. \tag{2.1}$$

We introduce a spinor dyad[8] $\{o^A, \iota^A\}$ in space–time with o^A chosen to be parallel to ϕ^A:

$$\phi^A(x^a) = \phi(x^a)o^A, \tag{2.2}$$

where $\phi(x^a)$ is a complex function. The corresponding null tetrad[8] $\{k^a, n^a, m^a, \bar{m}^a\}$ is thus adapted to the neutrino field in the sense that $k^a = \sigma^{aAA'}o_A o_{A'}$ is tangent to the pnc of the neutrino field.

For a neutrino field subject to (1.1), it has been proved that the energy tensor assumes the form[11]

$$T_{ab} = Ak_a k_b + B[4k_{(a}n_{b)} - g_{ab}] + 2Ek_{(a}m_{b)} + 2\bar{E}k_{(a}\bar{m}_{b)}, \tag{2.3}$$

with

$$AB - E\bar{E} \geq 0 \tag{2.4}$$

in terms of the adapted null tetrad. The quantities A, B, and E are defined[12] in terms of the neutrino field and the null tetrad; in fact B is proportional to the twist ω of the neutrino pnc:

$$B = -2\,\phi\,\bar{\phi}\,\omega. \tag{2.5}$$

The gravitational field equations

$$R_{ab} = -T_{ab} \tag{2.6}$$

in terms of suitable units, thus impose the following algebraic restrictions on the Ricci tensor:

$$R_{ab}k^a k^b = R_{ab}k^a m^b = R_{ab}m^a m^b = 0, \tag{2.7}$$

$$\Longleftrightarrow$$

$$k_{[a}R_{b][c}k_{d]} = \lambda\, k_{[a}g_{b][c}k_{d]},$$

and

$$R_a^a = 0. \tag{2.8}$$

These conditions, together with (1.3), have two important consequences. First, *the neutrino principal null direction (pnd) k^a must be a repeated pnd[13] of the gravitational field, which is thus algebraically special.* This is a straightforward application of a result due to Kundt and Trümper.[14] Secondly, *the dependence of the metric tensor components and of the neutrino field on the affine parameter r along the pnc is completely determined.*

3. NONEXPANDING PNC

If the expansion of the pnc vanishes, then it follows[15] that the twist also does. Equations (2.4) and (2.5) then give $E = 0$, and the energy tensor of the neutrino field reduces to

$$T_{ab} = Ak_a k_b. \tag{3.1}$$

Consequently, the field equations (2.6) further restrict the Ricci tensor by

$$R_{ab}k^a n^b = R_{ab}m^a n^b = 0,$$

which, together with (2.7), (2.8), (1.6), and the neutrino field equations enable us to reduce the line element (1.5) to the form

$$ds^2 = -\tfrac{1}{2}dz\,d\bar{z} + 2du[dr - \tfrac{1}{2}W(dz + d\bar{z}) - U\,du],$$

where

$$U = K + r\frac{\partial W}{\partial x}, \qquad 0 = \frac{\partial W}{\partial r} = \frac{\partial K}{\partial r},$$

and the real function W must satisfy

$$\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} = 0.$$

The metric tensor and the neutrino field are related by the differential equation

$$\frac{\partial^2 K}{\partial x^2} + \frac{\partial^2 K}{\partial y^2} + \left[2W\frac{\partial}{\partial x} + \frac{\partial W}{\partial x} - \frac{\partial}{\partial u}\right]\left(\frac{\partial W}{\partial x}\right) - \left(\frac{\partial W}{\partial y}\right)^2 = -\tfrac{1}{2}A, \quad (3.2)$$

where

$$A = 2i\left(\phi\frac{\partial \bar{\phi}}{\partial u} - \bar{\phi}\frac{\partial \phi}{\partial u} - i\phi\,\bar{\phi}\frac{\partial W}{\partial y}\right),$$

and the neutrino field function ϕ, defined by (2.2), is an arbitrary function of u alone. This result was obtained by a straightforward application of the Newman–Penrose[8] formalism. Alternatively it could be obtained by appropriately specializing the line element of Kundt.[16]

We note that the Weyl conform tensor is of Petrov–Penrose[13] type $\{31\}$ or $\{4\}$, or is zero, the last two possibilities occurring if and only if W satisfies $\partial^2 W/\partial z^2 = 0$. In this case, W can be transformed to zero, and the line element describes plane-fronted gravitational waves with parallel rays[16] together with a neutrino radiation field. Further, (3.2) can be integrated to

$$K = -\tfrac{1}{2}A\,z\bar{z} + F(z,u) + \bar{F}(\bar{z},u),$$

since A is now a function of u alone. This class of solutions has been given by Audretsch and Graf.[5]

4. EXPANDING PNC

In this case, Restrictions (2.7) and (2.8) on the Ricci tensor, together with (1.3), determine the r-dependence of the line element (1.5) as follows:

$$G\bar{G} = P^{-2}(r^2 + \Sigma^2), \qquad W = W^0 - r\frac{\partial Q}{\partial u}, \quad (4.1)$$

$$U = U^0 + r\frac{\partial(\ln P)}{\partial u} + [mr + (M - f\bar{f})\Sigma](r^2 + \Sigma^2)^{-1}.$$

In addition, the neutrino field equations (2.1) imply that

$$\phi = -f(r + i\Sigma)^{-1}. \quad (4.2)$$

The functions Σ, P, m, M, U^0 which are real, and W^0, f which are complex, arise as "constants" of integration and are thus independent of r. They cannot however be assigned arbitrarily; in fact, Σ, U^0, W^0, and M are determined by P and Q. In order to describe these relationships it is convenient to introduce the following differential operators:

$$D_1 = \frac{\partial}{\partial u}, \qquad D_3 = Q\frac{\partial}{\partial u} + 2\frac{\partial}{\partial \bar{z}}, \qquad D_4 = \bar{D}_3.$$

Then we have

$$2i\Sigma = P^2(D_3\bar{Q} - D_4 Q),$$
$$W^0 = -i(D_3\Sigma + \Sigma D_1 Q), \quad (4.3)$$
$$U^0 = -Re[P^2 D_3(D_4 \ln P + D_1\bar{Q})].$$

The equation for M is most easily expressed by introducing a function $V(z,\bar{z},u)$ which satisfies

$$D_1 V = P, \quad (4.4)$$

(as does Robinson[17] in his discussion of the corresponding vacuum solutions). Then,

$$M = \tfrac{1}{2}iP^3(D_4D_4D_3D_3V - D_3D_3D_4D_4V). \quad (4.5)$$

The remaining gravitational field equations (corresponding to $R_{ab}n^a m^b$ and $R_{ab}n^a n^b$) relate the metric tensor components and the neutrino field according to

$$D_3(m + iM) + 3(m + iM)D_1 Q = if[D_3\bar{f} + \bar{f}(2D_1 Q + \tfrac{1}{2}D_3 \ln P)], \quad (4.6)$$

$$D_1[P^{-3}m - \tfrac{1}{2}(D_3D_3D_4D_4V + D_4D_4D_3D_3V)] + P^{-1}(D_1D_3D_3V)(D_1D_4D_4V) = -iP^{-3}(fD_1\bar{f} - \bar{f}D_1 f). \quad (4.7)$$

The remaining neutrino field equation restricts the function f by

$$D_3 f + f(D_1 Q - \tfrac{1}{2}D_3 \ln P) = 0. \quad (4.8)$$

We note that in the absence of a neutrino field ($f = 0$), Eqs. (4.1) and (4.3)–(4.7) reduce to the general vacuum equations obtained by Robinson,[4] as required.

For this class of solutions the gravitational field is restricted to be of Petrov–Penrose type $\{211\}$ or $\{22\}$. In the special case that the twist of the pnc vanishes ($\Sigma = 0, Q = 0$), which as in Sec. 3 implies that T_{ab} is of the form (3.1), it can only be of type $\{22\}$. In this case the line element (1.5) can be further reduced to the very simple form

$$ds^2 = -r^2\,dz\,d\bar{z} + 2\,du[dr - r^{-1}m(u)\,du], \quad (4.9)$$

where

$$\frac{\partial m}{\partial u} = -i\left(f\frac{\partial \bar{f}}{\partial u} - \bar{f}\frac{\partial f}{\partial u}\right),$$

and f in (4.2) is an arbitrary function of u alone. The line element (4.9) has been given by Griffith and Newing,[7] but its uniqueness as the only neutrino-gravitational field satisfying (1.1) with expanding but twist-free pnc was not established. An immediate consequence of this result is that *spherical symmetry of space–time is not compatible*[18] *with the presence of a neutrino field satisfying (1.1).* This has its origins in the explicit appearance of the twist of the neutrino pnc in the energy tensor, through Eq. (2.5), and is

in contrast with the case of the electromagnetic field [for which (1.1) is always valid].

We finally give the simplest solution of the system of Eq.(4.4)–(4.8), with nonzero twist:

$$V = 2^{-1/2}u, \quad P = 2^{-1/2}, \quad Q = -i\Sigma z, \quad M = 0,$$

with m, Σ, and f being constants. The corresponding line element is given by (1.5) with

$$G\overline{G} = 2(r^2 + \Sigma^2), \quad Q = -i\Sigma z,$$
$$W = 0, \quad U = (mr - f\bar{f}\Sigma)(r^2 + \Sigma^2)^{-1},$$

and the neutrino field by

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

In addition the condition (2.4) is satisfied, with equality holding.

Further explicit solutions have been found, and a systematic treatment will be given in a subsequent paper.

APPENDIX

We outline here the procedure leading to the line element (1.5) in the case that k^a is an expanding congruence. In order to arrive at this line element, we need only assume that space-time admits a geodesic and shear-free null congruence k^a and that the Ricci tensor satisfies (2.7) and (2.8) for some complex null vector field m^a orthogonal to k^a; the fact that (2.7) and (2.8) follow from the presence of a neutrino field is at this stage immaterial. As in Sec. 2, these conditions imply that k^a is a repeated pnd of the Weyl conform tensor.

We found it convenient to start with the line element (1.5) of Robinson[19] and then use the Newman–Penrose formalism.[20] In the notation of this formalism, conditions (1.3), (2.7), and (2.8) read

$$\kappa = \sigma = \epsilon + \bar{\epsilon} = 0, \quad \Phi_{00} = \Phi_{01} = \Phi_{02} = \Lambda = 0$$

and the fact that k^a is a repeated pnd of the Weyl tensor is expressed as

$$\Psi_0 = \Psi_1 = 0.$$

A suitable null tetrad for the line element (1.5) is

$$k^a = \delta^a_2, \quad n^a = \delta^a_1 + U\delta^a_2,$$
$$m^a = G^{-1}(Q\delta^a_1 + W\delta^a_2 + \delta^a_3 + i\delta^a_4),$$

which by virtue of the commutators implies

$$\tau + \bar{\pi} = \lambda = 0.$$

The remaining coordinate and tetrad freedom can then be used to achieve[21]

$$\tau = \pi = \epsilon = 0.$$

Then by a straightforward integration of the Ricci identities, we obtain the r dependence of the remaining nonzero spin coefficients and the nonzero tetrad components of the Ricci and Weyl tensors:

$$\rho = -(r + i\Sigma)^{-1}, \quad \alpha = (PD_1\overline{Q} + \tfrac{1}{2}D_4P)\rho,$$
$$\beta = -\tfrac{1}{2}D_3P\bar{\rho}, \quad \gamma = -\tfrac{1}{2}D_1 lnP + \tfrac{1}{2}(m + iM)\rho^2 + \Phi^0_{11}\rho^2\bar{\rho}, \quad \mu = \tfrac{1}{2}(m + iM)(\rho^2 + \rho\bar{\rho}) + \Phi^0_{11}\rho^2\bar{\rho} + (-U^0 + iD_1\Sigma - 2i\Sigma D_1 lnP)\bar{\rho},$$
$$\nu = P(\nu^0 + \rho\,\Psi^0_3 + \tfrac{1}{2}\rho^2 Y^0_1 + 6^{-1}\rho^3 Y^0_2) + \bar{\rho}^{-1}\Phi_{21},$$

where

$$\nu^0 = D_1[D_4 lnP + D_1\overline{Q}],$$
$$Y^0_1 = 3(m + iM)D_1\overline{Q} + D_4(m + iM),$$
$$Y^0_2 = 6i(m + iM)(D_4\Sigma + \Sigma D_1\overline{Q}),$$
$$\Phi_{11} = \Phi^0_{11}\,\rho^2\bar{\rho}^2,$$
$$\Phi_{21} = \bar{\rho}^2 P(\rho\,\Phi^0_{21} + \rho^2 B^0_1 + \rho^3 B^0_2),$$

where

$$B^0_1 = D_4\Phi^0_{11} + 4\,\Phi^0_{11}\,D_1\overline{Q},$$
$$B^0_2 = 2i\,\Phi^0_{11}(D_4\Sigma + \Sigma D_1\overline{Q}),$$

$$\Phi_{22} = P^2[\rho\bar{\rho}\,\Phi^0_{22} + \rho\bar{\rho}^2\overline{C}^0_1 + \rho^2\bar{\rho}\,C^0_1 + \rho^2\bar{\rho}^2\,C^0_2 + \Phi^{0\,-1}_{11}(\rho\bar{\rho}^3\Phi^0_{21}B^0_2 + \rho^3\bar{\rho}\,\overline{\Phi}^0_{21}B^0_2 + \rho^2\bar{\rho}^3 B^0_1\overline{B}^0_2 + \rho^3\bar{\rho}^2 B^0_2\overline{B}^0_1 + \rho^3\bar{\rho}^3 B^0_2\overline{B}^0_2)],$$

where

$$C^0_1 = D_4\overline{\Phi}^0_{21} + 3\,\overline{\Phi}^0_{21}\,D_1\overline{Q} - \tfrac{1}{2}P^2\,D_1(\Phi^0_{11}P^{-4}),$$
$$C^0_2 = \tfrac{1}{2}\,D_3B^0_1 + 2\,B^0_1D_1Q + \tfrac{1}{2}\,D_4\overline{B}^0_1 + 2\,\overline{B}^0_1D_1\overline{Q},$$
$$\Psi_2 = (m + iM)\rho^3 + 2\,\Phi^0_{11}\rho^3\bar{\rho},$$
$$\Psi_3 = P[\rho^2\,\Psi^0_3 + \rho^3\,Y^0_1 + \tfrac{1}{2}\rho^4\,Y^0_2 + \bar{\rho}(\rho^2\,\Phi^0_{21} + 2\rho^3\,B^0_1 + 3\rho^4\,B^0_2)],$$

where

$$\Psi^0_3 = 2i\,\Sigma\,\nu^0 + (D_4 + 2D_1\overline{Q}) \times (-U^0 + iD_1\Sigma - 2i\Sigma D_1\,lnP),$$
$$\Psi_4 = (3\alpha + \overline{\beta})\nu + \rho P(\overline{W}\frac{\partial}{\partial r} + D_4)\nu.$$

The components of the metric tensor are given by (4.1), (4.3)–(4.5) except that U now reads

$$U = U^0 + r\frac{\partial(\ln P)}{\partial u} + (mr + M\Sigma - \Phi^0_{11})(r^2 + \Sigma^2)^{-1}.$$

Further, the final differential Eqs. (4.6) and (4.7) become

$$D_3(m + iM) + 3(m + iM)D_1Q = 2\,\overline{\Phi}^0_{21} \qquad \text{(A1)}$$

and

$$D_1[P^{-3}m - \tfrac{1}{2}(D_3D_3D_4D_4V + D_4D_4D_3D_3V)] + P^{-1}(D_1D_3D_3V)(D_1D_4D_4V) = -P^{-1}\,\Phi^0_{22}. \qquad \text{(A2)}$$

The quantities $\Phi^0_{11}, \Phi^0_{21}, \Phi^0_{22}$ are arbitrary functions of $z, \bar{z}, u$, and represent the freedom available in specifying the source. The source-free case arises when we set

$$\Phi^0_{11} = \Phi^0_{21} = \Phi^0_{22} = 0.$$

Then the preceding equations reduce to those given by Talbot,[20] apart from certain notational differences,[22] and the simplifications arising from the introduction of the function V.

If we assume that the source is a neutrino field of the type under consideration in Sec. 4, then the neutrino field equations yield (4.2) and (4.8), and from the form (2.3) of the neutrino energy tensor we obtain

$$\Phi_{11}^0 = \Sigma f\bar{f},$$
$$\Phi_{21}^0 = -\tfrac{1}{2}i\bar{f}[D_4 f + f(2D_1\overline{Q} + \tfrac{1}{2}D_4 \ln P)],$$
$$\Phi_{22}^0 = iP^{-2}(fD_1\bar{f} - \bar{f}D_1 f).$$

Then the expression for U and Eqs. (A1) and (A2) reduce to those given in Sec. 4.

* Work supported in part by the National Research Council of Canada, through a postgraduate scholarship (D. T.) and a research grant (J. W.).

† Present Address: Computer Science Department, University of Manitoba, Winnipeg 19, Manitoba, Canada.

1 J. Wainwright, J. Math. Phys. **12**, 828 (1971), see Sec. 1.

2 It has been shown by J. Audretsch, Commun. Math. Phys. **21**, 303 (1971), that the second of the conditions in (1.1) is a consequence of the first (provided $T_{ab} \neq 0$).

3 See Ref. 1, Theorem 3.5, and Eqs. (2.9)–(2.11).

4 I. Robinson, J. Robinson, and J. Zund, J. Math. Mech. **18**, 881 (1969), pp. 881-83. Also I. Robinson and J. Robinson, Intern. J. Theoret. Phys. **2**, 231 (1969). For convenience we list the following identifications of notation: $z \leftrightarrow 2\zeta$, $\bar{z} \leftrightarrow 2\bar{\zeta}$, $r \leftrightarrow \rho$, $u \leftrightarrow -\sigma$, $\overline{Q} \leftrightarrow b$, $U \leftrightarrow S$, $\overline{W} \leftrightarrow -Z$, $G\overline{G} \leftrightarrow P^2$. We have used the coordinate freedom to set $a = 1$.

5 J. Audretsch and W. Graf, Commun. Math. Phys. **19**, 315 (1970).

6 D. R. Brill and J. M. Cohen, J. Math. Phys. **7**, 238 (1966).

7 J. B. Griffiths and R. A. Newing, J. Phys. A **3**, 136 (1970).

8 E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

9 See I. Robinson, A. Schild, and H. Strauss, Intern. J. Theoret. Phys. **2**, 243 (1969) in this connection.

10 Spinor indices are denoted $A, B, \ldots, A', B', \ldots$, and assume the values 1, 2. Our conventions as regards the 2-spinor formalism are those of R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960). The metric tensor has signature $(- - - +)$.

11 See Ref. 3. We are here using a noncanonical tetrad ($E \neq 0$), which accounts for the appearance of the additional terms involving E in (2.3) and (2.4).

12 See Ref. 1, Eqs. (2.10)–(2.12).

13 R. Penrose, in *Battelle Recontres: 1967 Lectures in Mathematics and Physics*, edited by B. DeWitt and J. Wheeler (Benjamin, New York, 1968), pp. 162-63.

14 W. Kundt and M. Trümper, Akad. Wiss. Mainz, No. 12 (1962). The required result is an obvious combination of Theorem 3.4 and Lemma 3.5 on pp. 984–85.

15 This follows from the first of the Ricci identities in Newman–Penrose form [Eq. (4.2a) in Ref. 8]. In the problem under consideration we have $\kappa = \sigma = \Phi_{00} = 0$, so that zero expansion ($\rho + \bar{\rho} = 0$) implies zero twist ($\rho - \bar{\rho} = 0$).

16 W. Kundt, Z. Physik **163**, 77 (1961).

17 See Ref. 4, p. 886. Note that our V and P correspond to Robinson's U and e^{-u}.

18 The reason is that the two repeated gravitational pnc's in a spherically symmetric space–time (one of which must coincide with the neutrino pnc) are geodesic, shear-free, expanding but twist-free. However, the line element (4.9) is not compatible with spherical symmetry. A special case of this result, corresponding to a pure radiation neutrino field, has been proved by Griffiths and Newing (see Ref. 7, p. 145).

19 In contrast to the integration procedure used by Robinson in the vacuum case, we regard the six real equations (2.7) and (2.8) as the "main equations." The "trivial equation" of Robinson does not arise, and the "subsidiary conditions," which in Robinson's approach correspond to $R_{ab}m^a n^b$ and $R_{ab}n^a n^b$, are here augmented by $R_{ab}k^a n^b$.

20 Alternatively, we could have directly generalized the approach of Talbot, Commun. Math. Phys. **13**, 45 (1969), who derived the class of algebraically special vacuum solutions with twisting rays, using the Newman–Penrose formalism, but without directly making use of Robinson's result.

21 By a procedure analogous to that of Talbot. See Ref. 19, pp. 51–52.

22 Equations (A1) and (A2) are the generalizations of Talbot's Eqs. (5.15) and (5.16), respectively ($m + iM$ is to be identified with Ψ_2^0). Equation (4.5) is equivalent to Talbot's (5.16), which is not affected by the nonvanishing of Φ_{11}^0, Φ_{12}^0, and Φ_{22}^0.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Some Considerations of Entropy Change*

S. Okubo and A. Isihara
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
(Received 10 May 1971)

An inequality for two positive operators is used to discuss entropy theorems for time smoothing, mixing, and other processes. Especially, it is proved that neglect of nondiagonal matrix elements for the density matrix causes the entropy to increase.

1. INTRODUCTION

The problem of proving temporal entropy increase has been ordinarily investigated on the basis of a coarse-graining process, a perturbation treatment of the Liouville or Bloch equation, and so on.[1] It is the purpose of this paper to present some exact considerations of entropy increase for relatively simple but well-defined cases. In order to illustrate our method, it is perhaps worthwhile to outline the classical case[2] in this introduction.

Let $f(x)$ and $g(x)$ be two nonnegative distribution functions in phase space. The variable x represents all phase space variables which are needed to describe a given system. The domain of x integration is irrelevant and need not be specified here. Let $f(x)$ and $g(x)$ be connected by

$$g(y) = \int K(y,x)f(x)dx, \tag{1.1}$$

where $K(y,x)$ is a transition probability. In case a time variable is explicit, Eq. (1.1) may be rewritten as

$$g(y,t_2) = \int K(y,t_2;x,t_1)\,f(x,t_1)dx, \tag{1.2}$$

where the integration does not include the time t_1

Then the preceding equations reduce to those given by Talbot,[20] apart from certain notational differences,[22] and the simplifications arising from the introduction of the function V.

If we assume that the source is a neutrino field of the type under consideration in Sec. 4, then the neutrino field equations yield (4.2) and (4.8), and from the form (2.3) of the neutrino energy tensor we obtain

$$\Phi_{11}^0 = \Sigma f\bar{f},$$
$$\Phi_{21}^0 = -\tfrac{1}{2}i\bar{f}[D_4 f + f(2D_1\overline{Q} + \tfrac{1}{2}D_4 \ln P)],$$
$$\Phi_{22}^0 = iP^{-2}(fD_1\bar{f} - \bar{f}D_1 f).$$

Then the expression for U and Eqs. (A1) and (A2) reduce to those given in Sec. 4.

* Work supported in part by the National Research Council of Canada, through a postgraduate scholarship (D. T.) and a research grant (J. W.).

† Present Address: Computer Science Department, University of Manitoba, Winnipeg 19, Manitoba, Canada.

1 J. Wainwright, J. Math. Phys. **12**, 828 (1971), see Sec. 1.

2 It has been shown by J. Audretsch, Commun. Math. Phys. **21**, 303 (1971), that the second of the conditions in (1.1) is a consequence of the first (provided $T_{ab} \neq 0$).

3 See Ref. 1, Theorem 3.5, and Eqs. (2.9)–(2.11).

4 I. Robinson, J. Robinson, and J. Zund, J. Math. Mech. **18**, 881 (1969), pp. 881-83. Also I. Robinson and J. Robinson, Intern. J. Theoret. Phys. **2**, 231 (1969). For convenience we list the following identifications of notation: $z \leftrightarrow 2\zeta$, $\bar{z} \leftrightarrow 2\bar{\zeta}$, $r \leftrightarrow \rho$, $u \leftrightarrow -\sigma$, $\overline{Q} \leftrightarrow b$, $U \leftrightarrow S$, $\overline{W} \leftrightarrow -Z$, $G\overline{G} \leftrightarrow P^2$. We have used the coordinate freedom to set $a = 1$.

5 J. Audretsch and W. Graf, Commun. Math. Phys. **19**, 315 (1970).

6 D. R. Brill and J. M. Cohen, J. Math. Phys. **7**, 238 (1966).

7 J. B. Griffiths and R. A. Newing, J. Phys. A **3**, 136 (1970).

8 E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

9 See I. Robinson, A. Schild, and H. Strauss, Intern. J. Theoret. Phys. **2**, 243 (1969) in this connection.

10 Spinor indices are denoted $A, B, \ldots, A', B', \ldots$, and assume the values 1, 2. Our conventions as regards the 2-spinor formalism are those of R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960). The metric tensor has signature $(- - - +)$.

11 See Ref. 3. We are here using a noncanonical tetrad ($E \neq 0$), which accounts for the appearance of the additional terms involving E in (2.3) and (2.4).

12 See Ref. 1, Eqs. (2.10)–(2.12).

13 R. Penrose, in *Battelle Recontres: 1967 Lectures in Mathematics and Physics*, edited by B. DeWitt and J. Wheeler (Benjamin, New York, 1968), pp. 162-63.

14 W. Kundt and M. Trümper, Akad. Wiss. Mainz, No. 12 (1962). The required result is an obvious combination of Theorem 3.4 and Lemma 3.5 on pp. 984–85.

15 This follows from the first of the Ricci identities in Newman–Penrose form [Eq. (4.2a) in Ref. 8]. In the problem under consideration we have $\kappa = \sigma = \Phi_{00} = 0$, so that zero expansion ($\rho + \bar{\rho} = 0$) implies zero twist ($\rho - \bar{\rho} = 0$).

16 W. Kundt, Z. Physik **163**, 77 (1961).

17 See Ref. 4, p. 886. Note that our V and P correspond to Robinson's U and e^{-u}.

18 The reason is that the two repeated gravitational pnc's in a spherically symmetric space–time (one of which must coincide with the neutrino pnc) are geodesic, shear-free, expanding but twist-free. However, the line element (4.9) is not compatible with spherical symmetry. A special case of this result, corresponding to a pure radiation neutrino field, has been proved by Griffiths and Newing (see Ref. 7, p. 145).

19 In contrast to the integration procedure used by Robinson in the vacuum case, we regard the six real equations (2.7) and (2.8) as the "main equations." The "trivial equation" of Robinson does not arise, and the "subsidiary conditions," which in Robinson's approach correspond to $R_{ab}m^a n^b$ and $R_{ab}n^a n^b$, are here augmented by $R_{ab}k^a n^b$.

20 Alternatively, we could have directly generalized the approach of Talbot, Commun. Math. Phys. **13**, 45 (1969), who derived the class of algebraically special vacuum solutions with twisting rays, using the Newman–Penrose formalism, but without directly making use of Robinson's result.

21 By a procedure analogous to that of Talbot. See Ref. 19, pp. 51–52.

22 Equations (A1) and (A2) are the generalizations of Talbot's Eqs. (5.15) and (5.16), respectively ($m + iM$ is to be identified with Ψ_2^0). Equation (4.5) is equivalent to Talbot's (5.16), which is not affected by the nonvanishing of Φ_{11}^0, Φ_{12}^0, and Φ_{22}^0.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Some Considerations of Entropy Change*

S. Okubo and A. Isihara
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
(Received 10 May 1971)

An inequality for two positive operators is used to discuss entropy theorems for time smoothing, mixing, and other processes. Especially, it is proved that neglect of nondiagonal matrix elements for the density matrix causes the entropy to increase.

1. INTRODUCTION

The problem of proving temporal entropy increase has been ordinarily investigated on the basis of a coarse-graining process, a perturbation treatment of the Liouville or Bloch equation, and so on.[1] It is the purpose of this paper to present some exact considerations of entropy increase for relatively simple but well-defined cases. In order to illustrate our method, it is perhaps worthwhile to outline the classical case[2] in this introduction.

Let $f(x)$ and $g(x)$ be two nonnegative distribution functions in phase space. The variable x represents all phase space variables which are needed to describe a given system. The domain of x integration is irrelevant and need not be specified here. Let $f(x)$ and $g(x)$ be connected by

$$g(y) = \int K(y,x)f(x)dx, \tag{1.1}$$

where $K(y,x)$ is a transition probability. In case a time variable is explicit, Eq. (1.1) may be rewritten as

$$g(y,t_2) = \int K(y,t_2;x,t_1)\,f(x,t_1)dx, \tag{1.2}$$

where the integration does not include the time t_1

because the normalization of the distribution functions is made in phase space. In view of its probability character, $K(y,x)$ must satisfy

$$K(y,x) \geq 0, \tag{1.3}$$

$$\int dy K(y,x) = 1. \tag{1.4}$$

Moreover, the total probability of the system to be in a state described by y starting with any initial state x must not exceed one[3]:

$$\int K(y,x)dx \leq 1. \tag{1.5}$$

Under these conditions, one can show

$$\int g(y)\ \log g(y)dy \leq \int f(x)\ \log f(x)dx \tag{1.6}$$

First, we notice that Eq. (1.4) gives

$$\int g(y)dy = \int f(x)dx, \tag{1.7}$$

i.e., both distribution functions have the same normalization. Hence, if we so wish, we can set the common integral in Eq. (1.7) to be unity.

Now, using Eqs. (1.1) and (1.4), we find

$$\begin{aligned}\int f(x)\ \log f(x)dx - \int g(y)\ \log g(y)dy \\ = \iint dxdy\ \ K(y,x)[f(x)\ \log f(x) - f(x)\ \log g(y) \\ - f(x) + g(y)] + \int dy[1 - \int dx\ \ K(y,x)]g(y)\end{aligned} \tag{1.8}$$

The integrands are both nonnegative due to our conditions (1.3) and (1.5) and the Gibbs inequality,[2] which is equivalent to

$$z\log z - z + 1 \geq 0, \quad z \geq 0,$$

with $z = f(x)/g(y)$. Hence, we have proved the validity of Eq. (1.6).

Actually, we can relax our condition (1.5) to a weaker one,

$$\iint dydz\ \ K(y,z)\ \ K(y,x) \leq 1, \tag{1.5'}$$

without spoiling our conclusion. However, the physical significance of Eq. (1.5') is not clear, and we shall not discuss this point further.

Also, Eq. (1.5) shows that we have

$$0 \leq g(y) \leq \max_x f(x).$$

This would imply that the range of $g(y)$ becomes more restricted than that of $f(x)$. This may mean a loss of available information in our stochastic process, which may be partly responsible for the increase of the entropy.

In the above consideration, we have used a continuous variable, but discrete cases can be handled without any difficulty. The source of inequality expressed by Eq. (1.6) is in the integration process involved in Eq. (1.2). This represents an averaging process over all possible initial states. Keeping the above considerations in mind, let us now try to consider quantum mechanical cases.

2. INEQUALITY FOR OPERATORS

For two Hermitian matrices A and B, let the eigenvalue equations be

$$\begin{aligned}A|j\rangle &= a_j|j\rangle, \\ B|\alpha\rangle &= b_\alpha|\alpha\rangle,\end{aligned} \tag{2.1}$$

where we have used italic and Greek suffixes for distinction. Let us assume that the eigenvalues are nonnegative,

$$a_j \geq 0, \quad b_\alpha \geq 0, \tag{2.2}$$

and that they are connected with each other by

$$b_\alpha = \sum_j K_{\alpha j} a_j. \tag{2.3}$$

In analogy to Eqs. (1.3), (1.4) and (1.5), we require[4]

$$K_{\alpha j} \geq 0, \tag{2.4}$$

$$\sum_\alpha K_{\alpha j} = 1, \tag{2.5}$$

$$\sum_j K_{\alpha j} \leq 1. \tag{2.6}$$

The eigenvectors are normalized so that

$$\begin{aligned}\langle j|k\rangle &= \delta_{j,k}, \\ \langle\alpha|\beta\rangle &= \delta_{\alpha,\beta}.\end{aligned} \tag{2.7}$$

Without loss of generality, we can assume that the state vectors are complete:

$$\begin{aligned}\sum_j |j\rangle\langle j| &= 1, \\ \sum_\alpha |\alpha\rangle\langle\alpha| &= 1.\end{aligned} \tag{2.8}$$

The two matrices A and B may not commute. If they do, the classical inequality (1.6) holds for the present case with the correspondence

$$\begin{aligned}|j\rangle &\longleftrightarrow x, \quad a_j \longleftrightarrow f(x), \\ |\alpha\rangle &\longleftrightarrow y, \quad b_\alpha \longleftrightarrow g(y), \\ K_{\alpha j} &\longleftrightarrow K(y,x),\end{aligned} \tag{2.9}$$

while the integrations over x and y are replaced by a summation on j and α, respectively. The interpretation of the resulting inequality in terms of entropy is the same as in the classical case.

If the matrices A and B do not commute, we need some slight modification of the statement. However, we can still prove an analogous theorem, although the physical interpretation of $K_{\alpha j}$ as the

probability for a stochastic process is no longer possible.

First, from Eq. (2.5), we must have

$$\sum_{\alpha} b_{\alpha} = \sum_{j} a_{j} \tag{2.10}$$

or, equivalently, the correct normalization condition

$$\mathrm{Tr}B = \mathrm{Tr}A\,. \tag{2.11}$$

Moreover,

$$\begin{aligned} \mathrm{Tr}(A\log A) &= \sum_{j} a_{j} \log a_{j}, \\ \mathrm{Tr}(B\,\log B) &= \sum_{\alpha} b_{\alpha} \log b_{\alpha}. \end{aligned} \tag{2.12}$$

Then, we can prove the following inequality (without assuming the commutability of A and B):

$$\mathrm{Tr}(B\log B) \leq \mathrm{Tr}(A\log A). \tag{2.13}$$

The proof of this inequality is essentially the same as that of Eq. (1.6) with the correspondence Eq. (2.9) and may not be repeated. However, a simple inspection of the proof shows that the equality in Eq. (2.13) is possible if and only if two sets of eigenvalues $\{a_j\}$ and $\{b_\alpha\}$ coincide with each other as a whole with possible exception of extra zero eigenvalues $a_j = 0$ or $b_\alpha = 0$. Also, we can relax, if we so wish, the condition Eq. (2.6) by a weaker one

$$\sum_{\alpha,k} K_{\alpha k} K_{\alpha j} \leq 1 \tag{2.6'}$$

as in Eq. (1.5′).

It is obvious that Eq. (2.13) is a natural generalization of the classical case, Eq. (1.6). On the basis of this inequality, we can derive several interesting theorems.

Theorem 1: Let P be a projection operator

$$P^2 = P \tag{2.14}$$

and let us specify the operator $\overline{A}$ by

$$\overline{A} = PAP + (1-P)A(1-P) \tag{2.15}$$

in terms of a nonnegative operator A. Then, we have

$$\mathrm{Tr}(\overline{A}\log\overline{A}) \leq \mathrm{Tr}(A\log A). \tag{2.16}$$

The proof is simple. Due to Eq. (2.15), the transformation matrix $K_{\alpha j}$ is easily calculated to be

$$K_{\alpha j} = |\langle j|P|\alpha\rangle|^2 + |\langle j|(1-P)|\alpha\rangle|^2$$

when we identify B as $\overline{A}$. It is easy to check that this expression satisfies the required conditions (2.4), (2.5), and (2.6). Hence, Eq. (2.16) follows.

We may physically interpret this theorem that neglect of certain nondiagonal matrix elements of A causes the entropy increase. Hence, our theorem is an analog of the entropy theorem based on coarse graining, and is related to a loss of information when we pass from A to $\overline{A}$.

We may easily generalize our theorem by a mathematical induction to:

Theorem 2: Let P_n, $n = 1, 2, \ldots$, be a complete set of projection operators, i.e.,

$$P_n P_m = \delta_{n,m} P_n, \qquad \sum_{n} P_n = I \tag{2.17}$$

and let us introduce

$$\overline{A} = \sum_{n} P_n A P_n; \tag{2.18}$$

then

$$\mathrm{Tr}(\overline{A}\log\overline{A}) \leq \mathrm{Tr}(A\log A) \tag{2.19}$$

provided that A is nonnegative.

As a special case, we have

Theorem 3: Let ψ_n, $n = 1, 2, \ldots$, be a set of arbitrary orthonormal complete state vectors. Then, for a nonnegative matrix A, we have

$$\mathrm{Tr}(A\log A) \geq \sum_{n} \langle\psi_n|A|\psi_n\rangle \log\langle\psi_n|A|\psi_n\rangle. \tag{2.20}$$

The proof is simple. Choose operator P_n in Theorem 2 to be the projection operator for the state ψ_n. Equation (2.20) then follows immediately from Eq. (2.19). This inequality may be considered as a generalization of the well-known Peierls inequality[5] for the density matrix. Indeed, we shall prove shortly that the latter will result from Eq. (2.20).

The following special case of Eq. (2.13) is interesting. Suppose that B and A are related by

$$B = UAU^{+}, \tag{2.21}$$

where U is an unitary operator

$$U^{+}U = UU^{+} = 1; \tag{2.22}$$

then we have

$$\mathrm{Tr}\{(UAU^{+})\log(UAU^{+})\} = \mathrm{Tr}(A\log A). \tag{2.23}$$

Notice that Eq. (2.23) is an equality instead of an inequality. This is obvious since both operators A and UAU^{+} have exactly the same set of eigenvalues. Hence, by a remark after Eq. (2.13) we must have the equality in Eq. (2.13).

The equality (2.23) represents simply the dynamical reversibility. Particularly for $A \equiv \rho(0)$ representing the initial density matrix at $t = 0$, let

$$\begin{aligned} U &= U(t) \equiv \exp(iHt/\hbar), \\ B &= \rho(t) \equiv U(t)AU^{+}(t). \end{aligned} \tag{2.24}$$

Then, Eq. (2.23) is reduced to the Liouville theorem

$$\mathrm{Tr}\{\rho(t)\ \log\rho(t)\} = \mathrm{Tr}\{\rho(0)\ \log\rho(0)\} \tag{2.25}$$

Since the entropy $S(t)$ may be defined by

$$S(t) = -k\mathrm{Tr}\{\rho(t)\ \log\rho(t)\}, \tag{2.26}$$

k being the Boltzmann constant, it is rewritten as

$$S(t) = S(0). \tag{2.27}$$

In order to obtain some mechanism for a genuine increase of the entropy, it is obvious that we have to modify some aspects of theory. One possibility is to assume that U is really not unitary but instead satisfies

$$U^+U = 1, \quad UU^+ \leq 1. \tag{2.28}$$

Such a behavior is, of course, not possible[4] for finite-dimensional space. But Eq. (2.28) could be possible in an infinite-dimensional Hilbert space, as we see in the scattering theory where the Møller wave matrix satisfies analogous conditions.[6] In such a situation, a genuine reversibility could occur since we can prove only $S(t) \geq S(0)$ now. However, for an infinite-dimensional space, our formal proof must be carefully reinvestigated in a mathematically more rigorous fashion.

We will now consider some other mechanisms for entropy increase.

(i) Suppose that, in Eq. (2.24), B is replaced by its diagonal part:

$$B = \langle U^+(t)AU(t)\rangle_{\mathrm{diag}} \equiv \sum_j |j\rangle\langle j|\rho(t)|j\rangle\langle j| \tag{2.29}$$

where $\rho(t)$ is defined by Eq. (2.24) with $A = \rho(0)$. By definition, A and B commute and two sets of states $\{|\alpha\rangle\}$ and $\{|j\rangle\}$ coincide as a whole. We can connect A and B by introducing a transition matrix $K = (K_{\alpha j})$ by

$$K_{\alpha j} = |\langle\alpha|U(t)|j\rangle|^2, \tag{2.30}$$

which satisfies all properties (2.4)–(2.6). Thus, we arrive at

$$\mathrm{Tr}\{[\rho(t)]_D\ \log[\rho(t)]_D\} \leq \mathrm{Tr}\{\rho(0)\ \log\rho(0)\}, \tag{2.31}$$

where the suffix D stands for "diagonal" with respect to eigenstates of $\rho(0)$. Equation (2.31) implies

$$S(t) \geq S(0) \tag{2.32}$$

if we redefine $S(t)$ by

$$S(t) = -k\ \mathrm{Tr}\{[\rho(t)]_D\ \log[\rho(t)]_D\}. \tag{2.33}$$

We conclude that the entropy defined by $[\rho(t)]_D$ increases in general in comparison to $S(0)$.

As a particular case, let

$$\rho(0) = \exp(-\beta H_0) \tag{2.34}$$

be the density matrix at $t = 0$, and let us assume that it is diagonal. If, for $t > 0$, the system is subject to interactions, if the Hamiltonian is

$$H = H_0 + H_1, \tag{2.35}$$

and if the density matrix

$$\rho = \exp(-\beta H) \tag{2.36}$$

is replaced by its diagonal matrix

$$\rho_D = \exp(-\beta H_0)[\exp(-\beta H_1)]_D, \tag{2.37}$$

then the same inequality (2.31) holds.

This diagonal approximation is not generally good. However, if the time interval $(0, t)$ is small in comparison with $\hbar^{-1}$ times the magnitude of the interaction, then the diagonalized density matrix can be reasonably close to the correct density matrix defined by the perturbation from the Liouville equation,

$$\frac{\partial}{\partial t}\rho + \frac{1}{i\hbar}[\rho, H] = 0.$$

Van Hove[1] has shown that this is indeed the case.

(ii) Let $A = \rho(0)$ and let B be a time-smoothed density matrix defined by

$$B = \frac{1}{\tau}\int_0^\tau d.t\ \ U(t)AU^+(t) \equiv \langle\rho\rangle_\tau; \tag{2.38}$$

then

$$\mathrm{Tr}\{\langle\rho\rangle_\tau\ \log\langle\rho\rangle_\tau\} \leq \mathrm{Tr}\{\rho(0)\ \log\rho(0)\} \tag{2.39}$$

In other words, the time smoothing operation leads to entropy increase. The importance of introducing time averaging has been discussed by Kirkwood and others.[7] Since the proof of Eq. (2.39) is similar to the previous cases, it will not be repeated.

(iii) For an open system, a description in terms of an ensemble of Hamiltonians $\{H_n\}$ is necessary. We define

$$A = \rho(0), \quad U_n = \exp(iH_nt/\hbar), \tag{2.40}$$

and

$$B = \rho(t) = \sum_n C_nU_n(t)AU_n^+(t) \tag{2.41}$$

If the numerical weights C_n satisfy the condition

$$C_n \geq 0, \quad \sum_n C_n = 1, \tag{2.42}$$

we find that the transformation function $K_{\alpha j}$ now given by

$$K_{\alpha j} = \sum_n C_n|\langle\alpha|U_n(t)|j\rangle|^2$$

satisfies the conditions (2.4)–(2.6), and hence we conclude that

$$\mathrm{Tr}\{\rho(t)\ \log\rho(t)\} \leq \mathrm{Tr}\{\rho(0)\ \log\rho(0)\}. \tag{2.43}$$

Next, we shall prove that we can derive the Peierls inequality from Theorem 3. To obtain this result, we will first show that it gives us the generalized Gibbs inequality.

$$\mathrm{Tr}\,(A\,\log A - A\log B - A + B) \geq 0 \tag{2.44}$$

for two arbitrary nonnegative matrices A and B. Let ψ_n be the eigenstate of B, with

$$B|\psi_n\rangle = b_n|\psi_n\rangle.$$

Then, the inequality (2.20) can be used to derive

$$\mathrm{Tr}(A\,\log A - A\,\log B - A + B) \geq \sum_n (a_n\,\log a_n - a_n\,\log b_n - a_n + b_n),$$

where we have set $a_n = \langle\psi_n|A|\psi_n\rangle$. But the right-hand side of this equation is nonnegative due to the classical Gibbs inequality $z\,\log z - z + 1 \geq 0$ for $z \geq 0$ with identification $z = a_n/b_n$. Therefore, we find Eq. (2.44).

