JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 9, NUMBER 1 JANUARY 1968

Symmetrized Tensor "Trace" Operations

DONALD R. TOMPKINS*
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana

(Received 29 November 1966)

Extracting the "trace" from a Cartesian tensor requires working with the whole tensor. Extracting the trace from a symmetrized tensor is shown to allow one to work with separate invariant subspaces (hence yielding more convenient analytical results), and thus, reduces the number of coupled equations to be solved. Symmetrized trace extraction is described for extended symmetrization procedures based respectively on modified Young symmetrizers, Wigner projection operator symmetrizers, and Young symmetrizers. Symmetrized traceless tensors usually need to be further symmetrized to obtain fully symmetrized traceless tensors. Also described is a method where each rank of trace extraction is performed in a separate step and is accompanied by a step of symmetrization. This method yields fully symmetrized traceless tensors and the least coupling of equations.

1. INTRODUCTION

Tensors with vanishing mixed variance "traces" [i.e., $\sum_{ij} \delta^i_j T^{j\cdots}_{i\cdots} = 0$] constitute subspaces which are invariant under the full linear group $GL(n, c)$ and its subgroups.[1] For subgroups of $GL(n, c)$ which possess a metric tensor $G_{ij}(G^{ij})$ the condition of vanishing mixed variance traces is equivalent to the conditions

$$\sum_{ij} G_{ij} T^{ijk\cdots} = 0,$$

$$\sum_{ik} G_{ik} T^{ijk\cdots} = 0, \cdots \left(\sum_{ij} G^{ij} T_{ijk\ldots} = 0, \right.$$

$$\left. \sum_{ik} G^{ik} T_{ijk\ldots} = 0, \cdots \right).$$

For tensors carrying indices of both a group H and its cover group G (such tensors are respectively denoted by Latin and Greek indices) where H has some metric tensor, we can use a vector operator $(\sigma_i)^\alpha_\beta$ to write

$$T_{ijk\cdots h\beta} = \sum_\alpha (\sigma_i)^\alpha_\beta T_{jk\cdots h\alpha}.$$

Using the metric tensor to impose a vanishing trace condition among all vector indices $i, j, \cdots, h$ yields in general a nontrivial decomposition into invariant subspaces.[2]

Symmetrization of tensor indices also furnishes invariant subspaces and, for the classical groups, use of both symmetrization and simple trace operations yields irreducible tensors. Although the order of application is arbitrary, the usual approach is to first extract the traces and then symmetrize the traceless

* Present address: Department of Physics and Astronomy, University of Georgia, Athens, Georgia.

[1] We use the word "trace" to collectively describe tensor operations of the kind discussed in the text.

[2] It is exceptional for trace operations to only trivially supplement symmetrization, but the trace operations based on the skew metric tensors of $SL(2, C)$ and $SU(2)$ do present such exceptions.

indices.[3,4] The extraction of the symmetric (antisymmetric) trace from symmetric (antisymmetric) tensors has also been described.[3] For such one-dimensional representations of the permutation group, the traditional symmetrization procedures provide sufficient state organization. The purpose of this article is to describe symmetrized trace extraction for *all* representations appearing in an arbitrary rank tensor. We seek to obtain symmetrized trace extractions in a decoupled form, i.e., with no coupling between tensors derived from equivalent (or inequivalent) representations of the permutation group. To achieve this goal we need the state organization provided by extended symmetrization. [Weyl's symmetrization procedure describes each invariant subspace of a tensor with a single formal state and allows all arrangements of all sets of index values.[5] Extended symmetrization procedures describe each invariant subspace of a tensor with a set of formal states forming a basis representing the permutation group and allow a single arrangement of each set of index values.[6]]

The disadvantages of Cartesian tensor trace extractions are that (a) it is necessary to always deal with the whole tensor, (b) to evaluate the traceless components, a number of coupled linear equations must be solved, and (c) cumbrous analytical results are obtained. These features are illustrated by considering the mixed variance tensor T^h_{ijk}. To obtain the traceless tensor $(T^h_{ijk})^0$ it is necessary to solve six coupled linear equations. Using three-dimensional indices the result is

$$\begin{aligned}(T^h_{ijk})^0 = T^h_{ijk} &- \delta^h_i \sum_{a,b} (\delta^a_b/40)[3T^b_{kaj} - 2T^b_{akj} - 7T^b_{jak} \\ &+ 18T^b_{ajk} - 7T^b_{kja} + 3T^b_{jka}] \\ &- \delta^h_j \sum_{a,b} (\delta^a_b/40)[-2T^b_{kai} - 7T^b_{aik} + 3T^b_{aki} \\ &+ 18T^b_{iak} + 3T^b_{kia} - 7T^b_{ika}] \\ &- \delta^h_k \sum_{a,b} (\delta^a_b/40)[18T^b_{ija} + 3T^b_{aij} + 3T^b_{jai} \\ &- 7T^b_{aji} - 2T^b_{jia} - 7T^b_{iaj}], \qquad (1.1)\end{aligned}$$

which has yet to be symmetrized. Using symmetrized trace extractions it is possible completely to decouple these six equations and obtain results in a much more convenient form.

2. METHOD

For notational convenience the discussion is given in terms of tensors of a single variance and carrying indices of only one group. With appropriate modification the results of this section clearly apply to any of the trace operations described in the Introduction.

The decomposition of a Cartesian tensor into traceless tensors can formally be written[7]

$$\begin{aligned}T_{i_1\cdots i_r} = (T_{i_1\cdots i_r})^0 &+ G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 \\ &+ G_{i_1i_3}(T^{24\cdots r}_{i_2i_4\cdots i_r})^0 + \cdots \\ &+ G_{i_1i_r}(T^{23\cdots r-1}_{i_2i_3\cdots i_{r-1}})^0 + G_{i_2i_3}(T^{14\cdots r}_{i_1i_4\cdots i_r})^0 + \cdots \\ &+ G_{i_2i_r}(T^{13\cdots r-1}_{i_1i_3\cdots i_{r-1}})^0 + \cdots \\ &+ G_{i_{r-1}i_r}(T^{1\cdots r-2}_{i_1\cdots i_{r-2}})^0 + G_{i_1i_2}\langle G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 \\ &+ G_{i_3i_5}(T^{3\cdots r,46\cdots r}_{i_4i_6\cdots i_r})^0 + \cdots\rangle \\ &+ G_{i_1i_3}\langle G_{i_2i_4}(T^{24\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots\rangle + \cdots \\ &+ G_{i_{r-1}i_r}\langle G_{i_1i_2}(T^{1\cdots r-2,3\cdots r-2}_{i_3\cdots i_{r-2}})^0 + \cdots.\rangle + \cdots. \qquad (2.1)\end{aligned}$$

In the notation $(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0$ the $3\cdots r$ indicates that this traceless tensor is derived from $T^{3\cdots r}_{i_3\cdots i_r}$, and the $5\cdots r$ indicates which of the original indices the tensor carries. Thus we may have

$$(T^{3\cdots r}_{i_3\cdots i_r}) = (T^{3\cdots r}_{i_3\cdots i_r})^0 + G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots.$$

This notation is not redundant in actual calculations because indices then take specific values which, due to repeated values, may not unambiguously indicate the indices' origin. A superscript such as $3\cdots r$, $5\cdots r$ is called a *label configuration.*

It has been shown that with the extended symmetrization concept, symmetrized tensor bases for $GL(n, c)$ and $GL(n, c)$ subgroups can be obtained using not only Young symmetrizers but also using modified Young symmetrizers or Wigner projection operators[8] based on the permutation group. These symmetrization procedures are considered separately in the following parts.

A. An Approach Using Modified Young Symmetrizers

Operating on Eq. (2.1) with $(PQ')^\mu_j$ on the left yields

$$\begin{aligned}(PQ')^\lambda_j T_{i_1\cdots i_r} = (PQ')^\lambda_j[&(T_{i_1\cdots i_r})^0 + G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 \\ &+ \cdots. + G_{i_1i_2}\langle G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots\rangle + \cdots]. \qquad (2.2)\end{aligned}$$

[3] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).
[4] M. Ikeda, S. Ogawa, and Y. Ohnuki, Progr. Theoret. Phys. (Kyoto) **23**, 1073 (1960).
[5] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946).
[6] D. Tompkins, J. Math. Phys. **8**, 1502 (1967).
[7] If, instead of being Cartesian, the initial tensor $T_{i_1\cdots i_r}$ has some symmetry conditions among its indices, then such conditions need to be considered when writing an equation expressing a decomposition with trace extractions. Specifically, one needs to use appropriately symmetrized metric tensors instead of the Cartesian metric tensors shown in Eq. (2.1).
[8] E. P. Wigner, *Group Theory* (Academic Press Inc., New York 1959).

Using the permutations $\{S_{ij}\}_{(\text{all } i)}$ relating standard tableaux as $\mathcal{C}_i^\mu = S_{ij}\mathcal{C}_j^\mu$, we operate on Eqs. (2.2) to form the set of N^λ equations (extended symmetrization)

$$\{S_{ij}(PQ')_j^\lambda T_{i_1\cdots i_r} = S_{ij}(PQ')_j^\lambda[(T_{i_1\cdots i_r})^0 + G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 + \cdots + G_{i_1i_2}\langle G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots\rangle + \cdots]\}_{(\text{all } i)}. \tag{2.3}$$

For any given arrangement of indices $i_1\cdots i_r$ these components span a subspace irreducibly invariant under $\mathcal{S}_r$. Contracting *each* of Eqs. (2.3) with $G_{i_1i_2}$, $G_{i_1i_3}$, $\cdots$, and $G_{i_{r-1}i_r}$ gives the

$$[(r-1)+(r-2)+\cdots+(1)]N^\lambda$$

equations.[9]

$$\left\{\sum_{i_1i_2} G^{i_1i_2}S_{ij}(PQ')_j^\lambda T_{i_1\cdots i_r} = \sum_{i_1i_2} G^{i_1i_2}S_{ij}(PQ')_j^\lambda[G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 + \cdots + G_{i_1i_2}\langle G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_r\cdots i_r})^0 + \cdots\rangle + \cdots]\right\}_{(\text{all } i)}$$

$$\vdots$$

$$\left\{\sum_{i_{r-1}i_r} G^{i_{r-1}i_r}S_{ij}(PQ')_j T_{i_1\cdots i_r} = \sum_{i_{r-1}i_r} G^{i_{r-1}i_r}S_{ij}(PQ')_j^\lambda[G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 + \cdots + G_{i_1i_2}G_{i_3i_4}\langle(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots\rangle + \cdots]\right\}_{(\text{all } i)}, \tag{2.4}$$

where vanishing traceless terms were dropped. The remaining symmetrized trace equations are obtained by repeating this process, gaining

$$[(r-1)+(r-2)+\cdots+(1)]N^\lambda$$

additional trace equations each time, until insufficient free indices remain for further contractions. This process enlarges Eqs. (2.4) to Eqs. (2.4) *et seq.* On the right side of Eqs. (2.3) the tensors having common indices (perhaps rearranged) but different label configurations are collected to form expressions which are identified by their assembly of label configurations. The same expressions appear in Eqs. (2.4) *et seq.*, and these are the symmetrized traceless tensors.

By the full reducibility of $\mathcal{S}_r$ the invariant subspaces do not intersect, so all tensors in Eqs. (2.3) are given as solutions of Eqs. (2.4) *et seq.* These equations do not couple the trace operations of independent (but possibly equivalent) representations of $\mathcal{S}_r$.[10] Since all symmetrized traces are obtained by solving Eqs. (2.4) *et seq.*, we see that, associated with representation λ, there are at most $[(r-1)+(r-2)+\cdots+(1)]N^\lambda$ independent symmetrized traces of rank $r-2$, the same number of independent symmetrized traces of rank $r-4$, etc. Each solution of Eqs. (2.4) *et seq.* belongs to a set of tensors [also solutions of Eqs. (2.4) *et seq.*] spanning a subspace which is invariant under the group of permutations of traceless indices. Such a subspace (of traceless rank r') is usually reducible because the symmetry of $\mathcal{S}_r$ often does not coincide with irreducible representation symmetry of $\mathcal{S}_{r'} \subset \mathcal{S}_r$. The tensors which the final symmetrizations yield are those which we call fully symmetrized traceless tensors.

Rather than initially extracting all ranks of symmetrized traces, we can extract each rank of appropriately symmetrized traces in a separate step. We now explain how this is done. We start as usual[11] with Eq. (2.1) and arrive at Eqs. (2.4). Previously we used additional metric tensor contractions to get Eqs. (2.4) *et seq.*, which were solved to get symmetrized traceless tensors of rank $r-2, r-4, \cdots$, 1 (or 0) and by subtraction we obtain the symmetrized traceless tensors of rank r. Instead we now solve Eqs. (2.4) [not Eqs. (2.4) *et seq.*] for symmetrized tensors of rank $r-2$, which are formally expressed as

$$T_{r-2}[(T^{3\cdots r}_{i_3\cdots i_r})^0, \cdots, G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0, \cdots],$$

and are explicitly expressed as

$$\hat{T}_{r-2}\left[\sum_{i_1i_2} G_{i_1i_2}T_{i_1\cdots i_r}, \sum_{i_3i_4} G_{i_3i_4}T_{i_1\cdots i_r}, \cdots\right],$$

where the brackets indicate (linear) functional dependence here. Thus we have

$$\hat{T}_{r-2}\left[\sum_{i_1i_2} G_{i_1i_2}T_{i_1\cdots i_r}, \sum_{i_3i_4} G_{i_3i_4}T_{i_1\cdots i_r}, \cdots\right] = T_{r-2}[(T_{i_3\cdots i_r})^0, \cdots, G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0, \cdots]. \tag{2.5}$$

[9] If a metric tensor G_{ij} is invariant under a matrix group $\mathcal{G}$, then its symmetric and antisymmetric parts are each also invariant. This follows from the fact that symmetrization decomposes tensors of $\mathcal{G}$, and G_{ij} is such a tensor. If both G_{ij}, and $G_{i,j}$ are invariant then the associated trace conditions lead only to vanishing traceless tensors. Hence, we need only consider symmetric and antisymmetric metric tensors. Thus trace equations are formed by contractions with the $[(r-1)+(r-2)+\cdots+(1)]$ tensors $G_{i_1i_2}, G_{i_1i_3}, \cdots, G_{i_1i_r}; G_{i_2i_3}, G_{i_2i_4}, \cdots, G_{i_2i_r}; \cdots; G_{i_{r-1}i_r}$.

[10] The independent representations meant here are those of $\{S_{ij}(PQ')_j^\lambda\}_{\text{all } i}$ belonging to distinct values of j and/or λ. The sums in Eqs. (2.5) *et seq.* do, of course, include representations of S_r which are independent by virtue of belonging to different sets of values of $i_1, \cdots, i_r$.

[11] The reason we use Eq. (2.1) rather than the form

$$T_{i_1\cdots i_r} = (T_{i_1\cdots i_r})^0 + G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r}) + G_{i_1i_3}(T^{24\cdots r}_{i_2i_4\cdots i_r}) + \cdots + G_{i_{r-1}i_r}(T^{1\cdots r-2}_{i_1\cdots i_{r-2}})$$

[the tensors $(T^{3\cdots r}_{i_3\cdots i_r}), \cdots$ are generally not traceless] is because the explicitly symmetrized metric tensors are needed for the subsequent steps of symmetrized trace extraction.

By subtraction we obtain the symmetrized traceless tensors of rank r. Individually considering each one of Eqs. (2.5) [one equation for each tensor T_{r-2}], we then multiply through by $\{S_{ij}(PQ')^{\lambda}_{j}\}_{(\text{all } i)}$ of $\mathcal{S}_{r-2}$ and contract the resulting equations with $\{G_{ij}\}$ to obtain equations which we solve for symmetrized tensors of rank $r-4$. These tensors are expressed formally as

$$T_{r-4}[(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0, \cdots, G_{i_5i_6}(T^{3\cdots r,5\cdots r,7\cdots r}_{i_7\cdots i_r})^0, \cdots],$$

and explictly as

$$\hat{T}_{r-4}\Big[\sum_{i_3i_4} G_{i_3i_4}\sum_{i_1i_2} G_{i_1i_2}T_{i_1\cdots i_r}, \cdots\Big].$$

By subtraction we obtain the symmetrized traceless tensors of rank $r-2$. We continue this process to obtain, by subtraction, traceless tensors having the symmetry of irreducible representations of $\mathcal{S}_r$, $\mathcal{S}_{r-2}, \cdots, \mathcal{S}_2$, plus a (traceless) remainder of rank 1 or 0. This method can also be used with the symmetrization procedures of Parts B and C. In fact, since the symmetrizations of rank r, $r-2, \cdots, 2$ are independent, they can be carried out by different procedures.

This method yields the least possible coupling among trace equations because no further decoupling of the trace extractions is possible and irreducible representation symmetrizers are used. This method is illustrated in the second example of Sec. 4.

B. An Approach Using Wigner Projection Operators

Multiplying Eq. (2.1)

$$U^{\lambda}_{jq} \equiv (N^{\lambda}/r!)^2 \sum_{\xi,\gamma} [D^{\lambda}_{\xi^{-1}}]_{qj}[D^{\lambda}_{\gamma^{-1}}]_{qq}L_{\xi}L_{\gamma}$$

and allowing all j gives the N^{λ} equations

$$\{U^{\lambda}_{jq}T_{i_1\cdots i_r} = U^{\lambda}_{jq}[(T_{i_1\cdots i_r})^0 + G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0 + \cdots + G_{i_1i_2}\langle G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 + \cdots\rangle + \cdots]\}_{(\text{all } j)}. \tag{2.6}$$

The states $\{U^{\lambda}_{jq}T_{i_1\cdots i_r}\}_{(\text{all } j)}$ span the invariant subspace of row q, so Eq. (2.6) is of the same form as Eq. (2.3) and the rest of the analysis proceeds as before.

C. An Approach Using Young Symmetrizers

Unlike the idempotents used in Parts A and B, the usual (unmodified) Young symmetrizers do not all annul one another. However, we have not used this property, so the formal analysis for this part is the same as that for Part A except that the modified Young symmetrizers (PQ') of Part A are replaced by Young symmetrizers (PQ).

3. INNER PRODUCT STRUCTURE

In addition to Eq. (2.1), we have

$$\begin{aligned} T_{i_1\cdots i_r} &= (T_{i_1\cdots i_r})^0 \oplus (T^{3\cdots r}_{i_3\cdots i_r})^0 \oplus (T^{24\cdots r}_{i_2i_4\cdots i_r})^0 \oplus \cdots \\ &\oplus (T^{23\cdots r-1}_{i_2i_3\cdots i_{r-1}})^0 \oplus (T^{14\cdots r}_{i_1i_4\cdots i_r})^0 \oplus \cdots \oplus (T^{13\cdots r-1}_{i_1i_3\cdots i_{r-1}})^0 \\ &\oplus \cdots \oplus (T^{1\cdots r-2}_{i_1\cdots i_{r-2}})^0 \oplus \langle (T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0 \\ &\oplus (T^{3\cdots r,46\cdots r}_{i_4i_6\cdots i_r})^0 \oplus \cdots\rangle \oplus \langle (T^{24\cdots r,5\cdots r}_{i_5\cdots i_r})^0 \oplus \cdots\rangle \\ &\oplus \cdots \oplus \langle (T^{1\cdots r-2,3\cdots r-2}_{i_3\cdots i_{r-2}})^0 \oplus \cdots\rangle \oplus \cdots. \end{aligned} \tag{3.1}$$

The subspaces of this direct sum partition are invariant. What Eq. (2.1) does is embed each of the invariant subspaces in a Cartesian tensor space of rank r. Relative to the metric $(G)^r$ [r-fold Kronecker product of the rank-1 space metric tensor] of the Cartesian tensor space, the embedded subspaces of distinct traceless rank are mutually orthogonal.[3] This orthogonality is an important and general property of tensors decomposed by trace extraction. For the unitary symplectic group $Sp(n)$, trace extraction is carried out with the symplectic metric, and yet the embedded subspaces of distinct traceless rank are mutually orthogonal relative to either a symplectic metric or a Hermitian product.

Symmetrization refines the partition obtained by trace extraction, while symmetrized trace extraction refines the partition obtained by symmetrization. The same final partition is obtained in both cases, because otherwise the final partition would never be irreducible. It could be further shown that the same states are obtained in both cases, if the operations of trace extraction and symmetrization are shown to commute. However, it is not necessary to investigate this, because, although only symmetrized tensors are handled in symmetrized trace extraction calculations, Eqs. (2.1) show that actually the operation of trace extraction still precedes that of symmetrization. Thus the states obtained will be the same as those obtained for explicit Cartesian trace extraction followed by symmetrization. The discussion proceeds from the viewpoint of Cartesian trace extraction followed by symmetrization.

Carrying out symmetrization with unitary representation Wigner projection operators, we formally obtain subspaces of the form

$$\{U^{\lambda}_{jq}(T_{i_1\cdots i_r})^0\}, \quad \{U^{\lambda}_{jq}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})\}, \cdots.$$

If $(T_{i_1\cdots i_r})^0$ is orthogonal, then the subspaces $\{U^{\lambda}_{jp}(T_{i_1\cdots i_r})^0\}$ are fully orthogonal. (See note added in proof.) For the other subspaces, it is to be noted that, for example, we actually do not solve for $\{U^{\lambda}_{jq}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0\}$, but rather solve for certain linear

forms[12]:

$$\begin{aligned}\Gamma_{3\cdots r}\{U^{\lambda}_{jq}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0\}&\\ +\Gamma_{24\cdots r}\{U^{\lambda}_{jq}G_{i_1i_3}(T^{24\cdots r}_{i_2i_4\cdots i_r})^0\}+\cdots&\\ +\Gamma_{23\cdots r-1}\{U^{\lambda}_{jq}G_{i_1i_r}(T^{23\cdots r-1}_{i_2i_3\cdots i_{r-1}})^0\}&\\ +\Gamma_{14\cdots r}\{U^{\lambda}_{jq}G_{i_2i_3}(T^{14\cdots r}_{i_1i_4\cdots i_r})^0\}+\cdots&\\ +\Gamma_{13\cdots r-1}\{U^{\lambda}_{jq}G_{i_2i_r}(T^{13\cdots r-1}_{i_1i_3\cdots i_{r-1}})^0\}+\cdots&\\ +\Gamma_{12\cdots r-2}\{U^{\lambda}_{jq}G_{i_{r-1}i_r}(T^{12\cdots r-2}_{i_1i_2\cdots i_{r-2}})^0\}.&\quad(3.2)\end{aligned}$$

The form

$$\{U^{\lambda}_{jq}[\Gamma_{3\cdots r}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0+\cdots+\Gamma_{12\cdots r-2}G_{i_{r-1}i_r}(T^{12\cdots r-2}_{i_1i_2\cdots i_{r-2}})^0]\}$$

shows that here we get full orthogonality, if the tensors

$$\Gamma_{3\cdots r}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0+\cdots+\Gamma_{12\cdots r-2}G_{i_{r-1}i_r}(T^{12\cdots r-2}_{i_1i_2\cdots i_{r-2}})^0$$

are orthogonal.[13] A similar discussion applies to

$$\{U^{\lambda}_{jq}G_{i_1i_2}G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0\},\cdots,\text{ etc.}$$

If symmetrization is instead carried out with Young symmetrizers, then, using the fact that each set of equivalent representations $\{A^{\mu}\},\{B^{\mu}\},\cdots,\{D^{\mu}\}$ of $\mathcal{S}_r$ can be directly obtained in a form having the inner product structure[6] (the inner product structure for symmetrized non-Cartesian tensors is given as trace extraction generally yields such tensors),

$$\begin{aligned}\langle A^{\mu}_t(i_1,\cdots,i_r)\mid A^{\mu}_f(j_1,\cdots,j_r)\rangle&=\zeta(A,B)\\ \times\langle B^{\mu}_t(i_1,\cdots,i_r)\mid B^{\mu}_f(j_1,\cdots,j_r)\rangle&=\cdots\\ =\zeta(A,D)\langle D^{\mu}_t(i_1\cdots i_r)\mid D^{\mu}_f(j_1\cdots j_r)\rangle,&\quad(3.3)\end{aligned}$$

where ζ is always independent of state labels t and f and of indices $i_1\cdots i_r$ and $j_1\cdots j_r$, and

$$\begin{aligned}\langle A^{\mu}_t(i_1,\cdots,i_r)\mid A^{\mu}_f(j_1,\cdots,j_r)\rangle&=\lambda(A,B)\\ \times\langle A^{\mu}_t(i_1,\cdots,i_r)\mid B^{\mu}_f(j_1,\cdots,j_r)\rangle&=\cdots\\ =\lambda(A,D)\langle A^{\mu}_t(i_1,\cdots,i_r)\mid D^{\mu}_f(j_1,\cdots,j_r)\rangle,&\\ \lambda(B,A)\langle B^{\mu}_t(i_1,\cdots,i_r)\mid A^{\mu}_f(j_1,\cdots,j_r)\rangle&\\ =\langle B^{\mu}_t(i_1,\cdots,i_r)\mid B^{\mu}_f(j_1,\cdots,j_r)\rangle=\cdots&\\ =\lambda(B,D)\langle B^{\mu}_t(i_1,\cdots,i_r)\mid D^{\mu}_f(j_1,\cdots,j_r)\rangle,&\\ \vdots&\\ \lambda(D,A)\langle D^{\mu}_t(i_1,\cdots,i_r)\mid A^{\mu}_f(j_1,\cdots,j_r)\rangle&\\ =\lambda(D,B)\langle D^{\mu}_t(i_1,\cdots,i_r)\mid B^{\mu}_f(j_1,\cdots,j_r)\rangle=\cdots&\\ =\langle D^{\mu}_t(i_1,\cdots,i_r)\mid D^{\mu}_f(j_1,\cdots,j_r)\rangle,&\quad(3.4)\end{aligned}$$

where λ is always independent of state labels t and f and of indices $i_1\cdots i_r$ and $j_1\cdots j_r$, and may vanish. The subspaces $\{S_{im}(PQ)^{\lambda}_m(T_{i_1\cdots i_r})^0\}$, $(m=1,2,\cdots,N^{\lambda})$ can then be reasonably made fully orthogonal by using Eqs. (3.3) and (3.4). It is noted that, for example, rather than solving for $\{S_{im}(PQ)^{\lambda}_mG_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0\},\cdots$, we solve for certain linear forms:

$$\begin{aligned}\Lambda_{3\cdots r}\{S_{im}(PQ)^{\lambda}_mG_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0\}&\\ +\Lambda_{24\cdots r}\{S_{im}(PQ)^{\lambda}_mG_{i_1i_3}(T^{24\cdots r}_{i_2i_4\cdots i_r})^0\}&\\ +\cdots+\Lambda_{12\cdots r-1}\{S_{im}(PQ)^{\lambda}_mG_{i_1i_r}(T^{23\cdots r-1}_{i_2i_3\cdots i_{r-1}})^0\}&\\ +\Lambda_{14\cdots r}\{S_{im}(PQ)^{\lambda}_mG_{i_2i_3}(T^{14\cdots r}_{i_1i_4\cdots i_r})^0\}+\cdots&\\ +\Lambda_{13\cdots r-1}\{S_{im}(PQ)^{\lambda}_mG_{i_2i_r}(T^{13\cdots r-1}_{i_1i_3\cdots i_{r-1}})^0\}+\cdots&\\ +\Lambda_{12\cdots r-2}\{S_{im}(PQ)^{\lambda}_mG_{i_{r-1}i_r}(T^{12\cdots r-2}_{i_1i_2\cdots i_{r-2}})^0\}.&\quad(3.5)\end{aligned}$$

Writing expressions (3.5) as

$$\{S_{im}(PQ)^{\lambda}_m[\Lambda_{3\cdots r}G_{i_1i_2}(T^{3\cdots r}_{i_3\cdots i_r})^0+\cdots+\Lambda_{12\cdots r-2}G_{i_{r-1}i_r}(T^{12\cdots r-2}_{i_1i_2\cdots i_{r-2}})^0]\}$$

shows that the results of Eqs. (3.3) and (3.4) are applicable and this makes reasonable the task of orthogonalization. A similar discussion applies to $\{S_{im}(PQ)^{\lambda}_mG_{i_1i_2}G_{i_3i_4}(T^{3\cdots r,5\cdots r}_{i_5\cdots i_r})^0\},\cdots$, etc.

It should not be overlooked that the invariant subspaces due to trace extraction can have their embeddings described either by appropriately adding indices to the traceless tensors [e.g., Eqs. (2.3)] or else by contracting indices off of $T_{i_1\cdots i_r}$ [e.g., Eqs. (2.4) *et seq.*]. The important point is that the inner product structure results obtained apply for both descriptions of embedding.

4. EXAMPLES

The following examples should adequately illustrate a number of useful points.

[12] This may be clarified by a short example. For Young pattern (2, 1) and a symmetric metric tensor, Eqs. (2.6) appear as

$$U^{2,1}_{211,211}T_{ijk}=U^{2,1}_{211,211}[(T_{ijk})^0+G_{ij}T^3_k+G_{ik}T^2_j+G_{jk}T^1_i],$$
$$U^{2,1}_{121,211}T_{ijk}=U^{2,1}_{121,211}[(T_{ijk})^0+G_{ij}T^3_k+G_{ik}T^2_j+G_{jk}T^1_i],$$

and expand to

$$\begin{aligned}U^{2,1}_{211,211}T_{ijk}&=U^{2,1}_{211,211}(T_{ijk})^0+\tfrac{1}{3}G_{ij}(2T^3_k-T^2_k-T^1_k)\\ &\quad-\tfrac{1}{6}G_{jk}(2T^3_i-T^2_i-T^1_i)-\tfrac{1}{6}G_{ik}(2T^3_j-T^2_j-T^1_j),\\ U^{2,1}_{121,211}T_{ijk}&=U^{2,1}_{121,211}(T_{ijk})^0\\ &\quad+\frac{1}{2\sqrt{3}}G_{ik}(2T^3_j-T^2_j-T^1_j)-\frac{1}{2\sqrt{3}}G_{jk}(2T^3_i-T^2_i-T^1_i).\end{aligned}$$

We solve for $(2T^3_k-T^2_k-T^1_k),\cdots$, which is equivalent to solving for $G_{ij}(2T^3_k-T^2_k-T^1_k),\cdots$, as the metric tensor is assumed already given. Clearly the expressions $G_{ij}(2T^3_k-T^2_k-T^1_k),\cdots$, are each linearly dependent on $\{U^{2,1}_{p,211}G_{ij}T^3_k\},\{U^{2,1}_{p,211}G_{ik}T^2_j\}$ and $\{U^{2,1}_{p,211}G_{jk}T^1_i\}$.

[13] If only the original tensor $T_{i_1\cdots i_r}$ is orthogonal, then multiplets are mutually orthogonal if they are embedded in representations obtained with different values of q and/or λ in U^{λ}_{iq}.

Example 1

Here we investigate the embedding of the spherical bases of $SO(3)$ in $SU(3)$. We also use this example to illustrate some of the recoupling described in Ref. 6.

For the self-representation of $SU(3)$ with diagonal generators

$$H_1 = \begin{bmatrix} \frac{1}{2}\sqrt{3} & & \\ & -\frac{1}{2}\sqrt{3} & \\ & & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} \frac{1}{6} & & \\ & \frac{1}{6} & \\ & & -\frac{1}{3} \end{bmatrix},$$

the diagonal generator of $SO(3)$ is

$$\sqrt{3}\,H_1 + 3H_2 = \begin{bmatrix} 1 & & \\ & 0 & \\ & & -1 \end{bmatrix}.$$

With the other $SU(3)$ generators described as $E_{i,j}|j\rangle = (1/\sqrt{6})|i\rangle$, the $SO(3)$ raising and lowering operators are, respectively, $\sqrt{3}(E_{1,2} + E_{2,3})$ and $\sqrt{3}(E_{2,1} + E_{3,2})$. The symmetry (2, 1) provides two $SU(3)$ [$SO(3)$] octets. To irreducibly decompose each of these $SO(3)$ octets, we must use trace extraction. Denoting the spherical metric by M_{ij} and using $M_{ij} = M_{ji}$ and $\sum_j M_{ij}M^{jk} = \delta_i^k$, we have for the first octet

$$(PQ)_{ij,k}T_{ijk} = (PQ)_{ij,k}(T_{ijk})^0 + 2M_{ij}(T_k^3 - T_k^1)^0 - M_{ik}(T_j^3 - T_j^1)^0 - M_{jk}(T_i^3 - T_i^1)^0,$$

$$(PQ)_{ik,j}T_{ikj} = (PQ)_{ik,j}(T_{ijk})^0 - M_{ij}(T_k^3 - T_k^1)^0 + 2M_{ik}(T_j^3 - T_j^1)^0 - M_{jk}(T_i^3 - T_i^1)^0,$$

and for the second octet

$$(PQ)_{ij,k}T_{ikj} = (PQ)_{ij,k}(T_{ikj})^0 + 2M_{ij}(T_k^2 - T_k^1)^0 - M_{ik}(T_j^2 - T_j^1)^0 - M_{jk}(T_i^2 - T_i^1)^0,$$

$$(PQ)_{ik,j}T_{ijk} = (PQ)_{ik,j}(T_{ijk})^0 + 2M_{ik}(T_j^2 - T_j^1)^0 - M_{jk}(T_i^2 - T_i^1)^0 - M_{ij}(T_k^2 - T_k^1)^0,$$

where $(T_h^a)^0 - (T_h^b)^0 = (T_h^a - T_h^b)^0$ was used.

We can recouple these to obtain two octets which are orthogonal to each other. Adding corresponding states of the two octets above yields

$$(PQ)_{ij,k}(T_{ijk} + T_{ikj}) = (PQ)_{ij,k}(T_{ijk} + T_{ikj})^0 + 2M_{ij}(T_k^3 + T_k^2 - 2T_k^1)^0 - M_{ik}(T_j^3 + T_j^2 - 2T_j^1)^0 - M_{jk}(T_i^3 + T_i^2 - 2T_i^1)^0, \tag{4.1}$$

$$(PQ)_{ik,j}(T_{ikj} + T_{ijk}) = (PQ)_{ik,j}(T_{ikj} + T_{ijk})^0 - M_{ij}(T_k^3 + T_k^2 - 2T_k^1)^0 + 2M_{ik}(T_j^3 + T_j^2 - 2T_j^1)^0 - M_{jk}(T_i^3 + T_i^2 - 2T_i^1)^0, \tag{4.2}$$

where $(T_{ijk})^0 + (T_{ikj})^0 = (T_{ijk} + T_{ikj})^0$ was used.

Subtracting corresponding states of the two initial octets yields the second octet, which is easily seen to be orthogonal to that of Eqs. (4.1) and (4.2). If the symmetrized trace is viewed from $SO(3)$ rather than S_3, then Eqs. (4.1) and (4.2) describe one rather than three symmetrized traces. This is equivalent to viewing the symmetrized trace in the traditional way rather than by extended symmetrization. Equations (4.1) and (4.2) yield the trace equation

$$\sum_{ik} M^{ik}(PQ)_{ij,k}(T_{ijk} + T_{ikj}) = -2(T_j^3 + T_j^2 - 2T_j^1)^0.$$

Because the indices carrying the transformation properties also characterize the inner product structure within such a subspace and because the self-representation of $SO(3)$ is orthogonal, we know the resulting $SO(3)$ triplet is orthogonal. Using subscripts $+1$, 0, -1 to indicate weights and using $\cong$ because equality may be off by normalization, we have

$$A_{+1}^3 \equiv (T_1^3 + T_1^2 - 2T_1^1)^0 \cong (1/\sqrt{12}) \times (T_{113} + T_{131} + T_{221} + T_{212} - 2T_{311} - 2T_{122}),$$

$$A_0^3 \equiv (T_2^3 + T_2^2 - 2T_2^1)^0 \cong (1/\sqrt{12}) \times (-T_{123} - T_{132} - T_{312} - T_{321} + 2T_{213} + 2T_{231}),$$

$$A_{-1}^3 \equiv (T_3^3 + T_3^2 - 2T_3^1)^0 \cong (1/\sqrt{12}) \times (T_{223} + T_{232} + T_{331} + T_{313} - 2T_{133} - 2T_{332}).$$

By subtracting off the symmetrized trace, we obtain a quintet which is orthogonal, since all weights of $SO(3)$ are simple. The quintet is

$$A_{+2}^5 \equiv (PQ)_{11,2}(T_{112} + T_{121})^0 \cong (1\sqrt{6})(T_{112} + T_{121} - 2T_{211}),$$

$$A_{+1}^5 \equiv (PQ)_{11,3}(T_{113} + T_{131})^0 \cong (1/\sqrt{12})(T_{113} - 2T_{311} + T_{131} + 2T_{122} - T_{212} - T_{221}),$$

$$A_0^5 \equiv (PQ)_{12,3}(T_{123} + T_{132})^0 \cong \tfrac{1}{2}(T_{123} + T_{132} - T_{321} - T_{312}),$$

$$A_{-1}^5 \equiv (PQ)_{22,3}(T_{223} + T_{232})^0 = (1/\sqrt{12})(T_{223} - 2T_{322} + 2T_{133} - T_{331} - T_{313} + T_{232}),$$

$$A_{-2}^5 \equiv (PQ)_{23,3}(T_{233} + T_{233})^0 = (1/\sqrt{6})(2T_{233} - T_{323} - T_{332}).$$

The (tensor) phase of the states is such that all elements of the triplet and quintet raising operators are real and positive. Independent state selection is aided by using the traceless conditions. Describing a second-rank $SO(3)$ quintet $\{Y^0\}$ and singlet $\{T\}$ with

the notation

$$Y(ij)^0 = (1/\sqrt{2})(T_{ij} + T_{ji}) \quad (i \neq j),$$
$$Y(11)^0 = (1/\sqrt{6})(2T_{11} + T_{22} - T_{33}),$$
$$Y(22)^0 = (1/\sqrt{6})(2T_{22} + T_{11} + T_{33}),$$
$$T = (1/\sqrt{3})(-T_{11} + T_{22} - T_{33}),$$

we find

$$A^3_{+1} = (1/\sqrt{12})[\sqrt{2}\,T_1 Y(13)^0 - (\tfrac{2}{3})T_3\langle\sqrt{6}\,Y(11)^0 - \sqrt{3}\,T\rangle - (\tfrac{2}{3})T_1\langle\sqrt{6}\,Y(22)^0 + \sqrt{3}\,T\rangle + \sqrt{2}\,T_2 Y(12)^0].$$

It turns out that this is as close as we can get to writing A^3_{+1} in terms of simply-coupled irreducible representations (a form involving the product of only two lower-rank irreducible representations). This illustrates how trace extraction can conflict with simple coupling. This conflict occurs because we want simply to couple irreducible, rather than merely symmetrized, tensors. Denoting the second-rank triplet by $Z(ij)^0$, we find that the triplet

$$B^3_{+1} \equiv (T^3_1 - T^2_1)^0 \cong \tfrac{1}{2}(T_{113} - T_{131} + T_{221} - T_{212}),$$
$$B^3_0 \equiv (T^3_2 - T^2_2)^0 \cong \tfrac{1}{2}(T_{123} - T_{132} + T_{321} - T_{312}),$$
$$B^3_{-1} \equiv (T^3_1 - T^2_1)^0 \cong \tfrac{1}{2}(T_{223} - T_{232} + T_{331} - T_{313})$$

has the simply coupled form

$$B^3_{+1} = (1/\sqrt{2})[T_1 Z(13)^0 + T_2 Z(21)^0],$$
$$B^3_0 = (1/\sqrt{3})[T_1 Z(23)^0 - T_3 Z(21)^0],$$
$$B^3_{-1} = (1/\sqrt{2})[T_2 Z(23)^0 + T_3 Z(31)^0].$$

This simple coupling resulted because $Z(ij) = Z(ij)^0$.

Example 2

This example illustrates the minimal coupling obtained by extracting symmetrized traces in steps. Using $G_{ij} = \delta_{ij}$ and writing

$$T_{hijk} = (T_{hijk})^0 + \delta_{hi}[(T^{34}_{jk})^0 + \delta_{jk}T^{34}] + \delta_{hj}[(T^{24}_{ik})^0 + \delta_{ik}T^{24}] + \delta_{hk}[(T^{23}_{ij})^0 + \delta_{ij}T^{23}] + \delta_{ij}[(T^{14}_{hk})^0 + \delta_{hk}T^{14}] + \delta_{ik}[(T^{13}_{hj})^0 + \delta_{hj}T^{13}] + \delta_{jk}[(T^{12}_{hi})^0 + \delta_{hi}T^{12}],$$

we find

$$(PQ)_{hi,jk}T_{hijk} = (PQ)_{hi,jk}(T_{hijk})^0 + 2[\delta_{hi}\psi_{jk} - \delta_{ij}\psi_{hk} - \delta_{jh}\psi_{ik} - \delta_{ki}\psi_{hj} - \delta_{kh}\psi_{ij} + \delta_{jk}\psi_{hi}],$$
$$(PQ)_{hj,ik}T_{hjik} = (PQ)_{jh,ik}(T_{hjik})^0 + 2[\delta_{hj}\psi_{ik} - \delta_{hi}\psi_{jk} - \delta_{hk}\psi_{ij} - \delta_{ij}\psi_{hk} - \delta_{jk}\psi_{ih} + \delta_{ik}\psi_{hj}],$$

where, using $[ab]^{mn} \equiv [(T^{mn}_{ab})^0 + \delta_{ab}T^{mn}]$, we wrote

$$\psi_{ij} = [ij]^{34} + [ji]^{34} - [ij]^{23} - [ji]^{23} - [ij]^{14} - [ji]^{14} + [ij]^{12} + [ji]^{12}.$$

Using n-dimensional indices we obtain the symmetrized trace equation

$$\sum_{ih} \delta^{ih}(PQ)_{hi,jk}T_{hijk} = (2n - 8)\psi_{jk}.$$

By subtraction we evaluate the symmetrized traceless tensor $\{(PQ)_{hi,jk}(T_{hijk})^0, (PQ)_{hj,ik}(T_{hjik})^0\}$. It is necessary to carry out one further step of symmetrized trace extraction. For Young pattern (2) we have

$$(PQ)_{jk,}\frac{1}{(2n-8)}\sum_{ih}\delta^{ih}(PQ)_{hi,jk}T_{hijk} = (PQ)_{jk,}(T^{34}_{jk} + T^{34}_{kj} - T^{23}_{jk} - T^{23}_{kj} - T^{14}_{jk} - T^{14}_{kj} + T^{12}_{jk} + T^{12}_{kj})^0 + 2\delta^{jk}(T^{34} - T^{23} - T^{14} + T^{12}),$$

which yields the trace equation

$$\frac{1}{n(2n-8)}\sum_{jk}\delta^{jk}(PQ)_{jk,}\sum_{ih}\delta^{ih}(PQ)_{hi,jk}T_{hijk} = 2(T^{34} - T^{23} - T^{14} + T^{12}).$$

By subtraction we obtain the symmetrized traceless tensor

$$(PQ)_{jk,}(T^{34}_{jk} + T^{34}_{kj} - T^{23}_{jk} - T^{23}_{kj} - T^{14}_{jk} - T^{14}_{kj} + T^{12}_{jk} + T^{12}_{kj})^0 = \frac{1}{(2n-8)}\left(1 - \frac{1}{n}\sum_{jk}\delta^{ik}\right)(PQ)_{jk,}\sum_{ih}\delta^{ih}(PQ)_{hi,jk}T_{hijk}.$$

This completes the symmetrized trace extractions because for Young pattern (1, 1) we find

$$(PQ)_{j,k}\sum_{ih}\delta^{ih}(PQ)_{hi,jk}T_{hijk} = 0.$$

If instead we had removed both of these traces initially, then we would have obtained a pair of coupled equations.

5. DISCUSSION

Both the initial symmetrization and trace extraction are achieved by solving the symmetrized trace equations. The initial symmetrization accomplishes two objectives: first, it gives the traceless tensors of rank r (the rank of the original tensor) their final symmetry; and second, it decouples trace extraction into symmetrized trace extractions. For each lower-rank (less than r) traceless tensor a second symmetrization (without trace extraction) yields the final symmetry partition. Extracting the different rank traces in separate steps, each accompanied by symmetrization, makes maximal use of the decoupling provided by symmetrization.

The problem of obtaining an orthogonal decomposition with symmetrized trace extraction based on either unitary representation Wigner projection operators or Young symmetrizers was described in detail.

The visibly simple symmetry properties due to the use of Young symmetrizers simplify some calculations. This is especially true when using $PQ(QP)$ symmetrizers with an antisymmetric (symmetric) metric tensor.

Note Added in Proof: To remove an oversight in the proof of Eq. (3.2) of Ref. 6, we reconsider the result

$$\langle T_{i_1 \cdots i_r} \mid (U^{\mu}_{gv})^{\dagger} T_{j_1 \cdots j_r} \rangle = \delta_{gv} \langle T_{i_1 \cdots i_r} \mid T_{j_1 \cdots j_r} \rangle.$$

The reason for using $i_1 \cdots i_r$ and $j_1 \cdots j_r$, rather than $i_1 \cdots i_r$ on both sides, is to allow the two sides of the inner product to have different *sets* of index values. (If the sets of index values coincide, then we must have $i_1 = j_1, \cdots, i_r = j_r$.) The proof in Ref. 6 used the fact that if the tensor was orthogonal; and if all indices have different values, then we get zero unless $g = v$ because then only

$$\langle T_{i_1 \cdots i_r} \mid [\mathfrak{D}^{\mu}_{e}]^{*}_{gv} L_e T_{j_1 \cdots j_r} \rangle$$

can contribute. If the rank exceeds the space dimension, then no component $T_{i_1 \cdots i_r}(T_{j_1 \cdots j_r})$ appears with all indices having different values. To extend the proof to this situation we note that we can embed such a tensor in another tensor of the same rank but on a space of arbitrary dimension n'. When $n' \geq r$, the theorem clearly holds for the embedding tensor. The operations of S_r are here fundamentally on index position rather than index value and so are not affected by the embedding. Hence the theorem must also hold for the embedded tensor.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Related First Integral Theorem: A Method for Obtaining Conservation Laws of Dynamical Systems with Geodesic Trajectories in Riemannian Spaces Admitting Symmetries*

GERALD H. KATZIN AND JACK LEVINE
Departments of Physics and Mathematics, North Carolina State University, Raleigh, North Carolina

(Received 5 June 1967)

In this paper we develop in detail a unified method, referred to as the Related First Integral Theorem, for obtaining "derived" first integrals (i.e., constants of the motion) of mass-pole test particles with geodesic trajectories in a Riemannian space. By this method, which is based upon a process of Lie differentiation, additional conservation laws in the form of mth order first integrals can be generated from a given mth order first integral (conservation law), provided the space admits symmetries in the form of continuous groups of projective collineations (which include affine collineations and motions as special cases). We give in tensor form a reformulation of the well-known Poisson's theorem on constants of the motion for particles with geodesic trajectories. We then show for this class of trajectories that, as a method for generating mth order first integrals from a given mth order first integral, Poisson's theorem is a special case of the Related First Integral Theorem. It is also shown that dependency relations between generically related first integrals obtained by the Related First Integral Theorem are expressible in terms of the structure constants of the underlying continuous group of symmetries.

1. INTRODUCTION

It is generally recognized that the knowledge of conservation laws is of fundamental importance in the physical description of nature. It has also been observed that the existence of certain geometric symmetries described by continuous groups of motions or collineations lead to conservation laws expressible in the form of first integrals (i.e., constants of motion) of a dynamical system.

In this paper, we wish to consider the mth order first integrals (mFI) of dynamical systems whose trajectories are geodesics in a Riemannian space V_n. This class of trajectories is of particular importance in the general theory of relativity in that it includes the description of the motion of mass-pole test particles.

The geodesics are solutions of the equations[1]

$$\frac{Dp^i}{ds} \equiv \frac{dp^i}{ds} + \begin{Bmatrix} i \\ jk \end{Bmatrix} p^j p^k = 0, \quad \left(p^i \equiv \frac{dx^i}{ds} \right), \quad (1.1)$$

where the parameter s represents arclength and $\{^{i}_{jk}\}$ are the Christoffel symbols. These geodesics will admit

* This work was supported by a National Science Foundation Grant No. GP 6876.

[1] Latin indices take on values 1, 2, $\cdots$, n; and Greek indices take on values 1, 2, $\cdots$, r. The Einstein summation convention is used.

The visibly simple symmetry properties due to the use of Young symmetrizers simplify some calculations. This is especially true when using $PQ(QP)$ symmetrizers with an antisymmetric (symmetric) metric tensor.

Note Added in Proof: To remove an oversight in the proof of Eq. (3.2) of Ref. 6, we reconsider the result

$$\langle T_{i_1 \cdots i_r} \mid (U^{\mu}_{gv})^{\dagger} T_{j_1 \cdots j_r} \rangle = \delta_{gv} \langle T_{i_1 \cdots i_r} \mid T_{j_1 \cdots j_r} \rangle.$$

The reason for using $i_1 \cdots i_r$ and $j_1 \cdots j_r$, rather than $i_1 \cdots i_r$ on both sides, is to allow the two sides of the inner product to have different *sets* of index values. (If the sets of index values coincide, then we must have $i_1 = j_1, \cdots, i_r = j_r$.) The proof in Ref. 6 used the fact that if the tensor was orthogonal; and if all indices have different values, then we get zero unless $g = v$ because then only

$$\langle T_{i_1 \cdots i_r} \mid [\mathfrak{D}^{\mu}_{e}]^{*}_{gv} L_e T_{j_1 \cdots j_r} \rangle$$

can contribute. If the rank exceeds the space dimension, then no component $T_{i_1 \cdots i_r}(T_{j_1 \cdots j_r})$ appears with all indices having different values. To extend the proof to this situation we note that we can embed such a tensor in another tensor of the same rank but on a space of arbitrary dimension n'. When $n' \geq r$, the theorem clearly holds for the embedding tensor. The operations of S_r are here fundamentally on index position rather than index value and so are not affected by the embedding. Hence the theorem must also hold for the embedded tensor.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Related First Integral Theorem: A Method for Obtaining Conservation Laws of Dynamical Systems with Geodesic Trajectories in Riemannian Spaces Admitting Symmetries*

GERALD H. KATZIN AND JACK LEVINE
Departments of Physics and Mathematics, North Carolina State University, Raleigh, North Carolina

(Received 5 June 1967)

In this paper we develop in detail a unified method, referred to as the Related First Integral Theorem, for obtaining "derived" first integrals (i.e., constants of the motion) of mass-pole test particles with geodesic trajectories in a Riemannian space. By this method, which is based upon a process of Lie differentiation, additional conservation laws in the form of mth order first integrals can be generated from a given mth order first integral (conservation law), provided the space admits symmetries in the form of continuous groups of projective collineations (which include affine collineations and motions as special cases). We give in tensor form a reformulation of the well-known Poisson's theorem on constants of the motion for particles with geodesic trajectories. We then show for this class of trajectories that, as a method for generating mth order first integrals from a given mth order first integral, Poisson's theorem is a special case of the Related First Integral Theorem. It is also shown that dependency relations between generically related first integrals obtained by the Related First Integral Theorem are expressible in terms of the structure constants of the underlying continuous group of symmetries.

1. INTRODUCTION

It is generally recognized that the knowledge of conservation laws is of fundamental importance in the physical description of nature. It has also been observed that the existence of certain geometric symmetries described by continuous groups of motions or collineations lead to conservation laws expressible in the form of first integrals (i.e., constants of motion) of a dynamical system.

In this paper, we wish to consider the mth order first integrals (mFI) of dynamical systems whose trajectories are geodesics in a Riemannian space V_n. This class of trajectories is of particular importance in the general theory of relativity in that it includes the description of the motion of mass-pole test particles.

The geodesics are solutions of the equations[1]

$$\frac{Dp^i}{ds} \equiv \frac{dp^i}{ds} + \begin{Bmatrix} i \\ jk \end{Bmatrix} p^j p^k = 0, \quad \left(p^i \equiv \frac{dx^i}{ds} \right), \quad (1.1)$$

where the parameter s represents arclength and $\{^{i}_{jk}\}$ are the Christoffel symbols. These geodesics will admit

* This work was supported by a National Science Foundation Grant No. GP 6876.

[1] Latin indices take on values 1, 2, $\cdots$, n; and Greek indices take on values 1, 2, $\cdots$, r. The Einstein summation convention is used.

an mth order first integral (mFI)

$$A_{i_1\cdots i_m}p^{i_1}\cdots p^{i_m} = k = \text{const} \tag{1.2}$$

(where $A_{i_1\cdots i_m}$ is completely symmetric on all indices), if and only if[2]

$$P\{A_{i_1 i_2\cdots i_m;i_{m+1}}\} = 0 \tag{1.3}$$

(see Ref. 3, p. 84).

It is well known that if a V_n admits symmetries in the form of motions, then the geodesics admit linear first integrals (LFI) and conversely (see Ref. 4, p. 128 and Chap. VI. For a review of the physical interpretation of these LFI see Ref. 5.)

In a recent physics paper largely devoted to the study of particle conservation laws in general relativity, Davis and Moss[6] pointed out that infinitesimal projective and affine collineations admitted by a V_n lead to quadratic first integrals (QFI) of the geodesics. They thus established that geometric symmetries (other than motions) in the form of collineations lead to additional conservation laws.[7]

In the present paper we extend this work of Davis and Moss[6] by means of the Related First Integral Theorem (RIT) to be proved in Secs. 3 and 4. This theorem provides a unifying method for obtaining conservation laws associated with symmetries defined by collineations, and in addition gives a geometric explanation of why the existence of certain mFI is to be associated with projective and affine collineations.

It is shown that if a V_n admits a group of projective or affine collineations, or motions, and if, in addition, the space admits an mFI, then, in general, the geodesics admit additional mFI which can be derived by a process of Lie differentiation with respect to the generating vectors of the stated symmetry group.

By applying the RIT to generate QFI in a V_n, we obtain the QFI found by Davis and Moss.[6] Moreover, by use of this theorem, additional QFI can in general be generated, showing that more than one linearly independent QFI may be associated with each collineation. In this way, by systematic application of the RIT in a V_n, a sequence of mFI for any m can, in general, be generated (which we anticipate will lead to new conservation laws of physical significance).

As a means of examining the linear independence of the sequence of derived integrals generated by the RIT, we derive in Sec. 5 dependency relations involving the structure constants of the underlying group of symmetries.

An earlier well-known method for obtaining first integrals is given by Poisson's theorem on constants of the motion. We derive in tensor form a restatement of Poisson's theorem for mass particles with geodesic trajectories. We then compare Poisson's theorem with the RIT as methods for generating mFI from a given mFI, and show that (for geodesic trajectories) Poisson's theorem may be considered as a special case of the RIT.

Although in this paper we are mainly concerned with the mathematical development of the RIT, we give several examples of physical interest.

To achieve mathematical generality, much of the work presented in this paper is based on affine spaces A_n which include the Riemannian spaces V_n as special cases. For an A_n, Eq. (1.1) defines the "paths" of the space, provided the Christoffel symbols, $\{^{i}_{jk}\}$ are replaced by the symmetric components of affine connection Γ^i_{jk} and s is taken to be a path parameter.

2. COLLINEATIONS AND MOTIONS

By definition, a projective collineation is a point transformation which transforms paths into paths. The condition that the point transformation

$$\bar{x}^i = x^i + \xi^i(x)\delta t \tag{2.1}$$

(where δt is considered an infinitesimal) defines an infinitesimal projective collineation is expressible in terms of the Lie derivative[8,9] with respect to the vector ξ^i in the form[10]

$$\pounds_\xi\left(\frac{Dp^i}{ds}\right) = 0. \tag{2.2}$$

It follows from (2.2) (see Ref. 10, p. 454 or Ref. 3, p. 126) that an A_n admits an r-parameter group of proper projective collineations if

$$\xi^h_{(\alpha);ij} = \xi^k_{(\alpha)}B^h_{ijk} + \delta^h_j\phi_{(\alpha)i} + \delta^h_i\phi_{(\alpha)j}, \tag{2.3}$$

where the r linearly independent vectors $\xi^i_{(\alpha)}$ generate the group, the $\phi_{(\alpha)i}$ is a set of r vectors at least one

[2] Covariant differentiation is indicated by a semicolon (;), and partial differentiation by a comma (,). The symbol $P\{\ \}$ indicates the sum of the terms obtained by forming all cyclic permutations of the indices that are not summed completely within the braces. For example,

$$P\{A_{ij;k}\}\lambda^j \equiv (A_{ij;k} + A_{jk;i} + A_{ki;j})\lambda^j,$$

whereas

$$P\{A_{ij;k}\lambda^j\} \equiv (A_{ij;k} + A_{kj;i})\lambda^j.$$

[3] L. P. Eisenhart, *Non-Riemannian Geometry* (American Mathematical Society Publications, No. 8, New York, 1927).

[4] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1926).

[5] W. R. Davis and G. H. Katzin, Am. J. Phys, **30**, 750 (1962).

[6] W. R. Davis and M. K. Moss, Nuovo Cimento **38**, 1558 (1965).

[7] They also indicated that, under certain circumstances, projective collineations would lead to higher-order first integrals.

[8] K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland Publishing Company, Amsterdam, 1957).

[9] J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954), 2nd ed.

[10] K. Yano, K. Takano, and Y. Tomonaga, Japan. J. Math. **19**, 433 (1948).

of which is not zero, and the B^h_{ijk} are the curvature tensor components (see Ref. 3, p. 8).

For the case in which all $\phi_{(\alpha)i} = 0$ in (2.3), the corresponding set of vectors $\xi^i_{(\alpha)}$ is said to define an r-parameter group of affine collineations.

A V_n is said to admit an r-parameter group of motions generated by the r vectors $\xi^i_{(\alpha)}$ if Killing's equations $\xi_{(\alpha)i;j} + \xi_{(\alpha)j;i} = 0$ are satisfied where $\xi_{i(\alpha)} \equiv g_{ij}\xi^j_{(\alpha)}$ (see Ref. 4, p. 234). In this case, it follows that (2.3) with $\phi_{(\alpha)i} = 0$ must be satisfied as integrability conditions (see Ref. 4, p. 237).

An affine collineation will be called proper if it is not a motion. (This, of course, applies only to V_n spaces.)

Notation: We denote by PC_r, AC_r, and M_r an r-parameter group of proper projective collineations, proper affine collineations, and motions, respectively, generated by the r vectors $\xi^i_{(\alpha)}$, $(\alpha = 1, 2, \cdots, r)$.

An affine space A_n will be called an A^*_n space if its affine Ricci tensor is symmetric, i.e., $B_{ij} = B_{ji}$, where $B_{ij} \equiv B^h_{ijh}$. We are mainly concerned with collineations in A^*_n spaces (which include V_n spaces as special cases).

Assume now that an A^*_n admits a PC_r. It follows from Ref. 3 [Eq. (46.8)] and the symmetry of B_{ij}, that vectors $\phi_{(\alpha)i}$ in Eq. (2.3) are gradients defined by $\phi_{(\alpha)i} \equiv \phi_{(\alpha);i}$ with scalar $\phi_{(\alpha)}$ defined by

$$\phi_{(\alpha)} \equiv (n+1)^{-1}\xi^i_{(\alpha);i}. \tag{2.4}$$

It is known that, under projective collineations in A^*_n spaces, the pathparameters s and $\bar{s}_{(\alpha)}$ along corresponding paths are related by (Ref. 3, p. 107)

$$d\bar{s}_{(\alpha)} = e^{2\phi_{(\alpha)}+c}\, ds, \tag{2.5}$$

where c is an arbitrary constant.

In the case of an infinitesimal projective collineation, we may write

$$d\bar{s}_{(\alpha)} = ds + (2\phi_{(\alpha)} + c)\, ds\delta t. \tag{2.6}$$

3. DERIVATION OF THE RELATED FIRST INTEGRAL THEOREM

Let us assume that the paths (1.1) of an A^*_n admit an mFI (1.2), and in addition the A^*_n admits a PC_r.

As a result of a collineation transformation, the mFI (1.2) is "dragged along" in the sense of Schouten (see Ref. 9, p. 102) as paths transform into paths. The mFI is deformed by this process and takes the form

$$(\bar{A}_{i_1\cdots i_m}(x))_\alpha \bar{p}^{i_1}_{(\alpha)} \cdots \bar{p}^{i_m}_{(\alpha)} = \bar{k}_\alpha, \quad \left(\bar{p}^i_{(\alpha)} \equiv \frac{dx^i}{d\bar{s}_{(\alpha)}}\right), \tag{3.1}$$

where

$$(\bar{A}_{i_1\cdots i_m})_\alpha \bar{p}^{i_1}_{(\alpha)} \cdots \bar{p}^{i_m}_{(\alpha)} = A_{i_1\cdots i_m}p^{i_1}\cdots p^{i_m} + \mathfrak{L}_\alpha(A_{i_1\cdots i_m}p^{i_1}\cdots p^{i_m})\delta t \tag{3.2}$$

and

$$\bar{k}_\alpha \equiv k + (\mathfrak{L}_\alpha k)\delta t, \tag{3.3}$$

where $\mathfrak{L}_\alpha \equiv \mathfrak{L}_{\xi(\alpha)}$.

It therefore follows by Eqs. (3.2), (3.3), and (1.2) that the deformed mFI (3.1) can be expressed as

$$\mathfrak{L}_\alpha(A_{i_1\cdots i_m}p^{i_1}\cdots p^{i_m}) = \mathfrak{L}_\alpha k. \tag{3.4}$$

If Eq. (3.4) is expanded, there results

$$(\mathfrak{L}_\alpha A_{i_1\cdots i_m})p^{i_1}\cdots p^{i_m} - mA_{i_1\cdots i_m}p^{i_1}\cdots p^{i_{m-1}} \times (\mathfrak{L}_\alpha p^{i_m}) = \mathfrak{L}_\alpha k. \tag{3.5}$$

From Ref. 8, p. 89, we know that

$$\mathfrak{L}_\alpha p^i = -p^i\frac{\mathfrak{L}_\alpha\, ds}{ds}. \tag{3.6}$$

From the definition

$$\mathfrak{L}_\alpha\, ds = \lim_{\delta t\to 0}\frac{d\bar{s}_\alpha - ds}{\delta t}$$

and Eq. (2.6), we find that

$$\mathfrak{L}_\alpha\, ds = (2\phi_{(\alpha)} + c)\, ds. \tag{3.7}$$

Substitution from Eqs. (3.6) and (3.7) into Eq. (3.5) and use of (1.2) gives

$$[(\mathfrak{L}_\alpha A_{i_1\cdots i_m}) - 2m\phi_{(\alpha)}A_{i_1\cdots i_m}]p^{i_1}\cdots p^{i_m} = cmk + \mathfrak{L}_\alpha k \equiv k_{(\alpha)}. \tag{3.8}$$

Since $(\mathfrak{L}_\alpha k)\delta t$ represents the change in k induced by an infinitesimal point transformation of the form (2.1), we may consider $\mathfrak{L}_\alpha k$ expressible in the form $\mathfrak{L}_\alpha k = f_{(\alpha)}(k)$, and hence $\mathfrak{L}_\alpha k$ may be considered as a constant along a path. Thus, the right side of (3.8) is constant along a path, and hence the left side of (3.8) defines an mFI of the paths which is generically related to the original mFI (1.2).

In the following section, we give an alternate proof that the left side of (3.8) defines an mFI.

4. ALTERNATE PROOF

Define

$$A_{(\alpha)i_1\cdots i_m} \equiv \mathfrak{L}_\alpha A_{i_1\cdots i_m} - 2m\phi_{(\alpha)}A_{i_1\cdots i_m}. \tag{4.1}$$

We show that

$$P\{A_{(\alpha)i_1\cdots i_m;\, i_{m+1}}\} = 0, \tag{4.2}$$

which, by Eq. (1.3), will prove that

$$A_{(\alpha)i_1\cdots i_m}p^{i_1}\cdots p^{i_m} = \text{const} \tag{4.3}$$

is an mFI of the paths (1.1) of an A^*_n.

From (4.1) and the left side of (4.2), we obtain, by

use of (1.3),

$$P\{A_{(\alpha)i_1\cdots i_m;i_{m+1}}\} = P\Big\{A_{i_1\cdots i_m;ji_{m+1}}\xi^j_{(\alpha)} + A_{i_1\cdots i_m;j}\xi^j_{(\alpha);i_{m+1}} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m;i_{m+1}}\xi^j_{(\alpha);i_h} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m}\xi^j_{(\alpha);i_hi_{m+1}} - 2m\phi_{(\alpha);i_{m+1}}A_{i_1\cdots i_m}\Big\}. \quad (4.4)$$

In Eq. (4.4), we note the right-hand side contains a sum of $(m+1)^2$ terms which add to zero by (1.3). As a consequence, we can write (4.4) in the form

$$P\{A_{(\alpha)i_1\cdots i_m;i_{m+1}}\} = P\Big\{A_{i_1\cdots i_m;ji_{m+1}}\xi^j_{(\alpha)} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m}\xi^j_{(\alpha)i_hi_{m+1}} - 2m\phi_{(\alpha);i_{m+1}}A_{i_1\cdots i_m}\Big\}. \quad (4.5)$$

Employing the Ricci identity

$$A_{i_1\cdots i_m;ji_{m+1}} = A_{i_1\cdots i_m;i_{m+1}j} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ki_{h+1}\cdots i_m}B^k_{i_hji_{m+1}}, \quad (4.6)$$

we substitute for the first term on the right-hand side of (4.5) and obtain

$$P\{A_{(\alpha)i_1\cdots i_m;i_{m+1}}\} = P\Big\{A_{i_1\cdots i_m;i_{m+1}j}\xi^j_{(\alpha)} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m} \times (\xi^k_{(\alpha)}B^j_{i_hki_{m+1}} + \xi^j_{(\alpha);i_hi_{m+1}}) - 2m\phi_{(\alpha);i_{m+1}}A_{i_1\cdots i_m}\Big\}. \quad (4.7)$$

The first term on the right-hand side of (4.7) vanishes by (1.3), and by (2.3) the second covariant derivatives of $\xi^j_{(\alpha)}$ can be eliminated from the remaining terms. The resulting expression can be reduced to

$$P\{A_{(\alpha)i_1\cdots i_m;i_{m+1}}\} = P\Big\{mA_{i_1\cdots i_m}\phi_{(\alpha);i_{m+1}} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}i_{m+1}i_{h+1}\cdots i_m}\phi_{(\alpha);i_h} - 2m\phi_{(\alpha);i_{m+1}}A_{i_1\cdots i_m}\Big\}. \quad (4.8)$$

Finally, if the right-hand side of (4.8) be expanded by permuting indices as indicated, we obtain (4.2).