Now, in order to prove the Peierls inequality, let us make a substitution $A \to e^B$ and $B \to e^A$ in Eq. (2.44):

$$\mathrm{Tr}(Be^B - Ae^B - e^B + e^A) \geq 0. \tag{2.45}$$

For an arbitrary complete orthonormal state vector ψ_n, $n = 1, 2, \ldots$, and for a given Hermitian operator A, let us choose the operator B in Eq. (2.45) to be

$$B = \sum_n |\psi_n\rangle\langle\psi_n|A|\psi_n\rangle\langle\psi_n|.$$

It is then apparent that Eq. (2.45) leads immediately to

$$\mathrm{Tr}\,(e^A) \geq \sum_n \exp(\langle\psi_n|A|\psi_n\rangle). \tag{2.46}$$

This is the Peierls inequality.

Concluding this paper, we will make some applications of the Gibbs inequality (2.44). First, if we normalize A and B to be unity, i.e.,

$$\mathrm{Tr}A = \mathrm{Tr}B = 1, \tag{2.47}$$

then it reduces to

$$\mathrm{Tr}(A\,\log A) \geq \mathrm{Tr}(A\,\log B). \tag{2.48}$$

Let us choose A to be the exact distribution function

$$A = e^{-\beta H}/\mathrm{Tr}(e^{-\beta H}). \tag{2.49}$$

Also, we set B to be an uncorrelated trial distribution function of the form

$$B = \lambda \prod_{j=1}^{3N} \exp\{-F_j(x_j, p_j)\}, \tag{2.50}$$

where λ is the normalization constant, N is the total number of particles and F_j is a function of jth particle variables x_j and p_j. Then Eq. (2.48) gives us the following expression for the entropy:

$$S \leq k(-\log\lambda + \sum_{j=1}^{3N} \langle F_j\rangle), \tag{2.51}$$

where the expectation value of an operator Q is defined as usual by

$$\langle Q\rangle = \mathrm{Tr}(Qe^{-\beta H})/\mathrm{Tr}(e^{-\beta H}). \tag{2.52}$$

Especially, choose F_j to be

$$F_j(x_j, p_j) = \tfrac{1}{2}u(x_j)^2 + \tfrac{1}{2}v(p_j)^2, \tag{2.53}$$

where u and v are arbitrary positive constants. Since F_j is easily diagonalizable, the normalization constant λ for N identical particles is readily calculated to be

$$\lambda = N!\,(2\,\sinh\tfrac{1}{2}\omega)^{3N}, \qquad \omega = (uv\hbar^2)^{1/2} \tag{2.54}$$

so that Eq. (2.51) is reduced to

$$S \leq k\{-\log N! - 3N\,\log(2\,\sinh\tfrac{1}{2}\omega) + \tfrac{3}{2}N[u\langle x^2\rangle + v\langle p^2\rangle]\}, \tag{2.55}$$

where $\langle x^2\rangle$ ($\langle p^2\rangle$) is the common value of $\langle x_j^2\rangle$ ($\langle p_j^2\rangle$) for the isotropic medium. Since u and v are arbitrary, we can minimize the right-hand side of Eq. (2.55) by changing these variables, and we easily find

$$S \leq k\{-\log(2^{3N}N!) + \tfrac{3}{2}N[(z+1)\log(z+1) - (z-1)\log(z-1)]\}, \tag{2.56}$$

where z is given by

$$z = (2/\hbar)[\langle x^2\rangle\langle p^2\rangle]^{1/2}. \tag{2.57}$$

Notice that, by the usual uncertainty principle, we must have $z \geq 1$ automatically. Actually, we can replace $\langle x^2\rangle$ and $\langle p^2\rangle$ in Eq. (2.57) by $\langle(x - \langle x\rangle)^2\rangle$ and $\langle(p - \langle p\rangle)^2\rangle$, respectively. This is due to the fact that we could have replaced x_j and p_j in Eq. (2.53) by $x_j - \langle x_j\rangle$ and $p_j - \langle p_j\rangle$, respectively, from the beginning.

For a classical system, the situation is much simpler. If $f(x)$ and $g(x)$ are two classical distribution function normalized to be

$$\int f(x)dx = \int g(x)dx = 1, \tag{2.58}$$

then Eq. (2.48) is now rewritten as

$$\int dx\ f(x)\,\log f(x) \geq \int dx\ f(x)\,\log g(x). \tag{2.59}$$

Since the contribution due to the kinetic energy is easily subtracted out, we discuss only the quantity associated with the space coordinates. Again choose $f(x)$ to be the correct distribution and set

$$g(x) = \prod_{j=1}^{3N} \left(\frac{\lambda}{\pi}\right)^{1/2} \exp[-\lambda(x_j - b)^2], \tag{2.60}$$

where λ and b are arbitrary constants. With this choice, Eq. (2.59) leads to

$$\int f(x)\ \log f(x)dx \geq \sum_{j=1}^{3N}\left[\frac{1}{2}\log\left(\frac{\lambda}{\pi}\right) - \lambda\langle(x_j - b)^2\rangle\right].$$

Again maximizing the right-hand side with respect to λ and b, we find

$$S \leq \tfrac{3}{2}kN\{1 + \log[2\pi(\Delta x)^2]\}, \tag{2.61}$$

where we have set

$$(\Delta x)^2 = \langle(x_j - \langle x_j\rangle)^2\rangle. \tag{2.62}$$

* This work was supported in part by the U.S. Atomic Energy Commission. One of the authors (A.I.) was supported in part by the National Science Foundation through Grant GP-19740 and in part by the Advanced Research Project Agency through Contract #N00014-67-0398 while taking a leave of absence from the State University of New York at Buffalo.

1 T. Ehrenfest and P. Ehrenfest, Sitzber. Kaiserl. Akad. Wiss. Wien, Math. Naturw. Klasse, **15**, 89 (1906); L. Van Hove, Physica **21**, 512 (1955); **22**, 243 (1956); **23**, 441 (1957); I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience, New York, 1962); E. W. Montroll, in *Lectures in Theoretical Physics* (Gordon and Breach, New York, 1960), Vol. III, p. 221; P. H. E. Meijer and J. C. Edwards, Phys. Rev. A **1**, 1527 (1970).

2 A. Isihara, J. Phys. A. (Proc. Phys. Soc.) **1**, 539 (1968); J. Math. Anal. Appl., to be published.

3 This is weaker than the constraint $\int K(y,x)dx = 1$ imposed in the second paper of Ref. 2. The sufficiency of this relaxed condition has been independently observed by J. M. Richardson (private communication).

4 When we notice a trivial identity $\sum_\alpha(\sum_j K_{\alpha j}) = \sum_j(\sum_\alpha K_{\alpha j})$, then Eqs. (2.5) and (2.6) imply dim $B \geq$ dim A for the dimensions of matrices A and B. For a finite-dimensional space, the condition (2.6) automatically implies a stronger result $\sum_j K_{\alpha j} = 1$, provided that we insist dim B = dim A. In other words, the genuine inequality in Eq. (2.6) is physically possible only for infinite-dimensional cases.

5 R. E. Peierls, Phys. Rev. **54**, 918 (1938).

6 E.g., see, M. L. Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1965), p. 211.

7 J. G. Kirkwood, J. Chem. Phys. **14**, 180 (1940); M. S. Green, *ibid.* **25**, 836 (1956); J. L. Lebowitz and R. J. Rubin, Phys. Rev. **131**, 2381 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Perturbation of Statistical Semigroups in Quantum Statistical Mechanics

D. A. Uhlenbrock*

Department of Mathematics, University of Wisconsin, Madison, Wisconsin 53706

(Received 9 June 1971)

The perturbation theory of Hille and Phillips for semigroups of bounded linear operators on a Banach space is modified to apply to the semigroups of positive traceclass operators encountered in quantum statistical mechanics.

INTRODUCTION

Quantum statistical mechanics aims to clarify the macroscopic aspects of a system composed of a large number of microscopic subsystems which interact with each other. The division between macroscopic and microscopic systems is prompted by atomic theory. A frequently used model takes the microscopic system to be a "particle" (atom, molecule, or the like) and treats it as an irreducible component. Correlations among particles are attributed to interactions which *perturb* the behavior of the idealized free particles.

Mathematically, free particles are described by the free (Fock–Cook) representation $\pi_F : \mathfrak{A} \to \mathfrak{B}_\infty(\mathcal{H})$ of the C^*-algebra of fields $\mathfrak{A}$ in the C^*-algebra of bounded linear operators on the representation (H)-space $\mathcal{H}$.[1] Let $\mathfrak{B}_1(\mathcal{H})$[resp., $\mathfrak{B}_1^+(\mathcal{H})$] denote the ($B$)-space of traceclass operators (resp., the set of positive traceclass operators) on $\mathcal{H}$, normed by the tracenorm $\|\cdot\|_1$.[2] If $\mathfrak{A}$ is the C^*-algebra of a *bounded* system the physically interesting states on $\mathfrak{A}$ are the states which are *normal* with respect to the free representation.[3–5] Such states are of the form $\mathfrak{A} \ni A \rightsquigarrow \mathrm{tr}_{\mathcal{H}}(S\cdot\pi(A))\cdot(\mathrm{tr}_{\mathcal{H}}(S))^{-1}$ with a *statistical operator* $S \in \mathfrak{B}_1^+(\mathcal{H})$.

Let $\mathbb{R}_+ = (0, \infty)$ and define $\overline{\mathbb{R}}_+ = [0, \infty) = \{0\} \cup \mathbb{R}_+$. Then $\overline{\mathbb{R}}_+$ is an additive semigroup with identity element 0 under addition of real numbers. A (C_0)-*semigroup on* $\mathcal{H}$ is any semigroup map from $\overline{\mathbb{R}}_+$ into $\mathfrak{B}_\infty(\mathcal{H})$ which assigns $1_{\mathcal{H}}$ to the 0 of $\overline{\mathbb{R}}_+$ and which is continuous in the strong operator topology of $\mathfrak{B}_\infty(\mathcal{H})$.[6] Such a semigroup is said to be *self-adjoint* (resp., *positive; invertible*) in case its values in $\mathfrak{B}_\infty(\mathcal{H})$ are selfadjoint (resp., positive; invertible) operators. A bounded operator is called *invertible* if 0 is not an eigenvalue.

Proposition 1: Let $S \in \mathfrak{B}_\infty(\mathcal{H})$ be positive and invertible. Then there is one and only one self-adjoint (C_0)-semigroup $S(\cdot)$ such that $S = S(1)$. This semigroup, called the *power semigroup of* S, is also positive and invertible.

Proof: If $P(\cdot)$ is the spectral family of S, then $S = \int \sigma dP(\sigma)$. Since spec$S \subset \overline{\mathbb{R}}_+$, $H = -\int(\log\sigma)\cdot dP(\sigma) \equiv -\log S$ is a well-defined, possibly unbounded, selfadjoint operator. Let its spectral representation be given as $H = \int \eta \cdot dE(\eta)$. Then the spectral calculus implies that $s \rightsquigarrow \exp(-sH) \equiv \int e^{-s\eta}dE(\eta) = \int \sigma^s dP(\sigma)$ yields the existence part of the proposition. It is clear that $S(s) = \exp(-sH)$ is positive and invertible for all $s \in \overline{\mathbb{R}}_+$. For uniqueness, let $s \rightsquigarrow T(s)$ be any self-adjoint (C_0)-semigroup with $T(1) = S$. Then, by Theorem 1 of HP 22.3 (see Ref. 6), T has the representation

where λ and b are arbitrary constants. With this choice, Eq. (2.59) leads to

$$\int f(x)\ \log f(x)dx \geq \sum_{j=1}^{3N}\left[\frac{1}{2}\log\left(\frac{\lambda}{\pi}\right) - \lambda\langle(x_j - b)^2\rangle\right].$$

Again maximizing the right-hand side with respect to λ and b, we find

$$S \leq \tfrac{3}{2}kN\{1 + \log[2\pi(\Delta x)^2]\}, \tag{2.61}$$

where we have set

$$(\Delta x)^2 = \langle(x_j - \langle x_j\rangle)^2\rangle. \tag{2.62}$$

* This work was supported in part by the U.S. Atomic Energy Commission. One of the authors (A.I.) was supported in part by the National Science Foundation through Grant GP-19740 and in part by the Advanced Research Project Agency through Contract #N00014-67-0398 while taking a leave of absence from the State University of New York at Buffalo.

1 T. Ehrenfest and P. Ehrenfest, Sitzber. Kaiserl. Akad. Wiss. Wien, Math. Naturw. Klasse, **15**, 89 (1906); L. Van Hove, Physica **21**, 512 (1955); **22**, 243 (1956); **23**, 441 (1957); I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience, New York, 1962); E. W. Montroll, in *Lectures in Theoretical Physics* (Gordon and Breach, New York, 1960), Vol. III, p. 221; P. H. E. Meijer and J. C. Edwards, Phys. Rev. A **1**, 1527 (1970).

2 A. Isihara, J. Phys. A. (Proc. Phys. Soc.) **1**, 539 (1968); J. Math. Anal. Appl., to be published.

3 This is weaker than the constraint $\int K(y,x)dx = 1$ imposed in the second paper of Ref. 2. The sufficiency of this relaxed condition has been independently observed by J. M. Richardson (private communication).

4 When we notice a trivial identity $\sum_\alpha(\sum_j K_{\alpha j}) = \sum_j(\sum_\alpha K_{\alpha j})$, then Eqs. (2.5) and (2.6) imply dim $B \geq$ dim A for the dimensions of matrices A and B. For a finite-dimensional space, the condition (2.6) automatically implies a stronger result $\sum_j K_{\alpha j} = 1$, provided that we insist dim B = dim A. In other words, the genuine inequality in Eq. (2.6) is physically possible only for infinite-dimensional cases.

5 R. E. Peierls, Phys. Rev. **54**, 918 (1938).

6 E.g., see, M. L. Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1965), p. 211.

7 J. G. Kirkwood, J. Chem. Phys. **14**, 180 (1940); M. S. Green, *ibid.* **25**, 836 (1956); J. L. Lebowitz and R. J. Rubin, Phys. Rev. **131**, 2381 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Perturbation of Statistical Semigroups in Quantum Statistical Mechanics

D. A. Uhlenbrock*

Department of Mathematics, University of Wisconsin, Madison, Wisconsin 53706

(Received 9 June 1971)

The perturbation theory of Hille and Phillips for semigroups of bounded linear operators on a Banach space is modified to apply to the semigroups of positive traceclass operators encountered in quantum statistical mechanics.

INTRODUCTION

Quantum statistical mechanics aims to clarify the macroscopic aspects of a system composed of a large number of microscopic subsystems which interact with each other. The division between macroscopic and microscopic systems is prompted by atomic theory. A frequently used model takes the microscopic system to be a "particle" (atom, molecule, or the like) and treats it as an irreducible component. Correlations among particles are attributed to interactions which *perturb* the behavior of the idealized free particles.

Mathematically, free particles are described by the free (Fock–Cook) representation $\pi_F : \mathfrak{A} \to \mathfrak{B}_\infty(\mathcal{H})$ of the C^*-algebra of fields $\mathfrak{A}$ in the C^*-algebra of bounded linear operators on the representation (H)-space $\mathcal{H}$.[1] Let $\mathfrak{B}_1(\mathcal{H})$[resp., $\mathfrak{B}_1^+(\mathcal{H})$] denote the ($B$)-space of traceclass operators (resp., the set of positive traceclass operators) on $\mathcal{H}$, normed by the tracenorm $\|\cdot\|_1$.[2] If $\mathfrak{A}$ is the C^*-algebra of a *bounded* system the physically interesting states on $\mathfrak{A}$ are the states which are *normal* with respect to the free representation.[3–5] Such states are of the form $\mathfrak{A} \ni A \rightsquigarrow \mathrm{tr}_{\mathcal{H}}(S\cdot\pi(A))\cdot(\mathrm{tr}_{\mathcal{H}}(S))^{-1}$ with a *statistical operator* $S \in \mathfrak{B}_1^+(\mathcal{H})$.

Let $\mathbb{R}_+ = (0, \infty)$ and define $\overline{\mathbb{R}}_+ = [0, \infty) = \{0\} \cup \mathbb{R}_+$. Then $\overline{\mathbb{R}}_+$ is an additive semigroup with identity element 0 under addition of real numbers. A (C_0)-*semigroup on* $\mathcal{H}$ is any semigroup map from $\overline{\mathbb{R}}_+$ into $\mathfrak{B}_\infty(\mathcal{H})$ which assigns $1_{\mathcal{H}}$ to the 0 of $\overline{\mathbb{R}}_+$ and which is continuous in the strong operator topology of $\mathfrak{B}_\infty(\mathcal{H})$.[6] Such a semigroup is said to be *self-adjoint* (resp., *positive; invertible*) in case its values in $\mathfrak{B}_\infty(\mathcal{H})$ are selfadjoint (resp., positive; invertible) operators. A bounded operator is called *invertible* if 0 is not an eigenvalue.

Proposition 1: Let $S \in \mathfrak{B}_\infty(\mathcal{H})$ be positive and invertible. Then there is one and only one self-adjoint (C_0)-semigroup $S(\cdot)$ such that $S = S(1)$. This semigroup, called the *power semigroup of* S, is also positive and invertible.

Proof: If $P(\cdot)$ is the spectral family of S, then $S = \int \sigma dP(\sigma)$. Since spec$S \subset \overline{\mathbb{R}}_+$, $H = -\int(\log\sigma)\cdot dP(\sigma) \equiv -\log S$ is a well-defined, possibly unbounded, selfadjoint operator. Let its spectral representation be given as $H = \int \eta \cdot dE(\eta)$. Then the spectral calculus implies that $s \rightsquigarrow \exp(-sH) \equiv \int e^{-s\eta}dE(\eta) = \int \sigma^s dP(\sigma)$ yields the existence part of the proposition. It is clear that $S(s) = \exp(-sH)$ is positive and invertible for all $s \in \overline{\mathbb{R}}_+$. For uniqueness, let $s \rightsquigarrow T(s)$ be any self-adjoint (C_0)-semigroup with $T(1) = S$. Then, by Theorem 1 of HP 22.3 (see Ref. 6), T has the representation

$$T(s) = \int e^{-s\eta} d\bar{E}(\eta),$$

where $\bar{E}(\cdot)$ is the spectral family of its (negative) infinitesimal generator. Putting $s = 1$,

$$S = S(1) = \int e^{-\eta} dE(\eta) = T(1) = \int e^{-\eta} d\bar{E}(\eta)$$

shows, by the uniqueness of the spectral representation, that $E(\cdot) = \bar{E}(\cdot)$ and hence $S(\cdot) = T(\cdot)$.///

We call the operator $H = -\log S$ constructed above the *Hamiltonian* of the power semigroup of S. It is the (negative) generator of $S(\cdot)$.

In particular, if $S \in \mathfrak{B}_1^+(\mathfrak{K})$ is invertible, it is of the form

$$S = \sum_{n\geq 1} \sigma_n \cdot P_n$$

with eigenvalues $\| S \|_\infty = \sigma_1 > \sigma_2 > \cdots > \sigma_n > \cdots > 0$, $\sigma_n \searrow 0$ as $n \nearrow \infty$ and projectors P_n such that $\dim(P_n\mathfrak{K}) < \infty$. It follows that

$$H = \sum_{n\geq 1} (-\log\sigma_n)\cdot P_n \geq -\log\sigma_1 \cdot 1_{\mathfrak{K}} \text{ and } \operatorname{tr}_{\mathfrak{K}}(S) = \| S \|_1 = \| \exp(-H) \|_1 = \sum_{n\geq 1} \sigma_n \dim(P_n\mathfrak{K}) < \infty.$$

For $s > 1$ it then follows that $S(s) = \exp(-sH)$ is in $\mathfrak{B}_1^+(\mathfrak{K})$ also, since $\| S(s) \|_1 = \| S(1)\cdot S(s-1) \|_1 \leq \| S(1) \|_1 \cdot \| S(s-1) \|_\infty = \| S \|_1 \cdot (\sigma_1)^{s-1} < \infty$. If $0 < s < 1$, however, it does not follow in general that $S(s) \in \mathfrak{B}_1^+(\mathfrak{K})$. For example, if $(P_n)_{n\geq 1}$ is a family of one-dimensional pairwise orthogonal projectors in $\mathfrak{K}$ and $\sigma_n = n^{-2}$,

$$\operatorname{tr}_{\mathfrak{K}}(S) = \| S \|_1 = \sum_{n\geq 1} n^{-2} < \infty,$$

while $\operatorname{tr}_{\mathfrak{K}} S(\tfrac{1}{2}) = \| S^{1/2} \|_1 = \sum_{n\geq 1} n^{-1} = \infty$

Definition 1: A linear operator $S \in \mathfrak{B}_\infty^+(\mathfrak{K})$ is a *Gibbs operator* if it is invertible and $S(s) \in \mathfrak{B}_1^+(\mathfrak{K})$ for all $s > 0$. If S is a Gibbs operator, the associated Hamiltonian $H = -\log S$ is called a *Gibbs Hamiltonian*, while the power semigroup $S(\cdot)$ of S is a *Gibbs semigroup*.

By proposition 1 any datum among $S, H, S(\cdot)$ uniquely determines the other two. It would be of interest to have explicit spectral criteria for H to be a Gibbs Hamiltonian.

We note that for any $c > 0$: $(c\cdot S)(s) = c^s \cdot S(s)$ so that the power semigroups which extend S and $c\cdot S$, respectively, are simply related by multiplication with $c: s \rightsquigarrow c^s$. The corresponding Hamiltonians are H and $H - \log c$, respectively. Thus by choosing $c \geq \sigma_1^{-1} = \| S \|_\infty^{-1}$ (or $c > \sigma_1^{-1}$, respectively) it is often possible to restrict oneself to operators S with $\operatorname{spec} S \subset [0, 1]$ (or $\operatorname{spec} S \subset [0, 1)$, resp.) or, equivalently, to H with $\operatorname{spec} H \subset [0, \infty)$ (or $\operatorname{spec} H \subset (0, \infty)$, resp.). In any case $\| S(s) \|_\infty = \sigma_1^s$, so that for $\sigma_1 = 1$ $S(s)$ is always of norm 1, while for $0 < \sigma_1 < 1$, $S(s)$ is a contraction whose norm decays exponentially with $s \nearrow \infty$. The number $\log\sigma_1$ is called the *type* of the semigroup.

Proposition 2: Let $S(\cdot)$ be a Gibbs semigroup. Then its restriction to $\mathbb{R}_+ = \bar{\mathbb{R}}_+ \setminus \{0\}$ is continuous w.r.t. the tracenorm on $\mathfrak{B}_1^+(\mathfrak{K})$.

Proof: There is no loss of generality in assuming $S(\cdot)$ to be of strictly negative type. Let $0 < \frac{1}{2}s < s_1 \leq s \leq s_2$ be given with s fixed. Define a function

$$f_{s_1,s_2}: [0,1] \to \mathbb{R}: \sigma \rightsquigarrow \sigma^{s_1} - \sigma^{s_2}.$$

It is easily proved to assume its maximum

$$\frac{\delta}{s_2} \cdot \left(1 + \frac{\delta}{s_1}\right)^{-s_1/\delta}$$

at $\sigma_{\max} = (1 + \delta/s_1)^{-1/\delta}$, where $\delta = s_2 - s_1$.

Now, if

$$S = \sum_{n\geq 1} \sigma_n P_n \text{ (as above)}, \quad \| S(s_1) - S(s_2) \|_\infty = \sup_{n\geq 1} f_{s_1,s_2}(\sigma_n) \leq \sup_{\sigma\in[0,1]} f_{s_1,s_2}(\sigma) = \frac{\delta}{s_2}\left(1 + \frac{\delta}{s_1}\right)^{-s_1/\delta}$$

As $s_1 \nearrow s$ and $s_2 \searrow s$, $\delta \to 0$ so that the rhs tends to zero and $S(\cdot)$ is $\|\cdot\|_\infty$ – continuous on $\mathbb{R}_+$.

Finally, in the same limit

$$\| S(s_1) - S(s_2) \|_1 = \| S(\tfrac{1}{2}s)\cdot[S(s_1 - \tfrac{1}{2}s) - S(s_2 - \tfrac{1}{2}s)] \|_1 \leq \| S(\tfrac{1}{2}s) \|_1 \cdot \| S(s_1 - \tfrac{1}{2}s) - S(s_2 - \tfrac{1}{2}s) \|_\infty \to 0,$$

as $S(\cdot)$ is $\|\cdot\|_\infty$– continuous at $\frac{1}{2}s$ by the preceding step.///

Assume $S(\cdot)$ is a Gibbs semigroup. Thus it is clear from Stone's theorem that the self-adjoint operator H generates a strongly continuous unitary group by $\mathbb{R} \ni t \rightsquigarrow T(t) \equiv \exp(-itH)$. Furthermore $\bar{C}_\gamma \equiv \mathbb{R}_+ + i\mathbb{R} \ni s + it \rightsquigarrow S(s)\cdot T(t) \equiv U(s + it)$ defines the maximal semigroup extension of $S(\cdot)$ in $\mathfrak{B}_\infty(\mathfrak{K})$. It is strongly continuous on $\bar{C}_\gamma$ and (uniformly) holomorphic on the interior C_γ of $\bar{C}_\gamma$.[7] $U(\cdot)$ is the *full Gibbs semigroup* of S. Its knowledge is equivalent to the pair $(S(\cdot), T(\cdot))$ which is also characterized by the Hamiltonian H.

If we interpret H as the unperturbed Hamiltonian, an expression of the form $H + P$, involving a suitable perturbation P, can be regarded as the Hamiltonian of an interacting system. In the following we find sufficient conditions on P such that $H + P$ generates a perturbed Gibbs semigroup. Moreover, we represent the perturbed Gibbs semigroup in terms of P and the unperturbed Gibbs semigroup as a tracenorm convergent infinite series whose nth term involves precisely n factors of P.

With $S(\cdot)$ a Gibbs semigroup, the states $\mathfrak{A} \ni A \rightsquigarrow \operatorname{tr}_{\mathfrak{K}}(S(s)\pi_F(A))/\operatorname{tr}_{\mathfrak{K}}(S(s))$ are called the *Gibbs states* associated with S (or equivalently with its Gibbs Hamiltonian H). In more customary notation $s \equiv \beta = (kT)^{-1}$ so that $S(s) = e^{-\beta H}$. In sub-

sequent work we use the perturbation theory for $S(\cdot)$ which is developed here in order to derive corresponding expansions for the Gibbs states associated with $S(\cdot)$.

It should be apparent to the attentive reader of the relevant literature that the fundamental (nowadays predominatly algebraical) aspects of local quantum theory have received plenty of attention and are quite polished. In contrast to this the computational aspects appear to be still underdeveloped. A workable, mathematically controlled perturbation theory can serve to bridge this gap. One might well recall that the "applied part" of quantum electrodynamics, i.e., the part which produces numbers to be compared to experimental numbers, is almost totally dependent on perturbation ideas. In fact, the corresponding perturbation formalism has been *formally* transcribed into the context of quantum statistical mechanics.[8] Generally these treatments do not contain enough mathematical detail to allow one to arrive at convergence proofs and thus remain formal. An exception to this statement is formed by the work of Gruber[9] on Euclidean Green's functions. While similar in basic motivation his approach to the perturbation problem is mathematically entirely distinct from ours.

Outline: In Sec. 1, we present, as an example, a typical unperturbed semigroup which will figure in later applications. Section 2 gives pertinent facts arising from the semigroup perturbation theory of Hille and Phillips (HP—Chap. XIII), which forms the basis of our treatment. Since their perturbation theory is formulated in a Banach space and applies to more general classes of semigroups (not only (C_0)-semigroups) it becomes necessary to specialize to our context. In particular, the self-adjointness notion which is peculiar to (H)-spaces can be exploited. Naturally, somewhat stronger results can be claimed. Finally, in Sec. 3 we extend these results to Gibbs semigroups.

1. UNPERTURBED SEMIGROUP

As a typical example of the unperturbed setup we summarize the Fock–Cook construction for a fermion field in a "periodic box". Its details differ little from standard constructions[1,10–12] and are germaine to applications in subsequent work only.

A. Single Particle Space $\mathfrak{h}$

Let $\nu \geq 1$ be an integer and $b = (b_\alpha)_{1 \leq \alpha \leq \nu}$ an element of $\mathbb{R}^\nu_+$. The subset

$$B = \prod_{\alpha=1}^{\nu} [0, b_\alpha)$$

of $\mathbb{R}^\nu$ represents a "periodic box" of volume

$$|B| = \prod_{\alpha=1}^{\nu} b_\alpha.$$

If μ_B denotes the restriction of Lebesgue measure on $\mathbb{R}^\nu$ to B, the single particle space is given as $\mathfrak{h} = L^2_{\mathbb{C}}(\mathbb{R}^\nu, \mu_B)$. Its inner product and norm are written as $(f,g) = \int d\mu_B f \cdot \bar{g}$ and $\|f\| = (f,f)^{1/2}$. A CONB is obtained as follows: with

$$K = \left\{ \left(\frac{2\pi}{b_\alpha} m_\alpha \right)_{1 \leq \alpha \leq \nu} \mid m_\alpha \text{ integer} \right\}$$

define the functions (i.e., classes)

$$e_k : \mathbb{R}^\nu \to \mathbb{C} : x \rightsquigarrow |B|^{-1/2} \cdot \exp(ik \cdot x),$$

where $k \in K$.

B. Fock–Cook Multiparticle Space $\mathfrak{K}$

Let $\wedge \mathfrak{h} = \oplus_{m \geq 0} \wedge^m \mathfrak{h}$ be the exterior algebra of the complex vector space $\mathfrak{h}$. Here $\wedge^0 \mathfrak{h} = \mathbb{C} \cdot 1, \wedge^1 \mathfrak{h} = \mathfrak{h}$. $\wedge^m \mathfrak{h}$ is the mth exterior power of $\mathfrak{K}$ for $m \geq 2$, which is spanned by decomposable vectors of the form $f_1 \wedge \cdots \wedge f_m$ with $f_i \in \mathfrak{h}$.

The assignment

$$\langle f_1 \wedge \cdots \wedge f_m, g_1 \wedge \cdots \wedge g_n \rangle \equiv \delta_{mn} \cdot \det(\langle f_i, g_j \rangle)_{ij}$$

for all m, $n \geq 0$ and $f_i, g_j \in \mathfrak{h}$ extends uniquely to an inner product on $\wedge \mathfrak{h}$, which becomes a Hausdorff prehilbert space as a consequence. Let $\overline{\wedge} \mathfrak{h}$ denote its norm completion, which no longer is an algebra, since $\dim \mathfrak{h} = \infty$. $\overline{\wedge}\, \mathfrak{h}$ is the (H)-space direct sum of the completions $\overline{\wedge}^m \mathfrak{h}$ of $\wedge^m \mathfrak{h}$ and constitutes the multiparticle space $\mathfrak{K}$.

If K is assumed totally ordered and K' is any finite subset of K, let $e_{K'} \equiv e_{k_1} \wedge \cdots \wedge e_{k_{|K'|}}$, where $\{k_1, \ldots, k_{|K'|}\}$ is the ordered presentation of K'.

The collection $(e_{K'})_{K' \subset K \text{ finite}}$ thus obtained forms a CONB of $\mathfrak{K}$.

C. Field Algebra $\pi_F(\mathfrak{A})$

For $f \in \mathfrak{h}$, let $f\wedge$ be the linear operator on $\wedge \mathfrak{h}$ which sends $f_1 \wedge \cdots \wedge f_m$ into $f \wedge f_1 \wedge \cdots \wedge f_m$. Since it is bounded on $\wedge \mathfrak{h}$, it extends by continuity to an operator $c(f) \in \mathfrak{B}_\infty(\mathfrak{K})$. The map $f \rightsquigarrow c(f)$ is $\mathbb{C}$-linear and satisfies $c(f)c(g) + c(g)c(f) = 0$.

The *field algebra* $\pi_F(\mathfrak{A})$ in the Fock–Cook representation is then defined as the smallest C^*-subalgebra of $\mathfrak{B}_\infty(\mathfrak{K})$ which contains $1_{\mathfrak{K}}$ and all $c(f)$, $f \in \mathfrak{h}$.

The *creation operators* $c(f)$ and their adjoints, the *annihilation operators* $c(f)^*$, are characterized by the canonical anticommutation relations (CAR):

(i) $c(f) \cdot c(g) + c(g) \cdot c(f) = 0$,

(ii) $c(f) \cdot c(g)^* + c(g)^* \cdot c(f) = (f,g) \cdot 1_{\mathfrak{K}}$, the "vacuum condition",

(iii) $c(f)^* 1_{\mathfrak{K}} = 0$ together with the irreducibility of $\pi_F(\mathfrak{A})$.

If u is unitary on $\mathfrak{h}$, then $f_1 \wedge \cdots \wedge f_m \rightsquigarrow u f_1 \wedge \cdots \wedge u f_m$ extends uniquely to an automorphism $\wedge u$ on $\wedge \mathfrak{h}$. Since it is isometric and has an inverse $\wedge u^{-1}$

which is also isometric, Λu again extends uniquely to a unitary operator $\bar{\Lambda}u$ on $\bar{\Lambda}\mathfrak{h} = \mathfrak{K}$. The map $u \rightsquigarrow \bar{\Lambda}u$ is easily seen to be a strongly continuous representation of the unitary group of $\mathfrak{h}$ in the unitary group of $\mathfrak{K}$. Thus if $\mathbb{R} \ni t \rightsquigarrow u(t) = \exp(-iht)$ is a strongly continuous unitary group on $\mathfrak{h}$ with generator h then $\mathbb{R} \ni t \rightsquigarrow \bar{\Lambda}u(t)$ is such a group on $\mathfrak{K}$. In particular, it is of the form $\bar{\Lambda}u(t) = \exp(-iHt)$. It is convenient to write $H = d\bar{\Lambda}(h)$ to emphasize the connection between the *single particle Hamiltonian* h and the *many particle Hamiltonian* H. $d\bar{\Lambda}$ is the *second quantization map* for self-adjoint single particle operators. We note that $c(uf) = (\bar{\Lambda}u)\cdot c(f)\cdot(\bar{\Lambda}u^{-1})$.

In terms of this notation it was shown by Shale and Stinespring[12] that $\exp(-sh)$ is in $\mathfrak{B}_1^+(\mathfrak{h})$ if and only if $\exp(-sd\bar{\Lambda}(h))$ is in $\mathfrak{B}_1^+(\mathfrak{K})$, where $s > 0$. In particular it follows, that $s \rightsquigarrow \exp(-sh)$ is a Gibbs semigroup if and only if $s \rightsquigarrow \exp(-sd\bar{\Lambda}(h))$ is a Gibbs semigroup. Within this setup it is hence sufficient to assume that one is given a Gibbs semigroup on the single particle space $\mathfrak{h}$: $s \rightsquigarrow \exp(-sh)$. The perturbation problem is then posed for the corresponding semigroup $s \rightsquigarrow \exp(-sH)$, where $H = d\bar{\Lambda}(h)$.

2. HILLE-PHILLIPS PERTURBATION THEORY

Let $\mathfrak{K}$ be any separable complex (H)-space. Throughout this section we assume given a self-adjoint (C_0)-semigroup S: $\bar{\mathbb{R}}_+ \to \mathfrak{B}_\infty(\mathfrak{K})$: $s \rightsquigarrow S(s)$. The following lemma lists some automatic consequences. Let

$$\mathfrak{D}_{\bar{H}} = \{f \in \mathfrak{K} \mid \text{s-}\lim_{s \searrow 0_+} s^{-1}(S(s) - 1)f \text{ exists}\}.$$

For $f \in \mathfrak{D}_{\bar{H}}$ the *generator* $\bar{H} \equiv -H$ of $S(\cdot)$ is defined as

$$\bar{H}f = \text{s-}\lim_{s \searrow 0_+} s^{-1}(S(s) - 1)f.$$

Briefly, $\bar{H} = \lim(S(s) - 1)/s$ with the "natural" maximal domain $\mathfrak{D}_{\bar{H}}$ is the right derivative of $S(\cdot)$ at 0.

Lemma 2.1: (i) The limit $\lim_{s \to \infty} s^{-1}\cdot \| S(s) \|_\infty \equiv \omega_0$ exists and is finite and

$$\| S(s) \|_\infty = \exp(\omega_0 \cdot s);$$

(ii) The generator $\bar{H}$ is self-adjoint with $\bar{H} \leq \omega_0 \cdot 1$;

(iii) If $\bar{H}$ is bounded, $\mathbb{C} \ni s + it \rightsquigarrow \exp(s\bar{H})\cdot\exp(it\bar{H})$ defines the (entire) maximal holomorphic extension $U(\cdot)$ of $S(\cdot)$. If $\bar{H}$ is unbounded, $\bar{\mathbb{C}}_\gamma = \bar{\mathbb{R}}_+ + i\mathbb{R} \ni s + it \rightsquigarrow \exp(s\bar{H})\cdot\exp(it\bar{H})$ defines the maximal bounded extension $U(\cdot)$ of $S(\cdot)$. The restriction of $U(\cdot)$ to $\mathbb{C}_\gamma = \mathbb{R}_+ + i\mathbb{R}$ is the maximal holomorphic extension. For $s > 0$, $S(s)$ is positive and has an unbounded densely defined inverse;

(iv) For any $s + it$ in the domain of the maximal holomorphic extension $\mathfrak{D}_{\bar{H}} \supset U(s + it)\mathfrak{K}$ and $U(\cdot)$ is complex differentiable [in the uniform topology on $\mathfrak{B}_\infty(\mathfrak{K})$] with derivative $\bar{H}\circ U(s + it)$ at $s + it$.

Proof: Items (i), (ii), and (iii) are readily deduced from Theorem 1 of HP 22.3. (iv) is an easy consequence of the spectral representation of $U(\cdot)$ induced by the spectral representation of the self-adjoint $\bar{H}$.///

A linear operator $\bar{P} \equiv -P$ with domain $\mathfrak{D}_{\bar{P}} = \mathfrak{D}_{\bar{H}}$ is an *allowed perturbation* of $\bar{H}$ if $\bar{H} + \bar{P}$ generates a (C_0)-semigroup, given that $\bar{H}$ does.

A linear operator Q with domain $\mathfrak{D}_Q$ is $\bar{H}$-*bounded* if $\mathfrak{D}_Q \supset \mathfrak{D}_{\bar{H}}$ and there are constants $a \geq 0$, $b \geq 0$ with $\| Qf \| \leq a \| f \| + b \| \bar{H}f \|$ for all $f \in \mathfrak{D}_{\bar{H}}$.[2]

Let $\rho(\bar{H})$ be the resolvent set of $\bar{H}$ and for $\lambda \in \rho(\bar{H})$ set $R(\lambda; \bar{H}) \equiv R(\lambda) \equiv (\lambda \cdot 1 - \bar{H})^{-1}$.

Lemma 2.2: Let $\bar{P}$ have domain $\mathfrak{D}_{\bar{H}}$. Then the following are equivalent conditions on $\bar{P}$:

(i) For some $\lambda \in \rho(\bar{H})$, $\bar{P}\circ R(\lambda) \in \mathfrak{B}_\infty(\mathfrak{K})$;

(ii) For all $\lambda \in \rho(\bar{H})$, $\bar{P}\circ R(\lambda) \in \mathfrak{B}_\infty(\mathfrak{K})$;

(iii) $\bar{P}$ is $\bar{H}$-bounded;

(iv) $\bar{P}$ is a bounded operator from $\mathfrak{D}_{\bar{H}}$ (with the graph norm of $\bar{H}$) to $\mathfrak{K}$.

Proof: (i) $\Rightarrow$ (ii): If $\bar{P}\circ R(\lambda_0) \in \mathfrak{B}_\infty(\mathfrak{K})$ and $\lambda \in \rho(\bar{H})$ the first resolvent equation $\bar{P}\circ R(\lambda) = \bar{P}\circ R(\lambda_0) + (\lambda_0 - \lambda)\bar{P}\circ R(\lambda_0)\cdot R(\lambda)$ implies $P\circ R(\lambda) \in \mathfrak{B}_\infty(\mathfrak{K})$ also. (ii) $\Rightarrow$ (iii): Let $f \in \mathfrak{D}_{\bar{H}}$ and $\lambda \in \rho(\bar{H})$, then $\| \bar{P}f \| = \| \bar{P}\circ R(\lambda)\circ R(\lambda)^{-1}f \| \leq \| \bar{P}\circ R(\lambda) \|_\infty \cdot \| (\lambda 1 - \bar{H})f \| \leq \| \bar{P}\circ R(\lambda) \|_\infty \cdot \{|\lambda| \cdot \| f \| + \| \bar{H}f \|\}$. (iii) $\Rightarrow$ (iv): $\mathfrak{D}_{\bar{H}} \ni f \rightsquigarrow (f, \bar{H}f) \in$ graph $\bar{H}$ is a bijection. Since graph $\bar{H}$ is a (H) subspace of $\mathfrak{K} \oplus \mathfrak{K}$, $\mathfrak{D}_{\bar{H}}$ is a (H) space via this bijection. The graph *norm* of $f \in \mathfrak{D}_{\bar{H}}$ is $\| f \|_{\bar{H}} = \sqrt{\| f \|^2 + \| \bar{H}f \|^2}$. It is equivalent to the norm $\| f \|'_{\bar{H}} = \| f \| + \| \bar{H}f \|$. Let $f \in \mathfrak{D}_{\bar{H}}$, then

$$\begin{aligned} \| \bar{P}f \| &\leq a \| f \| + b \| \bar{H}f \| \\ &\leq \max(a, b)\cdot(\| f \| + \| \bar{H}f \|) \\ &\leq \max(a, b) \| f \|'_{\bar{H}}. \end{aligned}$$

(iv) $\Rightarrow$ (i): Let $\lambda \in \rho(\bar{H})$ and $g \in \mathfrak{K}$, then

$$\begin{aligned} \| R(\lambda)g \|'_{\bar{H}} &= \| R(\lambda)g \| + \| \bar{H}R(\lambda)g \| \\ &= \| R(\lambda)g \| + \| (\bar{H} - \lambda 1)R(\lambda)g + \lambda g \| \\ &\leq \| R(\lambda) \|_\infty \cdot \| g \| + |\lambda - 1| \cdot \| g \| \end{aligned}$$

Hence $R(\lambda)$: $\mathfrak{K} \to \mathfrak{D}_{\bar{H}}$ is a continuous bijection between (B) spaces and even a linear homeomorphism by the open mapping theorem. In particular this map is bounded (in both direction). $\bar{P}\circ R(\lambda)$ is the composition of bounded maps and is therefore bounded.///

Remark 2.1: Hille and Phillips denote the class of operators $\bar{P}$ which satisfy any (and therefore all) of the conditions of Lemma 2.2 by $\mathfrak{s}(\bar{H})$.