We may thus state the theorem to follow.

Theorem 4.1 (Related First Integral Theorem): If an affine space A_n^* admits a PC_r (group of proper projective collineations), and if the paths of the A_n^* admit an mFI (mth order first integral) (1.2), then the paths will admit the r additional mFI

$$A_{(\alpha)i_1\cdots i_m}p^{i_1}\cdots p^{i_m} = \text{const}, \quad (\alpha = 1, \cdots, r), \quad (4.9)$$

where $A_{(\alpha)i_1\cdots i_m}$ is defined by (4.1).

Corollary 4.1: If a space A_n admits an AC_r (group of proper affine collineations) and an mFI (1.2), then the A_n will admit the r additional mFI (4.9), where $A_{(\alpha)i_1\cdots i_m}$ is defined by (4.1) with all $\phi_{(\alpha)} = 0$. Theorem 4.1 also holds for a V_n admitting a PC_r, an AC_r, or an M_r (motions).

In Corollary 4.1, with the case of a V_n admitting a PC_r and where $A_{ij} = g_{ij}$ (the components of the metric tensor) and $m = 2$, the quadratic first integrals $(\pounds_\alpha g_{ij} - 4\phi_{(\alpha)}g_{ij})p^ip^j = \text{const}$ [given by (4.1)] have been obtained by Davis and Moss[6] in a slightly different notation.

Definition: Integrals (4.9) will be called "first derived integrals" with respect to the integral (1.2). In a like manner kth derived first integrals can be obtained with respect to $(k-1)$-derived integrals, for $k = 1, 2, 3, \cdots$. Dependency relations between such derived integrals are considered in the section to follow. Examples of such integrals are given in Sec. 9.

5. DEPENDENCY RELATIONS BETWEEN DERIVED INTEGRALS

Let an A_n^* satisfy the conditions of Theorem 4.1 and define the second derived integral coefficient

$$A_{(\beta\alpha)i_1\cdots i_m} \equiv \pounds_\beta A_{(\alpha)i_1\cdots i_m} - 2m\phi_{(\beta)}A_{(\alpha)i_1\cdots i_m} \quad (5.1)$$

with $A_{(\alpha)i_1\cdots i_m}$ defined by (4.1). It follows from Eq. (4.1) that

$$A_{(\beta\alpha)i_1\cdots i_m} - A_{(\alpha\beta)i_1\cdots i_m} = (\pounds_\beta\pounds_\alpha - \pounds_\alpha\pounds_\beta)A_{i_1\cdots i_m} - 2m(\pounds_\beta\phi_{(\alpha)} - \pounds_\alpha\phi_{(\beta)})A_{i_1\cdots i_m}. \quad (5.2)$$

From Ref. 8, pp. 29 and 35, we can write

$$\pounds_\beta\pounds_\alpha - \pounds_\alpha\pounds_\beta = C^\gamma_{\beta\alpha}\pounds_\gamma, \quad (5.3)$$

where $C^\gamma_{\beta\alpha}$ are the structure constants of the PC_r admitted by the A_n^*.

By the use of Eqs. (2.4), (5.3), and the Ricci identity, we can write Eq. (5.2) in the form

$$A_{(\beta\alpha)i_1\cdots i_m} - A_{(\alpha\beta)i_1\cdots i_m} = C^\gamma_{\beta\alpha}\pounds_\gamma A_{i_1\cdots i_m} - \frac{2m}{n+1}(\pounds_\beta\xi^j_{(\alpha)})_{;j}A_{i_1\cdots i_m}. \quad (5.4)$$

It follows (Ref. 8, p. 29) that

$$A_{(\beta\alpha)i_1\cdots i_m} - A_{(\alpha\beta)i_1\cdots i_m} = C^\gamma_{\beta\alpha}[\pounds_\gamma A_{i_1\cdots i_m} - 2m\phi_{(\gamma)}A_{i_1\cdots i_m}] = C^\gamma_{\beta\alpha}A_{(\gamma)i_1\cdots i_m}. \quad (5.5)$$

If the extreme sides of Eq. (5.5) be multiplied by $p^{i_1} \cdots p^{i_m}$ (and summed), the resulting equation may be expressed in the form

$$k_{(\beta\alpha)} - k_{(\alpha\beta)} = C^{\gamma}_{\beta\alpha} k_{(\gamma)}, \tag{5.6}$$

where $k_{(\alpha)}$, $k_{(\alpha\beta)}$ are the respective constants of motion defined by the mFI

$$A_{(\alpha)i_1 \cdots i_m} p^{i_1} \cdots p^{i_m}, \quad A_{(\alpha\beta)i_1 \cdots i_m} p^{i_1} \cdots p^{i_m}.$$

We can thus state:

Theorem 5.1: If an A_n^* admits a PC_r and an mFI (1.2), then the first and second derived integrals based on (1.2) satisfy a linear dependence relation of the form (5.5), and the corresponding constants of the motion satisfy (5.6).

Corollary 5.1: If an A_n admits an AC_r and an mFI, then dependency relations (5.5) and (5.6) hold with $\phi_{(\alpha)} = 0$. If a V_n admits a PC_r, or an AC_r, or an M_r, and also an mFI, then (5.5) and (5.6) hold.

We now consider some special types of dependence relations in a V_n. If, in Corollary 4.1 for a V_n admitting an M_r (motions), we take $A_{ij} = g_{ij}$, we immediately have that $A_{(\alpha)ij} = \mathcal{L}_\alpha g_{ij} = 0$. ($\mathcal{L}_\alpha$ of course refers to vector $\xi_{(\alpha)}$ of the motion.) We thus see that the first derived integrals, and hence those of all orders, are identically zero in this instance.

Consider again the above V_n admitting an M_r. From the fact that $\xi_{(\alpha)i}$ satisfy Killing's equation it follows that $\xi_{(\alpha)i}p^i \equiv L_\alpha$ are linear first integrals. By Ref. 8, p. 29, we have that the first derived integrals $(\mathcal{L}_\beta \xi_{(\alpha)i})p^i = C^{\gamma}_{\beta\alpha}(\xi_{(\gamma)i}p^i) = C^{\gamma}_{\beta\alpha} L_\gamma$. It follows that all derived integrals based on the L_α are linearly dependent on these L_α. It also follows that derived integrals based on mFI which are products of the L_α's will be linearly dependent on these products. (For examples of mFI expressible as products of LFI, see Ref. 11.)

Finally, consider a V_n admitting an AC_r. Based on the fundamental quadratic first integral $g_{ij}p^ip^j$, we form the first derived integral with coefficients

$$A_{(\alpha)ij} = \mathcal{L}_\alpha g_{ij} = \xi_{(\alpha)i;j} + \xi_{(\alpha)j;i}. \tag{5.7}$$

Next, form the second derived integral with coefficients

$$A_{(\beta\alpha)ij} = \mathcal{L}_\beta A_{(\alpha)ij} = \mathcal{L}_\beta [\xi_{(\alpha)i;j} + \xi_{(\alpha)j;i}]. \tag{5.8}$$

By Ref. 8, p. 16, we can interchange the order of Lie differentiation and covariant differentiation to obtain

$$A_{(\beta\alpha)ij} = (\mathcal{L}_\beta \xi_{(\alpha)i})_{;j} + (\mathcal{L}_\beta \xi_{(\alpha)j})_{;i} = C^{\gamma}_{\beta\alpha}[\xi_{(\gamma)i;j} + \xi_{(\gamma)j;i}] = C^{\gamma}_{\beta\alpha} A_{(\alpha)ij}. \tag{5.9}$$

This may also be expressed as $A_{(\beta\alpha)ij} = \mathcal{L}_\beta \mathcal{L}_\alpha g_{ij} = C^{\gamma}_{\beta\alpha}(\mathcal{L}_\gamma g_{ij})$. This gives us the theorem:

Theorem 5.2: If a V_n admits an AC_r, the second (and higher-order) derived integrals based on the fundamental quadratic first integral $g_{ij}p^ip^j$ are linearly dependent on the first derived integrals.

In general, a similar result for V_n admitting a PC_r will not hold for the dependence of second-derived on first-derived integrals (see Sec. 9).

6. COMPARISON WITH POISSON'S THEOREM

In this section we consider relations between Poisson's theorem on constants of motion and the Related First Integral Theorem.

The well-known Poisson's theorem[12,13] states that if U and V are constants of the motion of a trajectory of a dynamical system, then the Poisson bracket $[U, V] \equiv W$ is also a constant of the motion.

Let us now consider the formulation of Poisson's theorem for mass particles with geodesic trajectories in a V_n. Assume then that a V_n admits an mFI and a qFI given, respectively, by

$$A_{i_1 \cdots i_m} p^{i_1} \cdots p^{i_m} \equiv A^{i_1 \cdots i_m} p_{i_1} \cdots p_{i_m} \equiv U(x^j, p_k), \tag{6.1}$$

$$B_{i_1 \cdots i_q} p^{i_1} \cdots p^{i_q} \equiv B^{i_1 \cdots i_q} p_{i_1} \cdots p_{i_q} \equiv V(x^j, p_k), \tag{6.2}$$

where

$$p_i \equiv g_{ij}p^j. \tag{6.3}$$

In the Poisson bracket

$$[U, V] \equiv \sum_{k=1}^{n} \left(\frac{\partial U}{\partial x^k} \frac{\partial V}{\partial p_k} - \frac{\partial U}{\partial p_k} \frac{\partial V}{\partial x^k} \right), \tag{6.4}$$

we substitute for U and V from Eqs. (6.1) and (6.2) and obtain

$$[U, V] = \left(q \frac{\partial A^{i_1 \cdots i_m}}{\partial x^k} B^{k i_{m+1} \cdots i_{m+q-1}} - m \frac{\partial B^{i_1 \cdots i_q}}{\partial x^k} A^{k i_{q+1} \cdots i_{m+q-1}} \right) p_{i_1} \cdots p_{i_{m+q-1}}, \tag{6.5}$$

where the complete symmetry of $A^{i_1 \cdots i_m}$ and $B^{i_1 \cdots i_q}$ on all their respective indices has been used.

By use of the definition of covariant derivative, the partial derivatives in (6.5) can be replaced to give the following explicit tensor formulation of the Poisson

[11] G. H. Katzin and J. Levine, Tensor, New Series, **16**, 97 (1965); G. H. Katzin, J. Levine, and J. Halsey, *ibid*, **18**, 32 (1967).

[12] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1953), Chap. VIII.

[13] E. T. Whittaker, *Analytical Dynamics of Particles* (Cambridge University Press, London, 1964), 4th ed.

bracket:

$$[U, V] = (qA^{i_1\cdots i_m}_{;k}B^{ki_{m+1}\cdots i_{m+q-1}} - mB^{i_1\cdots i_q}_{;k}A^{ki_{q+1}\cdots i_{m+q-1}})p_{i_1}\cdots p_{i_{m+q-1}}. \quad (6.6)$$

We note that the constants of motion $W = [U, V]$ as given by the right side of (6.6) are of order $m + q - 1$.

The above results, which we consider as a new formulation of Poisson's theorem for first integrals of geodesic motion, can be stated as follows:

Theorem 6.1: If the geodesics of a V_n admit the mFI (6.1) and the qFI (6.2), then the first integral defined by the Poisson bracket (6.4) is of order $m + q - 1$ and is expressible in the tensor form (6.6).

It is observed that the order $m + q - 1$ as given by Theorem 6.1 is, in general, greater than the orders m or q of either of the original first integrals. However, the derived first integrals as given by the RIT are always of the same order as the original integral (see Theorem 4.1). Thus, in order to compare the two theorems as methods for generating first integrals (of geodesics) of order m from a given first integral of order m, we take $q = 1$ in Theorem 6.1.

We therefore consider a V_n which admits an M_r of motions. The geodesics then admit LFI ($q = 1$),

$$\xi_{(\alpha)i}p^i = \xi^i_{(\alpha)}p_i \equiv B^i_{(\alpha)}p_i \equiv V_{(\alpha)}. \quad (6.7)$$

In addition, assume the geodesics admit an mFI (1.2), which we express in the form (6.1).

If in (6.6) with V replaced by $V_{(\alpha)}$ and with $q = 1$, we make use of the skew symmetry of $\xi_{(\alpha)i;j}$ in i and j (due to Killing's equation; see Sec. 2), then it is found that (6.6) is expressible in the form

$$[U, V_\alpha] = \left(A_{i_1\cdots i_m;k}\xi^k_{(\alpha)} + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ki_{h+1}\cdots i_m}\xi^k_{(\alpha);i_h}\right)p^{i_1}\cdots p^{i_m}. \quad (6.8)$$

In terms of the Lie derivative, (6.8) can be written

$$[U, V_\alpha] = (\mathcal{L}_\alpha A_{i_1\cdots i_m})p^{i_1}\cdots p^{i_m}. \quad (6.9)$$

From Theorem 4.1, it is seen that the right side of (6.9) is an mFI of the geodesics. Hence we can state the following theorem:

Theorem 6.2: If a V_n admits a group M_r of motions [so that the V_n admits the LFI (6.7)] and an mFI (6.1), then the first integrals generated by the Poisson brackets (6.6) with $q = 1$ are of order m and are identical with the first derived integrals obtained by the Related First Integral Theorem using the mFI (6.1) and with Lie differentiation taken with respect to the vectors $\xi^i_{(\alpha)}$ of the M_r of motions.

If a V_n admits a PC_r or an AC_r of collineations and an mFI, then the RIT will generate additional mFI, in general not obtainable by Poisson's Theorem 6.1. Hence it follows that Poisson's theorem with $q = 1$ may be considered as a special case of the RIT. However, if $q > 1$, $m > 1$, then Poisson's theorem may be regarded as another method, distinct from the RIT, for obtaining additional first integrals.

7. SPECIAL FIRST INTEGRALS

The integral (1.2) is called a special first integral[14] if

$$A_{i_1\cdots i_m;i_{m+1}} = 0. \quad (7.1)$$

Assume an A_n admits an AC_r and also a special mFI. From the results of Sec. 4, we obtain the derived integrals (4.9), where

$$A_{(\alpha)i_1\cdots i_m} = \mathcal{L}_\alpha A_{i_1\cdots i_m}, \quad (7.2)$$

and $A_{i_1\cdots i_m}$ satisfies (7.1).

Since we are dealing with an AC_r, it follows that Lie differentiation and covariant differentiation commute (see Ref. 8, p. 16); hence from (7.2) we may write

$$A_{(\alpha)i_1\cdots i_m;i_{m+1}} = \mathcal{L}_\alpha(A_{i_1\cdots i_m;i_{m+1}}) = 0. \quad (7.3)$$

It follows that the first derived integrals will also be special integrals, and thus, this will likewise be true for all higher-order derived integrals. We may then state the following theorem:

Theorem 7.1: If the paths of an A_n admit a special mFI (1.2), and if the A_n admits an AC_r, then all derived mFI $A_{(\alpha_1\cdots\alpha_k)i_1\cdots i_m}p^{i_1}\cdots p^{i_m} = \text{const}$ ($k = 1, 2, \cdots$) are also special.

8. DERIVED $(m-1)$FI OBTAINED FROM mFI

Here we assume the paths of an A_n admit the mFI (1.2), and also that the A_n admits a vector λ^j such that

$$\mathcal{L}_\lambda A_{i_1\cdots i_m} \equiv A_{i_1\cdots i_m;j}\lambda^j + \sum_{h=1}^{m} A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m}\lambda^j_{;i_h} = 0. \quad (8.1)$$

If the $\sum$ term in (8.1) is replaced by its equivalent

$$\sum_{h=1}^{m}[(A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m}\lambda^j)_{;i_h} - A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m;i_h}\lambda^j]$$

and (1.3) is used, then (8.1) can be expressed as

$$\sum_{h=1}^{m}(A_{i_1\cdots i_{h-1}ji_{h+1}\cdots i_m}\lambda^j)_{;i_h} + 2A_{i_1\cdots i_m;j}\lambda^j = 0. \quad (8.2)$$

[14] For additional information on special QFI, see our paper "On the Number of Special Quadratic First Integrals in Affinely Connected and Riemannian Spaces," Tensor, New Series, **19** (1968).

Since $A_{i_1 \cdots i_m}$ is completely symmetric, we can define another completely symmetric tensor of order $m-1$ by

$$B_{i_1 \cdots i_{m-1}} \equiv A_{i_1 \cdots i_{m-1} j} \lambda^j. \tag{8.3}$$

By use of (8.3), we express (8.2) in the form

$$P\{B_{i_1 i_2 \cdots i_{m-1} ; i_m}\} + 2A_{i_1 \cdots i_m ; j} \lambda^j = 0. \tag{8.4}$$

Now if the second term on the left side of (8.4) vanishes, then (8.4) reduces to the necessary and sufficient conditions for the existence of an $(m-1)$FI. Hence we have the following:

Theorem 8.1: If the paths of an A_n admit the mFI (1.2), and if the A_n admits a vector λ^i such that

$$A_{i_1 \cdots i_m ; j} \lambda^j = 0$$

and

$$\pounds_\lambda A_{i_1 \cdots i_m} = 0,$$

then the A_n will admit an $(m-1)$FI given by

$$B_{i_1 \cdots i_{m-1}} p^{i_1} \cdots p^{i_{m-1}} = \text{const},$$

where $B_{i_1 \cdots i_m}$ is defined by (8.3).

9. ILLUSTRATIONS OF THE RELATED FIRST INTEGRAL THEOREM

In this section, we illustrate the various theorems obtained above.

Example 1: To illustrate the Related First Integral Theorem, consider the V_4 of Minkowski space-time of special relativity with metric

$$ds^2 = -(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2. \tag{9.1}$$

From Ref. 3, p. 127, it is known that this V_4 admits a 24-parameter group G_{24} generated by the vectors

$$\xi^i = a_h x^h x^i + b^i_h x^h + c^i, \tag{9.2}$$

which include a PC_4, an AC_{10}, and an M_{10}

Let us assume a particle with rest mass m_0 moves along a geodesic in this space-time. It is well known[5] that there exists a set of 10 LFI associated with the motion of this particle. We select from this set the LFI which describes the constancy of linear momentum in the x^1 direction:

$$A_i p^i \equiv m_0 p^i = \text{const}, \quad (A_i = m_0 \delta^1_i). \tag{9.3}$$

We take for a projective collineation vector and related scalar,

$$\xi^i_{(2)} \equiv x^i x^2, \quad \phi_{(2)} = x^2, \tag{9.4}$$

obtained from (9.2) and (2.4), respectively, [with $a_j = \delta^2_j$, $b^i_j = 0$, $c^i = 0$ in (9.2)].

If we substitute from (9.3) and (9.4) into (4.1), we obtain the first-derived LFI (4.3) in the form

$$A_{(2)i} p^i \equiv m_0(x^1 p^2 - x^2 p^1) = \text{const}. \tag{9.5}$$

It is observed that this derived integral infers the constancy of angular momentum of the particle about the x^3 axis.

In a similar manner, we can obtain as derived LFI all six LFI associated with the 6-parameter homogeneous Lorentz group by starting with the four LFI associated with the 4-parameter translation group.

Example 2: We again consider the Minkowski V_4 [Eq. (9.1)] and the G_{24} defined by Eq. (9.2). The vector $\xi^i_{(1)} \equiv x^1 x^i$ obtained from Eq. (9.2) gives a projective collineation with $\phi_{(1)} = x^1$. In Eq. (1.2), we take $m = 2$ and $A_{ij} = g_{ij}$ as determined from (9.1), so that (1.2) is the quadratic first integral

$$A_{ij} p^i p^j = -(p^1)^2 - (p^2)^2 - (p^3)^2 + (p^4)^2 = \text{const}. \tag{9.6}$$

From Theorem 4.1, we calculate the first derived integral (4.9) using (4.1), and obtain

$$\begin{aligned} A_{(1)ij} p^i p^j &= (\pounds_1 g_{ij} - 4\phi_{(1)} g_{ij}) p^i p^j \\ &= 2[x^1(p^2)^2 + x^1(p^3)^2 - x^1(p^4)^2 - x^2 p^1 p^2 \\ &\qquad - x^3 p^1 p^3 + x^4 p^1 p^4]. \end{aligned} \tag{9.7}$$

The second derived integral is calculated to be

$$\begin{aligned} A_{(1,1)ij} p^i p^j &= (\pounds_1 A_{(1)ij} - 4\phi_{(1)} A_{(1)ij}) p^i p^j \\ &= 2\{[-(x^2)^2 - (x^3)^2 + (x^4)^2](p^1)^2 \\ &\quad - (x^1)^2(p^2)^2 - (x^1)^2(p^3)^2 + (x^1)^2(p^4)^2 \\ &\quad + 2x^1 x^2 p^1 p^2 + 2x^1 x^3 p^1 p^3 - 2x^1 x^4 p^1 p^4\}. \end{aligned} \tag{9.8}$$

The third (and consequently all higher-order) derived integrals (derived from use of $\xi^i_{(1)}$) turn out to be identically zero. By inspection, it can be seen that the three quadratic first integrals [Eqs. (9.6)–(9.8)] are linearly independent.

It is easily shown that the above derived QFI [Eqs. (9.7) and (9.8)] are expressible as sums of products of the LFI admitted by the V_4 defined by Eq. (9.1). This fact is a consequence of a known property of V_n spaces of constant curvature (see Ref. 11). However, this property will not hold in general in V_n spaces not of constant curvature. This is illustrated in Example 3.

Example 3: We consider the V_2 (of nonconstant curvature) defined by

$$ds^2 = \frac{y^2}{y^2+1} dx^2 + \frac{y^2}{(y^2+1)^2} dy^2,$$

where we are now using the notation $x^1 = x$, $x^2 = y$. It is known (by Levine[15]) that this V_2 admits a G_2 defined by contravariant vectors

$$\xi_{(1)} = (1, 0); \quad \xi_{(2)} = (x, y), \tag{9.9}$$

where $\xi_{(1)}$ defines a motion, and $\xi_{(2)}$ a proper projective collineation.

Corresponding to $\xi^i_{(2)}$, we calculate scalar $\phi_{(2)}$ to be

$$\phi_{(2)} = (y^2 + 4)/3(y^2 + 1).$$

Based on the fundamental quadratic first integral

$$g_{ij}p^ip^j \equiv A_{ij}p^ip^j = \frac{y^2}{y^2+1}(p^1)^2 + \frac{y^2}{(y^2+1)^2}(p^2)^2 \equiv A, \tag{9.10}$$

we calculate the following derived integrals:

$$A_{(2)ij}p^ip^j = (\pounds_2 A_{ij} - 4\phi_{(2)}A_{ij})p^ip^j = \frac{2y^2(y^2-2)}{3(y^2+1)^2}(p^1)^2 - \frac{4y^2}{3(y^2+1)^2}(p^2)^2 \equiv A_{(2)}, \tag{9.11}$$

$$A_{(2,2)ij}p^ip^j = (\pounds_2 A_{(2)ij} - 4\phi_{(2)}A_{(2)ij})p^ip^j = \frac{4y^2(y^2+4)}{9(y^2+1)^2}(p^1)^2 + \frac{16y^2}{9(y^2+1)^2}(p^2)^2 \equiv A_{(2,2)}. \tag{9.12}$$

It can be shown that the first integrals [(9.10–(9.12)] are linearly dependent according to the relation

$$6A_{(2)} + 9A_{(2,2)} - 8A = 0. \tag{9.13}$$

It is found that the derived integrals are

$$A_{(1)ij}p^ip^j \equiv 0, \quad A_{(1,2)ij}p^ip^j \equiv 0. \tag{9.14}$$

It follows from Eqs. (9.13) and (9.14) that all other derived integrals based on Eq. (9.10) and G_2 defined by Eq. (9.9) are either identically zero or are linearly dependent on A and $A_{(2)}$.

[15] J. Levine, Ann. Math. (Princeton, New Jersey) **52**, 465 (1950). Note that in Theorem 5.2 of this reference, $[p, xp + q]$ should be replaced by $[p, xp + yq]$.

The only linearly independent LFI admitted by this V_2 is given by

$$\xi_{(1)i}p^i = \xi_{(1)1}p^1 = \frac{y^2}{y^2+1}p^1 \equiv L.$$

It is evident that none of the QFI, A, $A_{(2)}$, or $A_{(2,2)}$ is expressible as a constant times the square of L.

Example 4: From Ref. 15, p. 471, it is known that the A_2, defined by

$$\Gamma^1_{11} = \frac{2}{y^2}, \quad \Gamma^1_{12} = \frac{-1}{y}, \quad \Gamma^1_{22} = 0, \quad \Gamma^2_{11} = \frac{4}{y^3}, \quad \Gamma^2_{12} = \frac{1}{y^2}, \quad \Gamma^2_{22} = \frac{-2}{y},$$

admits the AC_3 generated by

$$\xi_{(1)} = (1, 0), \quad \xi_{(2)} = (2x, y), \quad \xi_{(3)} = (x^2, xy), \quad (x^1 = x, x^2 = y). \tag{9.15}$$

It can be shown that this A_2 admits the LFI

$$A_ip^i \equiv -\frac{2}{y^3}p^1 + \frac{1}{y^2}p^2 \equiv A. \tag{9.16}$$

The first derived integral based on the LFI (9.16) and (9.15) are calculated to be

$$A_{(1)i}p^i \equiv 0, \quad A_{(2)i}p^i = -A, \quad A_{(3)i}p^i = \left(\frac{2x}{y^3} + \frac{1}{y}\right)p^1 - \frac{x}{y^2}p^2 \equiv A_{(3)}. \tag{9.17}$$

All higher-order derived integrals derived from (9.16) are thus linearly dependent on A and $A_{(3)}$. The space A_2 is not projectively flat, and hence cannot admit more than two linearly independent linear first integrals (Ref. 3, p. 123). This implies all linear first integrals are linearly dependent on A and $A_{(3)}$.

ACKNOWLEDGMENT

The authors wish to thank Professor William R. Davis for his helpful comments in preparing the manuscript.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Reduction of Relativistic Wavefunctions to the Irreducible Representations of the Inhomogeneous Lorentz Group. Part II. Zero-Mass Components

H. E. MOSES
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*

(Received 23 January 1967)

In a previous paper, we showed how wavefunctions which transform in a relativistic manner in configuration space can be expanded in terms of amplitudes, which for nonzero mass transform like the wavefunctions for irreducible representations of the proper, orthochronous, inhomogeneous Lorentz group. A simple algorithm was given to obtain the expansion. In the present paper, we extend the results to include zero-mass amplitudes. It is shown that for wavefunctions which are required to transform under the homogeneous Lorentz group such that the matrices which involve the spinor indices are finite dimensional, the zero-mass amplitudes transform under nonunitary representations of the inhomogeneous Lorentz group. However, it is possible to split up each such nonunitary representation into a part which corresponds to a unitary representation for finite spin and into a part which corresponds to an unphysical change of wavefunction. As examples of the technique, we consider wavefunctions which transform as an antisymmetric real tensor (i.e., as an electromagnetic field) as a four-vector with and without the Lorentz condition, and as a Dirac spinor. The results offer interesting contrasts with the reductions of Part I where only nonzero-mass components were considered. It is shown that the expansion of the present paper, when applied to the solution of Maxwell's equations, leads to an expansion in terms of photon wavefunctions and that the unphysical change of wavefunction is zero. For a real vector potential with the Lorentz condition (i.e., the electromagnetic vector potential), the expansion corresponds to the sum of an expansion in terms of photon wavefunctions and a wavefunction which sets the gauge of the vector potential. The nonphysical part of the transformation of the electromagnetic vector potential is merely a gauge change. Finally, solutions of the massless Dirac equation are expanded in terms of wavefunctions for massless particles of spin $\frac{1}{2}$ for which the nonphysical part of the change is zero. In the present paper, we also show how invariant inner products are to be introduced, how negative-energy representations can be replaced by positive-energy representations ("antiparticles"), and show the connection with the usual canonical formalism. Finally, second quantization of the theory is given.

1. INTRODUCTION

The present paper was written as a direct extension of Part I.[1] When we require an equation which appears in Part I, we refer to that equation by prefacing the equation number by I.

In the present paper, we consider wavefunctions $\Psi(x)$, which transform under the transformations of the proper, orthochronous, inhomogeneous Lorentz group as in (I 2.10), (I 2.11), and (I 2.12), and assume the existence of mass-zero representations of the inhomogeneous Lorentz group are contained in the reduction formulas. However, the situation for the mass-zero case is rather different than for the nonzero case for most cases of physical interest; namely for those cases in which the operators M_i and N_i are finite-dimensional matrices. We require the matrices M_i to have the reduced form of (I 2.22) and be Hermitian. As mentioned in Part I, the matrices N_i cannot be Hermitian for the finite-dimensional case. This fact has a consequence. We show that, though it is possible to expand $\Psi(x)$ into modes which transform like *nonunitary* massless representations of the inhomogeneous Lorentz group such that under the transformations of the inhomogeneous Lorentz group, the transformed wavefunction $\Psi'(x)$ can be expanded in terms of the transformed modes; such an expansion is not, in general, possible if we require the modes to transform under the *irreducible unitary* representations. Thus, the modes which appear in the expansion cannot be interpreted as wavefunctions of massless relativistic particles in the same simple way as the modes which appear in the expansion of the wavefunction when we assume nonzero mass.

For such cases, which are the most important ones for physics, we are thus compelled to give up the requirement that the expansion of Ψ' in terms of relativistic particle wavefunctions be such that the transformed function Ψ' has the same expansion in terms of the transformed particle wavefunctions. However, if we loosen our requirement, it is possible to expand Ψ' into massless particles of finite spin in Wigner's classification of the irreducible representations of the inhomogeneous Lorentz group.[2] We

* Operated with support from the U.S. Advanced Research Projects Agency.

[1] H. E. Moses, J. Math. Phys. **8**, 1134 (1967).

[2] E. P. Wigner, Ann. Math. **40**, 149 (1939).

show that under the transformations of the inhomogeneous Lorentz group, the transformed Ψ has a part which we may interpret as being nonphysical or corresponding to a sort of "gauge change." If the unphysical part of the transformed function is ignored, the remaining "physical" part of Ψ is indeed the same expansion in terms of the transformed massless particle wavefunctions as the untransformed function Ψ was of the untransformed particle wavefunctions.

We have applied our formalism to the wavefunction which transforms like an antisymmetric real tensor (or, what is the same, as the electromagnetic field tensor) to the wavefunction which transforms like a four-vector with and without reality conditions, and with and without the Lorentz condition, and to the wavefunction which transforms as a Dirac spinor. It is found that for real antisymmetric tensors which satisfy Maxwell's equations and spinors which satisfy Dirac's equations for zero mass, the unphysical part of the transformed wavefunction is zero. For the real four-vector which satisfies the Lorentz condition (and can thus be interpreted as being the electromagnetic four-vector potential), the unphysical part of the transformed wavefunction is just a gauge change. Indeed, for this reason, the name "gauge change" seems justified for the unphysical parts of the transformed wavefunction.

Our expansion is a direct application of Ref. 3, just as was the expansion of Part I. In the present paper, we also show how the negative-energy wavefunctions which appear in the expansion can be replaced by positive-energy wavefunctions (wavefunctions for "antiparticles") so that one obtains expansions in terms of positive-energy wavefunctions only. We also show the connection to the canonical formalism for the examples treated herein and indicate how second quantization can be introduced to give the usual covariant commutation rules in configuration space.

The interesting results for the electromagnetic fields and electromagnetic vector potential have already been given in Ref. 4, where the expansions of the electromagnetic vector potentials and fields were presented and verified from a different point of view. In the present paper, we show how these expansions are obtained from a general procedure and effectively that the expansions of Ref. 4 are the only possible relativistic expansions.

It should be mentioned that the reduction of solutions of Maxwell's equations is given from a different point of view and using rather different techniques in Ref. 5. In Ref. 6, the reduction of second-quantized wave equations, including Maxwell's equations, is discussed from quite a different point of view than that of the present paper. A more closely related approach is given in Ref. 7, where wave equations are found for wavefunctions in the momentum representation when irreducible sets of matrices M_i, N_i are prescribed. In that reference (see also Ref. 8), it is also found necessary to regard part of the transformed wavefunction as being unphysical or corresponding to a gauge change.

2. MASS-ZERO REPRESENTATIONS FOR THE INFINITESIMAL GENERATORS OF THE INHOMOGENEOUS LORENTZ GROUP: ALGORITHM FOR THE REDUCTION OF WAVEFUNCTIONS

In the present section, we give the algorithm for the reduction of wavefunctions into the zero-mass representations of the inhomogeneous Lorentz group. But before we can present this algorithm, some preliminary material on "standard" realizations for the representations must be given. These realizations are generalizations of the realizations of Ref. 9 (see also Ref. 3).

A. Generators of the Two-Dimensional Euclidean Group

Let us consider a vector space of functions of a real variable λ which can be continuous, discrete, or finite dimensional. Let us assume the variable λ is discrete (as it will be for our applications) for the sake of concreteness. Any more general sets of values for λ can be treated similarly. A member of the vector space is denoted by $F(\lambda)$. Let us define three operators, T_1, T_2, and K, which satisfy the commutation relations,

$$\begin{aligned} [T_1, T_2] &= 0, \\ [T_1, K] &= -iT_2, \\ [T_2, K] &= iT_1. \end{aligned} \tag{2.1}$$

These three operators satisfy the commutation rules for the infinitesimal generators of the two-dimensional Euclidean group. Then we define the matrices $T_i(\lambda \mid \lambda')$, $K(\lambda \mid \lambda')$ by

$$\begin{aligned} T_iF(\lambda) &= \sum_{\lambda'} T_i(\lambda \mid \lambda')F(\lambda'), \\ KF(\lambda) &= \sum_{\lambda'} K(\lambda \mid \lambda')F(\lambda'). \end{aligned} \tag{2.2}$$

[3] J. S. Lomont and H. E. Moses, J. Math. Phys. **8**, 837 (1967).
[4] H. E. Moses, Nuovo Cimento **42**, 757 (1966).
[5] C. Fronsdal, Phys. Rev. **113**, 1367 (1959).
[6] S. Weinberg, Phys. Rev. **138B**, 989 (1965).
[7] A. McKerrell, Ann. Phys. (N.Y.) **40**, 237 (1966).
[8] A. McKerrell and D. L. Pursey, Bull. Am. Phys. Soc. **10**, 89 (1965).
[9] J. S. Lomont and H. E. Moses, J. Math. Phys. **3**, 405 (1962).

The matrices of Eq. (2.2) constitute a representation of the infinitesimal generators of the two-dimensional Euclidean group. The forms that these matrices can take when they are an irreducible Hermitian set of matrices such that the eigenvalues of K are either an integer or half-odd integer are discussed in Refs. 2 and 9. For most of our applications, such restrictions on the representations are too severe.

B. Massless Representations of the Infinitesimal Generators of the Inhomogeneous Lorentz Group

Let us now introduce a space of complex functions $f(\mathbf{p}, \lambda)$, where $\mathbf{p} = (p_1, p_2, p_3)$, and where the range of each variable p_i extends over the entire real axis. The range of λ is as before. We now introduce a realization of the infinitesimal generators of the inhomogeneous Lorentz group when any representation of the infinitesimal generators of the two-dimensional Euclidean group is given. For simplicity, we suppress the appearance of the variable λ and use $Kf(\mathbf{p})$ to mean

$$\sum_{\lambda'} K(\lambda \mid \lambda') f(\mathbf{p}, \lambda')$$

and similarly for $T_i f(\mathbf{p})$. Then our realization of the infinitesimal generators of the inhomogeneous Lorentz group is

$$\hat{P}^0 f(\mathbf{p}) = Hf(\mathbf{p}) = \epsilon p f(\mathbf{p}) \quad (p = |\mathbf{p}|),$$

$$\hat{P}_i f(\mathbf{p}) = p_i f(\mathbf{p}),$$

$$\hat{J}_1 f(\mathbf{p}) = \left[-i \sum_{j,k} \epsilon_{1jk} p_j \frac{\partial}{\partial p_k} + \frac{p_1}{p + p_3} K \right] f(\mathbf{p}),$$

$$\hat{J}_2 f(\mathbf{p}) = \left[-i \sum_{j,k} \epsilon_{2jk} p_j \frac{\partial}{\partial p_k} + \frac{p_2}{p + p_3} K \right] f(\mathbf{p}),$$

$$\hat{J}_3 f(\mathbf{p}) = \left[-i \sum_{j,k} \epsilon_{3jk} p_j \frac{\partial}{\partial p_k} + K \right] f(\mathbf{p}),$$

$$\hat{\delta}_1 f(\mathbf{p}) = \epsilon \left\{ ip \frac{\partial}{\partial p_1} + \frac{p_2}{p + p_3} K + \left[\frac{p_1^2}{p^2(p + p_3)} - \frac{1}{p} \right] T_1 + \frac{p_1 p_2}{p^2(p + p_3)} T_2 \right\} f(\mathbf{p}),$$

$$\hat{\delta}_2 f(\mathbf{p}) = \epsilon \left\{ ip \frac{\partial}{\partial p_2} - \frac{p_1}{p + p_3} K + \frac{p_1 p_2}{p^2(p + p_3)} T_1 + \left[\frac{p_2^2}{p^2(p + p_3)} - \frac{1}{p} \right] T_2 \right\} f(\mathbf{p}),$$

$$\hat{\delta}_3 f(\mathbf{p}) = \epsilon \left\{ ip \frac{\partial}{\partial p_3} + \frac{1}{p^2} [p_1 T_1 + p_2 T_2] \right\} f(\mathbf{p}). \tag{2.3}$$

In (2.3), the quantity ϵ (which is the sign of the energy) may take on either the value $+1$ or -1.

It is easily verified that the infinitesimal generators do indeed satisfy the required commutation relations (I 2.17). It is not difficult to show that a necessary and sufficient condition for the existence of an upper product which makes the above representation of the generators of the inhomogeneous Lorentz group Hermitian is that the representation of the generators of the Euclidean group be equivalent to a Hermitian representation of that group.

Certain kinds of representations of the generators of the Euclidean group are of particular interest. We call them Type I, Type II, and Type III representations.

Type I representations: By Type I representations, we mean those which are irreducible Hermitian representations for which the eigenvalues of K are all integer or half-odd integer. As is well known, these representations are characterized by the value of the scalar, real and positive r^2, which is given by

$$(T_1^2 + T_2^2) f(\mathbf{p}) = r^2 f(\mathbf{p}).$$

The infinitesimal generators of the Lorentz group defined by (2.3) are then Hermitian when we define the inner product of two vectors $|\Phi\rangle$ and $|\Phi^{(1)}\rangle$ represented by $f(\mathbf{p})$ and $f^{(1)}(\mathbf{p})$, respectively, by

$$\langle \Phi^{(1)} \mid \Phi \rangle = \sum_{\lambda} \int \frac{d\mathbf{p}}{p} f^{(1)*}(\mathbf{p}, \lambda) f(\mathbf{p}, \lambda). \tag{2.4}$$

[In (2.4), it was convenient to show the variable λ explicitly, though in (2.3) it is not necessary to do so.]

If $r^2 > 0$, then the variable λ takes on all integer or half-odd integer values, and the matrices which represent the generators of the Euclidean group are infinite dimensional. The dynamical system is identified with particles of "continuous" or infinite spin.

The corresponding representations of the inhomogeneous Lorentz group are irreducible.

Type II representations: These representations are special Type I representations and correspond to the case $r^2 = 0$. In this case, λ is either an integer or half-odd integer. The matrices corresponding to the generators T_i are zero and the operator corresponding to K is simply multiplication by the scalar λ. The dynamical systems corresponding to these representations of the infinitesimal generators of the inhomogeneous Lorentz group are identified with particles of finite spin $|\lambda|$. The value of λ, which includes the sign, gives the helicity of the representation. In practice, these representations are the most important ones for physics.

Type III representations: These representations are generalizations of Type I representations. These representations are characterized by the requirement that

the operator K be Hermitian and has only integer or half-odd integer eigenvalues. The Type III representations, with which we are most concerned, are those for which the matrices representing T_1 and T_2 are finite dimensional (but not zero). In this case, it can be shown that the matrices representing T_i are not unitarily equivalent to Hermitian matrices. For such representations of the infinitesimal generators of the Euclidean group, the generators of the Lorentz group given by (2.3) cannot be Hermitian nor reduced to Hermitian representations.

C. Unphysical Changes in Wavefunction: Gauge Change

Let us now consider a Type III representation of the infinitesimal generators of the inhomogeneous Lorentz group for which the operator K is diagonal. By carrying out a transformation in the λ variable, the operator K can always be brought to this diagonal form. Then,

$$Kf(\mathbf{p}, \lambda) = \lambda f(\mathbf{p}, \lambda). \tag{2.5}$$

Under the various infinitesimal changes of frames of reference, the wavefunction $f(\mathbf{p}, \lambda)$ changes by an infinitesimal amount given by $\hat{A}f(\mathbf{p}, \lambda)$, where $\hat{A}$ is the appropriate linear combination of infinitesimal generators of (2.3). Let us consider those cases for which $\hat{A}$ is restricted to being one of the infinitesimal generators. It is easily seen that for general Type III representations of the Euclidean group, when $\hat{A}$ is one of the operators $\hat{\mathfrak{J}}_i$, $\hat{A}$ will not be Hermitian generally and $\hat{A}f(\mathbf{p})$ will thus not correspond to a change of wavefunction which conserves probabilities. These considerations lead us to considering $\hat{A}f(\mathbf{p})$ as consisting of two parts, one of which corresponds to a true physical change of wavefunction and the other part of which corresponds to an unphysical change of wavefunction of "change of gauge." This view is combined with the notion of only working with finite-spin representations generated by Type II representations of the Euclidean group.

Let us consider $f(\mathbf{p}, \lambda)$ for each λ as being the wavefunction for a finite-spin representation characterized by λ (and ϵ). It is readily seen from (2.3) that if $\hat{A}$ is any of the infinitesimal generators P_i, H, and J_i, the function $f(\mathbf{p}, \lambda)$ transforms properly as a finite-spin wavefunction for each value of λ, separately. However, if $\hat{A}$ is $\hat{\mathfrak{J}}_i$, this statement is no longer true. In fact (suppressing λ for simplicity), we have

$$\hat{\mathfrak{J}}_i f(\mathbf{p}) = g_i(\mathbf{p}) + \hat{\mathfrak{J}}_i^{(1)} f(\mathbf{p}), \tag{2.6}$$

where $\hat{\mathfrak{J}}_i^{(1)}$ is the finite-spin operator (for which $T_i = 0$) and $g_i(\mathbf{p})$ is the inessential change in wavefunction or "gauge change" given by

$$g_1(\mathbf{p}) = \epsilon\left\{\left[\frac{p_1^2}{p^2(p + p_3)} - \frac{1}{p}\right] T_1 + \frac{p_1 p_2}{p^2(p + p_3)} T_2\right\} f(\mathbf{p}), \tag{2.7}$$

and so on [see (2.3)].

When the operator K is not diagonal, we introduce the unitary operator V which diagonalizes K:

$$V^{-1}KV = K_d, \tag{2.8}$$

where K_d is the diagonal matrix.

We also introduce the wavefunction $r(\mathbf{p}, \lambda)$ by

$$f(\mathbf{p}, \lambda) = \sum_{\lambda'} V(\lambda \mid \lambda') r(\mathbf{p}, \lambda'), \tag{2.9}$$

where $V(\lambda \mid \lambda')$ is the notation for the matrix element of V. We now regard $r(\mathbf{p}, \lambda)$ as the wavefunction for finite spin characterized by the helicity λ (by picking the label properly for λ) and sign of energy given by ϵ. Then, if $\hat{A}$ is any of the infinitesimal generators $\hat{P}_i$, $\hat{H}$, $\hat{J}_i$, we have (on suppressing the label λ)

$$\hat{A}f(\mathbf{p}) = V\hat{A}^{(1)} r(\mathbf{p}), \tag{2.10}$$

where $\hat{A}^{(1)}$ is the corresponding operator for the finite-spin representation. For the space–time infinitesimal generators $\hat{\mathfrak{J}}_i$, we have, instead of (2.6),

$$\hat{\mathfrak{J}}_i f(\mathbf{p}) = g_i(\mathbf{p}) + V\hat{\mathfrak{J}}_i^{(1)} r(\mathbf{p}). \tag{2.11}$$

The gauge changes g_i are given by (2.7) and similar expressions as before, but we replace $f(\mathbf{p})$ in these expressions by $Vr(\mathbf{p})$ in accordance with (2.9).

Finally, we note that the matrix V is not unique. But this lack of uniqueness is reflected in trivial choices of phase of $r(\mathbf{p}, \lambda)$.

D. Construction of "Antiparticle" Wavefunctions from Negative-Energy Wavefunctions for Finite-Spin Representations

As in Part I, we find that the expansions of the wavefunctions $\Psi(x)$ contain amplitudes which correspond to negative-energy representations. For physical applications, we want to replace these negative-energy representations by positive-energy representations. In this context, we let $f(\epsilon, \mathbf{p}, \lambda)$ be the wavefunction for a finite-spin particle characterized by the helicity λ and by the sign of energy given by ϵ. In Appendix A, it is shown that the function $h(\epsilon, \mathbf{p}, \lambda)$, defined by

$$h(\epsilon, \mathbf{p}, \lambda) = e^{-2i\lambda\varphi} f^*(-\epsilon, -\mathbf{p}, \lambda), \tag{2.12}$$

where φ is the angle given by

$$\tan \varphi = p_2/p_1, \tag{2.12'}$$

transforms like the wavefunction of a particle with finite spin characterized by the helicity λ and the sign of energy ϵ. The asterisk means, as usual, complex

conjugate. Equation (2.12) thus enables us to carry out our program of replacing negative-energy wavefunctions by those of positive energy.

E. First Form of the Algorithm for the Reduction of the Wavefunction

We now give the algorithm which enables one to reduce the wavefunction $\Psi(x)$ into representations of the proper orthochronous, inhomogeneous group for zero mass.

Let us define the column vector $\chi(\epsilon, \mathbf{p}, \lambda)$, whose components are denoted by $\chi(\gamma \mid \epsilon, \mathbf{p}, \lambda)$, by

$$\chi(\gamma \mid \epsilon, \mathbf{p}, \lambda) = \{\exp[i\boldsymbol{\omega} \cdot \mathbf{M}] \exp[i\nu N_3]\}_{\gamma\lambda}, \quad (2.13)$$

where $\{\exp[i\boldsymbol{\omega} \cdot \mathbf{M}] \exp[i\nu N_3]\}_{\gamma,\gamma'}$ denotes a matrix element of $\exp[i\boldsymbol{\omega} \cdot \mathbf{M}] \exp[i\nu N_3]$ and where the vector $\mathbf{p}$ is in a one-to-one correspondence with $\boldsymbol{\omega}$ and ν through the expressions

$$\begin{aligned} p &= e^{\epsilon\nu}, \\ p_1 &= -p(\sin\omega/\omega)\omega_2, \\ p_2 &= p(\sin\omega/\omega)\omega_1, \\ p_3 &= \cos\omega, \\ \omega_3 &= 0, \quad \omega = |\boldsymbol{\omega}|. \end{aligned} \quad (2.14)$$

Then our expansion is

$$\Psi(x) = \sum_{\lambda}\sum_{\epsilon}\int \frac{d\mathbf{p}}{p}\chi(\epsilon, \mathbf{p}, \lambda) f(\epsilon, \mathbf{p}, \lambda) \times \exp[i(\mathbf{p} \cdot \mathbf{x} - \epsilon p t)]. \quad (2.15)$$

This expansion is such that if A is one of the infinitesimal generators of Eqs. (I 2.13)–(I 2.15), then $A\Psi(x)$ has the same expansion (2.15) above, where $f(\epsilon, \mathbf{p}, \lambda)$ is replaced by $\hat{A}f(\epsilon, \mathbf{p}, \lambda)$, the operators $\hat{A}$ being given by (2.3), the value of ϵ being the same as that in $f(\epsilon, \mathbf{p}, \lambda)$. The representation of the infinitesimal generators of the Euclidean group which occurs in (2.3) is given in terms of M_i, N_i by

$$\begin{aligned} T_1 &= -M_2 - \epsilon N_1, \\ T_2 &= M_1 - \epsilon N_2, \\ K &= M_3. \end{aligned} \quad (2.15')$$

In (2.15′), we have used the notation that T_i and K are the matrices whose elements were previously denoted by $T_i(\lambda \mid \lambda')$ and $K(\lambda \mid \lambda')$, respectively [see (2.2)]. This algorithm is proved in Appendix B.

In the exceptional case that M_i and N_i are Hermitian (a necessary condition being that this set of matrices be infinite dimensional if they are not zero), the infinitesimal generators of the Euclidean group are Hermitian and reducible to the irreducible representations of the Euclidean group, which are representations of Type I. One can then introduce a unitary operator acting on the λ variable in the function $f(\epsilon, \mathbf{p}, \lambda)$ such that in the new basis the matrices T_i, K are reduced. One then has expanded Ψ in terms of the amplitudes or wavefunctions $f(\epsilon, \mathbf{p}, \lambda)$ which transform under the irreducible unitary ray representations of the proper, orthochronous, inhomogeneous Lorentz group with mass zero and sign of energy ϵ. One may then introduce the inner product (2.4) as an invariant inner product, where $\Psi(x)$ and $\Psi^{(1)}(x)$ are regarded as configuration-space representatives of the states in momentum representation given by $f(\epsilon, \mathbf{p}, \lambda)$, $f^{(1)}(\epsilon, \mathbf{p}, \lambda)$, respectively.

There seem, however, to be no physically interesting examples (except for the scalar case) for which the matrices M_i and N_i are all Hermitian. Indeed the only examples which we can think of are given in Ref. 10, where wave equations for functions of infinite spin are discussed. In principle, the wavefunctions of Ref. 10 can be put into the Lomont–Moses form through the use of the algorithm. There seems, however, no point in doing so.

Since we wish to treat the more physically interesting cases where the matrices M_i and N_i are finite dimensional, we derive from the present algorithm, variant forms which reduce the wavefunction to amplitudes which transform under the finite-spin representations of the inhomogeneous Lorentz group.

F. Second Form for the Algorithm

Let us introduce a unitary matrix V which diagonalizes $M_3 = K$:

$$V^{-1}M_3V = M_{3d}, \quad (2.16)$$

where M_{3d} is the diagonal matrix, which we choose to have the reduced form (I 2.22), in which the matrix $S_3^{(r)}$ is diagonal such that the eigenvalues take on the values $j^{(r)}, j^{(r)} - 1, \cdots, -j^{(r)} + 1, -j^{(r)}$ from top to bottom along the main diagonal. Let us now define the column vector $\chi^{(r)}(\epsilon, \mathbf{p}, \lambda)$, whose components are denoted by $\chi^{(r)}(\gamma \mid \epsilon, \mathbf{p}, \lambda)$, by

$$\chi^{(r)}(\gamma \mid \epsilon, \mathbf{p}, \lambda) = \{\exp[i\boldsymbol{\omega} \cdot \mathbf{M}] \exp[i\nu N_3]V\}_{\gamma\lambda}, \quad (2.17)$$

where λ is restricted to columns corresponding to the rth block in the reduced form of M_3. Furthermore, we relable λ so that when it corresponds to a column of the rth block, its value is taken to be the eigenvalue of $S_3^{(r)}$ associated with that column. That is, for example, if λ corresponds to the second column of the rth block, we take λ to have the value $j^{(r)} - 1$. If it corresponds to the third column, we give λ the value $j^{(r)} - 2$ and so on.

[10] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) 34, 211 (1948).

The vector $\mathbf{p}$ is in the one-to-one correspondence with ν and $\boldsymbol{\omega}$ given by (2.14).

Then our expansion is

$$\Psi(x) = \sum_{\epsilon}\sum_{r}\sum_{\lambda} C^{(r)}(\epsilon,\lambda)\int \frac{d\mathbf{p}}{p}\chi^{(r)}(\epsilon,\mathbf{p},\lambda)f^{(r)}(\epsilon,\mathbf{p},\lambda) \times \exp\left[i(\mathbf{p}\cdot\mathbf{x} - \epsilon pt)\right]. \tag{2.18}$$

In (2.18), the quantities $C^{(r)}(\epsilon,\lambda)$ are arbitrary positive constants. The expansion is such that if A is any of the infinitesimal generators P_i, H, and J_i, then $A\psi(x)$ has the same expansion in terms of $\hat{A}f^{(r)}(\epsilon,\mathbf{p},\lambda)$, where $\hat{A}$ is the corresponding finite-spin generator for sign of energy ϵ and helicity λ. However, the expansion for $\mathcal{J}_i\psi$ is somewhat more complicated:

$$\mathcal{J}_i\Psi(x) = \sum_{r}\sum_{\epsilon}\sum_{\lambda} C^{(r)}(\epsilon,\lambda)\int \frac{d\mathbf{p}}{p}\chi^{(r)}(\epsilon,\mathbf{p},\lambda)\hat{\mathcal{J}}_i f^{(r)}(\epsilon,\mathbf{p},\lambda) \times \exp\left[i(\mathbf{p}\cdot\mathbf{x} - \epsilon pt)\right] + G_i(x). \tag{2.19}$$

In (2.19), $\hat{\mathcal{J}}_i$ is the finite-spin generator and $G_i(x)$ is the unphysical change which is added to Ψ when the frame of reference is changed by an infinitesimal space–time transformation. Explicitly, $G_i(x)$ is given as follows. Let us introduce the column vectors $\Theta_i^{(r)}(\epsilon,\mathbf{p},\lambda)$, whose components are denoted by $\Theta_i^{(r)}(\gamma \mid \epsilon,\mathbf{p},\lambda)$, as being defined in the following way:

$$\Theta_i^{(r)}(\gamma \mid \epsilon,\mathbf{p},\lambda) = \{\exp(i\boldsymbol{\omega}\cdot\mathbf{M})\exp(i\nu N_3)B_i V\}_{\gamma\lambda}, \tag{2.20}$$

where

$$B_1 = \left[\frac{p_1^2}{p^2(p+p_3)} - \frac{1}{p}\right]T_1 + \frac{p_1p_2}{p^2(p+p_3)}T_2,$$
$$B_2 = \frac{p_1p_2}{p^2(p+p_3)}T_1 + \left[\frac{p_2^2}{p^2(p+p_3)} - \frac{1}{p}\right]T_2,$$
$$B_3 = \frac{1}{p^2}[p_1T_1 + p_2T_2]. \tag{2.20'}$$

In (2.20), our convention for labeling λ and r is the same as that used for Eq. (2.17). The matrices T_i are given in terms of M_i, N_i by the first two of Eq. (2.15').

Now we can write $G_i(x)$ explicitly as

$$G_i(x) = \sum_{r}\sum_{\epsilon}\sum_{\lambda} C^{(r)}(\epsilon,\lambda)\int \frac{d\mathbf{p}}{p}\Theta_i^{(r)}(\epsilon,\mathbf{p},\lambda)f^{(r)}(\epsilon,\mathbf{p},\lambda) \times \exp\left[i(\mathbf{p}\cdot\mathbf{x} - \epsilon pt)\right]. \tag{2.21}$$

This form of the algorithm follows from the previous form together with the discussion that leads to Eq. (2.11). The constants $C^{(r)}(\epsilon,\lambda)$ are introduced for the sake of convenience, though they are not essential.

The functions $G_i(x)$ can be given in a neater form. For $n = 1$ and 3, let us define the column vector $\Xi_n(x)$ by

$$\Xi_n(x) = \sum_{r}\sum_{\epsilon}\sum_{\lambda} C^{(r)}(\epsilon,\lambda)\int \frac{d\mathbf{p}}{p^{n+1}}\chi^{(r)}(\epsilon,\mathbf{p},\lambda)f^{(r)}(\epsilon,\mathbf{p},\lambda) \times \exp\left[i(\mathbf{p}\cdot\mathbf{x} - \epsilon pt)\right]. \tag{2.22}$$

[It is clear that $(\partial^2/\partial t^2)\Xi_3(x) = -\Xi_1(x)$.] Then using $\nabla_i = \partial/\partial x_i$, it is shown in Appendix C that

$$G_j(x) = i\left[-(\mathbf{M}\times\nabla)_j + \frac{\partial}{\partial t}N_j\right]\Xi_1(x) + i(\mathbf{N}\cdot\nabla)\nabla_j\frac{\partial}{\partial t}\Xi_3(x). \tag{2.23}$$

Now we are in a position to define an invariant inner product of two wavefunctions Ψ and $\Psi^{(1)}$ which have the same transformation properties. Let Ψ be expanded as in (2.18) and let $\Psi^{(1)}$ have the same expansion but with the functions $f^{(r)}(\epsilon,\mathbf{p},\lambda)$ replaced by $g^{(r)}(\epsilon,\mathbf{p},\lambda)$. Then the inner product of $\Psi^{(1)}$ and Ψ, which we denote by $(\Psi^{(1)},\Psi)$, is given by

$$(\Psi^{(1)},\Psi) = \sum_{r}\sum_{\epsilon}\sum_{\lambda} D^{(r)}(\epsilon,\lambda) \times \int \frac{d\mathbf{p}}{p} g^{(r)*}(\epsilon,\mathbf{p},\lambda)f^{(r)}(\epsilon,\mathbf{p},\lambda), \tag{2.24}$$

where the quantities $D^{(r)}(\epsilon,\lambda)$ are any convenient set of nonnegative real numbers. This definition of inner product makes no reference whatever to the gauge change.

G. Third Form for the Algorithm

The third form of the algorithm is simply a rewriting of the previous form in which the negative-energy wavefunctions are replaced by the complex conjugate of positive-energy wavefunctions so that the expansion can be interpreted entirely in terms of "physical" particles. Our applications are stated in this form of the expansion.

Let us then define the following:

$$\begin{aligned}\chi^{(r)}(\mathbf{p},\lambda) &= \chi^{(r)}(+1,\mathbf{p},\lambda),\\ C^{(r)}(\lambda) &= C^{(r)}(+1,\lambda),\\ f^{(r)}(\mathbf{p},\lambda) &= f^{(r)}(+1,\mathbf{p},\lambda),\\ \zeta^{(r)}(\mathbf{p},\lambda) &= e^{2i\lambda\varphi}\chi^{(r)*}(-1,-\mathbf{p},\lambda)\end{aligned} \tag{2.25}$$

(where $\tan\varphi = p_2/p_1$),

$$D^{(r)}(\lambda) = C^{(r)}(-1,\lambda),$$
$$h^{(r)}(\mathbf{p},\lambda) = e^{-2i\lambda\varphi}f^{(r)*}(-1,-\mathbf{p},\lambda).$$

The expansion (2.18) becomes

$$\Psi(x) = \sum_{r}\sum_{\lambda} C^{(r)}(\lambda)\int \frac{d\mathbf{p}}{p}\chi^{(r)}(\mathbf{p},\lambda)f^{(r)}(\mathbf{p},\lambda) \times \exp\left[i(\mathbf{p}\cdot x - pt)\right] + \sum_{r}\sum_{\lambda} D^{(r)}(\lambda)\int \frac{d\mathbf{p}}{p}\zeta^{(r)*}(\mathbf{p},\lambda)h^{(r)*}(\mathbf{p},\lambda) \times \exp\left[-i(\mathbf{p}\cdot\mathbf{x} - pt)\right]. \tag{2.26}$$

The function $\Xi_n(x)$ out of which the gauge change is constructed is written analogously:

$$\Xi_n(x) = \sum_r \sum_\lambda C^{(r)}(\lambda) \int \frac{d\mathbf{p}}{p^{n+1}} \chi^{(r)}(\mathbf{p}, \lambda) f^{(r)}(\mathbf{p}, \lambda) \times \exp[i(\mathbf{p}\cdot\mathbf{x} - pt)] + \sum_r \sum_\lambda D^{(r)}(\lambda) \int \frac{d\mathbf{p}}{p^{n+1}} \zeta^{(r)*}(\mathbf{p}, \lambda) h^{(r)*}(\mathbf{p}, \lambda) \times \exp[-i(\mathbf{p}\cdot\mathbf{x} - pt)]. \quad (2.27)$$

The inner product (2.24) now reads

$$(\Psi^{(1)}, \Psi) = \sum_r \sum_\lambda E^{(r)}(\lambda) \int \frac{d\mathbf{p}}{p} g^{(r)*}(\mathbf{p}, \lambda) f^{(r)}(\mathbf{p}, \lambda) + \sum_r \sum_\lambda F^{(r)}(\lambda) \int \frac{d\mathbf{p}}{p} h^{(r)*}(\mathbf{p}, \lambda) k^{(r)}(\mathbf{p}, \lambda), \quad (2.28)$$

where $E^{(r)}(\lambda)$ and $F^{(r)}(\lambda)$ are arbitrary positive constants which we may choose for our convenience, and $g^{(r)}(\mathbf{p}, \lambda)$ and $k^{(r)}(\mathbf{p}, \lambda)$ are the wavefunctions which take the place of $f^{(r)}(\mathbf{p}, \lambda)$ and $h^{(r)}(\mathbf{p}, \lambda)$ in the expansion of $\Psi^{(1)}$ according to (2.26).

3. WAVEFUNCTIONS WHICH TRANSFORM AS A REAL ANTISYMMETRIC TENSOR OR ELECTROMAGNETIC FIELD

We now apply the procedure given in the previous section. The simplest application is to wavefunctions which transform as a scalar. The various results for the expansion and the connection to the canonical formalism, as well as the introduction of second quantization, are identical to corresponding results for the scalar case in Part I when the mass μ is set equal to zero. Hence, we do not consider this case.

Instead, we do consider the important case of wavefunctions which transform as a real antisymmetric tensor or, what is equivalent, as an electromagnetic field. We shall thus expand the electromagnetic field wavefunctions in terms of the irreducible representations and, when the wavefunctions are required to satisfy Maxwell's equations in the vacuum, obtain the expansion of the wavefunction in terms of photon wavefunctions. This latter result has been obtained from a different point of view in Ref. 4 where the reduction was obtained as a sufficient condition for reduction to photon wavefunctions. By contrast, in the present paper the expansion is obtained as a necessary condition.

A. The Expansion and Gauge Change

We follow Sec. 4 of Part I very closely for the sake of brevity. Accordingly, we define the column vector $\Psi(x)$ by (I 4.1) and (I 4.2). The matrices M_i, N_i are given by (I 4.3) and (I 4.4). For the matrix V, we use

$$V = \begin{pmatrix} -(2)^{-\frac{1}{2}} & 0 & (2)^{-\frac{1}{2}} \\ -i(2)^{-\frac{1}{2}} & 0 & -i(2)^{-\frac{1}{2}} \\ 0 & 1 & 0 \end{pmatrix}. \quad (3.1)$$

The matrices M_i form an irreducible representation of the rotation group with $j = 1$. Hence, the label r takes on only one value, which we shall not indicate. Furthermore, the variable λ takes on the values $\pm 1, 0$. On using (2.17), (3.1), (I 4.6), and (I 4.8), we obtain the following results for $\chi(\epsilon, \mathbf{p}, \lambda)$:

$$\chi(\epsilon, \mathbf{p}, 0) = \begin{pmatrix} p_1/p \\ p_2/p \\ p_3/p \end{pmatrix}, \quad (3.2)$$

$$\chi(\epsilon, \mathbf{p}, \lambda) = [\lambda p^{-\epsilon\lambda}/(2)^{\frac{1}{2}}]\sigma(\mathbf{p}, \lambda) \quad \text{for} \quad \lambda = \pm 1, \quad (3.3)$$

where $\sigma(\mathbf{p}, \lambda)$ is the column vector (defined only for $\lambda = \pm 1$);

$$\sigma(\mathbf{p}, \lambda) = \begin{pmatrix} \dfrac{p_1(p_1 + i\lambda p_2)}{p(p + p_3)} - 1 \\ \dfrac{p_2(p_1 + i\lambda p_2)}{p(p + p_3)} - i\lambda \\ \dfrac{p_1 + \lambda i p_2}{p} \end{pmatrix}. \quad (3.3')$$

From (2.25) $\zeta(\mathbf{p}, \lambda)$ is given by

$$\zeta(\mathbf{p}, 0) = -\chi(\mathbf{p}, 0) = -\chi(+1, \mathbf{p}, 0), \quad (3.4)$$

$$\zeta(\mathbf{p}, \lambda) = -[\lambda p^{\lambda}/(2)^{\frac{1}{2}}]\sigma(\mathbf{p}, \lambda) \quad \text{for} \quad \lambda = \pm 1. \quad (3.5)$$

As in Sec. 4 of Part I, it is convenient to introduce a vector notation for Ψ and related quantities. Thus, we define the vector $\boldsymbol{\Psi}(x)$ as the vector whose components are the components of the column vector $\Psi(x)$. Similarly, we define the vectors $\boldsymbol{\chi}(\mathbf{p}, \lambda)$, $\boldsymbol{\zeta}(\mathbf{p}, \lambda)$, $\boldsymbol{\sigma}(\mathbf{p}, \lambda)$, $\boldsymbol{\Xi}_n(x)$, and $\mathbf{G}_i(x)$.