Given any $Q \in \mathcal{G}(\overline{H})$ define $Q^{\sim} = \lim_{\lambda\to\infty} \lambda \cdot Q \cdot R(\lambda)$. More precisely, let

$$\mathcal{D}_{Q^{\sim}} = \{f \in \mathcal{K} \mid \lim_{\lambda\to\infty} \lambda \cdot Q \cdot R(\lambda) \cdot f \text{ exists}\}$$

and put $Q^{\sim} f = \lim \lambda\, QR(\lambda) f$ for $f \in \mathcal{D}_{Q^{\sim}}$ as $\lambda \to \infty$. $Q^{\sim}$ is called the *$\overline{H}$-extension of Q*. Let $\mathcal{G}^{\sim}(H)$ be the class of all $\overline{H}$ extensions. Theorem 1 of HP 13.3 then asserts that $\mathcal{G}(\overline{H}) \ni Q \rightsquigarrow Q^{\sim} \in \mathcal{G}^{\sim}(\overline{H})$ is a bijection, whose inverse is the restriction map to $\mathcal{D}_{\overline{H}}$. Furthermore, if Q has a closed extension Q_1, then $Q \subset Q^{\sim} \subset Q_1$. Also, if Q is bounded, $Q^{\sim} \in \mathcal{B}_{\infty}(\mathcal{K})$ with

$$\|Q^{\sim}\|_{\infty} = \sup_{0 \neq f \in \mathcal{D}_{\overline{H}}} \frac{\|Qf\|}{\|f\|}.$$

Lemma 2.3: Let Q have $\mathcal{D}_Q \supset \mathcal{D}_{\overline{H}}$:

(i) If Q is closable then it is $\overline{H}$-bounded;

(ii) if Q is symmetric then it is $\overline{H}$-bounded.

In either case the restriction $Q|_{\mathcal{D}_{\overline{H}}} \in \mathcal{G}(\overline{H})$.

Proof: (i) Put $Q' = Q|_{\mathcal{D}_{\overline{H}}}$. Then, since $\mathcal{D}_{\overline{H}}$ is dense, Q'^{*} exists and $Q^{*} \subset (Q')^{*}$. Since Q is closable Q^{*} is densely defined, so that $(Q')^{*}$ is densely defined and $(Q')^{**}$ exists and is the closure of Q'. Thus it is sufficient to assume $\mathcal{D}_Q = \mathcal{D}_{\overline{H}}$.

Assume $f_n \in \mathcal{D}_Q$, $g \in \mathcal{K}$ with $\|f_n\|_{\overline{H}}^{\iota} = \|f_n\| + \|\overline{H} f_n\| \to 0$ and $\|Qf_n - g\| \to 0$. Since also $\|f_n\| \to 0$ and Q is closable as a map in $\mathcal{K}$ it follows that $g = 0$. This means also that Q is closable as a map from the (B)-space $\mathcal{D}_{\overline{H}}$ to the (B)-space $\mathcal{K}$. Since it is defined on all of $\mathcal{D}_{\overline{H}}$ it is closed and then bounded by the closed graph theorem. By Lemma 2.2 it is therefore $\overline{H}$-bounded.
(ii) $Q \subset Q^{*}$ implies $Q \subset Q^{**} \subset Q^{*}$ and Q is closable. Apply (i).///

Lemma 2.4: Let $\mathcal{D}_{\overline{P}} = \mathcal{D}_{\overline{H}}$. If $\overline{H} + \overline{P}$ also generates a self-adjoint (C_o)-semigroup then $\overline{P}$ is necessarily symmetric. If so, its $\overline{H}$ extension satisfies $\overline{P} \subset \overline{P}^{\sim} \subset \overline{P}^{**} \subset \overline{P}^{*}$.

Proof: By Lemma 2.1, $\overline{H} + \overline{P}$ is self-adjoint on its domain $\mathcal{D}_{\overline{H}}$. $\overline{P}$ is densely defined, so that $\overline{P}^{*}$ exists. Finally $\overline{P}^{*} = [(\overline{H} + \overline{P}) + (-\overline{H})]^{*} \supset (\overline{H} + \overline{P})^{*} + (-\overline{H})^{*} = (\overline{H} + \overline{P}) - \overline{H} = \overline{P}$. If $\overline{P} \subset \overline{P}^{*}$, then $\overline{P} \subset \overline{P}^{**} \subset \overline{P}^{*}$ and $\overline{P} \subset \overline{P}^{\sim} \subset \overline{P}^{**} \subset \overline{P}^{*}$ follows by remark 2.1, since $\overline{P}^{**}$ is the closure of $\overline{P}$. ///

As a result, if we are concerned with perturbations which do not destroy the self-adjointness of the semigroup, any perturbing operator $\overline{P}$ with $\mathcal{D}_{\overline{P}} = \mathcal{D}_{\overline{H}}$ will have to be symmetric and will automatically be $\overline{H}$-bounded due to Lemma 2.3. It is then sufficient to work with $\overline{P}^{*}$ which extends all relevant operators. Furthermore, if a symmetric Q with $\mathcal{D}_Q \supset \mathcal{D}_{\overline{H}}$ is given at the start, then $\overline{P} = Q|_{\mathcal{D}_{\overline{H}}}$ will do also, since $\overline{P} \subset Q \subset Q^{*} \subset \overline{P}^{*}$ so that $\overline{P}$ is symmetric.

Lemma 2.5: Let Q have $\mathcal{D}_Q \supset \mathcal{D}_{\overline{H}}$.

(i) $Q \circ S(s)$ is defined on $\mathcal{K}$ for $s \in \mathbb{R}_+$.

(ii) If Q is closable, then $Q \circ S(s) \in \mathcal{B}_{\infty}(\mathcal{K})$ for $s > 0$ and

$$\|Q \circ S(s)\|_{\infty} = \sup_{\|f\|=1,\, f \in \mathcal{D}_{\overline{H}}} \|Q \circ S(s) f\| < \infty.$$

Furthermore, $\mathbb{R}_+ \ni s \rightsquigarrow Q \circ S(s)$ is $\|\cdot\|_{\infty}$— continuous, while $\mathbb{R}_+ \ni s \rightsquigarrow \|Q \circ S(s)\|_{\infty}$ is continuous.

(iii) if Q is closable and S is a Gibbs semigroup then the last continuity statements hold in terms of $\|\cdot\|_1$.

Proof: (i) By Lemma 2.1, (iv) $\mathcal{D}_{\overline{H}} \supset S(s)\mathcal{K}$ for $s > 0$. (ii) $S(s)$ is closed and bounded. Therefore $Q \circ S(s)$ is closable, since Q is. Hence it is closed, since it is everywhere defined by (i), and therefore bounded by the closed graph therorem. Since $\mathcal{D}_{\overline{H}}$ is dense in $\mathcal{K}$ the equality statement follows.

For $s > 0$, let $0 < \frac{1}{2}s < s_1 \leq s \leq s_2$. Then $\|Q \circ S(s_2) - Q \circ S(s_1)\|_{\infty} = \|Q \circ S(\frac{1}{2}s) \cdot (S(s_2 - \frac{1}{2}s) - S(s_1 - \frac{1}{2}s))\|_{\infty} \leq \|Q \circ S(\frac{1}{2}s)\|_{\infty} \cdot \|S(s_2 - \frac{1}{2}s) - S(s_1 - \frac{1}{2}s)\|_{\infty}$ as $\substack{s_1 \\ s_2} \to s$ the rhs goes to 0 due to the $\|\cdot\|_{\infty}$— continuity of $s' \rightsquigarrow S(s')$ at $s' = \frac{1}{2}s$ (Lemma 2.1). Therefore (ii) follows. For (iii): As in (ii), $\|Q \circ S(s_2) - Q \circ S(s_1)\|_1 \leq \|Q \circ S(\frac{1}{2}s)\|_{\infty} \cdot \|S(s_2 - \frac{1}{2}s) - S(s_1 - \frac{1}{2}s)\|_1 \to 0$ due to $\|\cdot\|_1$— continuity of $S(\cdot)$ at $\frac{1}{2}s$ (Proposition 2).///

Remark 2.2: In the situation of Lemma 2.5, (ii) if ω_o is the type of $S(\cdot)$ (Lemma 2.1), then with $\delta > 0$ and $s > 0$,

$$\|Q \circ S(s + \delta)\|_{\infty} \leq \|Q \circ S(\delta)\|_{\infty} \cdot \|S(s)\|_{\infty} \leq \|Q \circ S(\delta)\|_{\infty} \cdot \exp(\omega_o s).$$

Thus, if $\omega_o < 0$, $s \rightsquigarrow \|Q \circ S(s)\|_{\infty}$ is integrable on any interval $[\delta, \infty)$ with $\delta > 0$.

In the general case (i), the continuity of $s \rightsquigarrow \|Q \circ S(s) f\|$ for $f \in \mathcal{K}$ implies the lower semicontinuity and a fortiori the measureability of $s \rightsquigarrow \sup_{\|f\|=1,\, f \in \mathcal{D}_{\overline{H}}} \|Q \circ S(s) f\|$.

Definition 2: A linear operator $\overline{P}$ is said to be in the *Phillips perturbation class* $\mathcal{P}(\overline{H})$ *of* $\overline{H}$, if

(i) $\mathcal{D}_{\overline{P}} = \mathcal{D}_{\overline{H}}$

(ii) $\overline{P}$ is $\overline{H}$-bounded,

(iii) for $s > 0$ $\sup_{\|f\|=1,\, f \in \mathcal{D}_{\overline{H}}} \|P \circ S(s) f\| < \infty$,

(iv) $\int_0^1 ds \sup_{\|f\|=1,\, f \in \mathcal{D}_{\overline{H}}} \|P \circ S(s) f\| < \infty$.

Let $\mathcal{P}_C(\overline{H})$ [resp., $\mathcal{P}_s(\overline{H})$] denote the class of $\overline{P} \in \mathcal{P}(\overline{H})$ which are closable (resp., which are symmetric or, equivalently, which possess a symmetric extension).

It is clear from this definition that $\mathcal{P}(\overline{H})$ is a complex linear space [$\mathcal{P}_s(\overline{H})$ is a real linear subspace]. Also, $\mathcal{P}_c(\overline{H}) \supset \mathcal{P}_s(\overline{H})$. For closable (or symmetric) P satisfying (i) it follows that $P \in \mathcal{P}(\overline{H})$ if and only if $s \rightsquigarrow \|P \circ S(s)\|_\infty$ is integrable at 0, since (ii) and (iii) are automatically satisfied according to lemmas 2.3 and 2.5. In particular, $\mathcal{P}(\overline{H})$ contains the restrictions to $\mathfrak{D}_{\overline{H}}$ of all $Q \in \mathcal{B}_\infty(\mathcal{K})$. Alternate conditions for (iii) and (iv) are in HP 13.5.

Theorem 2.1: Let $\overline{H}$ on $\mathfrak{D}_{\overline{H}}$ generate a self-adjoint (C_o)-semigroup $S(\cdot)$ on $\mathcal{K}$ and assume $\overline{P} \in \mathcal{P}(H)$. Let Q be either $\overline{P}$ or $1|_{\mathfrak{D}_{\overline{H}}}$. Then

(i) $\overline{H} + \overline{P}$ generates a (C_O)-semigroup

$$s \rightsquigarrow S^{\overline{P}}(s) = \sum_{n \geq 0} S_n^{\overline{P}}(s),$$

where $S_o^{\overline{P}}(s) = S(s)$ and $S_n^{\overline{P}}(s) = \int_0^s ds_1 S(s - s_1) \cdot P^\sim \circ S_{n-1}^{\overline{P}}(s_1)$ for $n \geq 1$ are defined as strong Bôchner integrals. The infinite sum converges absolutely (in norm) and uniformly so on intervals $(0, s_o)$ with $0 < s_o < \infty$. The functions $s \rightsquigarrow Q \circ S_n^{\overline{P}}(s)$ are strongly continuous for $s > 0$ as well as Bôchner integrable on compact subsets of $\overline{\mathbb{R}}_+$;

(ii) if $S(\cdot)$ is norm continuous for $s > 0$, then $S^{\overline{P}}(\cdot)$ is likewise norm continuous;

(iii) let D be a domain in C with elements z. Assume that $\overline{P}(z) \in \mathcal{P}(\overline{H})$ for $z \in D$ and that $z \rightsquigarrow \overline{P}(z)^\sim \circ S(s)$ is holomorphic in D such that

$$s \rightsquigarrow \sup_{z \in D} \|P(z)^\sim \cdot S(s)\|_\infty$$

is integrable on $[0, 1]$. Then $z \rightsquigarrow S^{\overline{P}(z)}(s)$ is holomorphic in D for any $s > 0$;

(iv) if (ii) applies or if $\overline{P} \in \mathcal{P}_c(H)$ the definition of $S_n^{\overline{P}}(s)$ holds as a norm Bôchner integral and $s \rightsquigarrow Q^\sim \circ S_n^{\overline{P}}(s)$ is norm continuous as well as integrable on compact subsets of $\mathbb{R}_+$.

Proof: (i) results from the proof of Theorem 1 and Corollary 1 of HP 13.4. (ii) follows from Theorem 2(5) of HP 13.4 since a (C_o)-semigroup is also a $(1, A)$-semigroup. (iii) Theorem 3 of HP 13.4. (iv) Lemma 2.5 (ii) yields the norm continuity of $s \rightsquigarrow \overline{Q}^\sim \cdot S(s)$. A straightforward modification of the proof of (i) then results in (iv). ///

Remark 2.3: Assume $\overline{P}_1$ and $\overline{P}_2$ are in $\mathcal{P}(\overline{H})$ and define

$$d(\overline{P}_1, \overline{P}_2) = \int_0^1 ds \, \| (\overline{P}_1 - \overline{P}_2)^\sim \circ S(s) \|_\infty < \infty.$$

Then according to HP 13.5, d is a metric on $\mathcal{P}(\overline{H})$ which makes it into a *complete* metric space. Also if $d(P_1, P_2) \to 0$ then

$$\|S^{\overline{P}_1}(s) - S^{\overline{P}_2}(s)\|_\infty \to 0$$

uniformly for s in compact subsets of $\overline{\mathbb{R}}_+$ (HP 13.5).

If $s \in \mathbb{R}$ and $(s)_n = (s_1, \ldots, s_n) \in \mathbb{R}^n$, define

$$\sigma_o = s - s_1, \quad s = \sigma_o + \sigma_1 + \cdots + \sigma_n,$$
$$\sigma_1 = s_1 - s_2, \quad s_1 = \sigma_1 + \sigma_2 + \cdots + \sigma_n,$$
$$\sigma_{n-1} = s_{n-1} - s_n, \quad s_{n-1} = \sigma_{n-1} + \sigma_n,$$
$$\sigma_n = s_n - 0, \quad s_n = \sigma_n.$$

Let $\Delta_n^s = \{(s)_n \in \mathbb{R}^n \mid s \geq s_1 \geq s_2 \geq \cdots \geq s_n \geq 0\}$
$= \{(s)_n \in \mathbb{R}^n \mid \sigma_i \geq 0 \ \& \ \sum \sigma_i = s\}$

and write $\chi_{\Delta_n^s}$ for the characteristic function of this set.

Lemma 2.6: With the notation of Theorem 2.1, let $f, g \in \mathcal{K}$: Then

$$\langle \overline{Q}^\sim S_n^{\overline{P}}(s) f, g \rangle = \int_0^s ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \times \langle \overline{Q}^\sim S(\sigma_o) \overline{P}^\sim \cdots \overline{P}^\sim S(\sigma_n) f, g \rangle = \int \chi_{\Delta_n^s}((s)_n) \cdot \langle \overline{Q}^\sim S(\sigma_o) \overline{P}^\sim \cdots \overline{P}^\sim S(\sigma_n) f, g \rangle d(s)_n.$$

Proof: $\langle \overline{Q}^\sim S_n^{\overline{P}}(s) f, g \rangle$

$$= \langle \overline{Q}^\sim \int_0^s ds_1 (S(\sigma_o) \overline{P}^\sim S_{n-1}^{\overline{P}}(s_1) f), g \rangle$$
$$= \int_0^s ds_1 \langle \overline{Q}^\sim S(\sigma_o) \overline{P}^\sim S_{n-1}^{\overline{P}}(s_1) f, g \rangle$$
$$= \int_0^s ds_1 \langle \overline{P}^\sim S_{n-1}^{\overline{P}}(s_1) f, (\overline{Q}^\sim S(\sigma_o))^* g \rangle$$
$$= \int_0^s ds_1 \int_0^{s_1} ds_2 \langle \overline{P}^\sim S(\sigma_1) \overline{P}^\sim S_{n-2}^{\overline{P}}(s_2) f, (\overline{Q}^\sim S(\sigma_o))^* g \rangle$$
$$= \int_0^s ds_1 \int_0^{s_1} ds_2 \langle \overline{Q}^\sim S(\sigma_0) \overline{P}^\sim S(\sigma_1) \overline{P}^\sim S_{n-2}^{\overline{P}}(s_2) f, g \rangle$$
$$= \cdots \text{ etc.}$$
$$= \int_0^s ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \langle \overline{Q}^\sim S(\sigma_o) \overline{P}^\sim \cdots \overline{P}^\sim S(\sigma_n) f, g \rangle$$

by Lemma 5 HP 13.3, and since any strong Bôchner integral is a weak integral.

The integrand is majorized as (* denotes convolution)

$$|\langle \overline{Q}^\sim S(\sigma_o) \overline{P}^\sim \cdots \overline{P}^\sim S(\sigma_n) f, g \rangle| \leq \|\overline{Q}^\sim S(s - s_1)\|_\infty \cdot \|\overline{P}^\sim S(s_1 - s_2)\|_\infty \cdot \cdots \cdot \|\overline{P}^\sim S(s_n) f\| \cdot \|g\|.$$

For fixed $s > 0$ these functions are clearly measurable on $\mathbb{R}^n$. Furthermore, the last term is integrable on Δ_n^s since its integral is equal to $(\|\overline{Q}^\sim S(\cdot)\|_\infty * \|\overline{P}^\sim S(\cdot)\|_* \ldots * \|\overline{P}^\sim S(\cdot) f\|)(s) \cdot \|g\|$. Note that as multiple convolution (*) of integrable functions the last of which is continuous this is a realvalued continuous function of s. Consequently, $(s)_n \rightsquigarrow \langle \overline{Q}^\sim S(\sigma_o) \cdots \overline{P}^\sim S(\sigma_n) f, g \rangle$ is integrable on Δ_n^s and by Fubini's theorem the iterated integral

equals the multiple integral of the assertion in Lemma 2.6. ///

Note also, that in case $S(\cdot)$ is uniformly continuous the estimate

$$|\langle \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n) f, g\rangle| \leq (\|\overline{Q}^{\sim} S(\cdot)\|_\infty * \cdots * \|\overline{P}^{\sim} S(\cdot)\|_\infty)(s) \cdot \|f\| \cdot \|g\|$$

shows that Fubini's theorem applies to the iterated norm-Bôchner integrals defining $S^{\overline{P}}_n(\cdot)$ and therefore

$$\overline{Q}^{\sim} S^{\overline{P}}_n(s) = \int \chi_{\Delta^s_n} \overline{Q}^{\sim} \cdot S(\sigma_o) \cdots \overline{P}^{\sim}(\sigma_n) d(s)_n.$$

Corollary 2.1: If $\overline{P} \in \mathcal{P}_s(\overline{H})$ then $S^{\overline{P}}(\cdot)$ is a *symmetric* semigroup.

Proof: By Theorem 2.1, it is sufficient to show that all $S^{\overline{P}}_n(s)$ are symmetric, which in turn is equivalent to $\langle S^{\overline{P}}_n(s) f, f\rangle$ being real for all $f \in \mathcal{H}$. By Lemmas 2.5 and 2.6,

$$\langle S^{\overline{P}}_n(s) f, f\rangle = \int \chi_{\Delta^s_n} \langle S(\sigma_o) \cdot \overline{P} S(\sigma_1) \cdots \overline{P} S(\sigma_n) f, f\rangle$$

and

$$\begin{aligned}\overline{\langle S^{\overline{P}}_n(s) f, f\rangle} &= \int \chi_{\Delta^s_n} \langle f, S(\sigma_o) \overline{P} S(\sigma_1) \cdots \overline{P} S(\sigma_n) f\rangle \\ &= \int \chi_{\Delta^s_n} \langle S(\sigma_n) \overline{P}^* S(\sigma_{n-1}) \overline{P}^* \cdots \overline{P}^* S(\sigma_o) f, f\rangle \\ &= \int \chi_{\Delta^s_n} \langle S(\sigma_n) \overline{P} S(\sigma_{n-1}) \overline{P} \cdots \overline{P} S(\sigma_o) f, f\rangle \\ &= \int \chi_{\Delta^s_n} \langle S(\sigma_o) \overline{P} S(\sigma_1) \cdots \overline{P} S(\sigma_n) f, f\rangle \\ &= \langle S^{\overline{P}}_n(s) f, f\rangle.\end{aligned}$$

The last step is allowed due to the symmetric form of the domain of integration $\Delta^s_n := \{(s_n)_n | \sigma_i \geq 0 \ \sum \sigma_i = s\}$. ///

3. PERTURBATION OF GIBBS SEMIGROUPS

The results of Sec. 2 do not permit us to conclude that $S^{\overline{P}}(\cdot)$ is a Gibbs semigroup, if $S(\cdot)$ is. Nor is it valid to conclude that

$$\mathrm{tr}_{\mathcal{H}}(S^{\overline{P}}(s) T) = \sum_{n \geq 0} \mathrm{tr}\ (S^{\overline{P}}_n(s) \cdot T) \quad \text{if } T \in \mathfrak{B}_\infty(\mathcal{H}).$$

Throughout this section, $S(\cdot)$ will be a given Gibbs semigroup on $\mathcal{H}$ with Hamiltonian $\overline{H}$. $\overline{P}$ will be any element of $\mathcal{P}(\overline{H})$. $\overline{Q}$ will denote either $\overline{P}$ or $1|_{\mathcal{D}_{\overline{H}}}$. Δ^s_n, s_i, σ_i, $S^{\overline{P}}_n(\cdot)$, and $S^{\overline{P}}(\cdot)$ retain their previous meaning. By Proposition 2, Theorem 2.1 (ii) applies and hence by (iv) all functions $s \mapsto \overline{Q}^{\sim} \cdot S^{\overline{P}}_n(s)$ are norm continuous on $\mathbb{R}_+$ and integrable on compact subsets of $\overline{\mathbb{R}}_+$. Also in terms of (norm) Bôchner integrals,

$$\overline{Q}^{\sim} S^{\overline{P}}_n(s) = \int \chi_{\Delta^s_n} \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n) d(s)_n.$$

Lemma 3.1: Assume $T \in \mathfrak{B}_\infty(\mathcal{H})$ and let $(f) = (f_m)_{1 \leq m \leq M < \infty}$, $(g) = (g_m)_{1 \leq m \leq M < \infty}$ be any orthonormal sets in $\mathcal{H}$ (M variable).

(i) $T \in \mathfrak{B}_1(\mathcal{H})$ if and only if τ

$$\equiv \sup_{(f)(g)} \left| \sum_{m=1}^{M} \langle T f_m, g_m\rangle \right| < \infty.$$

(ii) If $\tau < \infty$, $|\mathrm{tr}_{\mathcal{H}}(T)| \leq \|T\|_1 = \mathrm{tr}_{\mathcal{H}}(T^* T)^{1/2} = \tau$.

(iii) If $T \in \mathfrak{B}^+_1(\mathcal{H})$, $0 \leq \mathrm{tr}_{\mathcal{H}}(T) = \|T\|_1 = \tau < \infty$.

Proof: (i) If $\tau < \infty$, then

$$\sum_{n \geq 1} \langle T h_n, h_n\rangle$$

is (absolutely) convergent for any orthonormal base $(h_n)_{n \geq 1}$ of $\mathcal{H}$ and $T \in \mathfrak{B}_1(\mathcal{H})$ follows. Conversely, if $0 \neq T \in \mathfrak{B}_1(\mathcal{H})$, then T has a normconvergent representation

$$T f = \sum_{n \geq 1} t_n \langle f, h_n\rangle \cdot h'_n$$

with orthonormal sets $(h_n)_{n \geq 1}$, $(h'_n)_{n \geq 1}$ and $t_n > 0$.

$$\begin{aligned}&\left| \sum_{m=1}^{M} \langle T f_m, g_m\rangle \right| \\ &\quad = \left| \sum_{m=1}^{M} \sum_{n \geq 1} t_n \langle f_m, h_n\rangle \langle h'_n, g_m\rangle \right| \\ &\quad \leq \sum_{n \geq 1} t_n \left| \sum_{m=1}^{M} \langle f_m, h_n\rangle \langle h'_n, g_m\rangle \right| \\ &\quad \leq \sum_{n \geq 1} t_n = \|T\|_1 < \infty.\end{aligned}$$

Taking $f_m = h_m$ and $g_m = h'_m$ for $m \geq 1$,

$$\sup_{M < \infty} \left| \sum_{m=1}^{M} \langle T h_m, h'_m\rangle \right| = \sup_{M < \infty} \sum_{m=1}^{M} t_m = \|T\|_1.$$

(ii) and (iii) are obvious. ///

Lemma 3.2: For any $s > 0$, $\overline{Q}^{\sim} \cdot S^{\overline{P}}_n(s)$ is in $\mathfrak{B}_1(\mathcal{H})$ and

$$\|\overline{Q}^{\sim} \cdot S^{\overline{P}}_n(s)\|_1 \leq \int \chi_{\Delta^s_n} \|\overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)\|_1 d(s)_n < \infty.$$

Proof: Let (f) and (g) be as in the previous lemma. Using lemma 2.6, we have

$$\begin{aligned}&\left| \sum_{m=1}^{M} \langle \overline{Q}^{\sim} \cdot S^{\overline{P}}_n(s) f_m, g_m\rangle \right| \\ &\quad \leq \int \chi_{\Delta^s_n} \left| \sum_{m=1}^{M} \langle \overline{Q}^{\sim} S(\sigma_o) \cdots \right. \\ &\qquad \left. \overline{P}^{\sim} S(\sigma_n) f_m, g_m\rangle \right| d(s)_n \\ &\quad \leq \int \chi_{\Delta^s_n} \|\overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)\|_1.\end{aligned}$$

Note, that for any $(s)_n \in \Delta^s_n$, there is at least one index j with $\sigma_j \geq s/(n+1)$, since otherwise $\sum \sigma_j < s$.

Define

$$\begin{aligned}\Delta^s_{n,n} &= \{(s)_n \in \Delta^s_n | \sigma_n \geq s/(n+1)\}, \\ \Delta^s_{n,n-1} &= \{(s)_n \in \Delta^s_n | \sigma_n < s/(n+1), \sigma_{n-1} \\ &\qquad \geq s/(n+1)\},\end{aligned}$$

$$\Delta^s_{n,j} = \{(s)_n \in \Delta^s_n \mid \sigma_n < s/(n+1), \cdots, \sigma_{j+1} < s/(n+1), \sigma_j \geq s/(n+1)\},$$
$$\Delta^s_{n,o} = \{(s)_n \in \Delta^s_n \mid \sigma_n < s/(n+1), \cdots, \sigma_o \geq s/(n+1)\}.$$

Then $\Delta^s_n = \bigcup_{j=o}^{n} \Delta^s_{n,j}$ as a disjoint union and

$$\chi_{\Delta^s_n} = \sum_{j=o}^{n} \chi_{\Delta^s_{n,j}}.$$

Let $j > 0$; then on $\Delta^s_{n,j}$:

$$\|\overline{Q}^{\sim} \cdot S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)\|_1 \leq \|\overline{Q}^{\sim} S(\sigma_o)\|_\infty \cdots \|\overline{P}^{\sim} S(\sigma_{j-1})\|_\infty \cdot \|\overline{P}^{\sim} S(\sigma_j)\|_1 \cdots \|\overline{P}^{\sim} S(\sigma_n)\|_\infty.$$

Note that $\sigma_j \rightsquigarrow \|P^{\sim} S(\sigma_j)\|_1$ is generally nonintegrable at 0. However on $\Delta^s_{n,j}$, $\sigma_j \geq s/(n+1) > 0$ and this singularity will not contribute. Since the factors on the right-hand side are continuous and integrable on $\Delta^s_{n,j}$, it follows that the dominated function is integrable there. Proceeding similarly for $j = 0$, one deduces

$$\int_{\Delta^s_n} d(s)_n \|\overline{Q}^{\sim} S(\sigma_o) \ldots P^{\sim} s(\sigma_n)\|_1 = \sum_{j=o}^{n} \int_{\Delta^s_{n,j}} d(s)_n \|\overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)\|_1 < \infty.$$

Varying (f) and (g) in the first estimate and using Lemma 3.1, Lemma 3.2 follows.///

Lemma 3.3: $\mathbb{R}_+ \ni s \rightsquigarrow Q^{\sim} \cdot S^{\overline{P}}_n(s)$ is $\|\cdot\|_1$-continuous.

Proof: This is true for $n = o$ by assumption. Assume Lemma 3.3 holds for $1 \leq j \leq n$. Fix $s_o > 0$ and choose $\delta > 0$ with $s_o - 2\delta > 0$.

Let s_1, s_2 satisfy $s_o - \delta < s_i < s_o + \delta, i = 1, 2$.

$$\begin{aligned}
&\|\overline{Q}^{\sim} \cdot S^{\overline{P}}_{n+1}(s_2) - \overline{Q}^{\sim} S^{\overline{P}}_{n+1}(s_1)\|_1 \\
&= \|\int_0^{s_2} ds\, \overline{Q}^{\sim} S(s_2 - s) \cdot \overline{P}^{\sim} S^{\overline{P}}_n(s) \\
&\quad - \int_0^{s_1} ds\, \overline{Q}^{\sim} S(s_1 - s) \cdot \overline{P}^{\sim} S^{\overline{P}}_n(s)\|_1 \\
&\leq \|\int_0^{s_o - 2\delta} ds\, \overline{Q}^{\sim} [S(s_2 - s) - S(s_1 - s)] \cdot \overline{P}^{\sim} S^{\overline{P}}_n(s)\|_1 \\
&\quad + \|\int_{s_o - 2\delta}^{s_2} ds\, \overline{Q}^{\sim} S(s_2 - s) \overline{P}^{\sim} S^{\overline{P}}_n(s)\|_1 \\
&\quad + \|\int_{s_o - 2\delta}^{s_1} ds\, \overline{Q}^{\sim} S(s_1 - s) \overline{P}^{\sim} S^{\overline{P}}_n(s)\|_1. \\
&\leq \int_0^{s_o - 2\delta} ds \|\overline{Q}^{\sim} [S(s_2 - s) - S(s_1 - s)]\|_1 \cdot \|\overline{P}^{\sim} S^{\overline{P}}_n(s)\|_\infty \\
&\quad + \int_{s_o - 2\delta}^{s_2} ds\, \|Q^{\sim} S(s_2 - s)\|_\infty \cdot \|\overline{P}^{\sim} S^{\overline{P}}_n(s)\|_1 \\
&\quad + \int_{s_o - 2\delta}^{s_1} ds\, \|\overline{Q}^{\sim} S(s_1 - s)\|_\infty \cdot \|P^{\sim} S^{\overline{P}}_n(s)\|_1.
\end{aligned}$$

The integrand of the first term is dominated by

$$s \rightsquigarrow 2(\sup_{o < \delta < \sigma < s_o + \delta} \|\overline{Q}^{\sim} S(\sigma)\|_1) \cdot \|\overline{P}^{\sim} S^{\overline{P}}_n(s)\|_\infty,$$

which is integrable on $[0, s_o - 2\delta]$. As $s_i \to s_o$ the integrand of this term tends to 0 for every $s \in [0, s_o - 2\delta]$ due to the $\|\cdot\|_1$-continuity of $s \rightsquigarrow \overline{Q}^{\sim} \times S(s')$ at $s' = s_o - s > \delta > 0$. Hence the first integral converges to 0 as $s_i \to s_o$ by the dominated convergence theorem.

The second integral is dominated by

$$\sup_{0 < s_o - 2\delta \leq \sigma \leq s_o + \delta} (\|\overline{P}^{\sim} S^{\overline{P}}_n(\sigma)\|_1) \cdot \int_0^{3\delta} ds \|\overline{Q}^{\sim} \cdot S(s)\|_\infty.$$

The first factor in this expression is finite by the induction hypothesis and decreases with δ, while the second factor goes to 0 as $\delta \to 0$.

The third integral is treated analogously. Thus, given $\epsilon > 0$, find $\delta > 0$ small enough to hold each of the last two contributions below $\frac{1}{3}\epsilon$ and then choose s_1, s_2 close enough to s_o to depress the first term below $\frac{1}{3}\epsilon$ also. By the preceding estimates this shows continuity for $j = n + 1$ and hence induction takes care of any n.///

Lemma 3.4: Let $T \in \mathfrak{B}_\infty(\mathcal{K})$, then for $s > 0$ $\mathrm{tr}_{\mathcal{K}}(T\, \overline{Q}^{\sim} S^{\overline{P}}_n(s)) = \int_{\Delta^s_n} d(s)_n \mathrm{tr}_{\mathcal{K}}(T\, \overline{Q}^{\sim} \cdot S(\sigma_o) \cdots P^{\sim} S(\sigma_n))$.

Proof: Let $(f_m)_{m \geq 1}$ be a CONB of $\mathcal{K}$ and let E_M denote the projector onto the subspace spanned by the first M elements of this basis. Clearly, $E_M \nearrow I_{\mathcal{K}}$ as $M \to \infty$:

$$\begin{aligned}
&\mathrm{tr}_{\mathcal{K}}(T\, \overline{Q}^{\sim} S^{\overline{P}}_n(s)) \\
&= \lim_{M \to \infty} \mathrm{tr}_{\mathcal{K}}(E_M T\, \overline{Q}^{\sim} S^{\overline{P}}_n(s)) \\
&= \lim_{M \to \infty} \sum_{m=1}^{M} \int_{\Delta^s_n} d(s)_n \langle T\, \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n) f_m, f_m \rangle \\
&= \lim_{M \to \infty} \int_{\Delta^s_n} d(s)_n \mathrm{tr}_{\mathcal{K}}(E_M T\, \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)).
\end{aligned}$$

Now

$$|\int_{\Delta^s_n} d(s)_n \mathrm{tr}_{\mathcal{K}}(E_M T \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n))| \leq \int_{\Delta^s_n} d(s)_n \|E_M T \overline{Q}^{\sim} S(\sigma_o) \cdots \overline{P}^{\sim} S(\sigma_n)\|_1$$

$$\leq \int_{\Delta_n^s} d(s)_n \, \|E_M\|_\infty \cdot \|T\|_\infty \cdot \|\overline{Q}^\sim S(\sigma_o) \cdots \times \overline{P}^\sim S(\sigma_n)\|_1$$

$$\leq \int_{\Delta_n^s} d(s)_n \|\overline{Q}^\sim S(\sigma_o) \cdots \overline{P}^\sim S(\sigma_n)\|_1 \cdot \|T\|_\infty < \infty,$$

with the last term independent of M.

Thus the integrands in the trace formula are bounded by an integrable function independently of M. By the dominated convergence theorem the limit $M \to \infty$ can be interchanged with the integral and the lemma follows. ///

Collecting these results we have:

Theorem 3.1: The functions $\mathbb{R}_+ \ni s \rightsquigarrow \overline{Q}^\sim S_n^{\overline{P}}(s)$ are $\mathfrak{B}_1(\mathfrak{K})$-valued and $\|\cdot\|_1$ — continuous. They allow the estimates

$$\|Q^\sim \cdot S_n^{\overline{P}}(s)\|_1 \leq \int_{\Delta_n^s} d(s)_n \cdot \|\overline{Q}^\sim S(\sigma_o) \cdots \overline{P}^\sim S(\sigma_n)\|_1 < \infty.$$

If $T \in \mathfrak{B}_\infty(\mathfrak{K})$,

$$\mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim S_n^{\overline{P}}(s)) = \int_{\Delta_n^s} d(s)_n \, \mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim S(\sigma_o) \cdots \overline{P}^\sim S(\sigma_n)).$$

Lemma 3.5: Assume $\overline{P}$ is bounded and $T \in \mathfrak{B}_\infty(\mathfrak{K})$. Then for $s > 0$

$$|\mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim S_n^{\overline{P}}(s)| \leq \frac{(\|\overline{P}^\sim\|_\infty \cdot s)^n}{n!} \cdot \|T\overline{Q}^\sim\|_\infty \cdot \mathrm{tr}_{\mathfrak{K}} S(s)$$

and

$$\|T\overline{Q}^\sim S_n^{\overline{P}}(s)\|_1 \leq \frac{(\|\overline{P}^\sim\|_\infty \cdot s)^n}{n!} \|T\overline{Q}^\sim\|_\infty \cdot \|S(s)\|_1.$$

Proof: Since $\|T\overline{Q}^\sim S_n^{\overline{P}}(s)\|_1$ is finite and equal to $|\mathrm{tr}_{\mathfrak{K}}(WT\overline{Q}^\sim S_n^{\overline{P}}(s)|$ for a suitable isometry W on $\mathfrak{K}$, the second inequality follows from the first.

If A_i, $i = 0, 1, \ldots, n$, are elements of $\mathfrak{B}_\infty(\mathfrak{K})$, Ginibre and Gruber[13] have proved the formula

$$|\mathrm{tr}_{\mathfrak{K}}(A_o \cdot S(\sigma_o) A_1 \cdots A_n \cdot S(\sigma_n))| \leq \|A_o\|_\infty \cdot \cdots \cdot \|A_n\|_\infty \cdot \mathrm{tr}_{\mathfrak{K}}(S(s))$$

where $\sum_{i=0}^n \sigma_i = s$.

Then, using Theorem 3.1, we have

$$|\mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim S_n^{\overline{P}}(s))|$$

$$= |\int_{\Delta_n^s} d(s)_n \, \mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim S(\sigma_o) \cdot \cdots \cdot \overline{P}^\sim S(\sigma_n))|$$

$$\leq \int_{\Delta_n^s} d(s)_n \, |\mathrm{tr}_{\mathfrak{K}}(T\overline{Q}^\sim \cdots S(\sigma_n))|$$

$$\leq \int_{\Delta_n^s}^n d(s)_n \, \|T\overline{Q}^\sim\|_\infty \cdot \|\overline{P}^\sim\|_\infty^n \cdot \mathrm{tr}_{\mathfrak{K}} S(s)$$

$$\leq \|T\overline{Q}^\sim\|_\infty \cdot \frac{(\|\overline{P}^\sim\|_\infty \cdot s)^n}{n!} \cdot \mathrm{tr}_{\mathfrak{K}} S(s),$$

since $\int_{\Delta_n^s} d(s)_n \, 1 = \frac{s^n}{n!}$. ///

Theorem 3.2: Let $\overline{P}$ and $T \in \mathfrak{B}_\infty(\mathfrak{K})$. Then $S^{\overline{P}}(s) = \sum_{n \geq 0} S_n^{\overline{P}}(s)$ is $\|\cdot\|_1$-convergent and

$$\mathrm{tr}_{\mathfrak{K}}(TS^{\overline{P}}(s)) = \sum_{n \geq 0} \mathrm{tr}_{\mathfrak{K}}(TS_n^{\overline{P}}(s)).$$

Proof: By lemma 3.5,

$$\|\sum_{N \leq n \leq N' < \infty} S_n^{\overline{P}}(s)\|_1 \leq \mathrm{tr}_{\mathfrak{K}}(S(s)) \cdot \sum_{N \leq n \leq N' < \infty} \frac{(\|\overline{P}^\sim\|_\infty \cdot s)^n}{n!},$$

so that $N \rightsquigarrow \sum_{0 \leq n \leq N} S_n^P(s)$ is a Cauchy sequence in $\|\cdot\|_1$. By the completeness of $\mathfrak{B}_1(\mathfrak{K})$, it converges to a limit in $\mathfrak{B}_1(\mathfrak{K})$ which has to be $S^{\overline{P}}(s)$. The second formula follows, since $\mathfrak{B}_1(\mathfrak{K}) \ni S \rightsquigarrow \mathrm{tr}_{\mathfrak{K}}(TS)$ is $\|\cdot\|_1$-continuous. ///

4. CONCLUSION

Theorem 3.2 should be improved, since the restriction to bounded $\overline{P}$ is too severe. Since all previous results in Sec. 3 are derived without this assumption only the convergence proof for the series $\sum_{n \geq 0} S_n^{\overline{P}}$ needs to be extended, however. The difficulty in doing this stems from the fact that in cases of interest $s \rightsquigarrow \|\overline{Q}^\sim S(s)\|_1$ is generally not integrable at 0, even though $s \rightsquigarrow \|\overline{Q}^\sim S(s)\|_\infty$ is (by assumption).

The individual terms of $\sum S_n^{\overline{P}}$ are in general not of a definite sign, so that

$$A \rightsquigarrow \mathrm{tr}_{\mathfrak{K}}(AS_n^{\overline{P}}(s))$$

$$= \sum_{n \geq 0} \mathrm{tr}_{\mathfrak{K}}(AS_n^{\overline{P}}(s)) = \mathrm{tr}_{\mathfrak{K}}(AS_0^{\overline{P}}(s)) + \lim_{N \to \infty} \sum_{n=1}^N \mathrm{tr}_{\mathfrak{K}}(AS_n^{\overline{P}}(s))$$

and only the first term is a state. Thus the perturbed state is approximated in the set of continuous linear functionals only and not within the set of states, as one might wish.

This development is applied to self-interacting fermion systems in a forthcoming paper. We expect to deal with bosons systems in a similar fashion.

---

* Supported in part by grant GP No. 18641 of the NSF.

1 J. M. Cook, Trans. Am. Math. Soc. 74, 222 (1953).

2 T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).

3 D. Ruelle in *Applications of Mathematics to Problems in Theoretical Physics*, Summer School 1965 Cargèse, Corsica, edited by F. Lurçat (Gordon-Breach, New York, 1967).

4 D. Ruelle, Commun. Math. Phys. 3, 133 (1966).

5 S. Miracle-Sole and D. W. Robinson, Commun. Math. Phys. 14, 235 (1969).

6 The basic reference for semigroups of interest is E. Hille and R. S. Phillips, *Functional Analysis and Semigroups*, revised edition (Am. Math. Soc., Providence, R. I., 1957). Our detailed references will generally be given as HP a.b. which will refer to Sec. a.b. of this treatise.