Then the expansion corresponding to (2.26) can be written as follows:

$$\boldsymbol{\Psi}(x) = \int \frac{d\mathbf{p}}{p} \boldsymbol{\chi}(\mathbf{p}, 0)[C(0) f(\mathbf{p}, 0) e^{i(\mathbf{p}\cdot\mathbf{x} - pt)} - D(0) h^*(\mathbf{p}, 0) e^{-i(\mathbf{p}\cdot\mathbf{x} - pt)}] + \sum_{\lambda=\pm 1} \frac{\lambda}{(2)^{\frac{1}{2}}} \times \left\{ C(\lambda) \int \frac{d\mathbf{p}}{p} p^{-\lambda} \boldsymbol{\sigma}(\mathbf{p}, \lambda) f(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x} - pt)} - D(\lambda) \times \int \frac{d\mathbf{p}}{p} p^{\lambda} \boldsymbol{\sigma}^*(\mathbf{p}, \lambda) h^*(\mathbf{p}, \lambda) e^{-i(\mathbf{p}\cdot\mathbf{x} - pt)} \right\}. \quad (3.6)$$

The expression for $\boldsymbol{\Xi}_n(x)$ is the same as (3.6), but $\mathbf{d}p/p$ is replaced by $d\mathbf{p}/p^{n+1}$.

The expressions for the infinitesimal gauge changes can now be written in the following form. Let us denote the unit vector in the direction of the ith space axis as $\mathbf{e}_i$ [e.g., $\mathbf{e}_1 = (1, 0, 0)$].

Then,

$$\mathbf{G}_j(x) = -\nabla \Xi_{1,j}(x) + \mathbf{e}_j[\nabla \cdot \Xi_1(x)] - i\frac{\partial}{\partial t}[\mathbf{e}_j \times \Xi_1(x)] - i\frac{\partial^2}{\partial x_j \partial t}[\nabla \times \Xi_2(x)], \quad (3.7)$$

where $\Xi_{1,j}$ denotes the jth component of the vector Ξ_1, and ∇ is the gradient operator in $\mathbf{x}$ space.

Finally, we note certain properties of $\chi(\mathbf{p}, 0)$ and $\sigma(\mathbf{p}, \lambda)$ which are useful, to wit,

$$\begin{aligned} [\nabla \cdot \chi(\mathbf{p}, 0)e^{i\mathbf{p}\cdot\mathbf{x}}] &= ip, \\ [\nabla \cdot \sigma(\mathbf{p}, \lambda)e^{i\mathbf{p}\cdot\mathbf{x}}] &= 0, \\ [\nabla \times \chi(\mathbf{p}, 0)\exp(i\mathbf{p}\cdot\mathbf{x})] &= 0, \\ [\nabla \times \sigma(\mathbf{p}, \lambda)\exp(i\mathbf{p}\cdot\mathbf{x})] &= p\lambda\sigma(\mathbf{p}, \lambda) \end{aligned} \quad (3.8)$$

B. Solutions of Maxwell's Equations: Connection with the Canonical Formalism, Second Quantization, Gauge Change

In terms of $\Psi(x)$, Maxwell's equations in the vacuum for $\mathbf{E}$ and $\mathbf{H}$ become

$$\begin{aligned} \nabla \cdot \Psi &= 0, \\ \nabla \times \Psi &= -i\frac{\partial}{\partial t}\Psi. \end{aligned} \quad (3.9)$$

The first of Eqs. (3.9) leads to

$$f(\mathbf{p}, 0) = h(\mathbf{p}, 0) = 0, \quad (3.10)$$

when the first two of Eqs. (3.8) are used. Hence, there can be no spin-zero massless particles in the expansion for the electromagnetic fields.

The second of Eqs. (3.9) leads to

$$f(\mathbf{p}, +1) = h(\mathbf{p}, -1) = 0, \quad (3.11)$$

on using the second two of Eqs. (3.8). Thus, in the expansion (3.6), only the wavefunctions $f(\mathbf{p}, -1)$ and $h(\mathbf{p}, +1)$ need not be identically zero for Ψ to satisfy Maxwell's equations (3.9). It is natural to regard these functions as being wavefunctions of the photon where λ is the circular polarization. The value $\lambda = +1$ corresponds to circular polarization in the direction of propagation given by $\mathbf{p}$, while the value $\lambda = -1$ corresponds to circular polarization in the opposite direction.

To find the constants $C(-1)$ and $D(+1)$ in the expansion (3.6), we connect the canonical formalism to the particle nature of the field as discussed in Part I. The Hamiltonian density of the field, which in the present case is identical to the electromagnetic energy density, is given by

$$H(\mathbf{x}) = (8\pi)^{-1}(\mathbf{E}^2 + \mathbf{H}^2) = (8\pi)^{-1}\Psi^* \cdot \Psi. \quad (3.12)$$

For the mode $h(\mathbf{p}, +1) = 0$, we require that the total energy of the field equals the expectation energy when the wavefunction of the photon is given by $f(\mathbf{p}, -1)$, i.e.,

$$H = \int H(\mathbf{x})\, d\mathbf{x} = \int \frac{d\mathbf{p}}{p} f^*(\mathbf{p}, -1)pf(\mathbf{p}, -1).$$

Likewise, for the case that $f(\mathbf{p}, -1) = 0$, we require

$$H = \int d\mathbf{p}\, |h(\mathbf{p}, +1)|^2.$$

These requirements lead to

$$C(-1) = D(+1) = \pi^{-1}. \quad (3.13)$$

From

$$\begin{aligned} \mathbf{E} &= \tfrac{1}{2}(\Psi + \Psi^*), \\ \mathbf{H} &= \tfrac{1}{2}i(\Psi - \Psi^*), \end{aligned} \quad (3.14)$$

we can expand the electromagnetic fields $\mathbf{E}$ and $\mathbf{H}$ themselves in terms of the photon wavefunctions. The expansion takes a neater form if we change the phase of the wavefunctions in a trivial way. Toward this end, we introduce the wavefunctions $g(\mathbf{p}, \lambda)$ defined for $\lambda = \pm 1$ as follows:

$$\begin{aligned} g(\mathbf{p}, +1) &= -ih(\mathbf{p}, +1), \\ g(\mathbf{p}, -1) &= -if(\mathbf{p}, -1). \end{aligned} \quad (3.15)$$

Then from (3.14), (3.6), (3.13), (3.10), and (3.11), we obtain the following expansions for solutions of Maxwell's equations:

$$\begin{aligned} \mathbf{E}(x) &= i(8\pi^2)^{-\frac{1}{2}} \sum_\lambda \int d\mathbf{p}[g(\mathbf{p}, \lambda)\sigma(\mathbf{p}, \lambda)e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} \\ &\qquad - g^*(\mathbf{p}, \lambda)\sigma^*(\mathbf{p}, \lambda)e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}], \\ \mathbf{H}(x) &= (8\pi^2)^{-\frac{1}{2}} \sum_\lambda \lambda \int d\mathbf{p}[g(\mathbf{p}, \lambda)\sigma(\mathbf{p}, \lambda)e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} \\ &\qquad + g^*(\mathbf{p}, \lambda)\sigma^*(\mathbf{p}, \lambda)e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}]. \end{aligned} \quad (3.16)$$

Equations (3.16) are identical to Eqs. (4.2) and (4.3) of Ref. 4 except for some slight differences of notation. The significance of λ as a parameter describing circular polarization of the electromagnetic field is described in some detail in Ref. 4.

It is readily shown, from the fact that $\Xi_n(x)$ satisfies Maxwell's equations (3.9) when $\Psi(x)$ does, that the gauge changes $\mathbf{G}_i(x)$ are identically zero. Hence, when the wavefunctions $g(\mathbf{p}, \lambda)$ transform like single-spin relativistic wavefunctions, $\mathbf{E}$, $\mathbf{H}$, and Ψ transform as required without the necessity of introducing unphysical portions to the transformed wavefunction.

To second quantize the theory, we consider $g(\mathbf{p}, \lambda)$ and $g^*(\mathbf{p}, \lambda)$ to be destruction and creation operators,

respectively, which satisfy the boson commutation rules,

$$[g(\mathbf{p}, \lambda), g(\mathbf{p}', \lambda')] = [g^*(\mathbf{p}, \lambda), g^*(\mathbf{p}', \lambda')] = 0,$$
$$[g(\mathbf{p}, \lambda), g^*(\mathbf{p}', \lambda')] = p\delta_{\lambda,\lambda'}\delta(\mathbf{p} - \mathbf{p}'). \quad (3.17)$$

It is readily shown that the electromagnetic fields **E** and **H** satisfy the usual commutation rules. It is also an easy matter to show that the second quantized theory is relativistically invariant. Let $\hat{A}$ be any of the single-spin infinitesimal generators and let $\hat{A}g(\mathbf{p}, \lambda)$ mean that the operator $\hat{A}$ acts upon the destruction operator $g(\mathbf{p}, \lambda)$ through $\mathbf{p}$ and λ (which takes on only the values ± 1) as though $g(\mathbf{p}, \lambda)$ were a wavefunction of positive energy and a helicity given by λ. Then for every operator $\hat{A}$, we can define a second quantized operator $[A]$ by

$$[A] = \sum_{\lambda} \int \frac{d\mathbf{p}}{p} g^*(\mathbf{p}, \lambda)\hat{A}g(\mathbf{p}, \lambda). \quad (3.18)$$

As described in Part I, the operators $[A]$ are the infinitesimal generators for the second-quantized theory. Let us now regard $\Psi(x) = \mathbf{E}(x) - i\mathbf{H}(x)$ as a vector whose components are operators. Then under the translation $T(a^\alpha)$, the set of operators Ψ transforms to Ψ' by

$$\Psi'(x) = \Psi(x + a)$$
$$= \exp\left\{-i \sum_{\alpha} a^\alpha [P_\alpha]\right\} \Psi(x) \exp\left\{i \sum_{\alpha} a^\alpha [P_\alpha]\right\}. \quad (3.19)$$

Likewise, under the rotation $R(\boldsymbol{\theta})$, the new set of operators Ψ' is given by both

$$\Psi'(x) = \exp\{-i\boldsymbol{\theta} \cdot [\mathbf{J}]\}\Psi(x) \exp\{i\boldsymbol{\theta} \cdot [\mathbf{J}]\} \quad (3.20)$$

and (I 4.11). Under the pure Lorentz transformation $L(\boldsymbol{\beta})$, the new set Ψ' is given by both

$$\Psi'(x) = \exp\{-i\boldsymbol{\beta} \cdot [\mathbf{\mathfrak{J}}]\}\Psi(x) \exp\{i\boldsymbol{\beta} \cdot [\mathbf{\mathfrak{J}}]\} \quad (3.21)$$

and (I 4.12).

4. WAVEFUNCTIONS WHICH TRANSFORM AS A FOUR-VECTOR

In the present section we use the notation of Sec. 5 of Part I for the four-vector.

A. The Expansion and Gauge Change

In the present case, the matrices $M_i = \hat{M}_i$ are given in the reduced form (I 2.9), where one of the sets of reduced matrices corresponds to $j = 0$ and the other corresponds to the infinitesimal generators for ordinary rotations in three-dimensional space for which $j = 1$. Accordingly, the matrix V is written in a similarly reduced form as

$$V = \begin{pmatrix} 1 & 0 \\ 0 & V_3 \end{pmatrix}, \quad (4.1)$$

where V_3 is the three-by-three dimensional matrix called V in (3.1), and the zeros in (4.1) are the row or column vectors needed to bring V into the reduced form.

The label r can now take on two values corresponding to the two irreducible representations of M_i which appear in (I 2.9). We take $r = 0$ to label the scalar representation and $r = 1$ to label the vector representation. For $r = 0$, λ can take on only one value, namely zero. For $r = 1$, λ can take on the values $0, \pm 1$.

The rotation matrix $\exp(i\boldsymbol{\theta} \cdot \mathbf{M})$ also takes on a reduced form corresponding to (I 2.9); namely,

$$\exp(i\boldsymbol{\theta} \cdot \mathbf{M}) = \begin{pmatrix} 1 & 0 \\ 0 & \hat{R}(\boldsymbol{\theta}) \end{pmatrix}, \quad (4.2)$$

where $\hat{R}(\boldsymbol{\theta})$ is the ordinary 3×3 rotation matrix whose elements are given by (I 4.6).

We now can find the transformation column vectors $\chi^{(r)}$ from (2.13) and (2.14) on using (4.1), (3.1), (4.2), (I 4.6), (I 5.3), and (I 2.6). On labeling γ from 0 to 3 as in Part I and λ as indicated above, we have the following results:

$$\chi^{(0)}(\epsilon, \mathbf{p}, 0) = \epsilon \frac{p^2 - 1}{2p} \begin{pmatrix} \epsilon \dfrac{p^2 + 1}{p^2 - 1} \\ \dfrac{p_1}{p} \\ \dfrac{p_2}{p} \\ \dfrac{p_3}{p} \end{pmatrix}; \quad (4.3)$$

$$\chi^{(1)}(\epsilon, \mathbf{p}, 0) = \frac{p^2 + 1}{2p} \begin{pmatrix} \epsilon \dfrac{p^2 - 1}{p^2 + 1} \\ \dfrac{p_1}{p} \\ \dfrac{p_2}{p} \\ \dfrac{p_3}{p} \end{pmatrix}; \quad (4.4)$$

$$\chi^{(1)}(\epsilon, \mathbf{p}, \lambda) = \frac{\lambda}{2^{\frac{1}{2}}} \begin{pmatrix} 0 \\ \sigma_1(\mathbf{p}, \lambda) \\ \sigma_2(\mathbf{p}, \lambda) \\ \sigma_3(\mathbf{p}, \lambda) \end{pmatrix} \quad \text{for} \quad \lambda = \pm 1. \quad (4.5)$$

In (4.5), the quantities σ_i are the components of the column vector (3.3a).

For the column vectors $\zeta^{(r)}$, we have

$$\zeta^{(0)}(\mathbf{p}, 0) = \chi^{(0)}(\mathbf{p}, 0) = \chi^{(0)}(+1, \mathbf{p}, 0), \tag{4.6}$$

$$\zeta^{(1)}(\mathbf{p}, 0) = -\chi^{(1)}(\mathbf{p}, 0) = -\chi^{(0)}(+1, \mathbf{p}, 0), \tag{4.7}$$

$$\zeta^{(1)}(\mathbf{p}, \lambda) = -\chi^{(1)}(\mathbf{p}, \lambda) = -\chi^{(1)}(+1, \mathbf{p}, \lambda) \quad \text{for} \quad \lambda = \pm 1. \tag{4.8}$$

Now we can expand the column vector $A(x)$, whose components A^ν are the components of the four-vector, in accordance with (2.26). It is possible to rewrite this expansion in a neater form. First of all, we write

$$A = A_0 + A_1, \tag{4.9}$$

where A_0 contains that part of the expansion of A which contains zero-spin wavefunctions only, and A_1 contains only the spin-1 wavefunctions. In the expansion for A_1, it is convenient to replace the wavefunctions $f^{(1)}(\mathbf{p}, \lambda)$ by $\lambda f^{(1)}(\mathbf{p}, \lambda)$, and $h^{(1)}(\mathbf{p}, \lambda)$ by $-\lambda h^{(1)}(\mathbf{p}, \lambda)$. These changes are, of course, trivial changes of phase. Then the expansion for A_1 takes the form

$$A_1(x) = \begin{pmatrix} 0 \\ \mathbf{A}_1(x) \end{pmatrix}, \tag{4.9'}$$

where $\mathbf{A}_1$ is the vector whose components are the space components of the four-vector $A_1(x)$. $\mathbf{A}_1$ is given by

$$\mathbf{A}_1 = \sum_{\lambda=\pm 1} \Big\{ C(\lambda) \int \frac{d\mathbf{p}}{p} f^{(1)}(\mathbf{p}, \lambda) \boldsymbol{\sigma}(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x} - pt)} + D(\lambda) \int \frac{d\mathbf{p}}{p} h^{(1)*}(\mathbf{p}, \lambda) \boldsymbol{\sigma}^*(\mathbf{p}, \lambda) e^{-i(\mathbf{p}\cdot\mathbf{x} - pt)} \Big\}, \tag{4.9''}$$

where in (4.9'') we have absorbed a factor of $(2)^{\frac{1}{2}}$ into the yet undetermined constants $C(\lambda)$ and $D(\lambda)$, and where $\boldsymbol{\sigma}$ is the vector whose components are σ_i.

The expansion for A_0 can also be made simpler by expressing the wavefunctions for zero spin as linear combinations of other zero-spin wavefunctions. Accordingly, we introduce the wavefunctions $g^{(r)}(\mathbf{p})$, $k^{(r)}(\mathbf{p})$ by

$$\begin{aligned} E^{(0)} g^{(0)}(\mathbf{p}) &= \tfrac{1}{2}[C^{(1)}(0) f^{(1)}(\mathbf{p}, 0) + C^{(0)}(0) f^{(0)}(\mathbf{p}, 0)], \\ E^{(1)} g^{(1)}(\mathbf{p}) &= \tfrac{1}{2}[C^{(1)}(0) f^{(1)}(\mathbf{p}, 0) - C^{(0)}(0) f^{(0)}(\mathbf{p}, 0)], \\ F^{(0)} k^{(0)}(\mathbf{p}) &= -\tfrac{1}{2}[D^{(1)}(0) h^{(1)}(\mathbf{p}, 0) - D^{(0)} h^{(0)}(\mathbf{p}, 0)], \\ F^{(1)} k^{(1)}(\mathbf{p}) &= -\tfrac{1}{2}[D^{(1)}(0) h^{(1)}(\mathbf{p}, 0) + D^{(0)} h^{(0)}(\mathbf{p}, 0)], \end{aligned} \tag{4.9'''}$$

where $E^{(r)}$ and $F^{(r)}$ are arbitrary positive constants. It is clear that the functions $g^{(r)}(\mathbf{p})$ and $k^{(r)}(\mathbf{p})$ transform as massless particles of spin zero.

We also introduce the column vectors $\xi^{(r)}(\mathbf{p})$ by

$$\xi^{(0)}(\mathbf{p}) = \chi^{(1)}(\mathbf{p}, 0) + \chi^{(0)}(\mathbf{p}, 0) = \begin{pmatrix} p \\ p_1 \\ p_2 \\ p_3 \end{pmatrix},$$

$$\xi^{(1)}(\mathbf{p}) = \chi^{(1)}(\mathbf{p}, 0) - \chi^{(0)}(\mathbf{p}, 0) = \frac{1}{p^2} \begin{pmatrix} -p \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}. \tag{4.9''''}$$

Then A_0 has the expansion

$$A_0(x) = \sum_r \int \frac{d\mathbf{p}}{p} \xi^{(r)}(\mathbf{p}) [E^{(r)} g^{(r)}(\mathbf{p}) e^{i(\mathbf{p}\cdot\mathbf{x} - pt)} + F^{(r)} k^{(r)*}(\mathbf{p}) e^{-i(\mathbf{p}\cdot\mathbf{x} - pt)}]. \tag{4.9'''''}$$

Then to summarize: The four-vector written as the column vector A is split into two parts and in (4.9) with A_1 given by (4.9') and (4.9''), while A_0 is given by (4.9''''').

Now we write the expressions for the gauge change. We introduce the column vector Ξ_n. We can split it up as (4.9) corresponding to the separation into contributions from zero-spin wavefunctions and spin-1 wavefunctions. The expansions for each part of Ξ_n are the same as for the corresponding part of A [Eqs. (4.9'') and (4.9''''')] except that $d\mathbf{p}/p$ is replaced by $d\mathbf{p}/p^{n+1}$.

Let us denote the top component of $\Xi_n(x)$ by $\Xi_n^0(x)$ and use the three remaining space components to form the vector $\boldsymbol{\Xi}_n(x)$. Likewise, if $G_i(x)$ denotes the column vector which gives the infinitesimal gauge change, we introduce $G_i^0(x)$ and $\mathbf{G}_i(x)$ as the top component and vector formed of the space components, respectively. Then the gauge change is given by

$$\begin{aligned} G_i^0(x) &= \frac{\partial^2}{\partial x_i \partial t} [\boldsymbol{\nabla} \cdot \boldsymbol{\Xi}_3(x)] + \frac{\partial}{\partial t} \Xi_{1,i}(x), \\ \mathbf{G}_i(x) &= \frac{\partial^2}{\partial x_i \partial t} \boldsymbol{\nabla} \Xi_3^0(x) - \boldsymbol{\nabla} \Xi_{1,i}(x) + \mathbf{e}_i [\boldsymbol{\nabla} \cdot \boldsymbol{\Xi}_1(x)] + \mathbf{e}_i \frac{\partial}{\partial t} \Xi_1^0(x). \end{aligned} \tag{4.10}$$

In (4.10), the quantity $\Xi_{n,i}(x)$ is the ith component of the vector $\boldsymbol{\Xi}_n(x)$.

B. The Electromagnetic Vector Potential

Now we specialize the results to the electromagnetic vector potential. First we show the effect of imposing the Lorentz condition and the reality condition on the column vector A separately. Then we impose these conditions together. In this latter case, the

components of the column vector can be considered as the components of an electromagnetic vector potential.

First of all, we note that the components of A_1 already satisfy the Lorentz condition on account of the second of Eq. (3.8). Thus, only the zero-spin wavefunctions contained in A_0 are affected by the Lorentz condition. Indeed, it is seen that the imposition of the Lorentz condition leads to

$$g^{(1)}(\mathbf{p}) = k^{(1)}(\mathbf{p}) = 0. \tag{4.11}$$

The simplicity of (4.11) was the chief reason for the introduction of $g^{(r)}$ and $k^{(r)}$.

The components of A_0 then take on a particularly simple form:

$$A_0^\gamma(x) = -i\frac{\partial}{\partial x_\gamma} F(x), \tag{4.12}$$

where $F(x)$ is a scalar function given by

$$F(x) = \int \frac{d\mathbf{p}}{p} [E^{(0)} g^{(0)}(\mathbf{p}) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} - F^{(0)} k^{(0)*}(\mathbf{p}) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}]. \tag{4.12'}$$

Since the scalar $F(x)$ satisfies the zero-mass wave equation, it is clearly seen that A_0 represents a gauge change.

It is easy to show that, when the Lorentz condition is satisfied, A_0 makes no contribution whatever to the gauge change column vectors G_i. The expressions for the gauge change simplify considerably and become

$$G_i^\gamma(x) = -\frac{\partial}{\partial x_\gamma} \Xi_{1,i}(x), \tag{4.13}$$

where the vector Ξ_1 has the same expansion as $\mathbf{A}_1$ [Eq. (4.9b)] but where $d\mathbf{p}/p$ is replaced by $d\mathbf{p}/p^2$.

We now see the effect of imposing the reality condition (without, however, imposing the Lorentz condition). It is readily seen that the requirement that the components of the column vector A be real, leads to

$$\begin{aligned} E^{(r)} g^{(r)}(\mathbf{p}) &= F^{(r)} k^{(r)}(\mathbf{p}), \\ C(\lambda) f^{(1)}(\mathbf{p}, \lambda) &= D(\lambda) h^{(1)}(\mathbf{p}, \lambda). \end{aligned} \tag{4.14}$$

We now impose both the reality condition and the Lorentz condition on the components of A. We can summarize the results as follows:

$$\begin{aligned} A(x) &= A_0(x) + A_1(x), \\ A_0(x) &= E\int \frac{d\mathbf{p}}{p} \xi(\mathbf{p})[g(\mathbf{p}) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} + g^*(\mathbf{p}) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}], \\ A_1(x) &= \begin{pmatrix} 0 \\ \mathbf{A}_1(x) \end{pmatrix}, \\ \mathbf{A}_1(x) &= \sum_{\lambda=\pm 1} C(\lambda) \int \frac{d\mathbf{p}}{p} [f(\mathbf{p}, \lambda)\boldsymbol{\sigma}(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} \\ &\quad + f^*(\mathbf{p}, \lambda)\boldsymbol{\sigma}^*(\mathbf{p}, \lambda) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}], \\ \Xi_1(x) &= \sum_{\lambda=\pm 1} C(\lambda) \int \frac{d\mathbf{p}}{p^2} [f(\mathbf{p}, \lambda)\boldsymbol{\sigma}(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} \\ &\quad + f^*(\mathbf{p}, \lambda)\boldsymbol{\sigma}^*(\mathbf{p}, \lambda) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}], \end{aligned} \tag{4.15}$$

where

$$\xi = \xi^{(0)}, \tag{4.15'}$$

and E is a positive constant. The complex functions $g(\mathbf{p})$ and $f(\mathbf{p}, \lambda)$ are functions which transform as single-spin wavefunctions of spin zero and spin 1 with helicity λ, respectively. The gauge change is given by (4.13). Now we can find the electromagnetic fields by

$$\begin{aligned} \mathbf{E}(x) &= -\frac{\partial}{\partial t} \mathbf{A}_1(x), \\ \mathbf{H}(x) &= \nabla \times \mathbf{A}_1(x). \end{aligned} \tag{4.16}$$

(In obtaining the electromagnetic fields from the vector potential, we recollect that part of the four-vector given by A_0 is simply a gauge term.)

We find the constants $C(\lambda)$ by requiring that, when $f(\mathbf{p}, -1) = 0$ or $f(\mathbf{p}, +1) = 0$, we have

$$(8\pi)^{-1} \int [\mathbf{E}^2 + \mathbf{H}^2]\, d\mathbf{x} = \int |f(\mathbf{p}, +1)|^2\, d\mathbf{p}$$

or

$$\int |f(\mathbf{p}, -1)|^2\, d\mathbf{p},$$

respectively, in accordance with our requirement that the total energy of the field expressed in terms of field variables equals the expectation value of the energy in terms of wavefunctions when only one mode is present. We find that

$$C(\lambda) = (8\pi^2)^{-\frac{1}{2}}. \tag{4.17}$$

On picking the constant E to be

$$E = [4(\pi)^{\frac{3}{2}}]^{-1}, \tag{4.17'}$$

we see that the expansion given by (4.15) is identical to the expansion (3.14) of Ref. 4 when a slightly different notation is used.

We can compare the expansion of the electromagnetic fields, as given by (4.16) when the constants are given by (4.17) and (4.17') with the expansion (3.16) of the present paper. We see that

$$f(\mathbf{p}, \lambda) = g(\mathbf{p}, \lambda). \tag{4.18}$$

Equation (4.18) enables us to obtain the vector potential (within a gauge) corresponding to any field which satisfies Maxwell's equations.

The second quantization of the vector potential and the relativistic character of the second-quantized theory is discussed in great detail in Ref. 4, and we do not repeat it here.

5. WAVEFUNCTIONS WHICH TRANSFORM AS A DIRAC SPINOR

We use the notation of Sec. 6 of Part I. In particular, it is useful to note that the Dirac α matrices are given

in terms of the Pauli σ matrices and the matrices N_i by

$$\alpha_j = 2iN_j = \begin{pmatrix} 0_j & \sigma_j \\ \sigma_j & 0_j \end{pmatrix}, \tag{5.1}$$

and that the Dirac Hamiltonian H is given by

$$H = -i\boldsymbol{\alpha} \cdot \nabla. \tag{5.2}$$

A. Forms for the Expansion of the Wavefunction: Solutions of the Massless Dirac Equation

Since from (I 6.1) and (I 6.5), M_3 already appears in a reduced, diagonal form, we may take the matrix V of (2.17) to be the identity matrix. We note that the label r can take on two values which we call 1 and 2. The variable λ also takes on the two values $\frac{1}{2}$ and $-\frac{1}{2}$. From (2.17), (I 6.6), and (I 6.1), we obtain the following expressions for the column vectors $\chi^{(r)}$:

$$\chi^{(1)}(\epsilon, \mathbf{p}, +\tfrac{1}{2}) = \{2p[2(p+p_3)]^{\frac{1}{2}}\}^{-1} \begin{pmatrix} (p+1)(p+p_3) \\ (p+1)(p_1+ip_2) \\ \epsilon(p-1)(p+p_3) \\ \epsilon(p-1)(p_1+ip_2) \end{pmatrix}, \tag{5.3}$$

$$\chi^{(1)}(\epsilon, \mathbf{p}, -\tfrac{1}{2}) = \{2p[2(p+p_3)]^{\frac{1}{2}}\}^{-1} \begin{pmatrix} -(p+1)(p_1-ip_2) \\ (p+1)(p+p_3) \\ \epsilon(p-1)(p_1-ip_2) \\ -\epsilon(p-1)(p+p_3) \end{pmatrix}, \tag{5.4}$$

$$\chi^{(2)}(\epsilon, \mathbf{p}, +\tfrac{1}{2}) = \{2p[2(p+p_3)]^{\frac{1}{2}}\}^{-1} \begin{pmatrix} \epsilon(p-1)(p+p_3) \\ \epsilon(p-1)(p_1+ip_2) \\ (p+1)(p+p_3) \\ (p+1)(p_1+ip_2) \end{pmatrix}, \tag{5.5}$$

$$\chi^{(2)}(\epsilon, \mathbf{p}, -\tfrac{1}{2}) = \{2p[2(p+p_3)]^{\frac{1}{2}}\}^{-1} \begin{pmatrix} \epsilon(p-1)(p_1-ip_2) \\ -\epsilon(p-1)(p+p_3) \\ -(p+1)(p_1-ip_2) \\ (p+1)(p+p_3) \end{pmatrix}. \tag{5.6}$$

As usual, we define $\chi^{(r)}(\mathbf{p}, \lambda)$ by

$$\chi^{(r)}(\mathbf{p}, \lambda) = \chi^{(r)}(+1, \mathbf{p}, \lambda)$$

and find that the column vectors $\zeta^{(r)}$ are given by

$$\zeta^{(r)}(\mathbf{p}, \lambda) = -2\lambda\chi^{(r)*}(\mathbf{p}, -\lambda). \tag{5.7}$$

For future convenience, we note the important relations

$$\begin{aligned} \mathbf{p} \cdot \boldsymbol{\alpha}\chi^{(1)}(\mathbf{p}, \lambda) &= 2\lambda p\chi^{(2)}(\mathbf{p}, \lambda), \\ \mathbf{p} \cdot \boldsymbol{\alpha}\chi^{(2)}(\mathbf{p}, \lambda) &= 2\lambda p\chi^{(1)}(\mathbf{p}, \lambda). \end{aligned} \tag{5.8}$$

The expansion (2.26) can be simplified [in view of (5.7)] by replacing the wavefunction $h^{(r)}(\mathbf{p}, \lambda)$ by $-2\lambda h^{(r)}(\mathbf{p}, \lambda)$. This substitution represents a trivial change in the phase of the wavefunction. Then the expansion (2.26) can be written

$$\Psi(x) = \sum_r \sum_\lambda \int \frac{d\mathbf{p}}{p} \{C^{(r)}(\lambda) f^{(r)}(\mathbf{p}, \lambda)\chi^{(r)}(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} + D^{(r)}(\lambda) h^{*}(\mathbf{p}, \lambda)\chi^{(r)}(\mathbf{p}, -\lambda) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}\}. \tag{5.9}$$

We now impose the requirement that $\Psi(x)$ satisfies the Dirac equation

$$H\Psi = i\frac{\partial}{\partial t}\Psi.$$

We find from (5.8) that to satisfy the Dirac equation, we must have

$$\begin{aligned} C^{(2)}(\lambda) f^{(2)}(\mathbf{p}, \lambda) &= 2\lambda C^{(1)}(\lambda) f^{(1)}(\mathbf{p}, \lambda), \\ D^{(2)}(\lambda) h^{(2)}(\mathbf{p}, \lambda) &= -2\lambda D^{(1)}(\lambda) h^{(1)}(\mathbf{p}, \lambda). \end{aligned} \tag{5.10}$$

Equations (5.10) suggest another way of writing the expansion (5.9) so that the wavefunction $\Psi(x)$ can be expressed as the sum of two column vectors, each of which transforms like a Dirac spinor, but only one of which satisfies the Dirac equation. Accordingly, let us define the wavefunctions $g^{(r)}(\mathbf{p}, \lambda)$ and $k^{(r)}(\mathbf{p}, \lambda)$ by

$$\begin{aligned} E^{(1)}(\lambda) g^{(1)}(\mathbf{p}, \lambda) &= \tfrac{1}{2}[C^{(1)}(\lambda) f^{(1)}(\mathbf{p}, \lambda) + 2\lambda C^{(2)}(\lambda) f^{(2)}(\mathbf{p}, \lambda)], \\ E^{(2)}(\lambda) g^{(2)}(\mathbf{p}, \lambda) &= \tfrac{1}{2}[C^{(1)}(\lambda) f^{(1)}(\mathbf{p}, \lambda) - 2\lambda C^{(2)}(\lambda) f^{(2)}(\mathbf{p}, \lambda)], \end{aligned} \tag{5.11}$$

$$\begin{aligned} F^{(1)}(\lambda) k^{(1)}(\mathbf{p}, \lambda) &= \tfrac{1}{2}[D^{(1)}(\lambda) h^{(1)}(\mathbf{p}, \lambda) - 2\lambda D^{(2)}(\lambda) h^{(2)}(\mathbf{p}, \lambda)], \\ F^{(2)}(\lambda) k^{(2)}(\mathbf{p}, \lambda) &= \tfrac{1}{2}[D^{(1)}(\lambda) h^{(1)}(\mathbf{p}, \lambda) + 2\lambda D^{(2)}(\lambda) h^{(2)}(\mathbf{p}, \lambda)]. \end{aligned} \tag{5.11'}$$

In Eqs. (5.11) and (5.11'), the quantities $E^{(r)}(\lambda)$ and $F^{(r)}(\lambda)$ are arbitrary positive constants which we introduce for later convenience.

The expansion (5.9) in terms of the new wavefunctions can be written as

$$\Psi(x) = \Psi^{(1)}(x) + \Psi^{(2)}(x), \tag{5.12}$$

where

$$\Psi^{(r)}(x) = \sum_\lambda \int \frac{d\mathbf{p}}{p} \{E^{(r)}(\lambda) g^{(r)}(\mathbf{p}, \lambda)\xi^{(r)}(\mathbf{p}, \lambda) e^{i(\mathbf{p}\cdot\mathbf{x}-pt)} + F^{(r)}(\lambda) k^{(r)*}(\mathbf{p}, \lambda)\xi^{(r)}(\mathbf{p}, -\lambda) e^{-i(\mathbf{p}\cdot\mathbf{x}-pt)}\}, \tag{5.12'}$$

where $\xi^{(r)}(\mathbf{p}, \lambda)$ is a column vector defined by

$$\begin{aligned} \xi^{(1)}(\mathbf{p}, \lambda) &= \chi^{(1)}(\mathbf{p}, \lambda) + 2\lambda\chi^{(2)}(\mathbf{p}, \lambda), \\ \xi^{(2)}(\mathbf{p}, \lambda) &= \chi^{(1)}(\mathbf{p}, \lambda) - 2\lambda\chi^{(2)}(\mathbf{p}, \lambda). \end{aligned} \tag{5.12''}$$

From (5.8), it is easily shown that

$$\begin{aligned}\mathbf{p}\cdot\boldsymbol{\alpha}\xi^{(1)}(\mathbf{p},\lambda) &= p\xi^{(1)}(\mathbf{p},\lambda),\\ \mathbf{p}\cdot\boldsymbol{\alpha}\xi^{(2)}(\mathbf{p},\lambda) &= -p\xi^{(2)}(\mathbf{p},\lambda).\end{aligned} \tag{5.13}$$

From (5.13), it is seen that $\Psi^{(1)}$ satisfies Dirac's equation but that $\Psi^{(2)}$ does not. Hence, the condition that Ψ satisfies Dirac's equation is

$$g^{(2)}(\mathbf{p},\lambda) = k^{(2)}(\mathbf{p},\lambda) = 0,$$

which is entirely equivalent to the condition that $\Psi^{(2)}(x) = 0$.

The wavefunction $\psi^{(2)}$ satisfies the "anti-Dirac equation"

$$H\Psi^{(2)}(x) = -i\frac{\partial}{\partial t}\Psi^{(2)}(x).$$

Because the column vectors $\xi^{(r)}(\mathbf{p},\lambda)$ play such an important role, we shall give them explicitly:

$$\xi^{(1)}(\mathbf{p},\tfrac{1}{2}) = [2(p+p_3)]^{-\frac{1}{2}}\begin{pmatrix} p+p_3\\ p_1+ip_2\\ p+p_3\\ p_1+ip_2\end{pmatrix},$$

$$\xi^{(1)}(\mathbf{p},-\tfrac{1}{2}) = [2(p+p_3)]^{-\frac{1}{2}}\begin{pmatrix} -(p_1-ip_2)\\ p+p_3\\ p_1-ip_2\\ -(p+p_3)\end{pmatrix},$$

$$\xi^{(2)}(\mathbf{p},\tfrac{1}{2}) = [2p^2(p+p_3)]^{-\frac{1}{2}}\begin{pmatrix} p+p_3\\ p_1+ip_2\\ -(p+p_3)\\ -(p_1+ip_2)\end{pmatrix},$$

$$\xi^{(2)}(\mathbf{p},-\tfrac{1}{2}) = [2p^2(p+p_3)]^{-\frac{1}{2}}\begin{pmatrix} -(p_1-ip_2)\\ p+p_3\\ -(p_1-ip_2)\\ p+p_3\end{pmatrix}. \tag{5.14}$$

B. Gauge Change

Let us split up the column vector $\Xi_n(x)$ in a manner analogous to the splitting up of $\Psi(x)$ given by Eqs. (5.12) and (5.12′):

$$\Xi_n(x) = \Xi_n^{(1)}(x) + \Xi_n^{(2)}(x), \tag{5.15}$$

where $\Xi_n^{(r)}(x)$ is given by expressions analogous to those to (5.12′) for $\Psi^{(r)}(x)$ but where, as usual, $d\mathbf{p}/p$ is replaced by $d\mathbf{p}/p^{n+1}$. From the fact that $\Xi_n^{(1)}$ satisfies Dirac's equation and that $\Xi_n^{(2)}$ satisfies the anti-Dirac equation, we obtain the following very simple expression for the infinitesimal gauge changes $G_i(x)$; namely,

$$G_j(x) = -2i(\mathbf{M}\times\boldsymbol{\nabla})_j\Xi_1^{(2)}(x). \tag{5.16}$$

This result leads to the important theorem that *for solutions $\Psi(x)$ of Dirac's equation, the gauge change is zero.* Thus, space-time transformations of the wavefunctions $g^{(1)}(\mathbf{p},\lambda)$, $k^{(1)}(\mathbf{p},\lambda)$ map into the transformed Ψ without the necessity of taking into account a gauge change.

C. Relation to the Canonical Formalism and Second Quantization

In the present section, we restrict our discussion to solutions of Dirac's equations. Thus, we take $\Psi(x) = \Psi^{(1)}(x)$, where $\Psi^{(1)}(x)$ is expanded as in (5.12′). On using the Hamiltonian density (I 6.17) and equating the energy of the field with the expectation value of the energy in the particle picture, when only single modes are used, we obtain

$$E^{(1)}(\lambda) = F^{(1)}(\lambda) = (4\pi^{\frac{3}{2}})^{-1}. \tag{5.17}$$

To second quantize the theory, we replace $g^{(1)}(\mathbf{p},\lambda)$ and $k^{(1)}(\mathbf{p},\lambda)$ by destruction operators and their complex conjugates by creation operators which satisfy the commutation relations,

$$\begin{aligned}[g^{(1)}(\mathbf{p},\lambda), g^{(1)}(\mathbf{p}',\lambda')]_+ &= [k^{(1)}(\mathbf{p},\lambda), k^{(1)}(\mathbf{p}',\lambda')]_+ = 0,\\ [g^{(1)}(\mathbf{p},\lambda), k^{(1)}(\mathbf{p}',\lambda')]_+ &= [g^{(1)}(\mathbf{p},\lambda), k^{(1)*}(\mathbf{p}',\lambda')]_+ = 0,\\ [g^{(1)}(\mathbf{p},\lambda), g^{(1)*}(\mathbf{p}',\lambda')]_+ &= [k^{(1)}(\mathbf{p},\lambda), k^{(1)*}(\mathbf{p}',\lambda')]_+\\ &= p\delta_{\lambda,\lambda'}\delta(\mathbf{p}-\mathbf{p}').\end{aligned} \tag{5.18}$$

The column vector $\Psi(x)$ now has operators for components. We shall designate these operators by $\Psi(\mathbf{x},t,\gamma)$. Then from (5.12′), (5.17), and (5.18), we obtain the commutation rules for these operators:

$$\begin{aligned}&[\Psi(\mathbf{x},t,\gamma), \Psi(\mathbf{x}',t',\gamma')]_+ = 0,\\ &[\Psi(\mathbf{x},t,\gamma), \Psi^*(\mathbf{x}',t',\gamma')]_+\\ &\quad= \left[(\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})_{\gamma,\gamma'} - \delta_{\gamma,\gamma'}\frac{\partial}{\partial t}\right]D_0(\mathbf{x}-\mathbf{x}',t-t'),\end{aligned} \tag{5.19}$$

where $D_0(\mathbf{x},t)$ is the well-known Pauli–Jordan invariant function

$$D_0(x,t) = (4\pi|\mathbf{x}|)^{-1}[\delta(|\mathbf{x}|-t) - \delta(|\mathbf{x}|+t)], \tag{5.19′}$$

and $(\boldsymbol{\alpha}\cdot\boldsymbol{\nabla})_{\gamma,\gamma'}$ is the matrix element of $\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}$.

To show the invariance of the second quantized theory, we introduce for every single-particle infinitesimal generator $\hat{A}$, a second quantized operator $[A]$ defined by

$$\begin{aligned}[A] = \sum_\lambda\int\frac{d\mathbf{p}}{p}\{&g^{(1)*}(\mathbf{p},\lambda)\hat{A}g^{(1)}(\mathbf{p},\lambda)\\ &+ k^{(1)*}(\mathbf{p},\lambda)\hat{A}k^{(1)}(\mathbf{p},\lambda)\},\end{aligned} \tag{5.20}$$

where the expressions $\hat{A}g^{(1)}(\mathbf{p},\lambda)$ and $\hat{A}k^{(1)}(\mathbf{p},\lambda)$ indicate that the operator $\hat{A}$ acts on the variables $\mathbf{p}$, λ

as though $g^{(1)}$ and $k^{(1)}$ were single-particle wavefunctions of positive energy and helicity λ.

Then one requires that under the transformations $T(a^\alpha)$, $R(\boldsymbol{\theta})$, $L(\boldsymbol{\beta})$, the column vector $\Psi(x)$ constructed from the four operators $\Psi(\mathbf{x}, t, \gamma)$ transform to the column vector $\Psi'(x)$ with corresponding operator components, where $\Psi'(x)$ is given by (I 2.10)–(I 2.12), respectively, with (I 6.6) to be used to simplify the expressions. Invariance is shown by noting that Ψ' can be obtained from Ψ through the use of unitary transformations constructed from the second quantized infinitesimal generators.

Thus, under the transformation $T(a^\alpha)$, $R(\boldsymbol{\theta})$, and $L(\boldsymbol{\beta})$, the column vector $\Psi'(x)$ is given by

$$\begin{aligned}
&\Psi'(x) = \exp\left\{-i\sum_\alpha a^\alpha[P_\alpha]\right\}\Psi(x)\exp\left\{i\sum_\alpha a^\alpha[P_\alpha]\right\},\\
&\Psi'(x) = \exp\{-i\boldsymbol{\theta}\cdot[\mathbf{J}]\}\Psi(x)\exp\{i\boldsymbol{\theta}\cdot[\mathbf{J}]\},\\
&\Psi'(x) = \exp\{-i\boldsymbol{\beta}\cdot[\mathfrak{J}]\}\Psi(x)\exp\{i\boldsymbol{\beta}\cdot[\mathfrak{J}]\},
\end{aligned} \tag{5.21}$$

respectively, where the equations of (5.21) are meant to hold for each component of the four-vector.

APPENDIX A: INTRODUCTION OF ANTIPARTICLE WAVEFUNCTIONS

Let us consider a space of functions $\{f(\mathbf{p})\}$, such that each member is a function of three continuous variables p_i (collectively denoted by the vector $\mathbf{p}$), each of which can take on any value on the real axis.

Then the realization of finite-spin representations of the infinitesimal generators of the inhomogeneous Lorentz group which is given in Ref. 9 is (for helicity λ and sign of energy ϵ)

$$\begin{aligned}
\hat{P}^0 f(\mathbf{p}) &= \epsilon p f(\mathbf{p}),\\
\hat{P}^i f(\mathbf{p}) &= p_i f(\mathbf{p}),\\
\hat{J}_1 f(\mathbf{p}) &= \left[-i(\mathbf{p}\times\boldsymbol{\nabla})_1 + \frac{p_1}{p+p_3}\lambda\right]f(\mathbf{p}),\\
\hat{J}_2 f(\mathbf{p}) &= \left[-i(\mathbf{p}\times\boldsymbol{\nabla})_2 + \frac{p_2}{p+p_3}\lambda\right]f(\mathbf{p}),\\
\hat{J}_3 f(\mathbf{p}) &= [-i(\mathbf{p}\times\boldsymbol{\nabla})_3 + \lambda]f(\mathbf{p}),\\
\hat{\mathfrak{J}}_1 f(\mathbf{p}) &= \epsilon\left[ip\nabla_1 + \frac{p_2}{p+p_3}\lambda\right]f(\mathbf{p}),\\
\hat{\mathfrak{J}}_2 f(\mathbf{p}) &= \epsilon\left[ip\nabla_2 - \frac{p_1}{p+p_3}\lambda\right]f(\mathbf{p}),\\
\hat{\mathfrak{J}}_3 f(\mathbf{p}) &= \epsilon ip\nabla_3 f(\mathbf{p}).
\end{aligned} \tag{A1}$$

In (A1), we use $\nabla_i = \partial/\partial p_i$. If we introduce as the inner product of two functions $f^{(1)}$ and f of the space, the expression $\int (d\mathbf{p}/p) f^{(1)*}(\mathbf{p}) f(\mathbf{p})$, the infinitesimal generators are Hermitian. Of course, the realization which we have just given is identical to the realizations given by (2.3) when we use Type II representations for the Euclidean group. Let us regard the functions of the space as being wavefunctions. Then, as shown in Ref. 11, under the transformation $T(a^\alpha)$, the wavefunction which in the original frame of reference was $f(\mathbf{p})$ becomes in the new frame $f'(\mathbf{p})$, where

$$f'(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha \hat{P}_\alpha\right]f(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha p_\alpha\right]f(\mathbf{p}),$$
$$p_0 = -p^0 = -\epsilon p. \tag{A2}$$

Under the rotation $R(\boldsymbol{\theta})$, the wavefunction in the new frame is given by

$$f'(\mathbf{p}) = \exp(i\boldsymbol{\theta}\cdot\hat{\mathbf{J}})f(\mathbf{p}) = \exp[2i\lambda\zeta(\boldsymbol{\theta},\mathbf{p})]f(\mathbf{p}'), \tag{A3}$$

where $\mathbf{p}'$ is given by

$$\mathbf{p}' = \mathbf{p}\cos\theta + [(1-\cos\theta)/\theta^2]\times(\boldsymbol{\theta}\cdot\mathbf{p})\boldsymbol{\theta} + (\sin\theta/\theta)(\boldsymbol{\theta}\times\mathbf{p}), \tag{A3'}$$

[cf. (I 2.2)] and where $\zeta(\boldsymbol{\theta},\mathbf{p})$ is given by

$$\tan\zeta(\boldsymbol{\theta},\mathbf{p}) = \frac{[\boldsymbol{\theta}\cdot\mathbf{p}+\theta_3 p]\tan(\theta/2)}{\theta(p+p_3)+(\boldsymbol{\theta}\times\mathbf{p})_3\tan(\theta/2)}. \tag{A3''}$$

Under the pure Lorentz transformation $L(\boldsymbol{\beta})$, the wavefunction in the new frame is given by

$$f'(\mathbf{p}) = \exp(i\boldsymbol{\beta}\cdot\hat{\mathfrak{J}})f(\mathbf{p}) = \exp[2i\lambda\epsilon\sigma(\boldsymbol{\beta},\mathbf{p},\epsilon)]f(\mathbf{p}'), \tag{A4}$$

where p' is given by

$$\mathbf{p}' = \mathbf{p} + \boldsymbol{\beta}(\boldsymbol{\beta}\cdot\mathbf{p})[(\cosh\beta - 1)/\beta^2] - \boldsymbol{\beta}\epsilon p(\sinh\beta/\beta) \tag{A4'}$$

[cf. the second of Eqs. (I 2.3)] and where $\sigma(\boldsymbol{\beta},\mathbf{p},\epsilon)$ is given by

$$\tan\sigma(\boldsymbol{\beta},\mathbf{p},\epsilon) = \frac{(\boldsymbol{\beta}\times\mathbf{p})_3\tanh(\beta/2)}{\beta(p+p_3)-\epsilon(p\beta_3+\mathbf{p}\cdot\boldsymbol{\beta})\tanh(\beta/2)}. \tag{A4''}$$

There is a second realization of the infinitesimal generators of the inhomogeneous Lorentz group for finite spin with given sign of energy ϵ and helicity λ which is useful. This realization is, of course, unitarily equivalent to the previous one. It is best introduced through the use of a second space of functions $\{g(\mathbf{p})\}$ which is in a one-to-one correspondence with the functions of the set $\{f(\mathbf{p})\}$ through the relation

$$g(\mathbf{p}) = e^{2i\lambda\varphi}f(\mathbf{p}), \tag{A5}$$

where φ is defined by

$$\tan\varphi = p_2/p_1. \tag{A5'}$$

Let $\hat{A}$ be any of the infinitesimal generators as given in the realization (A1). Let us denote the corresponding operator in the new realization by $\tilde{A}$ which

[11] H. E. Moses, Ann. Phys. (N.Y.) **41**, 158 (1967).

is defined by

$$\tilde{A}g(\mathbf{p}) = e^{2i\lambda\varphi}\hat{A}f(\mathbf{p}). \qquad (A5'')$$

Then the second realization of the infinitesimal generators is

$$\begin{aligned}
\tilde{P}^0 g(\mathbf{p}) &= \epsilon p g(\mathbf{p}),\\
\tilde{P}^i g(\mathbf{p}) &= p_i g(\mathbf{p}),\\
\tilde{J}_1 g(\mathbf{p}) &= \{-i(\mathbf{p}\times\nabla)_1 + [p_1/(p-p_3)]\lambda\}g(\mathbf{p}),\\
\tilde{J}_2 g(\mathbf{p}) &= \{-i(\mathbf{p}\times\nabla)_2 + [p_2/(p-p_3)]\lambda\}g(\mathbf{p}),\\
\tilde{J}_3 g(\mathbf{p}) &= [-i(\mathbf{p}\times\nabla)_3 - \lambda]g(\mathbf{p}),\\
\tilde{\mathcal{J}}_1 g(\mathbf{p}) &= \epsilon\{ip\nabla_1 - [p_2/(p-p_3)]\lambda\}g(\mathbf{p}),\\
\tilde{\mathcal{J}}_2 g(\mathbf{p}) &= \epsilon\{ip\nabla_2 + [p_1/(p-p_3)]\lambda\}g(\mathbf{p}),\\
\tilde{\mathcal{J}}_3 g(\mathbf{p}) &= \epsilon ip\nabla_3 g(\mathbf{p}).
\end{aligned} \qquad (A6)$$

If we consider $g(\mathbf{p})$ to be a wavefunction, we can find from (A6) the wavefunction $g'(\mathbf{p})$ in the new frame of reference under finite transformations of the inhomogeneous Lorentz group. One simply integrates the infinitesimal generators of the realization of (A6) after the manner of Ref. 11. Thus, for the translation $T(a^\alpha)$, the new wavefunction is

$$g'(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha \tilde{P}_\alpha\right]g(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha p_\alpha\right]g(\mathbf{p}); \qquad (A7)$$

for the rotation $R(\boldsymbol{\theta})$, the new function is

$$g'(\mathbf{p}) = e^{i\boldsymbol{\theta}\cdot\tilde{\mathbf{J}}}g(\mathbf{p}) = \exp\left[-2i\lambda\zeta(\boldsymbol{\theta}, -\mathbf{p})\right]g(\mathbf{p}'), \qquad (A8)$$

where $\mathbf{p}'$ and ζ are given by (A3′) and (A3″), respectively, and finally, for the pure Lorentz transformation $L(\boldsymbol{\beta})$, the new wavefunction is

$$\begin{aligned}
g'(\mathbf{p}) &= \exp(i\boldsymbol{\beta}\cdot\tilde{\mathcal{J}})g(\mathbf{p})\\
&= \exp\left[2i\lambda\epsilon\sigma(\boldsymbol{\beta}, -\mathbf{p}, -\epsilon)\right]g(\mathbf{p}'),
\end{aligned} \qquad (A9)$$

where $\mathbf{p}'$ and σ are given by (A4′) and (A4″), respectively.

By comparing the finite transformation formulas for the two realizations, one sees that the function $f^*(-\mathbf{p})$ transforms like one of the functions of the set $\{g(\mathbf{p})\}$ for which the helicity is still λ but *for which the sign of energy is* $-\epsilon$. Then from (A5), we see that $e^{-2i\lambda\varphi}f^*(-\mathbf{p})$ transforms like wavefunctions belonging to a realization of the first kind with helicity λ, but with the sign of energy equal to $-\epsilon$. This observation is the reason for the use of Eq. (2.12).

APPENDIX B: DERIVATION OF THE FIRST FORM OF THE REDUCTION ALGORITHM

The algorithm is a direct consequence of the "recipe" for reduction of reducible representations of the inhomogeneous Lorentz group of Ref. 3. We want to consider Eq. (1.14) of Ref. 3 for the case that $\mu = 0$. This expansion and the transformation function $\langle\zeta \mid 0, \epsilon, \mathbf{p}, \lambda\rangle$ have been derived under the assumption that all of the infinitesimal generators of the inhomogeneous Lorentz group are Hermitian as expressed in the ζ representation. The formulas for the matrices which are a representation of the Euclidean group were also derived under this assumption.

However, when one examines the proof of the reduction, one sees that the expansion (1.14) of Ref. 3 and the technique of obtaining the transformation functions and Euclidean group are valid under the less severe restrictions that only the Hamiltonian H, the momentum operators P_i, and the angular momentum operator J_i need be Hermitian. The requirement on the operators $\mathcal{J}_i$ are weakened to the requirement that they can be exponentiated. In the expansion (1.14) of Ref. 3, we need only replace the transformation functions $\langle\zeta \mid 0, \epsilon, \mathbf{p}, \lambda\rangle$ by the transformation function $\langle\zeta \mid 0, \epsilon, p, \lambda\rangle$ which is constructed using Eq. (1.24) of Ref. 3. The wavefunctions $F(0, \epsilon, \mathbf{p}, \lambda)$ still transform under the infinitesimal Lorentz transformations according to (1.9) of Ref. 3; that is, they constitute a Hilbert space such that the realizations of the infinitesimal generators are of the form (2.3) of the present paper, where the matrices T_i and K ($K \equiv M$ in Ref. 3) are of Type III.

We proceed to derive (2.15) from (1.14) and (1.22) of Ref. 3. First of all, we identify the variable ζ of Ref. 3 with the variables $\mathbf{x}$, t, γ, collectively, and the function $f(\zeta)$ of Eq. (1.14) of Ref. 3 with the component $\Psi(\mathbf{x}, t, \gamma)$ of the column vector $\Psi(x)$. The function $F(0, \epsilon, \mathbf{p}, \lambda)$ of Ref. 3 is identical to the function $f(\epsilon, \mathbf{p}, \lambda)$ of Eq. (2.15). The transformation function $\langle\zeta \mid 0, \epsilon, \mathbf{p}, \lambda\rangle$ to be used in Eq. (1.14) of Ref. 3 is now denoted by $\langle\mathbf{x}, t, \gamma \mid \epsilon, \mathbf{p}, \lambda\rangle$. The function $f(\zeta; \epsilon, \lambda)$ is now denoted by $f(\mathbf{x}, t, \gamma; \epsilon, \lambda)$. Equations (1.22) of Ref. 3 become, on using (I 2.13),

$$\begin{aligned}
\frac{\partial}{\partial x_1}f(\mathbf{x}, t, \gamma; \epsilon, \lambda) &= \frac{\partial}{\partial x_2}f(\mathbf{x}, t, \gamma; \epsilon, \lambda) = 0,\\
\frac{\partial}{\partial x_3}f(\mathbf{x}, t, \gamma; \epsilon, \lambda) &= if(\mathbf{x}, t, \gamma; \epsilon, \lambda),\\
\frac{\partial}{\partial t}f(\mathbf{x}, t, \gamma; \epsilon, \lambda) &= -i\epsilon f(\mathbf{x}, t, \gamma; \epsilon, \lambda).
\end{aligned} \qquad (B1)$$

The general solution of (B1) is

$$f(x, t, \gamma; \epsilon, \lambda) = \exp\left[i(x_3 - \epsilon t)\right]k(\gamma; \epsilon, \lambda), \qquad (B2)$$

where $k(\gamma; \epsilon, \lambda)$ is essentially a constant of integration. We must choose λ to obtain all linearly independent solutions. The simplest choice, which is the one we take, is

$$k(\gamma; \epsilon, \lambda) = \delta_{\gamma\lambda}. \qquad (B3)$$

Thus,

$$f(\mathbf{x}, t, \gamma; \epsilon, \lambda) = \exp[i(x_3 - \epsilon t)]\delta_{\gamma\lambda}. \quad \text{(B4)}$$

From (1.24) of Ref. 3, we have in terms of the notation of the present paper on using (I 2.11), (I 2.12), (I 2.11a), and (I 2.2a)

$$\langle \mathbf{x}, t, \gamma \mid \epsilon, \mathbf{p}, \lambda) = \exp[i(\mathbf{p}\cdot\mathbf{x} - \epsilon p t)]\chi(\gamma \mid \epsilon, \mathbf{p}, \lambda), \quad \text{(B5)}$$

where $\chi(\gamma \mid \epsilon, \mathbf{p}, \lambda)$ is given by (2.13). Then Eq. (1.14) of Ref. 3 becomes Eq. (2.15) of the present paper.

Equations (2.15′) for the generators of the Euclidean group follow directly from (1.23) of Ref. 3, Eq. (B4) of the present paper, and (I 2.14) and (I 2.15).

APPENDIX C: DERIVATION OF THE EXPRESSION FOR THE INFINITESIMAL GAUGE CHANGES

We shall now prove that Eq. (2.21) for the infinitesimal gauge changes $G_i(x)$ is identical to Eq. (2.23).

Let us define the matrix A_i by

$$A_i = \exp(i\boldsymbol{\omega}\cdot\mathbf{M}) \exp(ivN_3)B_i. \quad \text{(C1)}$$

Then the components of the column vector $\mathfrak{H}_i^{(r)}$ are constructed from the matrix elements of the matrix A_iV in accordance with Eq. (2.20). The principal part of our proof consists in the rewriting of the matrix A_i.

On expressing the matrices T_i in terms of N_i and M_i through (2.15a), we obtain, on using exponentiation algorithm (I 3.26) and the commutation rules for the matrices M_i and N_i which are the same as those for the matrices in (I 2.7), the following result:

$$\exp(ivN_3)T_i \exp(-ivN_3) = pT_i. \quad \text{(C2)}$$

Thus, from (2.20′)

$$\exp(ivN_3)B_i = pB_i \exp(ivN_3). \quad \text{(C3)}$$

In a similar way, one can show that

$$\exp(i\boldsymbol{\omega}\cdot\mathbf{M})B_i = \frac{1}{p^2}[-(\mathbf{p}\times\mathbf{M})_i - \epsilon(p_i/p)(\mathbf{p}\cdot\mathbf{N}) + \epsilon pN_i] \exp(i\boldsymbol{\omega}\cdot\mathbf{M}). \quad \text{(C4)}$$

Thus, our expression for the matrix A_i is

$$A_i = (p)^{-1}[-(\mathbf{p}\times\mathbf{M})_i - \epsilon(p_i/p)(\mathbf{p}\cdot\mathbf{N}) + \epsilon pN_i] \exp(i\boldsymbol{\omega}\cdot\mathbf{M}) \exp(ivN_3). \quad \text{(C5)}$$

Equation (2.23) follows on using (2.20), (2.21), (2.17), and (2.22).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Padé Approximants and Bounds to Series of Stieltjes

A. K. COMMON
Department of Applied Mathematics, University of Kent at Canterbury

(Received 20 May 1967)

Series of Stieltjes with nonzero radius of convergence R have been considered in this paper. It is well known that sequences of Padé approximants to these series may be defined which converge in the complex plane cut from $-R$ to $-\infty$. It is shown that the Padé approximants satisfy inequalities between 0 and $-R$ which are much more general than those already proved on the positive real axis. A new sequence of approximants is defined, which are closely related to the Padé approximants and which have very similar properties. The two sets of approximants may be used to determine the series of Stieltjes within certain limits for points on the real axis between 0 and $-R$, given only the first few coefficients of the power series expansion. The result is then extended to all points in the interior of a circle with center at the origin and radius R.

1. INTRODUCTION

The importance of Padé approximants lies in the fact that they can be used to obtain directly an analytical continuation of a function of a complex variable outside the region of convergence of the Taylor series expansion for this function. Indeed, they can be used to interpret series whose radius of convergence is zero.

The known mathematical properties and, in particular, the convergence of sequences of these approximants have been discussed in a review article by Baker, Jr.,[1] (which will be referred to as GB for the rest of this paper). The problem of convergence when the degree of the denominator and numerator both go to infinity has only been solved completely for series of Stieltjes.[2] In particular, it can be shown that if such a series has radius of convergence R, then sequences of Padé approximants in any closed finite region of the complex plane cut along the negative real axis from $-R$ to $-\infty$ converge uniformly to the analytical function defined by the power series.

It is proved in GB that the Padé approximants form both upper and lower bounds to the series of Stieltjes on the positive real axis, and similar relations hold between the corresponding first derivatives. These properties which are presented in Sec. 2 are very important for the following reason: Given a function is a series of Stieltjes and given the first few coefficients of its power series expansion, then one can determine the values of the function on the positive real axis, to within certain limits given by the Padé approximants, without knowing any more about the detailed form of the function.

The main purpose of this paper is to discuss the properties of the Padé approximants on the negative real axis between 0 and $-R$. In Sec. 3 it is shown that the approximants once again give bounds to the series of Stieltjes, but in this case they are all lower bounds. Another important difference is that similar properties may be obtained for derivatives of all orders—not only the first. The fact that Padé approximants only give lower bounds means that they cannot be used by themselves to obtain limits on the exact function. However, in Sec. 4 it is shown that one can define a new sequence of functions which converge to the exact function in the cut plane and which form upper bounds to the exact function on the negative axis between 0 and $-R$, and any one function of the sequence depends on only the first few of the coefficients in the series expansion of the exact function. The new approximants are related to the derivatives of Padé approximants of a new series of Stieltjes obtained from the original one by integration. With this new type of approximant and the Padé approximant one may once again determine (to within certain limits) the values of the exact function on the negative real axis between 0 and $-R$, given only the first few coefficients of the series expansion. Indeed, one can extend this result to all points contained in the circle with center at the origin and of radius R, as is shown in Sec. 4.

To illustrate the above results, a particular example of a series of Stieltjes with nonzero radius of convergence is considered in Sec. 5. This series is related to a spectral integral of a kinematical factor which arises in the approximate solution of the N/D equations in partial wave-dispersion relations.[3] Several tables of numerical results are presented which demonstrate how the above function satisfies the properties discussed in the previous paragraphs. The results of this paper are discussed in Sec. 6.

[1] G. A. Baker, Jr., Advan. Theoret. Phys. **1**, 1 (1965).
[2] T. J. Stieltjes, Ann. Fac. Sci. Univ. Toulouse Sci. Math. Sci. Phys. **8, 9**, 1 (1894).
[3] A. K. Common, J. Math. Phys. **8**, 1669 (1967).

2. SERIES OF STIELTJES AND THEIR PADÉ APPROXIMANTS AS BOUNDS ON THE POSITIVE REAL AXIS

A series of Stieltjes is defined by

$$f(z) = \sum_{j=0}^{\infty} f_j(-z)^j, \tag{2.1}$$

if and only if there is a bounded nondecreasing function $\phi(u)$ taking on infinitely many values in the interval $0 \leq u \leq \infty$ such that

$$f_j = \int_0^{\infty} u^j \, d\phi(u). \tag{2.2}$$

This definition is equivalent to the conditions

$$D(0, n) > 0, \quad D(1, n) > 0, \quad (n = 0, 1, 2, \cdots,), \tag{2.3}$$

where

$$D(m, n) = \det \begin{vmatrix} f_m & f_{m+1} & \cdots & f_{m+n} \\ f_{m+1} & f_{m+2} & \cdots & f_{m+n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ f_{m+n} & f_{m+n+1} & \cdots & f_{m+2n} \end{vmatrix}. \tag{2.4}$$

From (2.3) it can be proved that $D(m, n) > 0$ for all m and n.

A Padé approximant to the series $f(z)$ is the quotient of one polynomial $P(z)$ of degree M by another $Q(z)$ of degree N such that

$$f(z)Q(z) - P(z) = Az^{M+N+1} + Bz^{M+N+2} + \cdots, \tag{2.5}$$

where $A, B, \cdots$, are constants. The polynomials $P(z)$ and $Q(z)$ are determined by Eq. (2.5) so long as the normalizing condition

$$Q(0) = 1 \tag{2.6}$$

is added, and we say that $P(z)/Q(z)$ is the $[N, M]$ Padé approximant to $f(z)$. In fact, $[N, M]$ may be written explicitly in terms of the coefficients f_j as

$$[N, M] = \frac{\begin{vmatrix} f_{M-N+1} & f_{M-N+2} & \cdots\cdots\cdots & f_{M+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ f_M & \cdots\cdots\cdots\cdots\cdots\cdots\cdots & & f_{M+N} \\ \sum_{j=N}^{M} f_{j-N}(-z)^j & \sum_{j=N-1}^{M} f_{j-N+1}(-z)^j & \cdots & \sum_{j=0}^{M} f_j(-z)^j \end{vmatrix}}{\begin{vmatrix} f_{M-N+1} & f_{M-N+2} & \cdots\cdots\cdots & f_{M+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ f_M & f_{M+1} & \cdots\cdots\cdots & f_{M+N} \\ (-z)^N & (-z)^{N-1} & \cdots\cdots\cdots & 1 \end{vmatrix}}, \tag{2.7}$$

where $f_j = 0$ if $j < 0$ and sums for which the initial point is larger than the terminal point are omitted.

Infinite sequences of Padé approximants of the form $[N, N + j]$ where $N \to \infty$ and $j \geq -1$ are considered in GB. Let $P_N^{(j)}(z)$ be the numerator and $Q_N^{(j)}(z)$ be the denominator of $[N, N + j]$ in Eq. (2.7). The inequalities (2.3) determine the position of the zeros of $Q_N^{(j)}(z)$ which lie on the negative real axis between $-R$ and ∞, where R is the radius of convergence of the series (2.1). It is then proved in GB that sequences of Padé approximants of the form $[N, N + j]$ with $j \geq -1$ converge in the cut plane $(-\infty \leq z \leq R)$ to the function $f(z)$ defined by the power series. It is also shown in GB that the Padé approximants satisfy the following inequalities when $z \geq 0$:

$$(-1)^{(1+j)}\{[N + 1, N + 1 + j] - [N, N + j]\} \geq 0, \tag{2.8}$$

$$(-1)^{(1+j)}\{[N, N + j] - [N - 1, N + j + 1]\} \geq 0, \tag{2.9}$$

$$(-1)^{(1+j)}\{[N, N + j + 1] - [N, N + j]\} \geq 0, \tag{2.10}$$

and

$$[N, N] \geq f(z) \geq [N, N - 1]. \tag{2.11}$$

Finally, it is proved that (2.11) may be differentiated with respect to z giving

$$[N, N]' \geq f'(z) \geq [N, N - 1]'. \tag{2.12}$$

Similarly, inequalities (2.9) and (2.10) may be differentiated so long as $j \geq 0$. These results were obtained from the following relations:

$$\frac{P_{N+1}^{(j)}(z)}{Q_{N+1}^{(j)}(z)} - \frac{P_N^{(j)}(z)}{Q_N^{(j)}(z)} = \frac{(-z)^{2N+1+j}[D(1+j,N)]^2}{Q_N^{(j)}(z)Q_{N-1}^{(j+2)}(z)}, \tag{2.13}$$

$$\frac{P_N^{(j)}(z)}{Q_N^{(j)}(z)} - \frac{P_{N-1}^{(j+2)}(z)}{Q_{N-1}^{(j+2)}(z)} = \frac{(-z)^{2N+1+j}[D(1+j,N-1)D(3+j,N-1) - D(1+j,N)D(3+j,N-2)]}{Q_N^{(j)}(z)Q_{N-1}^{(j+2)}(z)}, \tag{2.14}$$

and a special case (when $j = -1$) of the relation

$$\frac{P_N^{(j+1)}(z)}{Q_N^{(j+1)}(z)} - \frac{P_N^{(j)}(z)}{Q_N^{(j)}(z)} = \frac{(-z)^{2N+1+j}D(j+2,N-1)D(j+1,N)}{Q_N^{(j+1)}(z)Q_N^{(j)}(z)}. \tag{2.15}$$

3. PADÉ APPROXIMANTS AS LOWER BOUNDS OF $f(z)$ ON THE NEGATIVE REAL AXIS

It is proved in GB that the zeros of $Q_N^{(j)}(z)$ lie between $-R$ and $-\infty$, thus $Q_N^{(j)}(z)$ is positive for $-R \leq z \leq 0$, since $Q_N^{(j)}(0) = D(1+j, N-1) > 0$. Hence, from Eqs. (2.13), (2.14), and (2.15), for $j \geq -1$ and $-R < z \leq 0$,

$$[N+1, N+1+j] - [N, N+j] \geq 0, \tag{3.1}$$

$$[N, N+j] - [N-1, N+j+1] \geq 0, \tag{3.2}$$

$$[N, N+j+1] - [N, N+j] \geq 0. \tag{3.3}$$

The inequality (3.2) follows from the fact that

$$[D(1+j,N-1)D(3+j,N-1) - D(1+j,N) \times D(3+j,N-2)] > 0,$$

as is proved in GB. The sequence $[N, N+j]$ is monotonic increasing from (3.1), and since it converges to $f(z)$,

$$f(z) \geq [N, N+j] \tag{3.4}$$

for $-R < z \leq 0$. From Eqs. (3.1) and (3.2), the best lower bound to the function $f(z)$ obtainable from a Padé approximant with a given even number of coefficients is $[N, N]$; with a given odd number of coefficients it is $[N, N-1]$. Using Eq. (3.3), it may be proved that these "best" approximants satisfy the following inequalities:

$$[1,1] \leq [2,1] \leq [2,2] \leq \cdots \leq [N, N-1] \leq [N,N] \leq [N+1, N] \leq f(z). \tag{3.5}$$

Now in the case where $z \geq 0$, the inequalities (2.8)–(2.11) could be differentiated at most once; i.e., similar relations cannot be found for second or higher derivatives of the Padé approximants. However, when $-R < z \leq 0$, the situation is quite different. It is now proved that the inequalities (3.1)–(3.4) may be differentiated any number of times, so long as a factor (-1) is introduced with each differentiation, so that

$$\{[N+1, N+j+1]^{(n)} - [N, N+j]^{(n)}\}(-1)^n \geq 0, \tag{3.6}$$

$$\{[N, N+j]^{(n)} - [N-1, N+j+1]^{(n)}\}(-1)^n \geq 0, \tag{3.7}$$

and

$$\{[N, N+j+1]^{(n)} - [N, N+j]^{(n)}\}(-1)^n \geq 0. \tag{3.8}$$

The proof of (3.8) is given below; Eqs. (3.6) and (3.7) may be proved in a similar way. Since the zeros of $Q_N^{(j)}(z)$ lie between $-R$ and $-\infty$, the function

$$F(z) = \frac{D(j+2,N-1)D(j+1,N)}{Q_N^{(j+1)}(z)Q_N^{(j)}(z)}$$

may be written in the form

$$F(z) = C/[(z+\alpha_1)(z+\alpha_2)\cdots(z+\alpha_{2N})], \tag{3.9}$$

where C is a positive constant and $-\alpha_i$ are the zeros of $Q_N^{(j+1)}(z)$ and $Q_N^{(j)}(z)$.

Since $z + \alpha_i > 0$ for $-R < z \leq 0$,

$$(-1)^n \frac{d^n F(z)}{dz^n} > 0 \tag{3.10}$$

for these same values of z and any integer n. Using Leibniz's theorem, the nth derivative of the right-hand side of (2.15) is

$$\begin{aligned} &= (-1)^{j+1}[F^{(n)}(z)z^{2N+j+1} \\ &\quad + nF^{(n-1)}(z)z^{2N+j}(2N+j+1) + \cdots \\ &\quad + {}^nC_rF^{(n-r)}(z) \times (2N+j+1) \times \cdots \\ &\quad \times (2N+j-r+2)z^{2N+j-r+1} + \cdots], \\ &= (-1)^{[j+1+(n+2N+j+1)]} \times [\text{positive quantity}] \\ &= (-1)^n \times [\text{positive quantity}]. \end{aligned} \tag{3.11}$$

Inequality (3.8) then follows immediately from Eq. (2.15). Since $[N, N+j]$ tends uniformly to $f(z)$ in any finite closed region of the cut z plane, $[N, N+j]^{(n)}$ tends uniformly to $f^{(n)}(z)$ in this same region.

Therefore, using (3.7),

$$\{f^{(n)}(z) - [N, N+j]^{(n)}\}(-1)^n \geq 0. \tag{3.12}$$

This last set of inequalities tells us that the magnitude of the error between $f(z)$ and the $[N, N+j]$ Padé approximant increases as z decreases from 0 to $-R$.

The same is true for the corresponding derivatives of any order.

4. NEW SEQUENCE OF APPROXIMANTS TO A SERIES OF STIELTJES

It has been shown in Sec. 2 that Padé approximants give both upper and lower bounds to a series of Stieltjes $f(z)$ when $z \geq 0$. Thus, given the first few coefficients of a power series expansion, $f(z)$ can be determined to within certain limits provided by these bounds without knowing any more about its detailed form. When $0 \geq z > -R$, the Padé approximants only give lower bounds; $f(z)$ cannot be determined to within limits in the above way. It is now shown that a new sequence of functions may be defined which converge to $f(z)$, are upper bounds for $0 \geq z > -R$, and where any one function of the sequence depends only on the first few coefficients of a power series expansion of $f(z)$. Now

$$f(z) = \int_0^{1/R} \frac{1}{1+uz}\, d\phi(u) = f_0 - z\int_0^{1/R} \frac{u\, d\phi(u)}{1+uz}. \tag{4.1}$$

Let

$$F(z) = \int_0^{1/R} \frac{u\, d\phi(u)}{1+uz} \tag{4.2}$$

and let $G(z)$ be the integral of $F(z)$ so that

$$G(z) = \int_0^z F(t)\, dt. \tag{4.3}$$

Then

$$\begin{aligned} G(z) &= \int_0^{1/R} \log(1+uz)\, d\phi(u), \\ &= f_0 \log_e(1+z/R) - z\int_0^{1/R} \frac{\phi(u)\, du}{1+uz} \end{aligned} \tag{4.4}$$

after changing the order of integration[4] and then integrating by parts.

Without loss of generality we may take $\phi(0) = 0$. Then, if the function $\gamma(u)$ is defined by the relations

$$\gamma(u) = \begin{cases} \displaystyle\int_0^u \phi(t)\, dt & 0 \leq u \leq \frac{1}{R}, \\ \gamma\left(\dfrac{1}{R}\right) & \dfrac{1}{R} \leq u < \infty, \end{cases} \tag{4.5}$$

it is easily seen to be nondecreasing, bounded, and to assume infinitely many values. Therefore the function $K(z)$ defined by the relation

$$K(z) = \int_0^{1/R} \frac{\phi(u)\, du}{1+uz} = \int_0^{\infty} \frac{d\gamma(u)}{1+uz}$$

is a series of Stieltjes with radius of convergence R.

[4] The interchange of the order of integration is allowable for all measure functions $\phi(u)$.

Substituting in (4.4),

$$G(z) = f_0 \log_e(1 + z/R) - zK(z). \tag{4.6}$$

Then differentiating this equation with respect to z,

$$F(z) = f_0/(R+z) - [zK'(z) + K(z)]. \tag{4.7}$$

Finally,

$$f(z) = \frac{f_0 R}{R+z} + z[zK'(z) + K(z)]. \tag{4.8}$$

An approximate expression for $f(z)$ can be obtained by replacing $K(z)$ with its $[N, M]$ Padé approximant. Let (N, M) be the corresponding approximate form for $f(z)$. Then it is obvious from (4.8) that $(N, N+j)$ converges uniformly to $f(z)$ as $N \to \infty$ in any finite closed region of the z-plane cut from $-R$ to $-\infty$, for $j \geq -1$.

The inequalities analogous to Eqs. (3.6)–(3.8) and (3.12) are easily obtained. They are

$$[(N+1, N+j+1)^{(n)} - (N, N+j)^{(n)}](-1)^{n+1} \geq 0, \tag{4.9}$$

$$[(N, N+j)^{(n)} - (N-1, N+j+1)^{(n)}](-1)^{n+1} \geq 0, \tag{4.10}$$

$$[(N, N+j+1)^{(n)} - (N, N+j)^{(n)}](-1)^{n+1} \geq 0, \tag{4.11}$$

$$[f^{(n)}(z) - (N, N+j)^{(n)}](-1)^{n+1} \geq 0. \tag{4.12}$$

The last set of inequalities tell us that the $(N, N+j)$ are upper bounds to $f(z)$ for $-R < z \leq 0$, and that the magnitude of the error between $f^{(n)}(z)$ and $(N, N+j)^{(n)}$ increases as z decreases from 0 to $-R$. To find a particular $(N, N+j)$ one needs to know the corresponding Padé approximant for $K(z)$. This is determined from the coefficients of its power series expansion. Let

$$K(z) = \sum_{i=0}^{\infty} k_i(-z)^i. \tag{4.13}$$

The k_i are determined by substituting for $f(z)$ and $K(z)$ in Eq. (4.8) their power series expansions, and then equating coefficients. This gives

$$k_i = \frac{1}{(i+1)}\{f_0/R^{i+1} - f_{i+1}\}. \tag{4.14}$$

The Padé approximants of $K(z)$, and hence the approximants (N, M), can therefore be obtained from the first few coefficients of the power series expansion of $f(z)$, as was required. It also follows from the inequalities (4.9)–(4.12) that, for a given odd number of coefficients used in the Padé approximant to $K(z)$,

$(N, N-1)$ is the best approximant to $f(z)$; for a given even number of coefficients, the best approximant is (N, N). From Eqs. (4.11) and (4.12),

$$(N, N-1) \geq (N, N) \geq (N+1, N) \geq f(z), \quad (4.15)$$

so that the "best approximants" get closer to $f(z)$ as the number of coefficients used is increased.

Consider now the function $E(z)$ which is the error in replacing $f(z)$ by its $[N, M]$ Padé approximant, i.e.,

$$E(z) = [N, M] - f(z). \quad (4.16)$$

From (3.12), $(-1)^{n+1}E^{(n)}(z) \geq 0$ for $-R < z \leq 0$. Take any circle with center at the origin and radius $r < R$. Then since the Taylor series expansion of $E(z)$ converges when $|z| \leq r$, for these values of z

$$|E(z)| \leq |E(-r)| \quad (4.17)$$

by the maximum modulus theorem. But since $[N, M]$ and (N, M) form lower and upper bounds to $f(z)$ for $-R < z \leq 0$,

$$|E(-r)| \leq |\{[N, M] - (N, M)\}_{z=-r}|. \quad (4.18)$$

Combining (4.17) with (4.18), it follows that an upper bound on the modulus of the error in replacing $f(z)$ by $[N, M]$ may be obtained from the first few coefficients in the power series expansion of $f(z)$ when z lies inside or on a circle of radius $r < R$. Since the magnitude of $[N, M] - (N, M)$ increases as z decreases from 0 to $-R$, the best estimate of the error at a point $z = w$ (where w is complex and inside the circle $|z| = R$) is

$$= |\{[N, M] - (N, M)\}_{z=-|w|}|. \quad (4.19)$$

Similar bounds can be put on the error in replacing $f(z)$ by (N, M).

5. NUMERICAL RESULTS FOR A SERIES OF STIELTJES OCCURRING IN PARTIAL WAVE DISPERSION RELATIONS

Consider the scattering of scalar particles with equal mass μ and at a total energy squared equal to s. The corresponding partial wave amplitudes $A_l(s)$ are then analytic functions in the complex s plane with cuts along the positive real axis from $4\mu^2$ to ∞, and from some value $-s_1$ to $-\infty$ along the negative real axis. The usual problem is to solve for $A_l(s)$, given its discontinuity across the negative axis and using the unitary condition on the right-hand cut. The N/D method is probably the easiest way of doing this, but the resultant integral equations in N or D, in general, can only be solved approximately. One way is to approximate $F(s)$, the spectral integral of the kinematical factor $\rho(s)$, over the right-hand cut so that

$$F(s) = \frac{1}{\Pi} \int_{4\mu^2}^{\infty} \frac{\rho(x)}{(x-s)x}\, dx \quad (5.1)$$

by a rational function of s.[5] This particular case [when this rational function is a Padé approximant of $F(s)$] has been discussed by the author.[3]

The above function is not a series of Stieltjes as it stands, but, upon making the substitutions

$$s = \frac{4\mu^2 z}{(1+z)} \quad \text{and} \quad x = \frac{4\mu^2}{(1-u)}, \quad (5.2)$$

it takes on the new form

$$F(s) = \frac{(1+z)}{4\Pi\mu^2} \int_0^1 \frac{\rho[4\mu^2/(1-u)]}{(1+zu)}\, du,$$
$$= (1+z)f(z), \quad (5.3)$$

where this last equation defines the function $f(z)$. Then, since $\rho(x) > 0$, $f(z)$ is a series of Stieltjes in z and, from (5.3), it has a radius of convergence 1.

The particular case when

$$\rho(x) = \left[\frac{x - 4\mu^2}{x}\right]^{\frac{1}{2}} = u^{\frac{1}{2}}$$

is considered here. In this case the integral in (5.3) may be evaluated giving

$$f(z) = \frac{1}{2\Pi\mu^2 z}\left\{1 - \frac{1}{2(-z)^{\frac{1}{2}}} \log\left[\frac{1+(-z)^{\frac{1}{2}}}{1-(-z)^{\frac{1}{2}}}\right]\right\}. \quad (5.4)$$

The coefficients f_i in the power series expansion (2.1) for $f(z)$ are then given by

$$f_i = [2\Pi\mu^2(2i+3)]^{-1}. \quad (5.5)$$

Values of $f(z)$ and a selection of its approximants are given in[6] Table I for z between 0 and -1. It can be verified from the table that for these values of z,

$$[2, 1] \leq [2, 2] \leq [3, 2] \leq f(z)$$
$$\leq (3, 2) \leq (2, 2) \leq (2, 1),$$

which are consistent with Eqs. (3.5) and (4.15). The data given in Table I may then be used to prove that, for the above values of z,

$$[2, 1]' \geq [2, 2]' \geq [3, 2]' \geq f'(z)$$
$$\geq (3, 2)' \geq (2, 2)' \geq (2, 1)',$$

which are again in agreement with Eqs. (3.6), (3.12), (4.9), and (4.12). One may also notice that for values of z close to -1, the Padé approximants $[N, M]$ are much better approximations to $f(z)$ than the approximants (N, M). This result however may well depend

[5] H. Pagels, Phys. Rev. **140**, B1599 (1965).

[6] All numerical values in this paper have been calculated with $\mu = 1$.

Table I. $f(z)$ and its approximants. (Values are multiplied by 10^3.)

z	[2, 1]	[2, 2]	[3, 2]	$f(z)$	(3, 2)	(2, 2)	(2, 1)
−0.1	5.6481	5.6481	5.6481	5.6481	5.6481	5.6481	5.6481
−0.2	6.0496	6.0497	6.0497	6.0497	6.0497	6.0497	6.0497
−0.3	6.5277	6.5282	6.5282	6.5282	6.5283	6.5283	6.5285
−0.4	7.1095	7.1113	7.1116	7.1117	7.1118	7.1121	7.1134
−0.5	7.8366	7.8430	7.8445	7.8448	7.8452	7.8474	7.8537
−0.6	8.7776	8.7975	8.8036	8.8051	8.8081	8.8208	8.8485
−0.7	10.054	10.112	10.136	10.144	10.168	10.239	10.356
−0.8	11.901	12.076	12.168	12.216	12.448	12.865	13.387
−0.9	14.854	15.418	15.829	16.213	19.770	22.921	25.721

on the particular function $f(z)$ considered in this section and may not be a general property.

In Table II the values of $|[2, 2]|$, $|(2, 2)|$, and the corresponding errors are given for a selection of points in the complex z plane inside the circle $|z| = 1$. Only points with $\operatorname{Im} z > 0$ have been considered, but as $f(z)$ and its approximants are real functions of z, their values at points with $\operatorname{Im} z < 0$ are obtained by complex conjugation.

It is easily seen from Table II that the values of $|f(z) - [2, 2]|$ satisfy (4.17), i.e.,

$$|f(z) - [2, 2]| \leq |f(-r) - [2, 2]|_{z=-r},$$

where $|z| \leq r$. The same is also true for $|f(z) - (2, 2)|$.

As was shown in Sec. 4, one may put an upper bound on the error between $f(z)$ and the $[N, M]$ or (N, M) approximants without knowing any more about the form of $f(z)$ than the first $N + M + 2$ coefficients in its power series expansion which gives these approximants. For instance, when $|z| \leq r < 1$ and the first six coefficients of the power series expansion of $f(z)$ are given,[7]

$$|f(z) - [2, 2]| \leq |\{[2, 2] - (2, 2)\}_{z=-r}|,$$

and the right-hand side of this inequality can be calculated.

For instance, when $|z| \leq 0.6$,

$$|f(z) - [2, 2]| < 2.4 \times 10^{-4}.$$

Since $|[2, 2]| \sim 10^{-2}$ for $|z| = 0.6$, this is a useful estimate of the error involved in replacing $f(z)$ by its [2, 2] approximant. However, when $|z| \leq 0.9$, the estimate of the error is

$$|f(z) - [2, 2]| < 7.5 \times 10^{-2},$$

which is of the same order as $|[2, 2]|$ and is therefore not so useful. To have useful approximations to $f(z)$ for these values of z, one must use higher-order approximants.

[7] Given the first six coefficients of the power series, one can calculate [3, 2] as well as [2, 2]. Better estimates for $f(z)$ will be obtained by replacing [2, 2] with [3, 2] in the following work.

The particular series of Stieltjes considered in this section has unit radius of convergence, and so the results proved in this paper may be applied in the region $|z| < 1$, which is small compared with the whole complex z plane. However, this region of the z plane maps onto the much larger region of the complex s plane defined by $\operatorname{Re} s < 2\mu^2$. The interior of the circle $|z| = r < 1$ maps onto the interior of the circle with center at $s = -4\mu^2 r^2/(1 - r^2)$ and with radius $4\mu^2 r/(1 - r^2)$ in the complex s plane. The error between $f(z)$ and (N, M) or $[N, M]$ for values of s inside this circle is less than the error at the point $s = 4\mu^2 r/(1 - r^2)$ on the s circumference of the circle which corresponds to $z = -r$. This point on the circle is the one farthest away from the right-hand cut in the complex s plane.

6. CONCLUSIONS

Series of Stieltjes with nonzero radius of convergence have been considered in this paper. If such a series has radius of convergence R, then sequences of Padé approximants may be defined which converge uniformly in any finite closed region of the complex plane cut from $-R$ to $-\infty$. Properties of the Padé approximants on the positive real axis have been discussed in GB. An important result is that they provide both upper and lower bounds to the exact function in this region. This means that values of the series of Stieltjes may be determined within certain limits on the positive real axis, given the first few coefficients of the power series expansion.

This paper has been concerned with the behavior of the Padé approximants on the real axis between 0 and $-R$. The investigation of such behavior has led to the definition of a new type of approximant to the series of Stieltjes and the extension of the region of interest to the interior of a circle with center at the origin and of radius R.

Various inequalities between the series of Stieltjes $f(z)$ and its Padé approximants $[N, M]$ were obtained in Sec. 3 for z on the real axis between 0 and $-R$.

Table II. Values of [2, 2] and (2, 2) for complex z with the corresponding errors.

$\|z\|$	Arg z (radians)	$\|[2, 2]\|$	$\|f(z) - [2, 2]\|$	$\|(2, 2)\|$	$\|f(z) - (2, 2)\|$
0.3	0	4.5165 E-2	7.9401 E-8	4.5165 E-2	2.6892 E-7
0.3	0.2Π	4.6406 E-2	1.0479 E-7	4.6406 E-2	2.0406 E-7
0.3	0.4Π	4.9440 E-2	1.2803 E-7	4.9440 E-2	3.3491 E-7
0.3	0.6Π	5.4916 E-2	2.3520 E-7	5.4916 E-2	3.1326 E-7
0.3	0.8Π	6.1649 E-2	4.2468 E-7	6.1649 E-2	4.2524 E-7
0.3	Π	6.5282 E-2	5.2576 E-7	6.5283 E-2	9.1317 E-7
0.6	0	3.9567 E-2	1.3544 E-6	3.9567 E-2	1.1115 E-6
0.6	0.2Π	4.0893 E-2	1.6162 E-6	4.0893 E-2	9.4739 E-7
0.6	0.4Π	4.5289 E-2	2.7950 E-6	4.5289 E-2	2.4674 E-6
0.6	0.6Π	5.4241 E-2	7.2611 E-6	5.4236 E-2	6.9472 E-6
0.6	0.8Π	7.0880 E-2	2.7836 E-5	7.0904 E-2	4.1109 E-5
0.6	Π	8.7975 E-2	7.5778 E-5	8.8208 E-2	1.5764 E-4
0.9	0	3.5350 E-2	5.6592 E-6	3.5350 E-2	5.2075 E-6
0.9	0.2Π	3.6688 E-2	6.9494 E-6	3.6688 E-2	6.2891 E-6
0.9	0.4Π	4.1297 E-2	1.3276 E-5	4.1305 E-2	1.4117 E-5
0.9	0.6Π	5.1530 E-2	4.3881 E-5	5.1464 E-2	5.9592 E-5
0.9	0.8Π	7.5732 E-2	3.31194 E-4	7.6547 E-2	7.7226 E-4
0.9	Π	15.418 E-2	7.9492 E-3	22.921 E-2	6.7078 E-2

They are similar to corresponding inequalities holding on the positive real axis, but are much more general since they may be differentiated any number of times, while the latter inequalities may be differentiated at most once.

However, there is a disadvantage because the Padé approximants give only lower bounds to the series of Stieltjes $f(z)$ on the negative real axis, so that the sum of the series can no longer be determined (within certain limits) from the first few coefficients of the power series expansion. To remedy this defect, a new type of approximant (N, M) to $f(z)$ was defined in Sec. 4, which has properties quite similar to those of the Padé approximants. For example, it may be obtained from the first few coefficients of the power series expansion for $f(z)$, sequences of the form $(N, N + j)$ converging to $f(z)$ as $N \to \infty$ for $j \geq -1$, and all z in the complex plane cut from $-R$ to $-\infty$.

The properties of this new set of approximants were discussed in Sec. 4, and, in particular, their behavior was examined on the real axis between 0 and $-R$ for which various inequalities were obtained similar to those holding for the Padé approximants. However, they give upper bounds to $f(z)$ on the above part of the real axis and so can be used with the Padé approximants to determine $f(z)$, within certain limits, from the first few coefficients of the power series expansion. This result may be extended off the real axis to the interior of a circle with center at the origin and radius R, using the maximum modulus theorem, as was shown in Sec. 4. In Sec. 5 a series of Stieltjes, which is of some interest in the solution of partial-wave dispersion relations, was considered. Numerical results were obtained which amply demonstrate the results of the previous sections.

The new approximants (N, M) have only been defined for a series of Stieltjes. However, one can define them for any power series with nonzero radius of convergence by using (4.8), (4.13), and (4.14). Presumably they will have properties quite similar to the Padé approximants of the same power series.

ACKNOWLEDGMENTS

The author would like to thank Professor J. S. R. Chisholm for his continued interest in this work and for several useful discussions which have led to improvements in it. Thanks are also due to Dr. A. F. Beardon for comments on several mathematical points arising in the paper. Finally, the author is indebted to Dr. B. Spratt for the use of the facilities of the computer laboratory of the University of Kent at Canterbury.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Invariants and Scalars of Compact Inhomogeneous Unitary Algebras*

RONALD MIRMAN
Physics Department, The Brooklyn Center, Long Island University, Brooklyn, New York

(Received 14 February 1967)

The independent invariants of the fundamental and adjoint inhomogeneous algebras are explicitly constructed. Some classification of inhomogeneous algebras is given and their scalars and invariants are discussed.

1. INTRODUCTION

The first step in computing the representations of Lie algebras is usually that of finding the invariants which label the representations. Beyond this use, studying the invariants may also be helpful in understanding the structure of the algebra itself. Here we wish to discuss some general considerations concerning the polynomial invariants and scalars of inhomogeneous unitary algebras with compact homogeneous parts (compact inhomogeneous unitary algebras).

By an inhomogeneous algebra we mean the direct sum of a semisimple algebra (restricted here to be a unitary algebra, that is A_l), which we refer to as the homogeneous part, with an Abelian invariant algebra, the inhomogeneous part. As shown by Rosen[1] the inhomogeneous part transforms according to some representation of the homogeneous part. The generators are labeled E and p for homogeneous and inhomogeneous parts, respectively, and sometimes T is used for a generator which could be either E or p.

By a scalar we mean a polynomial in the generators which commutes with all E's, and by an invariant a scalar which commutes with all p's.

We shall be considering the algebra A_l and the group $SU(n)$, so that l and n shall have this fixed meaning (except when used as indices), and $n = l + 1$. $ISU(n)$ and $IU(n)$ shall refer to the inhomogeneous unimodular unitary and unitary groups, respectively, with unspecified inhomogeneous parts.

There are two special algebras, which will be given names: the adjoint (inhomogeneous unitary) algebra, and the fundamental (inhomogeneous unitary) algebra. The terms in parentheses may be left out. These are the algebras with an inhomogeneous part transforming according to the adjoint and fundamental representations of the simple part, respectively.

Tensor notation is used, with upper and lower indices. The commutation relations in this notation are reviewed for convenience in the Appendix. The indices will be arranged in Young patterns (except if indicated) and we will refer to mixed form for a generator or polynomial which has both upper and lower indices, and (completely) lower form if it is written with only lower indices. The summation convention will be used for latin letters appearing once as an upper index and once as a lower; Greek indices are not summed and are to be understood as representing any arbitrary index in the set.

If an index appears on a p and on an E, for example, we shall say that the p and E are linked or connected; thus $p_i \cdots E^i$, or we might say that the E is linked to a lower pattern. We might also have p connected to another p through an E: $p_i \cdots E^i_j \cdots p^j$.

In Sec. 2 we consider how to construct scalars for the case of a simple homogeneous part and an inhomogeneous part which transforms according to a single irreducible representation of the simple part (an irreducible simple inhomogeneous algebra). A few general remarks are also made about invariants. In Sec. 3 we find the independent invariants for the fundamental algebra, and in Sec. 4 we do the same for the adjoint algebra. Section 5 gives a brief discussion of the algebra whose inhomogeneous part is reducible and Sec. 6 describes the algebras which result when the homogeneous part is semisimple but not simple. Some remarks are made about scalars in this type of algebra.

2. IRREDUCIBLE SIMPLE INHOMOGENEOUS ALGEBRAS

A discussion of how scalars are formed is useful not only because it is a first step toward finding the invariants, which must be scalars, but also because it sheds light on the structure of the algebra. In this section we shall be concerned with the irreducible simple inhomogeneous algebra. Other cases are considered later.

First we wish to show that for the algebras considered here and in the notation in which an expression contains an equal number of upper and lower indices, an expression is a scalar if all the indices are saturated, that is, each appears once as an upper index and once as a lower, and is summed over.

* Work supported by a grant from Long Island University.
[1] J. Rosen, Nuovo Cimento **45A**, 234 (1966).

The scalars are to be found by taking all possible products of any number of the E's with any of the mixed forms of the p's and distributing the indices in all possible ways. Not all of these expressions will be independent of course. For example, for the adjoint representation of A_2 the different forms of the p's are

$$p^i_j,\ p^{\ m}_{r\ s},\ p_{ri},\ p^{si}_{\ \ s}.$$

In order that an expression be a scalar it will be required to commute with all the E^α_α separately, thereby not distinguishing explicitly between $U(n)$ and $SU(n)$. This can be done because by requiring an equal number of upper and lower indices the value of the commutator of E^i_i with some expression is equal to zero since it is proportional to the difference between the number of upper and lower indices. Therefore the independent generators H_i of $SU(n)$ which are each a sum of the E^α_α are completely determined by and completely determine the commutator of E^α_α. So either set can be used and it is simpler to use E^α_α. This requirement cannot always be imposed for other algebras as we shall show below.

We first note that saturated indices contribute nothing to the commutator; if all the indices were saturated, clearly the term would be a scalar:

$$[E^\alpha_\beta,\ T^{\cdots m\cdots}_{\cdots m\cdots}] = T^{\cdots\alpha\cdots}_{\cdots m\cdots}\delta^m_\beta - T^{\cdots m\cdots}_{\cdots\beta\cdots}\delta^\alpha_m + \cdots, \quad (2.1)$$

and these two terms cancel.

There is another possible way of forming invariants; by the use of the term $\epsilon_{abc\ldots}$, the completely antisymmetric symbol. Assuming that a complete set of polynomial invariants exist, then invariants with the ϵ are not independent of those that are formed without it. Let us consider $IU(n)$ in which case this symbol is not defined. Then, under the assumption, any representation is completely labeled by the invariants that do not include the ϵ. Now with two inequivalent, irreducible representations of $ISU(n)$ we can associate two inequivalent, irreducible representations of $IU(n)$; since the latter are completely labeled, the former are also.

That two inequivalent representations of $ISU(n)$ are inequivalent for $IU(n)$ is clear: the only difference between the latter and the former algebras is that the latter contains one more generator. Thus if there is no matrix which will transform all the representatives of $ISU(n)$ in one representation into the representatives in another representation, adding one more generator will not produce such a matrix.

An irreducible representation of $ISU(n)$ is an irreducible representation of $IU(n)$, with one exception, as we shall show below. So two inequivalent, irreducible representations of $ISU(n)$ give two inequivalent, irreducible representations of $IU(n)$. The converse need not be true. The exception occurs when E^i_i links two equivalent representations of $ISU(n)$. But this does not change the fact two inequivalent, irreducible representations of $ISU(n)$ are correlated with two inequivalent, irreducible representations of $IU(n)$, and so are completely labeled.

We must also consider the possibility that there are more independent invariants which can be formed for $IU(n)$ than can for $ISU(n)$. This may occur because there are different ways of writing the p's and because of the occurence of the operator E^i_i. However, as we shall show in Sec. 6, E^i_i can occur in an invariant only if the p's are chosen with equal number of upper and lower indices, and we can always chose p's not of this form. Furthermore, if we form two invariants using different forms of the p's, then their value is determined by the matrices of $ISU(n)$, as E^i_i does not appear in the invariant. Therefore the form of p is irrelevant in determining its value. So if an invariant has different values for different representations of $IU(n)$, it will have different values for the representations of $IUS(n)$ which are correlated with the former representations.

Accordingly we shall not consider the ϵ symbol further.

The proof that an irreducible representation of $IU(n)$ is an irreducible representation of $ISU(n)$ is the following. For $IU(n)$ there is one extra operator E^i_i which commutes with all the semisimple operators and whose commutator with any of the p's is proportional to that p. Suppose now that A and B are two different representatives of any generator M of $ISU(n)$ belonging to two different inequivalent, irreducible representations of $ISU(n)$, and suppose that E^i_i connects them. Letting E^i_i have the form $\begin{pmatrix} y & u \\ v & z \end{pmatrix}$ and M the form $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, then the commutator of E^i_i and M gives

$$\begin{pmatrix} [y, A] & uB - Au \\ vA - Bv & [z, B] \end{pmatrix} \approx \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad (2.2)$$

which means that

$$\begin{aligned} vA - Bv &= 0, \\ uB - Au &= 0, \end{aligned} \quad (2.3)$$

requiring that $u = v = 0$, since Schur's lemma goes there without change here.

The weakness of this proof (aside from the assumption of complete reducibility) is that many of the representations of inhomogeneous groups must be realized on spaces larger than Hilbert (for example,

on rigged Hilbert spaces) and the representatives may not be functions but (for example) functionals. We know of no general proof that the usual matrix manipulations can be performed for all the spaces on which a representation may be realized. We shall explicitly exclude from our present considerations representations for which it has not been proven that we can perform the "usual manipulations"; we leave open the questions as to how large a class of representations, if any, is thus excluded.

We can generally consider the invariants completely symmetrical with respect to the position of all the generators. To go from a symmetric to a nonsymmetric term requires commutation of the generators, which introduces expressions of lower order. (By "order" we mean the total number of generators multiplied together.) When we take the commutator of an expression and a p, the order of the expression is not changed. Let us now assume that the symmetrized version commutes with the p's, but the unsymmetrized one does not (or vice versa). Then the commutator of the symmetrized term equals the commutator of the unsymmetrized term plus the commutator of the difference, and this sum equals zero. However, the first term in the sum is of higher order than the other terms, and so it cannot be canceled by them. Hence we arrive at a contradiction. Therefore, if an expression is, or is not, an invariant, it remains so under any rearrangement of the generators in it. Notice that an invariant obtained by rearrangement from another invariant is not independent because the difference is an invariant of lower order (which is taken as independent).

Although the ordering of the generators in a scalar does not determine whether it is an invariant, nevertheless it is possible that a symmetrical expression (being the sum of several terms) might be zero, even though the individual invariants in the sum are not.

Finally, we mention the obvious point that any scalar consisting only of p's is an invariant.

3. FUNDAMENTAL INHOMOGENEOUS ALGEBRAS

The fundamental inhomogeneous algebra has one independent invariant.[2] Before exhibiting it, we shall make some remarks on the construction of scalars.

First we note that there are no scalars consisting only of p's, for such a scalar would have the unique form

$$\begin{matrix} i \\ j \\ \vdots \\ \vdots \end{matrix}$$
$$p_i p_j \cdots p_m p^m,$$

but this is zero because of the antisymmetry of the upper indices and the commutivity of the p's.

Furthermore, if there is a p connected to the upper pattern,

$$\begin{matrix} a \\ \vdots \\ \vdots \end{matrix}$$
$$p_j p^j,$$

the term is zero. In order to show this we shall take a particular component, without loss of generality, giving for the above term:

$$\begin{matrix} 3 & & 3 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ n & & n \end{matrix}$$
$$p^1 p_1 + p^2 p_2.$$

The first p equals $-p_2$ while the third is $-p_1$. The sum is thus zero.

A scalar containing E's must contain only one p having an upper pattern because (as is easily seen) it is impossible to connect the upper patterns together as the lower p's contain only a single index. Therefore a scalar with more than one upper pattern is just a product of scalars. Furthermore, if there is a p whose index appears the second time in the upper pattern, the scalar is zero by the above argument.

We shall now exhibit the invariant explicitly and prove that it is indeed an invariant. For A_1 it is $E_i^i p_i p^j$, and for any arbitrary A it is

$$\begin{matrix} a \\ b \\ \vdots \\ \vdots \end{matrix}$$
$$p^j E_f^k p_k [E_a^m E_b^n (E_n^j p_m - E_m^j p_n) p_j] \cdots, \qquad (3.1)$$

where there is a term similar to the one in the square brackets for each pair of indices in the upper pattern, and the $E_f^k p_k$ appears if and only if the number of upper indices which equals (l) is odd.

By taking a particular set of components of the E's and p's, it can be immediately seen that the expression is not identically zero.

To show that it is an invariant we commute it with p_α. This results in a sum of terms with each E replaced in turn by a p whose index is the lower index of the E times a Kronecker delta. The contribution to the sum due to the E_f^k and the E's in the first parenthesis is zero because they will give terms with a p connected to the upper pattern. The second parenthesis will give

$$(p_n p_m - p_m p_n) p_\alpha, \qquad (3.2)$$

so that the entire sum, and therefore the commutator, is zero. This proves that the expression is an invariant.

The requirement that the E's be traceless was not used in the above proof, and we now wish to show that

[2] R. Mirman, J. Math. Phys. 9, 47 (1968) (following paper).

the expression is not changed if this requirement is introduced; it remains an invariant when the E's are explicitly traceless. Substituting

$$E_j^i = E_j'^i + (1/l+1)\sum_k E_k^k \delta_j^i,$$

where E' is traceless, and denoting the second term by $E\delta_j^i$, we get

$$\begin{matrix} a \\ b \\ \vdots \\ \vdots \end{matrix}$$

$$p^j(E_f'^k p_k + Ep_k\delta_f^k) \times [(E_a'^m E_b'^n + EE_b'^n\delta_a^m + E_a'^m E\delta_b^n + E^2\delta_a^m\delta_b^n) \times (E_n'^i p_m - E_m'^i p_n + Ep_m\delta_n^i - Ep_n\delta_m^i)p_j]\cdots. \tag{3.3}$$

The first and fourth δ's gives zero because they result in a p connected directly to an upper pattern. The last two cancel. The second and third δ's gives a term of the form

$$\begin{matrix} a \\ b \\ \vdots \\ \vdots \end{matrix}$$

$$p\cdot EE_a'^n E_b'^j p_n p_j,$$

which is zero by the antisymmetry of the upper indices and because a p connected to an upper pattern gives zero. Hence the expression with the E' is equal to the expression with the E, and is thus an invariant.

4. ADJOINT INHOMOGENEOUS ALGEBRAS

We shall find all the independent invariants for the adjoint inhomogeneous algebra. The number of independent invariants can be found from the formula given by Beltrametti and Blasi.[3] The formula states that the number of invariants equals the number of parameters minus the rank of the matrix formed from the structure constants. The number of parameters for the inhomogeneous adjoint A_l is $2(l+1)^2 - 2$, which is twice as many as for A_l alone. We shall find the rank of the matrix by triangularizing it and using the well-known fact that the rank of a triangular matrix is equal to the number of nonzero elements in the diagonal. The commuting variables of Beltrametti and Blasi that replace the E's of the homogeneous part are here denoted by a's, and those which replace the p's are denoted by b's. The required matrix is then

$$\begin{pmatrix} b & a \\ 0 & -b \end{pmatrix},$$

where the a's and b's are both $[(l+1)^2 - 1]$ square matrices, and in fact, except for the different letters used, are identical. On the top line, the b on the left results from the commutation of E's and p's, and the a on the right from the commutation of E's; on the bottom line the zero results from the commutation of p's, and the $-b$ from commutation of p's and E's.

If a matrix is a triangular form, then it can be immediately seen that the determinant is the product of the diagonal elements. If the determinant is zero, then we strike out a row and a column of the matrix, so as to remove a zero from the diagonal. We continue this process until all the zeros have been removed to get a nonzero determinant. Thus we can immediately tell the rank of the original matrix by counting the number of nonzero elements in its diagonal. However this is only correct if, in striking out rows and columns, the diagonal is not shifted; otherwise we shall be counting the number of nonzero elements in the wrong diagonal. An example is the matrix

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which is obviously of rank 2, although there is one nonzero element in the diagonal. However in striking out the row and column with all zeros, we have shifted the diagonal.

Applying this to the present situation, we perform the operations of adding the various rows and columns to other rows and columns to triangularize the b matrices. In doing this the zero matrix will be affected, but this is meaningless and it will remain a zero matrix. Hence, after the b matrices have been triangularized, the whole matrix is triangular. We now strike out the required rows and columns, eliminating the zeros in the diagonal after rearranging the rows and columns if necessary, to keep the entire matrix in triangular form. Such rearrangement of one b matrix will not affect the other matrix nor will it change the zero matrix. Hence, after this series of operations, the entire matrix will be triangular with nonzero elements along the diagonal, and the number of such elements is the rank of the original matrix.

Now the b matrix is the same as the a matrix, and we know the number of nonzero diagonal elements in the entire matrix will be twice this number since there are two b matrices. But the number of parameters is also twice that of the homogeneous algebra; thus the number of invariants for the inhomogeneous adjoint A algebra is twice that of the A algebra, or $2l$. For the case of A_1 this gives the well-known result[4] of two

[3] E. G. Beltrametti and A. Blasi, Phys. Letters **20**, 62 (1966).

[4] W. Pauli, CERN 56-31 (1956), p. 13.

invariants. This result agrees with that obtained in a somewhat different manner.[2]

We shall now show that the scalars that we will use are invariants. The first type of invariant is a scalar containing only p's. This is clearly an invariant. The second type is a scalar containing only one E and any number of p's. Because the ordering is irrelevant, as other orderings differ only by scalars containing all p's and are not independent, for a given number of p's, it can be written in unique form and its commutation relations are

$$\begin{aligned}[p^\alpha_\beta, E^j_i p^k_j p^i_m \cdots p^s_k] &= -p^j_\beta p^k_j p^\alpha_m \cdots p^s_k + p^\alpha_i p^k_\beta p^i_m \cdots p^s_k \\ &= p^\alpha_m p^j_\beta(-p^m_n \cdots p^s_k p^k_j + p^m_q p^q_n \cdots p^k_j) = 0. \quad (4.1)\end{aligned}$$

(Some indices have been renamed in going to the second equation.)

Thus any scalar containing one or no E's is an invariant, and these are the only invariants we need.

The rule for obtaining independent invariants is to take each independent invariant of the homogeneous algebra and replace all the E's by p's to get the first set; to get the second set, we replace all but one of the E's by p's. Clearly this provides twice as many invariants as the homogeneous algebra has, which was required. It now remains to be proven that all these invariants are independent.

Firstly, in writing down any scalar, we shall take the completely symmetric form because this differs from a partially symmetric form only by terms of lower order, as discussed above. Suppose now that the completely symmetric form is not independent (so that it is possible to find a relation between it and the independent invariants), but suppose that the partially ordered form is independent. This would give a contradiction because we can substitute the equation for the partially symmetric form (which we write symbolically as $I_s = I_p + L$, where I_s and I_p are the symmetric and partially symmetric terms and L is a sum of terms of lower order) into the relationship that holds for the completely symmetric form. The relation for I_s we write as $I^n_s + J = 0$, where n is the power and J represents the other terms. After substitution, we get $I^n_p + J + nLI^{n-1}_p + \cdots = 0$; thus $I^n_p + J = 0$, as the other terms are of lower order; therefore these two terms must cancel each other. Hence we can take any form as independent; and we shall take the most symmetric one. Because the ordering of the terms is now irrelevant, we do not have to concern ourselves with whether or not the terms commute.

Turning now to the proof of the independence of the invariants, we first assume that it is possible to take an independent invariant of the homogeneous algebra, to replace all the E's by p's and to get an invariant which is not independent. This means that we can write an expression containing this invariant (and others) such that for each set of indices on any term, it is possible to find another term which will have a set of indices in its sum such that the term will be exactly the same as the first one (with opposite sign). However, this would be the case whether the term contained p's or E's; all that matters are the particular indices and their arrangement. Thus in the expression we could replace the p's by E's and still get cancellation. Therefore there is a relationship between invariants of the homogeneous algebra containing all E's. This is contrary to the original assumption that the invariant we are considering is independent. Hence we must conclude that the invariant containing the p's is independent.

Next consider an invariant containing one E, obtained from an independent invariant of the homogeneous algebra in the same way. Once again the argument applies: it too must be independent.

What can we now conclude about an invariant with the same structure as the ones above containing more than one E (including those containing no p's)? The argument again applies: these also are independent. But there are no further independent invariants; the ones described above (with one or no E's) exhaust the set. Therefore scalars similar to the independent ones with more than one E are not invariants.

As in the previous section we wish to consider what happens when the E's are made traceless. The invariant then becomes

$$E'^j_j p^j_k \cdots p^s_i - E p^i_k \cdots p^s_i. \quad (4.2)$$

E is an invariant because the p's have an equal number of upper and lower indices. The second term of the above expression is thus an invariant, so the term with the E' is invariant.

The independence arguments remain unchanged, and the scalars should be taken with traceless E's because the p's are traceless.

Therefore the invariants are those scalars with one or no E's, and all the independent ones are found by taking the independent scalars of the homogeneous algebra and replacing all the E's and all but one E, by p's.

5. REDUCIBLE INHOMOGENEOUS PART

When the p's form a reducible representation of the homogeneous part, there is a new way to form a scalar.

Before discussing this let us list in what way

reducible p's can occur in our notation. The p's will be reducible if the indices are not completely symmetrized in Young patterns, if the p's are not completely traceless, if two different types of representations appear (for example, p_i and p^i), and if one representation occurs more than once. This latter case will be handled by using different letters for the inhomogeneous operators (i.e., p_i and q_i).

In the discussion on the construction of scalars, it was not assumed that the indices were completely symmetrized and that the p's were traceless or of the same type; nor would the discussion be affected if some of the p's were replaced by q's, etc. So any scalar formed by completely saturating the indices is a scalar for a reducible representation also.

Of course many more scalars can be formed by this method if the p's are reducible. For example, the fundamental representation has no invariants containing only p's; but if we consider the fundamental representation and its contragradient, we get p_ip^i. The exception is A_1 for which the representation and its contragradient are equivalent and p_ip^i is zero.

While scalars formed according to the previously given rules are still scalars when the p's are reducible, this is no longer true of invariants. An expression that commutes with one irreducible representation of the reducible inhomogeneous part does not necessarily commute with other irreducible parts.

Since the adjoint case is important, it is worth noting that its p's need not be written in traceless form because any expression commutes with its trace, the scalar p:

$$[p_m^m, E_j^i \cdots] = p_j^m\delta_m^i \cdots - p_m^i\delta_j^m \cdots + \cdots,$$
$$= p_j^i \cdots - p_j^i \cdots = 0, \tag{5.1}$$

as might have been expected.

We now consider why there can be scalars other than those formed according to the previously stated rules. For an irreducible representation, a scalar, in completely lower notation, must have all its indices arranged in columns n indices long. This is not true for reducible representations.

To show this we take the sum of products of two similar irreducible representations (such that none of the terms is a scalar) and get a scalar. If the representations are identical, the value of the scalar is zero; the fact that they are different representations is crucial. For example, for $SU(2)$ consider the sum of two fundamental representations. Neither of them has an invariant, but $p_1q_2 - q_1p_2$ is an invariant. Commuting it with E_1^2, we get $p_1q_1 - q_1p_1$, which is zero, and similarly for E_1^2. Note that since we are not using saturated indices here, the expression is not invariant under $U(2)$ and so not under E_1^1 or E_2^2 individually; however, it is invariant under $E_z = E_1^1 - E_2^2$. So it is a scalar under $SU(2)$.

6. SEMISIMPLE HOMOGENEOUS PART

The algebra whose homogeneous part is semisimple but not simple remains to be discussed.

We first make a few remarks on the situation in which we deal with $IU(n)$ rather than $ISU(n)$. Denoting the operator E_j^i by E and letting M_k be an arbitrary scalar function of all operators other than E, an invariant must have the form $\sum_k (E)^kM_k$. Denoting any component of p by p_α and letting the difference of the number of upper from lower indices on p be z, we get

$$[p_\alpha, \textstyle\sum_k (E)^kM_k] = \sum \{E^k[p_\alpha, M_k] - zkE^{k-1}p_\alpha M_k\}$$
$$+ \text{lower order terms in } E. \tag{6.1}$$

Equating powers of E, we have

$$[p_\alpha, M_k] = z(k+1)p_\alpha M_{k+1} + \cdots \tag{6.2}$$

for each k and each component of p, as the requirement that the expression be an invariant. If z is zero, each M_k must be an invariant; in addition, E is an invariant.

Now we can always write p in more than one form. Changing the form will not change its commutation relations with M, but will change the value of z. Therefore we conclude that the only way an expression containing an E can be an invariant is if all the M_k are invariants and p is written with equal numbers of upper and lower indices (z is zero), and in this case, E is an invariant. This is not always possible.

It is also important to note that the various forms of the generators are not equivalent under $U(n)$ even though they are equivalent under $SU(n)$. An invariant under $SU(n)$ may not be an invariant under $U(n)$, although it may become an invariant if the generators are written in different form.

We now describe all compact inhomogeneous unitary algebras. They will consist of a semisimple algebra which is a direct sum of the simple algebras (with indices suppressed) E, $F, \cdots$, and Abelian invariant parts transforming according to representations of the homogeneous parts. Since E_j^i does not mix components but only multiplies them by a number, it is irrelevant here whether we consider $U(n)$ as simple or not.

The description of the homogeneous part is given by listing its simple components. We now turn to the description of the inhomogeneous part.

The entire inhomogeneous part will form a representation (in general, reducible) of E. Let us consider

one irreducible component which we label as p_i, p_j, etc., where each index represents some group of indices arranged in a Young pattern and

$$[E, p_i] = p_j. \tag{6.3}$$

We now consider the action of F on the p's:

$$[F, p_i] \equiv q_i, \tag{6.4}$$

and similarly for the other indices. Also

$$[E, q_i] \equiv r_i, \text{ etc.} \tag{6.5}$$

The indices on the E and F are suppressed and the letters represent any arbitrary generator in each simple part. Now the Jacobi identity gives

$$[[E, F], p_i] + [[F, p_i], E] + [[p_i, E], F] = 0. \tag{6.6}$$

Because E and F commute, we get $r_i = q_j$. Thus the representation q is isomorphic to the representation p. We can therefore label it by $p_{m,n}$, where the first index gives the component and the second the representation. The first index is affected only by the E, and the second only by the F. Furthermore, for fixed m the p's form a representation of F, by the definition of an inhomogeneous algebra, whose components are given by n. Without affecting the first index we can now symmetrize the indices represented by the n and reduce the representation with respect to F.

This process can be continued for the other simple algebras and the final result can be described as follows. We have a set of p's each transforming as the product of terms, each of the terms being an irreducible representation under one simple algebra and a scalar under the others. The inhomogeneous part will be a direct sum of the members of the set, each term in the direct sum being irreducible under all the simple algebras and each irreducible representation of a given simple algebra appearing with any other arbitrary representations of all other simple algebras any number of times (including zero). The enumeration of the irreducible representations, and the description of how they are combined, defines the Abelian part.

A simple special case is that in which the p's can be divided into sets, each a representation of one simple algebra and a scalar under all the others. Here the inhomogeneous algebra is a direct sum of what might be called completely simple inhomogeneous algebras.

The indices on the p's are arranged in Young patterns, each corresponding to an irreducible representation of a single simple algebra. Symmetrization of indices corresponding to different simple algebras is not invariant; thus the Young pattern is a product of patterns.

A scalar formed by saturation has each set of indices corresponding to a simple algebra separately saturated; scalars will not result from the saturation of indices appearing in two Young patterns which belong to different simple parts.

For the semisimple homogeneous case there are new ways to form scalars similar to the case of reducible p's. For one thing, a p will, in general, transform according to a sum of irreducible representations of each simple part; the different irreducible representations are labeled by indices going with the other simple parts.

Another consideration is that a term, in general, consists of products of the p's with the generators from different simple parts. Each generator will affect the p being commuted with that term. Thus a scalar which commutes with all the p's may not do so if the generators from all but one simple part are set equal to zero.

An example is the Poincaré algebra whose homogeneous part is the complex extension of a direct sum of two A_1 algebras and whose inhomogeneous part is the product of the fundamental representation of each of the two simple algebras.[5] Letting the simple algebras be A and B and the inhomogeneous operators be p_{ij}, where A acts on the first index and B on the second, we get for the first invariant $p_{12}p_{21} - p_{11}p_{22} = p^2$, the invariant distance. Notice the similarity between this invariant and the one discussed in the section on reducible p's. Under commutation with A_1^2, for example, it becomes $p_{12}p_{11} - p_{11}p_{12} = 0$. If the second index had been suppressed, the invariant would be zero.

Of the sum of terms making up the second invariant, let us consider just the following:

$$B_1^2B_2^1(p_{12}p_{21} - p_{11}p_{22}) + B_1^2(-A_2^1p_{12}p_{12} + (A_1^1 - A_2^2)p_{12}p_{22} - A_1^2p_{22}p_{22}). \tag{6.7}$$

Commuting this with p_{11}, we find that the contribution from the first term is needed to cancel the second term. The second term is a scalar under the A algebra; but in order to make it an invariant, the term from the B algebra is needed.

APPENDIX

We wish here to derive the expression

$$[E_i^j, T_\Upsilon] = -\sum_k T_{\Upsilon_i}\delta^j_{\alpha_k} \tag{A1}$$

and

$$[E_i^j, T^\Upsilon] = \sum_k T^{\Upsilon_j}\delta_i^{\beta_k}, \tag{A2}$$

[5] V. Heine, *Group Theory in Quantum Mechanics* (The Macmillan Company, New York, 1960), Sec. 31.

where the (Y's) indicate some Young diagram describing the representation to which the tensor T belongs. The indices in the pattern are β_j and α_j on the top and bottom, respectively. On the right-hand side the sum is over all boxes in the pattern and α_k is replaced by i, as indicated by the i subscript to the Y, [and similarly for j and β_k in Eq. (A2)]. The commutation relation for a mixed tensor is similar, and gives a sum over all boxes with the i replacing all lower indices in turn, with a minus sign, and with the j replacing all upper indices in turn.

To derive these relations we start with the commutation relation for the fundamental representation

$$[E_i^j, p_k] = -p_i\delta_k^j \tag{A3}$$

and note that the commutation relation for its contragradient must be

$$[E_i^j, p^k] = p^j\delta_i^k \tag{A4}$$

in order that the product $p_m p^n$ transform like the adjoint representation

$$[E_i^j, p_m p^n] = p_m p^j\delta_i^n - p_i p^n\delta_m^j. \tag{A5}$$

A product of tensors $\prod_g T_g$, which transforms according to any representation, possibly different for different values of the index, has the commutation relation

$$[E_i^j, \prod T_g] = \sum T_1 T_2 \cdots T_{m-1} C^{\beta}_{(ij)m} T_\beta T_{m+1} \cdots, \tag{A6}$$

where the C is the relevant structure constant and the sum is over all subscripts. This can be proven by noting that

$$[E_i^j, T_1 T_2] = T_1[E_i^j, T_2] + [E_i^j, T_1]T_2, \tag{A7}$$

which can be shown by writing out the commutators. Now each T can itself be a product of tensors, and so we get Eq. (A6) by induction.

In particular,

$$[E_\alpha^\beta, p_1 p_2 \cdots p_k] = p_\alpha p_2 \cdots p_k\delta_1^\beta + p_1 p_\alpha \cdots \delta_2^\beta + \cdots. \tag{A8}$$

The tensor $T_{ijk\cdots m}$ (no symmetrization of indices) transforms like $p_i p_j p_k \cdots p_m$, and similarly for upper and mixed indices. A tensor symmetrized according to some Young diagram is a sum, with proper signs, of unsymmetrized terms of the form $T_{i\cdots m}$. (Over-all multiplicative factors are irrelevant.) Each term of the sum transforms according to Eq. (A8).

The commutator of a symmetrized tensor equals a sum (with proper sign) of terms each with a different arrangement of indices, and each of these terms is itself a sum of terms similar to the right-hand side of Eq. (A8). Let us take the subset of this double sum which has the index s on the delta symbol. This subset is a sum of terms with proper sign, with all arrangements of the indices, and with α replacing s. Thus it is a tensor symmetrized according to the Young pattern which described the tensor inside the commutator, and since we had a double sum, we now have a sum of terms of symmetrized tensors, each with different indices on the delta symbol. This is Eq. (A1) [or Eq. (A2)].

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Number of Polynomial Invariants of Adjoint and Fundamental Compact Inhomogeneous Unitary Algebras*

RONALD MIRMAN
Physics Department, Long Island University, The Brooklyn Center Zeckendorf Campus, Brooklyn, New York

(Received 9 February 1967)

The number of invariants for two types of inhomogeneous Lie algebras is computed.

A formula for the number of independent polynomial invariants for any Lie algebra has been given by Beltrametti and Blasi,[1] and by Pauri and Prosperi.[2] Here we wish to use the formula to find the number of such invariants for two compact inhomogeneous algebras; those Lie algebras which are a direct sum of a compact Lie algebra of type A (whose generators are denoted by E), plus an Abelian algebra transforming according to a representation of the simple part (whose generators are denoted by p)—here specifically the adjoint and fundamental representations.

The formula gives the number of invariants as the number of parameters of the algebra minus the rank of a matrix formed from the structure constants and the generators. It is the evaluation of this rank which we consider here. The number of parameters is obvious.

It might be mentioned that we do not believe that the proof of the formula is complete. The difficulty is that in the proof a correspondence is used between the invariants and the solutions of certain partial differential equations. However, the solutions of the equations can conceivably involve negative powers of the variables. But the concept of negative power of a generator is undefined, and an invariant containing such a negative power is meaningless. So the number of invariants may be less than that given by the formula which we shall take to be a formula for the maximum number of invariants, a concept which is still of great use.[3]

For the fundamental inhomogeneous algebra the number of the parameters is $(l+1)^2 - 1 + l + 1$.

The number of parameters is also the order of the matrix, and it is always odd. Now the matrix is antisymmetric, if the same order of rows and columns is chosen, because the structure constants are antisymmetric. An antisymmetric determinant of odd order always vanishes. Hence the rank must be at least one less than the order. We show that it is exactly one less than the order.

To do this, we set all the parameters except a few equal to zero and compute the value of the determinant containing the remaining parameters. This will be equal to some of the terms in the expansion of the determinant of the complete matrix, with the other terms in that expansion all containing parameters that were set equal to zero, none of which appear in the terms we shall write down. Hence, if our result is nonzero, there will be no terms in the expansion of the determinant which will cancel the terms given below and so the determinant will not be zero.

The only parameters that we will not set equal to zero are those corresponding to the generators of the Abelian part, and of the Cartan subalgebra of the simple part.

In writing the matrix we first write on the extreme left the generator appearing first in the commutator, and on the top line the generator appearing second in the commutator. The matrix element in the row whose most extreme term (which is of course not an element of the matrix) is E_a, and the column whose highest term is E_b, has the value of the commutator of E_a and E_b. The matrix (see top of next page) is then $(n = l + 1)$.

We now strike out the column headed by p_1 and the row whose left-most member is p_1, and wish to show that the remaining matrix has a nonzero determinant.

The only nonzero elements in the subdeterminant formed by the E's and H's are the H's, one nonzero linear combination in each row and column of that submatrix at the intersection of the row E^i_j and column E^j_i.

To evaluate the determinant, we expand by minors starting with the first row, which contains one element $-p_2$. Striking out this row and column we go on to the second row, which now contains one element $-2p_3$. We continue this way for the first l rows, and we find that each in turn contains exactly one element. We next evaluate the minor of the last term $-p_n$ which is the determinant of the submatrix to the left of the column headed by p_1, and below the row whose left-most element is H_l. This we evaluate by columns, starting with the first which contains one element p_2. We continue as above giving, up to a numerical factor, a product of the p's times the determinant of the matrix formed by the columns

* Work supported by a grant from Long Island University.

[1] E. G. Beltrametti and A. Blasi, Phys. Letters **20**, 62 (1966).

[2] M. Pauri and G. M. Prosperi, Nuovo Cimento **43**, 533 (1966).

[3] E. Angelopoulos, Compt. Rend. **264**, 585 (1967).

	H_1	$H_2 \cdots H_l$	E_2^1	$E_3^1 \cdots E_n^1$	$E_1^2 \cdots E_{n-1}^n$	p_1	p_2	p_3	$p_4 \cdots p_n$
H_1	0	$0 \cdots 0$	0	$\cdots$	0	p_1	$-p_2$	0	$0 \cdots 0$
H_2	0	$0 \cdots 0$			0	p_1	p_2	$-2p_3$	$0 \cdots 0$
$\cdot$	$\cdot$								
$\cdot$	$\cdot$								
$\cdot$	$\cdot$								
H_l	0	$\cdots 0$			0	p_1			$-np_n$
E_2^1	0	$\cdots$		H_1 $\cdots$	0	p_2	0	$\cdots$	0
E_3^1	0	$\cdots$				p_3	0	$\cdots$	
$\cdot$									
$\cdot$									
$\cdot$									
E_{n-1}^n		0	$\cdots$			0	0	$\cdots$	p_{n-1}
p_1	$-p_1$	$-p_1$		$\cdots$		0	0		0
p_2	p_2	$-p_2$		$\cdots$					
$\cdot$	$\cdot$								
$\cdot$	$\cdot$								
$\cdot$	$\cdot$								
p_n	0	$\cdots$	$np_n 0$		$\cdots$				0

and rows which are headed by an E. But this contains one, and only one, nonzero element in each row and column and so is nonzero. Therefore the entire determinant is nonzero.

The number of parameters is equal to the order of the matrix, the rank is one less than the order, so the number of independent invariants is one.

In order to compute the number of invariants for the adjoint inhomogeneous algebra, we need the number of invariants for the homogeneous part A_l. Because of our previous comment that the number of invariants predicted may be greater than the actual number of invariants, we must show that for the homogeneous part, the number predicted equals the correct number l. This is to say we must show that the rank of the matrix is not less than $n^2 - 1 - l$, since if it were, more than l would be predicted. Now the matrix is the above one, ignoring the columns and rows headed by p's. If we strike out the columns and rows headed by H's, we are left with a matrix that we have just shown is nonzero. But the rank of the matrix is just $n^2 - 1 - l$, as required.

Turning now to the adjoint inhomogeneous algebra we can use the known results for the simple part to evaluate the matrix. The number of parameters for A is $(l + 1)^2 - 1$, and for the inhomogeneous algebra it is twice this or $2(l + 1)^2 - 2$.