[7] D. A. Uhlenbrock, J. Math. Phys. 7, 885 (1966).
[8] Some relevant publications are T. Matsubara, Progr. Theoret. Phys. 14, 351, 1955; L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962); A. A. Abrikosov, L. P. Gor'kov and J. E. Dzyaloshinsky, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N. J., 1963); C. Bloch, Diagram Expansions in Quantum Statistical Mechanics in Studies in Statistical Mechanics (Wiley, New York, 1965), Vol. III. A more detailed list is to be found in Ref. 9.
[9] C. W. Gruber, thesis (Princeton University, 1968).
[10] R. T. Powers, thesis (Princeton University, 1967).
[11] I. E. Segal, Ann. Math. 63, 160 (1956).
[12] D. Shale and F. Stinespring, Ann. of Math. 80, 365 (1964).
[13] J. Ginibre and C. Gruber, Commun. Math. Phys. 11, 198 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Minimal Uncertainty States for Bounded Observables

Evangelos K. Ifantis
Nuclear Research Center "Democritos", Aghia Paraskevi, Attikis, Athens, Greece
(Received 7 June 1971)

The procedure of finding the states, which minimize the uncertainty product of two noncompatible observables, leads to the study of the approximate point spectrum of a non-self-adjoint operator. The spectrum of a large class of such bounded non-self-adjoint operators is studied. The results are applied to the theory of the oscillator phase operators.

I. INTRODUCTION

The question "what sort of states is associated with the minimum uncertainty product $(\Delta A)\cdot(\Delta B)$ of two noncompatible observables, represented by the self-adjoint operators A and B in Hilbert space" was one of the earliest problems of quantum mechanics. This problem arose recently in the theory of the oscillator phase operators[1-4] C and S ("cosine" and "sine") and was studied for a simple model of phase operators C and S.

In order to extend the results[1,2] for the uncertainty product $(\Delta C)(\Delta S)$ of the "simplest phase operators" to a large class of physically motivated generalized phase operators C and S, we have been led to study the uncertainty product $(\Delta A)\cdot(\Delta B)$ for a class of bounded observables A and B. More specifically, the procedure of finding the states, which minimize the uncertainty product of two observables leads to the study of the spectrum of a non-self-adjoint operator $T = rM + M^*$ $(r \neq 1)$. In this paper we study the spectrum of T for the class of hyponormal operators M.

The results, having a quite general character, are applied to the case mentioned above.

In Sec. II we formulate the problem and derive the general form of the non-self-adjoint operator T. The point spectrum of this operator is associated with normalizable states, which minimize the uncertainty product $(\Delta A)\cdot(\Delta B)$ of the observables $A = (M^* + M)/2, B = (M^* - M)/2i$, and the continuous spectrum is associated with normalizable sequences, which tend to minimize the uncertainty product and which we call "minimal uncertainty sequences".

In Sec. III we prove that if M is hyponormal, then T and M have the same minimal uncertainty sequences. This result is very useful for the determination of nonnormalizable states minimizing an uncertainty product.

In Sec. IV it is found that if the spectrum of M has the same structure as the spectrum of the unilateral shift operator V, then the spectrum of T has the same structure as the well-known spectrum of the operator $rV + V^*$, provided that the spectrum of T satisfies certain conditions.

In Sec. V the above results are applied in the theory of the oscillator phase operators.

II. FORMULATION OF THE PROBLEM

The procedure for the determination of states, which minimize the uncertainty product $(\Delta A)\cdot(\Delta B)$ of two noncompatible observables A, B, is the following.

Let $ABf - BAf = iCf$, where $f \in D[(A,B)]$, the definition domain of the commutator $[A,B] = AB - BA$. Denote by $\langle A\rangle$ the expectation value (Af,f) for $\|f\| = 1$ and set $A_0 = A - \langle A\rangle$, $B_0 = B - \langle B\rangle$, $T_0 = A_0 + i\rho B_0$, $T_0^* = A_0 - i\rho B$. Then $T_0T_0^* = A_0^2 + \rho C + \rho^2 B_0^2$ and assuming that

$$\|A_0 f\| \neq 0, \qquad \|B_0 f\| \neq 0, \qquad (Cf,f) \neq 0, \tag{1}$$

we get

$$\|T_0^* f\|^2 = (f, T_0T_0^* f) = \|A_0 f\|^2 + \rho^2\|B_0 f\|^2 + \rho\langle C\rangle \tag{2}$$

for every real ρ. Thus $\langle C\rangle^2 - 4\|A_0 f\|^2\cdot\|B_0 f\|^2 \leq 0$, i.e., with $\Delta A = \|A_0 f\|$, $\Delta B = \|B_0 f\|$, we have

$$(\Delta A)^2\cdot(\Delta B)^2 \geq \tfrac{1}{4}\langle C\rangle^2. \tag{3}$$

Setting $\rho = -\langle C\rangle/2\|B_0 f\|^2$ in Eq. (2), we obtain

$$\|T^* f\|^2 = \frac{4\|A_0 f\|^2\cdot\|B_0 f\|^2 - \langle C\rangle^2}{4\|B_0 f\|^2}. \tag{4}$$

We are interested in states which minimize the uncertainty product (3), i.e., according to (4) we are interested in the following.

[7] D. A. Uhlenbrock, J. Math. Phys. 7, 885 (1966).
[8] Some relevant publications are T. Matsubara, Progr. Theoret. Phys. 14, 351, 1955; L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962); A. A. Abrikosov, L. P. Gor'kov and J. E. Dzyaloshinsky, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N. J., 1963); C. Bloch, Diagram Expansions in Quantum Statistical Mechanics in Studies in Statistical Mechanics (Wiley, New York, 1965), Vol. III. A more detailed list is to be found in Ref. 9.
[9] C. W. Gruber, thesis (Princeton University, 1968).
[10] R. T. Powers, thesis (Princeton University, 1967).
[11] I. E. Segal, Ann. Math. 63, 160 (1956).
[12] D. Shale and F. Stinespring, Ann. of Math. 80, 365 (1964).
[13] J. Ginibre and C. Gruber, Commun. Math. Phys. 11, 198 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Minimal Uncertainty States for Bounded Observables

Evangelos K. Ifantis
Nuclear Research Center "Democritos", Aghia Paraskevi, Attikis, Athens, Greece
(Received 7 June 1971)

The procedure of finding the states, which minimize the uncertainty product of two noncompatible observables, leads to the study of the approximate point spectrum of a non-self-adjoint operator. The spectrum of a large class of such bounded non-self-adjoint operators is studied. The results are applied to the theory of the oscillator phase operators.

I. INTRODUCTION

The question "what sort of states is associated with the minimum uncertainty product $(\Delta A)\cdot(\Delta B)$ of two noncompatible observables, represented by the self-adjoint operators A and B in Hilbert space" was one of the earliest problems of quantum mechanics. This problem arose recently in the theory of the oscillator phase operators[1-4] C and S ("cosine" and "sine") and was studied for a simple model of phase operators C and S.

In order to extend the results[1,2] for the uncertainty product $(\Delta C)(\Delta S)$ of the "simplest phase operators" to a large class of physically motivated generalized phase operators C and S, we have been led to study the uncertainty product $(\Delta A)\cdot(\Delta B)$ for a class of bounded observables A and B. More specifically, the procedure of finding the states, which minimize the uncertainty product of two observables leads to the study of the spectrum of a non-self-adjoint operator $T = rM + M^*$ $(r \neq 1)$. In this paper we study the spectrum of T for the class of hyponormal operators M.

The results, having a quite general character, are applied to the case mentioned above.

In Sec. II we formulate the problem and derive the general form of the non-self-adjoint operator T. The point spectrum of this operator is associated with normalizable states, which minimize the uncertainty product $(\Delta A)\cdot(\Delta B)$ of the observables $A = (M^* + M)/2, B = (M^* - M)/2i$, and the continuous spectrum is associated with normalizable sequences, which tend to minimize the uncertainty product and which we call "minimal uncertainty sequences".

In Sec. III we prove that if M is hyponormal, then T and M have the same minimal uncertainty sequences. This result is very useful for the determination of nonnormalizable states minimizing an uncertainty product.

In Sec. IV it is found that if the spectrum of M has the same structure as the spectrum of the unilateral shift operator V, then the spectrum of T has the same structure as the well-known spectrum of the operator $rV + V^*$, provided that the spectrum of T satisfies certain conditions.

In Sec. V the above results are applied in the theory of the oscillator phase operators.

II. FORMULATION OF THE PROBLEM

The procedure for the determination of states, which minimize the uncertainty product $(\Delta A)\cdot(\Delta B)$ of two noncompatible observables A, B, is the following.

Let $ABf - BAf = iCf$, where $f \in D[(A,B)]$, the definition domain of the commutator $[A,B] = AB - BA$. Denote by $\langle A\rangle$ the expectation value (Af,f) for $\|f\| = 1$ and set $A_0 = A - \langle A\rangle$, $B_0 = B - \langle B\rangle$, $T_0 = A_0 + i\rho B_0$, $T_0^* = A_0 - i\rho B$. Then $T_0T_0^* = A_0^2 + \rho C + \rho^2 B_0^2$ and assuming that

$$\|A_0 f\| \neq 0, \qquad \|B_0 f\| \neq 0, \qquad (Cf,f) \neq 0, \tag{1}$$

we get

$$\|T_0^* f\|^2 = (f, T_0T_0^* f) = \|A_0 f\|^2 + \rho^2\|B_0 f\|^2 + \rho\langle C\rangle \tag{2}$$

for every real ρ. Thus $\langle C\rangle^2 - 4\|A_0 f\|^2\cdot\|B_0 f\|^2 \leq 0$, i.e., with $\Delta A = \|A_0 f\|$, $\Delta B = \|B_0 f\|$, we have

$$(\Delta A)^2\cdot(\Delta B)^2 \geq \tfrac{1}{4}\langle C\rangle^2. \tag{3}$$

Setting $\rho = -\langle C\rangle/2\|B_0 f\|^2$ in Eq. (2), we obtain

$$\|T^* f\|^2 = \frac{4\|A_0 f\|^2\cdot\|B_0 f\|^2 - \langle C\rangle^2}{4\|B_0 f\|^2}. \tag{4}$$

We are interested in states which minimize the uncertainty product (3), i.e., according to (4) we are interested in the following.

(1) normalizable states, which are eigenstates of T^* with eigenvalue zero and

(2) normalizable sequences $\{f_n\}$ such that

$$\lim \|T_0^* f_n\| = 0 \quad \text{as} \quad n \to \infty. \tag{5}$$

Setting

$$\gamma = \langle C \rangle / 2 \cdot \|B_0 f\|^2, \tag{6}$$

the relations $\|T_0^* f\| = 0$ and (5) take the forms

$$Af + i\gamma Bf - \lambda I f = 0, \quad \lambda = \langle A \rangle + i\gamma \langle B \rangle \tag{7}$$

and

$$\lim \|(A + i\gamma B - \lambda I) f_n\| = 0 \quad \text{as } n \to \infty. \tag{8}$$

In previous work,[1,2] Eq. (7) was considered as an eigenvalue equation with three free parameters γ, Reλ, and Imλ. In the following we consider the spectrum of the operator

$$L = A + i\gamma B, \tag{9}$$

with the free parameter γ. According to the relations (7) and (8) we are interested in the approximate point spectrum of the operator (9), i.e., the points λ in the spectrum of L, for which the operator $L - \lambda I$ is not bounded from below. In particular, we are interested in the point spectrum of L, i.e., the eigenvalues of L.

We set $M = A - iB$, $M^* = A + iB$. Then $L = \frac{1}{2}[(1 - \gamma)M + (1 + \gamma)M^*] = [(1 + \gamma)/2] \cdot T$, where

$$T = rM + M^*, \qquad r = (1 - \gamma)/1 + \gamma) \tag{10}$$

Due to (1) and (6) we have $r \neq 1$. Thus, we have to study the spectrum of a non-self-adjoint operator. Since the approximate point spectrum of every operator is not empty it follows that:

There exist always normalizable states or normalizable sequences of states or both, minimizing the uncertainty product (3).

III. MINIMAL UNCERTAINTY SEQUENCES

In the following we shall study the approximate point spectrum of T for a large class of bounded non-self-adjoint operators M. This class is the class of hyponormal operators, i.e., operators for which $\|Mf\| \geq \|M^* f\|$ for every f or equivalently the self-commutator $M^*M - MM^*$ is nonnegative.

Definition: A (nonconstant) sequence of states f_n is called a minimal uncertainty sequence for T if there exists a λ such that $\lim \|(T - \lambda I) f_n\| = 0$ as $n \to \infty$.

Theorem 1: If $r > 1$ ($r < 1$) and M is hyponormal, then M and $T(T^*)$ have the same uncertainty sequences.

Proof: From the relation

$$T^*T - TT^* = (r^2 - 1)(M^*M - MM^*),$$

we see that T is also hyponormal, i.e., $\|Tf\| \geq \|Tf^*\|$ in case $r > 1$ (and $\|T^*f\| \geq \|Tf\|$ in case $r < 1$).

Let $\lambda = \mu e^{i\varphi}$, $\varphi \in [0, 2\pi]$, and $\mu \geq 0$ such that $\{f_n\}_1^\infty$ is a minimal uncertainty sequence for M, i.e., $\lim \|(M - \mu e^{i\varphi} I) f_n\| = 0$ as $n \to \infty$. Then, since $M - \lambda I$ is hyponormal, we have

$$\begin{aligned} \|(T - \{r\mu e^{i\varphi} + \mu e^{-i\varphi}\} I) f_n\| &= \|r(M - \mu e^{i\varphi} I) f_n \\ &+ (M - \mu e^{-i\varphi} I) f_n\| \leq r\|(M - \mu e^{i\varphi} I) f_n\| \\ &+ \|(M^* - \mu e^{-i\varphi} I) f_n\| \\ &\leq (r + 1)\|(M - \mu e^{i\varphi} I) f_n\| \to 0, \end{aligned}$$

i.e., f_n is a minimal uncertainty sequence for T.

Let conversely $\lim \|(T - \mu e^{i\varphi}) f_n\| = 0$ as $n \to \infty$. Then, from the relation

$$M = [1/(r^2 - 1)](rT - T^*),$$

we have

$$(M - kI) f_n = [1/(r^2 - 1)]\,[r(T - \mu e^{i\varphi} I) f_n - (T^* - \mu e^{-i\varphi} I) f_n]$$

where

$$k = [1/(r^2 - 1)](r\mu e^{i\varphi} - \mu e^{-i\varphi})$$

Since $r > 1$ and $\|(T^* - \mu e^{-i\varphi} I) f_n\| \leq \|(T - \mu e^{i\varphi} I) f_n\|$,

we have

$$\|(M - kI) f_n\| \leq [1/(r - 1)]\,\|(T - \mu e^{i\varphi} I) f_n\| \to 0.$$

(The proof for the case $r < 1$ is similar).

We observe that if the approximate point spectrum of M covers the entire unit disk, then the approximate point spectrum of T covers the entire ellipse:

$$\lambda = re^{i\varphi} + e^{-i\varphi}, \quad 0 \leq \varphi \leq 2\pi. \tag{11}$$

IV. A THEOREM FOR THE SPECTRUM OF T

Let M be a hyponormal operator ($M^*M - MM^* \geq 0$), whose spectrum is the entire unit disk, and let us assume that:

(a) The spectrum of M has the same structure as the spectrum of the unilateral shift V, i.e., the interior of the unit disk is the residual spectrum and the periphery the continuous spectrum of M.

(b) The interior D and the boundary R of the ellipse (11) are included in $\sigma(T)$, the spectrum of T, i.e.,

$$D \cup R \subseteq \sigma(T). \tag{12}$$

As we shall see later, Assumption (b) is always satisfied if the operator $M - V$ is compact. It also follows easily, as in the proof of Theorem 1, in case the spectrum of M is purely continuous.

Under the Assumptions (a) and (b) we shall prove the following theorem:

Theorem 2: R is the continuous spectrum of T and D is in case $r > 1$ the residual spectrum and in case $r < 1$ the point spectrum.

The proof is divided into six lemmas.

Lemma 1: $D \cup R = \sigma(T)$.

Proof: Denote by $W(T)$ the closure of the numerical range of T and set

$$(Mf,f) = \mu e^{i\varphi},\ 0 \le \mu \le 1,\ 0 \le \varphi \le 2\pi,\ \|f\| = 1$$

[this is possible because for hyponormal operators[5]

$\|M\|$ = spectral radius of $M \equiv \max|\lambda| : \lambda \in \sigma(M)$].

Then $(M^*f,f) = (f,Mf) = (Mf,f)^* = \mu e^{-i\varphi}$ and because of (10) $(Tf,f) = r\mu e^{i\varphi} + \mu e^{-i\varphi}$, $\mu \in [0,1]$, $\varphi \in [0,2\pi]$. Thus

$$W(T) \subseteq D \cup R. \tag{13}$$

Since always $\sigma(T) \subseteq W(T)$, we conclude from (13) that

$$\sigma(T) \subseteq D \cup R. \tag{14}$$

From (12) and (14) it follows, $\sigma(T) = D \cup R$.

Corollary: R belongs to the approximate point spectrum of T.

This follows from Lemma 1 and the fact that for every operator the boundary of the spectrum is included in the approximate point spectrum (see Ref. 6, problem 63).

Lemma 2: For $r > 1$ the point spectrum of T is empty.

Proof: Assume that there exists an element f such that $Tf = \lambda f$ or $(rM + M^*)f = (r\mu e^{i\varphi} + \mu e^{-i\varphi})f$ $\mu \in [0,1]$, $\varphi \in [0,2\pi]$.

Then

$$r(M - \mu e^{i\varphi}I)f = -(M^* - \mu e^{-i\varphi}I)f \tag{15}$$

and $r\|(M - \mu e^{i\varphi}I)f\| = \|(M^* - \mu e^{-i\varphi}I)f\|$. But since M is hyponormal so is $M - \mu e^{i\varphi}I$. Thus

$$\|(M^* - \mu e^{-i\varphi}I)f\| \ge r\|(M^* - \mu e^{-i\varphi}I)f\|$$

or

$$(1 - r)\|(M^* - \mu e^{-i\varphi}I)f\| \ge 0. \tag{16}$$

Since $r > 1$, relation (16) is possible only if $M^*f = \mu e^{-i\varphi}f$. But then from (15) we must have $Mf = \mu e^{i\varphi}f$, which is impossible because the point spectrum of M is by the Assumption (a) empty.

Lemma 3: For $r < 1$ the residual spectrum of T is empty.

Proof: If there exists a point in the residual spectrum of T, it must be an eigenvalue of $T^* = r(1/rM + M^*)$, i.e., we must have $(1/rM + M^*)f = \lambda f$ for some f and $1/r > 1$. This is impossible, because of Lemma 2.

Corollary: For $r < 1$, T has a purely approximate point spectrum.

Lemma 4: Eigenvalues of T on the boundary R of the ellipse (11) do not exist.

Proof: If $r > 1$, Lemma 4 follows from Lemma 2. If $r < 1$, assume that $Tf_0 = \lambda f_0$, $\lambda = re^{i\varphi} + e^{-i\varphi}$, $\varphi \in [0,2\pi]$. Then since T has a purely approximate point spectrum, the value λ^* must be an eigenvalue of T^* because if λ^* belongs to the continuous spectrum of T^*, then

$$0 = (f,(T - \lambda I)f_0) = ((T^* - \lambda^* I)f, f_0) \quad \text{for every } f,$$

i.e., $f_0 = 0$.

Thus,

$$(1/rM + M^*)g = (\lambda^*/r)g \quad \text{for some } g \text{ and } 1/r > 1,$$

which is impossible.

Lemma 5: For $r > 1$, every point in D belongs to the residual spectrum of T and hence to the point spectrum of T^*.

Proof: According to the Assumption (b), we only have to prove that the operator

$$rM + M^* - (r\mu e^{i\varphi} + \mu e^{-i\varphi})I, \quad \mu \in [0,1], \quad \varphi \in [0,2\pi]$$

is bounded from below, i.e.,

$$\|(rM + M^* - (r\mu e^{i\varphi} + \mu e^{-i\varphi})I]f\| \ge \alpha\|f\|$$

for every f and $\mu \in [0,1]$ and some positive α. This follows easily because of the hyponormality of M, i.e., we have

$$\|[rM + M^* - (r\mu e^{i\varphi} + \mu e^{-i\varphi})I]f\| \ge r\|(M - \mu e^{i\varphi}I)f\| - \|(M^* - \mu e^{-i\varphi}I)f\| \ge (r - 1)\cdot\|(M - \mu e^{i\varphi}I)f\|.$$

This completes the proof because $r > 1$ and $M - \mu e^{i\varphi}I$ for $\mu \in [0,1]$, $\varphi \in [0,2\pi]$ is bounded from below by hypothesis (a).

Lemma 6: For every $r < 1$, every point in D belongs to the residual spectrum of T^* and, hence to the point spectrum of T.

Proof: We have

$$\|(rM^* + M - \{r\mu e^{i\varphi} + \mu e^{-i\varphi}\}I)f\| \ge \|(M - \mu e^{-i\varphi}I)f\| - r\|(M^* - \mu e^{i\varphi}I)f\| \ge (1 - r)\cdot\|(M - \mu e^{-i\varphi}I)f\|.$$

Lemmas 1–6 complete the proof of the theorem.

V. APPLICATION TO OSCILLATOR PHASE OPERATORS

First we examine the well-known results[2] for the simplest model of the oscillator phase operators C and S ("cosine" and "sine").

The simplest phase operators are the following[4]:

$$C = (V^* + V)/2, \quad S = (V^* - V)/2i, \tag{17}$$

where V is the unilateral shift operator V; $Ve_n = e_{n+1}$, $n = 1, 2, \ldots$, and $\{e_n\}_1^\infty$ is the orthonormal basis of eigenstates of the oscillator number operator N: $Ne_n = (n-1)e_n$, $n = 1, 2, \ldots$.

For the operators (17), we have

$$CS - SC = \tfrac{1}{2}i\cdot(I - P), \tag{18}$$

where $P = VV^*$ projects on the subspace, which is the orthogonal complement of the subspace spanned by the element e_1 (the ground state of N). We have $M = C - iS$, $M^* = C + iS$, and

$$T = rV + V^*. \tag{19}$$

The spectrum of the operator (19) is well known and can be found in Ref. 7. Also it can be obtained with the help of the representation of the operator (19) in the Hardy–Lebesgue space (see Ref. 8). The spectrum of the operator (19) is the entire ellipse (11). For $r > 1$, the point spectrum is empty and, for $r < 1$, the point spectrum is the interior of the ellipse (11). The boundary of the ellipse belongs to the continuous spectrum.

From (18) and (6) we see that $\gamma > 0$ (except for the states which are orthogonal to e_1, in which $\gamma = 0$)[9], i.e., $r < 1$. Thus, the point spectrum of (19) is not empty, i.e., the uncertainty product of the phase operators (17) can be minimized by normalizable states.

Due to Theorem 1, the minimal uncertainty sequences for the operator (19) are the same as those of the unilateral shift V, i.e., they have the form

$$f_n = n^{-1/2}\cdot\sum_{k=1}^{n} e^{i\varphi k}e_k, \quad 0 \le \varphi \le 2\pi. \tag{20}$$

The general form of the oscillator phase operators "cosine" and "sine" is the following[4]:

$$C = (V^*A + AV)/2, \quad S = (V^*A - AV)/2i, \tag{21}$$

where A is a self-adjoint diagonal operator (A: $Ae_n = a(n)e_n$), such that the diagonal $a(n) \neq 0$, $n = 2, 3, \cdots$, converges to unity and the spectrum of C and S is the entire interval $[-1, 1]$. We have

$$CS - SC = (i/2)(V^*A^2V - AVV^*A).$$

Thus, if AV is hyponormal, we have $\gamma > 0$ and $r < 1$. The operator (10) can be written

$$T = rAV + V^*A = rV + V^* + r(A - I)V + V^*(A - I), \tag{22}$$

where $r(A - I)V + V^*(A - I)$ is compact.[4]

Due to Weyl's theorem for non-self-adjoint operators, the operators (22) and (19) have the same spectrum except for eigenvalues (see Ref. 6, problem 143). Since operator (19), in case $r > 1$ does not have eigenvalues, we conclude that the entire ellipse (11) belongs to the spectrum of T.

We now consider the class of phase operators, which are constructed from sequences $a(n) \neq 0$, $n = 2, 3, \cdots$, which converges to unity monotonically from below.

Proposition 1: $M = AV$ is hyponormal ($M^*M - MM^* \geq 0$) if and only if the sequence $a(n)$ is absolutely monotone nondecreasing.

Proof: We observe that the relation $MM^* - M^*M \geq 0$ is impossible. In fact, if it holds, then we must have

$$[(AVV^*A - V^*A^2V)e_n, e_n] = a^2(n) - a^2(n+1) \geq 0 \quad \text{for } n > 1$$

and

$$[(AVV^*A - V^*A^2V)e_1, e_1] = -a^2(2) \geq 0,$$

i.e., $a(2) = a(3) = \cdots 0$. The proof of the relation $M^*M - MM^* \geq 0$ is obvious.

It remains to prove the following proposition in order to apply Theorem 2.

Proposition 2: For $a(n) \neq 0$, $n = 2, 3, \cdots$, and monotonically convergent to unity from below the spectrum of AV has the same structure as the spectrum of the unilateral shift operator.

The spectrum of AV was studied in Ref. 10 for general classes of sequences $a(n)$. In particular for monotonically convergent sequences it was studied in Ref. 3. We give below the proof that concerns us.

Proof: Denote by D and $\overline{D}$ the open and closed unit disk, respectively. Since the operator $M = AV$ can be written in the form $M = V + (A - I)V$, where $(A - I)V$ is compact, we have according to Weyl's theorem

$$D \subseteq \sigma(M). \tag{23}$$

On the other hand, since M is hyponormal, we have

$$\text{spectral radius of } M = \|M\| = 1 \quad [\text{since } \lim a(n) = 1].$$

Thus

$$\sigma(M) \subseteq \overline{D}. \tag{24}$$

From (23) and (24) it follows that $\sigma(M) = \overline{D}$.

Since $a(n) \neq 0$, $n = 2, 3, \cdots$, it is easy to see that the point spectrum of M is empty. Thus the residual spectrum of M^* is empty. From a relatively general theorem[11] concerning perturbation of the operator V^*, we know that for every compact op-

erator K the open unit disk D belongs to the point spectrum of $V^* + K$. Thus, from $M^* = V^*A = V^* + (A - I)V^*$, we conclude that D belongs to the point spectrum of M^* and hence to the residual spectrum of M (since the point spectrum of M is empty). We can prove as in Lemma 4 that $\bar{D} - D$ is the continuous spectrum of M^* and M.

Conclusion 1: Normalizable states minimizing the uncertainty product $(\Delta C)(\Delta S)$ do exist for every pair of phase operators C and S, constructed from sequences $a(n) \neq 0$, $n = 2, 3, \cdots$, which converge monotonically to unity from below.

Conclusion 2: According to Theorem 1, the minimal uncertainty sequences for T are the same as those[3] of AV. The knowledge of the continuous spectrum of $AV(\sigma_c(AV) = e^{i\varphi}, 0 \leq \varphi \leq 2\pi)$ enables us to find easily the minimal uncertainty sequences from the realization of the equation $AVf = e^{i\varphi}f$ in the space $l_2(1, \infty)$, i.e., from the solution of a first-order difference equation.

1 P. Carruthers and M. Nieto, Rev. Mod. Phys. 40, 411 (1965).
2 R. Jackiw, J. Math. Phys. 9, 339 (1968).
3 E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. 11, 1679 (1970).
4 E. K. Ifantis, J. Math. Phys. 12, 1021 (1971).
5 J. G. Stampfli, Pacific J. Math. 12, 1453 (1962).
6 P. R. Halmos, *A Hilbert Space Problem Book* (Nostrand, Princeton, N.J., 1967).
7 P. L. Duren, Trans. Am. Math. Soc. 99, 320 (1961).
8 E. K. Ifantis, J. Math. Phys. 11, 3138 (1970).
9 Without restriction of the generality, we assume $r > 0$. In case that r is negative or in general complex, we write $r = |r|e^{2ik}$. Then $T = rV + V^* = e^{ik}W(|r|V + V^*)W^*$, where W is the unitary operator W: $We_n = e^{ikn}e_n$, i.e., the spectrum of T is a rotation of the spectrum of $|r|V + V^*$, rotating by $k = \arg(r/2)$.
10 W. C. Ridge, Trans. Am. Math. Soc. 147, 349 (1970).
11 J. G. Stampfli, J. London Math. Soc. 40, 345 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

On Clifford Numbers, Dirac and Relativistic Hamilton–Jacobi Equations

Jerome K. Percus*
Courant Institute of Mathematical Sciences, New York University, New York, New York
and
Nikitas L. Petrakopoulos†
Department of Physics, New York University, New York, New York

(Received 8 June 1970)

Conditions are established under which the Dirac equation for an electron in an electromagnetic field has an exact semiclassical solution. In other words, the phase is identified with a solution of the corresponding relativistic Hamilton–Jacobi equation and the spinor amplitude has no explicit dependence upon $\hbar$. The complete set of admissible fields is determined for one-dimensional and stationary three-dimensional systems, while an extensive class is indicated for the general time-dependent problem.

1. INTRODUCTION

The Hamilton–Jacobi equation forms a bridge between classical particle mechanics and the quantum mechanics of spinless particles. This bridge becomes a computational tool in semiclassical quantum theory, or more properly from a modern viewpoint, in the WKB approximation in quantum mechanics. In fact, one can construct a classical potential which, used with WKB, exactly reproduces quantum mechanics[1]—but the construction is, of course, as difficult as the solution of the true quantum mechanical problem. It is the purpose of this paper to show that a very similar situation exists for a Dirac electron in an electromagnetic field, based upon the relativistic Hamilton–Jacobi equation. However, a totally arbitrary field cannot now be subsumed by the classical format, with its insistence upon coherent phase for the full wavefunction, and we shall spell out in detail the nature of the required restrictions.

2. BASIC SOLUTION

According to classical relativistic dynamics,[2] the Hamilton–Jacobi equation for a charged particle of charge $-e$ and proper mass m is of the form

$$\xi_5^2 = \xi_\rho \xi_\rho + \xi_4^2, \quad \rho = 1, 2, 3 \text{ (summation convention)},$$

$$\text{where} \quad \xi_5 = \frac{1}{c}\left(\frac{\partial S}{\partial t} - eV\right), \quad \xi_4 = mc, \tag{1}$$

$$\text{and} \quad \xi_\rho = \frac{\partial S}{\partial x_\rho} + \frac{e}{c} A_\rho.$$

The electromagnetic field is given by $(\mathbf{A}, V)$ and $S(x_1, x_2, x_3, t; x^*)$ is the two-event characteristic function; we regard $(x^*) = (x_1^*, x_2^*, x_3^*, t^*)$ as constant.

Let us now "linearize" in precisely the same way that the Klein–Gordon equation gives rise to the Dirac equation, i.e., we introduce the Dirac matrices

$$\alpha_1 = \left(\begin{array}{cc|cc} & & 0 & 1 \\ & \bigcirc & 1 & 0 \\ \hline 0 & 1 & & \\ 1 & 0 & \bigcirc & \end{array}\right), \quad \alpha_2 = \left(\begin{array}{cc|cc} & & 0 & -i \\ & \bigcirc & i & 0 \\ \hline 0 & -i & & \\ i & 0 & \bigcirc & \end{array}\right),$$

erator K the open unit disk D belongs to the point spectrum of $V^* + K$. Thus, from $M^* = V^*A = V^* + (A - I)V^*$, we conclude that D belongs to the point spectrum of M^* and hence to the residual spectrum of M (since the point spectrum of M is empty). We can prove as in Lemma 4 that $\bar{D} - D$ is the continuous spectrum of M^* and M.

Conclusion 1: Normalizable states minimizing the uncertainty product $(\Delta C)(\Delta S)$ do exist for every pair of phase operators C and S, constructed from sequences $a(n) \neq 0$, $n = 2, 3, \cdots$, which converge monotonically to unity from below.

Conclusion 2: According to Theorem 1, the minimal uncertainty sequences for T are the same as those[3] of AV. The knowledge of the continuous spectrum of $AV(\sigma_c(AV) = e^{i\varphi}, 0 \leq \varphi \leq 2\pi)$ enables us to find easily the minimal uncertainty sequences from the realization of the equation $AVf = e^{i\varphi}f$ in the space $l_2(1, \infty)$, i.e., from the solution of a first-order difference equation.

1 P. Carruthers and M. Nieto, Rev. Mod. Phys. 40, 411 (1965).
2 R. Jackiw, J. Math. Phys. 9, 339 (1968).
3 E. C. Lerner, H. W. Huang, and G. E. Walters, J. Math. Phys. 11, 1679 (1970).
4 E. K. Ifantis, J. Math. Phys. 12, 1021 (1971).
5 J. G. Stampfli, Pacific J. Math. 12, 1453 (1962).
6 P. R. Halmos, *A Hilbert Space Problem Book* (Nostrand, Princeton, N.J., 1967).
7 P. L. Duren, Trans. Am. Math. Soc. 99, 320 (1961).
8 E. K. Ifantis, J. Math. Phys. 11, 3138 (1970).
9 Without restriction of the generality, we assume $r > 0$. In case that r is negative or in general complex, we write $r = |r|e^{2ik}$. Then $T = rV + V^* = e^{ik}W(|r|V + V^*)W^*$, where W is the unitary operator W: $We_n = e^{ikn}e_n$, i.e., the spectrum of T is a rotation of the spectrum of $|r|V + V^*$, rotating by $k = \arg(r/2)$.
10 W. C. Ridge, Trans. Am. Math. Soc. 147, 349 (1970).
11 J. G. Stampfli, J. London Math. Soc. 40, 345 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

On Clifford Numbers, Dirac and Relativistic Hamilton–Jacobi Equations

Jerome K. Percus*
Courant Institute of Mathematical Sciences, New York University, New York, New York
and
Nikitas L. Petrakopoulos†
Department of Physics, New York University, New York, New York

(Received 8 June 1970)

Conditions are established under which the Dirac equation for an electron in an electromagnetic field has an exact semiclassical solution. In other words, the phase is identified with a solution of the corresponding relativistic Hamilton–Jacobi equation and the spinor amplitude has no explicit dependence upon $\hbar$. The complete set of admissible fields is determined for one-dimensional and stationary three-dimensional systems, while an extensive class is indicated for the general time-dependent problem.

1. INTRODUCTION

The Hamilton–Jacobi equation forms a bridge between classical particle mechanics and the quantum mechanics of spinless particles. This bridge becomes a computational tool in semiclassical quantum theory, or more properly from a modern viewpoint, in the WKB approximation in quantum mechanics. In fact, one can construct a classical potential which, used with WKB, exactly reproduces quantum mechanics[1]—but the construction is, of course, as difficult as the solution of the true quantum mechanical problem. It is the purpose of this paper to show that a very similar situation exists for a Dirac electron in an electromagnetic field, based upon the relativistic Hamilton–Jacobi equation. However, a totally arbitrary field cannot now be subsumed by the classical format, with its insistence upon coherent phase for the full wavefunction, and we shall spell out in detail the nature of the required restrictions.

2. BASIC SOLUTION

According to classical relativistic dynamics,[2] the Hamilton–Jacobi equation for a charged particle of charge $-e$ and proper mass m is of the form

$$\xi_5^2 = \xi_\rho \xi_\rho + \xi_4^2, \quad \rho = 1, 2, 3 \text{ (summation convention)},$$

$$\text{where} \quad \xi_5 = \frac{1}{c}\left(\frac{\partial S}{\partial t} - eV\right), \quad \xi_4 = mc, \tag{1}$$

$$\text{and} \quad \xi_\rho = \frac{\partial S}{\partial x_\rho} + \frac{e}{c} A_\rho.$$

The electromagnetic field is given by $(\mathbf{A}, V)$ and $S(x_1, x_2, x_3, t; x^*)$ is the two-event characteristic function; we regard $(x^*) = (x_1^*, x_2^*, x_3^*, t^*)$ as constant.

Let us now "linearize" in precisely the same way that the Klein–Gordon equation gives rise to the Dirac equation, i.e., we introduce the Dirac matrices

$$\alpha_1 = \left(\begin{array}{cc|cc} & & 0 & 1 \\ & \bigcirc & 1 & 0 \\ \hline 0 & 1 & & \\ 1 & 0 & \bigcirc & \end{array}\right), \quad \alpha_2 = \left(\begin{array}{cc|cc} & & 0 & -i \\ & \bigcirc & i & 0 \\ \hline 0 & -i & & \\ i & 0 & \bigcirc & \end{array}\right),$$

$$\alpha_3 = \left(\begin{array}{cc|cc} & & 1 & 0 \\ & \bigcirc & 0 & -1 \\ \hline 1 & 0 & & \\ 0 & -1 & \bigcirc & \end{array}\right), \quad \alpha_4 = \left(\begin{array}{cc|cc} 1 & 0 & & \\ 0 & 1 & \bigcirc & \\ \hline & & -1 & 0 \\ \bigcirc & & 0 & -1 \end{array}\right) \tag{2}$$

as realizations of the quartet of Clifford numbers[3] satisfying the algebraic relations

$$\alpha_\lambda \alpha_\sigma + \alpha_\sigma \alpha_\lambda = \delta_{\sigma\lambda}, \tag{3}$$

where σ, λ will always run over the full set 1, 2, 3, 4. We then consider the set of equations

$$(-\xi_5 I + \alpha_\rho \xi_\rho + \alpha_4 \xi_4)u = 0,$$

$$\text{where} \quad u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \tag{4}$$

and claim that the Hamilton–Jacobi equation (1) is the condition under which the set (4) has a solution. This follows from the fact that $\det(\xi_5 I + \alpha_\sigma \xi_\sigma) = \det(\xi_5 I - \alpha_\sigma \xi_\sigma)$ (all algebraic relations are of even order), coupled with

$$(\xi_5 I + \alpha_\sigma \xi_\sigma)(\xi_5 I - \alpha_\sigma \xi_\sigma) = (\xi_5^2 - \xi_\sigma \xi_\sigma),$$

so that

$$\begin{aligned} \det(\xi_5 I - \alpha_\sigma \xi_\sigma) &= [\det(\xi_5^2 - \xi_\sigma \xi_\sigma) I]^{1/2} \\ &= (\xi_5^2 - \xi_\sigma \xi_\sigma)^2. \end{aligned}$$

What now is the relationship between (4) and the Dirac equation? In fact, for a suitably restricted electromagnetic potential, an exact solution of the Dirac equation (with initial conditions) can be constructed by setting

$$\psi = u e^{(i/\hbar)S}, \tag{5}$$

if the spinor u at the same time can be chosen to satisfy

$$\frac{1}{c}\frac{\partial u}{\partial t} = \alpha_\rho \frac{\partial u}{\partial x_\rho}. \tag{6}$$

To see this, we need only observe that the Dirac equation

$$\left[i\hbar \frac{\partial}{\partial ct} + \frac{eV}{c} + \alpha_\rho \left(\frac{\hbar}{i}\frac{\partial}{\partial x_\rho} + \frac{e}{c} A_\rho\right) + mc\alpha_4\right]\psi = 0 \tag{7}$$

in the form

$$\begin{aligned} e^{(i/\hbar)S}\left[i\hbar \frac{\partial u}{\partial ct} + \alpha_\rho \frac{\hbar}{i}\frac{\partial u}{\partial x_\rho}\right] + e^{(i/\hbar)S}\Bigg[\left(-\frac{1}{c}\frac{\partial S}{\partial t} + \frac{eV}{c}\right) \\ + \alpha_\rho \left(\frac{\partial S}{\partial x_\rho} + \frac{e}{c} A_\rho\right) + mc\alpha_4\Bigg] u = 0, \end{aligned}$$

is valid by virtue of (4) and (6) with the definitions of (1).

The requirement that the set (1, 4, 6) be consistent is that the ξ_σ be such that one of the sizable manifolds of u which satisfies (4) be a spinor satisfying (6) as well. By virtue of (1), one can in principle then eliminate the Hamilton–Jacobi S and rephrase this as a condition on the electromagnetic field $(\mathbf{A}, V)$. The first part of the program is very easily carried out. Equations (6) are of rank two in momentum space and hence are generated by two arbitrary functions. This fact may be expressed in many ways, a convenient one being the parametric form

$$\begin{aligned} &u = \left(\prod_\sigma \alpha_\sigma\right)\left(\frac{1}{c}\frac{\partial}{\partial t} - \alpha_\rho \frac{\partial}{\partial x_\rho}\right) v, \\ &\text{where} \quad \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_\rho}\frac{\partial}{\partial x_\rho}\right) v = 0, \end{aligned} \tag{8}$$

and it suffices to give v two nonvanishing components. Given the u_j, relation (4), written out as

$$\begin{aligned} u_4(\xi_1 - i\xi_2) + u_3\xi_3 &= u_1(\xi_5 - \xi_4), \\ u_2(\xi_1 + i\xi_2) - u_4\xi_3 &= u_2(\xi_5 - \xi_4), \\ u_2(\xi_1 - i\xi_2) + u_1\xi_3 &= u_3(\xi_5 + \xi_4), \\ u_1(\xi_1 - i\xi_2) - u_2\xi_3 &= u_4(\xi_5 + \xi_4), \end{aligned} \tag{9}$$

is of rank three, and hence may be solved for ξ_1, ξ_2, ξ_3:

$$\begin{aligned} \xi_1 - i\xi_2 &= \frac{(u_1^2 - u_3^2)\xi_5 - (u_1^2 + u_3^2)\xi_4}{u_1u_4 - u_2u_3} \\ \xi_1 + i\xi_2 &= \frac{(u_4^2 - u_2^2)\xi_5 + (u_4^2 + u_2^2)\xi_4}{u_1u_4 - u_2u_3} \\ \xi_3 &= \frac{(u_3u_4 - u_1u_2)\xi_5 + (u_3u_4 + u_1u_2)\xi_4}{u_1u_4 - u_2u_3} \end{aligned} \tag{10}$$

The final step of eliminating S from (10) is one which we shall consider piecemeal.

3. RELATION TO WKB APPROACH

Before proceeding with the reduction of (10), which will result in some unusual restrictions, we should point out the relationship between the situation being analyzed and a number of sophisticated and effective works[4] on the semiclassical Dirac equation. In fact, this relationship is most easily recognized in the semiclassical or WKB approximation to the one-dimensional stationary state Schrödinger equation

$$\begin{aligned} &\psi''(x) + [p(x)^2/\hbar^2]\psi(x) = 0, \\ &\text{where} \quad p(x)^2 = [E - V_0(x)]/2m. \end{aligned} \tag{11}$$

In one version of the WKB expansion, one seeks a solution in the perhaps asymptotic form

$$\psi(x) = \left(\sum_0^\infty (i\hbar)^j A_j(x)\right) e^{(i/\hbar)S(x)} \tag{12}$$

with $A_0 \neq 0$. Direct substitution into (11) then

yields, on equating powers of $(i\hbar)^{j-2}$,

$$\begin{aligned} (S')^2 - p^2 &= 0, \\ \left(S'' + 2S'\frac{d}{dx}\right) A_0 &= 0 \text{ or } \quad A_0 \propto (S')^{-1/2}, \\ \left(S'' + 2S'\frac{d}{dx}\right) A_{j+1} &= A_j'', \quad \text{for } j \geq 0, \end{aligned} \tag{13}$$

or $\frac{d}{dx}\left[(S')^{1/2} A_{j+1}\right] = \frac{1}{2}(S')^{-1/2} A_j''$,

solvable in quadratures.