The matrix whose rank we want is of the form

$$\begin{pmatrix} b & a \\ 0 & -b \end{pmatrix},$$

where the b submatrices are formed from the structure constants going with the commutator of the E's and p's. The a submatrix is formed from the structure constants of the E's with themselves and the zero results because the p's commute, so all the structure constants are zero. Note the a and b submatrices are identical, except for the different letters used for the generators.

Let us first replace the $-b$ submatrix by an arbitrary matrix c. Now we have just shown that the rank of the $((l + 1)^2 - 1)$ square submatrix a is $(l + 1)^2 - 1 - l$. Hence l columns are linearly dependent on the other columns. This must therefore be true for the b matrix, since it is identical to the a matrix, and for the whole matrix since the b matrix has zero under it. This means that without changing the rank we can replace l columns, by zeros. Also, since the row rank and the column rank are the same, l rows must be linearly dependent on the other rows, and these must be in the top half of the matrix since the submatrix c is arbitrary. Hence l of the top rows can be replaced by zeros.

We can now replace c by $-b$ and repeat the above argument, interchanging the words, rows and columns. This means that we will have replaced $2l$ rows, and $2l$ columns by zeros.

Suppose now that we could find another relation between the columns. This relationship could not be among the columns of the left-half of the matrix because we know that there are no further relations among the rows of the b matrix. Nor, can it be between the entire set of rows, for this would mean that there was another relation among the columns of b besides those already used. This would also be true if we could find a relation among the columns on the right side of the matrix. But we know that there are

only l relation among the columns. Hence the matrix we have finally achieved has a nonzero determinant. So the rank is $2(l+1)^2 - 2 - 2l$ and since the number of parameters is $2(l+1)^2 - 2$, the number of invariants of the adjoint inhomogeneous algebra is $2l$.

The algebra of the group $U(n)$ is worth some special consideration because there are different notations for the p's, which are equivalent for $SU(n)$ but not for $U(n)$, and it turns out that the number of invariants differs depending on the form.

Let us consider the fundamental algebra. There is now one more operator than for the inhomogeneous $SU(n)$, and the order of the matrix is therefore even. The matrix is, using the same notation as before,

	E_1^1	E_2^1	$E_3^1 \cdots E_1^2$	E_2^2	$\cdots$	E_n^n	p_1	p_2	$p_3 \cdots p_n$
E_1^1	0	0	$0 \cdots 0$	0	$\cdots$	0	p_1	0	$0 \cdots 0$
E_2^1	0	0 $\cdots$	$E_1^1 - E_2^2$	0	$\cdots$	0	p_2	0	$0 \cdots 0$
E_3^1	0	$\cdots$				0	p_3	0	$0 \cdots 0$
.	.								.
.	.								.
.	.								.
E_1^2	0	$-E_1^1 + E_2^2$		$\cdots$		0	0	p_1	$0 \cdots 0$
E_2^2	0	0		$\cdots$		0	0	p_2	$0 \cdots 0$
E_3^2	0			$\cdots$		0	0	p_3	$0 \cdots 0$
.	.								.
.	.								.
.	.								.
E_1^3	0			$\cdots$		0	0	0	$p_1 \cdots 0$
E_2^3	0			$\cdots$		0	0	0	$p_2 \cdots 0$
E_3^3	0			$\cdots$		0	0	0	$p_3 \cdots 0$
.	.								.
.	.								.
.	.								.
E_n^n			$\cdots$				0		$\cdots p_n$
p_1	$-p_1$	$-p_2$		$\cdots$			0		$\cdots 0$
p_2	0	0		$\cdots$			0		$\cdots 0$
p_3	0								
.	.						.		.
.	.						.		.
.	.						.		.
p_n	0			$\cdots$		$-p_n$	0		$\cdots 0$

To evaluate the determinant we expand by rows. The first row has one element p_1, the second one element $E_1^1 - E_2^2$, etc., for each of the first n rows. We now expand by columns in the same way, as the first n columns each have one element. After the first n columns we return to the rows, the second $n-1$ of which now have one and only one element each, as their second element was in the first n columns, which have been stricken out. Continuing in this way we see that we can always find one row or column with one and only one element, which is nonzero. Therefore the determinant is nonzero.

So the order of the determinant, which equals the number of parameters, equals the rank. There are no invariants.

Taking $SU(3)$ as an example, there are two ways of writing the p's—i.e., p_i and p^k_j. (The indices are arranged in Young patterns, so the upper two indices are antisymmetric.) Suppose that we were to write the invariant of $SU(3)$ with a mixture of these two forms, such that the number of upper and lower indices are equal. Going to $U(3)$ we add one operator E_i^i (sum), which when commuted with an expression gives the difference in the number of upper and lower indices, here zero. Therefore the invariant under $SU(3)$ remains an invariant under $U(3)$. This does not contradict the previous results because in order to apply the formula we have to treat p_i and p^k_j as being different since they are not equivalent under $U(3)$. The matrix whose rank we have to compute is not the one above, but a 15×15 matrix.

We shall not do this here because for $SU(n)$ the different forms of p are equivalent and we have already found the result for that algebra.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Unitary Representations of the Homogeneous Lorentz Group in an $O(2, 1)$ Basis*

N. MUKUNDA†
Physics Department, Syracuse University, Syracuse, New York

(Received 27 April 1967)

Unitary irreducible representations of the homogeneous Lorentz group, $SO(3, 1)$, belonging to the principal series and containing integral angular momenta, are reduced with respect to the subgroup $SO(2, 1)$. It is found that the representation $\{j_0, \rho\}$ of $SO(3, 1)$, where j_0 is a nonnegative integer and ρ a real number, contains each representation of $SO(2, 1)$ of the continuous class (nonexceptional and integral) twice, and each of the discrete representations $D_k^{(\pm)}$ of $SO(2, 1)$ once, for $k = 1, 2, \cdots, j_0$. The latter representations are absent for $j_0 = 0$. It is shown that the basis states of the representation $D_k^{(\pm)}$ (for $k \geq 2$) lie in the domain of those generators of $SO(3, 1)$ that are outside the $SO(2, 1)$ subalgebra, while the states of the representations $D_1^{(\pm)}$ do not lie in this domain. It is further shown that from the point of view of the nature of this domain, the representations $D_1^{(\pm)}$ of $SO(2, 1)$ are very intimately connected to the continuous class representations of $SO(2, 1)$, and that these two discrete representations act as a bridge between the remaining discrete representations on the one hand, and the continuous class representations on the other.

INTRODUCTION

Several recent developments in elementary particle physics indicate the growing usefulness of noncompact groups in understanding the properties of particle multiplets.[1] From the practical viewpoint, the treatment of the unitary irreducible representations (UIR's) of noncompact semisimple Lie groups is of varying complexity, the degree depending on the group involved; in a sense, the complexity increases with the "amount of noncompactness" of the group. For example, UIR's of the real pseudo-orthogonal groups of the type $SO(n, 1)$ are reasonably manageable, when expressed in a basis made up of finite dimensional UIR's of the maximal compact subgroup $SO(n)$.[2] The same is true for the UIR's of the pseudo unitary groups of the type $SU(n, 1)$ when reduced with respect to the compact subgroup $SU(n) \otimes U(1)$.[3] However, the analysis is much harder for the case of groups like $SO(p, q)$ or $SU(p, q)$ when both p and q are larger than unity.[4] In a previous paper[5] hereafter referred to as (I), we have considered an example of the reduction of UIR's of a noncompact group with respect to a noncompact subgroup, and have examined some of the new features which are involved in this process. The case considered was the group $SO(2, 1)$; UIR's of this group were expressed in a "basis" made up of irreducible unitary representations of the noncompact (Abelian) subgroup $O(1, 1)$.

In the present paper, we consider the reduction of UIR's of the homogeneous Lorentz group $SO(3, 1)$ with respect to its noncompact subgroup $SO(2, 1)$. [The Lie algebra is the same for the two groups $O(3, 1)$ and $SL(2, C)$; similarly $SO(2, 1)$, $SU(1, 1)$, and $SL(2, R)$ have the same Lie algebra.] The methods used are similar to those in (I), and an interesting point to understand is the way the generators of $SO(3, 1)$ act in the basis suited to reduction of the UIR's under $SO(2, 1)$. In the case treated in (I), we were interested in diagonalizing a generator of $SO(2, 1)$ corresponding to the subgroup $O(1, 1)$; this generator has a real continuous spectrum covering the real line either twice or once, depending on whether the UIR of $SO(2, 1)$ belongs to the continuous or the discrete class. In the present case, we wish to diagonalize both the quadratic Casimir operator Q of $SO(2, 1)$, and the compact generator J_3 of $SO(2, 1)$; and the situation turns out to be somewhat more complicated than before because the first of these

* Work supported in part by the U.S. Atomic Energy Commission.

† Present address: Tata Institute of Fundamental Research, Bombay, India.

[1] A partial list of references is: A. O. Barut, Phys. Rev. **135**, B839 (1964); Y. Dothan, M. Gell-Mann, and Y. Ne'eman, Phys. Letters **17**, 148 (1965); T. Cook, C. J. Goebel, and B. Sakita, Phys. Rev. Letters, **15**, 35 (1965); N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan, *ibid.* **15**, 1041 (1965); C. Fronsdal, in *Proceedings of the Third Coral Gables Conference*, (W. H. Freeman & Co., San Francisco, 1966); E. C. G. Sudarshan, *ibid.*; Y. Nambu, Progr. Theoret. Phys. **37**, 368 (1966); Phys. Rev. **160**, 1171 (1967).

[2] The well-known cases whose treatments are now classic are $SO(2, 1)$: V. Bargmann, Ann. Math. **48**, 568 (1947); $SO(3, 1)$: Harish-Chandra, Proc. Roy. Soc. **A189**, 272 (1947); I. M. Gel'fand and M. A. Naimark, Izv. Acad. Nauk. SSSR **11**, 411 (1947); $SO(4, 1)$: L. H. Thomas, Ann. Math. **42**, 113 (1941); T. D. Newton, *ibid.* **51**, 73 (1950); J. Dixmier, Bull. Soc. Math. (France) **89**, 9 (1961).

[3] See, for instance, the contribution by L. C. Biedenharn, in *Proceedings of the Milwaukee Conference on Non-Compact Groups in Particle Physics*, Y. Chow, Ed. (W. A. Benjamin & Co., New York, 1966).

[4] Some recent papers dealing with particular representations of these groups are: J. Fischer and R. Raczka, Commun. Math. Phys. **4**, 8 (1967); J. Niederle, ICTP preprint IC(66)99, Trieste (1966).

[5] N. Mukunda, Unitary Representations of the Group $O(2, 1)$ in an $O(1, 1)$ Basis, Syracuse University Preprint, 1967; J. Math. Phys. (to be published).

operators has both a continuous and a discrete spectrum.

The material in this paper is arranged as follows: Section 1 contains a collection of the pertinent facts concerning the group $SL(2, C)$, its Lie algebra, and its UIR's. In Sec. 2, we carry out the diagonalization of the operators Q and J_3; this leads to the reduction of the UIR's of $SL(2, C)$ with respect to $SU(1, 1)$. Attention will be restricted to the Principal Series of UIR's of $SL(2, C)$ and, among these, for the sake of simplicity, to single valued UIR's of $SO(3, 1)$. These UIR's contain only integral angular momenta. In Sec. 3 we examine the manner in which the generators of $SL(2, C)$ [outside of the subalgebra generating $SU(1, 1)$] act in the basis wherein Q and J_3 are diagonal, and how they perform the function of connecting different UIR's of $SU(1, 1)$ to one another. A similar treatment of the supplementary series of UIR's of $SL(2, C)$ will be presented elsewhere.

Note added in proof: [After this paper was written, the following two papers on the problem of reducing the principal series of UIR's of $SL(2, C)$ with respect to $SU(1, 1)$ have come to the author's attention: A. Sciarrino and M. Toller, J. Math. Phys. **8,** 1252 (1967); S. Ström, Arkiv Fysik **34,** 215 (1967).]

1. RESUMÉ OF $SL(2, C)$ AND ITS UNITARY REPRESENTATIONS

The group $SL(2, C)$[6] is the universal covering group of the homogeneous Lorentz group $SO(3, 1)$, the correspondence of elements from the former to the latter being two to one. Elements g of $SL(2, C)$ are in one-to-one correspondence with unimodular complex matrices in two dimensions:

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

Here $\alpha\beta\gamma\delta$ are complex numbers restricted only by the unimodularity condition. Elements of $SL(2, C)$ may then be parametrized by six real variables. The maximal compact subgroup $SU(2)$ of $SL(2, C)$ is made up of all elements corresponding to unimodular unitary matrices of the form

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

[The bar denotes complex conjugation.] $SU(2)$ is, of course, the universal covering group of $SO(3)$, the real orthogonal group in three dimensions. The noncompact $SU(1, 1)$ subgroup of $SL(2, C)$ consists of elements corresponding to matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad |\alpha|^2 - |\beta|^2 = 1. \tag{1.3}$$

There is a two-to-one homomorphism from $SU(1, 1)$ onto the real pseudo-orthogonal group in three dimensions, $SO(2, 1)$. [This is the homogeneous Lorentz group in two space and one time dimensions.] However, $SU(1, 1)$ is not the universal covering group of $SO(2, 1)$.

The Lie algebra of $SL(2, C)$ is six-dimensional, and is spanned by the elements J_j, K_j $(j = 1, 2, 3)$. The commutation relations are

$$[J_j, J_k] = i\epsilon_{jkl}J_l, \tag{1.4a}$$

$$[J_j, K_k] = i\epsilon_{jkl}K_l, \tag{1.4b}$$

$$[K_j, K_k] = -i\epsilon_{jkl}J_l. \tag{1.4c}$$

In a UIR of $SL(2, C)$, these six generators would be represented by self-adjoint operators. In the nonunitary representation of $SL(2, C)$ by two-dimensional matrices as in (1.1), the generator matrices may be taken to be

$$J_j = \tfrac{1}{2}\sigma_j, \quad K_j = (i/2)\sigma_j, \tag{1.5}$$

where the σ_j are the Pauli matrices. The Lie algebra of $SL(2, C)$ possesses two independent Casmimir invariants, both of the second degree in the generators. They are

$$C_1 = K_jK_j - J_jJ_j, \quad C_2 = K_jJ_j; \tag{1.6}$$

C_1 and C_2 commute with all the generators, and in a UIR of $SL(2, C)$ they are represented by real numbers.

The $SU(2)$ subgroup of $SL(2, C)$ is generated by the three elements J_j obeying (1.4a). The Casimir invariant of $SU(2)$ is

$$J^2 = J_jJ_j. \tag{1.7}$$

As is well known from elementary quantum mechanics, the eigenvalues of J^2 are of the form

$$J^2 = j(j + 1), \quad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots, \tag{1.8}$$

and the inequivalent UIR's of $SU(2)$ can be labeled by j. The UIR corresponding to a given j is $(2j + 1)$-dimensional, and in it the spectrum of J_3 is

$$J_3 = -j, -j + 1, \cdots, j - 1, j. \tag{1.9}$$

With respect to the $SU(2)$ subgroup generated by J_j, the remaining generators K_j behave as a vector operator. That is, under the finite transformations generated by the J_j, the K_j transform linearly among themselves according to the UIR of $SU(2)$ corresponding to $j = 1$.

[6] A very lucid discussion of this group and its representations may be found in: I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Group and their Applications* (The Macmillan Company, New York, 1963).

The $SU(1,1)$ subgroup of $SL(2,C)$ may be taken to be generated by the elements J_3, K_1, and K_2. They obey the commutation rules:

$$[J_3, K_1] = iK_2; \quad [J_3, K_2] = -iK_1; \quad [K_1, K_2] = -iJ_3. \tag{1.10}$$

The quadratic Casimir operator of $SU(1,1)$ is the operator Q given by

$$Q = K_1^2 + K_2^2 - J_3^2. \tag{1.11}$$

The different types of single-valued UIR's of $SU(1,1)$ have been listed in (I), and may be briefly described as follows:[7] They fall into three distinct classes, the continuous nonexceptional class, the continuous exceptional class, and the discrete class. In the first class, either $\frac{1}{4} \leq Q < \infty$ and the spectrum of J_3 is $0, \pm 1, \pm 2, \pm 3, \cdots, \pm\infty$, or $\frac{1}{4} < Q < \infty$ and the spectrum of J_3 is $\pm\frac{1}{2}, \pm\frac{3}{2}, \cdots, \pm\infty$. In the second class, the spectrum of J_3 consists of $0, \pm 1, \pm 2, \cdots, \pm\infty$, while Q lies in the interval $0 < Q < \frac{1}{4}$. In the third class, the possible values of Q are of the form

$$Q = k(1-k),$$
$$k = \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \cdots,$$

while J_3 increases from k to $+\infty$ in integer steps, or decreases from $-k$ to $-\infty$ in integer steps. In the present paper, we will have to deal only with the first and third classes of UIR's of $SU(1,1)$.

With respect to the $SU(1,1)$ subgroup generated by J_3, K_1, K_2, the remaining generators J_1, J_2, and K_3 behave as a three-dimensional "vector." That is, under finite $SU(1,1)$ transformations, they go over into real linear combinations of themselves according to the non-unitary three-dimensional irreducible representation of $SU(1,1)$. [This is the defining representation of the group $SO(2,1)$.] This transformation property under $SU(1,1)$ is shared by the generators of $SU(1,1)$.

We conclude this section by listing the different kinds of UIR's of the group $SL(2,C)$.[8] These may be specified by the values of the Casimir invariants C_1 and C_2, and the spectrum of UIR's of the compact subgroup $SU(2)$ (spectrum of j-values) that appear in the particular UIR of $SL(2,C)$. [Each UIR of $SU(2)$ appears no more than once in a given UIR of $SL(2,C)$.] There are two classes or series of UIR's of $SL(2,C)$:

(*A*) *Principal series*, $\{j_0, \rho\}$: Here the possible values of j are

$$j = j_0, j_0 + 1, j_0 + 2, \cdots, \textit{ad inf.},$$

where j_0 is a fixed positive integer, half odd integer, or zero:

$$j_0 = 0, \tfrac{1}{2}, 1, \tfrac{3}{2} \cdots$$

The values of C_1 and C_2 are

$$C_1 = 1 + \rho^2 - j_0^2, \quad C_2 = -\rho j_0;$$

ρ is a real parameter in the range $-\infty < \rho < \infty$. Therefore, UIR's of the principal series are characterized by j_0, the lowest UIR of $SU(2)$ present, and the real number ρ.

(*B*) *Supplementary series*, $\{0, i\rho\}$: Here the spectrum of j values is

$$j = 0, 1, 2, \cdots, \textit{ad inf.},$$

and the Casimir operators have the values

$$C_1 = 1 - \rho^2, \quad C_2 = 0, \quad 0 < \rho < 1.$$

Thus, these UIR's are labeled by the single real number ρ.

For both series of UIR's the generators of $SL(2,C)$ may be specified by their matrix elements in a basis made up of orthonormal eigenvectors of J^2 and J_3. These vectors are denoted in the usual fashion by the kets $|j, m\rangle$:

$$\langle j', m' \,|\, j, m\rangle = \delta_{j'j}\delta_{m'm};$$
$$J^2 |j, m\rangle = j(j+1)\,|j, m\rangle; \quad J_3 |j, m\rangle = m\,|j, m\rangle;$$
$$-j \leq m \leq j; \quad j = j_0, j_0 + 1 \cdots. \tag{1.12}$$

In place of the Cartesian components J_j, K_j, it is convenient to use the spherical components J_M, K_M, $M = 0, \pm 1$, defined by

$$J_{\pm 1} = \mp \frac{J_1 \pm iJ_2}{\sqrt{2}}, \quad J_0 = J_3;$$
$$K_{\pm 1} = \mp \frac{K_1 \pm iK_2}{\sqrt{2}}, \quad K_0 = K_3. \tag{1.13}$$

Then the matrix elements of J_M and K_M are given in terms of the $SU(2)$ Clebsch–Gordan coefficients and reduced matrix elements by

$$\langle j', m'|\, J_M\, |j, m\rangle = [j(j+1)]^{\frac{1}{2}} \delta_{j'j} C^{j1j}_{mMm'}$$
$$\langle j+1, m'|\, K_M\, |j, m\rangle = -[(j+1)(2j+1)]^{\frac{1}{2}} C^{j1j+1}_{mMm'} C_{j+1}$$
$$\langle j, m'|\, K_M\, |j, m\rangle = -[j(j+1)]^{\frac{1}{2}} C^{j1j}_{mMm'} A_j$$
$$\langle j-1, m'|\, K_M\, |j, m\rangle = -[j(2j+1)]^{\frac{1}{2}} C^{j1j-1}_{mMm'} C_j$$
$$A_j = \frac{ij_0 C}{j(j+1)}, \quad C_j = \frac{i}{j}\left[\frac{(j^2 - j_0^2)(j^2 - C^2)}{4j^2 - 1}\right]^{\frac{1}{2}}. \tag{1.14}$$

For UIR's of the Principal series, $C = -i\rho$, while for UIR's of the Supplementary series $C = \rho$. The selection rule, $\Delta j = 0, \pm 1$ for the non-zero matrix elements of K_M, is a consequence of the vector property of K_M.

We shall be concerned with the UIR's of the Principal series alone, for arbitrary ρ and integral j_0. These are single-valued UIR's of the homogeneous Lorentz group $SO(3,1)$.

[7] V. Bargmann, Ref. 2.
[8] See, for instance, Ref. 6, p. 199.

2. REDUCTION OF UIR'S OF $SO(3, 1)$ UNDER $SO(2, 1)$

It is possible to construct the UIR's of the principal series of $SL(2, C)$ via unitary transformations in a Hilbert space H of square integrable functions over a certain domain.[9] One considers elements of H to correspond to functions of a complex variable ξ, with ξ varying over the entire complex plane. [The specification of a vector f in H by means of a function $f(\xi)$ is only possible up to sets of measure zero.] These are not analytic functions, but complex valued functions of the real and imaginary parts of ξ. The scalar product and norm in H are defined by

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

Let g be an element of $SL(2, C)$ corresponding to the matrix on the right-hand side of (1.1). In the UIR $\{j_0, \rho\}$ of the principal series, the element g is represented by the unitary operator $U(g)$ which acts on functions $f(\xi)$ in the following way

$$[U(g)f](\xi) = (\delta + \beta\xi)^{j_0 - 1 + i\rho} (\bar{\delta} + \bar{\beta}\bar{\xi})^{-j_0 - 1 + i\rho} \times f\left(\frac{\alpha\xi + \gamma}{\beta\xi + \delta}\right). \tag{2.2}$$

By using the matrices for the generators as given in (1.5), and taking g to be an element of each of the one-parameter subgroups $\exp(i\tau J_1), \cdots, \exp(i\tau K_3)$ in turn, one can express the generators as differential operators in x and y.

One has

$$\begin{aligned}
[J_1 f](\xi) &= j_0 x f(\xi) - (\rho + i) y f(\xi) \\
&\quad - \frac{i}{2}(y^2 - x^2 + 1)\frac{\partial f}{\partial y} - ixy\frac{\partial f}{\partial x}, \\
[J_2 f](\xi) &= j_0 y f(\xi) + (\rho + i) x f(\xi) \\
&\quad + \frac{i}{2}(x^2 - y^2 + 1)\frac{\partial f}{\partial x} + ixy\frac{\partial f}{\partial y}, \\
[J_3 f](\xi) &= -j_0 f(\xi) + iy\frac{\partial f}{\partial x} - ix\frac{\partial f}{\partial y}, \\
[K_1 f](\xi) &= -j_0 y f(\xi) - (\rho + i) x f(\xi) \\
&\quad - \frac{i}{2}(x^2 - y^2 - 1)\frac{\partial f}{\partial x} - ixy\frac{\partial f}{\partial y}, \\
[K_2 f](\xi) &= j_0 x f(\xi) - (\rho + i) y f(\xi) \\
&\quad + \frac{i}{2}(x^2 - y^2 + 1)\frac{\partial f}{\partial y} - ixy\frac{\partial f}{\partial x}, \\
[K_3 f](\xi) &= (\rho + i) f(\xi) + ix\frac{\partial f}{\partial x} + iy\frac{\partial f}{\partial y}.
\end{aligned} \tag{2.3}$$

[9] Reference 6, p. 247.

Actually, for our purpose, this form for the UIR's is not the most convenient one. It would be better to have the representations in a Hilbert space of functions defined on the unit sphere; in that case, the effect of a transformation $U(g)$, where $g \in SU(2)$, is much easier to visualize, and the functions corresponding to the basis states $|j, m\rangle$ of (1.12), can be easily identified.

With this in mind, we define a mapping from the ξ plane ($\xi = x + iy$), to the points on the unit sphere, with polar angles θ and φ, as follows:

$$\tan\frac{\theta}{2} = \frac{1}{|\xi|} = [x^2 + y^2]^{-\frac{1}{2}} \qquad \tan\varphi = y/x; \quad 0 \le \theta \le \pi, \quad 0 \le \varphi \le 2\pi. \tag{2.4}$$

The region within the unit circle in the ξ-plane, $|\xi| \le 1$, gets mapped onto the lower hemisphere $\frac{1}{2}\pi \le \theta \le \pi$; and the exterior of the unit circle $|\xi| \ge 1$ is mapped onto the upper hemisphere $0 \le \theta \le \frac{1}{2}\pi$. Simultaneously, we must define the mapping from functions of ξ to functions of θ and φ. If an element f in H is specified by the function $f(\xi)$, then the same element f is also specified by a function $\tilde{f}(\theta, \varphi)$ given by

$$f \to \tilde{f}(\theta, \varphi) = \tfrac{1}{2} e^{-ij_0\varphi} [\sin^2\theta/2]^{-1+i\rho} f(\xi) \qquad \xi = \cot\theta/2 \cdot e^{i\varphi}. \tag{2.5}$$

Using this definition of $\tilde{f}(\theta, \varphi)$, we can rewrite the expression for the scalar product

$$(f, h) = \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} d\varphi\, \overline{\tilde{f}(\theta, \varphi)} \tilde{h}(\theta, \varphi). \tag{2.6}$$

This is the usual rotation–invariant integration over the unit sphere.

The generators J_j, K_j can now be expressed as differential operators in θ and φ acting on the functions $\tilde{f}(\theta, \varphi)$:

$$\begin{aligned}
J_1 &= j_0 \frac{\cos\varphi}{\sin\theta} + i\sin\varphi\frac{\partial}{\partial\theta} + i\cos\varphi\cot\theta\frac{\partial}{\partial\varphi}, \\
J_2 &= j_0 \frac{\sin\varphi}{\sin\theta} - i\cos\varphi\frac{\partial}{\partial\theta} + i\sin\varphi\cot\theta\frac{\partial}{\partial\varphi}, \\
J_3 &= -i\frac{\partial}{\partial\varphi}, \\
K_1 &= -(\rho + i)\sin\theta\cos\varphi - j_0\cot\theta\sin\varphi \\
&\quad + i\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - i\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}, \\
K_2 &= -(\rho + i)\sin\theta\sin\varphi + j_0\cot\theta\cos\varphi \\
&\quad + i\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + i\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}, \\
K_3 &= -(\rho + i)\cos\theta - i\sin\theta\frac{\partial}{\partial\theta}.
\end{aligned} \tag{2.7}$$

The Casimir invariant of $SU(2)$ is the operator

$$J^2 = -\frac{\partial^2}{\partial\theta^2} - \cot\theta\frac{\partial}{\partial\theta} - \frac{1}{\sin^2\theta}\left\{\frac{\partial^2}{\partial\varphi^2} - 2ij_0\cos\theta\frac{\partial}{\partial\varphi} - j_0^2\right\}. \quad (2.8)$$

The simultaneous eigenfunctions of J^2 and J_3 are well-known from elementary quantum mechanics, and are the D_j functions of angular momentum theory.[10] We have

$$J^2 D_j^{m,j_0}(\varphi,\theta,0) = j(j+1)D_j^{m,j_0}(\varphi,\theta,0)$$
$$J_3 D_j^{m,j_0}(\varphi,\theta,0) = m D_j^{m,j_0}(\varphi,\theta,0); \quad j \geq j_0, \quad |m| \leq j. \quad (2.9)$$

[Recall that the D_j functions are functions of three Euler angles; here we have set the third Euler angle equal to zero. We need these functions only for integral values of j.] Standard properties of these functions assure us of their orthogonality and completeness relations. In terms of three Euler angles, we have

$$\frac{1}{8\pi^2}\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\theta\,d\theta \times\int_0^{2\pi}d\psi\,\overline{D_{j'}^{m_1'm_2'}(\varphi,\theta,\psi)}D_j^{m_1m_2}(\varphi,\theta,\psi) = (2j+1)^{-1}\delta_{j'j}\delta_{m_1'm_1}\delta_{m_2'm_2};$$
$$\sum_{j=0}^{\infty}\sum_{m_1m_2}\frac{2j+1}{8\pi^2}D_j^{m_1m_2}(\varphi,\theta,\psi)\overline{D_j^{m_1m_2}(\varphi',\theta',\psi')} = \delta(\varphi'-\varphi)\delta(\cos\theta'-\cos\theta)\delta(\psi'-\psi). \quad (2.10)$$

Factoring out the trivial dependence of $D_j^{mm'}$ on ψ, and eliminating this variable, one establishes the completeness and orthogonality of the functions $D_j^{m,j_0}(\varphi,\theta,0)$ on the unit sphere. [For each fixed j_0, we obtain a complete orthogonal system.] Apart from j-dependent phase factors, the states $|j,m\rangle$ of (1.12) correspond to functions of θ, φ as follows:

$$|j,m\rangle \to \left[\frac{2j+1}{4\pi}\right]^{\frac{1}{2}} D_j^{m,j_0}(\varphi,\theta,0). \quad (2.11)$$

One can now express the effect of the operator $U(g)$ directly on functions $\tilde{f}(\theta,\varphi)$. For this one has only to combine (2.2), (2.4), and (2.5). We omit this calculation for arbitrary g here, but examine particular cases of $U(g)$ in the sequel. When g belongs to the $SO(3)$ subgroup of $SO(3,1)$, however, the effect of $U(g)$ is particularly simple. Let us denote an element of $SO(3)$ by R, since it is a rotation of the space coordinates only. Then one has

$$[U(R)f]^{\sim}(\theta,\varphi) = \mu(R,\theta,\varphi)\tilde{f}[R^{-1}(\theta,\varphi)]. \quad (2.12)$$

Here μ is a phase factor of unit modulus and the argument of the function $\tilde{f}$ on the right-hand side is just the point on the unit sphere obtained by applying the rotation R^{-1} to the point (θ,φ). The factor μ is present only when j_0 is nonzero; its meaning can be understood in the following way. If one takes the generators in the form (2.7), divides K_j by $-\rho$ and takes the limit $\rho \to \infty$, one ends up with the generators for a single UIR of the Euclidean group $E(3)$. In this process of contraction, the generator J_j and the operator $U(R)$ are unchanged. The factor μ in (2.12) then corresponds to the *helicity* in the UIR of $E(3)$.[11]

The realization of the UIR $\{j_0,\rho\}$ of $SO(3,1)$ in a Hilbert space of functions on the unit sphere is analogous to the realization of the (continuous class) UIR's of $SO(2,1)$ in a space of functions on the unit circle. [See (I).] This form is suited to the diagonalization of the operators J^2, J_3 pertaining to the maximal compact subgroup of $SO(3,1)$. If we wish to reduce the UIR under the $SO(2,1)$ subgroup of $SO(3,1)$, then, as in (I), we must cast the UIR in a form suited to this purpose. This is achieved as follows: We make a change of variable which maps the surface of the unit sphere onto two planes (it would be more precise to call them hyperboloids), the upper hemisphere onto one plane, the lower hemisphere onto the other. We introduce polar coordinates r, ψ on each plane, defined in terms of θ, φ by

$$\tan\frac{\theta}{2} = \tanh\frac{r}{2}, \quad \varphi = \psi: \quad 0 \leq \theta \leq \pi/2, \quad 0 \leq \varphi \leq 2\pi;$$
$$\tan\frac{\theta}{2} = \coth\frac{r}{2}, \quad \varphi = 2\pi - \psi: \quad \pi/2 \leq \theta \leq \pi, \quad 0 \leq \varphi \leq 2\pi. \quad (2.13)$$

As θ varies from 0 to $\frac{1}{2}\pi$ (upper hemisphere), the radial variable r goes from 0 to ∞; thus the upper hemisphere gets mapped onto one plane. As θ varies over the range $\frac{1}{2}\pi$ to π (lower hemisphere), r goes from ∞ to 0; and the lower hemisphere is mapped onto an entire plane. A single function $\tilde{f}(\theta,\varphi)$ is now to be replaced by a pair of functions $f_\lambda(r,\psi)$, $\lambda = 1,2$, to represent the same abstract element f in Hilbert space. We define these functions by

$$f_1(r,\psi) = [\cosh r]^{-1+i\rho}\tilde{f}(\theta,\varphi): \quad 0 \leq \theta \leq \frac{\pi}{2}, \quad 0 \leq \varphi \leq 2\pi;$$
$$f_2(r,\psi) = [\cosh r]^{-1+i\rho}\tilde{f}(\theta,\varphi): \quad \frac{\pi}{2} \leq \theta \leq \pi, \quad 0 \leq \varphi \leq 2\pi. \quad (2.14)$$

[10] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1957), Chap. 4.

[11] Representations of the Euclidean Group $E(3)$ are discussed in W. Pauli, CERN Report 56-31 (1956).

Using these definitions, the scalar product of two element f, h in H can be expressed in terms of their wavefunctions as

$$(f, h) = \sum_{\lambda=1}^{2} \int_0^{2\pi} d\psi \int_0^{\infty} \sinh r \, dr \overline{f_\lambda(r, \psi)} h_\lambda(r, \psi). \quad (2.15)$$

Needless to say, in order to recover *all* elements f in H, we must pick all pairs of functions on the planes, $f_\lambda(r, \psi)$, each being square integrable over the plane and the functions being chosen independently of one another. Thus the Hilbert space H is expressed as a direct sum of two Hilbert spaces $H = H_1 \oplus H_2$, elements of H_1 corresponding to vectors f with $f_2(r, \psi) = 0$, and those of H_2 being vectors f with $f_1(r, \psi) = 0$.

The generators J_j, and K_j as they act on the functions $f_\lambda(r, \psi)$ can be written as 2×2 matrix operators because we can write the functions $f_\lambda(r, \psi)$ as a column vector with two rows. A straightforward calculation yields

$$J_3 = \frac{1}{i} \frac{\partial}{\partial \psi} \otimes \sigma_3,$$

$$K_1 = \left\{ -j_0 \frac{\sin \psi}{\sinh r} + i \cos \psi \frac{\partial}{\partial r} - i \sin \psi \coth r \frac{\partial}{\partial \psi} \right\} \otimes 1,$$

$$K_2 = \left\{ j_0 \frac{\cos \psi}{\sinh r} + i \sin \psi \frac{\partial}{\partial r} + i \cos \psi \coth r \frac{\partial}{\partial \psi} \right\} \otimes \sigma_3,$$

$$K_3 = \left\{ -(\rho + i) \cosh r - i \sinh r \frac{\partial}{\partial r} \right\} \otimes \sigma_3,$$

$$J_1 = \left\{ (\rho + i) \sin \psi \sinh r + j_0 \cos \psi \coth r + i \frac{\cos \psi}{\sinh r} \frac{\partial}{\partial \psi} + i \sin \psi \cosh r \frac{\partial}{\partial r} \right\} \otimes 1,$$

$$J_2 = \left\{ -(\rho + i) \cos \psi \sinh r + j_0 \sin \psi \coth r + i \frac{\sin \psi}{\sinh r} \frac{\partial}{\partial \psi} - i \cos \psi \cosh r \frac{\partial}{\partial r} \right\} \otimes \sigma_3. \quad (2.16)$$

These expressions for the generators deserve comment since they seem to lead to a paradox. At first glance, it seems as though the subspaces H_1 and H_2 are invariant under all the generators and hence under the entire group $SO(3, 1)$. However, this certainly cannot be so, since we are dealing with an irreducible representation of $SO(3, 1)$. The point is that the differential operators given in (2.16) have been derived in a formal manner, and do not suffice by themselves to define unique self-adjoint operators: they must be supplemented by the knowledge of the domains of these operators. A simpler way to understand the situation is to examine the unitary operators $U(g)$ for specially chosen elements g. Let us first investigate the case when $g \in SO(2, 1)$. If g corresponds to a matrix of the form

$$g \to e^{i\tau J_3} = \begin{pmatrix} e^{i\tau/2} & 0 \\ 0 & e^{-i\tau/2} \end{pmatrix}, \quad (2.17)$$

then the operator $U(g)$ does in fact leave the subspaces H_1 and H_2 invariant. Next, if we take

$$g \to e^{i\tau K_1} = \begin{pmatrix} \cosh \tau/2 & -\sinh \tau/2 \\ -\sinh \tau/2 & \cosh \tau/2 \end{pmatrix}, \quad (2.18)$$

we can check first that, in the complex ξ-plane, the interior and the exterior of the unit circle are each mapped onto themselves. Now, the region $|\xi| \leq 1$ corresponds to the second of the two planes with coordinates r, ψ; and the region $|\xi| \geq 1$ to the first plane. This is enough to show that the operators $U(\exp i\tau K_1)$ leave the subspaces H_1 and H_2 invariant. The same is true for $U(\exp i\tau K_2)$ because actually every element g in $SO(2, 1)$ corresponds to a product of suitable matrices of the forms (2.17) and (2.18). Thus the transformations of the subgroup $SO(2, 1)$, as well as its self-adjoint generators J_3, K_1, K_2, leave the subspaces H_1 and H_2 invariant.

On the other hand, consider an element of the type

$$g \to e^{i\tau K_3} = \begin{pmatrix} e^{-\tau/2} & 0 \\ 0 & e^{\tau/2} \end{pmatrix}. \quad (2.19)$$

Depending on whether τ is positive or negative, points outside the unit circle in the ξ plane move to the interior of the unit circle or the other way around. It follows that the operators $U(\exp i\tau K_3)$ do not leave H_1 and H_2 invariant. The same conclusion holds for the operators $U(\exp i\tau J_1)$ and $U(\exp i\tau J_2)$. We conclude that in reality the self-adjoint generators K_3, J_1 and J_2 do not leave H_1 and H_2 invariant. This is exactly analogous to the situation encountered in (*I*). The paradox is then explained by the statement that if a vector f is in the domain of one of the operators K_3, J_1, or J_2, then its components f_1 and f_2 in H_1 and H_2 are constrained and cannot be chosen completely independently. The behaviors of $f_1(r, \psi)$ and $f_2(r, \psi)$ at $r = \infty$ are correlated. This is but a reflection of the continuity and differentiability of the function $\hat{f}(\theta, \varphi)$ representing such a vector f.

Having satisfied ourselves that the transformations representing elements of the subgroup $SO(2, 1)$ do in fact leave H_1 and H_2 invariant, we can be sure that the same is true of the Casimir invariant Q of $SO(2, 1)$. We can evaluate Q as a differential operator, using (2.16). We have

$$Q = \left[-\frac{\partial^2}{\partial r^2} - \coth r \frac{\partial}{\partial r} - \frac{1}{\sinh^2 r} \times \left\{ \frac{\partial^2}{\partial \psi^2} - 2ij_0 \cosh r \frac{\partial}{\partial \psi} - j_0^2 \right\} \right] \otimes \mathbf{1}. \quad (2.20)$$

To reduce the UIR $\{j_0, \rho\}$ of $SO(3, 1)$ with respect to $SO(2, 1)$, we seek a complete set of eigenfunctions of Q and J_3. We can look for eigenvectors of the two types

$$\begin{pmatrix} f_1(r, \psi) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ f_2(r, \psi) \end{pmatrix} \tag{2.21}$$

belonging to H_1 and H_2, respectively. Eigenvectors of J_3 for an eigenvalue m are

$$\begin{pmatrix} f_1(r)e^{im\psi} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ f_2(r)e^{-im\psi} \end{pmatrix}. \tag{2.22}$$

For each integral value of m, positive, negative, or zero, we must find a complete orthonormal set of eigenfunctions of Q. As one would expect, these turn out to be generalizations of the $d_j^{mm'}$ functions of angular momentum theory, with complex j and m' set equal to j_0. For a given m, the functions $f_\lambda(r)$ in (2.22) are solutions of the following eigenvalue equations:

$$\left[-\frac{d}{dz}(z^2-1)\frac{d}{dz} + \frac{m^2 + j_0^2 - 2mj_0z}{z^2-1}\right] f_1(z) = q f_1(z),$$

$$\left[-\frac{d}{dz}(z^2-1)\frac{d}{dz} + \frac{m^2 + j_0^2 + 2mj_0z}{z^2-1}\right] f_2(z) = q f_2(z),$$

$$z = \cosh r; \quad 1 \le z < \infty. \tag{2.23}$$

The eigenvalue of Q is denoted by q. The general solutions to Eq. (2.23) are[12]

$$\left.\begin{aligned} f_1(z) &= d_j^{m,j_0}(z) \\ f_2(z) &= d_j^{-m,j_0}(z) \end{aligned}\right\} q = -j(j+1). \tag{2.24}$$

These functions are of course solutions to (2.23) for all complex j. The eigenfunctions of the self-adjoint operator Q must, however, correspond to real q, be normalizable at least in the continuum sense, and form a complete orthonormal system; furthermore, the possible values of q are already known from the knowledge of UIR's of $SO(2, 1)$. Solutions of Eq. (2.23) with all these properties are well known.[13] The completeness and orthonormality properties of the d_j functions depend on m, and complete sets of solutions are the following:

$m \ge 1$:

$$\int_0^\infty s \tanh \pi s \, ds \, d^{m,j_0}_{-\frac{1}{2}+is}(z) \, \overline{d^{m,j_0}_{-\frac{1}{2}+is}(z')} + \frac{1}{2} \sum_{j=0,1,\cdots}^{\min(m-1, j_0-1)} (2j+1) \, d_j^{m,j_0}(z) \, \overline{d_j^{m,j_0}(z')} = \delta(z - z'); \tag{2.25a}$$

$m \le 0$:

$$\int_0^\infty s \tanh \pi s \, ds \, d^{m,j_0}_{-\frac{1}{2}+is}(z) \, \overline{d^{m,j_0}_{-\frac{1}{2}+is}(z')} = \delta(z - z'). \tag{2.25b}$$

In Eq. (2.25a), the discrete summation exists if and only if $j_0 \ge 1$. Taking Eq. (2.25) in conjunction with Eqs. (2.21) and (2.22), we find a complete orthonormal set of eigenfunctions of Q and J_3 to be the following:

$$\left.\begin{aligned} \Psi^{(1)}_{s,m}(r, \psi) &= \frac{e^{im\psi}}{(2\pi)^{\frac{1}{2}}} d^{m,j_0}_{-\frac{1}{2}+is}(z)\begin{pmatrix}1\\0\end{pmatrix} \\ \Psi^{(2)}_{s,m}(r, \psi) &= \frac{e^{-im\psi}}{(2\pi)^{\frac{1}{2}}} d^{-m,j_0}_{-\frac{1}{2}+is}(z)\begin{pmatrix}0\\1\end{pmatrix} \end{aligned}\right\}, \quad 0 \le s < \infty, \quad -\infty < m < \infty$$

$$\left.\begin{aligned} \Psi^{(+)}_{k,m}(r, \psi) &= \left(\frac{2k-1}{4\pi}\right)^{\frac{1}{2}} e^{im\psi} d^{m,j_0}_{k-1}(z)\begin{pmatrix}1\\0\end{pmatrix}, \quad m = k, k+1, \cdots, \infty \\ \Psi^{(-)}_{k,m}(r, \psi) &= \left(\frac{2k-1}{4\pi}\right)^{\frac{1}{2}} e^{-im\psi} d^{-m,j_0}_{k-1}(z)\begin{pmatrix}0\\1\end{pmatrix}, \quad m = -k, -k-1, \cdots, -\infty \end{aligned}\right\} \quad k = 1, 2, \cdots, j_0,$$

$$Q\Psi^{(\lambda)}_{s,m} = (\tfrac{1}{4} + s^2)\Psi^{(\lambda)}_{s,m}; \quad J_3\Psi^{(\lambda)}_{s,m} = m\Psi^{(\lambda)}_{s,m},$$
$$Q\Psi^{(\pm)}_{k,m} = k(1-k)\Psi^{(\pm)}_{k,m}; \quad J_3\Psi^{(\pm)}_{k,m} = m\Psi^{(\pm)}_{k,m}. \tag{2.26}$$

The eigenfunctions $\Psi^{(1)}_{s,m}$ and $\Psi^{(+)}_{k,m}$ span the subspace H_1, and are thus orthogonal to the eigenfunctions $\Psi^{(2)}_{s,m}$ and $\Psi^{(-)}_{k,m}$ which span H_2. Beyond this, we have the following orthogonality relations [the scalar product is (2.15)]:

$$(\Psi^{(+)}_{k',m'}, \Psi^{(+)}_{k,m}) = (\Psi^{(-)}_{k',m'}, \Psi^{(-)}_{k,m}) = \delta_{k'k}\delta_{m'm};$$
$$(\Psi^{(+)}_{k',m'}, \Psi^{(1)}_{s,m}) = (\Psi^{(-)}_{k',m'}, \Psi^{(2)}_{s,m}) = 0; \tag{2.27}$$
$$(\Psi^{(1)}_{s',m'}, \Psi^{(1)}_{s,m}) = (\Psi^{(2)}_{s',m'}, \Psi^{(2)}_{s,m}) = \delta_{m'm}\frac{\delta(s'-s)}{s\tanh \pi s}.$$

The result of the above considerations is the following

Theorem: The UIR $\{j_0, \rho\}$ of the principal series of $SO(3, 1)$, for integral $j_0 \ge 0$, contains each UIR of $SO(2, 1)$ of the continuous class (integral type and nonexceptional) *twice*, and each UIR of $SO(2, 1)$ of the discrete classes $D_k^\pm$, for $k = 1, 2, \cdots, j_0$ *once*. [The UIR's of the discrete classes are absent if $j_0 = 0$.]

It need hardly be pointed out that the continuous class UIR's of $SO(2, 1)$ appear in the UIR of $SO(3, 1)$ in the sense of a *direct integral*, with the multiplicity factor two.[14] Among the wavefunctions written down in Eq. (2.26), only the discrete ones

$$\Psi^{(+)}_{k,m}, \quad \Psi^{(-)}_{k,m}$$

[12] Reference 10, p. 65.

[13] V. Bargmann, Ref. 2, p. 624. M. Andrews and J. Gunson, J. Math. Phys. 5, 1391 (1964).

[14] This is similar to the multiplicity of eigenvalues of the $O(1, 1)$ generator in the continuous class UIR's of $SO(2, 1)$. See V. Bargmann, Ref. 2 and (I).

represent normalizable vectors in the Hilbert space H, being eigenvectors for the discrete spectrum of Q. The remaining wavefunctions

$$\Psi^{(\lambda)}_{s,m}, \quad \lambda = 1, 2$$

are nonnormalizable vectors, being "ideal" eigenfunctions of Q corresponding to the continuous part of the spectrum of Q. A normalizable vector f in H may be expanded in terms of these basis vectors in the form

$$f = \sum_{\lambda=1}^{2} \sum_{m=-\infty}^{\infty} \int_0^\infty ds\, s \tanh \pi s f^{(\lambda)}_m(s) \Psi^{(\lambda)}_{s,m} + \sum_{k=1}^{j_0} \sum_{m=k}^{\infty} [f^{(+)}_{k,m} \Psi^{(+)}_{k,m} + f^{(-)}_{k,-m} \Psi^{(-)}_{k,-m}]. \quad (2.28)$$

Thus each vector f determines a "wavefunction" in the Q, J_3 basis to be

$$f \to \{f^{(\lambda)}_m(s), f^{(+)}_{k,m}, f^{(-)}_{k,m}\}. \quad (2.29)$$

And the scalar product of f with h is

$$(f, h) = \sum_{\lambda=1}^{2} \sum_{m=-\infty}^{\infty} \int_0^\infty ds\, s \tanh \pi s \overline{f^{(\lambda)}_m(s)} h^{(\lambda)}_m(s) + \sum_{k=1}^{j_0} \sum_{m=k}^{\infty} [\overline{f^{(+)}_{k,m}} h^{(+)}_{k,m} + \overline{f^{(-)}_{k,-m}} h^{(-)}_{k,-m}]. \quad (2.30)$$

Of course, the functions of s in Eq. (2.29) are relevant only up to sets of vanishing measure. The relation between the wavefunction (2.29) and the wavefunctions $f_\lambda(r, \psi)$ in (2.14) is given by

$$f^{(\lambda)}_m(s) = \int_0^{2\pi} d\psi \int_0^\infty \sinh r\, dr\, \overline{\Psi^{(\lambda)}_{s,m}(r, \psi)} f_\lambda(r, \psi), \quad \text{(no sum on } \lambda)$$

$$f^{(+)}_{k,m} = \int_0^{2\pi} d\psi \int_0^\infty \sinh r\, dr\, \overline{\Psi^{(+)}_{k,m}(r, \psi)} f_1(r, \psi),$$

$$f^{(-)}_{k,m} = \int_0^{2\pi} d\psi \int_0^\infty \sinh r\, dr\, \overline{\Psi^{(-)}_{k,m}(r, \psi)} f_2(r, \psi). \quad (2.31)$$

[We have dropped the column vector notation for wavefunctions in writing down Eq. (2.31).]

Q being the Casimir invariant of $SO(2, 1)$, the generators J_3, K_1, and K_2 do not alter the values of s and k when they act on $\Psi^{(\lambda)}_{s,m}$ and $\Psi^{(\pm)}_{k,m}$, respectively. The matrix elements of K_1 and K_2 in a UIR of $SO(2, 1)$ and in a basis of eigenvectors of J_3 may be written in a standard form. [See (I).] The eigenfunctions $\Psi^{(m)}_{s,\lambda}$, $\Psi^{(\pm)}_{k,m}$ of Q and J_3 will generally have to be modified by m-dependent phase factors in order that the differential operators for K_1 and K_2 given in Eq. (2.16) have the standard matrix elements between these vectors.

3. BEHAVIOR OF THE GENERATORS IN THE $SO(2, 1)$ BASIS

In this last section we examine the manner in which the generators of $SO(3, 1)$—other than those which generate the subgroup $SO(2, 1)$—connect different UIR's of $SO(2, 1)$ to one another. The vector nature of the generators K_j results in their having matrix elements connecting different UIR's of $SO(3)$ to one another in a basis of eigenvectors of J^2 and J_3, according to the selection rule

$$\Delta j = 0, \pm 1. \quad (3.1)$$

We seek an analog to this situation when Q and J_3 are diagonalized, as in Sec. II.

It is convenient to write the $SO(3, 1)$ generators in the form

$$F_0 = J_3, F_1 = K_1, F_2 = K_2; \quad G_0 = K_3, G_1 = -J_1, G_2 = -J_2. \quad (3.2)$$

Then the commutation relations of $SO(3, 1)$ are

$$[F_j, F_k] = iC^l_{jk} F_l; \quad (3.3a)$$

$$[F_j, G_k] = iC^l_{jk} G_l; \quad (3.3b)$$

$$[G_j, G_k] = -iC^l_{jk} F_l. \quad (3.3c)$$

The nonvanishing components of the tensor C^l_{jk} are

$$C^2_{01} = -C^2_{10} = C^1_{20} = -C^1_{02} = C^0_{21} = -C^0_{12} = +1. \quad (3.4)$$

Equation (3.3b) exhibits the transformation property of the tensor G_j under $SO(2, 1)$.

Since the matrix elements of the operators F_j within any UIR of $SO(2, 1)$, in a basis of eigenvectors of F_0, are known, and since one can pass from one component of G_j to another by commutation with the F_j, it is sufficient to know how G_0, say, connects the different UIR's of $SO(2, 1)$ appearing in one UIR of $SO(3, 1)$. It turns out that the essential features of the situation are already seen if, instead of G_j, one considers the tensor operator $G^{(0)}_j$ obtained by taking the limit, as $\rho \to \infty$ of $-(1/\rho)\, G_j$ [as given in Eq. (2.16)], while at the same time restricting oneself to the subspace H_1 in H:

$$G^{(0)}_0 = \cosh r, \quad G^{(0)}_1 = \sinh r \sin \psi, \quad G^{(0)}_2 = -\sinh r \cos \psi. \quad (3.5)$$

These obey

$$[F_j, G^{(0)}_k] = iC^l_{jk} G^{(0)}_l, \quad [G^{(0)}_j, G^{(0)}_k] = 0. \quad (3.6)$$

F_j and $G^{(0)}_j$ generate a single UIR of the Poincaré group in two space and one time dimension, with

timelike three-momenta.[15] The subspace H_1 of H is in fact invariant under $G_j^{(0)}$; and, as we have seen, the operators F_j acting on H_1 contain each UIR of $SO(2,1)$ of the continuous integral nonexceptional class *once* and each UIR of the discrete class $D_k^{(+)}$ *once*, for $k = 1, 2, \cdots, j_0$.

We consider then the action of $G_0^{(0)}$ on H_1. A vector f in H_1 may be specified either by the wavefunction $f(r, \psi)$, or by the wavefunction

$$f_{k,m}^{(+)}, \quad f_m^{(1)}(s), \tag{3.7}$$

the connection between these two specifications being given by Eqs. (2.28) and (2.31) [suitably restricted to H_1]. It is clearly sufficient to consider vectors f which are eigenvectors of J_3:

$$f(r, \psi) = \frac{e^{im\psi}}{(2\pi)^{\frac{1}{2}}} f(z), \quad z = \cosh r. \tag{3.8}$$

For definiteness, we will assume $m \geq 1$; this is when the UIR's $D_k^{(+)}$ appear. Other cases for m can be handled in a way similar to what follows. Then the expansion theorem for f is

$$\begin{aligned} f &= \int_0^\infty ds\, s \tanh \pi s f^{(1)}(s) \Psi_{s,m}^{(1)} + \sum_{k=1,2,\cdots}^{\min(m,j_0)} f_k^{(+)} \Psi_{k,m}^{(+)}, \\ f^{(1)}(s) &= \int_1^\infty dz\, d_{-\frac{1}{2}+is}^{m,j_0}(z) f(z), \\ f_k^{(+)} &= (k - \tfrac{1}{2})^{\frac{1}{2}} \int_1^\infty dz\, d_{k-1}^{m,j_0}(z) f(z). \end{aligned} \tag{3.9}$$

[The complex conjugation of the d functions has been omitted, since they are real for $1 \leq z < \infty$].[16] We begin by proving that the basis vectors of the UIR's $D_k^{(+)}$ for $k \geq 2$ are in the domain of $G_0^{(0)}$. The functions $d_j^{m,j_0}(z)$ can be written as a linear combination of two other functions of z, one of which goes like z^j, the other like z^{-j-1} as $z \to \infty$.[17] In this sense, the asymptotic behavior of $d_j^{m,j_0}(z)$ is exhibited by the formula

$$d_j^{m,j_0}(z) \sim \frac{B}{2j+1} \sin(\pi j) \quad z^j + \frac{A}{\cos \pi j} z^{-j-1} \tag{3.10}$$

The coefficients A and B are analytic functions of j, and, at least in the region $-\frac{1}{2} \leq \mathrm{Re}\, j < m, j_0$, are finite. Equation (3.10) shows why the states $\Psi_{k,m}^{(+)}$ $(k = 1, 2, \cdots)$ are normalizable: The potentially dangerous first term is absent for $j = 0, 1, \cdots, \min(m-1, j_0 - 1)$. [More precisely, when $d_j^{m,j_0}(z)$ is written as a linear combination as mentioned above, the first function has a vanishing coefficient when $j = 0, 1, \cdots, \min(m-1, j_0 - 1)$; and the remaining term goes like z^{-j-1}.] Beyond this, since the effect of $G_0^{(0)}$ is just to multiply a wavefunction $f(z)$ by z, the states $G_0^{(0)}\Psi_{k,m}^{(+)}$ are also normalizable *for* $k \geq 2$. [The relation between k and j is $k = j + 1$.] The vector

$$G_0^{(0)}\Psi_{k,m}^{(+)}, \quad k \geq 2$$

can be expanded in terms of the $\Psi_{k,m}^{(+)}$ by using the identity:

$$\begin{aligned} z\, d_j^{m,j_0}(z) = &-[(m^2 - (j+1)^2)(j_0^2 - (j+1)^2)]^{\frac{1}{2}} \\ &\times \frac{d_{j+1}^{m,j_0}(z)}{(j+1)(2j+1)} + \frac{j_0 m}{j(j+1)} d_j^{m,j_0}(z) \\ &- [(m^2 - j^2)(j_0^2 - j^2)]^{\frac{1}{2}} \frac{d_{j-1}^{m,j_0}(z)}{j(2j+1)}, \\ &(j \neq 0, -\tfrac{1}{2}, -1). \end{aligned} \tag{3.11}$$

[See the Appendix for the derivation of (3.11).] Using this for $k \geq 2$, $(j \geq 1)$, we get

$$\begin{aligned} &G_0^{(0)}\Psi_{k,m}^{(+)} \\ &= -\frac{1}{k}\left[\frac{(m^2 - k^2)(j_0^2 - k^2)}{(2k+1)(2k-1)}\right]^{\frac{1}{2}} \Psi_{k+1,m}^{(+)} + \frac{j_0 m}{k(k-1)} \Psi_{k,m}^{(+)} \\ &\quad - \frac{1}{k-1}\left[\frac{(m^2 - (k-1)^2)(j_0^2 - (k-1)^2)}{(2k-1)(2k-3)}\right]^{\frac{1}{2}} \Psi_{k-1,m}^{(+)}. \end{aligned} \tag{3.12}$$

Thus $G_0^{(0)}$ [and so also $G_j^{(0)}$] can act on the discrete states $\Psi_{km}^{(+)}$, *for* $k \geq 2$, and yields linear combinations of the states $\Psi_{k\pm1,m}^{(+)}$ and $\Psi_{k,m}^{(+)}$. We see in Eq. (3.12) the selection rule $\Delta k = \pm 1, 0$, exactly like the selection rule (3.1).

In contrast to the UIR's $D_k^{(+)}$, $k \geq 2$, Eq. (3.10) shows that the normalizable vectors $\Psi_{1,m}^{(+)}$ belonging to the UIR $D_1^{(+)}$ *do not lie in the domain of* $G_0^{(0)}$, i.e., $G_0^{(0)}\Psi_{1,m}^{(+)}$ is not normalizable. Thus there is a sharp difference between the first of the discrete UIR's, $D_1^{(+)}$, and the succeeding ones, $D_2^{(+)}, D_3^{(+)}, \cdots$. The latter are in the domain of $G_j^{(0)}$; the former is not.

To understand the role of the UIR $D_1^{(+)}$, we then argue as follows: Suppose that a vector f is in the domain of $G_0^{(0)}$, and also that

$$(\Psi_{1,m}^{(+)}, f) \neq 0;$$

then what can be said about f? It is irrelevant whether

[15] The following point is worth noting. In Sec. 2, we saw that a single UIR of $SO(3,1)$ gives rise on contraction to a single UIR of the Euclidean group $E(3)$. However, if we wish to contract to UIR's of the Poincaré group in 3 dimensions [pseudo-Euclidean group $E(2,1)$], then we find that on contraction a single UIR of $SO(3,1)$ yields *two* UIR's of $E(2,1)$. Both involve timelike three-momenta; however, one has a positive time component of momentum and the other has a negative time component. A similar situation is encountered in higher dimensions. Thus a single UIR of $SO(4,1)$ can be contracted to yield two UIR's of the Poincaré group $E(3,1)$ both involving timelike momenta. Conversely, if one wants to do the reverse of this process of contraction and if one wants to find self-adjoint generators for the group $SO(4,1)$ in the enveloping algebra of the Lie algebra of the Poincaré group $E(3,1)$, one must start with *two* timelike UIR's of $E(3,1)$ corresponding to the two signs of the energy. Otherwise it will be impossible to integrate to a representation of the group $SO(4,1)$. We have already encountered such a situation in the discussion following Eq. (2.16).

[16] V. Bargmann, Ref. 2, p. 626. Bargmann's $V_{mn}(\zeta)$ is our $d_j^{j,j_0}(z)$.

[17] M. Andrews and J. Gunson, Ref. 13, p. 1392.

the coefficients $(\Psi^{(+)}_{k,m}, f)$ for $k \geq 2$ are zero or non-zero, and we can assume that they vanish, without taking f outside the domain of $G_0^{(0)}$. Then we conclude that the part of the wavefunction of f relating to the continuous class UIR's of $SO(2, 1)$, namely $f^{(1)}(s)$—*cannot vanish identically*. For if it did, then f would be proportional to $\Psi^{(+)}_{1,m}$ and could not lie in the domain of $G_0^{(0)}$. The same argument carried a stage further shows that $f_1^{(+)}$, the components of f in the UIR $D_1^{(+)}$, cannot be chosen independently of the function $f^{(1)}(s)$ [assuming that $G_0^{(0)}f$ exists]. We can even prove explicitly that $f^{(1)}(s)$ completely determines $f_1^{(+)}$. The two conditions on $f(z)$ are

$$\|f\|^2 = \int_1^\infty |f(z)|^2\, dz < \infty,$$

$$\|G_0^{(0)}f\|^2 = \int_1^\infty z^2\, |f(z)|^2\, dz < \infty; \qquad (3.13)$$

$f^{(1)}(s)$ and $f_1^{(+)}$ are given by

$$f^{(1)}(s) = \int_1^\infty dz\; d^{m,j_0}_{-\frac{1}{2}+is}(z) f(z),$$

$$f_1^{(+)} = \frac{1}{\sqrt{2}} \int_1^\infty dz\; d_0^{m,j_0}(z) f(z). \qquad (3.14)$$

Taking into account the asymptotic behavior of $d_j^{m,j_0}(z)$ as given by Eq. (3.10), and using Eq. (3.13), we see that the function $f^{(1)}(s)$ *can be continued off the real axis as an analytic function of the complex variable* s, *at least up to Im* $s = \pm 1$. Further, since such an analytic continuation is unique, and since the point $s = -i/2$ always lies within the minimum region of analyticity, $f_1^{(+)}$ is determined by $f^{(1)}(s)$:

$$f_1^{(+)} = \frac{1}{\sqrt{2}} f^{(1)}\left(-\frac{i}{2}\right). \qquad (3.15)$$

Thus we obtain a remarkable result: If a vector f, with $J_3 f = mf$, lies in the domain of $G_0^{(0)}$ and has a wavefunction

$$f \to f^{(1)}(s),\quad f_k^{(+)},\quad k = 1, 2, \cdots, \min(m, j_0),$$

then

(i) $f^{(1)}(s)$ is the boundary value on the real axis of an analytic function of s, analytic at least up to Im $s = \pm 1$;

(ii) $f_1^{(+)}$, the component of f in the subspace containing the UIR $D_1^{(+)}$, is determined by $f^{(1)}(s)$ via (3.15);

(iii) the component $f_2^{(+)}, f_3^{(+)}, \cdots$ are arbitrary, corresponding to the fact that the vectors $\Psi^{(+)}_{2,m}, \Psi^{(+)}_{3,m}, \cdots$ lie in the domain of $G_0^{(0)}$. Clearly, these results hold for all components of $G_j^{(0)}$ and they show again the distinction between the UIR $D_1^{(+)}$ and the remaining discrete UIR's. It should be pointed out that what we have proved is that, if a vector f is in the domain of $G_0^{(0)}$, then $f_1^{(+)}$ is determined by $f^{(1)}(s)$; we do not claim that for all such vectors f, the quantity $f_1^{(+)}$ is non-zero. Whether $f_1^{(+)}$ is zero or not depends on the particular function $f^{(1)}(s)$.

It remains to express the wavefunction of $h = G_0^{(0)}f$ in terms of that of f. We need Eq. (3.11) and its analog for $j = 0$. This is derived in the Appendix, and is

$$z\, d_0^{m,j_0} = -[(m^2-1)(j_0^2-1)]^{\frac{1}{2}} d_1^{m,j_0}(z) + m j_0\, d_0^{m,j_0}(z) + 2m j_0 \left[\frac{\partial}{\partial j} d_j^{m,j_0}(z)\right]_{j=0}. \qquad (3.16)$$

Using Eqs. (3.11), (3.12) and (3.16) we can express $h_m^{(1)}(s)$, $h_{k,m}^{(+)}$ in terms of $f_m^{(1)}(s), f_{k,m}^{(+)}$ as follows:

$$h = G_0^{(0)} f;$$

$$h_m^{(1)}(s) = [(m^2 - (\tfrac{1}{2}+is)^2)(j_0^2 - (\tfrac{1}{2}+is)^2)]^{\frac{1}{2}} \times \frac{f_m^{(1)}(s-i)}{(s-i)(2s-i)} - \frac{m j_0}{s^2+\frac{1}{4}} f_m^{(1)}(s) + [(m^2 - (\tfrac{1}{2}-is)^2)(j_0^2 - (\tfrac{1}{2}-is)^2)]^{\frac{1}{2}} \times \frac{f_m^{(1)}(s+i)}{(s+i)(2s+i)},$$

$$h_{1,m}^{(+)} = -\frac{1}{\sqrt{3}}[(m^2-1)(j_0^2-1)]^{\frac{1}{2}} f_{2,m}^{(+)} + m j_0 f_{1,m}^{(+)} - i\sqrt{2}\, m j_0 \left[\frac{d}{ds} f_m^{(1)}(s)\right]_{s=-\frac{i}{2}},$$

$$h_{k,m}^{(+)} = -\frac{1}{k}\left[\frac{(m^2-k^2)(j_0^2-k^2)}{(2k-1)(2k+1)}\right]^{\frac{1}{2}} f_{k+1,m}^{(+)} + \frac{j_0 m}{k(k-1)} f_{k,m}^{(+)} - \frac{1}{(k-1)} \times \left[\frac{(m^2-(k-1)^2)(j_0^2-(k-1)^2)}{(2k-1)(2k-3)}\right]^{\frac{1}{2}} f_{k-1,m}^{(+)}; \quad k \geq 2. \qquad (3.17)$$

Equations similar to the above can be derived in case $m \leq 0$. The behavior of $G_0^{(0)}$ in H_1 is typical of all the operators $G_j^{(0)}$, the only new possibility being that the eigenvalue of J_3 may change by ± 1.