On the other hand, one can assume

$$\psi(x) = A_0(x)\, e^{(i/\hbar)S(x)}, \tag{12'}$$

restricting the appearance of $\hbar$ to the classically unobservable phase factor, and ask for what potential

$$V(x) = \sum_0^2 \hbar^j V_j(x) \tag{14}$$

the postulated solution is valid. Direct substitution now affirms that the expression for A_0 and S in terms of V_0 is unchanged, and that $V_1 = 0$, but that one has an additional quantum mechanical modification

$$V_2(x) = (1/2m)[A_0''(x)/A_0(x)] \tag{15}$$

in order for the simple semiclassical form to hold. We can then judge whether this is to be regarded as a small correction or not.

In the same way, the Dirac equation (4) may be examined from two points of view. The WKB expansion consists of setting

$$\psi = \sum_0 (i\hbar)^j u^{(j)} e^{(i/\hbar)S}, \tag{16}$$

resulting at once in the sequence of relations

$$\begin{aligned} (-\xi_5 I + \alpha_\rho \xi_\rho + \alpha_4 \xi_4) u^{(0)} &= 0 \\ (-\xi_5 I + \alpha_\rho \xi_\rho + \alpha_4 \xi_4) u^{(j+1)} &= \left(\frac{\partial}{\partial ct} - \alpha_\rho \frac{\partial}{\partial x_\rho}\right) u^{(j)}, \end{aligned} \tag{17}$$

where the notation of (1) has been used. These can be solved sequentially when the Hamilton–Jacobi equation for S is established, the two arbitrary functions at each stage being determined by consistency for the following stage. In the alternative viewpoint, we choose

$$\begin{aligned} \psi &= u e^{(i/\hbar)S}, \\ V &= V^{(0)} + \hbar V^{(1)}, \quad A_\rho = A_\rho^{(0)} + \hbar A_\rho^{(1)}, \end{aligned} \tag{18}$$

and find on substitution into (4) that

$$(-\xi_5^{(0)} I + \alpha_\rho \xi_\rho^{(0)} + \alpha_4 \xi_4^{(0)}) u = 0, \tag{19a}$$

$$\left(\frac{\partial}{\partial ct} - \alpha_\rho \frac{\partial}{\partial x_\rho}\right) u = \frac{ie}{c}(V^{(1)} I - A_\rho^{(1)} \alpha_\rho) u. \tag{19b}$$

Since u, from (19a), depends upon two arbitrary functions, $V^{(1)}$ and $A_\rho^{(1)}$ are not uniquely determined, although the choice $u = u^{(0)}$ from (17) is a very natural one. Instead, we are allowed the option of choosing the parameters in u to minimize a desired physical effect of $V^{(1)}$ and $A_\rho^{(1)}$. It is our purpose in this paper to carry this quest to its extreme and ask under what circumstances the quantum corrections $V^{(1)}$ and $A_\rho^{(1)}$ can be set equal to zero. Then indeed (19a), (19b) reduce to (4) and (6).

4. THE ONE-DIMENSIONAL TIME-DEPENDENT PROBLEM

As prototype, albeit a very restricted one, let us consider a Dirac electron in a one-dimensional time-dependent potential $V(z, t)$, where $x_3 = z$. In the notation of (1), then,

$$\begin{aligned} &\xi_1 = \xi_2 = 0, \quad \xi_3 = \frac{\partial S}{\partial z}, \quad \xi_4 = mc, \\ &\xi_5 = \frac{\partial S}{\partial ct} - \frac{e}{c} V, \quad A_\rho = 0. \end{aligned} \tag{20}$$

The first two equations of (10) tell us that

$$\xi_5/mc = (u_1^2 + u_3^2)/(u_1^2 - u_3^2) = (u_2^2 + u_4^2)/(u_2^2 - u_4^2).$$

Thus $(u_1/u_3)^2 = (u_2/u_4)^2$. Since $u_1/u_3 = u_2/u_4$ is inconsistent with (10), we have

$$u_1/u_3 = -u_2/u_4 \equiv U. \tag{21}$$

It then follows from (10) that

$$\frac{\partial S}{\partial z} = \frac{2U}{U^2 - 1} mc, \quad \frac{\partial S}{\partial ct} = \frac{e}{c} V + \frac{U^2 + 1}{U^2 - 1} mc, \tag{22}$$

an integrable system if

$$\frac{e}{mc^2}\frac{\partial V}{\partial z} = \frac{\partial}{\partial ct}\frac{2U}{U^2 - 1} - \frac{\partial}{\partial z}\frac{U^2 + 1}{U^2 - 1}. \tag{23}$$

Now the condition (6) in the present case becomes simply

$$\begin{aligned} \frac{\partial u_1}{\partial ct} &= \frac{\partial u_3}{\partial z}, \quad \frac{\partial u_3}{\partial ct} = \frac{\partial u_1}{\partial z}, \\ \frac{\partial u_2}{\partial ct} &= -\frac{\partial u_4}{\partial z}, \quad \frac{\partial u_4}{\partial ct} = -\frac{\partial u_2}{\partial z}. \end{aligned}$$

Hence

$$\begin{aligned} u_1 &= f(z + ct) - g(z - ct), \\ u_3 &= f(z + ct) + g(z - ct), \end{aligned} \tag{24}$$

for arbitrary f, g, and we may choose, e.g., $u_2 = u_1$, $u_4 = -u_3$. On substituting into (23), we arrive at the parametric representation

$$\begin{aligned} -\frac{e}{mc^2}\frac{\partial V}{\partial z} &= g(z - ct)\frac{f'(z + ct)}{f(z + ct)^2} + f(z + ct)\frac{g'(z - ct)}{g(z - ct)^2}, \end{aligned} \tag{25}$$

of perhaps unexpected generality considering the imposition of the condition $A_\rho = 0$.

5. STATIONARY STATES

A class of problems of much greater interest and utility is that of the stationary states $\psi \propto e^{\epsilon t/i\hbar}$, where ϵ is then the total energy. According to (5), we can then write $S \to S - \epsilon t$, and all remaining quantities will be time independent. It will be helpful to use a somewhat more compact notation, namely,

$$u = \begin{pmatrix} v \\ w \end{pmatrix}, \quad \text{where } v = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad w = \begin{pmatrix} u_3 \\ u_4 \end{pmatrix}. \tag{26}$$

Then Eq. (4) becomes, with ξ the vector of components ξ_1, ξ_2, ξ_3,

$$\begin{aligned} \xi\cdot\sigma w &= (\xi_5 - \xi_4)v, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ & \qquad\qquad \xi_4 = mc, \\ \xi\cdot\sigma v &= (\xi_5 + \xi_4)w, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ & \qquad\qquad \xi_5 = -\frac{1}{c}(\epsilon + eV), \\ & \qquad\qquad \sigma_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{27}$$

and Eq. (6) (for time independence)

$$\sigma\cdot\nabla v = 0 = \sigma\cdot\nabla w. \tag{28}$$

To solve (27), we observe that (with the prime denoting transpose)

$$I = (wv' - vw')\sigma_2/v'\sigma_2 w \tag{29}$$

is the 2×2 identity matrix for any v, w. Hence

$$\begin{aligned} \xi\cdot\sigma &= \xi\cdot\sigma I \\ &= [(\xi_5 - \xi_4)vv' - (\xi_5 + \xi_4)ww']\sigma_2/v'\sigma_2 w, \end{aligned}$$

and from the algebraic identity

$$\xi = \tfrac{1}{2}\,\mathrm{Tr}\sigma(\xi\cdot\sigma) \tag{30}$$

it follows that

$$\xi = \frac{\xi_5 - \xi_4}{2v'\sigma_2 w}\, v'\sigma_2\sigma v - \frac{\xi_5 + \xi_4}{2v'\sigma_2 w}\, w'\sigma_2\sigma w, \tag{31}$$

which can be identified with (10).

It will be a bit more convenient to introduce the variables $\bar{v} = v + w$, $\bar{w} = v - w$. Then (28) and (31) read explicitly

$$\begin{aligned} &\nabla S + \frac{e}{c}A \\ &\quad = \frac{\epsilon + eV}{c}\,\frac{\bar{w}'\sigma_2\sigma\bar{v}}{\bar{w}'\sigma_2\bar{v}} + \frac{mc}{2}\,\frac{\bar{w}'\sigma_2\sigma\bar{w} + \bar{v}'\sigma_2\sigma\bar{v}}{\bar{w}'\sigma_2\bar{v}}, \\ &\quad \text{where} \quad \sigma\cdot\nabla\bar{v} = \sigma\cdot\nabla\bar{w} = 0 \end{aligned} \tag{32}$$

Regarded as a set of three equations for S, (32) is integrable if its curl is an identity. Since $\mathbf{E} = -\nabla V, \mathbf{B} = \nabla \times \mathbf{A}$, we thus have

$$\begin{aligned} \mathbf{B} = \frac{\bar{w}'\sigma_2\sigma\bar{v}}{\bar{w}'\sigma_2\bar{v}} \times \mathbf{E} + \left(V + \frac{\epsilon}{e}\right)\nabla \times \frac{\bar{w}'\sigma_2\sigma\bar{v}}{\bar{w}'\sigma_2\bar{v}} \\ + \frac{mc^2}{2e}\,\nabla \times \frac{\bar{w}'\sigma_2\sigma\bar{w} + \bar{v}'\sigma_2\sigma\bar{v}}{\bar{w}'\sigma_2\bar{v}}. \end{aligned} \tag{33}$$

In other words, given the potential V, we have available a wide class of magnetic fields depending parametrically upon $\bar{v}$ and $\bar{w}$, and hence upon two independent solutions of the Laplace (time-independent wave) equation. For each such combination of V and B, the semiclassical form (5) provides an exact solution to the one-particle Dirac equation.

The relation (33) is hardly transparent, nor is it transparently real. The latter can be rectified by setting (and is not quite equivalent to) $\bar{v} = a$, $\bar{w} = \sigma_2 a^*$, where $\sigma\cdot\nabla a = 0$. It follows that $\sigma\cdot\nabla\sigma_2 a^* = 0$, and the result

$$\begin{aligned} \mathbf{B} = \frac{a^+\sigma a}{a^+ a} \times \mathbf{E} + \left(V + \frac{\epsilon}{e}\right)\nabla \times \frac{a^+\sigma a}{a^+ a} \\ + \frac{mc^2}{2e}\,\nabla \times \frac{a^+\sigma\sigma_2 a^* - a'\sigma_2\sigma a}{a^+ a} \end{aligned} \tag{34}$$

is certainly real. Let us consider further subcases. The simplest solution of $\sigma\cdot\nabla a = 0$ is $a =$ const. Then (34) reduces to

$$\hat{\mathbf{B}} = \hat{\mathbf{n}} \times \mathbf{E}, \tag{35}$$

where $\hat{\mathbf{n}}$ is an arbitrary unit vector. This problem is quite degenerate since both the invariants $B^2 - E^2$ and $\mathbf{B}\cdot\mathbf{E}$ vanish. A more structured case involves choosing $a = f(x - iy)\binom{1}{0}$. Then

$$\begin{aligned} \mathbf{B} = \mathbf{E} \times \hat{\mathbf{z}} + \frac{mc^2}{2e}\,\nabla \\ \times \left((\hat{\mathbf{y}} + i\hat{\mathbf{x}})\frac{f^*(x + iy)}{f(x - iy)} + (\hat{\mathbf{y}} - i\hat{\mathbf{x}})\frac{f(x - iy)}{f^*(x + iy)}\right). \end{aligned} \tag{36}$$

In general, we see from (33) that $\mathbf{B} = \hat{\mathbf{n}} \times \mathbf{E} + \delta\mathbf{B}$, but $\hat{\mathbf{n}}$ can be a space-dependent unit vector and $\delta\mathbf{B}$ a function with considerable parametric dependence.

6. FULL SPACE–TIME DEPENDENCE

If no special assumption is made as to the nature of the time dependence, the solution (34) remains valid, with (28) replaced by $\partial v/\partial ct = \sigma\cdot\nabla w$, $\partial w/\partial ct = \sigma\cdot\nabla v$. The integrability conditions for the three first-order partial differential equations in four independent variables can be found by the standard Jacobi elimination[5] and are now quite involved in form. We shall therefore proceed instead by setting up a somewhat less general parametric form for the solution to (4). It is based upon the well-known[1] pairs of independent solutions

$$(u_1u_2u_3u_4) = \begin{cases} (w_0 \quad 0 \quad w_3 \quad w_1 + iw_2) \\ (0 \quad w_0 \quad w_1 - iw_2 \quad -w_3) \end{cases}$$

$$w_\rho \equiv w_0\xi_\rho/(\xi_4 + \xi_5), \tag{37a}$$

$$(u_1u_2u_3u_4) = \begin{cases} (-w_3 \quad -w_1 - iw_2 \quad -w_0 \quad 0) \\ (-w_1 + iw_2 \quad w_3 \quad 0 \quad -w_0) \end{cases}$$

$$w_\rho \equiv -w_0\xi_\rho/(\xi_4 - \xi_5). \tag{37b}$$

By using $\xi_4 = mc$ and the Hamilton–Jacobi $\xi_5^2 - \xi_4^2 = \xi_\rho\xi_\rho$, the definitions of w_ρ in (37) are readily inverted to yield

$$\begin{aligned} \pm\,\xi_\rho &= 2mc\; w_0w_\rho/(w_0w_0 - w_\lambda w_\lambda), \\ \pm\,\xi_5 &= 2mc\,[w_0w_0/(w_0w_0 - w_\lambda w_\lambda)] - mc \end{aligned} \tag{38}$$

with $\pm$ for cases a and b, respectively. On the other hand, the w's, in addition to satisfying the wave equation, are intimately connected through (6). We find

$$w_\rho = \frac{\partial\Omega}{\partial x_\rho}, \qquad w_0 = \frac{\partial\Omega}{\partial ct}, \tag{39}$$

for some solution Ω of the wave equation.

The standard[6] Lorentz metric $(1 - 1 - 1 - 1)$ with $x^0 \equiv ct$ and $\xi_5 \equiv \xi_0$ offers a still more compact notation for (38) and (39):

$$\pm\,\xi_\mu = \frac{2mc}{w_\nu w^\nu}\, w_0w_\mu + mc\delta_{\mu 0}, \qquad w_\mu = \frac{\partial\Omega}{\partial x^\mu}, \tag{40}$$

where $\dfrac{\partial^2\Omega}{\partial x_\nu \partial x^\nu} = 0.$

Now in the same notation

$$\xi_\mu = \frac{\partial S}{\partial x^\mu} + \frac{e}{c}\, A_\mu, \tag{41}$$

where $A_0 = -V$. The integrability conditions for the existence of a consistent solution S now consist simply of the curl of (41), leading at once to

$$\begin{aligned} F_{\mu\mu'} &= \pm\,\frac{2mc^2}{e}\left(\frac{\partial}{\partial x^{\mu'}}\,\frac{w_0w_\mu}{w_\nu w^\nu} - \frac{\partial}{\partial x^\mu}\,\frac{w_0w_{\mu'}}{w_\nu w^\nu}\right), \\ &= \pm\,\frac{2mc^2}{e}\left(w_\mu\,\frac{\partial}{\partial x^{\mu'}} - w_{\mu'}\,\frac{\partial}{\partial x^\mu}\right)\frac{w_0}{w_\nu w^\nu}, \end{aligned} \tag{42}$$

where $F_{\mu\nu} \equiv \partial A_\mu/\partial x^\nu - \partial A_\nu/\partial x^\mu$ is the electromagnetic field. For this explicit class of fields, then, the Dirac equation is solvable in semiclassical form. Of course, since only observable fields enter, the result is obviously gauge invariant.

---

* Supported in part by the U.S. Atomic Energy Commission, Contract AT(30-1)-1480.

† Present address: Concentration of Environmental Control, University of Wisconsin-Green Bay, Green Bay, Wisconsin 54302.

1 See, e.g., L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955), Chap. VII.

2 See, e.g., J. L. Synge, "Classical Dynamics," in *Handbuch der Physik, Vol. III/1* (Springer, Berlin, 1960).

3 See, e.g., H. Boerner, *Group Representations* (North-Holland, Amsterdam, 1963).

4 Bargmann, Michel, and Telegdi, Phys. Rev. Letters **2**, 435 (1959); R. Schiller, Phys. Rev. **125**, 1100 (1962); S. I. Rubinow and J. B. Keller, *ibid.* **131**, 2789 (1963).

5 See, e.g., Goursat, Hedrick, and Dunkel, *Differential Equations* (Ginn, Boston, 1971), Sec. 91.

6 See, e.g., P. G. Bergmann, *Theory of Relativity* (Prentice-Hall, New York, 1946), Chap. V.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Path Integrals in Curved Spaces*

David W. McLaughlin
New York University, Courant Institute of Mathematical Sciences
and
Mathematics Department, University Heights Campus, New York University, New York, New York
and
L. S. Schulman†
Physics Department, Indiana University, Bloomington, Indiana

(Received 9 June 1971)

In this paper we present a simplification of the path integral solution of the Schrödinger equation in terms of coordinates which need not be Cartesian. After presenting the existing formula, we discuss the relationship between the distance and time differentials. Making this relationship precise through the technique of stationary phase, we are able to simplify the path integral. The resulting expression can be used to obtain a Hamiltonian path integral. Finally, we comment on a similar phenomenon involving differentials in the Itô integral.

It is known that the path integral for a particle described by non-Cartesian coordinates is complicated by differentials of a kind neglected in most other types of integration. As a specific example, the path integral for a free particle moving in a plane, described by polar coordinates, involves the action integral S evaluated along the classical path connecting $\mathbf{r}_1$ and $\mathbf{r}_2$ in a time ϵ:

$$(u_1u_2u_3u_4) = \begin{cases} (w_0 \quad 0 \quad w_3 \quad w_1 + iw_2) \\ (0 \quad w_0 \quad w_1 - iw_2 \quad -w_3) \end{cases}$$

$$w_\rho \equiv w_0\xi_\rho/(\xi_4 + \xi_5), \tag{37a}$$

$$(u_1u_2u_3u_4) = \begin{cases} (-w_3 \quad -w_1 - iw_2 \quad -w_0 \quad 0) \\ (-w_1 + iw_2 \quad w_3 \quad 0 \quad -w_0) \end{cases}$$

$$w_\rho \equiv -w_0\xi_\rho/(\xi_4 - \xi_5). \tag{37b}$$

By using $\xi_4 = mc$ and the Hamilton–Jacobi $\xi_5^2 - \xi_4^2 = \xi_\rho\xi_\rho$, the definitions of w_ρ in (37) are readily inverted to yield

$$\begin{aligned} \pm\,\xi_\rho &= 2mc\; w_0w_\rho/(w_0w_0 - w_\lambda w_\lambda), \\ \pm\,\xi_5 &= 2mc\,[w_0w_0/(w_0w_0 - w_\lambda w_\lambda)] - mc \end{aligned} \tag{38}$$

with $\pm$ for cases a and b, respectively. On the other hand, the w's, in addition to satisfying the wave equation, are intimately connected through (6). We find

$$w_\rho = \frac{\partial\Omega}{\partial x_\rho}, \qquad w_0 = \frac{\partial\Omega}{\partial ct}, \tag{39}$$

for some solution Ω of the wave equation.

The standard[6] Lorentz metric $(1 - 1 - 1 - 1)$ with $x^0 \equiv ct$ and $\xi_5 \equiv \xi_0$ offers a still more compact notation for (38) and (39):

$$\pm\,\xi_\mu = \frac{2mc}{w_\nu w^\nu}\, w_0w_\mu + mc\delta_{\mu 0}, \qquad w_\mu = \frac{\partial\Omega}{\partial x^\mu}, \tag{40}$$

where $\dfrac{\partial^2\Omega}{\partial x_\nu \partial x^\nu} = 0.$

Now in the same notation

$$\xi_\mu = \frac{\partial S}{\partial x^\mu} + \frac{e}{c}\, A_\mu, \tag{41}$$

where $A_0 = -V$. The integrability conditions for the existence of a consistent solution S now consist simply of the curl of (41), leading at once to

$$\begin{aligned} F_{\mu\mu'} &= \pm\,\frac{2mc^2}{e}\left(\frac{\partial}{\partial x^{\mu'}}\,\frac{w_0w_\mu}{w_\nu w^\nu} - \frac{\partial}{\partial x^\mu}\,\frac{w_0w_{\mu'}}{w_\nu w^\nu}\right), \\ &= \pm\,\frac{2mc^2}{e}\left(w_\mu\,\frac{\partial}{\partial x^{\mu'}} - w_{\mu'}\,\frac{\partial}{\partial x^\mu}\right)\frac{w_0}{w_\nu w^\nu}, \end{aligned} \tag{42}$$

where $F_{\mu\nu} \equiv \partial A_\mu/\partial x^\nu - \partial A_\nu/\partial x^\mu$ is the electromagnetic field. For this explicit class of fields, then, the Dirac equation is solvable in semiclassical form. Of course, since only observable fields enter, the result is obviously gauge invariant.

---

* Supported in part by the U.S. Atomic Energy Commission, Contract AT(30-1)-1480.

† Present address: Concentration of Environmental Control, University of Wisconsin-Green Bay, Green Bay, Wisconsin 54302.

1 See, e.g., L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955), Chap. VII.

2 See, e.g., J. L. Synge, "Classical Dynamics," in *Handbuch der Physik, Vol. III/1* (Springer, Berlin, 1960).

3 See, e.g., H. Boerner, *Group Representations* (North-Holland, Amsterdam, 1963).

4 Bargmann, Michel, and Telegdi, Phys. Rev. Letters **2**, 435 (1959); R. Schiller, Phys. Rev. **125**, 1100 (1962); S. I. Rubinow and J. B. Keller, *ibid.* **131**, 2789 (1963).

5 See, e.g., Goursat, Hedrick, and Dunkel, *Differential Equations* (Ginn, Boston, 1971), Sec. 91.

6 See, e.g., P. G. Bergmann, *Theory of Relativity* (Prentice-Hall, New York, 1946), Chap. V.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

Path Integrals in Curved Spaces*

David W. McLaughlin
New York University, Courant Institute of Mathematical Sciences
and
Mathematics Department, University Heights Campus, New York University, New York, New York
and
L. S. Schulman†
Physics Department, Indiana University, Bloomington, Indiana

(Received 9 June 1971)

In this paper we present a simplification of the path integral solution of the Schrödinger equation in terms of coordinates which need not be Cartesian. After presenting the existing formula, we discuss the relationship between the distance and time differentials. Making this relationship precise through the technique of stationary phase, we are able to simplify the path integral. The resulting expression can be used to obtain a Hamiltonian path integral. Finally, we comment on a similar phenomenon involving differentials in the Itô integral.

It is known that the path integral for a particle described by non-Cartesian coordinates is complicated by differentials of a kind neglected in most other types of integration. As a specific example, the path integral for a free particle moving in a plane, described by polar coordinates, involves the action integral S evaluated along the classical path connecting $\mathbf{r}_1$ and $\mathbf{r}_2$ in a time ϵ:

$$S(\mathbf{r}_2, \epsilon\,|\mathbf{r}_1, 0) \equiv \frac{m}{2\epsilon}\,|\mathbf{r}_1 - \mathbf{r}_2\,|^2$$
$$\simeq \frac{m}{2\epsilon}\,[(r_1 - r_2)^2 + r_1 r_2(\phi_2 - \phi_1)^2$$
$$- \tfrac{1}{12}\,r_1 r_2(\phi_2 - \phi_1)^4]. \qquad (1)$$

This expression enters the Green's function for infinitesimal times, which is ultimately iterated to give the finite time propagator. As shown by Edwards and Gulyaev,[1] it is necessary to retain terms through the fourth order in Eq. (1), for if it or any lower-order term is neglected, the Green's function obtained by iteration is wrong! That is, it does not propagate solutions of the Schrödinger equation.

The requirement for this degree of accuracy in the expansion of the action is quite general. In 1957, DeWitt[2] formally established the general expansion of the Green's function for curved spaces. He begins with the classical system characterized by a Lagrangian of the general form

$$L = \tfrac{1}{2}\,g_{ij}\dot{q}^i\dot{q}^j + a_i\dot{q}^i - v, \qquad (2)$$

where g_{ij} is the metric tensor, dots represent derivatives with respect to the time t, a and v are functions of q and t, and the summation convention is adopted. The action function $S(q'', t''\,|q', t')$ may be expanded about the point (q', t') and put in the form[3]

$$S(q'', t''\,|q', t') = \frac{1}{\Delta t}\,\{\tfrac{1}{2}g_{ij}\overset{i}{\Delta}\overset{j}{\Delta} + \tfrac{1}{4}g_{ij,k}\overset{i}{\Delta}\overset{j}{\Delta}\overset{k}{\Delta}$$
$$+ \tfrac{1}{12}(g_{ij,kl} - \tfrac{1}{2}g^{mn}\,[ij, m][kl, n])\overset{i}{\Delta}\overset{j}{\Delta}\overset{k}{\Delta}\overset{l}{\Delta} + O(\overset{5}{\Delta})\}$$
$$+ a_i\overset{i}{\Delta} + \tfrac{1}{4}(2a_{i,j} + \frac{\partial}{\partial t}\,g_{ij})\overset{i}{\Delta}\overset{j}{\Delta} - v\Delta t + O(\overset{3}{\Delta})$$
$$+ O(\Delta(\Delta t)) + O((\Delta t)^2), \qquad (3)$$

where all functions are evaluated at (q', t'), $\Delta t \equiv (t'' - t')$, $\overset{i}{\Delta} \equiv (q''^i - q'^i)$, subscripts following a comma denote partial derivatives with respect to the component indicated, g^{mn} denotes the inverse of the metric tensor, and $[jk, i] \equiv \tfrac{1}{2}(g_{ij,k} + g_{ik,j} - g_{kj,i})$. In Eq. (3) we have simplified DeWitt's summations.

Extending the earlier work of Van Vleck,[4] Morette,[5] and Pauli;[6] DeWitt showed that, when exponentiated and iterated, the expansion (3) is sufficiently accurate to yield the propagator K of the Schrödinger equation. Explicitly he obtained

$$K(x'', t''\,|x', t') = \lim_{N\to\infty} \int d^n q_1 \cdots \int d^n q_{N-1}$$
$$\times \left(\prod_{j=1}^{N-1}\,[g(q_j, t_j)]^{1/2}\right)\left(\frac{1}{2\pi i\hbar\Delta t}\right)^{nN/2}$$
$$\times \exp\left[\frac{i}{\hbar}\left(\sum_{j=1}^{N}\tilde{S}(q_j, t_j\,|q_{j-1}, t_{j-1}) + \frac{\hbar^2}{12}\,R\,\Delta t\right)\right], \qquad (4)$$

where g is the determinant of the metric tensor, R is the curvature of the space, $\Delta t \equiv (t'' - t')/N$, $t_j = t' + j\Delta t$, $\tilde{S}$ is the truncated expansion (3), and $q_0 \equiv x'$, $q_N \equiv x''$.

However, for some purposes this may not be the most convenient form in which to express the path integral. Terms containing $\Delta^3/\Delta t$ and $\Delta^4/\Delta t$ are particularly unpleasant. One would prefer to eliminate such terms in favor of an approximate action of the form $S = \frac{1}{2}g_{ij}\Delta^i\Delta^j - u\Delta t$, for then a resemblance to the usual definition of the action integral is maintained. Merely writing, for example,

$$\Delta^3/\Delta t = (\Delta^3/(\Delta t)^3)\,(\Delta t)^2 \qquad (5)$$

formally accomplishes this replacement, but could give rise to the impression that terms of second order in (Δt) must be retained.[7] This impression is false. The basic fact of life about path integrals is that only terms of $O(\Delta t)$ need be retained. In fact, one may interpret the path integral as an iteration of an infinitesimal Green's function

$$e^{-iHt} = (e^{-iHt/N})^N = \lim_{N\to\infty}(1 - iHt/N)^N. \qquad (6)$$

This is based upon the formal extension of the elementary formula

$$e^x = \lim_{N\to\infty}(1 + x/N)^N. \qquad (7)$$

If x is replaced by $x + y_N$ where y_N is $O(1/N)$, the same limit e^x is obtained. It is this fact which forms the basis for neglecting terms of $O((\Delta t)^2)$. (The fact that H is an unbounded operator causes technical difficulties in the application of these formal statements; nevertheless, these difficulties have been overcome.[8])

Thus, rather than ask which powers of Δt explicitly must appear in the integrand, the appropriate question is which terms in this integrand yield terms of $O(\Delta t)$. The answer has been known for some time in this and other contexts.

Here, just as in Brownian motion, the operation of N-fold integration yields the relationship $\Delta^2 = O(\Delta t)$; that is, the integral transforms terms of $O(\Delta^2)$ into ones of $O(\Delta t)$. In Cartesian coordinates this fact may be easily seen. For large N(small Δt), the dominant term in the integrand is

$$\left(\frac{m}{2\pi i\hbar\Delta t}\right)^{Nn/2}\exp\left[\frac{im}{2\hbar\Delta t}\left(\sum_{j=1}^{N}|\mathbf{r}_j - \mathbf{r}_{j-1}|^2\right)\right]. \qquad (8)$$

Considering this dominant term a part of an N-fold measure, one immediately establishes that terms in the remaining integrand of $O(\Delta^2)$ are transformed into ones linear in Δt. Perhaps the clearest way to see this is to follow Feynman's original argument.[9] For example, consider

$$I \equiv \lim_{N\to\infty}\left(\frac{mN}{2\pi i\hbar t}\right)^{Nn/2}\int d^3r_1 \cdots \int d^3r_N$$
$$\times \exp\left[\frac{imN}{2\hbar t}\left(\sum_{j=1}^{N}|\mathbf{r}_j - \mathbf{r}_{j-1}|^2\right)\right]\frac{(x_N - x_{N-1})^k}{(t/N)},$$
$$k = 2, 4, 6, \qquad (9)$$

where $(x_N - x_{N-1})$ is the "x" component of $(\mathbf{r}_N - \mathbf{r}_{N-1})$. Evaluating the Gaussian integrals in Cartesian coordinates, one obtains

$$I = \lim_{N\to\infty}\left[\frac{N}{t}\,(k-1)!!\left(\frac{i\hbar t}{mN}\right)^{k/2},\right] \quad k = 2, 4, 6, \ldots. \tag{10}$$

Clearly all vanish except $k = 2$. This term equals $i\hbar/m$. This calculation shows that $(x_N - x_{N-1})^2$ acts as if it is $O(t/N) = O(\Delta t)$, when in the integrand of the path integral.[10]

The primary purpose of this paper is to give precise meaning to these order of magnitude relationships in non-Cartesian coordinates, and, in particular, to establish that sums of the form $\sum_i c_i\,\Delta^4/\Delta t$ can be replaced by sums of the form $\sum_i b_i \Delta t$, where the relationship between c_i and b_i is explicitly given. In this manner, higher order terms in the expansion of the action $\bar{S}$ are replaced by an additional potential, and the path integral is considerably simplified. We remark that this replacement is accomplished through the appropriate use of certain aspects of curved space path integration previously considered as complicating features of the non-Cartesian path integral. In the last section we discuss the relationship of this work with the Itô integral in the theory of Brownian motion.

Before presenting the simplification, we mention that Arthurs[7] has obtained a similar result for the specific example of the free particle in polar coordinates. His technique differs from ours. By quantizing a Hamiltonian description of classical mechanics, he replaces the fourth-order term with one linear in Δt.

Derivation of the Effective Potential

Note that all functions appearing in (3) are evaluated at (q', t'), the "left-hand end point". Evaluating these instead at the "midpoint" $[(q' + q'')/2, (t' + t'')/2]$, we obtain

$$S(q'', t''\,|\,q', t') = \frac{1}{\Delta t}\{\tfrac{1}{2} g_{ij}\,\Delta^i\Delta^j + \tfrac{1}{48}\,(g_{ij,\,kl} - 2g^{mn}[ij, m][kl, n])\Delta^i\Delta^j\Delta^k\Delta^l\} + a_i\Delta^i - v\Delta t + O\left(\frac{\Delta^5}{\Delta t}\right) + O(\Delta^3) + O(\Delta\cdot(\Delta t)) + O((\Delta t)^2). \tag{11}$$

Notice that terms of the form $\Delta^3/\Delta t$ and Δ^2 do not appear in (11).

Lemma 1: The following are identities:

(a) $$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} dx^1 \ldots dx^n\, \exp\left(\frac{i}{2\hbar\Delta t}\right) g_{ij}x^i x^j = (2\pi i\hbar\Delta t)^{n/2}\, g^{-1/2},$$

(b) $$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} dx^1\cdots dx^n \times\left[\exp\left(\frac{i}{2\hbar\Delta t}\, g_{ij}x^i x^j\right) x^\alpha x^\beta x^\gamma x^\delta\right]$$
$$= -\hbar^2(\Delta t)^2(2\pi i\hbar\Delta t)^{n/2}\, g^{-1/2}[g^{\alpha\beta}g^{\gamma\delta} + g^{\alpha\gamma}g^{\beta\delta} + g^{\alpha\delta}g^{\beta\gamma}],$$

where $\alpha, \beta, \gamma, \delta = 1, 2, \ldots, n$ and where the g_{ij} are fixed constants. These identities are proven by diagonalizing the matrix $[g_{ij}]$ and evaluating the resulting Gaussian integrals.[10]

Lemma 2: Let $\alpha, \beta, \gamma, \delta$ be fixed constants between 1 and n, $1 \leq \alpha, \beta, \gamma, \delta, \leq n$. Define $I^{\alpha\beta\gamma\delta}(\Delta t)$ by

$$I^{\alpha\beta\gamma\delta}(\Delta t) \equiv (2\pi i\hbar\Delta t)^{-n/2}\int d^n q'(g(q', t''))^{1/2} \times \exp\left(\frac{i}{2\hbar\Delta t}\, g_{ij}(q'', t'')\Delta^i\Delta^j\right)\frac{\Delta^\alpha\Delta^\beta\Delta^\gamma\Delta^\delta}{\Delta t}.$$

Then

$$\lim_{\Delta t\to 0}\frac{I^{\alpha\beta\gamma\delta}(\Delta t)}{\Delta t} = (-1)^n(-\hbar^2)[g^{\alpha\beta}g^{\gamma\delta} + g^{\alpha\delta}g^{\beta\gamma} + g^{\alpha\gamma}g^{\beta\delta}],$$

where the $g^{\alpha\beta}$ are evaluated (q'', t'').

To establish Lemma 2, one translates the coordinate system so that the point $(q_1'', \ldots, q_n'')$ lies at the origin, and, assuming g to be sufficiently smooth, expands $[g(q', t')]^{1/2}$ about (q'', t''). One then applies Lemma 1 to the leading term in $I^{\alpha\beta\gamma\delta}$. The remaining terms, being at least $O((\Delta t)^{3/2})$, vanish in the limit. Definitions:

$$\text{(a)}\quad v' \equiv v - a_i(\Delta^i/\Delta t) - \tfrac{1}{12}\hbar^2 R$$
$$\text{(b)}\quad F_{ijkl} \equiv \tfrac{1}{48}(g_{ij,\,kl} - 2g^{mn}[ij, m][kl, n]) \tag{12}$$
$$\text{(c)}\quad F \equiv -\hbar^2 F_{ijkl}(g^{ij}g^{kl} + g^{ik}g^{jl} + g^{il}g^{jk}).$$

Theorem: Define J by

$$J \equiv (2\pi i\hbar\Delta t)^{-n/2}\int d^n q'\,[g(q', t')]^{1/2} \times \exp\left[\left(\frac{i}{2\hbar\Delta t}\, g_{ij}\Delta^i\Delta^j - \frac{i}{\hbar}\, v'\Delta t\right) \times \exp\left(\frac{i}{\hbar\Delta t}\, F_{ijkl}\Delta^i\Delta^j\Delta^k\Delta^l\right)\right]$$

and I by

$$I \equiv (2\pi i\hbar\Delta t)^{-n/2}\int d^n q'\,[g(q', t')]^{1/2} \times\left[\exp\left(\frac{i}{2\hbar\Delta t}\, g_{ij}\Delta^i\Delta^j - \frac{i}{\hbar}\, v'\Delta t\right)\exp\left(\frac{i}{\hbar}F\,\Delta t\right)\right],$$

where all functions are evaluated at the midpoint unless explicitly stated otherwise. Then $\lim[(J - I)/\Delta t] = 0$ as $\Delta t \to 0$.

Proof: Consider $(J - I)$:

$$(J - I) = (2\pi i\hbar\Delta t)^{-n/2}\int d^n q'\,[g(q', t')]^{1/2} \times \exp\left[\frac{i}{2\hbar\Delta t}\, g_{ij}\left(\frac{q' + q''}{2}, \frac{t' + t''}{2}\right) - \frac{i}{\hbar}\, v'\Delta t\right] \times\left[\exp\left(\frac{i}{\hbar}\, F_{ijkl}\,\frac{\Delta^i\Delta^j\Delta^k\Delta^l}{\Delta t}\right) - \exp\left(\frac{i}{\hbar}\, F\,\Delta t\right)\right].$$

Expanding both g_{ij} and the term in brackets about (q'', t''), we obtain

$$(J-I) = (2\pi i\hbar\Delta t)^{-n/2}\int d^n q'[g(q',t')]^{1/2} \times \exp\left[\frac{i}{\hbar}\left(\frac{1}{2\Delta t}g_{ij}(q'',t'')\Delta^i\Delta^j - v'\Delta t\right)\right] \times \left(\frac{i}{\hbar}F_{ijkl}\frac{\Delta^i\Delta^j\Delta^k\Delta^l}{\Delta t} - \frac{i}{\hbar}F\Delta t\right) + O((\Delta t)^{3/2}).$$

The result follows immediately from Lemma 2.

This theorem indicates that J and I agree to first order in Δt. Thus, we may replace J by I in the integrand of the path integral. Doing so, we obtain

$$K(q'',t''\,|\,q',t') = \lim_{N\to\infty}\int d^n q_1\cdots\int d^n q_{N-1} \times\left(\prod_{j=1}^{N-1}[g(q_j,t_j)]^{1/2}\right)\left(\frac{1}{2\pi i\hbar\Delta t}\right)^{nN/2} \times\exp\left[\frac{i}{\hbar}\left(\sum_{k=0}^{N-1}\frac{g_{ij}(k)}{2\Delta t}(q^i_{k+1}-q^i_k) \times(q^j_{k+1}-q^j_k) - u(k)\Delta t\right)\right], \tag{13}$$

where

$$g_{ij}(k) \equiv g_{ij}\left(\frac{q_{k+1}+q_k}{2},\frac{t_{k+1}+t_k}{2}\right),$$

$$u(k) \equiv v'\left(\frac{q_{k+1}+q_k}{2},\frac{t_{k+1}+t_k}{2}\right) - F\left(\frac{q_{k+1}+q_k}{2},\frac{t_{k+1}+t_k}{2}\right),$$

and where $q_0 = q'$, $q_N = q''$. Equation (13) is the main result of this paper. Notice that the integrand contains only a "kinetic energy term," $\frac{1}{2}g_{ij}(\Delta^i\Delta^j/\Delta t)$, and u, the "effective potential."

Finally, Eq. (13) may be converted into a path integral based upon a Hamiltonian rather than a Lagrangian description of classical mechanics. We restrict ourselves to a flat space upon which a curvilinear coordinate system has been selected.[11] By applying the integral identity

$$\left(\frac{m}{2\pi i\hbar t}\right)^{n/2}\exp\left(\frac{im}{2\hbar t}\,|\mathbf{q}'-\mathbf{q}''|^2\right) = \frac{1}{(2\pi\hbar)^n}\int d^n p\,\exp\left[\frac{i}{\hbar}\left(\mathbf{p}\cdot(\mathbf{q}'-\mathbf{q}'') - \frac{t|p|^2}{2m}\right)\right], \tag{14}$$

N times, one may establish that (13) is equivalent to

$$K(\mathbf{q}'',t''\,|\,\mathbf{q}',t') = \lim_{N\to\infty}\left(\frac{1}{2\pi\hbar}\right)^{Nn}\int d^n q_1\int d^n q_{N-1}\int d^n p_1\cdots \times\int d^n p_N\exp\left[\frac{i}{\hbar}R_N\right], \tag{15}$$

where

$$R_N \equiv \sum_{j=1}^{N}\mathbf{p}_j\cdot(\mathbf{q}_j-\mathbf{q}_{j-1}) - H_j\,\Delta t$$

and

$$H_j \equiv (|\mathbf{p}_j|^2/2m) + u(k).$$

As before, all functions are evaluated at the midpoints. Equation (15) is of Hamiltonian form with an "effective" potential u.

Relation to the "Itô Integral"

As mentioned earlier, the significance of "high-order distance differentials" is evident in other contexts, in particular in the parallel study of Brownian motion. The clearest manifestation is in the Itô integral[12] which is analogous to the action integral appearing in the integrand of the path integral.[1] Itô wishes to define an integral,

$$\int_{t=0}^{t=1} f(x(t,\omega))\,dx(t,\omega), \tag{16}$$

where $x(t,\omega)$ is, as a function of time t, the position of a particle undergoing Brownian motion. Here ω is a point in the sample space Ω which indexes the various Brownian paths. This integral cannot be defined in the sense of Stieltjes since $x(t,\omega)$ is not of bounded variation. Nevertheless, Itô, by breaking the interval $[0,1]$ into N subintervals, defines the integral as

$$\int_0^1 f(x(t,\omega))\,dx(t,\omega) \equiv \underset{N\to\infty}{\text{l.i.m.}}\sum_{k=1}^{N} f\left(x\left(\frac{k-1}{N},\omega\right)\right) \times\left[x\left(\frac{k}{n},\omega\right) - x\left(\frac{k-1}{N},\omega\right)\right]. \tag{17}$$

Here l.i.m. is limit in the mean and f is assumed to be suitably restricted.[13] Notice that the function f is evaluated at the "left-hand end point" of each subinterval. A different value for the integral is obtained if f is evaluated at other points in the subintervals.

With this definition Itô proves the following "fundamental theorem of the Itô calculus":

Theorem (Itô): Let $\phi(u)$ have continuous second derivative. Then

$$\int_0^1\phi'(x(t,\omega))dx(t,\omega) = \phi(x(1,\omega)) - \phi(x(0,\omega)) - \tfrac{1}{2}\int_0^1\phi''(x(t,\omega))dt \tag{18}$$

almost everywhere in Ω. (This formula presumes the diffusion coefficient to be $\frac{1}{2}$.)

The last term in Eq. (18) arises precisely because sums of the form

$$\sum_{k=1}^{N}\phi''\left(x\left(\frac{k-1}{N},\omega\right)\right)\left[x\left(\frac{k}{N},\omega\right) - x\left(\frac{k-1}{N},\omega\right)\right]^2 \tag{19}$$

do not vanish in the limit, and, because of the stochastic nature of the paths, may be replaced by sums of the form

$$\sum_{k=1}^{N}\phi''\left(x\left(\frac{k-1}{N},\omega\right)\right)\Delta t, \qquad \Delta t \equiv \frac{1}{N}. \tag{20}$$

If the right-hand end point were selected in the definition, the opposite sign for the ϕ'' integral would result. If some average were selected, the integral would drop out altogether. In fact, in the "Stratonovich integral,"[13,14] the midpoint is selected.

In addition to the latter simplication, Stratonovich[14] noticed another advantage of the midpoint selection. Essentially, he uses the "Stratonovich integral" to define a diffusion process $\mathbf{x}(t)$ as a stochastic transformation of a Wiener process. The Kolmogorov equation for the probability density associated with this diffusion process $\mathbf{x}(t)$ appears in an invariant form with respect to an arbitrary change of variable $\mathbf{x} \Rightarrow \tilde{\mathbf{x}}(x)$. This invariance is not obtained if the Itô integral is used to define the process $\mathbf{x}(t)$. In quantum mechanics also, transformation properties dictate the choice of summation procedure. If one performs a gauge transformation, the new action involves a vector potential $A = \nabla\phi$, and the path integral for the propagator may be written symbolically as

$$K(x_f, t_f | x_i, t_i) = \sum_{\text{paths}} \exp\left(i S_0 + i \int_{t_i}^{t_f} \nabla\phi \cdot \frac{dx}{dt}\, dt\right) \tag{21}$$

(now taking $\hbar$ and other physical constants to be 1), where S_0 is the action computed with the old Lagrangian. The additional term in the action can be written as

$$\int_{t_i}^{t_f} \nabla\phi \cdot dx, \tag{22}$$

a line integral along the particular path. Now a gauge transformation is known to be equivalent to a multiplication of the wavefunction by a phase factor. In fact, it is the same as the unitary transformation

$$\psi' = U\psi, \quad \langle x | U | y \rangle = \delta(x - y) e^{i\phi(x)}. \tag{23}$$

Therefore, the effect of such a transformation on the propagator is

$$K'(x_f, t_f | x_i, t_i) = K(x_f, t_f | x_i, t_i) \exp\{i[\phi(x_f) - \phi(x_i)]\}. \tag{24}$$

In order for (24) and (22) to agree, it is necessary that

$$\int_{t_i}^{t_f} \nabla\phi \cdot dx = \phi(x_f) - \phi(x_i), \tag{25}$$

independent of the path from x_i to x_f. Just as in the Brownian case, it is the midpoint selection which yields (25). Quantum mechanics thus makes a definite choice of the summation procedure. Incidentally, it may be noted that once the path integral is defined (along with the rule for evaluating the action), the foregoing discussion is a particularly direct way of displaying the relation between gauge transformations and phase transformations.

The work described in the body of the present paper is concerned with manifestations of the $(\Delta x)^2 \sim \Delta t$ phenomena other than Eq. (18) (since this effect is eliminated in quantum mechanics anyway), in particular those arising from a spatial dependence of the coefficient of $\dot{x}^2$ in the path integral.

ACKNOWLEDGMENTS

We wish to thank Professor J. B. Keller and Professor G. Papanicolaou for helpful discussion.

* Work supported in part by Air Force Office of Scientific Research under Grant AFOSR-71-2013.

† NATO Fellow. Address during academic year 1970-1971: Physics Department, Technion, Haifa, Israel.

1 S. F. Edwards and Y. V. Gulyaev, Proc. Roy. Soc. (London) **A279**, 229 (1964).

2 Bryce DeWitt, Rev. Mod. Phys. **29**, 377 (1957).

3 DeWitt's expression differs from ours by a term $-g_{ij}a^i a^j \Delta t$. This term cannot be present, since in particular, it does not appear when g_{ij} is constant.

4 J. H. Van Vleck, Proc. Natl. Acad. Sci. (U.S.) **14**, 178 (1928).

5 C. Morette, Phys. Rev. **81**, 848 (1951).

6 W. Pauli, "Feldquantisierung," lecture notes, Zurich, 1951, Appendix.

7 A. M. Arthurs, Proc. Roy. Soc. (London) **A313**, 445 (1969).

8 E. Nelson, J. Math. Phys. **5** 332 (1964).

9 R. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

10 Calculations such as these make sense only if the integrals are properly defined. This may be done in several ways, one of which is to supply a convergence factor, say $m \to m + i\delta$, $\delta > 0$. This remark applies to all such integrals in this paper.

11 Actually this restriction to flat space is not necessary. In a curved space, one only need to replace Eq. (14) with the appropriate expression valid through order Δt. However, we do not include the case of q being a periodic coordinate.

12 K. Itô, Mem. Am. Math. Soc., No. 4 (1951).

13 R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control* (American Elsevier, New York, 1968), Chap. II.

14 R. L. Stratonovich, J. SIAM Control **4**, 362 (1966).

The Phase Shift. II. As a Continuous Functional

William M. Frank and David W. McLaughlin*
Naval Ordnance Laboratory, Silver Spring, Maryland 20910
(Received 16 March 1971)

The continuity of the phase shift and S-wave scattering length as a functional of the potential is considered. The question is investigated in the present article for sequences of square wells which converge pointwise to zero. The results show certain conditions to be sufficient for convergence of the phase shift and scattering length to zero. These results are generalized to certain integral norms.

I. INTRODUCTION

Recent investigations on the subject of singular potentials[1] have employed a procedure of calculating the phase shifts of a potential as the limit of the phase shifts of a sequence of potentials which converge pointwise to that potential. The presupposition underlying this approach is that if a sequence of potentials converges pointwise to another potential, then the corresponding phase shifts are likewise convergent. This assumption was called into question through a counterexample provided by Calogero[2] and others following him, as well. This raises the over-all question of the validity of such a presupposition for nonsingular as well as singular potentials. This question is investigated for the class of square well potentials in the present article, and for a variety of potentials in the subsequent article. A more or less general solution is presented in that article which provides a fresh perspective on certain ideas in potential theory.

We pose the more constructive question as to what an appropriate sense of convergence of potentials might be that would make the phase shift at any fixed energy a continuous functional of the potential. When will the phase shifts converge uniformly for all energies? The same question is also considered for the S-wave scattering length and for bound states. We find the convergence of the sequence $\int_0^\infty dr\,|V_n(r)|$ to be a sufficient but not a necessary condition for convergence of the phase shifts, though the convergence is not necessarily uniform. We find two norms, $\int_0^\infty dr\,r\,|V_n(r)|$ and $\int_0^\infty dr\,|V_n(r)|^{1/2}$, (the latter only for certain more restricted classes of potentials), whose convergence is a sufficient condition for uniform convergence of the phase shift, and which for repulsive potentials can also be shown to be a necessary condition in the neighborhood of $V(r) = 0$. The results suggest that the potentials might be viewed as the elements of a Banach space with the phase shift a functional which is generally continuous.

We first consider the question of when the convergence of a potential to zero implies the convergence of the corresponding phase shift to zero. This question is studied in the present article in the case of square wells. In the following article the question of when a vanishing potential implies vanishing phase shift and scattering length is studied for general potentials.

II. SQUARE WELLS

We consider scattering by a square well located in the interval $a \le r \le c$ $(c - a \equiv b)$. We adhere to the units $\hbar^2/2m = 1$. If the well is attractive and of depth g, one finds as the expression for the phase shift

$$\tan\delta = \frac{(k^2 + K^2 \tan ka \tan kc)\tan Kb + kK(\tan ka - \tan kc)}{(k^2 \tan kc - K^2 \tan ka)\tan Kb + kK(1 + \tan ka \tan kc)}, \tag{1}$$

where we set $K = (k^2 + |g|)^{1/2}$. For a repulsive potential of height $0 < g < k^2$ located between $r = a$ and $r = c$, one finds $[\kappa = (k^2 - g)^{1/2}]$

$$\tan\delta = \frac{(k^2 + \kappa^2 \tan ka \tan kc)\tan \kappa b + k\kappa(\tan ka - \tan kc)}{(k^2 \tan kc - \kappa^2 \tan ka)\tan \kappa b + k\kappa(1 + \tan ka \tan kc)}. \tag{2}$$

which is identical with (1) with the replacement $K \to \kappa$. For a repulsive potential of height $g > k^2$, one finds $[\gamma \equiv (g - k^2)^{1/2}]$

$$\tan\delta = \frac{(k^2 - \gamma^2 \tan ka \tan kc)\tanh \gamma b + k\gamma(\tan ka - \tan kc)}{(k^2 \tan kc + \gamma^2 \tan ka)\tanh \gamma b + k\gamma(1 + \tan ka \tan kc)}. \tag{3}$$

We pose the questions: If $V(r) \to 0$ pointwise, under what circumstances will it follow that $\delta \to 0$?; that the scattering length $A \to 0$?; that there are no bound states? Two extreme cases will be considered: (i) when the sequence of square wells becomes taller and narrower eventually approaching a "spike," i.e., $b \to 0, g \to \infty$, and (ii) when the sequence of square wells becomes broader and shallower eventually flattening out to no interaction, i.e., $b \to \infty, g \to 0$. In all these cases, we presume the sequence of potentials to approach zero pointwise. In the $b \to 0, g \to \infty$ case, this would mean that the center of the well moves in such a way that the limiting value of the potential at any fixed point is zero. An example would be the sequence of potentials

TABLE I. Limiting values for the case $b \to 0, g \to \infty$ with $a \neq 0$.

Case (i), $a \neq 0$	Limiting value	Sufficient condition for no effect
$k \neq 0$ Att. or rep.	$\tan\delta = -\dfrac{\zeta \sin^2 ka}{k + \frac{1}{2}\zeta \sin 2ka}$	$\zeta = 0$
$k = 0$ Att. or rep.	$A = \dfrac{-\zeta}{1 + \zeta a}$	$\zeta = 0$
Bound state threshold	$-\zeta = 1/a$, $\eta = 0$ or $\zeta = \infty$, $\eta \to n\pi + (ag^{1/2})^{-1}$ $(n \neq 0)$	$-\zeta < 1/a$

TABLE II. Limiting values for the case $b \to 0$, $g \to \infty$ with $a = 0$.

Case (i), $a = 0$	Limiting value	Sufficient condition for no effect
$k \neq 0$ Att.	$\tan\delta = k/f$ if $n \sim (n + \frac{1}{2}\pi - f\lvert g\rvert^{-1/2}$	$\eta < \frac{1}{2}\pi$
$k \neq 0$ Rep.	$\tan\delta = 0$	Always
$k = 0$ Att.	$A = 1/f$ if $\eta \sim (n + \frac{1}{2})\pi - f\lvert g\rvert^{-1/2}$	$\eta < \frac{1}{2}\pi$
$k = 0$ Rep.	$A = 0$	Always
Bound state threshold	$\eta \to (n + \frac{1}{2})\pi$	$\eta < \frac{1}{2}\pi$

$$V_n(r) \equiv g_n \theta\left(\frac{1}{2^{n+1}} \leq r \leq \frac{1}{2^n}\right),$$

where

$$\theta(a \leq r \leq c) \equiv \theta(r - a)\,\theta(c - r) = \begin{cases} 0 \text{ if } r < a \\ 1 \text{ if } a \leq r \leq c, \\ 0 \text{ if } r > c \end{cases}$$

with $\theta(x)$ the familiar step function, unity for non-negative values of x and zero otherwise. The typical potentials of the sequence under consideration will be of the form

$$V_n(r) \equiv g_n \theta(a_n \leq r \leq c_n) \tag{4}$$

and we write $b_n \equiv c_n - a_n$. The index n will frequently be omitted when confusion is unlikely. The openness or closedness of the interval of support of the square well is immaterial, and we shall generally assume support in a closed interval.

In what follows, the following limits will interest us:

$$\xi \equiv \lim_{n\to\infty} b_n g_n^{1/3},$$
$$\eta \equiv \lim_{n\to\infty} b_n \lvert g_n\rvert^{1/2}, \tag{5}$$
$$\zeta \equiv \lim_{n\to\infty} b_n g_n,$$

where the limiting procedures may correspond either to case (i) $b \to 0, g \to \infty$ or case (ii) $b \to \infty$, $g \to 0$.

Case (i), $b \to 0, g \to \infty$

In this case it is clear that $\zeta \to 0$ implies $\eta \to 0$ which implies $\xi \to 0$, and that $\xi < \infty$ implies $\eta = \infty$, while $\eta < \infty$ implies $\zeta = \infty$. The results of the appropriate limiting procedure applied to Eqs. (1) and (3) are summarized in Tables I and II. We have defined $a \equiv \lim a_n = \lim c_n$, as $n \to \infty$.

For the case $a \neq 0, k \neq 0$, one finds as the appropriate limiting expression

$$\tan\delta \to \frac{-\zeta \sin^2 ka}{k + \frac{1}{2}\zeta \sin 2ka} \tag{6}$$

valid for both the attractive and repulsive cases (ζ is negative in the attractive case). Clearly, the vanishing of ζ is a sufficient condition for no scattering and is necessary unless $ka = n\pi$ (n is an integer). Since this exception occurs only for a discrete set of accidental values of a parameter, we shall speak of the nonvanishing of ζ as a "necessary" condition for the nonvanishing of the phase shift. (The term "sufficient" will be used in a similar sense.) We note that

$$\lvert\zeta\rvert = \lim_{n\to\infty} \int_0^\infty dr\, \lvert V_n(r)\rvert, \tag{7}$$

which suggests that the vanishing of the quantity on the right-hand side of Eq. (7) might be a sufficient condition for no scattering for a general sequence of potentials. We shall see in the following article that this is indeed the case.

The finiteness of ζ corresponds to a δ-function type spike, and the phase shift as given in Eq. (6) is exactly what one would obtain from a δ-function potential of strength ζ located at $r = a \neq 0$. The condition $a \neq 0$ is crucial.

If we let $\lvert\zeta\rvert \to \infty$ in Eq. (6), we find

$$\lim_{\lvert\zeta\rvert\to\infty} \tan\delta = -\tan ka, \tag{8}$$

which corresponds to the phase shift due to an infinitely repulsive barrier at $r = a$. It will be noted that the boundary condition $u(0) = 0$ with $u(r) = r\psi(r)$ [$\psi(r)$ the wavefunction[3]] for the three-dimensional problem corresponds to a one-dimensional problem with an infinite repulsive barrier for $r < 0$. A δ function of infinite strength, attractive or repulsive (i.e., $\lvert\zeta\rvert = \infty$) at $r = a$ gives rise to an infinite slope on the emerging side of the potential and therefore to an effective boundary condition at $r = a$ of infinite logarithmic derivative. This is equivalent to a vanishing wavefunction at $r = a$, which is as if there were an infinite barrier at $r = a$. This is the statement in Eq. (8), since generally $\tan\delta = -\tan kr_0$, where r_0 is a node of the asymptotic wavefunction.

The scattering length obtained from Eqs. (1) and (3) for the case $a \neq 0$ agrees with that extrapolated from Eq. (6), *viz.*

$$A = -\zeta a^2/(1 + \zeta a). \tag{9}$$

In the attractive case $\zeta < 0, A$ has a pole at $-\zeta = 1/a$. This is precisely the condition for the appearance of the first bound state. One readily verifies that the condition for binding a bound state with energy 0 in the potential ($g > 0$)

$$V(r) = -g\theta(a \le r \le c) \tag{10}$$

is

$$g^{1/2} \tan\eta = 1/a. \tag{11}$$

The weakest coupling solution to Eq. (11) corresponds to $\eta \approx 0$, which gives the condition $|\zeta| = b|g| = 1/a$. On the other hand, the other solutions to Eq. (11) all correspond to infinite ζ, as would be expected in the presence of at least one bound state. The other solutions to Eq. (11) are of the form

$$\eta = n\pi + (ag^{1/2})^{-1} + o(g^{-1/2}) \tag{12}$$

for $n \neq 0$, corresponding to finite nonvanishing η, prescribing not a value but a rate of convergence to a value of η. The emergence of η in the criterion for binding is not surprising in view of the $L^{1/2}$ necessary condition for a bound state[4]

$$\tfrac{1}{2}\pi \le \int dr\, |V(r)|^{1/2}, \tag{13a}$$

or the Bargmann–Schwinger condition[5]

$$1 \le \int dr\, r|V(r)|. \tag{13b}$$

The quantity on the right of Eq. (13a) corresponds to η, while that in Eq. (13b) corresponds to η^2. The reason that η is the significant parameter for binding will be discussed subsequently.

The finiteness of η plays the leading role in the case that $a = 0$. We first consider the attractive case. From Eq. (1), one finds (remembering that $b = c$, if $a = 0$)

$$\tan\delta = \frac{k/K \tan Kb - \tan kb}{1 + (k/K)\tan Kb \tan kb} = \tan(\omega - kb), \tag{14}$$

where

$$\tan\omega = (k/K)\tan Kb. \tag{15}$$

Since $b \to 0$, a "sufficient" and necessary condition for nonvanishing scattering is the nonvanishing of ω. From Eq. (15), one sees that unless $Kb \to (n + \frac{1}{2})\pi$, ω will tend to zero. In fact one finds that if

$$Kb \to (n + \tfrac{1}{2})\pi - f|g|^{-1/2} + o(g^{-1/2}), \tag{16}$$

then $\tan\omega \to k/f$. The statement in Eq. (16) is a statement about how rapidly η approaches $(n + \frac{1}{2})\pi$. In order to understand this situation, it should be noted that f is merely the logarithmic derivative at $r = b$, the edge of the square well. This follows from Eqs. (14) and (15) and the expression for the logarithmic derivative

$$\text{log der} \equiv k \cot(kb + \delta) = k \cot\omega = K \cot Kb \to fK|g|^{-1/2} \to f. \tag{17}$$

As is well-known,[6] a resonance takes place when the logarithmic derivative at the "potential boundary" vanishes and the penetration of the potential region will be depressed as this quantity increases in magnitude. Thus scattering windows are present in the limit, only when Eq. (16) holds for finite values of f, with a resonance at $f = 0$.

The reason that the role of arbiter of scattering has changed from the parameter ζ to the parameter η is the vanishing boundary condition at $r = 0$. This "clamps" the wavefunction at $r = 0$ and a stronger spike than a δ function at the origin is necessary to produce scattering. One readily verifies that a δ-function potential located at the origin produces no scattering (unlike a δ function located elsewhere), while the more singular potential $\delta(r)/r$ gives resonance scattering (corresponding to $f = 0$). What actually happens is that the singularity selects the *irregular solution* which does not vanish at $r = 0$, but whose derivative vanishes at $r = 0$ corresponding to a vanishing logarithmic derivative. Thus infinite ζ causes the wavefunction to emerge from the "interaction region" with infinite slope, and therefore with infinite logarithmic derivative. The effect of a singularity corresponding to a finite value of η is to bend the function in the interaction region further and to make penetration from the outside region possible only when the slope at the boundary is zero. Zero logarithmic derivative is characteristic of the irregular solution, as well as threshold binding. Such a pitfall was indeed encountered by Calogero[2] in a case where a (singular) repulsive potential was approached by a sequence of potentials, which contained a deep attractive pocket which converged to an attractive spike at $r = 0$.

For $k = 0, a = 0$, one obtains the correct scattering length directly from Eq. (14) and one finds

$$A = g^{-1/2}(\tan\eta - \eta). \tag{18}$$

In complete analogy with the $k \neq 0$ case, one finds a nonvanishing scattering length if

$$\eta \to (n + \tfrac{1}{2})\pi - f|g|^{-1/2} + o(g^{-1/2}), \tag{19}$$

which results in $A = 1/f$. The interpretation is the same as before.

In the repulsive case where $k \neq 0, a = 0$, one finds

$$\tan\delta = \frac{k/\gamma \tanh\gamma b - \tan kb}{1 + k/\gamma \tan kb \tanh\gamma b}, \tag{20}$$

where $\gamma = (g - k^2)^{1/2}$ which in the limit $b \to 0$, $g \to \infty$ gives no scattering under any circumstances. This is readily understood in the light of earlier considerations. A repulsive potential makes the wavefunction show a convex face toward the u axis [$u(r) = r\psi(r)$] and one never approaches the penetration condition $f < \infty$. If $k = 0$, one can go to the limit directly in Eq. (20) and find that the scattering length vanishes unconditionally.

If $a = 0$ the criterion for binding at zero energy reads

TABLE III. Limiting values for the case $b \to \infty$, $g \to 0$.

Case (ii)	Limiting value	Sufficient condition for no effect
$k \neq 0$ Att. or rep.	$\tan\delta = \tan(\zeta/2k)$	$\zeta = 0$
$k = 0$ Att. or rep.	$A = -\frac{1}{3}\xi^3$	$\xi = 0$
Bound state threshold	$\eta \sim (n + \frac{1}{2})\pi - a\lvert g\rvert^{1/2}$	$\eta < \frac{1}{2}\pi$

$$g^{1/2}\cot\eta = 0, \tag{21}$$

which implies $\eta \to (n + \frac{1}{2})\pi + o(g^{-1/2})$ which is essentially the same as Eq. (19).

To summarize, we found that the important distinction for case (i) is whether a does or does not vanish. If $a \neq 0$, a criterion for (nontrivial) scattering, is the nonvanishing of ζ, while for $a = 0$ scattering was possible only in "small neighborhoods" of a set of discrete values of η.

The existence of a bound state was possible if $a \neq 0$ only for one finite value of ζ, but for an infinite number of discrete values of η, while for $a = 0$ only the latter type criterion in η was required for bound states. It was mentioned earlier that the limit $\zeta \to \infty$ corresponded effectively to an infinite barrier at $r = a$, at which the wavefunction takes on the value zero. In the case $a = 0$, we saw that a sufficiently strong singularity at the origin produced the irregular solution corresponding to vanishing slope. In the same way the effectively infinite barrier at $r = a$, may, if a sufficiently singular potential is present, i.e., $\eta \neq 0$, allow solutions with vanishing slope, which is the condition for a zero energy bound state.

The result that $\zeta \to 0$ is a sufficient condition for no scattering would be expected to generalize to the corresponding quantity in Eq. (7). Generalizations broad enough to include the results for the $a = 0$ case might have the forms

$$\int_0^\infty dr\, \frac{r}{1 + \beta r}\lvert V_n(r)\rvert \to 0 \tag{22a}$$

(β, some positive constant) or

$$\int_0^\infty dr\, r\lvert V_n(r)\rvert \to 0. \tag{22b}$$

We shall see in the following paper that these are indeed sufficient conditions for the vanishing of the phase shift.

Case (ii), $b \to \infty, g \to 0$

In this case it is clear that $\xi \to 0$ implies $\eta \to 0$, which implies $\zeta \to 0$, and that $\zeta < \infty$ implies $\eta = \infty$, while $\eta < \infty$ implies $\xi = \infty$. The results for the present case are summarized in Table III.

One finds in the present case that there is no distinction between $a \neq 0$ and $a = 0$. This is as one might expect since the range is asymptotically large and boundary effects at $r = 0$ are washed out. One finds for $k \neq 0$ that the phase shift is given from

$$\tan\delta = \tan(\zeta/2k), \tag{23}$$

a result valid for both the attractive and repulsive cases, for which, of course, ζ will differ in sign. For large k, this result is in agreement with the Born approximation. No limit is approached in the low frequency case. This is not difficult to understand in view of the approach to infinity of three distinct lengths $b, g^{-1/2}$, and $\lambda = 1/k$ whose commensurability determines the phase shift. One finds no limit approached as well in the case $\zeta \to \infty$. Either of the generalizations of ζ expressed in Eqs. (7) or (22a) are divergent for the Coulomb-like tail for which one indeed knows that the phase shift as conventionally defined does not exist.

In the $k = 0$ case, one finds the result

$$A = -\tfrac{1}{3}\,\xi^3 \tag{24}$$

valid for both the attractive and repulsive cases (where the sign of ξ is different) and for $a \neq 0$ as well as $a = 0$. The natural generalization of ξ^3 as a criterion would correspond to the integral $\int_M^\infty dr\, r^2 V(r)$ whose finiteness we know is a condition for a finite scattering length. The range $0 < \xi < \infty$ which corresponds to a finite scattering length corresponds to $\zeta = 0$. Thus while no scattering is present in $\zeta \to 0$ limit at nonzero energies, one generally finds a nonvanishing scattering length. The difference results from the damping in the finite wavelength case which is absent when $\lambda \to \infty$. We easily see that $\xi \to \infty$, which gives an infinite scattering length, does not correspond to the presence of a bound state at zero energy as in case (i). Rather, the infinite A derives from the infinite range of the interaction $b \to \infty$. The scattering length among other things is a measure of the range of the interaction.

The criterion for binding a bound state at zero energy is again expressed in terms of η. For $a \neq 0$ the condition for binding is again expressed by Eq. (11) which requires that

$$\eta \to (n + \tfrac{1}{2})\pi - a\lvert g\rvert^{1/2} + o(g^{1/2}). \tag{25}$$

In view of $a + b = c$, this can be written as

$$c\lvert g^{1/2}\rvert \to (n + \tfrac{1}{2})\pi + o(g^{1/2}). \tag{25'}$$

For $a = 0$ the condition is again expressed by Eq. (21), and the resulting condition is

$$\eta = b\lvert g\rvert^{1/2} \to (n + \tfrac{1}{2})\pi + o(g^{1/2}). \tag{26}$$

Since $b = c$ when $a = 0$, this condition coincides in form with Eq. (25'). The $a = 0$ and $a \neq 0$ cases also resemble each other in that the binding corresponds to the same asymptotic values of η, which was not true for case (i).

We conclude in this case that $\zeta_n \to 0$ is a sufficient condition for no scattering, a result which would be expected to generalize to the conditions $\int dr\,|V_n(r)| \to 0$ or $\int dr\,r\,|V_n(r)|/(1+\beta r) \to 0$. Similarly $\eta \to 0$ which implies no bound states, corresponds either to $\int dr\,|V_n(r)|^{1/2} \to 0$ or $\int dr\,r\,|V_n(r)| \to 0$. The latter condition emerges as the one applicable to all the cases considered, while the former applies to somewhat more restricted classes of potentials.

The results of the present article are generalized to the proper norms for ordinary potentials in the following article.

---

* Part of this work was carried out when author was at the Naval Ordance Laboratory during the summer of 1967. Present address: Courant Institute, New York University, New York, N.Y. and Mathematics Dept., University Heights Campus, New York University, Bronx, New York.

1 N. N. Khuri and A. Pais, Rev. Mod. Phys. **36**, 590 (1964).
2 F. Calogero, Phys. Rev. **139**, B602 (1965).
3 We shall call $u(r) = r\psi(r)$ the radial wave function.
4 F. Calogero, Nuovo Cimento **36**, 199 (1965). This result is valid only if $|V(r)|$ is monotonically nonincreasing.
5 V. Bargmann, Proc. Natl. Acad. Sci. (U.S.) **38**, 961 (1952); J. Schwinger, *ibid.* **47**, 122 (1961).
6 J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), Chap. VIII, Sec. 79.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Phase Shift. III. As a Continuous Functional

William M. Frank
U.S. Naval Ordnance Laboratory, Silver Spring, Maryland 20910
(Received 16 March 1971)

The continuity of the phase shift and scattering length as a functional of the potential is considered. We find that the convergence of $\int dr\,|V_n|$ to zero is a sufficient though not necessary condition for the vanishing of the phase shift, while the convergence of $\int dr\,r\,|V_n(r)|$ or $\int dr\,|V_n(r)|^{1/2}$ to zero would be a sharp condition. A number of theorems of this type are proven. Theorems are also proven demonstrating the continuity of the phase shift and scattering length as a sequence of potentials converging in the proper norms to a nonzero potential. Some implications of the existence of a Banach space of potentials are discussed briefly.

I. INTRODUCTION

In Paper I of this series[1] the phase shift and other "scattering functions" of potential theory considered as functionals of the potential were shown to have absolute bounds related to certain integral norms of the potential. Paper II studied in detail the question of when the convergence of a potential to zero implies the convergence of the corresponding phase shifts to zero for the case of square well potentials. It was found that such convergence applies when convergence of certain integral norms for sequences of these potentials holds. In this paper, these results are found to apply quite generally to nonsingular potentials. The behavior of the phase shifts and wavefunctions corresponding to a sequence of potentials which converge to a nonzero potential is also dealt with in this article. Certain aspects of continuity of the phase shift for cases of singular potentials are discussed in an appendix.

II. CONTINUITY IN THE NEIGHBORHOOD OF $V(r) = 0$

We consider potentials $V(r)$ which belong to the class L^1 in $[0,\infty)$,[1] i.e., potentials which satisfy $\zeta \equiv \int_0^\infty dr\,|V(r)| < \infty$.[2] The phase shift $\delta(k)$ at an energy $k^2(\hbar^2/2m = 1)$ is a functional of $V(r)$.

Given a sequence of potentials $\{V_n(r)\}$ which approach zero pointwise; under what circumstances does it follow that the sequence of phase shifts $\delta_n(k)$ vanishes for each value of k? Under what circumstances does $\delta_n(k) \to 0$ for each k imply that $V_n(r) \to 0$ pointwise, or uniformly, or in some other norm? (Our phase shifts are normalized by the convention that they are continuous for real k and that $\delta(k) \to 0$ as $k \to \infty$.) The methods of the present section lean on the results of the first of these papers which we shall refer to as I, and we shall refer to equations in I by attaching I to the equation number.

We prove the following theorem which generalizes the result of Paper II that $\zeta \to 0$ implies no scattering.

Theorem 1: Let $\{V_n(r)\}$ be a sequence of potentials such that

$$\zeta_n \equiv \int dr\,|V_n(r)| \to 0. \tag{1}$$

Then $\delta_n(k) \to 0$ for all $k \neq 0$, uniformly for all k in any open interval of the positive k axis which excludes a neighborhood of $k = 0$.

Proof: The result follows immediately from the inequalities (I 9) by setting $L = 0$. In Appendix A of I it was shown that the same inequality is valid for all partial waves. We see from Eq. (I. 9) that the convergence is uniform in any interval of the k axis excluding a neighborhood of $k = 0$. That $k = 0$ can be a *bona fide* exceptional point is seen

We conclude in this case that $\zeta_n \to 0$ is a sufficient condition for no scattering, a result which would be expected to generalize to the conditions $\int dr\,|V_n(r)| \to 0$ or $\int dr\,r\,|V_n(r)|/(1+\beta r) \to 0$. Similarly $\eta \to 0$ which implies no bound states, corresponds either to $\int dr\,|V_n(r)|^{1/2} \to 0$ or $\int dr\,r\,|V_n(r)| \to 0$. The latter condition emerges as the one applicable to all the cases considered, while the former applies to somewhat more restricted classes of potentials.

The results of the present article are generalized to the proper norms for ordinary potentials in the following article.

---

* Part of this work was carried out when author was at the Naval Ordance Laboratory during the summer of 1967. Present address: Courant Institute, New York University, New York, N.Y. and Mathematics Dept., University Heights Campus, New York University, Bronx, New York.

1 N. N. Khuri and A. Pais, Rev. Mod. Phys. **36**, 590 (1964).
2 F. Calogero, Phys. Rev. **139**, B602 (1965).
3 We shall call $u(r) = r\psi(r)$ the radial wave function.
4 F. Calogero, Nuovo Cimento **36**, 199 (1965). This result is valid only if $|V(r)|$ is monotonically nonincreasing.
5 V. Bargmann, Proc. Natl. Acad. Sci. (U.S.) **38**, 961 (1952); J. Schwinger, *ibid.* **47**, 122 (1961).
6 J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), Chap. VIII, Sec. 79.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Phase Shift. III. As a Continuous Functional

William M. Frank
U.S. Naval Ordnance Laboratory, Silver Spring, Maryland 20910
(Received 16 March 1971)

The continuity of the phase shift and scattering length as a functional of the potential is considered. We find that the convergence of $\int dr\,|V_n|$ to zero is a sufficient though not necessary condition for the vanishing of the phase shift, while the convergence of $\int dr\,r\,|V_n(r)|$ or $\int dr\,|V_n(r)|^{1/2}$ to zero would be a sharp condition. A number of theorems of this type are proven. Theorems are also proven demonstrating the continuity of the phase shift and scattering length as a sequence of potentials converging in the proper norms to a nonzero potential. Some implications of the existence of a Banach space of potentials are discussed briefly.

I. INTRODUCTION

In Paper I of this series[1] the phase shift and other "scattering functions" of potential theory considered as functionals of the potential were shown to have absolute bounds related to certain integral norms of the potential. Paper II studied in detail the question of when the convergence of a potential to zero implies the convergence of the corresponding phase shifts to zero for the case of square well potentials. It was found that such convergence applies when convergence of certain integral norms for sequences of these potentials holds. In this paper, these results are found to apply quite generally to nonsingular potentials. The behavior of the phase shifts and wavefunctions corresponding to a sequence of potentials which converge to a nonzero potential is also dealt with in this article. Certain aspects of continuity of the phase shift for cases of singular potentials are discussed in an appendix.

II. CONTINUITY IN THE NEIGHBORHOOD OF $V(r) = 0$

We consider potentials $V(r)$ which belong to the class L^1 in $[0,\infty)$,[1] i.e., potentials which satisfy $\zeta \equiv \int_0^\infty dr\,|V(r)| < \infty$.[2] The phase shift $\delta(k)$ at an energy $k^2(\hbar^2/2m = 1)$ is a functional of $V(r)$.

Given a sequence of potentials $\{V_n(r)\}$ which approach zero pointwise; under what circumstances does it follow that the sequence of phase shifts $\delta_n(k)$ vanishes for each value of k? Under what circumstances does $\delta_n(k) \to 0$ for each k imply that $V_n(r) \to 0$ pointwise, or uniformly, or in some other norm? (Our phase shifts are normalized by the convention that they are continuous for real k and that $\delta(k) \to 0$ as $k \to \infty$.) The methods of the present section lean on the results of the first of these papers which we shall refer to as I, and we shall refer to equations in I by attaching I to the equation number.

We prove the following theorem which generalizes the result of Paper II that $\zeta \to 0$ implies no scattering.

Theorem 1: Let $\{V_n(r)\}$ be a sequence of potentials such that

$$\zeta_n \equiv \int dr\,|V_n(r)| \to 0. \tag{1}$$

Then $\delta_n(k) \to 0$ for all $k \neq 0$, uniformly for all k in any open interval of the positive k axis which excludes a neighborhood of $k = 0$.

Proof: The result follows immediately from the inequalities (I 9) by setting $L = 0$. In Appendix A of I it was shown that the same inequality is valid for all partial waves. We see from Eq. (I. 9) that the convergence is uniform in any interval of the k axis excluding a neighborhood of $k = 0$. That $k = 0$ can be a *bona fide* exceptional point is seen

by considering the sequence

$$V_n(r) = g_n/(1 + r^\alpha),$$

with $1 < \alpha < 2$ and $g_n \to 0$. It can be shown that all the potentials of the sequence have an infinite number of bound states and hence, certainly, a nonvanishing phase shift at $k = 0$. The lack of uniformity is, however, not necessarily connected with the presence of boundary states. The expression for $\tan\delta$ in (II 23) for a sequence of repulsive potentials manifests an obvious non uniformity in k near $k = 0$. We shall term $\zeta \equiv \int dr\,|V(r)|$ the L^1 norm of $V(r)$.

The next theorem valid for all partial waves will be explicitly demonstrated only for S waves. The proof for other waves is easily constructed.

Theorem 2: $\{V_n(r)\}$ be a sequence of potentials such that

$$\omega_{\beta,n} \equiv \omega_n \equiv \int_0^\infty dr\,\frac{r\,|V_n(r)|}{1+\beta r} \to 0, \tag{2}$$

where the ω_n are uniformly bounded. Then $\delta_n(k) \to 0$ for all $k \neq 0$ uniformly in any open interval of the k axis, which excludes a neighborhood of $k = 0$. The sequence of wavefunctions converges uniformly pointwise to the free solution.

N.B. $\zeta_n \to 0$ implies that $\omega_n \to 0$.

Proof: The inequality Eq. (I 10a) sets a vanishing lower bound on the phase shifts. Let M be a uniform bound to ω_n. We proceed from[3]

$$\sin\delta = -(1/k)\int_0^\infty dr\ \sin kr\,V(r)\bar{u}(r), \tag{3}$$

where $\bar{u}(r)$ is the radial wavefunction normalized to have the asymptotic behavior $\bar{u}(r) \sim \sin(kr + \delta)$ which obeys the integral equation[4]

$$\bar{u}(r) = \sin kr + (1/k)\int_0^r dr'\,\sin k(r-r')V(r')\bar{u}(r'). \tag{4}$$

Iteration and the inequalities

$$\frac{\sin kr}{k} \le \frac{2r}{1+kr} \le \frac{2\beta r}{k(1+\beta r)} \tag{5}$$

yields for $k \neq 0$ to

$$|\bar{u}(r) - \sin kr| \le \exp\left(\frac{2\beta}{k}\,\omega_\beta[V]_0^r\right) - 1 \le C - 1 < \infty \tag{6}$$

in the notation of I (see Sec. II), where $C \equiv \exp(2\beta M/k)$. Alternately, we may derive that

$$|\bar{u}(r) - \sin kr| \le \exp 2\omega_k[V]_0^r - 1. \tag{7}$$

Equation (6) implies the uniform convergence of $\bar{u}(r)$ to the free solution when $\omega_n \to 0$ if $k \neq 0$. Equation (3) together with Eq. (6) imply that

$$|\sin\delta| \le \frac{2\beta C}{k}\int_0^\infty dr\,\frac{r|V(r)|}{1+\beta r} \le \frac{2\beta M}{k}\,\exp\left(\frac{2\beta M}{k}\right) \tag{8}$$

from which the theorem follows. By employing the inequality $(\sin\chi) \le 1$, we prove Theorem 1 again. The present theorem allows stronger singularities in the potential than Theorem 1 does as for the case of a Yukawa potential. These two theorems include as special cases all the conditions suggested in II for the vanishing of the phase shift. The lack of uniformity in the neighborhood of $k = 0$ is to be noted again.

The following theorem provides for uniform convergence in k.

Theorem 3: Let $\{V_n(r)\}$ be a sequence of potentials such that

$$\chi_n \equiv \int dr r\,|V_n(r)| \to 0, \tag{9}$$

then $\delta_n(k) \to 0$ uniformly for all real k.

N.B. $\chi_n \to 0$ implies that $\omega_{\beta,n} \to 0$. The present proof applies only to S waves. We denote by $L^{(1)}$ the class of potentials satisfying $\int dr r|V(r)| < \infty$.

Proof: The result is an immediate consequence of Eqs. [I (10b)] and [I (19)]. The convergence of the wavefunction for $k \neq 0$ is a consequence of Theorem 2 since $\chi_n \to 0$ implies $\omega_n \to 0$. A glance at Eq. (7) shows the same to be true at $k = 0$. The uniformity is related to the absence of bound states.[5]

We shall present another proof of this theorem by a method which will anticipate the arguments of the next section. The method employs the analytical properties of the Jost function in k, the "energy" variable. We imagine a common coupling constant coefficient g attached to all the $V_n(r)$. We eventually set $g = 1$. For not too pathological potentials,[6] a necessary and sufficient condition for the existence of the Jost function is the finiteness of the $L^{(1)}$ norm. A number of properties of the Jost function will be utilized. One is its analytic character in the entire g plane[7] and in the half-k plane ($\mathrm{Im}k \le \mu \ge 0$ for some μ)[8] if the potentials satisfy $\chi \equiv \int dr r|V(r)| < \infty$.

Another is the expression of $f(k;g)$ (the g dependence will frequently be suppressed in the notation) as a power series expansion[9]

$$f(k;g) = \sum_{n=0}^\infty \frac{g^n}{k^n}\int_0^\infty dr_1 \cdots \int_0^{r_{n-1}} dr_n \times V(r_1)\cdots V(r_n)e^{-ikr_1}\ \sin k(r_1 - r_2)\cdots \sin k(r_{n-1} - r_n)\,\sin kr_n, \tag{10}$$

and the third is the set of dispersion relations[10]

$$\ln|f(k)| = -\frac{P}{\pi}\int_{-\infty}^\infty dk'\,\frac{\delta(k')}{k'-k} + \sum_{j=1}^{n_B}\ln\frac{E-E_j}{E} \tag{11}$$

$$\delta(k) = \frac{P}{\pi}\int_{-\infty}^\infty dk'\,\frac{\ln|f(k')|}{k'-k} + 2\sum_{j=1}^{n_B}\cot^{-1}\left(\frac{k}{\kappa_j}\right), \tag{12}$$

where $E_j\,(j = 1, \ldots, n_B)$ is in the set of bound state energies and $E_j = -\kappa_j^2$. The condition $\int dr r|V(r)| < \infty$ need not require the convergence of the dispersion integrals; however, the once subtracted

dispersion relations (which we write now without bound state terms) are necessarily convergent. This follows from the uniform boundedness of the $\delta(k)$ when the $L^{(1)}$ norm is bounded, as established in the boundedness lemma of I.

The once substracted dispersion relations read

$$\ln\left|\frac{f(k)}{f(0)}\right| = \frac{-k}{\pi} P\int dk' \frac{\delta(k')}{k'(k'-k)}, \tag{13}$$

$$\delta(k) = \frac{k}{\pi} P\int dk' \frac{\ln|f(k')|}{k'(k'-k)}, \tag{14}$$

where we have dropped the bound state terms, since $\chi_n \to 0$ implies that eventually there are no bound states.[5] The coefficients $\phi_m(k)$ of g^m in Eq. (10) satisfy the inequality

$$|\phi_m(k)| \le \frac{1}{m!}\left[\int_0^\infty dr r|V(r)|\right]^m. \tag{15}$$

The convergence of the power series or $f_n(k)$[11] is uniform in n and k for k real. If $\chi_n \to 0$, then $|f_n(k)| \to 1$ and $\ln|f_n(k)| \to 0$ so that the right side of the dispersion relation Eq. (14), gives zero which implies that $\delta_n(k) \to 0$ uniformly in k.

We can infer similar results for potentials of the $\bar{L}^{(1/2)}$ class (see Sec. II of I).

Theorem 4: Let $V_n(r)$ be a sequence of $\bar{L}^{(1/2)}$ potentials, whose $L^{(1/2)}$ norms τ_n approach zero with n, then $\delta_n(k) \to 0$ uniformly with k. This result is valid for all partial waves.

Proof: The coefficients of the Jost function for $L^{(1/2)}$ potentials (which always exists) can be bounded by[12]

$$|\phi_m(k)| \le \tau^{2m}/(2m)!. \tag{16}$$

$\tau_n \to 0$ implies $f_n(k;g) \to 1$ and the argument proceeds as in Theorem 3.

The above theorems have supplied only sufficient conditions on the potential sequence. Necessary conditions have been found only for sequences of repulsive potentials, and are given in the following theorems. The writer believes that the same results are perhaps valid for sequences of attractive or mixed potentials, provided one has conditions on the absence of bound states. The following result is valid for all partial waves.

Theorem 5: Let $\{\delta_n(k)\}$ be a sequence of phase shift functions (of k) each corresponding to an $L^{(1)}$ potential, each of which for sufficiently large n is everywhere nonnegative, and suppose there exists a $\nu > 0$ such that for all n

$$\lim_{r\to\infty} e^{\nu r} V_n(r) \to 0. \tag{17}$$

If $\delta_n(k) \to 0$ for each k, the sequence of $L^{(1)}$ norms χ_n approaches zero.

Remark: It is not clear whether condition (17) reflects a limitation of mathematical technique or a true physical condition.