Returning to the case of interest to us, namely, the operators G_j of $SO(3, 1)$, there are two new points. First of all, the expressions (2.16) for the G_j differ from the expressions (3.5) by the presence of additional terms involving the differential operator $\partial/\partial z$. The second point, which is related to the first, is that the G_j operates in the Hilbert space H which is the direct sum of the two orthogonal subspaces H_1 and H_2; this is essential in order that G_j be self-adjoint.[18] In spite of these two new features, the basic mechanism by

[18] Compare with Ref. 15.

which the UIR's of $SO(2, 1)$ appearing in a given UIR of $SO(3, 1)$ are connected to one another by the G_j is the one we have explained above. The states of the UIR's $D_k^{(\pm)}$ for $k \geq 2$ lie in the domain of the G_j; the matrix elements of G_j among these vectors obey the $\Delta k = 0, \pm 1$ selection rule. The states of the two UIR'S $D_1^{(\pm)}$ are not in the domain of G_j. A linear combination f of the ideal vectors $\Psi_{s,m}^{(\lambda)}$ with wavefunction $f_m^{(\lambda)}(s)$ can be acted upon by G_j only if $f_m^{(\lambda)}(s)$ is the boundary value of an analytic function of s, analytic at least up to Im $s = \pm 1$. For such a vector f, the components $f_{1,m}^{(\pm)}$ in the UIR's $D_1^{(\pm)}$ are already fixed by the functions $f_m^{(\lambda)}(s)$. [Of course, the special role enjoyed by the first discrete UIR $D_1^{(+)}$ and its close connection to the UIR's of the continuous class is now extended to the UIR $D_1^{(-)}$ as well, when we consider the generators of $SO(3, 1)$ rather than $E(2, 1)$.] One has, in fact,

$$f_{1,m}^{(+)} = \frac{1}{\sqrt{2}} f_m^{(1)}\left(-\frac{i}{2}\right); \quad f_{1,m}^{(-)} = \frac{1}{\sqrt{2}} f_m^{(2)}\left(-\frac{i}{2}\right). \tag{3.18}$$

Actually, in making these statements, one has to be careful and pay attention to the question of the domains of the operators G_j, in the sense that if f is in the domain of G_j, then its components f_1 and f_2 in H_1 and H_2 are constrained and cannot be chosen completely independently of one another. These constraints are just a reflection of the continuity and differentiability of the wavefunction $\vec{f}(\theta, \varphi)$ on the unit sphere associated with the vector f. We conclude by writing down the analog to Eq. (3.17) for the operator G_0 [generator of $SO(3, 1)$]:

$$h = G_0 f;$$

$$\begin{aligned} h_m^{(\lambda)}(s) = &-[(m^2 - (\tfrac{1}{2} + is)^2)(j_0^2 - (\tfrac{1}{2} + is)^2)]^{\frac{1}{2}} \\ &\times \frac{\left(\rho + s - \frac{i}{2}\right)}{(s - i)(2s - i)} f_m^{(\lambda)}(s - i) + \frac{\rho m j_0}{s^2 + \frac{1}{4}} f_m^{(\lambda)}(s) \\ &- [(m^2 - (\tfrac{1}{2} - is)^2)(j_0^2 - (\tfrac{1}{2} - is)^2)]^{\frac{1}{2}} \\ &\times \frac{\left(\rho + s + \frac{i}{2}\right)}{(s + i)(2s + i)} f_m^{(\lambda)}(s + i), \quad (\lambda = 1, 2), \end{aligned}$$

$$\begin{aligned} h_{1,m}^{(\pm)} = &\frac{1}{\sqrt{3}} [(m^2 - 1)(j_0^2 - 1)]^{\frac{1}{2}} (\rho - i) f_{2,m}^{(\pm)} \\ &- m j_0 (\rho + i) f_{1,m}^{(\pm)} \\ &+ i\sqrt{2}\, m j_0 \rho \left[\frac{d}{ds} f_m^{(1,2)}(s)\right]_{s=-\frac{i}{2}}; \end{aligned}$$

$$\begin{aligned} h_{k,m}^{(\pm)} = &\frac{1}{k}\left[\frac{(m^2 - k^2)(j_0^2 - k^2)}{(2k - 1)(2k + 1)}\right]^{\frac{1}{2}} (\rho - ik) f_{k+1,m}^{(\pm)} \\ &- \frac{\rho m j_0}{k(k - 1)} f_{k,m}^{(\pm)} + \frac{1}{k - 1} \\ &\times \left[\frac{(m^2 - (k - 1)^2)(j_0^2 - (k - 1)^2)}{(2k - 1)(2k - 3)}\right]^{\frac{1}{2}} \\ &\times (\rho - i(k - 1)) f_{k-1,m}^{(\pm)}. \quad (k \geq 2) \end{aligned} \tag{3.19}$$

In the second equation above, expressing $h_{1,m}^{(\pm)}$ in terms of f, the plus sign goes with $\lambda = 1$ in the last term, and the minus sign with $\lambda = 2$.

4. CONCLUSION

We have carried out the reduction of the UIR's $\{j_0, \rho\}$ of the Principal series of $SO(3, 1)$ for integral j_0 with respect to the noncompact subgroup $SO(2, 1)$. We have found that, in this reduction, every UIR of $SO(2, 1)$ of the continuous class (integral type and nonexceptional) appears *twice*; and each of the UIR's of the discrete classes $D_k^{(+)}$ and $D_k^{(-)}$ appear *once*, for $k = 1, 2, \cdots, j_0$. These latter UIR's are absent if $j_0 = 0$.

We have examined how the generators G_j of $SO(3, 1)$ connect different UIR's of $SO(2, 1)$ to one another. The most striking result is that the basis states of the UIR's $D_k^{(\pm)}$, for $k \geq 2$, lie in the domain of G_j while the basis states of the two UIR's $D_1^{(\pm)}$ do not. Further, only those linear combinations of the ideal vectors $\Psi_{s,m}^{(\lambda)}$ [basis states of the UIR's of the continuous class] can lie in the domain of G_j, for which the wavefunctions $f_m^{(\lambda)}(s)$ are boundary values of analytic functions of s. For such vectors, the components of the vector in the space of the UIR's $D_1^{(\pm)}$ are determined by their components in the spaces of the UIR's of the continuous class. In a sense, then, the two UIR's $D_1^{(\pm)}$ act as bridges between the other discrete UIR's of $SO(2, 1)$ on the one hand, and the UIR's of the continuous class on the other.

ACKNOWLEDGMENTS

The author would like to thank Professor J. G. Kuriyan, Professor L. O'Raifeartaigh, and Professor E. C. G. Sudarshan for many enlightening discussions.

Most of this work was done while the author was at the Palmer Physical Laboratory of Princeton University, during the academic year 1965–1966. He would like to thank Professor W. Bleakney and Professor M. L. Goldberger for the hospitality of that Institution.

APPENDIX

We prove Eqs. (3.11) and (3.16) of the text. The d functions are defined in terms of the hypergeometric

functions by[19]

$$d_j^{m,j_0}(z) = \left[\frac{(j+m)!(j-j_0)!}{(j-m)!(j+j_0)!}\right]^{\frac{1}{2}} F_j,$$

$$F_j = \left(\frac{1+z}{2}\right)^{(m+j_0)/2}\left(\frac{1-z}{2}\right)^{(m-j_0)/2} \times \frac{F(-j+m, j+m+1; 1+m-j_0; \frac{1}{2}(1-z))}{(m-j_0)!}. \quad \text{(A1)}$$

[As given here, Eq. (A1) is valid for $m \geq j_0$, $m + j_0 \geq 0$; however, in the ensuing manipulations, m and j_0 do not change.]

We use Eqs. (31), (32), and (33) of Sec. 2.8 of Bateman.[20] These are, respectively,

$$zF(a,b;c;z) = \frac{c-2a}{b-a}F(a,b;c;z) - \frac{c-a}{b-a}F(a-1,b;c;z) + \frac{a}{b-a}(1-z)F(a+1,b;c;z), \quad \text{(A2)}$$

$$F(a,b;c;z) = \frac{b}{b-a}F(a,b+1;c;z) + \frac{a}{a-b}F(a+1,b;c;z), \quad \text{(A3)}$$

$$(1-z)F(a,b;c;z) = \frac{c-b}{a-1}F(a-1,b-1;c;z) + \frac{a+b-c-1}{a-1}F(a-1,b;c;z). \quad \text{(A4)}$$

Use Eq. (A4) in Eq. (A2) to get

$$zF(a,b;c;z) = F(a,b;c;z) + \frac{c-b}{b-a}F(a,b-1;c;z) + \frac{c-a}{a-b}F(a-1,b;c;z). \quad \text{(A5)}$$

Therefore,

$$(1-2z)F(a,b;c;z) = -F(a,b;c;z) + 2\frac{b-c}{b-a}F(a,b-1;c;z) + 2\frac{a-c}{a-b}F(a-1,b;c;z). \quad \text{(A6)}$$

Using Eq. (A3) twice in Eq. (A6) we get

$$\begin{aligned}(1-2z)F(a,b;c;z) &= -F(a,b;c;z) + 2\frac{b-c}{b-a}\frac{1}{b-a-1} \\ &\quad \times [(b-1)F(a,b;c;z) - aF(a+1,b-1;c;z) \\ &\quad + 2\frac{a-c}{a-b}\frac{1}{b-a+1} \\ &\quad \times [bF(a-1,b+1;c;z) - (a-1)F(a,b;c;z)].\end{aligned} \quad \text{(A7)}$$

We identify

$$a = -j+m, \quad b = j+m+1, \quad c = 1+m-j_0 \quad \text{(A8)}$$

to get

$$\begin{aligned}(1-2z)F(a,b;c;z) &= \frac{(j+m+1)(j-j_0+1)}{(j+1)(2j+1)}F(a-1,b+1;c;z) \\ &\quad + \frac{(j-m)(j+j_0)}{j(2j+1)}F(a+1,b-1;c;z) \\ &\quad + \frac{mj_0}{j(j+1)}F(a,b;c;z).\end{aligned}$$

Written in terms of F_j, this is

$$zF_j = \frac{(j+m+1)(j+1-j_0)}{(j+1)(2j+1)}F_{j+1} + \frac{mj_0}{j(j+1)}F_j + \frac{(j-m)(j+j_0)}{j(2j+1)}F_{j-1}, \quad j \neq 0, -\tfrac{1}{2}, -1. \quad \text{(A9)}$$

Equation (A9) in conjunction with Eq. (A1) leads to Eq. (3.11) in the text.

For the special case $j = 0$, we must start with Eq. (A.9) and take the limit as j approaches zero. This gives

$$zF_0 = -(j_0-1)(m+1)F_1 + [(j_0-1)(m+1)+1]F_0 + 2mj_0\left(\frac{\partial F_j}{\partial j}\right)_{j=0}. \quad \text{(A10)}$$

The derivative of F_j with respect to j can be related to that of d_j^{m,j_0} by a routine calculation. We find that

$$\frac{\partial}{\partial j}d_j^{m,j_0}(z)\bigg|_{j=0} = \frac{j_0-m}{2mj_0}d_0^{m,j_0}(z) + \phi_0\frac{\partial F_j}{\partial j}\bigg|_{j=0}, \quad \text{(A11)}$$

where ϕ_0 is the ratio $d_j^{m,j_0}/F_j$ evaluated at $j = 0$. Equations (A11), (A10), and (A1) lead to Eq. (3.16) of the text.

[19] M. Andrews and J. Gunson, Ref. 13, p. 1392.

[20] *Bateman Manuscript Project* (McGraw-Hill Book Co., Inc., New York, 1953), Vol. 1, p. 101.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

n-Particle Kinematics

ERIC H. ROFFMAN*
Courant Institute of Mathematical Sciences, New York, New York
and
The Institute for Advanced Study, Princeton, New Jersey

(Received 26 April 1967)

We suggest a very simple method for handling relativistic n-body kinematics. In fact the method works well for studying the n-fold Kronecker products of the representations of many groups, including all compact Lie groups as a simple case. The method is explicitly covariant, and treats all particles symmetrically.

In this article we suggest a very simple method for handling relativistic n-body kinematics.[1] In fact, the method works well for studying the n-fold Kronecker products of the representations of many groups, including all compact Lie groups as a simple case. The method is explicitly covariant, and treats all particles symmetrically.

The method is based on the observation that under right translation by g, $\mathcal{R}(g)$, on the space G^n:

$$\mathcal{R}(g): \quad G^n \to G^n,$$
$$g_1, \cdots, g_k \to g_1 g, g_2 g, \cdots, g_n g, \tag{1}$$

variables of the form $G_{ij} = g_i g_j^{-1}$ remain invariant; there are exactly $n-1$ independent variables of this type; the remaining nth independent variable can be chosen to be any one of the g_i, and will transform under right translation; and finally, the conversion between any one such set of variables and any other is very simple, because $G_{ik}G_{kj} = G_{ij}$ and $G_{ij}g_j = g_i$.

A general description of the method follows. We then apply it to the n-fold Kronecker product of representations of an arbitrary compact group, and after to the relativistic n-body problem. The method may be applied to other groups with varying degrees of success, depending on the structure of the group (e.g., the ease with which a function on the group can be expanded in terms of matrix elements of irreducible representations). However, we do not do so here.

Before describing our method, let us review some of the other approaches to n-body kinematics. The earliest method[2] was to couple the first two particles together, then add the third particle, and continue adding each particle to the system consisting of the previous particles. The parameters which specify the n-particle state are the quantum numbers of the subsystems. This method does not preserve a symmetry between the particles. It is complicated to transform between any one such system and any other.

The method of Werle[3] is close, in spirit, to the method we use. His procedure is to project out the total angular momentum in the center of mass system and observe that the coordinates, in a body fixed coordinate system, of the momenta of the individual particles are now invariant parameters which can be used to specify the state.

Our procedure retains all the virtues of the method of Werle, while being explicitly covariant and rather general. It is also quite versatile; for example, the other approaches to n-body kinematics can be obtained as variations of the basic method we describe.

A characteristic feature of our method is that we have realized the linear spaces on which our representations act, in terms of a space of functions whose domain is the group itself. This technique is found to be very useful, for example, in the study of the S matrix by the methods of harmonic analysis. We plan to give examples of how this method can be used in a separate article.

To translate the language of the space of functions defined on the group into the more common notation of an abstract vector space, it is only necessary to note the fact that a matrix element $D^J_{\alpha\beta}(g)$ of a compact group G may be considered as either a function on the group, or as an abstract object with indices $^J_{\alpha\beta}$ which label it as a basis of the usual (abstract) vector space. Equations (10) and (12) have the same interpretation for the Poincaré group.

Suppose that we are given a set of functions F, from a group G to the complex plane $\mathbb{C}$, $F = \{f: G \to \mathbb{C}\}$, which is a linear space and for which

$$f(gg_0) \in F \quad \text{iff} \quad f(g) \in F,$$

i.e., F is invariant under right translation of the

* Present address: State University of New York at Stony Brook.

[1] The kinematics of relativistic n-particle states have been discussed by those mentioned in Refs. 2 and 3.

[2] H. Joos, Fortschr. Physik **10**, 65 (1962); G. C. Wick, Ann. Phys. (N.Y.) **18**, 65 (1962); A. J. Macfarlane, Rev. Mod. Phys. **34**, 41 (1962).

[3] J. Werle, Phys. Rev. Letters **4**, 127 (1963); J. Werle, Nucl. Phys. **44**, 579, 637 (1963); S. M. Berman and N. Jacob, Phys. Rev. **139**, B1023 (1965) and others. A review with additional references is given in J. M. Lévy-Leblond, "Global and Democratic Methods for Classifying N-Particle Systems," University of Rochester Report UR-875-136.

domain. Then the map

$$T: F \times G \to F$$

defined by

$$T(g_0)f(g) = f(gg_0)$$

is a representation of G.

The tensor product of n such representations on $F_1, F_2, \cdots, F_n$ is defined, loosely speaking, on the tensor-product space

$$F_{1,\cdots,n} = F_1 \otimes \cdots \otimes F_n = \{f: G \otimes \cdots -n- \cdots \otimes G \to \mathbb{C}\},$$

where $f(g_1, \cdots, g_n)$ is a linear combination of functions of the form $f_1(g_1), f_2(g_2), \cdots, f_n(g_n)$, where $f_i \in F_i$. The n-fold Kronecker product representation is given on this space by

$$T(g)f_{1,\cdots,n}(g_1, \cdots, g_n) = f_{1,\cdots,n}(g_1 g, g_2 g, \cdots, g_n g). \tag{2}$$

We now define a new function $k: G^{n+1} \to \mathbb{C}$ by

$$k_{1,\cdots,n}(\gamma; \gamma_1, \cdots, \gamma_n) = f_{1,\cdots,n}(\gamma_1\gamma, \gamma_2\gamma, \cdots, \gamma_n\gamma) \tag{3a}$$

with the transformation law

$$T(g)k_{1,\cdots,n}(\gamma; \gamma_1, \cdots, \gamma_n) = k_{1,\cdots,n}(\gamma g; \gamma_1, \cdots, \gamma_n). \tag{3b}$$

We have

$$g_i = \gamma_i \gamma. \tag{4}$$

Note that if we set $\gamma_j = 1$ then $\gamma = g_j$ and $\gamma_i = g_i g_j^{-1}$. Thus, these variables correspond to those which we discussed in the introduction, except that we have maintained for the moment a symmetrical appearance by retaining a superfluous coordinate. This is a convenience, not an essential part of the method.

The function $k(\gamma, \gamma_1, \cdots, \gamma_n)$ may now be expanded, if one wishes (whenever the structure of the group allows this), in terms of matrix elements of irreducible representations of the variables γ, γ_1, $\gamma_2, \cdots, \gamma_n$. The expansion of k, as a function of γ into irreducible representation spaces, tells how it transforms under the n-fold Kronecker product representation. The behavior of k as a function of γ_j tells how it behaves as "particle" [4] j is removed from the other particles. (If γ_j has been set to 1, $\gamma_j = 1$, then $\gamma_i = g_i g_j^{-1}$ is the variable which describes the distance between particles i and j. If we have still retained the superfluous variable, then γ_j is a measure of the distance of particle j from the "central point"—all $\gamma_i = 1$.)

As one may guess, the transformation from a basis in the variables g to the basis in the variables γ, is performed rather easily, since g_i is just the product $\gamma_i\gamma$. There is only one nontrivial step.

Let us turn to the case of an arbitrary compact Lie group. Let $D^J_{\alpha\beta}(g)$ denote the matrix element of the irreducible representation J between the basis elements α and β of the space of that representation. It is well known, from the theory of harmonic analysis on compact groups, that, with proper normalization, we have

$$\int D^{*J}_{\alpha\beta}(g) D^{J'}_{\alpha'\beta'}(g)\, dg = \delta_{JJ'}\delta_{\alpha\alpha'}\delta_{\beta\beta'} \equiv \delta\begin{pmatrix} J & \alpha & \beta \\ J' & \alpha' & \beta' \end{pmatrix} \tag{5}$$

(* denotes complex conjugate),

where dg is the Haar measure on the group.

We denote by $|^J_{\alpha\beta}\rangle$ the function $D^J_{\alpha\beta}(g)$.

Then the Kronecker product basis may be written

$$\left| \begin{matrix} J_1 \\ \alpha_1\beta_1 \end{matrix}, \cdots, \begin{matrix} J_n \\ \alpha_n\beta_n \end{matrix} \right\rangle = D^{J_1}_{\alpha_1\beta_1}(g_1) \cdots D^{J_n}_{\alpha_n\beta_n}(g_n) \tag{6a}$$

which transforms according to

$$\left| \begin{matrix} J_1 \\ \alpha_1\beta_1 \end{matrix}, \cdots, \begin{matrix} J_n \\ \alpha_n\beta_n \end{matrix} \right\rangle \to \left| \begin{matrix} J_1 \\ \alpha_1\beta_1 \end{matrix}, \cdots, \begin{matrix} J_n \\ \alpha_n\beta_n \end{matrix} \right\rangle' = \sum_{\beta_1', \cdots, \beta_n'} D^{J_1}_{\beta_1'\beta_1}(g) \cdots D^{J_n}_{\beta_n'\beta_n}(g) \left| \begin{matrix} J_1 \\ a_1\beta_1' \end{matrix}, \cdots, \begin{matrix} J_n \\ \alpha_n\beta_n' \end{matrix} \right\rangle. \tag{6b}$$

Let us compute the transition amplitude into the state

$$\left| \begin{matrix} \kappa \\ ab \end{matrix}, \begin{matrix} \kappa_1 \\ a_1b_1 \end{matrix}, \cdots, \begin{matrix} \kappa_n \\ a_nb_n \end{matrix} \right\rangle \quad D^{\kappa}_{ab}(\gamma)\, D^{\kappa_1}_{a_1b_1}(\gamma_1) \cdots D^{\kappa_n}_{a_nb_n}(\gamma_n), \tag{7a}$$

which transforms according to

$$\left| \begin{matrix} \kappa \\ ab \end{matrix}, \begin{matrix} \kappa_1 \\ a_1b_1 \end{matrix}, \cdots, \begin{matrix} \kappa_n \\ a_nb_n \end{matrix} \right\rangle' = \sum_{b'} D^{\kappa}_{b'b}(g) \left| \begin{matrix} \kappa \\ ab' \end{matrix}, \begin{matrix} \kappa_1 \\ a_1b_1 \end{matrix}, \cdots, \begin{matrix} \kappa_n \\ a_nb_n \end{matrix} \right\rangle. \tag{7b}$$

$$\left\langle \begin{matrix} \kappa \\ ab' \end{matrix}, \begin{matrix} \kappa_1 \\ a_1b_1 \end{matrix}, \cdots, \begin{matrix} \kappa_n \\ a_nb_n \end{matrix} \middle| \begin{matrix} J_1 \\ \alpha_1\beta_1 \end{matrix}, \cdots, \begin{matrix} J_n \\ \alpha_n\beta_n \end{matrix} \right\rangle$$

$$= \int d\gamma\, d\gamma_1 \cdots d\gamma_n D^{*\kappa}_{ab}(\gamma)\, D^{*\kappa_1}_{a_1b_1}(\gamma_1) \cdots D^{*\kappa_n}_{a_nb_n}(\gamma_n);$$

$$\sum_{\beta_1', \cdots, \beta_n'} D^{J_n}_{\alpha_1\beta_1'}(\gamma_1) D^{J_1}_{\beta_1'\beta_1}(\gamma) \cdots D^{J_n}_{\alpha_n\beta_n'}(\gamma_n) D^{J_n}_{\beta_n'\beta_n}(\gamma)$$

$$= \int d\gamma\, \delta^{\kappa_1 a_1}_{J_1\alpha_1} \cdots \delta^{\kappa_n a_n}_{J_n\alpha_n} D^{*\kappa}_{ab}(\gamma) D^{J_1}_{b_1\beta_n}(\gamma) \cdots D^{J_n}_{b_n\beta_n}(\gamma). \tag{8}$$

The only nontrivial part of the calculation is the remaining integral. It depends, of course, on the structure of the group, and we do not compute it

[4] "Particle" is occasionally used in the paper to denote a function which transforms under a representation of some group. The physical particles are usually thought to be described by a wavefunction which transforms as a state of an irreducible representation of the Poincaré group, and our terminology is a convenient generalization.

explicitly. Note that it could, for example, be computed by using the Clebsch–Gordan coefficients to expand the product of two such D functions into a single one and iterating this procedure. Since we have already accounted for all the internal parameters necessary to specify this as an n-particle state, this iteration procedure is independent of the order in which it is performed.

We write

$$\begin{pmatrix} \kappa \\ ab \end{pmatrix} \begin{matrix} J_1, \cdots, J_n \\ b_1\beta_1, \cdots, b_n\beta_n \end{matrix} = \int d\gamma D^{*\kappa}_{ab}(\gamma) D^{J_1}_{b_1\beta_1}(\gamma_1) \cdots D_{b_n\beta_n}^{J_n}(\gamma_n).$$

So,

$$\left\langle \begin{matrix} \kappa & \kappa_1 & \cdots & \kappa_n \\ ab & a_1b_1 & \cdots & a_nb_n \end{matrix} \middle| \begin{matrix} J_1 & \cdots & J_n \\ \alpha_1\beta_1 & \cdots & \alpha_n\beta_n \end{matrix} \right\rangle = \delta^{\kappa_1,\cdots,\kappa_n}_{J_1,\cdots,J_n} \delta^{a_1,\cdots,a_n}_{\alpha_1,\cdots,\alpha_n} \begin{pmatrix} \kappa \\ ab \end{pmatrix} \begin{matrix} J_1, \cdots, J_n \\ b_1\beta_1, \cdots, b_n\beta_n \end{matrix}. \quad (9)$$

To treat representations of the Poincaré group by this method, we must think in terms of functions on the Poincaré group. The idea of using functions on the Poincaré group has been given by Lurçat.[5] Nghiem Xuan Hai and Lurçat[6] have defined and studied generalized functions on the Poincaré group which are analogous to the $D^J_{ab}(g)$ for the compact Lie groups. It is not difficult to work with these functions, and we shall see, by using them, that we avoid the usual complications of n-particle kinematics. The calculation is heuristic.

Define

$$f_{QP}{}^{J}_{rs}(aA) = e^{iQ_\mu a^\mu} \omega \delta^3(H_P A^{-1} H_Q^{-1} m - m) \times D^J_{rs}(H_Q A H_P^{-1}), \quad (10)$$

where H_P^{-1} is the standard boost from rest to momentum P; $\omega\delta^3$ is a covariant δ function which says that A^{-1} takes Q into P. Therefore, $H_Q A H_P^{-1}$ is a rotation and $D^J_{rs}(H_Q A H_P^{-1})$ is the corresponding rotation matrix.

Note that $f_{QP}{}^{J}_{rs}(aA)$ is essentially what one would expect for the matrix element of the Poincaré group between a state of momentum P, spin J, projection s, and a state of momentum Q, spin J, projection r.

These generalized matrix elements have the easily verified property

$$\sum_{s'} \int dQ' f_{QQ'}{}^{J}_{rs'}(aA) f_{Q'P}{}^{J}_{s's}(bB) = f_{QP}{}^{J}_{rs}[(aA)\cdot(bB)], \quad (11)$$

which is analogous to the usual matrix multiplication law. Note also that

$$f_{MP}{}^{J}_{om}(aA) \equiv |P\ J\ m\rangle \quad (12)$$

is the usual Dirac ket of momentum P, spin J, projection m.

Let us now use this to compute the reduction of the n-fold Kronecker product of representations (which is the same thing as computing the kinematics of the relativistic n-particle problem).

Define

$$|R_1 T^{J_1}_{1a_1\beta_1}, \cdots, R_n T^{J_n}_{na_n\beta_n}\rangle = f_{R_1 T_1}{}^{J_n}_{a_n\beta_n}(g_1) \cdots f_{R_n T_n}{}^{J_n}_{a_n\beta_n}(g_n), \quad (13a)$$

where the g_i are elements of the Poincaré group.

This transforms according to the n-fold Kronecker product representation. Define a state of the new basis

$$|QP^{\kappa}_{ab}, Q_1 P^{\kappa_1}_{1a_1b_1}, \cdots, Q_n P^{\kappa_n}_{na_nb_n}\rangle = f_{QP^{\kappa}_{ab}}(\gamma) f_{Q_1 P_1{}^{\kappa_1}_{a_1b_1}}(\gamma_1) \cdots f_{Q_n P_n{}^{\kappa_n}_{a_nb_n}}(\gamma_n). \quad (13b)$$

Then we have

$$\begin{aligned} &\langle QP^{\kappa}_{ab}, Q_i P^{\kappa_i}_{ia_ib_i}\ i = 1, \cdots, n \mid R_i T^{J_i}_{ia_i\beta_i}\ i = 1, \cdots, n\rangle \\ &= \sum_{\beta_i'} \int (\Pi_i\, d(a_iA_i)\, dS_i)\, d(bB) \\ &\quad \times e^{-iQb} \omega \delta(H_P B^{-1} H_Q^{-1} m - m) D^{*\kappa}_{ab}(H_Q B H_P^{-1}) \\ &\prod_{i=1}^{n} e^{-iQ_i a_i} \omega_i \delta(H_{P_i} A_i^{-1} H_{Q_i}^{-1} m - m) D^{*\kappa_i}_{a_ib_i}(H_{Q_i} A_i H_{P_i}^{-1}) \\ &\prod_{i=1}^{n} e^{iR_i a_i} \omega_i \delta(H_{S_i} A_i^{-1} H_{R_i}^{-1} m - m) D^{J_i}_{\alpha_i\beta_i'}(H_{R_i} A_i H_{S_i}^{-1}) \\ &\prod_{i=1}^{n} e^{iS_i b} \omega_i \delta(H_{T_i} B^{-1} H_{S_i}^{-1} m - m) D^{J_i}_{\beta_i'\beta_i}(H_{S_i} B H_{T_i}^{-1}) \\ &\quad = \delta(Q - \textstyle\sum P_i)\delta(P - \sum T_i)\Pi_i[\delta(Q_i - R_i)\delta^{\kappa_i a_i}_{J_i\alpha_i}] \\ &\int dB \omega \delta(H_P B^{-1} H_Q^{-1} m - m) D^{\kappa}_{ab}(H_Q B H_P^{-1}) \\ &\quad \Pi_i \omega_i \delta(H_{T_i} B^{-1} H_{P_i}^{-1} m - m) \Pi_i D^{J_i}_{b_i\beta_i}(H_{P_i} B^{-1} H_{T_i}). \end{aligned} \quad (14)$$

We do not perform the remaining integral, which is the same integral that normally occurs in the computation of the projection of the spin content of the n-particle state. Whether the integral is calculated directly or by an iteration procedure, the result is the same.

In the usual kinematics of an n-particle system, this reduction of the total spin content must also be performed, but it is only one of the many problems, and not at all the worst. In the method presented here, it is the only problem.

The internal parameters, those invariant under the

[5] F. Lurçat, Phys. 1, 95 (1964).

[6] Equations (10) and (11) and the decomposition into irreducible representations of two such functions are results due to F. Lurçat and Nghiem Xuan Hai. I am extremely grateful to them for showing their work to me in advance of publication.

Kronecker product representation, are handled by this method symmetrically, and simply. In the usual iteration procedure, or its variants,[1] the meaning of the internal coordinates is obfuscated, as well as complicated.

It is hoped that this method may be useful whenever Kronecker products of representations are needed. Many variations on this method, for particular purposes, are possible. For example, it may be useful to express the S matrix of an n-particle to m-particle process as a function of all the G_{ij}, i and j running over all n and m incoming and outgoing indices.

In this method we have embedded a function of G^n into the space of functions on $G_0 \otimes G_1^n$ (where G_0 and G_1 are isomorphic to G). Together with the group of permutations, G_1^n describes all the internal structure of the n-fold Kronecker product. We define G_1^n to be the "internal group" of the system. It would be natural to study the dependence of the dynamics of interacting particles, resonances, and unstable particles on the subgroups and factor spaces of the permutation and internal groups of the particles which produce them, and into which they decay.

ACKNOWLEDGMENTS

The author would like to thank F. Lurçat for his hospitality at the Laboratiore de Physique Théorique et Hautes Energies (Orsay) where part of this work was done with support from the Centre National de la Recherche Scientifique. The author is most grateful to F. Lurçat and Nghiem Xuan Hai for many discussions while he was there.[6] Most of this paper was written at the Courant Institute under National Science Foundation Grant NSF-GP-3465. At the Institute for Advanced Study, research was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR No. 42-65.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Weyl Transformation and the Classical Limit of Quantum Mechanics

BORIS LEAF

Department of Physics, State University of New York, College at Cortland, Cortland, New York

(Received 27 April 1967)

The Weyl correspondence for obtaining quantum operators from functions of classical coordinates and momenta is known to be incorrect. To calculate quantum-mechanical expectation values as phase-space averages with the Wigner density function, one cannot use classical functions but must use Weyl transforms. These transforms are defined and their properties derived from quantum mechanics. Their properties are expressed in terms of a Hermitian operator $\Delta(Q, K)$ whose Weyl transform is a δ function. The Wigner function is the transform of the density operator. Every Weyl transform is exhibited as a difference of two functions which are nonnegative on the phase space. Weyl transforms do not obey the algebra of classical functions. In the classical limit $\hbar \to 0$, Weyl transforms become classical functions, the Wigner function becomes nonnegative throughout the phase space, and the Hilbert space is spanned by an orthonormal set of vectors which are simultaneous eigenkets of the commuting coordinate and momentum operators.

1. INTRODUCTION

It was suggested by Weyl[1] in 1928 that an operator of quantum mechanics $A(\mathbf{q}, \mathbf{p})$ can be obtained from a classical dynamical function $\mathcal{A}_c(\mathbf{Q}, \mathbf{P})$ by expressing the latter as a Fourier integral

$$\mathcal{A}_c(\mathbf{Q}, \mathbf{P}) = \int \cdots \int d\mathbf{u}\, d\mathbf{v} \alpha(\mathbf{u}, \mathbf{v}) \times \exp\left[2\pi i(\mathbf{u} \cdot \mathbf{Q} + \mathbf{v} \cdot \mathbf{P})\right],$$

and then replacing in the integral the classical Cartesian coordinate and momentum variables $\mathbf{Q}$, $\mathbf{P}$ by the corresponding quantum operators $\mathbf{q}$, $\mathbf{p}$. This so-called Weyl correspondence has generated considerable interest,[2,3] particularly because of its apparently close relationship to the Wigner formalism of quantum mechanics. This formalism[4,5] is similar to classical

[1] H. Weyl, *Gruppentheorie und Quantenmechanik* (S. Hirzel Verlag, Leipzig, 1928).

[2] D. Massignon, *Mécanique Statistique des Fluides* (Dunod Cie., Paris, 1957).

[3] J. Vlieger, P. Mazur, S. R. DeGroot, Physica **27**, 353 (1961); **27**, 957 (1961).

[4] E. Wigner, Phys. Rev. **40**, 749 (1932).

[5] H. Mori, I. Oppenheim, and J. Ross, in *Studies in Statistical Mechanics*, J. De Boer and G. E. Uhlenbeck, Eds. (North-Holland Publishing Company, Amsterdam, 1962), Vol. I.

Kronecker product representation, are handled by this method symmetrically, and simply. In the usual iteration procedure, or its variants,[1] the meaning of the internal coordinates is obfuscated, as well as complicated.

It is hoped that this method may be useful whenever Kronecker products of representations are needed. Many variations on this method, for particular purposes, are possible. For example, it may be useful to express the S matrix of an n-particle to m-particle process as a function of all the G_{ij}, i and j running over all n and m incoming and outgoing indices.

In this method we have embedded a function of G^n into the space of functions on $G_0 \otimes G_1^n$ (where G_0 and G_1 are isomorphic to G). Together with the group of permutations, G_1^n describes all the internal structure of the n-fold Kronecker product. We define G_1^n to be the "internal group" of the system. It would be natural to study the dependence of the dynamics of interacting particles, resonances, and unstable particles on the subgroups and factor spaces of the permutation and internal groups of the particles which produce them, and into which they decay.

ACKNOWLEDGMENTS

The author would like to thank F. Lurçat for his hospitality at the Laboratiore de Physique Théorique et Hautes Energies (Orsay) where part of this work was done with support from the Centre National de la Recherche Scientifique. The author is most grateful to F. Lurçat and Nghiem Xuan Hai for many discussions while he was there.[6] Most of this paper was written at the Courant Institute under National Science Foundation Grant NSF-GP-3465. At the Institute for Advanced Study, research was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR No. 42-65.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Weyl Transformation and the Classical Limit of Quantum Mechanics

BORIS LEAF

Department of Physics, State University of New York, College at Cortland, Cortland, New York

(Received 27 April 1967)

The Weyl correspondence for obtaining quantum operators from functions of classical coordinates and momenta is known to be incorrect. To calculate quantum-mechanical expectation values as phase-space averages with the Wigner density function, one cannot use classical functions but must use Weyl transforms. These transforms are defined and their properties derived from quantum mechanics. Their properties are expressed in terms of a Hermitian operator $\Delta(Q, K)$ whose Weyl transform is a δ function. The Wigner function is the transform of the density operator. Every Weyl transform is exhibited as a difference of two functions which are nonnegative on the phase space. Weyl transforms do not obey the algebra of classical functions. In the classical limit $\hbar \to 0$, Weyl transforms become classical functions, the Wigner function becomes nonnegative throughout the phase space, and the Hilbert space is spanned by an orthonormal set of vectors which are simultaneous eigenkets of the commuting coordinate and momentum operators.

1. INTRODUCTION

It was suggested by Weyl[1] in 1928 that an operator of quantum mechanics $A(\mathbf{q}, \mathbf{p})$ can be obtained from a classical dynamical function $\mathcal{A}_c(\mathbf{Q}, \mathbf{P})$ by expressing the latter as a Fourier integral

$$\mathcal{A}_c(\mathbf{Q}, \mathbf{P}) = \int \cdots \int d\mathbf{u}\, d\mathbf{v} \alpha(\mathbf{u}, \mathbf{v}) \times \exp\left[2\pi i(\mathbf{u} \cdot \mathbf{Q} + \mathbf{v} \cdot \mathbf{P})\right],$$

and then replacing in the integral the classical Cartesian coordinate and momentum variables $\mathbf{Q}$, $\mathbf{P}$ by the corresponding quantum operators $\mathbf{q}$, $\mathbf{p}$. This so-called Weyl correspondence has generated considerable interest,[2,3] particularly because of its apparently close relationship to the Wigner formalism of quantum mechanics. This formalism[4,5] is similar to classical

[1] H. Weyl, *Gruppentheorie und Quantenmechanik* (S. Hirzel Verlag, Leipzig, 1928).

[2] D. Massignon, *Mécanique Statistique des Fluides* (Dunod Cie., Paris, 1957).

[3] J. Vlieger, P. Mazur, S. R. DeGroot, Physica **27**, 353 (1961); **27**, 957 (1961).

[4] E. Wigner, Phys. Rev. **40**, 749 (1932).

[5] H. Mori, I. Oppenheim, and J. Ross, in *Studies in Statistical Mechanics*, J. De Boer and G. E. Uhlenbeck, Eds. (North-Holland Publishing Company, Amsterdam, 1962), Vol. I.

mechanics in that expectation values of observables are computed as phase-space averages, with the Wigner function playing the role of the probability density distribution function. However, as is known, the Weyl correspondence is incorrect[6]; it does not produce the operators of quantum mechanics from classical functions. Quantum-mechanical expectation values cannot, in general, be calculated as phase-space averages of *classical* functions with the Wigner function giving the density of probability. Efforts have been made to replace the Wigner function with improved distribution functions.[7,8] Instead, in this paper the Wigner function is retained, but the classical functions are replaced by special functions in the phase space; these are the Weyl transforms [defined in Eq. (2.8) below]. Weyl transforms are indeed related to quantum operators by a transformation which is the Weyl correspondence, but they do not obey the commutative algebra of classical functions. The Wigner function is the Weyl transform of the quantum-mechanical density operator. The functions of classical dynamics are obtained from the Weyl transforms in the classical limit $\hbar \to 0$. Weyl transforms of operators which are functions of the coordinates only or functions of the momenta only are the same as their classical limits.

It is the purpose of this paper to derive from quantum mechanics the transformation which expresses quantum operators in terms of their Weyl transforms and the inverse transformation giving the transforms in terms of the operators,[9] and to develop the properties of the Weyl transforms. The transformation is characterized by a Hermitian operator $\Delta(\mathbf{Q}, \mathbf{K})$ (where $\mathbf{P} = 2\pi\hbar\mathbf{K}$). In Sec. II, the Weyl transforms and the operator $\Delta(\mathbf{Q}, \mathbf{K})$ are defined. $\Delta(\mathbf{Q}, \mathbf{K})$ is expressed in various alternative forms which lead to different expressions for the Weyl transform $A_w(\mathbf{Q}, \mathbf{K})$ of the quantum operator A. A is Hermitian if and only if $A_w(\mathbf{Q}, \mathbf{K})$ is real. One form of the Weyl transformation shows that it is the Weyl transform and not the classical function which is related to the quantum operator by the "Weyl correspondence" [Eqs. (2.22), (2.23)] thereby proving the incorrectness of the original statement of the Weyl correspondence. The matrix elements of $\Delta(\mathbf{Q}, \mathbf{K})$ are obtained, leading to various expressions for the inverse transformation. It is shown that $\Delta(\mathbf{Q}, \mathbf{K})$ can be expressed as the difference of positive-definite Hermitian operators; similarly, the Weyl transforms (including the Wigner function) appear as the difference of nonnegative terms. In Sec. 3, the algebraic relations for multiplication of Weyl transforms are obtained. The Weyl transform for the product of two operators AB is not the same as for the product BA. The trace of the product AB is the phase-space integral of the product of the corresponding Weyl transforms $A_w(\mathbf{Q}, \mathbf{K})$ and $B_w(\mathbf{Q}, \mathbf{K})$. In Sec. 4 the properties of the Wigner function are developed. It is the Weyl transform of the density operator.[10] Its general form for a mixed quantum state is given as the difference of nonnegative terms. As an application of the Weyl transformation, the time dependence of the Wigner function is obtained from the von Neumann equation for the density operator; the result agrees with a formulation by Prigogine.[11] In Sec. 5 the classical limit is considered. It is shown that in this limit the quantum operators operate in a Hilbert space spanned by orthonormal vectors $|\mathbf{Q}, \mathbf{K}\rangle$ which are simultaneous eigenkets of the commuting operators for coördinate and momentum. The Weyl transforms become classical functions obeying commutative multiplication. The Wigner function becomes nonnegative throughout the phase space, and its time dependence is governed by the classical Liouville equation. The Hilbert space formulation has been used in the solution of the Liouville equation for classical systems by Leaf and Schieve[12] to discuss the approach to equilibrium.

2. WEYL TRANSFORMATION

Consider the Hilbert space of a quantum-mechanical system of N degrees of freedom. If $\mathbf{q} = \{q_1, q_2, \cdots, q_N\}$ is the Cartesian coordinate operator and $\mathbf{p} = 2\pi\hbar\mathbf{k} = 2\pi\hbar\{k_1, \cdots, k_N\}$ is the conjugate momentum operator, the space is spanned by the eigenkets $|\mathbf{Q}\rangle$ of $\mathbf{q}$ or $|\mathbf{K}\rangle$ of $\mathbf{k}$. The completeness conditions are

$$\int d\mathbf{Q}\, |\mathbf{Q}\rangle\langle\mathbf{Q}| = \mathbf{1}, \quad \int d\mathbf{K}\, |\mathbf{K}\rangle\langle\mathbf{K}| = \mathbf{1}. \tag{2.1}$$

The scalar product $\langle \mathbf{Q} \mid \mathbf{K} \rangle$ is

$$\langle \mathbf{Q} \mid \mathbf{K} \rangle = \exp(2\pi i \mathbf{Q} \cdot \mathbf{K}). \tag{2.2}$$

The commutation relations hold that

$$q_i k_j - k_j q_i = (i/2\pi)\delta_{ij}. \tag{2.3}$$

(The notation used avoids normalization factors which depend upon N. Also, the element of phase

[6] J. R. Shewell, Am. J. Phys. **27**, 16 (1959).
[7] H. Margenau and R. N. Hill, Prog. Theoret. Phys. (Kyoto) **26**, 722 (1961).
[8] L. Cohen, J. Math. Phys. **7**, 781 (1966).
[9] A statement of the Weyl transform and its inverse was given by H. J. Groenewold, Physica **12**, 405 (1946), especially p. 450.
[10] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).
[11] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962), p. 260.
[12] B. Leaf and W. C. Schieve, Physica (to be published).

space $d\mathbf{Q}\,d\mathbf{K}$ is independent of the system of units.)

Consider the identity for an operator A of the Hilbert space:

$$A = \int\cdots\int d\mathbf{Q}'\,d\mathbf{Q}''\,d\mathbf{K}'\,d\mathbf{K}''\,|\mathbf{Q}''\rangle\langle\mathbf{Q}''\,|\,\mathbf{K}''\rangle \times \langle\mathbf{K}''|\,A\,|\mathbf{K}'\rangle\langle\mathbf{K}'\,|\,\mathbf{Q}'\rangle\langle\mathbf{Q}'|.$$

With the transformation, whose Jacobian is unity,

$$\mathbf{Q}'' = \mathbf{Q} + \tfrac{1}{2}\mathbf{v}, \quad \mathbf{Q}' = \mathbf{Q} - \tfrac{1}{2}\mathbf{v},$$
$$\mathbf{K}'' = \mathbf{K} + \tfrac{1}{2}\mathbf{u}, \quad \mathbf{K}' = \mathbf{K} - \tfrac{1}{2}\mathbf{u},$$

the identity becomes

$$A = \int\cdots\int d\mathbf{Q}\,d\mathbf{K}\,d\mathbf{u}\,d\mathbf{v}\,|\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\rangle \times \langle\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\,|\,\mathbf{K} + \tfrac{1}{2}\mathbf{u}\rangle\langle\mathbf{K} + \tfrac{1}{2}\mathbf{u}|\,A\,|\mathbf{K} - \tfrac{1}{2}\mathbf{u}\rangle \times \langle\mathbf{K} - \tfrac{1}{2}\mathbf{u}\,|\,\mathbf{Q} - \tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q} - \tfrac{1}{2}\mathbf{v}|. \tag{2.4}$$

But, according to (2.2),

$$\langle\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\,|\,\mathbf{K} + \tfrac{1}{2}\mathbf{u}\rangle\langle\mathbf{K} - \tfrac{1}{2}\mathbf{u}\,|\,\mathbf{Q} - \tfrac{1}{2}\mathbf{v}\rangle = \exp\,[2\pi i(\mathbf{u}\cdot\mathbf{Q} + \mathbf{v}\cdot\mathbf{K})]. \tag{2.5}$$

Therefore, if the operator $\Delta(\mathbf{Q}, \mathbf{K})$ is defined as

$$\Delta(\mathbf{Q}, \mathbf{K}) = \int d\mathbf{v}\,|\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q} - \tfrac{1}{2}\mathbf{v}|\,\exp\,(2\pi i\mathbf{v}\cdot\mathbf{K}), \tag{2.6}$$

the identity is equivalent to the transformation

$$A = \int\cdots\int d\mathbf{Q}\,d\mathbf{K}A_w(\mathbf{Q}, \mathbf{K})\Delta(\mathbf{Q}, \mathbf{K}), \tag{2.7}$$

with the inverse transformation

$$A_w(\mathbf{Q}, \mathbf{K}) = \int d\mathbf{u}\,\langle\mathbf{K} + \tfrac{1}{2}\mathbf{u}|\,A\,|\mathbf{K} - \tfrac{1}{2}\mathbf{u}\rangle\,\exp\,(2\pi i\mathbf{u}\cdot\mathbf{Q}). \tag{2.8}$$

The function $A_w(\mathbf{Q}, \mathbf{K})$ will be called the Weyl transform of the quantum operator A.

The Weyl transformation is characterized by the properties of the operator $\Delta(\mathbf{Q}, \mathbf{K})$. Various alternative expressions for it can be obtained. Since

$$\exp\,(-2\pi i\mathbf{v}\cdot\mathbf{k})\,|\mathbf{Q}\rangle = |\mathbf{Q} + \mathbf{v}\rangle, \tag{2.9}$$

therefore

$$\Delta(\mathbf{Q}, \mathbf{K}) = \int d\mathbf{v}\,\exp\,[-2\pi i\mathbf{v}\cdot(\mathbf{k} - \mathbf{K})] \times |\mathbf{Q} - \tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q} - \tfrac{1}{2}\mathbf{v}|. \tag{2.10}$$

But the projection operator $|\mathbf{Q}\rangle\langle\mathbf{Q}|$ is the same as the δ function operator

$$|\mathbf{Q}\rangle\langle\mathbf{Q}| = \delta(\mathbf{q} - \mathbf{Q}) \equiv \int d\mathbf{u}\,\exp\,[2\pi i\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q})], \tag{2.11}$$

since, for any ket $|\mathbf{Q}'\rangle$ in the complete set of eigenkets of $\mathbf{q}$,

$$|\mathbf{Q}\rangle\langle\mathbf{Q}\,|\,\mathbf{Q}'\rangle = \delta(\mathbf{Q}' - \mathbf{Q})\,|\mathbf{Q}'\rangle = \delta(\mathbf{q} - \mathbf{Q})\,|\mathbf{Q}'\rangle. \tag{2.12}$$

Therefore (2.10) becomes

$$\Delta(\mathbf{Q}, \mathbf{K}) = \int\cdots\int d\mathbf{u}\,d\mathbf{v}\,\exp\,(-i\pi\mathbf{u}\cdot\mathbf{v}) \times \exp\,[2\pi i\mathbf{v}\cdot(\mathbf{k} - \mathbf{K})]\exp\,[2\pi i\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q})]. \tag{2.13}$$

But for any two operators A and B whose commutator is a constant,[13]

$$\exp\,(A + B) = \exp A\cdot\exp B\cdot\exp\,[-\tfrac{1}{2}(AB - BA)]. \tag{2.14}$$

Therefore, according to (2.3),

$$\Delta(\mathbf{Q}, \mathbf{K}) = \int\cdots\int d\mathbf{u}\,d\mathbf{v}\,\exp\,\{2\pi i[\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q}) + \mathbf{v}\cdot(\mathbf{k} - \mathbf{K})]\} \tag{2.15}$$

$$= \int\cdots\int d\mathbf{u}\,d\mathbf{v}\,\exp\,(i\pi\mathbf{u}\cdot\mathbf{v})\,\exp\,[2\pi i\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q})] \times \exp\,[2\pi i\mathbf{v}\cdot(\mathbf{k} - \mathbf{K})]. \tag{2.16}$$

Since $\mathbf{q}$ and $\mathbf{k}$ are Hermitian, it follows from (2.15) that $\Delta(\mathbf{Q}, \mathbf{K})$ is Hermitian. From (2.7) and (2.8), the operator A is Hermitian if and only if the Weyl transform $A_w(\mathbf{Q}, \mathbf{K})$ is real. Equations (2.13) and (2.16) can also be written

$$\Delta(\mathbf{Q}, \mathbf{K}) = \exp\,[(i/4\pi)(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K})] \times \delta(\mathbf{k} - \mathbf{K})\delta(\mathbf{q} - \mathbf{Q}),$$
$$= \exp\,[(-i/4\pi)(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K})] \times \delta(\mathbf{q} - \mathbf{Q})\delta(\mathbf{k} - \mathbf{K}). \tag{2.17}$$

Since

$$\exp\,(2\pi i\mathbf{u}\cdot\mathbf{q})\,|\mathbf{K}\rangle = |\mathbf{K} + \mathbf{u}\rangle, \tag{2.18}$$

therefore, from (2.16), another form of $\Delta(\mathbf{Q}, \mathbf{K})$ can be obtained, namely,

$$\Delta(\mathbf{Q}, \mathbf{K}) = \int d\mathbf{u}\,\exp\,[2\pi i\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q})] \times \int d\mathbf{v}\,\exp\,[2\pi i\mathbf{v}\cdot(\mathbf{k} - \mathbf{K} + \tfrac{1}{2}\mathbf{u})]$$
$$= \int d\mathbf{u}\,\exp\,[2\pi i\mathbf{u}\cdot(\mathbf{q} - \mathbf{Q})]\,|\mathbf{K} - \tfrac{1}{2}\mathbf{u}\rangle\langle\mathbf{K} - \tfrac{1}{2}\mathbf{u}|$$
$$= \int d\mathbf{u}\,|\mathbf{K} + \tfrac{1}{2}\mathbf{u}\rangle\langle\mathbf{K} - \tfrac{1}{2}\mathbf{u}|\,\exp\,(-2\pi i\mathbf{u}\cdot\mathbf{Q}). \tag{2.19}$$

Correspondingly, the Weyl transform (2.8) can be written[9]

$$A_w(\mathbf{Q}, \mathbf{K}) = \mathbf{PK}\,[A\Delta(\mathbf{Q}, \mathbf{K})] = \int d\mathbf{v}\langle\mathbf{Q} + \tfrac{1}{2}\mathbf{v}|\,A\,|\mathbf{Q} - \tfrac{1}{2}\mathbf{v}\rangle\,\exp\,(-2\pi i\mathbf{v}\cdot\mathbf{K}). \tag{2.20}$$

[13] A derivation of Eq. (2.14) is given by Messiah and attributed to Glauber. A. Messiah, *Mécanique quantique* (Dunod Cie., Paris, 1959), Vol. I, p. 375.

The transform of any operator $A(\mathbf{q})$ which is independent of $\mathbf{k}$ is $A_w(\mathbf{Q}, \mathbf{K}) = A(\mathbf{Q})$ according to (2.20); similarly, the transform of any operator $A(\mathbf{p})$ is $A(\mathbf{P})$, where $\mathbf{P} = 2\pi\hbar\mathbf{K}$, according to (2.8). The transform of $\Delta(\mathbf{Q}', \mathbf{K}')$ is

$$\Delta_w(\mathbf{Q}'\mathbf{K}', \mathbf{QK}) = \delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}). \quad (2.21)$$

From (2.15) and (2.7),

$$A = \int \cdots \int d\mathbf{u}\, d\mathbf{v}\alpha(\mathbf{u}, \mathbf{v}) \exp [2\pi i(\mathbf{u} \cdot \mathbf{q} + \mathbf{v} \cdot \mathbf{k})], \quad (2.22)$$

where $\alpha(\mathbf{u}, \mathbf{v})$ is the Fourier transform of Weyl transform $A_w(\mathbf{Q}, \mathbf{K})$:

$$A_w(\mathbf{Q}, \mathbf{K}) = \int \cdots \int d\mathbf{u}\, d\mathbf{v}\alpha(\mathbf{u}, \mathbf{v}) \times \exp [2\pi i(\mathbf{u} \cdot \mathbf{Q} + \mathbf{v} \cdot \mathbf{K})]. \quad (2.23)$$

Accordingly, the prescription for obtaining the quantum operator $A(\mathbf{q}, \mathbf{k})$ corresponding to a Weyl transform $A_w(\mathbf{Q}, \mathbf{K})$ is to replace $\mathbf{Q}, \mathbf{K}$ in the Fourier representation (2.23) of $A_w(\mathbf{Q}, \mathbf{K})$ by the operators $\mathbf{q}, \mathbf{k}$. This is the same prescription originally proposed by Weyl[1] for obtaining a quantum operator from a function $A_c(\mathbf{Q}, \mathbf{K})$ of the classical Cartesian coordinates and momenta. Since the Weyl transform $A_w(\mathbf{Q}, \mathbf{K})$ is *not*, in general, the same as a classical function $A_c(\mathbf{Q}, \mathbf{K})$, (as will be shown in Sec. 3, they obey different algebras,) the original Weyl prescription is incorrect.[6] The Weyl transform, defined in (2.8), should not be confused with the classical function; the latter is obtained (see Sec. 5) by taking the classical limit $\hbar \to 0$ of the Weyl transform.

According to (2.7), the matrix elements of A are determined by those of $\Delta(\mathbf{Q}, \mathbf{K})$. From (2.17),

$$\begin{aligned}\langle \mathbf{Q}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{K}'\rangle &= \langle \mathbf{Q}' \mid \mathbf{K}'\rangle \exp [(-i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})] \times \delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}), \\ \langle \mathbf{K}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{Q}'\rangle &= \langle \mathbf{K}' \mid \mathbf{Q}'\rangle \exp [(i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})] \times \delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}).\end{aligned} \quad (2.24)$$

From (2.10) and (2.19),

$$\langle \mathbf{Q}''| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{Q}'\rangle = \delta[\mathbf{Q} - \tfrac{1}{2}(\mathbf{Q}' + \mathbf{Q}'')] \exp [2\pi i\mathbf{K} \cdot (\mathbf{Q}'' - \mathbf{Q}')], \quad (2.25)$$

$$\langle \mathbf{K}''| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{K}'\rangle = \delta[\mathbf{K} - \tfrac{1}{2}(\mathbf{K}' + \mathbf{K}'')] \exp [-2\pi i\mathbf{Q} \cdot (\mathbf{K}'' - \mathbf{K}')]. \quad (2.26)$$

Therefore

$$\begin{aligned}\langle \mathbf{Q}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{Q}'\rangle &= \delta(\mathbf{Q} - \mathbf{Q}'), \\ \langle \mathbf{K}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{K}'\rangle &= \delta(\mathbf{K} - \mathbf{K}').\end{aligned} \quad (2.27)$$

Since

$$\text{trace } \Delta(\mathbf{Q}, \mathbf{K}) = 1, \quad (2.28)$$

therefore, according to (2.7),

$$\text{trace } A = \int \cdots \int d\mathbf{Q}\, d\mathbf{K} A_w(\mathbf{Q}, \mathbf{K}). \quad (2.29)$$

It may be noted also that

$$\int \cdots \int d\mathbf{Q}\, d\mathbf{K}\Delta(\mathbf{Q}, \mathbf{K}) = \mathbf{1}. \quad (2.30)$$

From (2.24), since $\langle \mathbf{Q}' \mid \mathbf{K}'\rangle\langle \mathbf{K}' \mid \mathbf{Q}'\rangle = 1$,

$$\begin{aligned}&\delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}) \\ &= \exp [(i/4\pi)(\partial/\partial\mathbf{Q}') \cdot (\partial/\partial\mathbf{K}')] \times \langle \mathbf{Q}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{K}'\rangle\langle \mathbf{K}' \mid \mathbf{Q}'\rangle, \\ &= \exp [(-i/4\pi)(\partial/\partial\mathbf{Q}') \cdot (\partial/\partial\mathbf{K}')] \times \langle \mathbf{Q}' \mid \mathbf{K}'\rangle\langle \mathbf{K}'| \Delta(\mathbf{Q}, \mathbf{K}) |\mathbf{Q}'\rangle.\end{aligned} \quad (2.31)$$

Accordingly, another form of the Weyl transform is

$$\begin{aligned}&A_w(\mathbf{Q}, \mathbf{K}) \\ &= \int \cdots \int d\mathbf{Q}'\, d\mathbf{K}' A_w(\mathbf{Q}', \mathbf{K}')\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'), \\ &= \exp [(i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})]\langle \mathbf{Q}| A |\mathbf{K}\rangle\langle \mathbf{K} \mid \mathbf{Q}\rangle, \\ &= \exp [(-i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})]\langle \mathbf{Q} \mid \mathbf{K}\rangle\langle \mathbf{K}| A |\mathbf{Q}\rangle.\end{aligned} \quad (2.32)$$

Moyal[10] has defined phase-space eigenfunctions which are equivalent to matrix elements of $\Delta(\mathbf{Q}, \mathbf{K})$. Given an orthonormal set of eigenkets $|\psi_l\rangle$, these are the quantities:

$$f_{l,m}(\mathbf{Q}, \mathbf{K}) = \langle \psi_l| \Delta(\mathbf{Q}, \mathbf{K}) |\psi_m\rangle. \quad (2.33)$$

According to (2.30) they satisfy the "self-orthogonality" relations,

$$\int \cdots \int d\mathbf{Q}\, d\mathbf{K} f_{l,m}(\mathbf{Q}, \mathbf{K}) = \delta_{l,m}. \quad (2.34)$$

In addition, as shown by Moyal,

$$\int \cdots \int d\mathbf{Q}\, d\mathbf{K} f_{l,m}(\mathbf{Q}, \mathbf{K}) f_{l',m'}(\mathbf{Q}, \mathbf{K}) = \delta_{l,l'}\delta_{m,m'}. \quad (2.35)$$

The operator $\Delta(\mathbf{Q}, \mathbf{K})$ can also be exhibited as the difference of two positive definite Hermitian operators.