Proof: The lemma on the uniform boundedness of $|\delta(k)|$ implies from Eq. (13) that, if $\delta_n(k) \to 0$, then

$$\lim_{n\to\infty}\left|\frac{f_n(k)}{f_n(0)}\right| \to 1. \tag{18}$$

By hypothesis $\delta_n(k) \to \delta_n(0) = 0$. Hence the quantity $f_n(k)/f_n(0) \to 1$ for all real k. For purely repulsive potentials $f_n(0) \ge 1$. If $\lim e^{\nu r} V_n(r) \to 0$ as $r \to \infty$ for some $\nu > 0$, then the inequality[13]

$$|f(k) - 1| \le C\int_0^\infty \frac{dr r|V(r)|}{1 + |k|r} \exp[2\theta(k)\mathrm{Im}kr] \tag{19}$$

(C some constant) shows $f_n(k)/f_n(0)$ to be uniformly bounded in k and n in the half-k plane $\mathrm{Im}k < \frac{1}{2}\nu$. One concludes from Vitali's convergence theorem[14] that $\lim f_n(k)/f_n(0)$ as $n \to \infty$ is an analytic function of k in the aforementioned open half-plane and must therefore be identically 1. From Eq. (19), $f(k)$ approaches one as $|k| \to \infty$ within the interior of this analyticity domain. Consequently,

$$\lim_{n\to\infty} f_n(k) \equiv 1 \tag{20}$$

for all k. For purely repulsive potentials all the terms in the power series expansion [Eq. (10)] for $f(0)$ are positive, and $f_n(0) \to 1$ implies that

$$\lim_{n\to\infty} \int dr r V_n(r) \to 0, \tag{21}$$

since this quantity is the coefficient of g in the expansion of the Jost function at $k = 0$.

A variety of alternative conditions are possible. In place of Eq. (17), a condition that the potentials be analytic[15] in r in an appropriate sector would suffice to make the real k-axis (excluding possibly $k = 0$) interior to a domain in which $f_n(k)/f_n(0)$ is uniformly bounded, which would lead to the same result.

All the sufficient conditions of Paper II regarding square wells have been proven in the present section. In the following section we discuss continuity properties in the neighborhood of an arbitrary potential

III. CONTINUITY IN THE NEIGHBORHOOD OF ANY $V(r)$

In the present section, we shall prove some results concerning the convergence of a sequence of phase shifts when the corresponding sequence of potentials $\{V_n(r)\}$ converges to a potential $V(r)$, which is also L^1. These include the results of the previous section as special cases, though we can prove less in the general case. An alternative type of problem, which we shall deal with incidentally but not systematically, concerns the convergence of the phase shifts when the sequence of potentials constitutes a Cauchy sequence in some norm. The

existence of an appropriate limiting potential entails a space of potential functions which is complete. We do not propose to discuss this matter in this paper.

We prove the analog of Theorem 1. It is valid for all partial waves.

Theorem 6: Let $\{V_n(r)\}$ be a sequence of potentials which constitutes a Cauchy sequence in the L^1 norm; i.e., for given ϵ, when m, n are greater than an appropriate $N(\epsilon)$,

$$\int dr\,|V_n(r) - V_m(r)| < \epsilon. \tag{22}$$

We also presume $\int dr\,|V_n(r)| < M < \infty$ for all n. Then $\delta_n(k) \to \delta(k)$ for all $k \neq 0$ uniformly in any open interval of the k axis which excludes $k = 0$. The radial wavefunctions also converge pointwise and uniformly.

Proof: We follow the method used in proving Theorem 2. One derives from Eq. (3) and from the reasoning which proved Theorem 2 that

$$\begin{aligned}\sin\delta_m(k) - \sin\delta_n(k) &= -(1/k)\int_0^\infty dr\,\sin kr\cdot[V_m(r) - V_n(r)]\tilde{u}_m(r) \\ &\quad + (1/k)\int_0^\infty dr\,\sin kr\cdot V_n(r)[\tilde{u}_n(r) - \tilde{u}_m(r)],\end{aligned}$$

which implies the inequality

$$\begin{aligned}|\sin\delta_m(k) - \sin\delta_n(k)| &\le (C/k)\int_0^\infty dr\,|V_m(r) - V_n(r)| \\ &\quad + (1/k)\int_0^\infty dr\,|V_n(r)|\,|u_n(r) - \tilde{u}_m(r)|,\end{aligned} \tag{23}$$

where $C \equiv \exp(M/k)$. From the equation for the radial wavefunction,[4] one readily derives the inequality

$$\begin{aligned}|\tilde{u}_m(r) - \tilde{u}_n(r)| &\le W_{mn} \\ &\quad + (1/k)\int_0^r dr'\,|V_n(r')|\,|\tilde{u}_m(r') - u_n(r')|,\end{aligned} \tag{24}$$

where

$$W_{mn} \equiv (C/k)\int_0^\infty dr'\,|V_m(r') - V_n(r')|. \tag{25}$$

We also derive easily that iteration of the inequality, Eq. (24), leads to

$$\begin{aligned}|\tilde{u}_m(r) - u_n(r)| &\le W_{mm}\exp(1/k)\int_0^\infty dr'\,|V_n(r')| \le CW_{mn},\end{aligned} \tag{26}$$

and find easily that the sequence $\sin\delta_n(k)$ $(n = 1, 2, \ldots)$ for $k \neq 0$ is a Cauchy sequence which therefore converges to a limit. For physical purposes the convergence of $\sin\delta_n$ is sufficient. However the continuity of $\delta_n(k)$ together with $\delta_n(\infty) = 0$ implies that the $\delta_n(k)$ also converge. The point $k = 0$ may be an exceptional point as illustrated by the example of a sequence $\{V_n(r)\}$ of the form $V_n(r) \equiv g_n V(r)$, where the sequence $\{g_n\}$ may converge to a value g such that $gV(r)$ has a different number of bound states than all the $V_n(r)$. The pointwise and uniform (in r) convergence of the radial wavefunction $\tilde{u}(r)$ is a direct consequence of Eq. (26).

The next theorem is an analog of Theorem 2. It can be proven by the method employed in the proof of Theorem 2 modified along the lines in which the proof of Theorem 6 modified the proof of Theorem 1.

Theorem 7: Let $\{V_n(r)\}$ be a sequence of potentials such that

$$\int_0^\infty dr\,\frac{r|V_m(r) - V_n(r)|}{1 + \beta r} < \epsilon \tag{27}$$

for m, n larger than some $N(\epsilon)$. Then the sequence $\{\delta_n(k)\}$ converges for any $k \neq 0$ to a function $\delta(k)$ continuous in any open interval of the k axis which excludes $k = 0$. The radial wavefunction converges pointwise and uniformly.

The following theorem is the analog of Theorem 3.

Theorem 8: Let $\{V_n(r)\}$ be a sequence of potentials of uniformly bounded $L^{(1)}$ norm which converges to an L^1 potential $V(r)$ with finite $L^{(1)}$ norm, in the sense that

$$\lim_{n\to\infty}\int_0^\infty dr r|V_n(r) - V(r)| = 0. \tag{28}$$

Then the corresponding sequence of phase shifts $\{\delta_n(k)\}$ converges pointwise for $k \neq 0$ to $\delta(k)$, the phase shift function for $V(r)$. If also $f(0) \neq 0$, where $f(k)$ is the Jost function of $V(r)$, then the corresponding sequence of phase shifts $\{\delta_n(k)\}$ converges pointwise uniformly to $\delta(k)$, the phase shift function for $V(r)$. The corresponding wavefunctions also converge pointwise and uniformly in r.

N. B. The condition $f(0) \neq 0$ implies that $V(r)$ does not have a new bound state appearing at zero energy. Note further that this theorem requires not merely the existence of a Cauchy sequence in the $L^{(1)}$ norm, but the existence of a limiting potential in this norm.

Proof: Let χ be the common upper bound to the $L^{(1)}$ norms of $V(r)$ and the potentials $V_j(r)(j = 1, 2, \ldots)$. Then the Jost function $f_j(k)$ of each $V_j(r)$ exists as an entire function of g (a common coupling constant factor attached to all the potentials).

We show that the convergence of the $L^{(1)}$ norms χ_n implies uniform convergence of the Jost functions $f_j(k)$ to $f(k)$. In fact the existence of a limiting function could be inferred if only a Cauchy sequence in the $L^{(1)}$ norm were assumed, but it need not have the properties of a Jost function. For the sake of greater generality, we shall demonstrate the convergence of the Jost solutions $f_j(k, r)$ to the Jost solution $f(k, r)$. We introduce the notations

$$\begin{aligned}&I_a^b[x, y; V]_m \\ &\equiv \frac{e^{-ik(x-y)}}{k^{m+1}}\int_a^b dr_1\int_a^{r_1} dr_2\cdots\int_a^{r_{m-1}} dr_m\,V(r_1)\cdots\end{aligned}$$

$$V(r_m)\sin k(x-r_1)\sin k(r_1-r_2)\cdots \times \sin k(r_{m-1}-r_m)\sin k(r_m-y), \tag{29a}$$

$$I_a^b[y;V]_m = \frac{1}{k^m}\int_a^b dr_1\int_a^{r_1}dr_2\cdots\int_a^{r_{m-1}}dr_m V(r_1)\cdots \times V(r_m)e^{-ik(r_1-y)}\sin k(r_1-r_2)\cdots \times \sin k(r_{m-1}-r_m)\sin k(r_m-y). \tag{29b}$$

The coefficient $\phi_n(k,r)$ of g^n in the series for $f(k,r)$ corresponding to a potential $V(r)$ is

$$\phi_n(k,r) = I_r^\infty[0;V]_n. \tag{30}$$

We find for two potentials $V_1(r)$ and $V_2(r)$

$$\phi_n^{(1)}(k,r)-\phi_n^{(2)}(k,r) = I_r^\infty[0;V_1]_n - I_r^\infty[0;V_2]_n = \sum_{s=1}^{n}\int_r^\infty dr_s[V_1(r_s)-V_2(r_s)]I_{r_s}^\infty[r_s,V_1]_{s-1} \times I_r^{r_s}[r_s,0;V_2]_{n-s}, \tag{31}$$

where an empty factor is interpreted as unity. In the present case the following inequalities are valid;

$$I_a^b[y;V]_m \le (1/m!)[\int_a^b dr r|V(r)|]^m \tag{32a}$$

$$I_a^b[x,y;V]_m \le (x/m!)[\int_a^b dr r|V(r)|]^m, \tag{32b}$$

which result in

$$|\phi_n^{(1)}(k,r)-\phi_n^{(2)}(k,r)| \le \frac{(2\chi)^{n-1}}{(n-1)!}\int_0^\infty dr r|V_1(r)-V_2(r)|, \tag{33}$$

where χ is the universal $L^{(1)}$ bound to the $V^{(n)}(r)$ and $V(r)$. Consequently,

$$|f_i(k,r)-f_j(k,r)| \le |g|\exp(2|g|\chi)\int_0^\infty dr r|V_i(r)-V_j(r)|, \tag{34}$$

which shows the sequence $\{f_j(k,r)\}$ to be a Cauchy sequence. We can conclude from this, the uniform pointwise convergence of the Jost solutions to $f(k,r)$ as well as of the Jost function to $f(k)$. The phase shift $\delta(k) = Im\ \ln f(k)$ likewise converges if $f(k) \neq 0$. We have assumed that the $\{V_n(r)\}$ actually converge in the norm to an $L^{(1)}$ potential $V(r)$, so that we could conclude that $f(k)$ is itself a Jost function and therefore[16] $f(k) \neq 0$. If furthermore $f(0) \neq 0$, then the convergence is uniform for all k.

The uniform convergence of the wavefunction is a consequence of the uniform convergence of the Jost solutions $f(k,r)$ and $f(-k,r)$, for real k and the fact that $f(k) \neq 0$.

We can conclude from the uniform bounds

$$|f_j(k)| \le \exp(\chi|g|) \tag{35}$$

that the limiting Jost function $f(k)$ exists as an analytic function of g. It is not necessarily an analytic function of k for k real since the real axis may only be the boundary of the analyticity domain in k. (The above conditions have sufficed to imply analyticity in the lower half- k plane.) The dispersion relations, Eqs. (11)–(14), are valid for the boundary values of an analytic function.

It is not possible to obtain results to the effect that the covergence of phase shifts and bound state energies entails convergence of potentials in some norm. This follows from the results of Gelfand–Levitan theory[17] which imply that a knowledge of the residues of the Jost function at the bound state poles is necessary in addition to their location and the phase shifts at all energies in order to reconstruct the potential.

We conclude with a result on convergence of the $l = 0$ scattering lengths.

Theorem 9: Let $\{V_n(r)\}$ be a sequence of potentials for which the $L^{(1)}$ norms and

$$\gamma_n \equiv \int_0^\infty dr r^2|V_n(r)| \tag{36}$$

are uniformly bounded. Suppose moreover that the $\{V_n\}$ constitute a Cauchy sequence in the first ($L^{(1)}$) and second moment norms, i.e., for $p = 1$ and 2 and arbitrary $\epsilon > 0$,

$$\int_0^\infty dr r^p|V_m(r)-V_n(r)| < \epsilon \tag{37}$$

can be satisfied for all m, n sufficiently large, and that the corresponding Jost functions $f_n(0)$ do not converge to zero. Then the sequence of scattering lengths A_n is a convergent one.

Proof: We imagine a common coupling constant parameter g attached to each of the potentials $V_m(r)$. We may eventually set $g = 1$. We proceed from the expression for the S-wave scattering length[18]

$$-A(g) = \frac{\sum_{n=1}^\infty g^n\int_0^\infty dr_n\cdots\int_{r_2}^\infty dr_1 r_1(r_1-r_2)\cdots(r_{n-1}-r_n)r_n V(r_1)\cdots V(r_n)}{\sum_{n=0}^\infty g^n\int_0^\infty dr_n\cdots\int_{r_2}^\infty dr_1(r_1-r_2)\cdots(r_{n-1}-r_n)r_n V(r_1)\cdots V(r_n)} \equiv \frac{N(g)}{D(g)}. \tag{38}$$

$D(g)$ is merely the zero-energy Jost function $f(0)$ considered as a function of g. $N(g)$ and $D(g)$ are each entire functions of g if $V(r)$ has either bounded $L^{(1)}$ or $\overline{L}^{(1/2)}$ norm. We write

$$N(g) \equiv \sum_{n=1}^{\infty} g^n N_n, \qquad D(g) \equiv \sum_{n=0}^{\infty} g^n D_n.$$

The coefficient D_n is identical to $\phi_n(0, 0)$ of Eq. (30). The inequality, Eq. (33), for $|D_n^{(1)} - D_n^{(2)}|$ follows as for Theorem 8, and this implies the convergence of the $D_m(g)$ for each g. From a completely parallel argument utilizing the conditions of the present theorem, one finds similarly for the N_n

$$|N_n^{(1)} - N_n^{(2)}| \leq \frac{(2\chi)^{n-2}}{(n-2)!}\left[\gamma \int_0^\infty dr r |V_1 - V_2| + \frac{\chi}{2^{n-2}(n-1)} \int_0^\infty dr r^2 |V_1 - V_2|\right] \quad (39)$$

where χ and γ are, respectively, uniform bounds of the first and second moments. One finds that

$$|N_i(g) - N_j(g)| \leq g^2 \gamma e^{2g\chi} \int_0^\infty dr r |V_i - V_j| + g e^{g\chi} \int_0^\infty dr r^2 |V_i - V_j|, \quad (40)$$

which implies the convergence of the $N_m(g)$ for each g. If the $D_m(g)$, i.e., the $f_m(0)$, do not approach zero, then the $A_m(g)$ clearly approach a limit. The limiting functions $N(g)$ and $D(g)$ are analytic for all finite g so that $A(g)$ is meromorphic in g. We readily verify for $V(r) \geq 0$ that

$$N_n \equiv \int_0^\infty dr_n \cdots \int_{r_2}^\infty dr_1 r_1 (r_1 - r_2) \cdots r_n V(r_1) \cdots V(r_n)$$
$$\leq \int_0^\infty dr_1 r_1^2 V(r_1) \int_0^\infty dr_n \cdots \times \int_{r_3}^\infty dr_2 (r_2 - r_3) \cdots r_n V(r_2) \cdots V(r_n) = \gamma D_{n-1},$$

which immediately implies the useful inequality

$$A(g) \leq g\gamma. \quad (41)$$

This shows in particular that a bounded $L^{(1)}$ norm and the convergence of the γ_n to zero imply a vanishing scattering length.

IV. DISCUSSION

An immediate application of the above results is to the clarification of limiting procedures in potential theory[18a] some of which have led to paradoxical results or to specious reasoning. The emphasis would not be misplaced in calling the attention of physicists to the pitfalls in what seems a most plausible type of limiting procedure. Such caution is especially in order in elementary particle formalisms where regularizations and limiting procedures abound, though the present type of analysis is likely to prove not so feasible in these cases.

It seems difficult to offer any direct physical understanding of the conclusion that the norms $\int_0^\infty dr r|V(r)|$ and $\int_0^\infty dr |V(r)|^{1/2}$ are appropriate to the correct formation of limiting procedures beyond the observation that they are both dimensionless quantities (in the units $\hbar^2/2m = 1$) and therefore appropriate to measure the number of bound states. It should be noted that an infinite set of dimensionless quantities is available of the form

$$\Omega_q \equiv \int_0^\infty dr r^{2q-1} |V(r)|^q, \quad (42)$$

which, apart from $q = \frac{1}{2}$, $q = 1$, do not seem to play a special role. The inequality

$$\Omega_1 \leq \Omega_q (\Omega_{1/2})^{2(q-1)}$$

derivable from the Holder inequality[19] indicates the lack of independence of these norms. Since the role of the $L^{(1/2)}$ norm $\int_0^\infty dr |V(r)|^{1/2}$ seems to have been recently discovered,[12,19a] this paper has made a special effort to sponsor the role of this type of norm for which the demonstrated results have been generally weaker than for the more familiar $L^{(1)}$ norm. Time will show whether better results may not hold.

From the over-all point of view, the results of this article point to the concept of potentials as elements of a Banach space, with the phase shift as a continuous and bounded functional of the potential almost everywhere (with continuity failing to hold in the neighborhood of a potential which has a new bound state emerging at zero energy).

The condition of completeness requisite for a Banach space has not been discussed in this paper. One knows that[20] all metric space can be completed in unique fashion. Alternative topologies to norm topologies are also possible, such as weak topologies which would allow a wider space of potentials including in particular, δ functions. Within the framework of a Banach space, a variety of new questions and concepts can be dealt with: Bases for the space, everywhere dense subsets, compactness, etc. A possible projection of the usefulness of the new abstract structure may lie in the device of proving general properties by proving them for the elements of an everywhere dense set. Certain additive inequalities of the phase shift functional have been proven and will be presented in a subsequent paper.

It is clear that all that has been said of the phase shift as a continuous bounded functional applies equally well to $\ln|f(k)|$. This suggests generalization in the direction of determining other or more general continuous bounded functionals in the potential space.

In conclusion, one might enumerate the directions into which the present work may be extended. One might consider the continuity question for classical scattering, for one-dimensional potential problems, for non local and energy-dependent potentials, as well as for the total cross section in three dimensions.

ACKNOWLEDGMENTS

The author would like to acknowledge helpful conversations with J. M. Cook, R. G. Newton, and C. van Winter, and the cooperation and assistance of David

McLaughlin with some of the calculations. He would also like to acknowledge Mrs. Bernice Ruppalt for her help in preparing the manuscript and her patience with all the revisions it has undergone.

APPENDIX A: THE DISPLACEMENT LEMMA

A fundamental result upon which we base subsequent proofs of the continuity of scattering functions in cases of singular potentials is the following lemma. This lemma allows us to apply the ideas of the present paper to a sequence of potentials considered in the displaced interval (a,∞) $(a>0)$.

Displacement Lemma: Let $\{V_n(r)\}$ be a sequence of potentials of uniformly bounded $L_+^{(1)}$ norm in the interval (a,∞) $(a>0)$ and let them form a Cauchy sequence in the $L_+^{(1)}$ norm over this interval. Suppose the sequence $\{\delta_n(a,k)\}$ of values of the respective variable phase functions at $r=a$ for energy $k^2(\neq 0)$ converges to a finite limit $\delta(a,k)$ which satisfies $\delta(a,k)>-ka$. Then the sequence of phase shifts $\{\delta_n(k)\}$, as well as the sequence of radial wavefunctions, converges. If the sequence $\{V_n(r)\}$ converges pointwise to a potential $V(r)$ which is of bounded $L_+^{(1)}$ norm in (a,∞), then the limiting values correspond to this potential.

N.B. The value of a may be chosen to depend on k provided $a>0$ for each $k>0$.

Proof: We start by integrating from a to r $(r>a)$ the variable phase equation for a potential $V_n(r)$ of the sequence (we work with S waves for simplicity)

$$\delta_n(r,k) = \delta_n(a,k) - (1/k)\int_a^r ds\,V_n(s)\,\sin^2[ks+\delta_n(s,k)]. \tag{A1}$$

By hypothesis, the sequence $\{\delta_n(a,k)\}$ converges to a limit, say $\delta(a,k)$. We consider the quantity $\hat{\delta}_n(r,k)$ which solves the integral equation

$$\hat{\delta}_n(r,k) = \delta(a,k) - (1/k)\int_a^r ds\,V_n(s)\,\sin^2[ks+\hat{\delta}_n(s,k)]. \tag{A2}$$

Since $\delta(a,k)$ obeys the inequality $-ka<\delta(a,k)<\infty$, one can find a square well potential $S(r)$ of range a and finite depth (depending possibly on k) such that its variable phase function at $r=a$ for energy k^2 has the value $\delta(a,k)$. Then $\hat{\delta}_n(r,k)$ is merely the variable phase at energy k^2 due to the potential

$$Y_n(r) \equiv \begin{cases} S(r) & \text{for } r<a \\ V_n(r) & \text{for } r>a. \end{cases} \tag{A3}$$

The sequence of potentials $\{Y_n(r)\}$ is a Cauchy sequence in the ω_β norm and the $Y_n(r)$ have uniformly bounded ω_β norm from the hypothesis of the theorem and the boundedness of $S(r)$. We therefore conclude easily from Theorem 7 that the sequence of variable phase functions $\{\hat{\delta}_n(r,k)\}$ and correspondingly the sequence of radial wavefunctions $\{u_n(r,k)\}$ converges for any r in the interval (a,∞). The result follows for any $k>0$. It also follows that, if the sequence $\{V_n(r)\}$ in the interval (a,∞) converges pointwise to $V(r)$ of bounded $L_+^{(1)}$ norm, the $\hat{\delta}_n(r,k)$ converge to the corresponding quantities (at energy k^2) for the limiting potential

$$Y(r) = \begin{cases} S(r) & r<a \\ V(r) & r>a \end{cases} \tag{A4}$$

and therefore, of course, converge in the interval (a,∞) to the corresponding quantities for the limiting potential $V(r)$.

We now show that

$$\lim_{n\to\infty}[\delta_n(r,k)-\hat{\delta}_n(r,k)] = 0. \tag{A5}$$

Subtracting Eq. (A2) from Eq. (A1), we find

$$\begin{aligned}\delta_n(r,k)-\hat{\delta}_n(r,k) &= \delta_n(a,k)-\delta(a,k) \\ &+ (1/k)\int_a^r ds V_n(s)\,\sin[2ks+\delta_n(s,k) \\ &+ \hat{\delta}_n(s,k)]\,\sin[\hat{\delta}_n(s,k)-\delta_n(s,k)]\end{aligned} \tag{A6}$$

which implies the inequality

$$\begin{aligned}|\delta_n(r,k)-\hat{\delta}_n(r,k)| &\le |\delta_n(a,k)-\delta(a,k)| \\ &+ (1/k)\int_a^r ds\,|V_n(s)|\,|\delta_n(s,k)-\hat{\delta}_n(s,k)|\end{aligned} \tag{A7}$$

One infers by Titchmarsh's lemma[21] that

$$\begin{aligned}&|\delta_n(r,k)-\hat{\delta}_n(r,k)| \\ &\le |\delta_n(a,k)-\delta(a,k)|\exp[(1/k)\int_a^\infty ds\,|V_n(s)|]\end{aligned} \tag{A8}$$

From the uniform boundedness of the $L_+^{(1)}$ norms in the interval (a,∞) and the convergence of the $\delta_n(a,k)$ to $\delta(a,k)$, one infers that the sequences $\{\delta_n(r,k)\}$ and $\{\hat{\delta}_n(r,k)\}$ converge to the same quantity. The convergence of the variable phase function sequence $\{\delta_n(r,k)\}$ implies, as stated earlier, the convergence of the sequence of wavefunctions in (a,∞) for the potential sequence $Y_n(r)$ of Eq. (A3). This sequence of wavefunctions is identical with the sequence of wavefunctions corresponding to $\{V_n(r)\}$ in (a,∞) with the correct function and slope at $r=a$. This follows because the variable phase function at a point determines both wavefunction and slope at the point. Thus the lemma is proven.

APPENDIX B: CONVERGENCE THEOREM FOR A CLASS OF REPULSIVE SINGULAR POTENTIALS

In this appendix, we extend the results of the present article to cases where the boundedness of any of the norms $\omega_\beta[V]$, $\chi[V]$, or $\tau[V]$ is violated because of the behavior of the potential in the neighborhood of $r=0$. This situation arises for potential sequences which converge to a repulsive singular potential. We restrict the present considerations to the convergence of the S-wave phase

shift and wavefunction for potential sequences which approach the limiting potential monotonically within the fixed open interval $(0, c)$. A more general theorem without the restriction of a fixed interval will be treated in a subsequent article.

We now prove the following theorem which relies heavily on the displacement lemma presented in Appendix A.

Theorem: Let $\{V_n(r)\}$ be a sequence of potentials with the following properties:

(a) For each $r > 0$, $|V_n(r)| < \infty$ for each n.

(b) For each $r > 0$, the sequence $\{V_n(r)\}$ converges pointwise to a potential $V_\infty(r) \equiv V(r)$.

(c) For each $r > 0$, $|V(r)| < \infty$.

(d) The $L^{(1)}$ norms of the attractive parts are uniformly bounded in n; i.e., $\int_0^\infty dr r|V_n^-(r)| < M < \infty$.

(e) The norms $(\beta > 0)$

$$\omega_\beta[V_n]_c^\infty \equiv \int_c^\infty \frac{dr r|V_n(r)|}{1+\beta r} \tag{B1}$$

are uniformly bounded in n, including $n = \infty$, with c some fixed positive number.

(f) The potentials $V_n(r)$ form a Cauchy sequence in the $L_\beta^{(1)}$ norm over the interval (c, ∞); i.e., for any $\epsilon > 0$ one can make

$$\omega_\beta[V_m - V_n]_c^\infty \equiv \int_c^\infty \frac{dr r|V_m(r) - V_n(r)|}{1+\beta r} < \epsilon \tag{B2}$$

by choosing m, n larger than some $N(\epsilon)$.

(g) For any fixed $r < c$, the sequence of numbers $\{V_n(r)\}$ is nondecreasing.

Then the limit

$$\delta(k) = \lim_{n\to\infty} \delta_n(k)$$

exists for all real $k = 0$. If in (e) and (f) one allows $\beta = 0$, then the limit also exists for $k = 0$.

Remark: Nothing has been said concerning the quantities $V_n(0)$ and $V(0)$ which may be finite or infinite (positively or negatively).

Proof: In view of conditions (e) and (f) and the displacement lemma, it suffices to demonstrate that the "boundary values" $\delta_n(a, k)$ converge for each fixed $k \neq 0$ with $0 < a \leq c$ and that $\delta_n(a, k) > -k(a - \eta)$ for some (possible k dependent) value of $\eta > 0$ uniformly in n. While the point $r = c$ can be used as the reference point for application of the displacement lemma, we shall more generally choose $r = a$ with $0 < a \leq c$. We therefore consider the sequence of truncated potentials $\{V_n(r, a)\}$, where

$$V_n(r, a) \equiv V_n(r)\theta(a - r). \tag{B3}$$

$\delta_n(a, k)$ is clearly determined by the truncated potential. We denote the interval $0 \leq r \leq c$ by C. Since the $V_n(r, a)$ are nondecreasing with n for each r in C, one concludes from the comparison lemma in I that, for each k, the corresponding phase shift sequence $\{\delta_n(a, k)\}$ is monotonically nonincreasing. If the sequence is bounded from below, the limit must exist. Such a lower bound must exist, since the potentials can be bounded pointwise from above by the potential

$$X(r) = \max_n V_n(r; c), \tag{B4}$$

which is finite for all $r > 0$ due to conditions (a), (b), and (c). The phase shift $\delta_x(a, k)$ for $X(r)$ clearly obeys $\delta_x(a, k) > -k(a - \eta)$ for some η, since $\delta(a, k) = -ka$ for the hard core potential of radius a. We conclude from the monotonicity and uniform boundedness in n of $\delta_n(a, k)$ that a limit $\delta(a, k)$ exists satisfying $\delta(a, k) > -ka$. The sequence $\{V_n(r)\}$ is uniformly bounded in n and pointwise convergent for each $r > 0$ by conditions (a), (b), and (c). From this one easily concludes that conditions (e) and (f) are valid with c replaced by any $a > 0$. We now conclude from the displacement lemma that the sequence $\{\delta_n(k)\}$ is convergent to the phase shift of the limiting potential in $(0, \infty)$.

The displacement lemma also specifies the convergence of the corresponding wavefunctions to those of the limiting potential in the interval $a \leq r < \infty$. We can in fact conclude the convergence of the limiting wavefunction for any $r > 0$ since a is arbitrary. We can easily see that the limiting wavefunction vanishes at $r = 0$, so that the limit yields the regular solution. We can write from Eq. (I 10b) that

$$-ka \leq \delta_n(a, k) \leq \chi[\int_0^a dr r|V_n^-(r)|]. \tag{B5}$$

For fixed n, both bounds in Eq. (B5) shrink to zero as $a \to 0$. Since the $\{V_n(r)\}$ are nondecreasing with n for each $0 < r < c$, the quantity $\int_0^a dr r|V_n^-(r)|$ is nonincreasing with n. One therefore concludes that the limiting $\delta(a, k)$ vanishes as $a \to 0$. Since the radial wavefunction can be uniquely constructed from the variable phase function $\delta(r, k)$ by means of the relation[2]

$$u(r, k) = \exp\{(1/2k)\int_r^\infty ds V(s) \sin 2[ks + \delta(s, k)]\} \times [\sin kr + \delta(r, k)], \tag{B6}$$

we conclude that $u(0, k)$ vanishes. In Eq. (B6) the radial wavefunction is normalized for large r by

$$u(r, k) \sim \sin[kr + \delta(k)]. \tag{B7}$$

---

[1] W. M. Frank, J. Math. Phys. 12, 2348 (1971), referred to as I; W. M. Frank and D. W. McLaughlin, J. Math. Phys. 12, 2525 (1971) (preceeding paper), referred to as II. Various classes of potentials, e.g., L^1, $L^{(1)}$, $L^{(\frac{1}{2})}$, $L^{(1/2)}$ and $\tilde{L}^{(1/2)}$, which are used in the present article, are defined in I

[2] The theory for the existence of solutions of partial wave scat-

tering equation for continuous potentials $V(r)$ is a special case of the existence theorems in E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1944), Chap. III. See also E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), Chap. 3. The requirements of L^1 potentials on $[0,\infty]$ is indicated by the specific explicit expression for the Jost solution which exists for $k \neq 0$ if $V(r)$ is L^1 and exists for $k = 0$ if $V(r)$ is L^1.

3 See, e.g., P. Roman, *Advanced Quantum Theory* (Addison-Wesley, Reading, Mass. 1965), Sec. 3-2b, Eq. (3-52).

4 R. G. Newton, J. Math. Phys. 1, 319 (1960), Eq. (3.7), in a different notation with different normalization.

5 See V. Bargmann, Proc. Natl. Acad. Sci. **38**, 961 (1952). J. Schwinger, *ibid.* **47**, 122 (1961), for a bound on the number of bound states in a given potential, $n_B \leq \chi[V]$. Recall also the fact that the zero-energy phase shift is related to the number of bound states.

6 For all potentials $\chi < \infty$ implies the existence of the Jost function. If the Jost function exists and either one of $\chi_\pm = \int dr\, |V^{\pm}(r)| < \infty$ is finite, then χ is finite. Conceivably for potentials of the form $\sin r^m$ with $m > 2$, the Jost function might exist while $\chi = \infty$.

7 The Jost function is identical with the Fredholm determinant of the partial wave scattering equation. See Newton, Ref. 4. It is a standard result that the Fredholm determinant is entire in the coupling parameter; see, e.g., W. V. Lovitt, *Linear Integral Equations* (Dover, New York, 1950), Chap. III.

8 See Newton, Ref. 4, discussion following Eq. (4.15).

9 R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), Eq. (2.17).

10 R. G. Newton, Ref. 4, end of Sec. 5

11 The notation $f_n(k)$ refers to Jost solution corresponding to the potential V_n. The $f_n(k)$ are to be distinguished from the quantities $\phi_m(k)$ which are the coefficients of the expansion in g of $f(k)$.

12 W. M. Frank, J. Math. Phys. **8**, 466 (1967); **9**, 1890 (1968). The results in this reference were based on a representation of $|V(r)|^{1/2}$ as a Laplace-Stieltjes transform. The representation of $|V(r)^{1/2}$ as a Laplace-Stieltjes transform is not a real limitation since the space of L-S transforms is everywhere dense over the space of functions with finite first moment. I am indebted to Dr. H. Furstenberg for his observation. The constant $2/\sqrt{\pi}$ in the expression for $\tilde{\tau}$ following Eq. (15) in the last reference can be dispensed with. I am grateful to Dr. D. J. Land for this result.

13 See R. G. Newton, Ref. 4, Eq. (4.15).

14 E. C. Titchmarsh, *The Theory of Functions* (Oxford U. P., London, 1939), 2nd ed, 5.21.

15 R. G. Newton, Ref. 9, Chap. **12**, pp. 337–38.

16 R. G. Newton, Ref. 9, p. 346

17 R. G. Newton, Ref. 4, Sec. 8.

18 This expression for the scattering length can be obtained from the power series expansion for the Jost function $f(k)$ and the relation $A = -\lim_{k\to 0}[f(-k) - f(k)]/2ikf(k)$. The expression for the Jost function is given in Eq. (10).

18a F. Calogero, Phys. Rev. **139**, B602 (1965); E. M. Ferreia and F. F. Teixeira, J. Math. Phys. **7**, 1207 (1966).

19 The Holder inequality which appears in any book on functional analysis has the form for discrete sums $(1/p + 1/q = 1)$

$$|\sum a_i b_i| < [\sum (a_i)^p]^{1/p} [\sum (b_i)^q]^{1/q}.$$

19a D. Masson, J. Math. Phys. **8**, 2308 (1967); F. Calogero, Commun. Math. Phys. **1**, 80 (1965).

20 Kolmogorov and Fomin, *Functional Analysis* (Graylock Press, Rochester, N.Y., 1957), Translation of 1954 Russian Edition by L. F. Boron, Vol. I, Chap. II, Theorem 2, p. 40.

21 E. C. Titchmarsh, *Eigenfunction Expansions*, Part I (Clarendon Oxford, 1962), 2nd ed., p. 115, Lemma 5.6.

22 F. Calogero, *Variable Phase Approach to Potential Scattering* (Academic, New York, 1967), Chap. 6, Eqs. 1–10.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Inverse Decay Problem*

L. P. Horwitz, J. A. LaVita, and J.-P. Marchand
Department of Mathematics, University of Denver, Denver, Colorado 80210
(Received 17 May 1971)

Let $U_{\mathcal{K}}(t)$ be a one-parameter operator family of positive type in a Hilbert space $\mathcal{K}$ and $U(t)$ its minimal unitary dilation with infinitesimal generator H. If $U_{\mathcal{K}}(t)$ is a contractive semigroup, then H is not positive. If in addition $U_{\mathcal{K}}(t) \to 0$ for $t \to \infty$, then there exists a state $\phi \in \mathcal{K}$ on which H is not defined. We interpret these and other results in the context of the quantum-mechanical theory of unstable particles and the scattering theory of Lax and Phillips.

1. INTRODUCTION

We assume within the framework of quantum mechanics that unstable particles are represented by a definite subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$ and that their time evolution $U_{\mathcal{K}}(t)$ is therefore the contraction[1] of the unitary evolution $U(t) = e^{-iHt}$ to the subspace $\mathcal{K}$. *The inverse decay problem* then consists in reconstructing a pair $\{\mathcal{H}, U(t)\}$ provided $U_{\mathcal{K}}(t)$ is given.

Such an embedding of an evolution $U_{\mathcal{K}}(t)$ into an extended space $\mathcal{H}$ is possible if and only if $U_{\mathcal{K}}(t)$ is an operator family of positive type which is contractive, i.e., for which $U_{\mathcal{K}}(0) = I_{\mathcal{K}}$ and $U_{\mathcal{K}}(-t) = U_{\mathcal{K}}^{\dagger}(t)$. Furthermore the reconstruction is unique if $\mathcal{K}$ is supposed to generate $\mathcal{H}$ under $U(t)$, i.e., if

$$\overline{\bigcup_{t\geq 0} U(t)\mathcal{K}} = \mathcal{H}.$$

Since we are interested mainly in the phenomenon of decay, we assume that $U_{\mathcal{K}}(t)$ is strongly contractive:

$$(1) \quad \mathrm{s} - \lim_{t\to\infty} U_{\mathcal{K}}(t) = 0.$$

Together with (1) we also consider the impact of the following conditions which may or may not hold in decay:

(2a) $U_{\mathcal{K}}(t)$ is a semigroup.

(2b) $\mathcal{K}$ is finite-dimensional.

(2c) the states in $\mathcal{K}$ have finite total energy.

(2d) H is defined everywhere in $\mathcal{K}$.

The reconstruction of $\{\mathcal{H}, U(t)\}$ from $\{\mathcal{K}, U_{\mathcal{K}}(t)\}$ is based on the well-known and highly developed theory of extension of Hilbert spaces[1]; but it is remarkable that under Assumptions (1) and (or) (2) many striking results of physical interest can

tering equation for continuous potentials $V(r)$ is a special case of the existence theorems in E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1944), Chap. III. See also E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), Chap. 3. The requirements of L^1 potentials on $[0,\infty]$ is indicated by the specific explicit expression for the Jost solution which exists for $k \neq 0$ if $V(r)$ is L^1 and exists for $k = 0$ if $V(r)$ is L^1.

3 See, e.g., P. Roman, *Advanced Quantum Theory* (Addison-Wesley, Reading, Mass. 1965), Sec. 3-2b, Eq. (3-52).

4 R. G. Newton, J. Math. Phys. 1, 319 (1960), Eq. (3.7), in a different notation with different normalization.

5 See V. Bargmann, Proc. Natl. Acad. Sci. **38**, 961 (1952). J. Schwinger, *ibid.* **47**, 122 (1961), for a bound on the number of bound states in a given potential, $n_B \leq \chi[V]$. Recall also the fact that the zero-energy phase shift is related to the number of bound states.

6 For all potentials $\chi < \infty$ implies the existence of the Jost function. If the Jost function exists and either one of $\chi_\pm = \int dr\, |V^{\pm}(r)| < \infty$ is finite, then χ is finite. Conceivably for potentials of the form $\sin r^m$ with $m > 2$, the Jost function might exist while $\chi = \infty$.

7 The Jost function is identical with the Fredholm determinant of the partial wave scattering equation. See Newton, Ref. 4. It is a standard result that the Fredholm determinant is entire in the coupling parameter; see, e.g., W. V. Lovitt, *Linear Integral Equations* (Dover, New York, 1950), Chap. III.

8 See Newton, Ref. 4, discussion following Eq. (4.15).

9 R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), Eq. (2.17).

10 R. G. Newton, Ref. 4, end of Sec. 5

11 The notation $f_n(k)$ refers to Jost solution corresponding to the potential V_n. The $f_n(k)$ are to be distinguished from the quantities $\phi_m(k)$ which are the coefficients of the expansion in g of $f(k)$.

12 W. M. Frank, J. Math. Phys. **8**, 466 (1967); **9**, 1890 (1968). The results in this reference were based on a representation of $|V(r)|^{1/2}$ as a Laplace-Stieltjes transform. The representation of $|V(r)^{1/2}$ as a Laplace-Stieltjes transform is not a real limitation since the space of L-S transforms is everywhere dense over the space of functions with finite first moment. I am indebted to Dr. H. Furstenberg for his observation. The constant $2/\sqrt{\pi}$ in the expression for $\tilde{\tau}$ following Eq. (15) in the last reference can be dispensed with. I am grateful to Dr. D. J. Land for this result.

13 See R. G. Newton, Ref. 4, Eq. (4.15).

14 E. C. Titchmarsh, *The Theory of Functions* (Oxford U. P., London, 1939), 2nd ed, 5.21.

15 R. G. Newton, Ref. 9, Chap. **12**, pp. 337–38.

16 R. G. Newton, Ref. 9, p. 346

17 R. G. Newton, Ref. 4, Sec. 8.

18 This expression for the scattering length can be obtained from the power series expansion for the Jost function $f(k)$ and the relation $A = -\lim_{k\to 0}[f(-k) - f(k)]/2ikf(k)$. The expression for the Jost function is given in Eq. (10).

18a F. Calogero, Phys. Rev. **139**, B602 (1965); E. M. Ferreia and F. F. Teixeira, J. Math. Phys. **7**, 1207 (1966).

19 The Holder inequality which appears in any book on functional analysis has the form for discrete sums $(1/p + 1/q = 1)$

$$|\sum a_i b_i| < [\sum (a_i)^p]^{1/p} [\sum (b_i)^q]^{1/q}.$$

19a D. Masson, J. Math. Phys. **8**, 2308 (1967); F. Calogero, Commun. Math. Phys. **1**, 80 (1965).

20 Kolmogorov and Fomin, *Functional Analysis* (Graylock Press, Rochester, N.Y., 1957), Translation of 1954 Russian Edition by L. F. Boron, Vol. I, Chap. II, Theorem 2, p. 40.

21 E. C. Titchmarsh, *Eigenfunction Expansions*, Part I (Clarendon Oxford, 1962), 2nd ed., p. 115, Lemma 5.6.

22 F. Calogero, *Variable Phase Approach to Potential Scattering* (Academic, New York, 1967), Chap. 6, Eqs. 1–10.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 12, NUMBER 12 DECEMBER 1971

The Inverse Decay Problem*

L. P. Horwitz, J. A. LaVita, and J.-P. Marchand
Department of Mathematics, University of Denver, Denver, Colorado 80210
(Received 17 May 1971)

Let $U_{\mathcal{K}}(t)$ be a one-parameter operator family of positive type in a Hilbert space $\mathcal{K}$ and $U(t)$ its minimal unitary dilation with infinitesimal generator H. If $U_{\mathcal{K}}(t)$ is a contractive semigroup, then H is not positive. If in addition $U_{\mathcal{K}}(t) \to 0$ for $t \to \infty$, then there exists a state $\phi \in \mathcal{K}$ on which H is not defined. We interpret these and other results in the context of the quantum-mechanical theory of unstable particles and the scattering theory of Lax and Phillips.

1. INTRODUCTION

We assume within the framework of quantum mechanics that unstable particles are represented by a definite subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$ and that their time evolution $U_{\mathcal{K}}(t)$ is therefore the contraction[1] of the unitary evolution $U(t) = e^{-iHt}$ to the subspace $\mathcal{K}$. *The inverse decay problem* then consists in reconstructing a pair $\{\mathcal{H}, U(t)\}$ provided $U_{\mathcal{K}}(t)$ is given.