From Eqs. (2.6) and (2.9),

$$\begin{aligned}
&\Delta(\mathbf{Q},\mathbf{K})\\
&= \tfrac{1}{2}\int d\mathbf{v}[|\mathbf{Q}+\tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q}-\tfrac{1}{2}\mathbf{v}|\exp(2\pi i\mathbf{v}\cdot\mathbf{K})\\
&\qquad + |\mathbf{Q}-\tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q}+\tfrac{1}{2}\mathbf{v}|\exp(-2\pi i\mathbf{v}\cdot\mathbf{K})]\\
&= \tfrac{1}{4}\int d\mathbf{v}\{[\exp(i\pi\mathbf{v}\cdot\mathbf{K})|\mathbf{Q}+\tfrac{1}{2}\mathbf{v}\rangle\\
&\quad + \exp(-i\pi\mathbf{v}\cdot\mathbf{K})|\mathbf{Q}-\tfrac{1}{2}\mathbf{v}\rangle]\\
&\quad \times [\exp(i\pi\mathbf{v}\cdot\mathbf{K})\langle\mathbf{Q}-\tfrac{1}{2}\mathbf{v}|\\
&\quad + \exp(-i\pi\mathbf{v}\cdot\mathbf{K})\langle\mathbf{Q}+\tfrac{1}{2}\mathbf{v}|]\\
&\quad + [\exp(i\pi\mathbf{v}\cdot\mathbf{K})|\mathbf{Q}+\tfrac{1}{2}\mathbf{v}\rangle\\
&\quad - \exp(-i\pi\mathbf{v}\cdot\mathbf{K})|\mathbf{Q}-\tfrac{1}{2}\mathbf{v}\rangle]\\
&\quad \times [\exp(i\pi\mathbf{v}\cdot\mathbf{K})\langle\mathbf{Q}-\tfrac{1}{2}\mathbf{v}|\\
&\quad - \exp(-i\pi\mathbf{v}\cdot\mathbf{K})\langle\mathbf{Q}+\tfrac{1}{2}\mathbf{v}|]\}\\
&= \int d\mathbf{v}\{\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\langle\mathbf{Q}|\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]\\
&\quad - \sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\langle\mathbf{Q}|\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]\}.
\end{aligned} \tag{2.36}$$

It is readily shown that

$$\begin{aligned}
&\int d\mathbf{v}\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\langle\mathbf{Q}|\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]\\
&\qquad = \tfrac{1}{2}[1+\Delta(\mathbf{Q},\mathbf{K})],\\
&\int d\mathbf{v}\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\langle\mathbf{Q}|\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]\\
&\qquad = \tfrac{1}{2}[1-\Delta(\mathbf{Q},\mathbf{K})],\\
&\int d\mathbf{v}\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\langle\mathbf{Q}|\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})] = 0.
\end{aligned} \tag{2.37}$$

Alternatively, from Eqs. (2.18) and (2.19),

$$\begin{aligned}
\Delta(\mathbf{Q},\mathbf{K}) = \int d\mathbf{u}\{&\cos[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]|\mathbf{K}\rangle\langle\mathbf{K}|\\
&\times\cos[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})] - \sin[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]\\
&\times|\mathbf{K}\rangle\langle\mathbf{K}|\sin[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]\}.
\end{aligned} \tag{2.38}$$

Similarly, from Eqs. (2.8), (2.9), (2.18), and (2.20),

$$\begin{aligned}
&A_w(\mathbf{Q},\mathbf{K})\\
&= \int d\mathbf{v}\{\langle\mathbf{Q}|\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]A\cos[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\\
&\quad - \langle\mathbf{Q}|\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]A\sin[\pi\mathbf{v}\cdot(\mathbf{k}-\mathbf{K})]|\mathbf{Q}\rangle\}\\
&= \int d\mathbf{u}\{\langle\mathbf{K}|\cos[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]A\cos[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]|\mathbf{K}\rangle\\
&\quad - \langle\mathbf{K}|\sin[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]A\sin[\pi\mathbf{u}\cdot(\mathbf{q}-\mathbf{Q})]|\mathbf{K}\rangle\}.
\end{aligned} \tag{2.39}$$

3. ALGEBRAIC RELATIONSHIPS

The products of Weyl transforms can readily be obtained by the use of the relationship

$$\begin{aligned}
&\Delta(\mathbf{Q},\mathbf{K})\Delta(\mathbf{Q}',\mathbf{K}')\\
&\quad = \exp\{(i/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}') - (\partial/\partial\mathbf{K})\cdot(\partial/\partial\mathbf{Q}')]\}\\
&\qquad \times \Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}').
\end{aligned} \tag{3.1}$$

The proof of this equation follows: according to Eq. (2.15),

$$\begin{aligned}
&\Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&= \int\cdots\int d\mathbf{u}\,d\mathbf{v}\exp[2\pi i(\mathbf{u}\cdot\mathbf{q}+\mathbf{v}\cdot\mathbf{k})]\\
&\quad \times\exp[(-\pi i)(\mathbf{u}\cdot\mathbf{Q}+\mathbf{v}\cdot\mathbf{K})]\\
&\quad \times\exp[(-\pi i)(\mathbf{u}\cdot\mathbf{Q}'+\mathbf{v}\cdot\mathbf{K}')]\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&= \int\cdots\int d\mathbf{u}\,d\mathbf{v}\,d\mathbf{u}'\,d\mathbf{v}'\exp[2\pi i(\mathbf{u}\cdot\mathbf{q}+\mathbf{v}\cdot\mathbf{k})]\\
&\quad \times\exp\{-2\pi i[\mathbf{Q}\cdot(\mathbf{u}'+\tfrac{1}{2}\mathbf{u})+\mathbf{K}\cdot(\mathbf{v}'+\tfrac{1}{2}\mathbf{v})]\}\\
&\quad \times\exp\{2\pi i[\mathbf{Q}'\cdot(\mathbf{u}'-\tfrac{1}{2}\mathbf{u})+\mathbf{K}'\cdot(\mathbf{v}'-\tfrac{1}{2}\mathbf{v})]\}.
\end{aligned}$$

The change of variables, $\mathbf{u}''=\mathbf{u}'+\frac{1}{2}\mathbf{u}$, $\mathbf{u}'''=-\mathbf{u}'+\frac{1}{2}\mathbf{u}$, $\mathbf{v}''=\mathbf{v}'+\frac{1}{2}\mathbf{v}$, $\mathbf{v}'''=-\mathbf{v}'+\frac{1}{2}\mathbf{v}$, gives

$$\begin{aligned}
&\Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&= \int\cdots\int d\mathbf{u}''\,d\mathbf{v}''\,d\mathbf{u}'''\,d\mathbf{v}'''\\
&\quad \times\exp\{2\pi i[\mathbf{u}''\cdot(\mathbf{q}-\mathbf{Q})+\mathbf{v}''\cdot(\mathbf{k}-\mathbf{K})\\
&\qquad +\mathbf{u}'''\cdot(\mathbf{q}-\mathbf{Q}')+\mathbf{v}'''\cdot(\mathbf{k}-\mathbf{K}')]\},
\end{aligned}$$

and use of (2.14) gives

$$\begin{aligned}
&\Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&= \int\cdots\int d\mathbf{u}''\,d\mathbf{v}''\,d\mathbf{u}'''\,d\mathbf{v}'''\exp[i\pi(\mathbf{u}''\cdot\mathbf{v}'''-\mathbf{u}'''\cdot\mathbf{v}'')]\\
&\quad \times\exp\{2\pi i[\mathbf{u}''\cdot(\mathbf{q}-\mathbf{Q})+\mathbf{v}''\cdot(\mathbf{k}-\mathbf{K})]\}\\
&\quad \times\exp\{2\pi i[\mathbf{u}'''\cdot(\mathbf{q}-\mathbf{Q}')+\mathbf{v}'''\cdot(\mathbf{k}-\mathbf{K}')]\}\\
&= \exp\{(-i/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}')-(\partial/\partial\mathbf{K})\cdot(\partial/\partial\mathbf{Q}')]\}\\
&\quad \times\Delta(\mathbf{Q},\mathbf{K})\Delta(\mathbf{Q}',\mathbf{K}').
\end{aligned}$$

With the use of Eq. (3.1), the product of two quantum operators becomes

$$\begin{aligned}
AB &= \int\cdots\int d\mathbf{Q}\,d\mathbf{K}\,d\mathbf{Q}'\,d\mathbf{K}'A_w(\mathbf{Q},\mathbf{K})B_w(\mathbf{Q}',\mathbf{K}')\\
&\qquad \times\Delta(\mathbf{QK})\Delta(\mathbf{Q}',\mathbf{K}')\\
&= \int\cdots\int d\mathbf{Q}\,d\mathbf{K}\,d\mathbf{Q}'\,d\mathbf{K}'A_w(\mathbf{Q},\mathbf{K})B_w(\mathbf{Q}',\mathbf{K}')\\
&\quad \times\exp\{(i/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}')\\
&\qquad -(\partial/\partial\mathbf{K})\cdot(\partial/\partial\mathbf{Q}')]\}\\
&\quad \times\Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&= \int\cdots\int d\mathbf{Q}\,d\mathbf{K}\,d\mathbf{Q}'\,d\mathbf{K}'\\
&\quad \times\Delta(\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}')\\
&\quad \times\exp\{(i/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}')\\
&\quad -(\partial/\partial\mathbf{K})\cdot(\partial/\partial\mathbf{Q}')]\}A_w(\mathbf{Q},\mathbf{K})B_w(\mathbf{Q}',\mathbf{K}'),
\end{aligned} \tag{3.2}$$

$$AB = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}'\, d\mathbf{K}' \Delta(\mathbf{Q}, \mathbf{K}) \times \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \cdot A_w[\mathbf{Q} + (i/4\pi)(\partial/\partial\mathbf{K}') \quad \mathbf{K} - (i/4\pi)(\partial/\partial\mathbf{Q}')]B_w(\mathbf{Q}', \mathbf{K}'). \tag{3.3}$$

But since the Weyl transform of AB is $(AB)_w(\mathbf{Q}, \mathbf{K})$, Eq. (3.2) gives

$$(AB)_w(\mathbf{Q}, \mathbf{K}) = \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}' \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \times \exp\{(i/4\pi)[(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K}') - (\partial/\partial\mathbf{K}) \cdot (\partial/\partial\mathbf{Q}')]\} \times A_w(\mathbf{Q}, \mathbf{K})B_w(\mathbf{Q}', \mathbf{K}'). \tag{3.4}$$

If $A_w(\mathbf{Q}, \mathbf{K})$ and $B_w(\mathbf{Q}, \mathbf{K})$ are real, then $(AB)_w(\mathbf{Q}, \mathbf{K})$ is not real. Furthermore, the Weyl transform $(AB)_w(\mathbf{Q}, \mathbf{K})$ is not the same as $(BA)_w(\mathbf{Q}, \mathbf{K})$. From Eq. (3.2),

$$\tfrac{1}{2}(AB + BA) = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}'\, d\mathbf{K}' \Delta(\mathbf{Q}, \mathbf{K})\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \times \cos\{(4\pi)^{-1}[(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K}') - (\partial/\partial\mathbf{K}) \cdot (\partial/\partial\mathbf{Q}')]\} \times A_w(\mathbf{Q}, \mathbf{K})B_w(\mathbf{Q}', \mathbf{K}'), \tag{3.5}$$

$$AB - BA = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}'\, d\mathbf{K}' \Delta(\mathbf{Q}, \mathbf{K})\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \times 2i \sin\{(4\pi)^{-1}[(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K}') - (\partial/\partial\mathbf{K}) \cdot (\partial/\partial\mathbf{Q}')\} \times A_w(\mathbf{Q}, \mathbf{K})B_w(\mathbf{Q}', \mathbf{K}'). \tag{3.6}$$

According to Eqs. (3.1) and (2.28),

$$\text{trace } \Delta(\mathbf{Q}, \mathbf{K})\Delta(\mathbf{Q}', \mathbf{K}') = \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'). \tag{3.7}$$

Therefore,

$$\text{trace } AB = \int\cdots\int d\mathbf{Q}\, d\mathbf{K} A_w(\mathbf{Q}, \mathbf{K})B_w(\mathbf{Q}, \mathbf{K}). \tag{3.8}$$

The algebraic relationships derived in this section differ from, but are similar to, those in the literature which are based upon the incorrect Weyl correspondence. [See, for example, Ref. 2.]

4. WIGNER DENSITY FUNCTION

The state of a system in quantum mechanics is given by its density operator ρ. ρ is a positive definite Hermitian operator which has the general form

$$\rho = \sum_m |\psi_m\rangle w_m \langle\psi_m|, \quad 0 \le w_m \le 1, \quad \sum_m w_m = 1, \tag{4.1}$$

$$\text{trace } \rho = 1. \tag{4.2}$$

In the representation of wave mechanics, the Weyl transform of ρ given by (4.1) and (2.20) is

$$\rho_w(\mathbf{Q}, \mathbf{K}) = \sum_m w_m \int d\mathbf{v}\, \psi_m^*(\mathbf{Q} + \tfrac{1}{2}\mathbf{v})\psi_m(\mathbf{Q} - \tfrac{1}{2}\mathbf{v}) \exp(2\pi i \mathbf{K} \cdot \mathbf{v}), \tag{4.3}$$

the form originally used by Wigner.[4] The expectation value of an observable A for a quantum system in state ρ is, according to Eqs. (3.8), (2.29), and (4.2),

$$\langle A\rangle = \text{trace } \rho A = \int\cdots\int d\mathbf{Q}\, d\mathbf{K} \rho_w(\mathbf{Q}, \mathbf{K})A_w(\mathbf{Q}, \mathbf{K}), \tag{4.4}$$

with

$$\int\cdots\int d\mathbf{Q}\, d\mathbf{K} \rho_w(\mathbf{Q}, \mathbf{K}) = 1. \tag{4.5}$$

The Wigner function $\rho_w(\mathbf{Q}, \mathbf{K})$ can be considered to be a probability density in a phase space whose representative point has coordinates $(\mathbf{Q}, \mathbf{K})$. It is not, however, nonnegative at every point in the phase space, as is required of a probability density.

According to Eq. (2.7) every quantum state is specified by a density operator

$$\rho = \int\cdots\int d\mathbf{Q}\, d\mathbf{K} \rho_w(\mathbf{Q}, \mathbf{K})\Delta(\mathbf{Q}, \mathbf{K}). \tag{4.6}$$

Classically, a state of a system can be specified by a probability density

$$\rho_c(\mathbf{Q}, \mathbf{P}) = (2\pi\hbar)^N \delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{P} - \mathbf{P}_0) = \delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{K} - \mathbf{K}_0). \tag{4.7}$$

Therefore, it might be suggested that a quantum state could be specified by a density operator $\rho = \Delta(\mathbf{Q}_0, \mathbf{K}_0)$ whose Wigner function, according to Eq. (2.21), is $\delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{K} - \mathbf{K}_0)$. But such a quantum state does not exist; $\Delta(\mathbf{Q}_0, \mathbf{K}_0)$ does not meet the requirement of being a positive definite operator.[4] In fact, in Eqs. (2.36) and (2.38), $\Delta(\mathbf{Q}, \mathbf{K})$ is exhibited as the difference of two positive-definite operators. Since ρ must be positive definite while $\Delta(\mathbf{Q}, \mathbf{K})$ is not, therefore, in (4.6), the Wigner function $\rho_w(\mathbf{Q}, \mathbf{K})$ cannot be nonnegative; it must be negative at the points where $\Delta(\mathbf{Q}, \mathbf{K})$ is negative. There can be no Wigner function equal to $\delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{K} - \mathbf{K}_0)$; it is easy to show that for such a Wigner function the uncertainty relation of Heisenberg would not hold. The general form of the Wigner function for the state (4.1) is given by (2.39) as

$$\rho_w(\mathbf{Q}, \mathbf{K}) = \sum_m w_m \int d\mathbf{v}\, \langle\mathbf{Q}| \cos \pi\mathbf{v} \cdot (\mathbf{k} - \mathbf{K}) |\psi_m\rangle \times \langle\psi_m| \cos \pi\mathbf{v} \cdot (\mathbf{k} - \mathbf{K}) |\mathbf{Q}\rangle - \sum_m w_m \int d\mathbf{v} \langle\mathbf{Q}| \sin \pi\mathbf{v} \times (\mathbf{k} - \mathbf{K}) |\psi_m\rangle\langle\psi_m| \sin \pi\mathbf{v} \cdot (\mathbf{k} - \mathbf{K}) |\mathbf{Q}\rangle \tag{4.8}$$

which displays $\rho_w(\mathbf{Q}, \mathbf{K})$ as the difference of nonnegative terms. The fact that no Wigner function $\rho_w(\mathbf{Q}, \mathbf{K})$ can be nonnegative throughout the phase-space while all classical distribution functions $\rho_c(\mathbf{Q}, \mathbf{P})$

must be nonnegative [an example appears in (4.7)] illustrates clearly that the transform $A_w(\mathbf{Q}, \mathbf{K})$ of a quantum operator must not be confused with the corresponding classical function $A_c(\mathbf{Q}, \mathbf{K})$; $A_c(\mathbf{Q}, \mathbf{K})$ results from $A_w(\mathbf{Q}, \mathbf{K})$ only in the limit $\hbar \to 0$.

A good application of the transform (2.7) is to obtain the time dependence of the Wigner function from that of the density operator $\rho(t)$. In the Schrödinger picture $\rho(t)$ is a solution of the von Neumann equation (H is the Hamiltonian operator)

$$-i\hbar\partial\rho/\partial t = \rho H - H\rho. \tag{4.9}$$

According to Eqs. (2.7) and (3.6),

$$\begin{aligned}
&\partial\rho_w(t, \mathbf{Q}, \mathbf{K})/\partial t \\
&= \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \cdot (2/\hbar) \\
&\quad \times \sin\{(4\pi)^{-1}[(\partial/\partial\mathbf{Q}') \cdot (\partial/\partial\mathbf{K}) - (\partial/\partial\mathbf{K}') \cdot (\partial/\partial\mathbf{Q})]\} \\
&\quad \times H_w(\mathbf{Q}', \mathbf{K}')\rho_w(t, \mathbf{Q}, \mathbf{K}) \qquad (4.10) \\
&= \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \cdot (-i/\hbar) \\
&\quad \times \{H_w[\mathbf{Q}' + (i/4\pi)(\partial/\partial\mathbf{K}), \mathbf{K}' - (i/4\pi)(\partial/\partial\mathbf{Q})] \\
&\quad - H_w[\mathbf{Q}' - (i/4\pi)(\partial/\partial\mathbf{K}), \mathbf{K}' + (i/4\pi)(\partial/\partial\mathbf{Q})]\} \\
&\quad \times \rho_w(t, \mathbf{Q}, \mathbf{K}). \qquad (4.11)
\end{aligned}$$

It should be emphasized that this is an exact quantum-mechanical equation. With the Weyl transform $\rho_w(t, \mathbf{Q}, \mathbf{K})$ and $H_w(\mathbf{Q}, \mathbf{K})$ in the form given by (2.8), Eq. (4.11) becomes

$$\begin{aligned}
&\partial\langle\mathbf{K} + \tfrac{1}{2}\mathbf{u}|\, \rho\, |\mathbf{K} - \tfrac{1}{2}\mathbf{u}\rangle/\partial t \\
&= -(i/\hbar)\int d\mathbf{u}'\langle\mathbf{K} + \tfrac{1}{2}\mathbf{u}'|\, H\, |\mathbf{K} - \tfrac{1}{2}\mathbf{u}'\rangle \\
&\quad \times [\exp(-\tfrac{1}{2}\mathbf{u}' \cdot \partial/\partial\mathbf{K}) - \exp(\tfrac{1}{2}\mathbf{u}' \cdot \partial/\partial\mathbf{K})] \\
&\quad \times \langle\mathbf{K} + \tfrac{1}{2}(\mathbf{u} - \mathbf{u}')|\, \rho\, |\mathbf{K} - \tfrac{1}{2}(\mathbf{u} - \mathbf{u}')\rangle, \qquad (4.12)
\end{aligned}$$

in agreement with the formulation of Prigogine (Ref. 11).

5. CLASSICAL LIMIT

The classical limit results from the quantum formalism by letting $\hbar \to 0$, keeping $\mathbf{p} = 2\pi\hbar\mathbf{k}$ and $\mathbf{P} = 2\pi\hbar\mathbf{K}$ unchanged. Accordingly, the formalism must be expressed in terms of $\mathbf{p}$ and $\mathbf{P}$ in place of $\mathbf{k}$ and $\mathbf{K}$ before the limit $\hbar \to 0$ is applied.

Since $|\mathbf{K}\rangle = (2\pi\hbar)^{\frac{1}{2}N}|\mathbf{P}\rangle$ for a system with N degrees of freedom, therefore (2.32) gives, in the classical limit,

$$\begin{aligned}
A_w(\mathbf{Q}, \mathbf{K}) &= (2\pi\hbar)^{-N}\exp[(\tfrac{1}{2}i\hbar)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{P})] \\
&\qquad \times \langle\mathbf{Q}|\, A\, |\mathbf{P}\rangle\langle\mathbf{P}\,|\,\mathbf{Q}\rangle \\
&\to (2\pi\hbar)^{-N}\langle\mathbf{Q}|\, A\, |\mathbf{P}\rangle\langle\mathbf{P}\,|\,\mathbf{Q}\rangle \\
&= \text{trace}\, |\mathbf{K}\rangle\langle\mathbf{K}\,|\,\mathbf{Q}\rangle\langle\mathbf{Q}|\, A, \qquad (5.1)
\end{aligned}$$

and also

$$\begin{aligned}
A_w(\mathbf{Q}, \mathbf{K}) &\to (2\pi\hbar)^{-N}\langle\mathbf{Q}\,|\,\mathbf{P}\rangle\langle\mathbf{P}|\, A\, |\mathbf{Q}\rangle \\
&= \text{trace}\, |\mathbf{Q}\rangle\langle\mathbf{Q}\,|\,\mathbf{K}\rangle\langle\mathbf{K}|\, A. \qquad (5.2)
\end{aligned}$$

Similarly, (2.17) gives

$$\begin{aligned}
\Delta(\mathbf{Q}, \mathbf{K}) &= (2\pi\hbar)^{-N}\exp[(\tfrac{1}{2}i\hbar)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{P})] \\
&\qquad \times \delta(\mathbf{p} - \mathbf{P})\delta(\mathbf{q} - \mathbf{Q}) \\
&\to \delta(\mathbf{k} - \mathbf{K})\delta(\mathbf{q} - \mathbf{Q}) \\
&= \delta(\mathbf{q} - \mathbf{Q})\delta(\mathbf{k} - \mathbf{K}). \qquad (5.3)
\end{aligned}$$

These results are in accord with the fact that q and $\mathbf{p}$ (or $\mathbf{k}$) commute in the classical limit, since, according to Eq. (2.3),

$$q_i p_j - p_j q_i = i\hbar\delta_{ij} \to 0.$$

Since the order of the commuting operators is immaterial, the classical limit in Eq. (5.3) can be written [see Eq. (2.11)] as

$$\begin{aligned}
\Delta(\mathbf{Q}, \mathbf{K}) &\to |\mathbf{K}\rangle\langle\mathbf{K}\,|\,\mathbf{Q}\rangle\langle\mathbf{Q}| \\
&= |\mathbf{Q}\rangle\langle\mathbf{Q}\,|\,\mathbf{K}\rangle\langle\mathbf{K}| \equiv |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|. \qquad (5.4)
\end{aligned}$$

Similarly, the classical limit in Eqs. (5.1) and (5.2) can be written as

$$\begin{aligned}
A_w(\mathbf{Q}, \mathbf{K}) &\to \text{trace}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|\, A \\
&\equiv A_c(\mathbf{Q}, \mathbf{K}) = \mathcal{A}_c(\mathbf{Q}, \mathbf{P}). \qquad (5.5)
\end{aligned}$$

In the classical limit A becomes an operator in a Hilbert space of vectors $|\mathbf{Q}, \mathbf{K}\rangle$ which are simultaneous eigenkets of the commuting operators $\mathbf{q}$, $\mathbf{k}$. According to Eqs. (5.4) and (2.30),

$$\int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}| = \mathbf{1} \tag{5.6}$$

so that these eigenkets are complete in the space. Also, Eq. (5.6) implies the normalization condition

$$\langle\mathbf{Q}', \mathbf{K}'\,|\,\mathbf{Q}, \mathbf{K}\rangle = \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'). \tag{5.7}$$

Now, $|\mathbf{Q}, \mathbf{K}\rangle$ is an eigenfunction of A belonging to eigenvalue $A_c(\mathbf{Q}, \mathbf{K})$:

$$A\, |\mathbf{Q}, \mathbf{K}\rangle = A_c(\mathbf{Q}, \mathbf{K})\, |\mathbf{Q}, \mathbf{K}\rangle. \tag{5.8}$$

Accordingly,

$$\langle\mathbf{Q}, \mathbf{K}|\, A\, |\mathbf{Q}', \mathbf{K}'\rangle = A_c(\mathbf{Q}, \mathbf{K})\delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}), \tag{5.9}$$

so that, in agreement with Eq. (5.5)

$$\begin{aligned}
A_c(\mathbf{Q}, \mathbf{K}) &\equiv \text{trace}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|\, A \\
&= \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\langle\mathbf{Q}, \mathbf{K}|\, A\, |\mathbf{Q}', \mathbf{K}'\rangle. \qquad (5.10)
\end{aligned}$$

Because of the singularity of the diagonal matrix elements in Eq. (5.9) the integration over $(\mathbf{Q}', \mathbf{K}')$ in

Eq. (5.10) need include only the "diagonal" point $(\mathbf{Q}', \mathbf{K}') = (\mathbf{Q}, \mathbf{K})$.

From Eq. (5.9) and the identity

$$A = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}'\, d\mathbf{K}' \times |\mathbf{Q}', \mathbf{K}'\rangle\langle\mathbf{Q}', \mathbf{K}'|\, A\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|. \quad (5.11)$$

it follows that

$$A = \int\cdots\int d\mathbf{Q}\, d\mathbf{K} A_c(\mathbf{Q}, \mathbf{K})\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|, \quad (5.12)$$

which is the classical limit of the transformation (2.7) according to (5.4) and (5.5).

In the classical limit the algebraic relation (3.4) becomes

$$(AB)_c(\mathbf{Q}, \mathbf{K}) = A_c(\mathbf{Q}, \mathbf{K})B_c(\mathbf{Q}, \mathbf{K}) = (BA)_c(\mathbf{Q}, \mathbf{K}), \quad (5.13)$$

showing that the classical functions obey commutative multiplication. Since, for a system of N degrees of freedom,

$$|\mathbf{Q}, \mathbf{P}\rangle = (2\pi\hbar)^{-\frac{1}{2}N}\, |\mathbf{Q}, \mathbf{K}\rangle.$$

from Eqs. (3.5), (5.4), and (5.5), to lowest order in $\hbar$,

$$\begin{aligned}\tfrac{1}{2}(AB + BA) &\to \int\cdots\int d\mathbf{Q}\, d\mathbf{P}\, |\mathbf{Q}, \mathbf{P}\rangle\langle\mathbf{Q}, \mathbf{P}|\, \mathcal{A}_c(\mathbf{Q}, \mathbf{P})\mathcal{B}_c(\mathbf{Q}, \mathbf{P}) \\ &= \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|\, A_c(\mathbf{Q}, \mathbf{K})B_c(\mathbf{Q}, \mathbf{K}), \end{aligned} \quad (5.14)$$

and, from Eq. (3.6),

$$\begin{aligned} AB - BA &\to \int\cdots\int d\mathbf{Q}\, d\mathbf{P}\, d\mathbf{Q}'\, d\mathbf{P}'\, |\mathbf{Q}, \mathbf{P}\rangle\langle\mathbf{Q}, \mathbf{P}| \\ &\quad \times \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{P} - \mathbf{P}') \cdot i\hbar \\ &\quad \times [(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{P}') - (\partial/\partial\mathbf{P}) \cdot (\partial/\partial\mathbf{Q}')] \\ &\quad \times \mathcal{A}_c(\mathbf{Q}, \mathbf{P})\mathcal{B}_c(\mathbf{Q}', \mathbf{P}') \\ &= \int\cdots\int d\mathbf{Q}\, d\mathbf{P}\, |\mathbf{Q}, \mathbf{P}\rangle\langle\mathbf{Q}, \mathbf{P}| \\ &\quad \times i\hbar[(\partial\mathcal{A}_c/\partial\mathbf{Q}) \cdot (\partial\mathcal{B}_c/\partial\mathbf{P}) - (\partial\mathcal{A}_c/\partial\mathbf{P}) \cdot (\partial\mathcal{B}_c/\partial\mathbf{Q})] \\ &= i\int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}| \\ &\quad \times [(\partial A_c/\partial\mathbf{Q}) \cdot (\partial B_c/\partial\mathbf{K}) - (\partial A_c/\partial\mathbf{K}) \cdot (\partial B_c/\partial\mathbf{Q})]. \end{aligned} \quad (5.15)$$

Equations (5.14) and (5.15) express the well-known correspondence rules:

(a) The product $\mathcal{A}_c(\mathbf{Q}, \mathbf{P})\mathcal{B}_c(\mathbf{Q}, \mathbf{P})$ corresponds in the classical limit to the symmetrized quantum operator $\frac{1}{2}(AB + BA)$.

(b) The Poisson bracket $(\partial\mathcal{A}_c/\partial\mathbf{Q}) \cdot (\partial\mathcal{B}_c/\partial\mathbf{P}) - (\partial\mathcal{A}_c/\partial\mathbf{P}) \cdot (\partial\mathcal{B}_c/\partial\mathbf{Q})$ corresponds in the classical limit to the commutator $(AB - BA)/i\hbar$.

From Eq. (4.10), the classical limit of the equation for the time dependence of the Wigner function is the classical Liouville equation,

$$\begin{aligned}&\partial F_c(t, \mathbf{Q}, \mathbf{P})/\partial t \\ &\quad = [\partial\mathcal{H}_c(\mathbf{Q}, \mathbf{P})/\partial\mathbf{Q}] \cdot [\partial F_c(t, \mathbf{Q}, \mathbf{P})/\partial\mathbf{P}] \\ &\qquad - [\partial\mathcal{H}_c(\mathbf{Q}, \mathbf{P})/\partial\mathbf{P}] \cdot [\partial F_c(t, \mathbf{Q}, \mathbf{P})/\partial\mathbf{Q}], \end{aligned} \quad (5.16)$$

where

$$F_c(t, \mathbf{Q}, \mathbf{P}) = \rho_c(t, \mathbf{Q}, \mathbf{K}), \quad \mathcal{H}_c(\mathbf{Q}, \mathbf{P}) = H_c(\mathbf{Q}, \mathbf{K}).$$

From Eq. (3.8),

$$\begin{aligned}\text{trace } AB &\to (2\pi\hbar)^{-N}\int\cdots\int d\mathbf{Q}\, d\mathbf{P}\mathcal{A}_c(\mathbf{Q}, \mathbf{P})\mathcal{B}_c(\mathbf{Q}, \mathbf{P}) \\ &= \int\cdots\int d\mathbf{Q}\, d\mathbf{K} A_c(\mathbf{Q}, \mathbf{K})B_c(\mathbf{Q}, \mathbf{K}), \end{aligned} \quad (5.17)$$

so that the expectation of A in the state described by density operator ρ becomes

$$\begin{aligned}\langle A\rangle &\to (2\pi\hbar)^{-N}\int\cdots\int d\mathbf{Q}\, d\mathbf{P} F_c(t, \mathbf{Q}, \mathbf{P})\mathcal{A}_c(\mathbf{Q}, \mathbf{P}) \\ &= \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\rho_c(t, \mathbf{Q}, \mathbf{K})A_c(\mathbf{Q}, \mathbf{K}). \end{aligned} \quad (5.18)$$

A general classical state is given by

$$|\psi(t)\rangle = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K} \mid \psi(t)\rangle, \quad (5.19)$$

so that in the classical limit the density operator of Eq. (4.1) becomes

$$\begin{aligned}\rho(t) &\to \sum_m w_m \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}'\, d\mathbf{K}' \\ &\times |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K} \mid \psi_m(t)\rangle\langle\psi_m(t) \mid \mathbf{Q}', \mathbf{K}'\rangle\langle\mathbf{Q}', \mathbf{K}'|. \end{aligned} \quad (5.20)$$

But, according to (5.9),

$$\begin{aligned}\sum_m w_m\langle\mathbf{Q}, \mathbf{K} \mid \psi_m(t)\rangle\langle\psi_m(t) \mid \mathbf{Q}', \mathbf{K}'\rangle \\ = \rho_c(t, \mathbf{Q}, \mathbf{K})\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'). \end{aligned} \quad (5.21)$$

Therefore,

$$\rho(t) = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\rho_c(t, \mathbf{Q}, \mathbf{K})\, |\mathbf{Q}, \mathbf{K}\rangle\langle\mathbf{Q}, \mathbf{K}|, \quad (5.22)$$

and

$$\begin{aligned}&\rho_c(t, \mathbf{Q}, \mathbf{K}) \\ &\quad = \sum_m w_m \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\langle\mathbf{Q}, \mathbf{K} \mid \psi_m(t)\rangle \\ &\qquad \times \langle\psi_m(t) \mid \mathbf{Q}', \mathbf{K}'\rangle \geq 0, \end{aligned} \quad (5.23)$$

where the integration over $(\mathbf{Q}', \mathbf{K}')$ need only include the diagonal point $(\mathbf{Q}, \mathbf{K})$. In the classical limit the Wigner function becomes nonnegative throughout the phase-space.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Gauge Fields with Noninvariant Interactions

R. FITZGERALD* AND S. SCHIMINOVICH
Department of Physics, The City College of the City University of New York, New York, New York

(Received 20 March 1967)

Gauge fields interacting with source fields in a gauge-noninvariant way are examined in a purely classical framework. The existence of conserved currents for the solutions of such theories follows from the properties of the free gauge fields rather than from the usual requirement of global gauge covariance. One of the simplest gauge-noninvariant theories of electromagnetism is shown to be an extension of Dirac's theory of classical clouds of charge. In this extension the clouds exhibit some quantum features. Gauge-noninvariant interactions for Einstein's theory of gravitation are presented in which the properties of the sources are carried by the metric field. In the case of the Yang–Mills field, internal symmetries for a real scalar source field are a consequence of the properties of the gauge field.

1. INTRODUCTION

This paper deals with the physical content of theories in which a gauge field, such as the electromagnetic, Yang–Mills, or gravitational field, is coupled in a gauge-noninvariant way to a source field. Customarily, such theories are excluded by the requirement of global gauge covariance. This requirement is imposed either to assure the existence of conserved quantities or for heuristic reasons. Such motivations are questionable. Although gauge covariance and the existence of a conserved current are inseparably joined in single-field theories by Noether's procedures, we show that, for interacting fields, conserved currents can follow as a result of the existence of Bianchi identities satisfied by the gauge-field variables.

In proposing a study of the gauge-noninvariant interactions of gauge fields, our point of departure is opposite that taken by Utiyama,[1] who generates interactions using the requirements of global gauge covariance. Such an invariant theoretical approach to interactions has proven sterile up to the present and is subject to criticism.[2] In Utiyama's procedure, the source field, invariant under its constant-parameter gauge group, is the all-important initial element; in our procedure the gauge field invariant under its coordinate-dependent gauge group, is the basis upon which new theories are constructed.

In order to give meaning to the solutions of a gauge-invariant theory, one has to supplement the theory with coordinate conditions. These can be given at the level of the action principle with the aid of auxiliary Lagrange multiplier fields coupled through gauge-noninvariant terms. The multiplier field can be made a dynamical field by introducing its free Lagrangian into the action. It is hoped that such dynamical multiplier fields may describe properties of gauge-field sources known to exist in nature.

Because the theory is now gauge noninvariant, all the components of the original gauge field have physical meaning. This means that properties of the sources are described not only by the new Lagrange multiplier fields but also by the components of the gauge field. In this way our work will provide, when applied to the electromagnetic field, an extension to Dirac's classical theory of electrons.[3]

The existence of physically meaningful gauge-field components is further related to the discussion and work on the physical meaning of potentials by Aharanov and Bohm.[4] The interpretation of the physical meaning of potentials is bound to the determination of the role played by the phase variables; for, as we shall see, the potentials acquire physical meaning when the phase variables have been completely eliminated from the formulation. Additional points of contact exist with the work of Aharanov and Wisniveski.[5]

There are occasions when the Lagrange multiplier field can be omitted and the original gauge invariance broken merely by the introduction of noninvariant terms into the free gauge-field Lagrangian. In such a case the properties of the source fields are described solely by the gauge field. Thus our study, when applied to the gravitational field, will exhibit a relationship with the work of Wheeler, Rainich, and Misner[6] expressed in a particular coordinate frame.

---

* In partial fulfillment of the requirements for the Ph.D. degree at the City College of the City University of New York.

[1] R. Utiyama, Phys. Rev. **101**, 1597 (1956).

[2] V. I. Ogievetski and I. V. Polubarinov, Nuovo Cimento **23**, 173 (1962); V. Fock, *The Theory of Space Time and Gravitation* (The Macmillan Company, New York, 1964).

[3] P. A. M. Dirac, Proc. Roy. Soc. (London) **209A**, 291 (1951); **212A**, 330 (1952); **223A**, 438 (1954).

[4] Y. Aharanov and D. Bohm, Phys. Rev. **115**, 485 (1959); **123**, 1511 (1961).

[5] Y. Aharanov and D. Wisniveski (to be published).

[6] See, for example, Louis Witten, in *Gravitation: An Introduction To Current Research*, L. Witten, Ed. (John Wiley & Sons, Inc., New York, 1962).

For the present we limit our attention to theories that can be derived from an Utiyama-type theory. It is clear that any such theory can be cast in a frame that is manifestly noninvariant by choosing a particular gauge. This procedure is not trivial as a result of the following:

(a) We require that the particular gauge chosen leaves the theory written explicitly and solely in terms of physically meaningful fields (i.e., in terms of fields not subject to any remaining gauge conditions).

(b) We require that the choice of gauge reduces the number of dynamical variables. The possibility of decreasing the number of variables indicates that the original Utiyama-type theory contains an element of arbitrariness: Gauge covariance is accomplished by the introduction of fictitious fields and by the subsequent redefinition of the observable fields in a manner which compensates for these fictitious fields.

In general it is possible to formally choose the gauge which reduces the number of explicit variables in the Lagrangian. For particular solutions, however, this may require discontinuous or singular gauge transformations, especially when the solutions are defined in nonsimple connected domains. The physical content of the new and old theories may then be different.

To begin with we show how we can use the gauge freedom in the Utiyama-type theories to formally reduce the number of explicit dynamical variables and yet guarantee the existence of conserved quantities by reason of the Bianchi identities satisfied by the gauge-field variables. We then consider the example of the electromagnetic field coupled to a Klein–Gordon field and show how this formulation provides an extension to Dirac's considerations. Beams of classical charges described by this theory enjoy some of the properties of quantized charges. In particular, static distributions of charge in a Coulomb field are possible with the same multiplicity of solutions and energies as the stationary states of a hydrogen atom, provided self-energies are neglected. We then briefly consider the cases of the gravitational and Yang–Mills fields.

2. GENERAL THEORY

We first consider the properties of theories of the Utiyama type and how in such theories the gauge degrees of freedom can be used to reduce the number of explicit dynamical variables. These theories can be generally characterized by actions of the form

$$\text{ACTION} = \int_\Omega d^4x\{\mathfrak{L}_G(g) + \mathfrak{L}_I(g, p) + \mathfrak{L}_M(p)\}, \quad (2.1)$$

where $\mathfrak{L}_G(g)$ and $\mathfrak{L}_I(g, p) + \mathfrak{L}_M(p)$ are scalar densities under a group of global gauge transformations (i.e., coordinate-dependent transformations whose parameters vary arbitrarily in space–time). $\mathfrak{L}_G(g)$ is the free gauge-field Lagrangian, $\mathfrak{L}_M(p)$ is the free source-field Lagrangian, and $\mathfrak{L}_I(g, p)$ is the coupling term. The gauge invariance properties can be expressed as

$$\delta^* \int_\Omega d^4x \mathfrak{L}_G(g) = 0, \quad (2.2)$$

$$\delta^* \int_\Omega d^4x\{\mathfrak{L}_I(g, p) + \mathfrak{L}_M(p)\} = 0, \quad (2.3)$$

where δ^* means that the variables g and p are varied by an infinitesimal amount determined by an arbitrary infinitesimal transformation of the global gauge group. The variables g are the gauge-field variables and the variables p are the source-field variables.

In Utiyama's procedure one starts with an action $\mathfrak{L}_M(p)$ which is invariant only under the group of gauge transformations depending on n constant parameters. Imposing the requirement that the theory be invariant under the more general group of global gauge transformations, for which the n parameters may be arbitrary functions of space–time, forces the introduction of a set of gauge-field variables g through the term $\mathfrak{L}_I(g, p)$. The theory is then completed by adding a free gauge-field Lagrangian which is restricted in form by the requirement of global gauge covariance.

We denote the change induced in the field variable p by an infinitesimal global gauge transformation as δ^*p and write

$$\delta^*p = O_p\delta\xi, \quad (2.4)$$

where O_p is an operator which may depend on the p. We assume that O_p is a differential matrix operator applied on $\delta\xi$ which stands for the set of n infinitesimal parameters $\delta\xi_i(x)$ of the global gauge group. Thus $\delta\xi$ is arbitrary. Similarly we have

$$\delta^*g = O_g\delta\xi. \quad (2.5)$$

We now take advantage of the arbitrary variability of $\delta\xi$ over space–time to make it vanish over the boundaries of the volume of integration Ω. This restriction simplifies the calculation without impairing the generality of the conclusions provided we keep Ω arbitrary. Assuming locality for Eqs. (2.4) and (2.5), it follows that δ^*g and δ^*p also vanish at the boundaries of Ω.

From (2.1) we have

$$\int_\Omega d^4x \frac{\delta\mathfrak{L}_G(g)}{\delta g}\delta^*g = \int d^4x \frac{\delta\mathfrak{L}_G(g)}{\delta g} O_g\delta\xi, \quad (2.6)$$

or

$$\int_\Omega d^4x O_g^H\left(\frac{\delta\mathfrak{L}_G}{\delta g}\right)\delta\xi = 0, \quad (2.7)$$

where $\delta\mathfrak{L}_G/\delta g$ is the variational derivative of $\mathfrak{L}_G$ and O_g^H is the Hermitian conjugate of O_g. In the same way, from (2.2) we have

$$\int d^4x\left[O_g^H\left(\frac{\delta\mathfrak{L}_I}{\delta g}\right)+O_p^H\left(\frac{\delta\mathfrak{L}_I}{\delta p}+\frac{\delta\mathfrak{L}_M}{\delta p}\right)\right]\delta\xi=0, \quad (2.8)$$

which, because of the arbitrariness of $\delta\xi$, gives

$$O_g^H T=-O_p^H\left(\frac{\delta\mathfrak{L}_I}{\delta p}+\frac{\delta\mathfrak{L}_M}{\delta p}\right), \quad (2.9)$$

where $T\equiv\delta\mathfrak{L}_I/\delta g$. The equations of motion that follow from the action (2.1) are

$$\delta\mathfrak{L}_G/\delta g+\delta\mathfrak{L}_I/\delta g=0, \quad (2.10)$$

and

$$\delta\mathfrak{L}_I/\delta p+\delta\mathfrak{L}_M/\delta p=0. \quad (2.11)$$

Thus when Eq. (2.11) is satisfied, Eq. (2.8) implies

$$O_g^H T=0, \quad (2.12)$$

which is traditionally called the "conservation law" of the theory. Equation (2.12), being a consequence of the equations of motion (2.11) and the global gauge-invariance property (2.3), is identically satisfied in g.

This property of global gauge covariance is by no means necessary, however, for (2.12) to be true when p fields interact with g fields; for from (2.7) we also have

$$O_g^H\delta\mathfrak{L}_G/\delta g=0 \quad (2.13)$$

identically in g. These are the "Bianchi identities" of the theory and from them and (2.10) we have that, for any solution of the dynamical problem, (2.12) has to be satisfied as a consistency requirement of (2.10). The "conservation laws" follow for any theory in which fields are coupled to a gauge field whose equations of motion satisfy a set of "Bianchi identities," whether or not the total theory is globally gauge covariant. Relations (2.12) are true conservation laws only when $O_g=\partial_\nu$. A sufficient condition for the existence of a Bianchi identity with such an O_g is that there exists a term in the Lagrangian that is a function only of $F_{\mu\nu}\equiv(\partial_\mu g_\nu-\partial_\nu g_\mu)$. This same condition assures the existence of a conserved current.

Coming back to the Utiyama-type theories, we assume that a transformation among the p dynamical variables can be found which splits them into two groups, p^{ph} and p^{in}, such that O_p becomes

$$O_p=\begin{pmatrix}-1 & 0\\ 0 & 0\end{pmatrix},$$

for

$$p=\begin{pmatrix}p^{\text{ph}}\\ p^{\text{in}}\end{pmatrix} \quad (2.14)$$

with a suitable definition for $\delta\xi$. With such a redefinition the dynamical variables are split into two groups: the p^{in} are invariants of the constant parameter gauge group and the remaining p^{ph} alone are affected by the transformations of the group. We now choose the p^{ph} to be the very descriptors of the gauge group and we shall call them "phase" variables. As the descriptors of the gauge group, the p^{ph} can be reduced to zero by a suitable finite global gauge transformation. This means that, with a suitable choice of gauge, $p^{\text{ph}}=0$ and that this set of dynamical variables drops out of the description of the system.

If one wonders how the information in the equations of motion of the phase variables is to be preserved, it is sufficient to inspect Eq. (2.9) and notice that, by virtue of (2.14), we have

$$O_g^H T=\delta\mathfrak{L}_I/\delta p^{\text{ph}}+\delta\mathfrak{L}_M/\delta p^{\text{ph}}. \quad (2.15)$$

This tells us that the "conservation laws" of the theory are just the equations of motion obtained by variation of the phase variables. In the gauge-noninvariant formulation written in terms of the variables g' and p^{in}, the physical content of the equations of motion (2.15) is a consequence of the Bianchi identities of the free gauge field.

Even though the reduction in the explicit number of dynamical variables is in principle the result of a particular choice of gauge, one cannot be certain that the physical content of the covariant Utiyama-type theory is the same as that of the corresponding noncovariant theory, if one requires that in both cases the solutions to the equations of motion be single-valued and continuous. This problem arises whenever the transformation of variables from g and p to g' and p^{in} and p^{ph} is singular.

Assume that the freedom allowed by covariance can be used to remove the explicit appearance of the "phase variables" at least in some approximation. The noncovariant theory which results has such desirable properties as: (a) all variables are physically meaningful; (b) the number of variables is fewer than in the corresponding covariant theory; (c) the gauge field can in certain cases carry all the information previously carried by the source field.

If the noncovariant theory is exactly equivalent to the corresponding covariant theory, a reinterpretation of the covariant theory is in order. If the noncovariant theory is only an approximation to the covariant theory, then the noncovariant theory with its desirable properties can be used as the basis upon which new theories are constructed. In particular, the classical noncovariant theory to be presented in Sec. 3, which exhibits certain "quantum" properties and

which does not require the introduction of the charge e, can be made the basis for a quantized theory of interactions.

The elimination of phase variables from covariant theories leads to a class of gauge-invariant theories of the type studied in this paper. Still open is the question whether any noncovariant theory can be made covariant by the introduction of enough phase variables with suitable gauge-group transformation properties.

3. MAXWELL–SCHRÖDINGER THEORY

Instead of systematically reviewing all the possible gauge-noninvariant interactions for the Maxwell field A_μ, we consider only the theory determined by the action,

$$\text{ACTION} = \int d^4x[-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \tfrac{1}{2}R^2(A^\mu A_\mu - 1) + \tfrac{1}{2}R^{,\mu}R_{,\mu}], \quad (3.1)$$

which is one of the simplest Lorentz-covariant possibilities. We have chosen the Lorentz metric to be $\eta_{00} = 1$, $\eta_{rr} = -1$, and $F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$. This theory is related to the Utiyama-type theory in which the electromagnetic field is in minimal coupling with a complex scalar Klein–Gordon field. In such a theory the global gauge group of transformations is

$$\psi' = \psi e^{i\phi(x)}, \quad (3.2)$$

$$A'_\mu = A_\mu - i\partial_\mu\phi. \quad (3.3)$$

The transition from variables ψ, ψ^* to the new variables R, S through

$$\psi = R\exp(iS) \quad (3.4)$$

is a transition to a phase variable S and a field R invariant under the group of gauge transformations. If we formally set $S = 0$ in the complex Klein–Gordon theory, we recover the form of the action expressed in Eq. (3.1). This can be achieved by a gauge transformation whenever S is a continuous and single-valued function of position. The continuity and single valuedness of ψ, however, do not imply the same for R and S. Thus, for the theory described by Eq. (3.1), a physical content different from that of the original Utiyama theory may be expected.

We have written the action (3.1) in a completely nondimensional form. If one introduces a unit of length $1/k$, Eq. (3.1) can be reexpressed in terms of

$$x'_\mu = x_\mu/k \quad (3.5)$$

as

$$\text{ACTION} = \int d^4x'k^2[-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \tfrac{1}{2}k^2R^2(A^\mu A_\mu - 1) + \tfrac{1}{2}R^{,\mu}R_{,\mu}]. \quad (3.6)$$

We could further introduce the more customary definitions for the fields

$$A'_\mu = kA_\mu, \quad R' = kR \quad (3.7)$$

and set $k = m/e$, the ratio of mass to charge. Then (3.6) becomes

$$\text{ACTION} = \int d^4x[-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \tfrac{1}{2}R^2(A^\mu A_\mu - k^2) + \tfrac{1}{2}R^{,\mu}R_{,\mu}], \quad (3.8)$$

where we have dropped all primes. This is a clear extension of Dirac's work.[3] Here the Lagrange multiplier field R^2 has been given a dynamics of its own by including the $\frac{1}{2}R^{,\mu}R_{,\mu}$ term in the Lagrangian. Thus, our theory describes all the features of classical charges, as does Dirac's theory, in the limit of negligible fluctuations of the R field. On the other hand, the connection with the complex Klein–Gordon theory and, in the nonrelativistic limit, the Schrödinger theory makes it apparent that such fluctuations will incorporate some quantum features into the description of the classical clouds of charge. We present some of these features by studying the static configurations that such clouds may assume in the presence of an external Coulomb field.

The equations of motion derived from (3.1) are

$$F^{\mu\nu}_{\ ,\nu} = R^2A^\mu, \quad (3.9)$$

$$\Box R = R(A^\mu A_\mu - 1). \quad (3.10)$$

According to Eq. (3.9), the electric charge and current densities are

$$J^\mu = R^2A^\mu. \quad (3.11)$$

They are conserved because the Bianchi identities require

$$(R^2A^\mu)_{,\mu} = F^{\mu\nu}_{\ \ \mu\nu} \equiv 0. \quad (3.12)$$

In the nonrelativistic limit, the complex Klein–Gordon theory is equivalent to the Schrödinger theory. In such a limit, Eq. (3.10) is the real part and Eq. (3.11) is the imaginary part of the Schrödinger equation written in the gauge $S = 0$. This nonrelativistic limit is obtained by setting

$$A_0 = 1 + \phi, \quad \phi \ll 1, \quad (3.13)$$

where ϕ is a first-order field, and by neglecting ϕ^2 and second time derivatives.

To determine the observables in this classical field theory we compute the energy tensor $T_{\mu\nu}$ as well as the current J^μ. This task may prove delicate in our gauge-noninvariant theory because the integrals over infinite surfaces which are usually dropped now may become important. Indeed, constant terms in the potentials

are of physical significance because of the lack of gauge invariance. We compute the $T_{\mu\nu}$ tensor by writing (3.1) in a system of curvilinear coordinates with the help of the $g^{\mu\nu}$ metric and then finding

$$T_{\mu\nu} \equiv \frac{2}{(-g)^{\frac{1}{2}}} \frac{\delta}{\delta g^{\mu\nu}} [\mathfrak{L}(-g)^{\frac{1}{2}}]. \tag{3.14}$$

By this procedure we obtain

$$T_{\mu\nu} = -F_{\mu}^{\lambda}F_{\lambda\nu} + \tfrac{1}{4}g_{\mu\nu}F^{\sigma\lambda}F_{\sigma\lambda} + R_{,\mu}R_{,\nu} - \tfrac{1}{2}g_{\mu\nu}g^{\sigma\lambda}R_{,\sigma}R_{,\lambda} + \tfrac{1}{2}R^2(2A_{\mu}A_{\nu} - g_{\mu\nu}g^{\sigma\lambda}A_{\sigma}A_{\lambda} + g_{\mu\nu}). \tag{3.15}$$

Setting $g^{\mu\nu} = \eta^{\mu\nu}$ (the Lorentz metric), we obtain the energy of the system as the positive-definite quantity

$$E = \int d^3x T_{00} = \int d^3x [\tfrac{1}{2}(E^2 + H^2 + R_{,0}R_{,0} + R_{,r}R_{,r} + R^2A_0A_0 + R^2A_rA_r + R^2)]. \tag{3.16}$$

In order to study the static configurations that the clouds of charge may assume in the presence of external fields, we proceed in the same manner after modifying the original Lagrangian by introducing the static external contravariant current vector $J^{\mu(\text{ext})} = (0, 0, 0, \rho^{\text{ext}})$ and the fixed external field $A^{(\mu\text{ext})}$ generated by such a current. For static solutions the energy is

$$E = \int d^3x [\tfrac{1}{2}A_{0,r}A_{0,r} + R^2A_0A_0], \tag{3.17}$$

where $A_0 = A_0^{(s)} + A_0^{(\text{ext})}$, and we have made use of the equations of motion after integrating the R variable by parts.

Truly static situations will arise when the electric currents **J** are zero. We therefore look for solutions in which

$$A_{\mu}^{(s)} = (0, 0, 0, 1 + C + \phi^{(\text{cl})}), \tag{3.18}$$

where we allow for a constant $1 + C$ that may be significant and where $\phi^{(\text{cl})}$ is the potential generated by the cloud of charge:

$$-\nabla^2\phi^{\text{cl}} = -A_0R^2, \tag{3.19}$$

which is one of the equations of motion. The other equation of motion

$$-R_{,rr} = R(A_0A_0 - 1) \tag{3.20}$$

reduces to the nonrelativistic time-independent Schrödinger equation in the limit in which $A_0 = 1 + \phi$, $\phi \ll 1$, and squares of ϕ are neglected. Indeed, Eq. (3.20) under these conditions becomes

$$-\tfrac{1}{2}R_{,rr} - \phi^{\text{cl}}R - \phi^{\text{ext}}R = CR \tag{3.21}$$

so that C is the eigenvalue $C(n)$ of the Schrödinger equation.

We compute the energy (3.17) corresponding to a solution $R^{(n)}$ for a cloud with total charge equal to -1, so that

$$\int d^3x R^2A_0 = 1. \tag{3.22}$$

We have, from Eq. (3.17),

$$E^{(n)} = \int d^3x [\tfrac{1}{2}\phi_{,r}^{\text{cl}}\phi_{,r}^{\text{cl}} + \phi_{,r}^{\text{cl}}\phi_{,r}^{\text{ext}} + \tfrac{1}{2}\phi_{,r}^{\text{ext}}\phi_{,r}^{\text{ext}} + R^2A_0(1 + C^{(n)} + \phi^{\text{cl}} + \phi^{\text{ext}})], \tag{3.23}$$

or, by virtue of Eqs. (3.19) and (3.22),

$$E^{(n)} = 1 + C^{(n)} - \tfrac{1}{2}\int d^3x \phi_{,r}^{\text{cl}}\phi_{,r}^{\text{cl}} + \tfrac{1}{2}\int d^3x \phi_{,r}^{\text{ext}}\phi_{,r}^{\text{ext}}. \tag{3.24}$$

This result, of general validity for static solutions, shows that the energy is equal to a constant term, the rest-mass energy, plus $C^{(n)}$, the usual energy eigenvalue for a stationary state. This eigenvalue reflects the distortion of the cloud due to Coulomb self-interaction. The next term in Eq. (3.24) subtracts the Coulomb self-energy of the cloud.

In the case of $A_0^{\text{ext}} = 1(4\pi r)$, a complete set of solutions for R in the nonrelativistic case is given by the complete set of real solutions to the hydrogen atom

$$R^{n,l,m}(r, \theta, \phi) = R^{n,l}(r)P^{l,n}(\theta)\begin{cases}\cos m\phi \\ \sin m\phi.\end{cases} \tag{3.25}$$

Thus our extension of Dirac's theory of classical electrons allows the clouds of charge some quantum features explainable in the framework of a classical and consistent theory of electromagnetism. At the same time, the difficulties encountered in Dirac's theory are removed. For example, charges at rest in a constant magnetic field are described as satisfactorily by this theory as by the Schrödinger theory. In addition, one realizes that a large constant term in the A_0 component of the gauge field is important for the inertial properties of the cloud and that a rotational motion proportional to the external magnetic field intensity is related to the diamagnetic properties of the cloud.

The full extent to which quantum features can be explained in purely classical terms is outside the scope of this paper. We wish to point out, however, that the essential quantum properties seem to be particle number, spin, and statistics rather than the features for which the quantum formalism was originally developed. It is further possible that other ideas,[7] developed in a different context, may prove fruitful

[7] D. Finkelstein, J. Math. Phys. 7, 1218 (1966).

when applied to the topological properties of such classical field solutions.

4. GRAVITATIONAL FIELD

Einstein's theory of gravitation is a theory of the Utiyama type considered in Sec. 2. It is covariant under the group of general coordinate transformations and has the metric field as the associated gauge field. We consider two cases in which the gravitational field is coupled to a source characterized by a vector field: A^μ (the electromagnetic field) and

$$\sum_i \int \frac{d\xi^{\mu(i)}}{d\tau}\, \delta^4(x - \xi^{(i)})\, d\tau$$

(the field of matter-point singularities). Because all components of these fields are phase variables, their dynamical meaning, short of coordinate conditions, is completely arbitrary. The metric field is capable of absorbing all the information usually carried by such source-field variables.

A. Gravitational–Electromagnetic Field

We assume that the A^μ vector field is timelike in character. This may seem an arbitrary restriction especially if gauge transformations $A'_\mu = A_\mu + \partial_\mu \phi$ are allowed. We have in mind, however, the gauge-noncovariant theory of electromagnetism developed in the last section, for which it can be seen that A^μ is a timelike vector field, at least in the nonrelativistic limit in which $A^\mu = \delta^\mu_0 + \phi^\mu$, $\phi^\mu \ll 1$. All components of the A^μ field are phase variables, since they can be reduced to constant values and eliminated as dynamical variables. We call such a set of constant values a "standard form." The only restriction imposed on such a standard form is that it has the timelike character of A^μ. We choose

$$A^{\mu(\text{standard})} = \delta^\mu_0. \tag{4.1}$$

It should be noted that in general a timelike vector field can be brought to a standard form only in its contravariant components. Indeed, one chooses the field to be the time-coordinate field.

If a covariant vector field could be brought to such a form, it would obviously have a vanishing ordinary curl in the new system of coordinates. Because the ordinary curl of a covariant vector is a tensor, it follows that only rotationless covariant vector fields can be brought to standard form.

Condition (4.1) can, equivalently, be imposed as a coordinate condition on the $g_{\mu\nu}$:

$$g_{0\mu} = A_\mu. \tag{4.2}$$

We conclude then that the theory determined by

$$A = \int d^4x(-g)^{\frac{1}{2}}[G + \tfrac{1}{2}g_{0\lambda,\rho}g_{0\mu,\nu}(g^{\mu\rho}g^{\nu\lambda} - g^{\nu\rho}g^{\mu\lambda}) + \tfrac{1}{4}R^2(g_{00} - 1) + \tfrac{1}{2}g^{\mu\nu}R_{,\mu}R_{,\nu}] \tag{4.3}$$

describes the gravitational plus electromagnetic properties of the classical clouds of charge found in the last section. The expression (4.3) for the action was obtained from the expression for the action (3.1), made generally covariant through the introduction of the $g_{\mu\nu}$, supplemented with the Einstein free Lagrangian G, and with coordinate condition (4.2) taken into account. The content of the theory of Sec. 2 is a consequence here of the Bianchi identities satisfied by the Einstein equations.

The coordinates are, of course, no longer arbitrary but are restricted by the theory. Their physical meaning is disclosed by (4.2), which allows us to interpret the physical meaning of the $g_{0\mu}$ components of the metric. The meaning of the coordinate frame can be made clearer by noting that the electric charge-current 4-vector is now

$$J^\mu \equiv R^2 A^\mu = R^2 \delta^\mu_0. \tag{4.4}$$

Thus we are in the frame of reference in which all electric charges are at rest.

We do not claim to have a theory of gravitation–electromagnetism in the sense of unified theories. The inclusion of electromagnetism is unsatisfactory to the extent to which the addition to the Lagrangian of the gauge noninvariant term

$$g_{0\lambda,\rho}g_{0\mu,\nu}(g^{\mu\rho}g^{\nu\lambda} - g^{\nu\rho}g^{\mu\lambda})$$

is *ad hoc*. The situation is no worse, however, than the inclusion of the ordinary Maxwell Lagrangian in the usual theory in terms of the A_μ. The advantage of the formulation in terms only of the metric is that the phase variables A_μ are eliminated and physical meaning is imparted to the components of the metric field. In the following simple case of chargeless-point singularities, we have an example of how the metric components can carry the information usually shared between the metric field and the phase variables.

B. Gravitational Field Interacting with Matter-Point Singularities

It is conceivable that this theory may arise as an idealized limit of the one considered in the previous section. An analogous coordinate condition results from choosing a frame such that the timelike 4-vector $d\xi^{\mu(i)}/d\tau$ is

$$d\xi^{\mu(i)}/d\tau = \delta^{\mu(i)}_0 \tag{4.5}$$

or

$$d\xi_\mu^{(i)}/d\tau = g_{0\mu} \quad \text{at} \quad x^\mu = \xi^{\mu(i)}, \tag{4.6}$$

namely the particles remain always at rest. The only nonvanishing contravariant component of the energy-momentum tensor is the constant

$$T^{00} = m \tag{4.7}$$

defined over the world line of the particle. The covariant tensor for the matter-point sources is then

$$T_{\mu\nu} = m g_{0\mu}(\xi) g_{0\nu}(\xi) \quad \text{for} \quad x^\mu = \xi^\mu. \tag{4.8}$$

The dynamical variables ξ^μ of the particles are now fixed constants which determine the constant position of the particles. The physical content usually associated with the ξ^μ is now carried by the values of the $g_{0\mu}$ at the position of the particles and the dynamical development is determined solely by the Bianchi identities associated with Einstein's equations.

In this formulation, in which the particle coordinates are no longer dynamical variables, the problem of motion is most easily approached through

$$T^{\mu\nu}_{;\nu} = T^{\mu\nu}_{,\nu} + \Gamma^\nu_{\beta\nu} T^{\mu\beta} + \Gamma^\mu_{\beta\lambda} T^{\beta\lambda} = 0, \tag{4.9}$$

or, remembering that $T^{\mu\nu} = m\delta^\mu_0 \delta^\nu_0$,

$$\Gamma^r_{00} = 0, \quad r \neq 0; \tag{4.10}$$

$$\Gamma^0_{00} + \Gamma^\mu_{0\mu} = 0 \quad \text{for} \quad x^\mu = \xi^{\mu(i)}, \tag{4.11}$$

or

$$\tfrac{1}{2} g^{r\mu}(2g_{0\mu,0} - g_{00,\mu}) = 0, \tag{4.12}$$

$$\tfrac{1}{2} g^{0\mu}(2g_{0\mu,0} - g_{00,\mu}) + \frac{1}{2g} g_{,0} = 0, \tag{4.13}$$

for $x^\mu = \xi^{\mu(i)}$. Because the $d\xi^\mu/d\tau$ are not independent, we can impose the further coordinate condition

$$-g g_{00} = 1, \tag{4.14}$$

which satisfies (4.11) identically by virtue of (4.12) and gives, from (4.12),

$$g_{0r,0} = \tfrac{1}{2} g_{00,r} - \tfrac{1}{2} g_{0r}(\ln g_{00})_{,0}, \tag{4.15}$$

for $x^\mu = \xi^{\mu(i)}$, as the equations of motion for the point particles. This equation can be roughly interpreted as giving the acceleration on the left-hand side equal to a force derivable from a potential and a velocity-dependent force. In the nonrelativistic limit it gives Newton's law of gravitation when g_{00} is evaluated from the Einstein equations of the same order. This formulation may be derived from the noninvariant action,

$$\text{ACTION} = \int d^4x(-g)^{\frac{1}{2}} G + \int d\tau m^{(i)} (g_{00})^{\frac{1}{2}}, \tag{4.16}$$

which results from the usual action when the coordinate conditions (4.5) and (4.14) are used.

The above formulation of the motion of gravitating singularities makes it clear that the motion of a point particle cannot be an absolute attribute. One can speak only of locations in space where there is or is not such a particle. If we label positions by means of the point particles, what is usually called their equation of motion becomes the condition 4.15, which is satisfied by the $g_{0\mu}$ along the world lines of the particles. This is in accord with the identification of (4.6), which gives the physical meaning of such $g_{0\mu}$.

A noncovariant theory, which could not be made generally covariant by the introduction of new dynamical variables, might prove even more interesting than the case considered. Whether such theories exist and whether the existence of a limit of the form of (4.7) for the vanishing of some length parameter assures all the verifiable content of a theory of space, time, and gravitation,remain open questions.

5. YANG–MILLS FIELD

In the gauge noninvariant theory of Sec. 3, the electromagnetic field is coupled to a real scalar field R; yet, the theory describes charged particles of both signs. It might be thought that the introduction of more complicated internal symmetries or, equivalently, the description of higher multiplets will necessarily require that the R field be generalized to a multicomponent field. To show that this is not the case, we briefly consider the Yang–Mills field $\mathbf{B}_\mu$.

The action,

$$\text{ACTION} = \int d^4x[-\tfrac{1}{4}\mathbf{F}_{\mu\nu} \cdot \mathbf{F}_{\mu\nu} + \tfrac{1}{2}R^2(\mathbf{B}_\mu \cdot \mathbf{B}_\mu - 1) + \tfrac{1}{2}R_{,\mu}R_{,\mu}], \tag{5.1}$$

where

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{B}_\nu - \partial_\nu \mathbf{B}_\mu - \tfrac{1}{2}(\mathbf{B}_\mu \mathbf{B}_\nu - \mathbf{B}_\nu \mathbf{B}_\mu)$$

and R is a scalar real field, describes the charge-independent interaction of a Yang–Mills field with an isospinor scalar meson. This action results from the ordinary expression when the prescription of Sec. 2 is used. We have written the usual isospinor scalar field ψ as

$$\psi = \begin{pmatrix} a + ib \\ c + id \end{pmatrix} = e^{\tau \cdot \mathbf{s}} \begin{pmatrix} R \\ 0 \end{pmatrix},$$

where $\boldsymbol{\tau}$ are the three isospinor matrices and R^2 is the isospinor invariant

$$R^2 = (\psi, \psi) = a^2 + b^2 + c^2 + d^2.$$

Substitution into the usual action with $\mathbf{s} = 0$, which can always be accomplished by a gauge transformation

if s is a single-valued function of position, leads to expression (5.1).

Although the action (5.1) is not invariant under isospinor transformations, the symmetry implied by charge independence is assured by the symmetric way in which $\mathbf{B}_\mu$ appears. Conservation of the isospin current $\mathbf{J}^\mu = \mathbf{J}^\mu_{(B)} + R^2\mathbf{B}^\mu$, where $\mathbf{J}^\mu_{(B)}$ is the isospin current contributed solely by the Yang–Mills field, is a consequence of the Bianchi identities. The contribution of the meson field to the isospin current is $\mathbf{J}^{\mu(\text{meson})} = R^2\mathbf{B}^\mu$.

The action (5.1) is the simplest and most obvious generalization of the electromagnetic theory of Sec. 3. It shows how internal symmetries of the particles and multiplet structure can be described solely by an enrichment of the properties of the gauge field. In addition, the breaking of the symmetry and the appearance of an electromagnetic axis in the theory can be accomplished for both source and gauge fields by modifying only the gauge-field action in a way previously proposed.[8]

ACKNOWLEDGMENTS

The authors are grateful to Professor D. Finkelstein and Professor Y. Aharanov for the fruitful discussions they have shared with them. One of us (R. F.) is thankful for the support received under a National Aeronautics and Space Administration Traineeship.

[8] D. Finkelstein, J. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **4**, 788 (1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Green's Functions for the One-Speed Transport Equation in Spherical Geometry

R. C. ERDMANN
Department of Engineering, University of California, Los Angeles, California
AND
C. E. SIEWERT
Department of Nuclear Engineering, North Carolina State University, Raleigh, North Carolina

(Received 4 May 1967)

Several problems in one-speed neutron transport theory for spherically symmetrical systems are discussed. The singular eigenfunction expansion technique is used to construct a solution for a specific finite-slab Green's function. This slab solution is then used to construct the finite-medium spherical Green's function by extending the point-to-plane transformation concept. For the general case, the expansion coefficients are shown to obey a Fredholm equation, and first-order solutions are obtained; however, in the infinite-medium limit the solution is represented in closed form. In addition, the solution for the angular density in an infinite-medium due to an isotropic point source is developed directly from the set of normal modes of the transport equation. A proof that the result so obtained obeys the proper source condition at the origin is given.

I. INTRODUCTION

The singular eigenfunction expansion technique that was introduced by Case has been used extensively to develop exact solutions for many problems in neutron transport theory.[1–5] In addition, several applications of the method have been made in the field of radiative transfer in stellar atmospheres.[6–8] Although the class of problems for which the Case method has been used is a broad one, the major limitation appears to be the restriction to plane geometries. The purpose of this paper is to present an extension of this method in order to solve for the Green's function in spherical geometry and thus to establish a procedure by which rigorous solutions for such problems can be obtained.