Such an embedding of an evolution $U_{\mathcal{K}}(t)$ into an extended space $\mathcal{H}$ is possible if and only if $U_{\mathcal{K}}(t)$ is an operator family of positive type which is contractive, i.e., for which $U_{\mathcal{K}}(0) = I_{\mathcal{K}}$ and $U_{\mathcal{K}}(-t) = U_{\mathcal{K}}^{\dagger}(t)$. Furthermore the reconstruction is unique if $\mathcal{K}$ is supposed to generate $\mathcal{H}$ under $U(t)$, i.e., if

$$\overline{\bigcup_{t\geq 0} U(t)\mathcal{K}} = \mathcal{H}.$$

Since we are interested mainly in the phenomenon of decay, we assume that $U_{\mathcal{K}}(t)$ is strongly contractive:

$$(1) \quad \mathrm{s} - \lim_{t\to\infty} U_{\mathcal{K}}(t) = 0.$$

Together with (1) we also consider the impact of the following conditions which may or may not hold in decay:

(2a) $U_{\mathcal{K}}(t)$ is a semigroup.

(2b) $\mathcal{K}$ is finite-dimensional.

(2c) the states in $\mathcal{K}$ have finite total energy.

(2d) H is defined everywhere in $\mathcal{K}$.

The reconstruction of $\{\mathcal{H}, U(t)\}$ from $\{\mathcal{K}, U_{\mathcal{K}}(t)\}$ is based on the well-known and highly developed theory of extension of Hilbert spaces[1]; but it is remarkable that under Assumptions (1) and (or) (2) many striking results of physical interest can

be derived by a rather unsophisticated application of this theory.

The mathematical results are derived in Sec. 3. In Sec. 4 we give their physical interpretation. With respect to *quantum mechanics* the main points are as follows:

(1) If $U_{\mathcal{K}}(t)$ has the semigroup property, then the total energy is not positive.

(2) If $\mathcal{K}$ is finite dimensional and $U_{\mathcal{K}}(t)$ is a semigroup which decays to zero, then $\mathcal{K}$ contains states with infinite energy.

(3) In a decay-scattering system[2] the states in $\mathcal{K}$ have finite energy and the evolution $U_{\mathcal{K}}(t)$ decays to zero but does not have the semigroup property.

With respect to the theory of *acoustical scattering*, the same mathematical results lead to a somewhat different physical interpretation, since the generator H of $U(t)$ represents the square root of the energy rather than the energy itself.

2. BASIC MATHEMATICAL TOOLS

Let $\mathcal{K}$ be a separable Hilbert space and $U(t) = e^{-iHt}$ a strongly continuous unitary representation of the additive real line R, H its self-adjoint generator. We call $U(t)$ a *motion*. Let P be a nontrivial *projection in $\mathcal{K}$ with n-dimensional* range $\mathcal{K} = P\mathcal{H}$. The *contracted motion* $U_{\mathcal{K}}(t)$ is the restriction of $P\,U(t)\,P$ to $\mathcal{K}$. $\mathcal{K}$ is *generating under* $U(t)$ if the closed linear space spanned by the vectors $U(t)\phi$ for all $\phi \in \mathcal{K}$ and $t \in R$ is $\mathcal{K}$.

The contracted motion $U_{\mathcal{K}}(t)$ is a strongly continuous operator family in $\mathcal{K}$ of *positive type*, i.e.,

$$\sum_{i,k} (\phi_i, U_{\mathcal{K}}(t_k - t_i)\phi_k)_{\mathcal{K}} \geq 0$$

for all finite sequences $\phi_i \in \mathcal{K}$, $t_i \in R$. Furthermore it satisfies

$$U_{\mathcal{K}}(0) = I(-t) = U_{\mathcal{K}}^{\dagger}(t)$$

and its operator norm is bounded by 1.

Conversely the following holds:

Theorem 1[1]: If $U_{\mathcal{K}}(t)$ is a continuous operator family of positive type in a Hilbert space $\mathcal{K}$, then there exists a triplet $\{\mathcal{H}, U(t), P\}$, where $U(t)$ is a motion in $\mathcal{H}$ and P a projector, such that

$$\mathcal{K} = P\mathcal{H}, \qquad U_{\mathcal{K}}(t) = PU(t)P\big|_{\mathcal{K}}.$$

If $\mathcal{K}$ is required to be generating under $U(t)$, then the dilation[1] $\{\mathcal{H}, U(t)\}$ of $\{\mathcal{K}, U_{\mathcal{K}}(t)\}$ is minimal and essentially unique.

Next we review the concepts we shall need from the theory of semigroups. An operator family $Z(t)$ in $\mathcal{K}$ is a *semigroup*, if $Z(t_1 + t_2) = Z(t_1)Z(t_2)$ for all $t_1, t_2 \geq 0$. A semigroup $Z(t)$ is *contractive* if $Z(0) = I$ and the operator norm of $Z(t)$ is bounded by 1; a contractive semigroup $Z(t)$ is *strongly contractive* if $\text{s-lim}_{t\to\infty} Z(t) = 0$ and *strictly contractive* if the operator norm is strictly less than 1 for some $t > 0$. If $\mathcal{K}$ is finite dimensional, the concepts of strong and strict contractiveness coincide. But in infinite spaces this is not true; in fact, let ϕ_i be an orthonormal basis in $\mathcal{K}$, then the semigroup

$$Z_{ik}(t) = (\phi_i, Z(t)\phi_k) = \delta_{ik}e^{-t/k}$$

is strongly, but not strictly contractive.

If the semigroup $Z(t) = U_{\mathcal{K}}(t)$ is the contraction of a motion, then it is automatically strongly continuous and contractive, and it can be defined for $t < 0$ by $Z(-t) = PU(-t)P = PU^{\dagger}(t)P = (PU(t)P)^{\dagger} = Z^{\dagger}(t)$.

Finally, we remark that a contractive semigroup is of positive type and admits therefore a unitary dilation.

Example 1: Exponential decay: The exponential $Z(t) = e^{-\alpha t}$, $t \geq 0, \alpha > 0$, is a strictly contractive semigroup of operators in a one-dimensional space $\mathcal{K}$. Extending its definition to the negative axis, we obtain the following continuous function of positive type:

$$Z(t) = \begin{cases} e^{-\alpha t} & t \geq 0 \\ e^{\alpha t} & t < 0 \end{cases} = \int_{-\infty}^{\infty} e^{-i\lambda t} \frac{\alpha/\pi}{\lambda^2 + \alpha^2}\, d\lambda.$$

In reconstructing the unitary dilation $U(t)$ of $Z(t)$, we compare this expression with the contracted motion

$$PU(t)P = Pe^{-iHt}P = \int e^{-i\lambda t} d(PE(\lambda)P),$$

where $E(\lambda)$ is the spectral family of H and P the one-dimensional projector $P\mathcal{H} = \mathcal{K}$. From this comparison we see that $PE(\lambda)P$ admits the Radon-Nikodym derivative $(\alpha/\pi)/(\lambda^2 + \alpha^2)$, whose support is the entire line R. Hence the spectrum $\sigma(H)$ of H is R and absolutely continuous (and, of course, simple) (cf. proof of Theorem 2, Sec. 3) in the minimal extension, in which $PE(\Delta)P \neq 0$ for all Borel subsets Δ of R. The dilation $\{\mathcal{H}, U(t)\}$ is therefore realized in the space $L^2(R)$ of Lebesgue square integrable functions by

$$(U(t)\psi)(\lambda) = e^{-i\lambda t}\psi(\lambda)$$

and the subspace $\mathcal{K} = P\mathcal{H}$ is spanned by any normalized square-integrable functions $\phi(\lambda)$ satisfying (almost everywhere)

$$\big|\phi(\lambda)\big|^2 = \frac{\alpha/\pi}{\lambda^2 + \alpha^2},$$

since

$$Z(t) = (\phi, U(t)\phi) = \int e^{-i\lambda t}\,|\phi(\lambda)|^2\, d\lambda = \int e^{-i\lambda t} \frac{\alpha/\pi}{\lambda^2 + \alpha^2}\, d\lambda.$$

We finally remark that H is not defined on ϕ and

therefore on $\mathcal{K}$, since $(\mathcal{K}\phi)(\lambda) = \lambda\phi(\lambda)$ is not square integrable.

Example 2: Periodic motion: Let $Z(t) = \cos t$. This is again a continuous function of positive type, satisfying $Z(0) = 1$ and $Z(-t) = Z^\dagger(t)$ (not a semigroup). From

$$\begin{aligned} Z(t) &= \tfrac{1}{2}(e^{it} + e^{-it}) \\ &= \tfrac{1}{2}\int e^{-i\lambda t}[\delta(\lambda+1) + \delta(\lambda-1)]d\lambda \\ &= \int e^{-it}d(PE(\lambda)P), \end{aligned}$$

it follows that $\sigma(H)$ consists of two discrete, simple eigenvalues ± 1 in the minimal dilation of $Z(t)$. The triplet $\{\mathcal{K}, U(t), P\}$ is then represented by

$$\mathcal{K} = R^2, \quad U(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

3. UNITARY DILATION OF SEMIGROUPS

We establish relations between contracted motions $U_{\mathcal{K}}(t)$ and the spectrum of their minimal unitary dilations $U(t)$ or, respectively, the infinitesimal generators H of $U(t)$. Since we shall be mainly concerned with decay, we impose on $U_{\mathcal{K}}(t)$ to be strongly contractive semigroups.

Assumption 1: Let $U(t) = e^{-iHt}$ be a motion in a Hilbert space $\mathfrak{H}$, $\mathcal{K} = P\mathfrak{H}$ a proper, finite-dimensional subspace of $\mathfrak{H}$ which is generating under $U(t)$, and the contracted motion $U_{\mathcal{K}}(t)$ a strongly contractive semigroup.

Note that in this case $U_{\mathcal{K}}(t)$ is automatically continuous, contractive, and satisfies $U_{\mathcal{K}}(-t) = U_{\mathcal{K}}^\dagger(t)$ and $U_{\mathcal{K}}(0) = I_{\mathcal{K}}$; $U(t)$ is isomorphic to the minimal dilation of $U_{\mathcal{K}}(t)$. If B is the generator of $U_{\mathcal{K}}(t)$, $E_{\mathcal{K}}(\lambda)$, and $R_{\mathcal{K}}(\lambda)$ being the contractions of the spectral family $E(\lambda)$ of H and the resolvent $R(z) = (z - H)^{-1}$, then we have

$$\begin{aligned} U_{\mathcal{K}}(t) &= \begin{Bmatrix} e^{Bt} & t \geq 0 \\ e^{-B^\dagger t} & t < 0 \end{Bmatrix} = \int_{\sigma(H)} e^{-i\lambda t} dE_{\mathcal{K}}(\lambda) \\ &= \frac{1}{2\pi i} \oint e^{-izt} R_{\mathcal{K}}(z)dz, \end{aligned} \tag{3.1}$$

where the integration path $\circlearrowleft$ encloses the spectrum $\sigma(H)$.

In Theorem 2 we collect results which can be simply derived from the Assumptions 1. Although one or two of them are found in different contexts elsewhere, for instance part (1) in Ref. 3, we feel it useful to restate all proofs under our unified assumptions.

Theorem 2: Under the Assumption 1,

(1) $\sigma(H) = R$ (entire real line).

(2) $\mathrm{Re}\sigma(B) < 0$.

(3) $\sigma(H)$ absolutely continuous.

(4) H not defined on all $\mathcal{K}$.

(5) H defined on $\phi \in \mathcal{K}$ if and only if $\mathrm{Re}(\phi, B\phi) = 0$, $\mathrm{Im}(\phi, B^2\phi) = 0$.

(6) The expectation value $\langle H\rangle_\phi \equiv \int \lambda d(\phi, E(\lambda)\phi)$ of H in the state ϕ exists if and only if $\mathrm{Re}(\phi, B\phi) = 0$.

Parts (1) and (4) do not require the assumption that $\mathcal{K}$ is finite dimensional.

Proof: Part (1): $R_{\mathcal{K}}(z)$ is regular analytic for $z \notin \sigma(H)$ and can be expressed in terms of B and $B^\dagger$ as follows:

$$R_{\mathcal{K}}(z) = \begin{cases} (z - iB)^{-1}, & \mathrm{Im}z > 0, \\ (z + iB^\dagger)^{-1}, & \mathrm{Im}z < 0. \end{cases} \tag{3.2}$$

In fact, for $\mathrm{Im}z > 0$,

$$\begin{aligned} R_{\mathcal{K}}(z) &= -i\int_0^\infty Pe^{i(z-H)t}Pdt \\ &= -i\int_0^\infty e^{i(z-iB)t}dt = (z - iB)^{-1} \end{aligned}$$

and, for $\mathrm{Im}z < 0$,

$$\begin{aligned} R_{\mathcal{K}}(z) &= i\int_{-\infty}^0 Pe^{i(z-H)t}Pdt \\ &= i\int_{-\infty}^0 e^{i(z+iB^\dagger)t}dt = (z + iB^\dagger)^{-1}. \end{aligned}$$

Suppose now $\sigma(H) \neq R$. Then $R - \sigma(H)$ is an open subset of R and $R_{\mathcal{K}}(z)$ can be analytically continued from one half-plane to the other. We obtain therefore $(z - iB)^{-1} = (z + iB^\dagger)^{-1}$ for all z and this implies $B^\dagger = -B$. According to (3.1) we have then $U_{\mathcal{K}}(t) = e^{Bt}$ for all $t \in R$ with B anti-Hermitian. Hence $U_{\mathcal{K}}(t)$ would be a unitary group with minimal dilation $U(t) = U_{\mathcal{K}}(t)$ and $\mathcal{K} = \mathfrak{H}$ not a proper subspace of $\mathfrak{H}$. This proves (1).

Part (2): B is an operator in the finite-dimensional space $\mathcal{K}$; $\sigma(B)$ is therefore discrete. Suppose $z \in \sigma(B)$ with $\mathrm{Re}z \geq 0$. Then for a vector ϕ in the corresponding eigenspace,

$$\|U_{\mathcal{K}}(t)\phi\|^2 = \|e^{Bt}\phi\| = e^{(2\mathrm{Re}z)t}\|\phi\|;$$

hence $U_{\mathcal{K}}(t)$ is not strongly contractive.

Part (3): According to (3.1) and (3.2), for any $\epsilon > 0$,

$$\begin{aligned} U_{\mathcal{K}}(t) &= \frac{1}{2\pi i}\int_{-\infty}^\infty e^{-i\lambda t}[R_{\mathcal{K}}(\lambda - i\epsilon) - R_{\mathcal{K}}(\lambda + i\epsilon)]d\lambda \\ &= \frac{1}{2\pi i}\int_{-\infty}^\infty [(\lambda + iB^\dagger - i\epsilon)^{-1} - (\lambda - iB + i\epsilon)^{-1}] \\ &\quad \times e^{-i\lambda t}d\lambda. \end{aligned} \tag{3.3}$$

Since $\mathrm{Re}\sigma(B) < 0$, $\mathrm{Re}\sigma(B^\dagger) < 0$, the real axis R belongs to the regularity domain of both $(z + iB^\dagger)^{-1}$ and $(z - iB)^{-1}$. Hence the limit $\epsilon \to 0$ exists and can be carried out in the integrand. We then obtain

$$\begin{aligned} U_{\mathcal{K}}(t) &= \int_{\sigma(H)} e^{-i\lambda t}dE_{\mathcal{K}}(\lambda) \\ &= \frac{1}{2\pi i}\int_{-\infty}^\infty e^{-i\lambda t}[(\lambda + iB^\dagger)^{-1} - (\lambda - iB)^{-1}]d\lambda, \end{aligned}$$

from which we conclude that $E_{\mathcal{K}}(\lambda)$ has the Radon-Nikodym derivative (almost everywhere)

$$\frac{d}{d\lambda} E_{\mathcal{K}}(\lambda) = \frac{1}{2\pi i} [(\lambda + iB^\dagger)^{-1} - (\lambda - iB)^{-1}]. \qquad (3.4)$$

Hence $E_{\mathcal{K}}(\lambda)$ is absolutely continuous. Suppose now, $E(\lambda)$ were not absolutely continuous; then P would have to project away the nonabsolutely continuous part of $\mathcal{K}$ with respect to H. [The projection of the nonabsolutely continuous part of $E(\lambda)$ cannot be absolutely continuous unless it is zero.] But this is impossible, since $P\mathcal{K}$ is generating. In fact, if $P\mathcal{K}$ is generating, then $E(\Delta)P \neq 0$ for all Borel sets $\Delta \subset R$. Let Δ_0 be the support of the non-absolutely continuous part of $\sigma(H)$. $E(\Delta_0) \neq 0$ then implies $PE(\Delta_0) = (E(\Delta_0)P)^\dagger \neq 0$ and it follows that $E_{\mathcal{K}}(\Delta_0) = (PE(\Delta_0))(PE(\Delta_0))^\dagger$ is nonzero. Therefore, $E(\lambda)$ and the spectrum $\sigma(H)$ are absolutely continuous.

Part (4): If H is defined on all $\mathcal{K}$, then it is defined on an orbit $U_{\mathcal{K}}(t)\phi$, for all $t \geq 0$. But then it follows from the semigroup property that $PHe^{-iHt}P = PHPe^{-iHt}P$, and hence

$$\frac{d}{dt} \|U_{\mathcal{K}}(t)\phi\|^2 = \frac{d}{dt} \|Pe^{-iHt}\phi\|^2$$
$$= 2\mathrm{Re}[-i(Pe^{-iHt}\phi, HPe^{-iHt}\phi)] = 0$$

identically for all $t \geq 0$; $U_{\mathcal{K}}(t)$ would therefore not be strongly contractive.

Part (5): Since the spectrum $\sigma(H)$ is the entire real line R and absolutely continuous, there exists a spectral representation $\int \mathcal{K}_\lambda d\lambda$ with respect to H by square integrable vector functions. Let $\{\phi(\lambda)\}$ with $\phi(\lambda) \in \mathcal{K}_\lambda$ be the spectral representation of $\phi \in \mathcal{K}$. Then, using (3.4) and the relation

$$(\phi, U(t)\phi) = \int e^{-i\lambda t} \|\phi(\lambda)\|_\lambda^2 d\lambda = \int e^{-i\lambda t} d(\phi, E(\lambda)\phi),$$

we find

$$\|\phi(\lambda)\|_\lambda^2 = \frac{d}{d\lambda} (\phi, E_{\mathcal{K}}(\lambda)\phi)$$
$$= \frac{1}{2\pi i} (\phi, [(\lambda + iB^\dagger)^{-1} - (\lambda - iB)^{-1}]\phi)$$
$$= -\frac{1}{2\pi} (\phi, (\lambda + iB)^{-1}(B + B^\dagger)(\lambda - iB)^{-1}\phi)$$
$$= -\frac{1}{\pi} \mathrm{Re}(\phi, (\lambda + iB^\dagger)^{-1}(\lambda - iB)^{-1} B\phi).$$

Now, asymptotically for $|\lambda| \to \infty$,

$$(\lambda + iB^\dagger)^{-1}(\lambda - iB)^{-1}$$
$$= \lambda^{-2}(1 + iB^\dagger/\lambda)^{-1}(1 - iB/\lambda)^{-1}$$
$$= \lambda^{-2}[1 - iB^\dagger/\lambda + O(\lambda^{-2})]$$
$$\times [1 + iB/\lambda + O(\lambda^{-2})]$$
$$= \lambda^{-2} + i\lambda^{-3}(B - B^\dagger) + O(\lambda^{-4})$$

and, therefore,

$$\|\phi(\lambda)\|_\lambda^2 \sim -\lambda^{-2}\pi^{-1}\,\mathrm{Re}(\phi, B\phi)$$
$$-\lambda^{-3}\pi^{-1}\,\mathrm{Re}[i(\phi, (B - B^\dagger)B\phi)] + O(\lambda^{-4})$$

or since $\mathrm{Re}[i(\phi, B^\dagger B\phi)] = 0$,

$$\|\phi(\lambda)\|_\lambda^2$$
$$\sim -\lambda^{-2}\pi^{-1}\,\mathrm{Re}(\phi, B\phi) + \lambda^{-3}\pi^{-1}\,\mathrm{Im}(\phi, B^2\phi)$$
$$+ O(\lambda^{-4}). \qquad (3.5)$$

Since H is defined on ϕ if and only if

$$\|H\phi\|^2 = \int_{-\infty}^{\infty} \lambda^2 \|\phi(\lambda)\|_\lambda^2 d\lambda < \infty,$$

(3.5) implies as a necessary and sufficient condition

$$\mathrm{Re}(\phi, B\phi) = 0, \quad \mathrm{Im}(\phi, B^2\phi) = 0.$$

Remark 1: Necessity could have been derived in a much simpler way: if H is defined on $\phi \in \mathcal{K}$, then the equation $(\phi, e^{-iHt}\phi) = (\phi, e^{Bt}\phi)$ has first- and second-order derivatives at $t = 0$, i.e.,

$$(\phi, B\phi) = -i(\phi, H\phi) \quad \text{purely imaginary},$$
$$(\phi, B^2\phi) = -(\phi, H^2\phi) = -\|H\phi\|^2 \quad \text{real}.$$

Part (6): the expectation value $\langle H\rangle_\phi$ of H in ϕ is defined if and only if $\int_{-\infty}^{\infty} \lambda \|\phi(\lambda)\|_\lambda^2 d\lambda < \infty$, or, according to (3.5) if and only if $\mathrm{Re}(\phi, B\phi) = 0$.

■

We now construct an example of a strongly contractive semigroup which admits a state on which H is defined. The existence of such an example proves that Part (4) of Theorem 3 cannot be strengthened to the statement that the domain of H and $\mathcal{K}$ are disjoint, and Parts (5) and (6) are therefore not empty.

Example 3: Let $\mathcal{K}$ be a two-dimensional Hilbert space with orthonormal basis ϕ_1, ϕ_2. Let $B = \begin{pmatrix} -1 & 2\sqrt{2} \\ 0 & -2 \end{pmatrix}$. Then

$$U_{\mathcal{K}}(t) = \begin{cases} e^{-t}\begin{pmatrix} 1 & 2\sqrt{2}(1 - e^{-t}) \\ 0 & e^{-t} \end{pmatrix}, & t \geq 0 \\ e^{t}\begin{pmatrix} 1 & 0 \\ 2\sqrt{2}(1 - e^{-t}) & e^{t} \end{pmatrix}, & t < 0 \end{cases}$$

is a contractive semigroup satisfying

$$U_{\mathcal{K}}(0) = I, \quad U_{\mathcal{K}}(-t) = U_{\mathcal{K}}{}^\dagger(t), \quad \text{s-}\lim_{t\to\infty} U_{\mathcal{K}}(t) = 0.$$

$U_{\mathcal{K}}(t)$ is therefore a strongly (and strictly) contractive semigroup. Let $U(t) = e^{-iHt}$ be its minimal unitary dilation, and $E_{\mathcal{K}}(t)$ the contracted spectral family of H. Then we have, by Fourier transform,

$$U_{\mathcal{K}}(t) = \int_{-\infty}^{\infty} e^{-i\lambda t} dE_{\mathcal{K}}(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t} \frac{1}{\pi} \begin{pmatrix} 1/(\lambda^2 + 1) & -\sqrt{2}/[(\lambda + i)(\lambda + 2i)] \\ -\sqrt{2}/[(\lambda - i)(\lambda - 2i)] & 2/(\lambda^2 + 4) \end{pmatrix} d\lambda, \qquad (3.6)$$

and we conclude that $E_{\mathcal{K}}(\lambda)$ admits a Radon–Nikodym derivative with support R and that therefore the spectrum $\sigma(H)$ is the entire line, absolutely continuous and of uniform multiplicity 2.

Let $\phi_i(\lambda)$ be the spectral representation with respect to H of the basis vectors $\phi_i \in \mathcal{K}$. Then, comparing

$$(U_{\mathcal{K}})_{ik}(t) \equiv (\phi_i, U_{\mathcal{K}}(t)\phi_k) = \int_{-\infty}^{\infty} e^{-i\lambda t}\overline{\phi_i(\lambda)}\,\phi_k(\lambda)d\lambda$$

to (3.6), we obtain the normalized functions

$$\phi_1(\lambda) = \frac{1}{i\sqrt{\pi}}\,\frac{1}{\lambda - i}, \qquad \phi_2(\lambda) = i\sqrt{\frac{2}{\pi}}\,\frac{1}{\lambda + 2i}$$

in $L^2(R)$. This achieves the reconstruction of the minimal dilation, since we know how H and $e^{-iHt} = U(t)$ act in this representation and since we have identified the two-dimensional subspace $\mathcal{K}$ spanned by $\phi_1(\lambda)$ and $\phi_2(\lambda)$.

The particular feature of this example is the fact that although H is not defined on the basis ϕ_i, since $\lambda\phi_i(\lambda) \notin L^2(R)$, there exists a unique linear combination (up to a scalar multiple)

$$\phi = \sqrt{2}\phi_1 + \phi_2$$

or, in spectral representation,

$$\phi(\lambda) = \sqrt{2}\phi_1(\lambda) + \phi_2(\lambda) = 3\sqrt{\left(\frac{2}{\pi}\right)} \times \frac{1}{(\lambda - i)(\lambda + 2i)} \sim O(\lambda^{-2})$$

on which H and $\langle H\rangle_\phi$ are defined ($\|H\phi\|^2 = 6$, $\langle H\rangle_\phi = 0$). This implies, in particular, that

$$\frac{d}{dt}\,\|U_{\mathcal{K}}(t)\phi\|^2\Big|_{t=0} = 2\,\mathrm{Re}(\phi, B\phi) = 0 \qquad (3.7)$$

for this special vector ϕ, while for all other vectors in $\mathcal{K}$ this derivative is negative, consistent with contractiveness, as can easily be checked in this concrete example.

Relation (3.7) is interesting, because it means that in a semigroup with $\|\|U_{\mathcal{K}}(t)\|\| \leq 1$ a vector ϕ may exist whose norm has initial decay rate 0, although $U_{\mathcal{K}}(t)\phi \to 0$ as $t \to \infty$. This is clearly due to the nonnormality of B which makes it possible that the numerical range of B exceeds the spectrum $\sigma(B)$. [In our example, although $\sigma(B) = \{-1, -2\}$, $(\phi, B\phi) = 0$]. If B is normal the most general two-dimensional strongly contractive real semigroup is a logarithmic spiral and $(d/dt)\|e^{Bt}\phi\|^2\big|_{t=0} < 0$ for all $\phi \in \mathcal{K}$.

Remark 2: Theorem 2 has been formulated under the global assumption 1 for the sake of simplicity. It should, however, be noted that the full strength of these assumptions is not necessary for all of the assertions made. In particular, (1) and (4) do not require that the subspace $\mathcal{K}$ be generating or finite-dimensional. Furthermore, the necessity of (5) has been proved without using strong contractiveness of $U(t)$.

The sufficiency of the conditions in Parts (5) and (6) depend on Part (2) which in turn makes essential use of the finiteness of $\mathcal{K}$, since in infinite-dimensional subspaces the spectrum $\sigma(B)$ can be continuous and may extend to $\mathrm{Re}\sigma(B) = 0$. In that case, the limit $\epsilon \to 0$ in the integrand of (3.3) may not lead to an integrable function and absolute continuity of $\sigma(H)$ as expressed in (3.4) may fail (see Ref. 3).

If we content ourselves with ordinary continuity of $\sigma(H)$, the following modification of Part (3) can be proved without assuming the semigroup property:

Theorem 3: If $U_{\mathcal{K}}(t)$ is a contracted motion which is strongly contractive (indeed even weakly) (not necessarily a semigroup), then $\sigma(H)$ is continuous.

Proof: Let $E_c(\lambda)$ be the spectral family of the continuous part H_c of H and E_j the projections on the discrete eigenspaces with eigenvalues λ_j. For an arbitrary $\phi \in \mathcal{K}$ we then have, for $t \to \infty$,

$$(\phi, PU(t)P\phi) = \sum_j e^{-i\lambda_j t}\|E_jP\phi\|^2 + \int_{\sigma(H)} e^{-i\lambda t}d(\phi, PE_c(\lambda)P\phi) \to 0.$$

We must show that $E_jP\phi = 0$ for all $\phi \in \mathcal{K}$, since then $P\mathcal{K}$ is not generating under $U(t)$ unless $E_j = 0$. [Note that if $E_c(\lambda)$ were absolutely continuous, this would be an immediate consequence of the Riemann–Lebesgue lemma.]

Let

$$k(t) = \sum_j e^{-i\lambda_j t}\|E_jP\phi\|^2 = (\phi, PU(t)P\phi) - \int_{\sigma(H)} e^{-i\lambda t}dm(\lambda),$$

where $dm(\lambda) = d(\phi, PE_c(\lambda)P\phi)$ is of finite variation and has no point measure. According to Wiener's theorem (cf., for example, Ref. 3, p. 145),

$$\hat{m}(t) = \int e^{-i\lambda t}dm(\lambda)$$

has the property that

$$\lim_{t\to\infty}\frac{1}{T}\int_{-T}^{T}|\hat{m}(t)|^2\,dt = 0.$$

It is a corollary to this theorem that, given any $\epsilon, \delta > 0$ and T sufficiently large, the Lebesgue measure of the set of t in $(-T, T)$ for which $|\hat{m}(t)| \geq \epsilon$ is less than $T\delta$. There is, furthermore, by the contractive property, a T_0 such that $|(\phi, PU(t)P\phi)| < \epsilon$ for $t \geq T_0$. Since

$$|k(t)| \leq |(\phi, PU(t)P\phi)| + |\hat{m}(t)|,$$

it follows that $|k(t)| < 2\epsilon$ on a set of measure greater than $2T - T\delta - 2T_0$ in $(-T, T)$ (for T sufficiently large). $k(t)$ is, however, an almost periodic function, since the λ_j are denumerable and the series is absolutely convergent ($\sum_j \|E_jP\phi\|^2 \leq \|P\phi\|^2$).

Hence the coefficients are given by

$$\|E_j P\phi\|^2 = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{i\lambda_j t}\, k(t)\, dt$$
$$< \lim_{T\to\infty} \frac{1}{2T} [4\epsilon T + (T\delta + 2T_0)\|P\phi\|^2]$$
$$< 2\epsilon + \delta \|\phi\|^2,$$

i.e., they vanish, as was to be proven. ∎

4. PHYSICAL INTERPRETATION

We introduce the following language suitable for a quantum-mechanical interpretation. $\mathcal{K}$ is the space of *physical states*, $U(t)$ their *time evolution*, and $\mathcal{K}$ a subsystem of $\mathcal{K}$ with time evolution $U_{\mathcal{K}}(t)$. If $U_{\mathcal{K}}(t)$ is strongly contractive, $\mathcal{K}$ is a subspace of *decaying states* and $U_{\mathcal{K}}(t)$ describes their *decay*. The *decay law* of a state $\phi \in \mathcal{K}$ is defined as

$$p_\phi(t) = \|U_{\mathcal{K}}(t)\phi\|^2 = \mathrm{Tr} P U^\dagger(t) P_\phi U(t) P,$$

where P_ϕ is the projection on the one-dimensional subspace spanned by ϕ, and $(d/dt)p_\phi(t)$ is the *decay rate* of ϕ. $p_\phi(t)$ is interpreted physically as the probability that a state ϕ created at time $t = 0$ is still in the subspace $\mathcal{K}$ at time t.

The infinitesimal generator H of $U(t)$ is called the *Hamiltonian* and its physical significance is that of the total energy of the system. The *expectation value of the energy in the state* ϕ is defined by

$$\langle H\rangle_\phi = \int \lambda d(\phi, E(\lambda)\phi).$$

If H is defined on ϕ this equals simply $(\phi, H\phi)$. A state ϕ for which $\langle H\rangle_\phi$ is finite is called a *finite energy state*; otherwise it is called an *infinite energy state*.

Suppose now that the evolution $U_{\mathcal{K}}(t)$ of a subsystem $\mathcal{K}$ is given, then the minimal unitary dilation $\{\mathcal{K}, U(t)\}$ of $\{\mathcal{K}, U_{\mathcal{K}}(t)\}$ can be interpreted as the *minimal quantum-mechanical phenomenology* compatible with $U_{\mathcal{K}}(t)$.

With this terminology the results of Sec. 3 can be interpreted as follows:

Theorem 4: Let $U_{\mathcal{K}}(t)$ be the time evolution of a finite-dimensional subspace $\mathcal{K}$, $\{\mathcal{K}, U(t)\}$ the corresponding minimal quantum-mechanical phenomenology, and H the total energy.

(1) If $U_{\mathcal{K}}(t)$ is a semigroup, then the total energy is not positive.

(2) If $U_{\mathcal{K}}(t)$ is a semigroup which decays, then there exist infinite energy states in $\mathcal{K}$.

(3) If $U_{\mathcal{K}}(t)$ is a semigroup, finite energy states have initial decay rate zero.

(4) If all states in $\mathcal{K}$ have finite energy, then $U_{\mathcal{K}}(t)$ is not a semigroup.

Proof: Part (1) is an immediate consequence of Theorem 2, Part (1). Part (2): If all $\phi \in \mathcal{K}$ had finite energy, then, according to 2 (6), $\mathrm{Re}(\phi, B\phi) = 0$ for all $\phi \in \mathcal{K}$, and hence, in particular, $\mathrm{Re}(U_{\mathcal{K}}(t)\phi, BU_{\mathcal{K}}(t)\phi) = 0$ for all $t \geq 0$. But then

$$\frac{d}{dt}\|U_{\mathcal{K}}(t)\phi\|^2 = 2\,\mathrm{Re}(U_{\mathcal{K}}(t)\phi, BU_{\mathcal{K}}(t)\phi) \equiv 0$$

and $U_{\mathcal{K}}(t)$ would not be strongly contractive.

Part (3): If $\phi \in \mathcal{K}$ had finite energy, then again, according to 2 (6), $\mathrm{Re}(\phi, B\phi) = 0$ and hence

$$\frac{d}{dt} p_\phi(t)\Big|_{t=0} = \frac{d}{dt}\|U_{\mathcal{K}}(t)\phi\|^2\Big|_{t=0} = 2\,\mathrm{Re}(\phi, B\phi) = 0.$$

Part (4) is an immediate reformulation of 4 (2). We remark that an evolution which decays nowhere in $\mathcal{K}$ is unitary in $\mathcal{K}$ and is isomorphic to its minimal dilation. ∎

Next we abandon the semigroup property for $U_{\mathcal{K}}(t)$ and assume instead that the finite-dimensional subspace $\mathcal{K}$ has finite energy. We then decompose H as follows:

$$H = H_0 + V, \quad H_0 = PHP + \bar{P}H\bar{P},$$
$$V = PH\bar{P} + \bar{P}HP, \qquad (4.1)$$

where $\bar{P} = I - P$. This decomposition has the following properties

(1) H_0 and H are self-adjoint.

(2) The eigenstates ϕ_i of PHP have eigenvalues $m_i = \langle H\rangle_{\phi_i}$.

(3) V has finite rank $2n$, where n is the dimension of $\mathcal{K}$.

The decomposition (4.1) is canonical in the sense that the subspace $\mathcal{K}$ is invariant under H_0 while V contains only the part of H under which $\mathcal{K}$ decays (no self-interaction of $P\mathcal{K} = \mathcal{K}$ and $\bar{P}\mathcal{K} = \mathcal{K}\ominus\mathcal{K}$). With these conventions we have

Theorem 5: If $\mathcal{K}$ has finite energy [the decomposition (4.1) being therefore well defined], then $U_{\mathcal{K}}(t)$ has the form

$$U_{\mathcal{K}}(t) = \frac{1}{2\pi i} \oint e^{-izt} R_{\mathcal{K}}(z)\, dz,$$
$$R_{\mathcal{K}}(z) = \Big(\sum_i (z - m_i) P_i - PV\bar{P}R_0(Z)\bar{P}VP\Big)^{-1}, \qquad (4.2)$$

where P_i are the projections on the eigenstates ϕ_i of PHP with eigenvalues m_i and $P_0(z) = (z - H_0)^{-1}$ is the resolvent of H_0. Furthermore, $U_{\mathcal{K}}(t)$ satisfies the integro-differential equation

$$i\frac{d}{dt} U_{\mathcal{K}}(t)$$
$$= PHU_{\mathcal{K}}(t) - i\int_0^t d\tau PH\bar{P}e^{-i\bar{P}H\bar{P}\tau} PHU_{\mathcal{K}}(t-\tau). \qquad (4.3)$$

Proof: (4.2) follows from the second resolvent formula $R(z) = R_0(z) + R_0(z)VR(z)$ by projecting into the subspaces $P\mathcal{K}$ and $\bar{P}\mathcal{K}$ and noting that $P, \bar{P}$ reduce $R_0(z)$. Similarly (4.3) is obtained from the

identity $zR(z) - I = HR(z)$ by projection with $P, \bar{P}$ and by making the inverse Laplace transform in which products become convolutions. (For details see Ref. 2.) ■

In the particular case where $\sigma(H)$ is assumed to be absolutely continuous, the decomposition (4.1) leads to the so-called *decay scattering systems* $\{H_0, H\}$. Such systems, with physical applications, for example, to K-meson decay, have been treated in detail elsewhere. (See Ref. 2 and 5.) Due to the finite rank of the "potential" V, the spectrum of $\bar{P}H_0\bar{P}$ is absolutely continuous and coincides with $\sigma(H)$; a scattering theory with a unitary scattering matrix on $\bar{P}\mathcal{H}$ then becomes possible. It can be shown that if V is sufficiently small, the discrete eigenvalues m_i of H_0 are embedded in $\sigma(\bar{P}H_0\bar{P})$. The corresponding eigenvectors ϕ_i are then interpreted as unstable particles which decay into the continuum $\bar{P}\mathcal{H}$ through the potential V. The present analysis shows that these decays cannot be exponential; more generally: the decay of the subspace $\mathcal{K} = P\mathcal{H}$ of unstable particles is not a semigroup.

We also consider briefly the *classical scattering theory of Lax and Phillips*.[3] In this theory the total time evolution is again given by a unitary motion $U(t)$ in a Hilbert space $\mathcal{H}$. There exists furthermore a subspace $\mathcal{K}$ of $\mathcal{H}$ for which the contracted motion $U_{\mathcal{K}}(t)$ is a strongly contractive semigroup. The subspace $\mathcal{K}$ is, however, infinite dimensional in most practical cases. If we assume that $\mathcal{K}$ is generating under $U(t)$, then Theorem 2 immediately implies that in the Lax–Phillips system the generator H of $U(t)$

(1) is not positive

(2) is not defined on the entire space $\mathcal{K}$.

Note however, that in this theory the generator H is *not* interpreted as the energy. The latter corresponds rather to H^2. The physical significance of (1) and (2) (stated above) is, therefore, less striking in the classical context.

In the following paragraphs, we review the parts of the theory which are essential for our remarks, and show in what sense the subspace $\mathcal{K}$ can be considered as generating.

The theory of Lax and Phillips is characterized by the existence of orthogonal subspaces D_+ and D_- of $\mathcal{H}$, called outgoing and ingoing, and a unitary group $U(t)$, with the properties

(i) $U(t)D_+ \subset D_+$ and $U(-t)D_- \subset D_-, t > 0$,

(ii) $\cap\, U(t)D_+ = \{0\}$,

(iii) $\overline{\cup\, U(t)D_+} = \mathcal{H}, \ \forall t$.

The statements (ii) and (iii) are valid for D_- as well. Let P_+ and P_- be the orthogonal projections with null spaces D_+ and D_-. Then, the one-parameter family of operators $Z(t) = P_+U(t)P_-$ is a semigroup on the subspace $\mathcal{K} = (D_+ \oplus D_-)^\perp$. To verify this assertion, we note that

$Z(t+s) - Z(t)Z(s) = P_+U(t)[I - P_-P_+]U(s)P_-$ vanishes on $\mathcal{K}$. If $x \in \mathcal{K}$, then $P_-x \in D_-^\perp$; since $U(-t)D_- \subset D_-$, $U(t)D_-^\perp \subset D_-^\perp$, and $y = (I - P_-P_+)U(s)P_-x$ therefore belongs to D_+. It then follows from property (i) that $P_+U(t)y = 0$.

We now show that $Z(t)$ may be equivalently defined as

$$Z(t) = P_{\mathcal{K}}U(t)P_{\mathcal{K}},$$

i.e., as a contracted motion. To see this, we write $x = x_{D_+} + x_{D_-} + x_{\mathcal{K}}$; then

$$\begin{aligned}(P_+U(t)P_-)\,x &= P_+U(t)\,(x_{D_+} + x_{\mathcal{K}}) \\ &= P_+U(t)P_{\mathcal{K}}x \\ &= P_{\mathcal{K}}U(t)P_{\mathcal{K}}x,\end{aligned}$$

where the last equality follows from the fact that $P_{\mathcal{K}}x \in D_-^\perp$; $U(t)$ maps $D_-^\perp$ into itself, and on this subspace, P_+ is equal to $P_{\mathcal{K}}$.

The subspace $\mathcal{K}$ can be considered as generating in the sense of the following theorem:

Theorem 6: Let $\mathcal{H}'$ be the subspace of $\mathcal{H}$ generated by $\mathcal{K}$ under $U(t)$. Then the scattering operator S is reduced by the decomposition $\mathcal{H} = \mathcal{H}' \oplus (\mathcal{H} \ominus \mathcal{H}')$ and its part in $(\mathcal{H} \ominus \mathcal{H}')$ is trivial.

Proof: Let $\bar{D}_\pm = D_\pm \ominus (D_\pm \cap \mathcal{H}')$. Since $U(t)$ is unitary and invariant on $\mathcal{H}'$, it is invariant on $\bar{D}_+ \oplus \bar{D}_-$, and the scattering problem can therefore be considered on each part separately. Since $\bar{D}_+$, $\bar{D}_-$, $U(t)$ forms a scattering system in which $\bar{D}_+$ [$\bar{D}_-$] is outgoing [incoming] and $\mathcal{H} \ominus \mathcal{H}' = \bar{\mathcal{H}} = \bar{D}_+ \oplus \bar{D}_-$, we know that in this part the scattering matrix is equivalent to I. The study of scattering systems can therefore be restricted to situations in which $\mathcal{K}$ is generating. ■

Note added in proof: Meanwhile David Williams published an article in Commun. Math. Phys. **21**, 314 (1971), whose results (not methods) partly overlap with ours.

ACKNOWLEDGEMENT

We wish to thank Professor Calvin Wilcox for many interesting and useful discussions of the theory of Lax and Phillips.

* This work was supported in part by National Science Foundation grant No. GP-23304.

1 Bela Sz.-Nagy and Ciprian Foias, *Analyse harmonique des operateurs de l'espace de Hilbert* (Masson, Paris, 1967). F. Riesz and B. Sz.-Nagy, *Functional Analysis* (appendix).

2 L. P. Horwitz and J.-P. Marchand, Rocky Mountain J. Math. **1**, 225 (1971).

3 P. Lax and R. S. Phillips, *Scattering Theory* (Academic, New York, 1967).

4 R. McKelvey, J. Math. Anal. Appl. **11**, 447 (1966).

5 L. P. Horwitz and J.-P. Marchand, Helv. Phys. Acta **42**, 1039 (1969).