It has been shown that the integral form of the homogeneous Boltzmann equation in spherical geometry can be related to the integral equation for a corresponding problem in slab geometry.[9,10] An extension of this technique to include inhomogeneous source terms is described in this paper. Thus, several problems in spherical geometry can be solved by inspection once the slab solution is known.[9,10,11]

In addition to his work on the integral transport equation, Davison also found a set of solutions to the homogeneous version of the differential transport equation.[9,10] In a more recent work, Mitsis found two classes of eigensolutions of the transport equation in spherical geometry: one regular and the other highly singular at the origin.[12]

In Sec. II we solve the problem of a spherical-shell source in a finite medium by solving the integral equation that defines the density. Several limiting situations are investigated; and we obtain, as a special case, the density due to an isotropic point source in a finite sphere. For the finite-medium Green's function, one must solve a Fredholm equation for the expansion coefficients; hence an approximate solution, in the spirit of the "wide slab," is obtained.[13,14] In the infinite-medium limit the solution is expressed in closed form.

Although the integral equation approach is convenient for obtaining the density in the single-region problems considered here, finding the angular density, the current, or the higher moments requires further work. In order to illustrate a procedure by which the angular density can be obtained directly from the normal modes of the spherical transport equation, Sec. III is devoted to the solution of the infinite-medium point source problem. Thus, by making expansions in terms of the normal modes of the homogeneous equation and by properly determining the expansion coefficients from the boundary conditions, we are able to solve for the angular density in a manner analogous to that used in plane geometry.

[1] K. M. Case, Ann. Phys. (New York) **9**, 1 (1960).
[2] P. F. Zweifel, Michigan Memorial Phoenix Project Report, University of Michigan (1964).
[3] N. J. McCormick and I. Kuščer, J. Math. Phys. **6**, 1939 (1965).
[4] J. Mika, Nucl. Sci. Eng. **22**, 235 (1965).
[5] R. C. Erdmann, Ph.D. thesis, California Institute of Technology (1966).
[6] C. E. Siewert and P. F. Zweifel, Ann. Phys. (New York) **36**, 61 (1966).
[7] C. E. Siewert and P. F. Zweifel, J. Math. Phys. **7**, 2092 (1966).
[8] C. E. Siewert and S. K. Fraley, Ann. Phys. (New York) **43**, 338 (1967).
[9] B. Davison, Canadian Report MT-112, National Research Council of Canada, Division of Atomic Energy (1945).
[10] B. Davison, *Neutron Transport Theory* (Oxford University Press, London, 1957).
[11] A Leonard and T. W. Mullikin, Proc. Natl. Acad. Sci. **52**, 683 (1964).
[12] G. J. Mitsis, Argonne National Laboratory Report ANL-6787 (1963).
[13] M. R. Mendelson, Ph.D. thesis, University of Michigan (1964).
[14] K. M. Case, Michigan Memorial Phoenix Project Report, University of Michigan (1961).

The current and higher moments can then be obtained easily by integration. This procedure has additional merit since it may be a more useful way in which to approach the solution for multiregion problems where the integral equation approach becomes unmanageable.[15,16]

II. SOLUTION IN A FINITE SPHERE CONTAINING A SPHERICAL-SHELL SOURCE

A. Integral Equation in Spherical Geometry

We consider a single region of radius R in which a source of neutrons has been placed. The medium scatters neutrons isotropically in the laboratory system and no energy degradation of the neutrons is permitted. The source is taken to be spherically symmetric and has an isotropic emission character. Thus, we write the integral equation for the steady state neutron density as

$$\rho(r_0; r) = \int_0^R \int_0^\pi \int_0^{2\pi} r'^2\, dr' \sin\theta'\, d\theta'\, d\varphi' \times \left[\frac{c}{4\pi}\rho(r_0; r') + \frac{S}{4\pi r_0^2}\delta(r' - r_0)\right] \frac{e^{-|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|^2}, \quad 0 \le r, \quad r_0 \le R. \quad (1)$$

The total source strength in Eq. (1) is S and distances are measured in units of mean free paths. In addition, the mean number of secondary neutrons per collision is denoted by c. It is possible to relax the requirement of an isotropic emission character for the source by treating the first collision neutrons as the source in Eq. (1) and then adding in the uncollided source neutrons.[17] The point source at the center of a sphere is a limiting case of Eq. (1).

The technique used to solve Eq. (1) was suggested by Davison and has been used by Mitsis for solving critical problems.[10,12] Its basis lies in the similarity of a modified form of Eq. (1) and the integral transport equation in slab geometry. Since solutions to the latter are readily obtained, one need only determine the relationship between spherical and planar problems in order to obtain an explicit solution to Eq. (1).[11] To cast Eq. (1) into a more appropriate form, we perform the two angular integrations. This yields

$$r\rho(r_0; r) = \frac{c}{2}\int_0^R r'\rho(r_0; r') \times \{E_1(|r - r'|) - E_1(|r + r'|)\}\, dr' + \frac{S}{2r_0}\{E_1(|r - r_0|) - E_1(|r + r_0|)\}, \quad 0 \le r, \quad r_0 \le R, \quad (2)$$

where

$$E_1(x) = \int_0^1 \frac{e^{-x/u}}{y}\, dy. \quad (3)$$

Extending the range of r to $-R \le r \le R$ and demanding that $\rho(r_0; -r) = \rho(r_0; r)$ permits us to write

$$r\rho(r_0; r) = \frac{c}{2}\int_{-R}^R t\rho(r_0; t)E_1(|r - t|)\, dt + \frac{S}{2r_0}\{E_1(|r - r_0|) - E_1(|r + r_0|)\}, \quad 0 \le r_0 \le R, \quad -R \le r \le R. \quad (4)$$

In the limit as r_0 approaches zero, Eq. (4) yields, for a centrally located point source,

$$r\rho(0; r) = \frac{c}{2}\int_{-R}^R t\rho(0; t)E_1(|r - t|)\, dt + \frac{Se^{-|r|}}{r}, \quad -R \le r \le R. \quad (5)$$

Thus, by solving Eq. (4), we can obtain the point source solution by taking the limit used to construct Eq. (5).[18,19]

In the finite slab, $-R \le r \le R$, the integral equation that describes the neutron density resulting from a unit plane isotropic source at $r_0 > 0$ is

$$\phi(r_0; r) = \frac{c}{2}\int_{-R}^R \phi(r_0; t)E_1(|r - t|)\, dt + E_1(|r - r_0|). \quad (6)$$

Hence, a linear combination of slab solutions can be used to construct an integral equation identical in form to Eq. (4), i.e.,

$$[\phi(r_0; r) - \phi(-r_0; r)] = \frac{c}{2}\int_{-R}^R [\phi(r_0; t) - \phi(-r_0; t)]E_1(|r - t|)\, dt + E_1(|r - r_0|) - E_1(|r + r_0|). \quad (7)$$

Equating the dependent variables of Eqs. (4) and (7) and taking into account the source normalization, we obtain

$$r\rho(r_0; r) = (S/2r_0)[\phi(r_0; r) - \phi(-r_0; r)]. \quad (8)$$

We note from Eq. (6) that $\phi(-r_0; r) = \phi(r_0; -r)$; this allows us to write Eq. (8) in the alternate form

$$\rho(r_0; r) = (S/2r_0r)[\phi(r_0; r) - \phi(r_0; -r)], \quad 0 \le r, \quad r_0 \le R. \quad (9)$$

This result contains the point-to-plane transformation

[15] The problem of a point source in one of two dissimilar half-spaces was solved in Ref. 16; however, the total cross section was taken to be the same in the two regions.

[16] R. C. Erdmann, Trans. Am. Nucl. Soc. **9**, 443 (1966).

[17] K. M. Case, F. de Hoffmann, and G. Placzek, *Introduction to the Theory of Neutron Diffusion* (United States Government Printing Office, Washington, 1953), Vol. 1.

[18] The solution to the point source problem in a finite sphere has been obtained independently by Smith (Ref. 19). He used a transform technique similar to that used by Mitsis (Ref. 12).

[19] O. J. Smith (private communication).

for the infinite medium as a special case[20]:

$$\rho(0; r) = -\frac{S}{r}\frac{d}{dr}\phi(0; r), \quad R \to \infty. \tag{10}$$

For finite R, correction terms can be added to Eq. (10) to account for the additional leakage.

B. General Solution for $\rho(r_0; r)$

An expression for $\rho(r_0; r)$ can be obtained once $\phi(r_0; r)$ is determined; this, in turn, is found by using the method suggested by Case.[1] Since this development parallels previous work, the discussion of it is brief and will be presented mainly as an aid in defining the notation.[14]

The finite-slab Green's function is defined by

$$\mu\frac{\partial}{\partial r}\Psi(r_0; r, \mu) + \Psi(r_0; r, \mu) = \frac{c}{2}\int_{-1}^{1}\Psi(r_0; r, \mu')\,d\mu' + \delta(r - r_0), \quad r_0 > 0, \quad -R \leq r \leq R, \tag{11}$$

with

$$\Psi(r_0; \pm R, \mu) = 0, \quad \mu \lessgtr 0. \tag{12}$$

The neutron density is

$$\phi(r_0; r) = \int_{-1}^{1}\Psi(r_0; r, \mu)\,d\mu. \tag{13}$$

The general solution to Eq. (11) is

$$\Psi(r_0; r, \mu) = (A_+ \pm K_+)\phi_+(\mu)e^{-r/\nu_0} + (A_- \mp K_-)\phi_-(\mu)e^{r/\nu_0} + \int_{-1}^{1}d\nu[A(\nu) \pm K(\nu)]\phi_\nu(\mu)e^{-r/\nu}, \quad r \gtrless r_0. \tag{14}$$

Here,[14]

$$K_\pm = \frac{e^{\pm r_0/\nu_0}}{2N_+}, \tag{15a}$$

$$K(\nu) = \frac{e^{r_0/\nu_0}}{2N(\nu)}, \tag{15b}$$

$$\phi_\pm(\mu) = \frac{c\nu_0}{2}\frac{1}{\nu_0 \mp \mu}, \tag{15c}$$

$$\phi_\nu(\mu) = \frac{c\nu}{2}\frac{P}{\nu - \mu} + \lambda(\nu)\delta(\nu - \mu), \tag{15d}$$

$$\lambda(\nu) = 1 - c\nu\tanh^{-1}\nu, \tag{15e}$$

$$1 - c\nu_0\tanh^{-1}1/\nu_0 = 0, \tag{15f}$$

$$N_+ = \frac{c\nu_0}{2}\left(\frac{c\nu_0^2}{\nu_0^2 - 1} - 1\right), \tag{15g}$$

and

$$N(\nu) = \nu\left[\lambda^2(\nu) + \frac{c^2\pi^2\nu^2}{4}\right]. \tag{15h}$$

To determine the coefficients $A_\pm$ and $A(\nu)$, we apply the boundary conditions at $\pm R$. Hence

$$0 = (A_+ \pm K_+)\phi_+(\mu)e^{\mp R/\nu_0} + (A_- \mp K_-)e^{\pm R/\nu_0}\phi_-(\mu) + \int_{-1}^{1}[A(\nu) \pm K(\nu)]\phi_\nu(\mu)e^{\mp R/\nu}\,d\nu, \quad \mu \lessgtr 0. \tag{16}$$

The conditions given by Eq. (16) are sufficient to specify uniquely all of the unknown expansion coefficients; however, they are not expressible in closed form. We defer a further discussion on the evaluation of $A_\pm$ and $A(\nu)$ to Sec. IIC.

Since $\phi_\pm(\mu)$ and $\phi_\nu(\mu)$ are normalized to unity, the expression for the density is obtained immediately from Eq. (14):

$$\phi(r_0; r) = (A_+ \pm K_+)e^{-r/\nu_0} + (A_- \mp K_-)e^{r/\nu_0} + \int_{-1}^{1}[A(\nu) \pm K(\nu)]e^{-r/\nu}\,d\nu, \quad r \gtrless r_0. \tag{17}$$

Substituting Eq. (17) into Eq. (9), we obtain the solution for the density in the finite sphere:

$$\rho(r_0; r) = \frac{S}{rr_0}\left[(A_- - A_+)\sinh r/\nu_0 + \frac{1}{2N_+}\left\{\sinh\left(\frac{r + r_0}{\nu_0}\right) - \sinh\frac{|r - r_0|}{\nu_0}\right\} + \int_0^1\left\{[A(-\nu) - A(\nu)]\sinh r/\nu + \frac{1}{2N(\nu)}\left[\sinh\left(\frac{r + r_0}{\nu}\right) - \sinh\frac{|r - r_0|}{\nu}\right]\right\}d\nu\right], \quad 0 \leq r, \quad r_0 \leq R. \tag{18}$$

C. First-Order Solution for the Expansion Coefficients

The two expressions for $A_\pm$ and $A(\nu)$ given by Eq. (16) can be simplified by using the half-range orthogonality theorem proved by Kuščer, McCormick, and Summerfield, i.e.,[21]

$$\int_0^1(\nu_0 - \mu)\gamma(\mu)\phi_\nu(\mu)\phi_{\nu'}(\mu)\,d\mu = 0, \quad \nu \neq \nu'; \ \nu, \nu' > 0. \tag{19}$$

Here,

$$(\nu_0 - \mu)\gamma(\mu) = c\mu/2(1 - c)(\nu_0 + \mu)X(-\mu), \tag{20a}$$

$$X(z) = \frac{1}{1 - z}\exp\left[\frac{1}{\pi}\int_0^1\arg\Lambda^+(\mu)\frac{d\mu}{\mu - z}\right], \tag{20b}$$

and

$$\Lambda^\pm(\mu) = \lambda(\mu) \pm i(\pi c\mu/2). \tag{20c}$$

[20] K. M. Case and P. F. Zweifel, *An Introduction to Linear Transport Theory* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1967).

[21] I. Kuščer, N. J. McCormick, and G. C. Summerfield, Ann. Phys. (New York) 30, 411 (1964).

To use directly the results for the various normalization integrals and cross products given by Kuščer, McCormick, and Summerfield, we make several changes of variables in Eq. (16) to obtain the two equations (one associated with the upper signs and one with the lower signs)

$$-\int_0^1 d\nu e^{-R/\nu}\phi_{-\nu}(\mu)[A(\pm\nu) \pm K(\pm\nu)]$$
$$= (A_+ \pm K_+)\phi_\mp(\mu)e^{\mp R/\nu_0} + (A_- \mp K_-)\phi_\pm(\mu)e^{\pm R/\nu_0}$$
$$+\int_0^1 d\nu e^{R/\nu}\phi_\nu(\mu)[A(\mp\nu) \pm K(\mp\nu)], \quad \mu > 0. \quad (21)$$

If we now multiply Eq. (21) by $(\nu_0 - \mu)\gamma(\mu)\phi_+(\mu)$, integrate over μ from 0 to 1, and make use of Eqs. (A2), (A4), and (A5) of Ref. 21, we find

$$A_\mp = K_\mp - (A_\pm + K_\pm)e^{-2(R+\delta)/\nu_0} - \frac{e^{-(R+2\delta)/\nu_0}}{\nu_0 X(-\nu_0)}$$
$$\times \int_0^1 e^{-R/\nu}\nu X(-\nu)[A(\pm\nu) \pm K(\pm\nu)]\, d\nu. \quad (22a)$$

In a similar manner, we take the scalar product of Eq. (21) with $\phi_{\nu'}(\mu)$, $\nu' > 0$, and use Eqs. (A1), (A3), and (A6) of Ref. 21, to obtain

$$A(\mp\nu) = \mp K(\mp\nu) - \frac{2(A_\pm + K_\pm)}{N(\nu)\gamma(\nu)}$$
$$\times e^{-R/\nu_0}e^{-R/\nu}\nu^2 X(-\nu_0)\phi_-(\nu)\phi_+(\nu)$$
$$- \left(\frac{c\nu}{2}\right)^2 \frac{e^{-R/\nu}}{\gamma(\nu)N(\nu)} \int_0^1 d\nu' e^{-R/\nu'}$$
$$\times [A(\pm\nu') + K(\pm\nu')] \frac{\nu'(\nu_0 + \nu')X(-\nu')}{(\nu + \nu')(\nu_0 - \nu')},$$
$$\nu > 0. \quad (22b)$$

Here,

$$\frac{X(-\nu_0)}{X(\nu_0)} = -e^{-2\delta/\nu_0}. \quad (23)$$

The quantity δ is the extrapolation distance $z_0(c)$ as defined in Ref. 17.

It is not possible to solve Eqs. (22) explicitly for the coefficients $A_\pm$ and $A(\nu)$. However, one can reduce these expressions to a Fredholm equation for $A(\nu)$; this suggests that an iterative type solution could be obtained.[22] Since an iterative approach has been used successfully for similar problems, we proceed in this manner.[12,13]

Firstly, we neglect terms of order e^{-R} in Eq. (22b) to obtain $A(\mp\nu) = \mp K(\mp\nu)$. This result is then substituted into Eq. (22a) to give $A_\pm$ accurate to first order in e^{-R}. Finally, when this expression is entered back into Eq. (22b), we find $A(\pm\nu)$ correct to first order. Thus

$$A_\pm = [1 - e^{-4(R+\delta)/\nu_0}]^{-1}\Big\{K_\pm(1 + e^{-4(R+\delta)/\nu_0})$$
$$- 2K_\mp e^{-2(R+\delta)/\nu_0} + \int_0^1 d\nu e^{-R/\nu}\nu X(-\nu)$$
$$\times \frac{2e^{-R/\nu_0}}{\nu_0 X(\nu_0)}\left[\mp K(\mp\nu) \mp K(\pm\nu)e^{-2(R+\delta)/\nu_0}\right]\Big\} \quad (24a)$$

and

$$A(\mp\nu) = \mp K(\mp\nu) - [1 - e^{-4(R+\delta)/\nu_0}]^{-1}$$
$$\times [K_\pm - K_\mp e^{-2(R+\delta)/\nu_0}][1 - c]$$
$$\times \left[\frac{2c\nu X(-\nu)}{N(\nu)}\nu_0^2 X(-\nu_0)e^{-R/\nu}e^{-R/\nu_0}\right]. \quad (24b)$$

When Eqs. (24) are substituted into Eq. (18), we find that the first-order solution for the neutron density becomes

$$\rho(r_0; r)$$
$$= \frac{S}{rr_0}\left[\frac{\sinh((R + \delta - r_0)/\nu_0)}{N_+ \sinh((R + \delta)/\nu_0)} \sinh r/\nu_0\right.$$
$$+ \int_0^1 \frac{d\nu}{N(\nu)} e^{-r_0/\nu} \sinh r/\nu$$
$$+ \frac{e^{\delta/\nu_0}}{\nu_0 X(\nu_0) \sinh((R + \delta)/\nu_0)} \int_0^1 d\nu \frac{\nu X(-\nu)}{N(\nu)} e^{-R/\nu}$$
$$\left.\times \left\{\sinh\frac{r_0}{\nu} \sinh\frac{r}{\nu_0} + \sinh\frac{r_0}{\nu_0} \sinh\frac{r}{\nu}\right\}\right],$$
$$0 \le r < r_0 \le R. \quad (25)$$

The result for $\rho(r_0; r)$, $r > r_0$, is obtained from Eq. (25) by interchanging r and r_0.

An estimate of the accuracy of this "wide-sphere" result can be obtained from comparisons with similar approximation procedures. Mendelson solves several plane geometry problems in the "wide-slab" approximation and shows that for slab widths greater than two mean free paths, the exact and wide-slab solutions are not discernible.[13] Since most spheres are at least four mean free paths in diameter, the "wide-sphere" analysis presented here should yield very accurate descriptions of the density.

D. Special Cases

The results described by Eq. (25) contain essentially all of the physics for solutions of the one-speed Boltzmann equation in spherical geometry.

Interchanging r and r_0 in Eq. (25) and taking the limit as r_0 approaches zero, we find the finite-medium solution (to first order) for a centrally located point

[22] The coefficients $A_\pm$ could then be found in terms of $A(\nu)$.

source[11]; thus

$$\rho(0;r)=\frac{S}{r}\left[\frac{\sinh((R+\delta-r)/\nu_0)}{\nu_0 N_+\sinh((R+\delta)/\nu_0)}+\int_0^1\frac{e^{-r/\nu}}{\nu N(\nu)}\,d\nu+\frac{e^{\delta/\nu_0}}{\nu_0 X(\nu_0)\sinh((R+\delta)/\nu_0)}\int_0^1\frac{\nu X(-\nu)e^{-R/\nu}}{N(\nu)}\times\left\{\frac{1}{\nu_0}\sinh\frac{r}{\nu}+\frac{1}{\nu}\sinh\frac{r}{\nu_0}\right\}d\nu\right]. \quad (26)$$

Taking the limit as R approaches infinity ($c<1$) in Eq. (25), we find

$$\rho_\infty(r_0;r)=\frac{S}{rr_0}\left(\frac{e^{-r_0/\nu_0}\sinh r/\nu_0}{N_+}+\int_0^1\frac{e^{-r_0/\nu}\sinh r/\nu}{N(\nu)}\,d\nu\right),\quad r<r_0, \quad (27a)$$

and

$$\rho_\infty(r_0;r)=\frac{S}{rr_0}\left(\frac{e^{-r/\nu_0}\sinh r_0/\nu_0}{N_+}+\int_0^1\frac{e^{-r/\nu}\sinh r_0/\nu}{N(\nu)}\,d\nu\right),\quad r>r_0. \quad (27b)$$

If one considers only the discrete part of Eq. (25), it is found to satisfy the Helmholtz equation

$$\nabla^2\rho(r_0;r)-\frac{1}{\nu_0^2}\rho(r_0;r)=-\frac{S}{\nu_0 N_+ r_0^2}\delta(r-r_0), \quad (28)$$

and the boundary condition

$$\rho(r_0;R+\delta)=0. \quad (29)$$

This is just the equation one would solve in the diffusion theory approximation to the Boltzmann equation.[23,24,25]

For a medium with no absorption ($c=1$), Eq. (25) reduces to

$$\rho(r_0;r)=S\left[\frac{3}{r_0}-\frac{3}{R+\delta}+\frac{1}{rr_0}\int_0^1\frac{e^{-r_0/\nu}}{N(\nu)}\sinh\frac{r}{\nu}\,d\nu-\frac{1}{R+\delta}\int_0^1\frac{\nu X(-\nu)e^{-R/\nu}}{N(\nu)}\times\left(\frac{1}{r_0}\sinh\frac{r_0}{\nu}+\frac{1}{r}\sinh\frac{r}{\nu}\right)d\nu\right],\quad 0\le r\le r_0\le R. \quad (30)$$

We note, also, that Eq. (25) can be used to describe the density in a multiplying sphere ($c>1$) provided, of course, that the sphere is subcritical. Taking cognizance of the fact that ν_0 is imaginary ($\eta_0\overset{\Delta}{=}|\nu_0|$) for $c>1$, we write Eq. (25) as

$$\rho(r_0;r)=\frac{S}{rr_0}\left[\frac{\sin((R+\delta-r_0)/\eta_0)\sin(r/\eta_0)}{|N_+|\sin((R+\delta)/\eta_0)}+\int_0^1\frac{e^{-r_0/\nu}}{N(\nu)}\sinh\frac{r}{\nu}\,d\nu+\frac{e^{\delta/\nu_0}}{\nu_0 X(\nu_0)\sin((R+\delta)/\eta_0)}\int_0^1\frac{\nu X(-\nu)e^{-R/\nu}}{N(\nu)}\times\left\{\sinh\frac{r_0}{\nu}\sin\frac{r}{\eta_0}+\sin\frac{r_0}{\eta_0}\sinh\frac{r}{\nu}\right\}d\nu\right],\quad 0\le r\le r_0\le R. \quad (31)$$

Again, the results for $r>r_0$ are obtained by interchanging r and r_0 in Eqs. (30) and (31).

III. INFINITE-MEDIUM GREEN'S FUNCTION AS OBTAINED FROM THE SPHERICAL NORMAL MODES

A. General Analysis

In the previous section the solution for a spherical-shell source in a finite medium was obtained. The procedure there was to solve the integral equation for the density; the angular density was then available by integration over free-flight paths. This approach obviously has merit since it generates solutions for the classical single-region problems in spherical geometry. However, it would be satisfying to be able to construct solutions directly from the normal modes of the transport equation. Then, by applying the proper boundary conditions, one would attempt to determine the expansion coefficients, perhaps in a manner analogous to the Case technique that has been used in plane geometry. If such a procedure could be established, multidimensional problems or problems with more than one mean free path might become amenable to solution.[15]

With the above philosophy in mind, we prescribe a method by which the infinite-medium Green's function can be obtained from the normal modes of the homogeneous transport equation in spherical geometry. The Green's function considered here is the solution of

$$\mu\frac{\partial}{\partial r}\Psi(r,\mu)+\frac{(1-\mu^2)}{r}\frac{\partial}{\partial\mu}\Psi(r,\mu)+\Psi(r,\mu)=\frac{c}{2}\int_{-1}^1\Psi(r,\mu)\,d\mu+\frac{\delta(r)}{4\pi r^2} \quad (32)$$

subject to the constraint that $r^2\Psi(r,\mu)$ must vanish as r increases without bound.[26] In the usual manner, we replace the source term in Eq. (32) by an equivalent

[23] It should be noted, however, that the diffusion parameters are improvements to the usual ones (Refs. 24 and 25).

[24] R. L. Murray, *Nuclear Reactor Physics* (Prentice-Hall. Englewood Cliffs, N.J. 1957).

[25] A. M. Weinberg and E. P. Wigner, *The Physical Theory of Neutron Chain Reactors* (University of Chicago Press, Chicago, 1958).

[26] Obviously, the medium must be nonmultiplying ($c<1$).

boundary condition, i.e.,[27]

$$\lim_{r\to 0} 4\pi r^2\Psi(r,\mu) = \delta(1-\mu). \tag{33}$$

Thus the density and current satisfy, respectively, the following:

$$\lim_{r\to 0} 4\pi r^2\rho(r) = 1, \tag{34a}$$

$$\lim_{r\to 0} 4\pi r^2 j(r) = 1. \tag{34b}$$

A set of normal modes for the homogeneous spherical transport equation was obtained by Davison.[9] Mitsis, in a more recent study, discussed this same set.[12] There are several interesting aspects to the manner in which Mitsis obtained his results; we simply state the results by expanding $\Psi(r,\mu)$ in terms of this basis set. Thus, we write

$$\begin{aligned}\Psi(r,\mu) &= \sum_{m=0}^{\infty}\left(\frac{2m+1}{2}\right)P_m(\mu)\\ &\times\Big\{c\nu_0 Q_m(\nu_0)[A_+k_m(r/\nu_0)+(-1)^m B_+ i_m(r/\nu_0)]\\ &+\int_0^1[c\nu Q_m(\nu)+\lambda(\nu)P_m(\nu)]\\ &\times[A(\nu)k_m(r/\nu)+(-1)^m B(\nu)i_m(r/\nu)]\,d\nu\Big\}.\end{aligned} \tag{35}$$

Here A_+, B_+, $A(\nu)$, and $B(\nu)$ are the arbitrary expansion coefficients; $P_m(\mu)$ and $Q_m(\nu)$ are, respectively, Legendre polynomials and the Legendre functions of the second kind. Also,

$$k_m(x) = (\pi/2x)^{\frac{1}{2}}K_{m+\frac{1}{2}}(x) \tag{36a}$$

and

$$i_m(x) = (\pi/2x)^{\frac{1}{2}}I_{m+\frac{1}{2}}(x). \tag{36b}$$

Following the notation of Watson, we have used $K_{m+\frac{1}{2}}$ and $I_{m+\frac{1}{2}}$ to denote the modified Bessel functions.[28] The fact that $I_{m+\frac{1}{2}}(x)$ diverges at infinity leads us immediately to equate B_+ and $B(\nu)$ to zero. Thus

$$\begin{aligned}\Psi(r,\mu) = \sum_{m=0}^{\infty}\tfrac{1}{2}(2m+1)P_m(\mu)\Big\{&A_+c\nu_0Q_m(\nu_0)k_m(r/\nu_0)\\ &+\int_0^1 A(\nu)[c\nu Q_m(\nu)+\lambda(\nu)P_m(\nu)]k_m(r/\nu)\,d\nu\Big\}.\end{aligned} \tag{37}$$

The expression for $\Psi(r,\mu)$ given by Eq. (37) has the correct behavior at infinity. The expansion coefficients, A_+ and $A(\nu)$, must be determined, therefore, so that the boundary condition at the origin, Eq. (33), is satisfied. For the sake of brevity, we introduce the notation

$$\begin{aligned}S_m(r) \triangleq\ &A_+c\nu_0Q_m(\nu_0)k_m(r/\nu_0)\\ &+\int_0^1 A(\nu)[c\nu Q_m(\nu)+\lambda(\nu)P_m(\nu)]k_m(r/\nu)\,d\nu.\end{aligned} \tag{38}$$

Thus,

$$\Psi(r,\mu) = \sum_{m=0}^{\infty}\tfrac{1}{2}(2m+1)P_m(\mu)S_m(r); \tag{39}$$

the source condition can be stated, therefore, as

$$\lim_{r\to 0} 4\pi r^2 S_m(r) = 1. \tag{40}$$

There are several interesting aspects of Eq. (40). We note that the expression must be true for *all* m. Since $k_m(r)$ diverges as r^{-m-1} in the limit of r tending to zero, it is not obvious how $r^2S_m(r)$ could exist in that limit. There must be, therefore, a very subtle interrelation between A_+ and $A(\nu)$.[29] The procedure we use here is firstly to determine A_+ and $A(\nu)$ such that Eq. (40) is satisfied for the *particular choice* of $m = 1$. This insures that the angular density will satisfy the "weak" or current boundary condition, i.e.,

$$\lim_{r\to 0} 4\pi r^2 j(r) = 1;$$

however, the complete boundary condition is stated by Eq. (33). Thus the $m = 1$ condition is necessary but *not* sufficient. In order to prove that the expansion coefficients, as determined from the "weak" boundary condition, are the correct ones (i.e., that Eq. (40) is valid for *all* m) is a formidable task; we prefer to devote section IIIB to this proof.

For the case $m = 1$, we must satisfy

$$\frac{1}{2\pi^2} = A_+\nu_0^2c\nu_0Q_1(\nu_0) + \int_0^1 A(\nu)\nu^2[c\nu Q_1(\nu)+\nu\lambda(\nu)]\,d\nu. \tag{41}$$

The form of Eq. (41) is suggestive of an expansion in terms of the eigenfunctions used by Case for problems in plane geometry.[1] We pursue the point further by considering the following full-range expansion:

$$\begin{aligned}\frac{1}{2\pi^2\mu} &= A_+\nu_0^2[\phi_+(\mu)-\phi_-(\mu)]\\ &+\int_0^1 A(\nu)\nu^2[\phi_\nu(\mu)-\phi_{-\nu}(\mu)]\,d\nu,\quad \mu\in[-1,1].\end{aligned} \tag{42}$$

The validity of this expansion follows from the full-range completeness theorem proved by Case.[1] Since Eq. (42) is a valid expansion, the coefficients, A_+ and

[27] One of the authors (C. E. S.) is indebted to Dr. Z. Akcasu and Dr. G. C. Summerfield for a discussion of this point.

[28] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, Cambridge, 1945).

[29] That this is true can be seen immediately by considering the cases $m = 0$, 1, and 2.

$A(\nu)$, can be determined immediately by taking full-range scalar products.[2] We find

$$A_+ = \frac{1}{2\pi^2 \nu_0^2 N_+} \tag{43a}$$

and

$$A(\nu) = \frac{1}{2\pi^2 \nu^3}\, g(c, \nu), \tag{43b}$$

where

$$g(c, \nu) = \frac{1}{\lambda^2(\nu) + (\tfrac{1}{2}c\nu\pi)^2}. \tag{44}$$

If one multiplies Eq. (42) by μ and integrates over μ from -1 to 1, the following results:

$$\frac{1}{2\pi^2} = A_+\nu_0^2 \int_{-1}^{1} \mu\phi_+(\mu)\, d\mu + \int_0^1 A(\nu)\nu^2 \int_{-1}^{1} \mu\phi_\nu(\mu)\, d\mu\, d\nu; \tag{45}$$

this is identical with Eq. (41), and thus Eqs. (43) do, in fact, satisfy the weak boundary condition.

Since the expansion coefficients have been determined (with the *proviso* that the necessary rigor is to be given in Sec. IIIB), the solution is complete; it can be written as

$$\Psi(r, \mu) = \sum_{m=0}^{\infty} \left(\frac{2m+1}{4\pi^2}\right) P_m(\mu) \left\{ \frac{c\nu_0 Q_m(\nu_0)}{\nu_0^2 N_+}\, k_m\left(\frac{r}{\nu_0}\right) + \int_0^1 g(c, \nu)[c\nu Q_m(\nu) + \lambda(\nu) P_m(\nu)] k_m\left(\frac{r}{\nu}\right) \frac{d\nu}{\nu^3} \right\}. \tag{46}$$

The density and current can be obtained from Eq. (46) by inspection; we find

$$\rho(r) = \frac{1}{4\pi r} \left[\frac{1}{\nu_0 N_+}\, e^{-r/\nu_0} + \int_0^1 g(c, \nu) e^{-r/\nu} \frac{d\nu}{\nu^2} \right] \tag{47a}$$

and

$$j(r) = \frac{1-c}{4\pi r} \left[\frac{1}{N_+} \left(1 + \frac{\nu_0}{r}\right) e^{-r/\nu_0} + \int_0^1 g(c, \nu) \left(1 + \frac{\nu}{r}\right) e^{-r/\nu} \frac{d\nu}{\nu} \right]. \tag{47b}$$

In addition, the higher moments of $\Psi(r, \mu)$, if so desired, are obtained trivially from Eq. (46).

It is interesting to note the form of $\Psi(r, \mu)$ in the limit as c tends to zero. It is a simple matter to obtain

$$\lim_{c \to 0} \Psi(r, \mu) \triangleq \Psi_0(r, \mu) = \frac{1}{2\pi^2} \sum_{m=0}^{\infty} \left(\frac{2m+1}{2}\right) P_m(\mu) G_m(r), \tag{48}$$

where

$$G_m(r) \triangleq \int_0^1 k_m\left(\frac{r}{\nu}\right) P_m(\nu) \frac{d\nu}{\nu^3}. \tag{49}$$

The integral above has been evaluated by Harrington, Siewert, and Murray.[30] They found

$$G_m(r) = \frac{\pi}{2r^2}\, e^{-r}. \tag{50}$$

Note that $G_m(r)$ is independent of m. When Eq. (50) is substituted into Eq. (48), we obtain the usual result, i.e.,[9]

$$\Psi_0(r, \mu) = \frac{e^{-r}}{4\pi r^2}\, \delta(1 - \mu). \tag{51}$$

This, of course, represents the uncollided angular density.

B. Boundary Condition at the Origin

The expansion coefficients A_+ and $A(\nu)$ were determined in the previous work by noting the similarity between the current boundary condition and a suitably chosen full-range expansion in terms of the eigenfunctions $\phi_\nu(\mu)$. We would like now to discuss further the analysis of the source condition and to prove that Eqs. (43) are the correct solutions for the expansion coefficients.

We have shown that A_+ and $A(\nu)$ are such that the "weak" boundary condition is satisfied, i.e., these expansion coefficients are solutions to Eq. (41); however, they are *not* the only possible solutions to Eq. (41). This equation obviously has no *unique* solution. We must keep in mind, however, that Eq. (41) is only one in an infinite set of conditions that must be satisfied (i.e., we must consider *all* m).

The necessary and sufficient condition that the complete boundary condition at the origin, viz.,

$$\lim_{r \to 0} 4\pi r^2 \Psi(r, \mu) = \delta(1 - \mu), \tag{52}$$

be satisfied is that

$$\lim_{r \to 0} 4\pi r^2 S_m(r) = 1, \tag{53}$$

where

$$S_m(r) \triangleq A_+ T_m(\nu_0) k_m\left(\frac{r}{\nu_0}\right) + \int_0^1 A(\nu) T_m(\nu) k_m\left(\frac{r}{\nu}\right) d\nu. \tag{54}$$

We have written Eq. (54) in the more tractable form by defining[12]

$$T_m(\nu_0) \triangleq c\nu_0 Q_m(\nu_0) \tag{55}$$

and

$$T_m(\nu) \triangleq c\nu Q_m(\nu) + \lambda(\nu) P_m(\nu). \tag{56}$$

Here $T_m(x)$ are mth degree polynomials of x; they satisfy the same recursion relation as the Legendre

[30] W. J. Harrington, C. E. Siewert, and R. L. Murray (to be published).

polynomials, i.e.,

$$mT_m(x) = (2m-1)xT_{m-1}(x) - (m-1)T_{m-2}(x); \tag{57}$$

they begin differently, however:

$$T_0(x) = 1, \quad T_1(x) = x(1-c). \tag{58}$$

We note that the $T_m(x)$ reduce to the $P_m(x)$ as c vanishes.

The method used to prove Eq. (53) is that of mathematical induction. We verify the validity of this condition for $m = 0, 1$, assume it to be true for $m = k-1, k-2$, and then deduce that it must be true for $k = m$. Firstly then, we must prove the $m = 0$ condition; this takes the form

$$\lim_{r\to 0} 2\pi^2 r \int_0^1 A(\nu)\nu e^{-r/\nu}\, d\nu = 1, \tag{59}$$

where

$$A(\nu) = g(c,\nu)/2\pi^2\nu^3. \tag{60}$$

For $g(c,\nu)$ we use the power series given by Case, de Hoffmann, and Plazcek, i.e.,[17]

$$g(c,\nu) = 1 + \sum_{\beta=1}^{\infty} \Gamma_\beta \nu^{2\beta}. \tag{61}$$

Thus Eq. (59) can be written as

$$\lim_{r\to 0} r\left[E_0(r) + \sum_{\beta=1}^{\infty} \Gamma_\beta E_{2\beta}(r)\right] = 1, \tag{62}$$

where $E_n(r)$ denotes the exponential integral function.[17] Clearly, Eq. (62) is satisfied. The $m = 1$ condition takes the form

$$\lim_{r\to 0}\left[2\pi^2\nu_0^3(1-c)A_+ + 2\pi^2(1-c)\int_0^1 A(\nu)\nu^3\, d\nu\right] = 1. \tag{63}$$

If we substitute the expressions for the expansion coefficients into Eq. (63), we find

$$\int_0^1 g(c,\nu)\, d\nu = \frac{1}{1-c} - \frac{\nu_0}{N_+}. \tag{64}$$

This result is given explicitly by Case *et al.*; Eq. (63) is thus valid.

The condition to be proved, Eq. (53), has been verified for $m = 0, 1$; we proceed, therefore, with the inductive proof for arbitrary m. If we utilize the recursion and differentiation formulas for the $k_m(x)$, viz.,[31]

$$k_m\left(\frac{r}{x}\right) = (2m-1)\frac{x}{r}k_{m-1}\left(\frac{r}{x}\right) + k_{m-2}\left(\frac{r}{x}\right) \tag{65a}$$

and

$$\frac{d}{dr}\left[r^{m+1}k_m\left(\frac{r}{x}\right)\right] = -\frac{r^{m+1}}{x}k_{m-1}\left(\frac{r}{x}\right), \tag{65b}$$

as well as Eq. (57), we obtain

$$m\frac{d}{dr}[r^{m+1}S_m(r)] = -(m-1)\frac{d}{dr}[r^{m+1}S_{m-2}(r)] - (2m-1)r^{m+1} \times S_{m-1}(r) + (2m-1)(m-1)r^m S_{m-2}(r). \tag{66}$$

Integrating Eq. (66) from 0 to r, we find[32]

$$mr^{m+1}S_m(r) - C_m = -(m-1)r^{m+1}S_{m-2}(r) - \lim_{r\to 0} r^{m+1}S_{m-2}(r) - (2m-1)\int_0^r t^{m+1}S_{m-1}(t)\, dt + (2m-1)(m-1)\int_0^r t^m S_{m-2}(t)\, dt. \tag{67}$$

We now make the inductive assumption that for all k, $0 \le k \le m-1$, $m \ge 2$,

$$\lim_{r\to 0} 4\pi r^2 S_k(r) = 1. \tag{68}$$

Thus the limit term in Eq. (67) is zero. The quantity C_m is a constant of integration and, as will be proved, must also be zero. We write, therefore, Eq. (67) as

$$mr^2 S_m(r) = -(m-1)r^2 S_{m-2}(r) - (2m-1)\frac{\int_0^r t^{m+1}S_{m-1}(t)\, dt}{r^{m-1}} + (2m-1)(m-1)\frac{\int_0^r t^m S_{m-2}(t)\, dt}{r^{m-1}} \tag{69}$$

If, in Eq. (69), we take the limit as r approaches zero and apply L'Hospital's rule to evaluate the two indeterminate forms,

$$\lim_{r\to 0}\frac{\int_0^r t^{m+1}S_{m-1}(t)\, dt}{r^{m-1}} = 0 \tag{70a}$$

and

$$\lim_{r\to 0}\frac{\int_0^r t^m S_{m-2}(t)\, dt}{r^{m-1}} = \frac{1}{4\pi(m-1)}, \tag{70b}$$

we obtain the desired result, viz.,

$$\lim_{r\to 0} 4\pi r^2 S_m(r) = 1. \tag{71}$$

[31] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions* (U.S. Department of Commerce, National Bureau of Standards, Washington, 1964), Appl. Math. Ser. 55, Chap. 10.

[32] This approach was suggested by W. J. Harrington.

The proof that the integration constant C_m is zero provides the bridge for the saltus used to develop Eq. (71). This proof not only is necessary, but has the additional merit that it removes the ambiguity associated with the uniqueness of the method used in Sec. IIIA to determine A_+ and $A(\nu)$. We note from Eq. (67) that

$$C_m = m \lim_{r\to 0} r^{m+1} S_m(r). \tag{72}$$

Returning to the definition of $S_m(r)$, Eq. (54), and using the explicit form of the Bessel functions, i.e.,[28]

$$k_m(r) = \frac{\pi e^{-r}}{2r} \sum_{\alpha=0}^{m} W_\alpha^m r^{-\alpha}, \tag{73}$$

we note that Eq. (72) can be written as

$$C_m = m W_m^m \frac{\pi}{2}\left[A_+ \nu_0^{m+1} T_m(\nu_0) + \int_0^1 A(\nu)\nu^{m+1} T_m(\nu)\, d\nu \right]. \tag{74}$$

To show that C_m is zero, we must prove that the term in brackets on the right-hand side of Eq. (74) vanishes. Thus, multiplying Eq. (42) by $\mu^{m-1}P_m(\mu)$ and integrating over all μ, we obtain (after a change of variables)

$$\frac{1}{4\pi^2}\int_{-1}^{1} d\mu \mu^{m-2} P_m(\mu) = A_+ \frac{c\nu_0^3}{2}\int_{-1}^{1} d\mu \mu^{m-1} \frac{P_m(\mu)}{\nu_0 - \mu} + \int_0^1 d\nu A(\nu)\nu^2 \int_{-1}^{1} d\mu \mu^{m-1} P_m(\mu)\phi_\nu(\mu). \tag{75}$$

The left-hand side of Eq. (75) is zero obviously ($m > 1$). Noting that

$$\frac{\mu^{m-1}}{\mu - \nu_0} = \mu^{m-2} + \nu_0\mu^{m-3} + \nu_0^2\mu^{m-4} + \cdots \frac{\nu_0^{m-1}}{\mu - \nu_0}, \tag{76}$$

we rewrite Eq. (75) as

$$A_+ \frac{c\nu_0^3}{2}\int_{-1}^{1} d\mu \left[-\mu^{m-2} - \nu_0\mu^{m-3} - \cdots + \frac{\nu_0^{m-1}}{\nu_0 - \mu}\right] P_m(\mu) + \int_0^1 d\nu A(\nu)\nu^2 \int_{-1}^{1} d\mu \mu^{m-1} P_m(\mu)\phi_\nu(\mu) = 0. \tag{77}$$

The continuum term in Eq. (77) also can be decomposed in the manner indicated by Eq. (76). Since to do so only complicates the notation, we do not explicitly write it out. The fact that the Legendre polynomials are orthogonal can be used again to reduce Eq. (77) to

$$A_+ \frac{c\nu_0^{m+2}}{2}\int_{-1}^{1} d\mu \frac{P_m(\mu)}{\nu_0 - \mu} + \int_0^1 d\nu A(\nu)\nu^{m+1} \int_{-1}^{1} d\mu P_m(\mu)\phi_\nu(\mu) = 0. \tag{78}$$

The final result is thus obtained, viz.,

$$A_+ \nu_0^{m+1} T_m(\nu_0) + \int_0^1 A(\nu)\nu^{m+1} T_m(\nu)\, d\nu = 0, \tag{79}$$

or

$$C_m \equiv 0. \tag{80}$$

ACKNOWLEDGMENTS

One of the authors (C. E. S.) would like to acknowledge the helpful discussions regarding Sec. III of this paper that were held with Dr. Z. Akcasu, Dr. G. C. Summerfield, Dr. R. Zelazny, and Dr. P. F. Zweifel at The University of Michigan; also, appreciation for many useful suggestions is extended to O. J. Smith and Dr. W. J. Harrington, and Dr. R. L. Murray of North Carolina State University.

This work was supported in part by the National Science Foundation.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Values of the Potential and Its Derivatives at the Origin in Terms of the S-Wave Phase Shift and Bound-State Parameters

F. CALOGERO
Istituto di Fisica dell'Università di Roma, Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy
AND
A. DEGASPERIS*
Istituto di Fisica dell'Università di Roma, Rome, Italy

(Received 5 June 1967)

The framework of this paper is the quantum-mechanical theory of nonrelativistic scattering of a particle by a spherically symmetrical, local, and energy-independent potential. The principle results are explicit expressions for the values at the origin of the potential and of its derivatives in terms of the S-wave scattering phase shift and of the parameters of the S-wave bound states (if any). Previous work in this direction is reviewed, and different routes to the establishing of these results are discussed. Some remarks concerning the characterization of S-wave physically equivalent potentials (i.e., potentials having the same S-wave phase shift and S-wave binding energies) are presented. The properties of the S-wave phase shift and bound state parameters which are necessary and sufficient to ensure that the corresponding potential be an even function of r are ascertained. The analysis is restricted to potentials represented by holomorphic functions of r.

1. INTRODUCTION

This paper is the first one of a series devoted to the derivation and discussion of exact, explicit relationships between the parameters which describe the scattering and the interaction which causes it. The philosophy of this approach is closer to the "inverse problem" of scattering theory than to the "direct problem"; however, our emphasis is on the derivation of *explicit* exact relationships rather than the discussion of general methods of solution and of existence and uniqueness proofs (although some new results in this direction are also presented).

It is our general impression that the potentialities of the approach we advocate have not yet been fully explored mainly because the general frame of mind of theoretical physicists investigating collision theory has been predominantly based on the "direct" version of the scattering problem. This fact undoubtedly has a historical, rather than a practical, justification; it was certainly proper so long as the interaction being investigated was mainly of electromagnetic origin, and therefore *a priori* essentially known; it appears much less reasonable in the realm of nuclear and elementary particle physics, where the interaction is *a priori* unknown and the inference of its nature from the experimental information is our principal aim.

In this paper we restrict the analysis to the simplest scattering problem: S-wave scattering on a smooth finite potential. In this case the interaction is completely described by a single function of the modulus of the radius vector $\mathbf{r}$, the potential $V(r)$, which is by assumption holomorphic (and vanishes asymptotically); thus the interaction is completely specified by the values of the potential and all its derivatives at the origin. On the other hand, the physical phenomenon is described by a single function, the scattering phase shift $\eta(k)$, of the modulus of the linear momentum k, having simple analyticity and asymptotic properties, and, in addition, by the parameters of the (S-wave) bound states (if any). Our main results are explicit expressions for the values of the potential and its derivatives at the origin in terms of the phase shift and the parameters of the bound states. In view of the very special nature of the interaction considered in this paper, we realize that the interest of these results lies more in their mathematical value than in their applicability to the physical world. We believe, however, that the extension of this approach to more general interactions will eventually lead to results which are relevant for direct phenomenological applications in atomic, nuclear, and elementary particle physics.

Results such as those obtained below have already appeared in the literature. In some sense (which will be made clear in the following), Levinson's theorem[1] already is an example. But the first real result of this type was obtained simultaneously and independently by Newton[2] and Faddeev,[3] and consists of the following exact relationship between the value of the potential at the origin, the scattering phase shift, and

* Supported by a CNR Fellowship for 1967.

[1] N. Levinson, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. **25**, No. 9 (1949).
[2] R. G. Newton, Phys. Rev. **101**, 1588 (1956).
[3] L. D. Faddeev, Dokl. Akad. Nauk SSSR **115**, 878 (1957).

the (negative) bound-state energies:

$$V(0) = -\frac{8}{\pi}\int_0^\infty dk\, k\frac{d}{dk}[k\eta(k)] - 4\sum_n E_n. \quad (1.1)$$

Subsequently, analogous expressions for the second derivative of the potential in the origin have been derived by Buslaev and Faddeev,[4] by Percival,[5] and by Roberts[6] (who used previous results of Verde[7] on the asymptotic behavior of the scattering phase shift). The methods employed by these authors, which are reviewed below, do not yield an expression for the value at the origin of the first derivative of the potential, nor for the subsequent derivatives, except in the special case of an even potential (in which case the explicit expressions for the fourth and sixth derivatives are easily obtained from the explicit results of Percival[5]). Note, however, that the conditions which a scattering phase shift must satisfy in order that there exist a corresponding potential which is even are far from trivial and cannot be investigated on the basis of their approaches. The principal achievement of the present paper is to introduce methods which make it possible to derive explicit expressions (in terms of the phase shift and of the bound state parameters) for all the derivatives of the potential evaluated at the origin.

It should be mentioned that certain results obtained in the papers quoted above extend to more general interactions than the case considered in this paper. We also note that there are a few other papers[8] where the same approach is used, but whose conclusions are not discussed here because they consider cases different from that treated in this paper.

In Sec. 2 (and Appendices A and B) we extend Newton's method,[2] which is based on the Gel'fand–Levitan approach to the inverse scattering problem,[9] and we derive explicit expressions for the potential and its first derivative at the origin. Although the method introduced also yields the subsequent derivatives, we prefer to postpone their explicit derivation to Sec. 4, where we develop a more powerful approach.

In Sec. 3 (and Appendices C, D, and E) we sketch the method of operator traces, whose mathematical treatment is due to Dikiĭ.[10] This is basically the approach used by Percival.[5] We also report, in this section, the treatment of Buslaev and Faddeev,[4] although their method is in fact closer to that of Sec. 4. Although we are able to provide some slight improvements, in this section we give mainly results which have been obtained by others.

In Sec. 4 (and Appendices F, G, and H) we discuss a method, introduced by Roberts[6] (and previously, to a certain extent, also by Buslaev and Faddeev[4]), which is based on the asymptotic expansion of the scattering parameters (whose structure is conveniently studied using a technique due to Verde[7]) and on their analyticity properties. This elegant and powerful method is modified in such a manner that it becomes possible to derive explicit expressions for the values at the origin of the potential and of all its derivatives. As the order of differentiation increases, the formulas become more complicated; however, their general structure is given, and the numerical coefficients entering in these formulas may be evaluated by means of simple recursion relations, which are given explicitly and which may be fed into a computer. A table of these numerical coefficients, which includes all those relevant for the evaluation of the derivatives, of order ten or less, of the potential at the origin, is presented.

In Sec. 5 (and Appendix I) the treatment of Sec. 4 is utilized to obtain some results on the question of uniqueness in the inverse scattering problem.

Section 6 contains a summary of the results obtained and Sec. 7, a concluding discussion. It may be a good idea to glance through these sections first.

Throughout this paper we use units such that $\hbar = 2m = 1$, m being the mass of the scattered (or bound) particle.

2. FUNCTIONAL DERIVATIVE APPROACH

On the basis of the Gel'fand–Levitan solution of the inverse scattering problem, Newton[2] obtained the following expression for the functional derivative of the potential $V(r)$ with respect to the scattering phase shift $\eta(k)$:

$$\frac{\delta V(r)}{\delta\eta(k)} = -\frac{4}{\pi}k\frac{d}{dr}\tilde{G}(k^2, r), \quad (2.1)$$

where

$$\tilde{G}(k^2, r) = G(k^2, r) - \sum_n \frac{C_n}{k^2 - E_n}\varphi^2(E_n, r) \quad (2.2)$$

[4] V. S. Buslaev and L. D. Faddeev, Dokl. Akad. Nauk SSSR **132**, 13 (1960); [English transl.: Soviet Math.—Doklady **1**, 451 (1960)].

[5] I. C. Percival, Proc. Phys. Soc. (London) **80**, 1290 (1962).

[6] M. J. Roberts, Proc. Phys. Soc. (London) **84**, 825 (1964).

[7] M. Verde, Nuovo Cimento **2**, 1001 (1955).

[8] V. S. Buslaev, Dokl. Akad. Nauk SSSR **143**, 1067 (1962); [English transl.: Soviet Phys.—Doklady **7**, 295 (1962)]; I. C. Percival and M. J. Roberts, Proc. Phys. Soc. (London) **81**, 519 (1963); M. J. Roberts, *ibid.* **83**, 503 (1964), **86**, 683 (1965).

[9] General references on the inverse scattering problem: L. D. Faddeev, Usp. Mat. Nauk **14**, 57 (1959) [English transl.: J. Math. Phys. **4**, 72 (1963)]; Z. S. Agranovich and V. A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach Science Publishers, New York 1963); R. G. Newton, J. Math. Phys. **1**, 319 (1960); V. de Alfaro and T. Regge, *Potential Scattering* (North-Holland Publishing Company, Amsterdam, 1965); R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Company, Inc., New York, 1966).

[10] L. A. Dikiĭ, Usp. Mat. Nauk **13**, 111 (1958) [English transl.: in Trans. Am. Math. Soc. (2) **18**, 81 (1961)].

and

$$G(k^2, r) = \varphi(k^2, r)\psi(k^2, r) \tag{2.3a}$$

$$= -\frac{1}{2k}\,\mathrm{Im}\,[S(k)f^2(-k, r)], \quad (\mathrm{Im}\,k = 0). \tag{2.3b}$$

In these equations the function $f(-k, r)$ is the "Jost" solution of the radial Schrödinger equation

$$f''(-k, r) + [k^2 - V(r)]f(-k, r) = 0, \tag{2.4}$$

i.e., it is the solution of this equation characterized by the asymptotic boundary condition

$$\lim_{r\to\infty}[f(-k, r)e^{-ikr}] = 1. \tag{2.5}$$

In Eq. (2.4) and in what follows, the prime stands for (partial) differentiation with respect to r. The functions $\varphi(k^2, r)$ and $\psi(k^2, r)$ are also solutions of the radial Schrödinger equation and are identified through their connection to the Jost solutions,

$$\varphi(k^2, r) = (i/2k)[f(-k)f(k, r) - f(k)f(-k, r)], \tag{2.6}$$

$$\psi(k^2, r) = -[2f(k)f(-k)]^{-1}[f(-k)f(k, r) + f(k)f(-k, r)], \tag{2.7}$$

where we have introduced the two Jost functions

$$f(\pm k) \equiv f(\pm k, 0). \tag{2.8}$$

[We note that this definition implies that $\varphi(k^2, r)$ vanishes in the origin, so that it corresponds, apart from normalization, to the radial wavefunction.] The sum in Eq. (2.2) extends over all S-wave bound states, which are characterized by the (negative) binding energies E_n, by the corresponding radial wavefunctions $\varphi(E_n, r)$, whose normalization is specified through the condition

$$\lim_{r\to 0}[r^{-1}\varphi(E_n, r)] = 1, \tag{2.9}$$

and by the constants

$$C_n = \left[\int_0^\infty dr\,\varphi^2(E_n, r)\right]^{-1}. \tag{2.10}$$

We assume throughout that the number of bound states is finite.[11,12] Finally, the S-matrix function $S(k)$, appearing in Eq. (2.3b), is given by the usual formula:

$$S(k) = \exp[2i\eta(k)] = f(k)/f(-k). \tag{2.11}$$

We have, of course, for real k,

$$f(k) = |f(k)|\exp[i\eta(k)], \quad (\mathrm{Im}\,k = 0) \tag{2.12}$$

and

$$f(-k) = [f(k)]^*, \quad (\mathrm{Im}\,k = 0). \tag{2.13}$$

At this point we do not elaborate on the validity of Eq. (2.1), which is discussed below when the problem of its integration is tackled, except to note that throughout this paper, whenever we write a functional derivative, we assume the momentum variable k to extend over the range $-\infty$, $+\infty$; the values of the scattering phase shift for negative k are obtained from the reflection property

$$\eta(-k) = -\eta(k). \tag{2.14}$$

It is now easy to show, by straightforward differentiation of Eq. (2.2) and using the radial Schrödinger equation (2.4), that the function $\tilde{G}(k^2, r)$ satisfies the linear third-order differential equation

$$\tilde{G}'''(k^2, r) = 4[V(r) - k^2]\tilde{G}'(k^2, r) + 2V'(r)\tilde{G}(k^2, r). \tag{2.15}$$

The following boundary conditions are also easily obtained:

$$\tilde{G}(k^2, 0) = 0, \tag{2.16}$$

$$\tilde{G}'(k^2, 0) = -1, \tag{2.17}$$

$$\tilde{G}''(k^2, 0) = -2\,\mathrm{Re}\,\tilde{g}(k), \quad (\mathrm{Im}\,k = 0), \tag{2.18}$$

with

$$\tilde{g}(k) = g(k) + \sum_n \frac{C_n}{k^2 - E_n} \tag{2.19}$$

and

$$g(k) = ik + \frac{f'(k, 0)}{f(k, 0)} = ik + \frac{d}{dr}[\ln f(k, r)]\Big|_{r=0}. \tag{2.20}$$

The first condition, Eq. (2.16), follows immediately from Eqs. (2.3a) and (2.6); the second one, Eq. (2.17), from the Wronskian property

$$f'(-k, r)f(k, r) - f(-k, r)f'(k, r) = 2ik \tag{2.21}$$

and from Eqs. (2.3); the last one, Eqs. (2.18)–(2.20), from Eqs. (2.3b), (2.4), (2.11), (2.12), and (2.13).

The function $\tilde{g}(k)$ of Eq. (2.19) plays an important role in the following. It is easy to infer from the well-known analytic properties of $f(k, r)$ and the properties of $f(k)$ that it is a meromorphic function of k in the half-plane $\mathrm{Im}\,k \leq 0$. In fact, we prove in Appendix A that it is holomorphic in this half-plane (even if bound states are present), because the poles of $g(k)$ at $k = -i|E|^{\frac{1}{2}}$ are exactly canceled by the

[11] A sufficient condition for the finiteness of the number of bound states, first derived by Bargmann (see Ref. 12), is the existence of the integral $\int_0^\infty dr\,r\,V(r)$.

[12] V. Bargmann, Proc. Natl. Acad. Sci. U.S. **38**, 961 (1952).

sum in Eq. (2.19). Moreover, it satisfies, for real k, the equation

$$\operatorname{Im} \tilde{g}(k) = k[1 - |f(k)|^{-2}], \quad (\operatorname{Im} k = 0), \quad (2.22)$$

as implied by its definition, Eqs. (2.19) and (2.20), and by Eq. (2.21). Therefore its real part may be obtained, for real k, by the dispersion relation

$$\operatorname{Re} \tilde{g}(k) = \frac{1}{\pi} P\int_{-\infty}^{+\infty} dq \, \frac{q^2}{q^2 - k^2} [|f(q)|^{-2} - 1], \quad (\operatorname{Im} k = 0). \quad (2.23)$$

Incidentally, we note (also for future reference) that the function $|f(k)|$ is even:

$$|f(k)| = |f(-k)|, \quad (\operatorname{Im} k = 0), \quad (2.24)$$

while the reflection properties of $\tilde{g}(k)$ [or $g(k)$] are as follows:

$$\operatorname{Re} \tilde{g}(-k) = \operatorname{Re} \tilde{g}(k), \quad (\operatorname{Im} k = 0); \quad (2.25a)$$

$$\operatorname{Im} \tilde{g}(-k) = -\operatorname{Im} \tilde{g}(k), \quad (\operatorname{Im} k = 0); \quad (2.25b)$$

or, more generally,

$$\tilde{g}^*(-k^*) = \tilde{g}(k), \quad (\operatorname{Im} k \leq 0). \quad (2.25c)$$

The integral in Eq. (2.23) is convergent because, at large k,

$$|f(k)| = 1 + O(k^{-2}), \quad (2.26)$$

and for $\operatorname{Re} \tilde{g}(k)$ it implies the asymptotic behavior

$$\operatorname{Re} \tilde{g}(k) = O(k^{-2}), \quad (\operatorname{Im} k = 0), \quad (2.27)$$

while of course, from Eqs. (2.22) and (2.26), we get

$$\operatorname{Im} \tilde{g}(k) = O(k^{-1}), \quad (\operatorname{Im} k = 0). \quad (2.28)$$

A more detailed discussion of the asymptotic behavior of these quantities will be given below.

It should be emphasized that the function $|f(k)|$ is known once the phase shift $\eta(k)$ and the bound-state energies E_n are given; in fact it is expressed by the well-known formula[9]

$$|f(k)| = \left[\prod_n \left(1 - \frac{E_n}{k^2}\right)\right] \times \exp\left[-\frac{1}{\pi} P\int_{-\infty}^{+\infty} dq \, \frac{q}{q^2 - k^2} \eta(q)\right], \quad (\operatorname{Im} k = 0). \quad (2.29)$$

It follows, through Eqs. (2.22) and (2.23), that the function $\tilde{g}(k)$ may also be considered known once $\eta(k)$ and the E_n's are given. We record here, for future reference, the explicit expressions for $\operatorname{Re} \tilde{g}(k)$ and $\operatorname{Im} \tilde{g}(k)$:

$$\operatorname{Re} \tilde{g}(k) = \frac{1}{\pi} P\int_{-\infty}^{+\infty} dq \, \frac{q^2}{q^2 - k^2} \left\{\left[\prod_n \left(1 - \frac{E_n}{q^2}\right)\right]^{-2} \times \exp\left[\frac{2}{\pi} P\int_{-\infty}^{+\infty} dp \, \frac{p}{p^2 - q^2} \eta(p)\right] - 1\right\}, \quad (\operatorname{Im} k = 0), \quad (2.30a)$$

$$\operatorname{Im} \tilde{g}(k) = k\left\{1 - \left[\prod_n \left(1 - \frac{E_n}{k^2}\right)\right]^{-2} \times \exp\left[\frac{2}{\pi} P\int_{-\infty}^{+\infty} dq \, \frac{q}{q^2 - k^2} \eta(q)\right]\right\}, \quad (\operatorname{Im} k = 0). \quad (2.30b)$$

We now introduce the power expansion of the potential $V(r)$ and of the "Green's function" $\tilde{G}(k^2, r)$, setting

$$V(r) = \sum_{n=0}^{\infty} \frac{v_n}{n!} r^n, \quad (2.31)$$

$$\tilde{G}(k^2, r) = \sum_{n=1}^{\infty} \frac{\tilde{G}_n(k^2)}{n!} r^n. \quad (2.32)$$

The lower limit of the sum in the second equation is implied by Eq. (2.16). Moreover, the existence of these expansions and their convergence in the neighborhood of the origin is guaranteed by the assumed regularity of the potential function. Introducing these expansions in the differential equation (2.15), we secure the recursion relations

$$\tilde{G}_{n+3}(k^2) = -4k^2\tilde{G}_{n+1}(k^2) + 2\sum_{m=1}^{n+1} \frac{n + m + 1}{n + 1} \times \binom{n + 1}{m} \tilde{G}_m(k^2) v_{n+1-m}, \quad (2.33)$$

while the boundary conditions Eqs. (2.16)–(2.18) imply that

$$\tilde{G}_1(k^2) = -1, \quad (2.34)$$

$$\tilde{G}_2(k^2) = -2 \operatorname{Re} \tilde{g}(k), \quad (\operatorname{Im} k = 0). \quad (2.35)$$

But from Eq. (2.1) we infer that

$$\frac{\delta v_n}{\delta \eta(k)} = -\frac{4}{\pi} k\tilde{G}_{n+1}(k^2), \quad (2.36)$$

and, inserting this relation in Eq. (2.33), we finally obtain

$$\frac{\delta v_{n+2}}{\delta \eta(k)} = -4k^2 \frac{\delta v_n}{\delta \eta(k)} + 2\sum_{m=0}^{n} \frac{n + m + 2}{n + 1} \binom{n + 1}{m + 1} v_{n-m} \frac{\delta v_m}{\delta \eta(k)}. \quad (2.37)$$

In these equations the quantities v_n should be considered functionals of the phase shift $\eta(k)$. As is well known, however, the potential depends also on the parameters characterizing the bound states.[9] Therefore our hypothesis that the v_n's be considered functionals only of the phase shift restricts (for the time being) our consideration to families of potentials having the same number of bound states characterized by the same parameters.

This system of recursive equations together with the two starting conditions

$$\frac{\delta v_0}{\delta \eta(k)} = \frac{4}{\pi} k, \tag{2.38}$$

$$\frac{\delta v_1}{\delta \eta(k)} = \frac{8k}{\pi^2} P\int_{-\infty}^{+\infty} dq \, \frac{q^2}{q^2 - k^2} \left\{ \left[\prod_n \left(1 - \frac{E_n}{q^2}\right) \right]^{-2} \times \exp\left[\frac{2}{\pi} P\int_{-\infty}^{+\infty} dp \, \frac{p}{p^2 - q^2} \eta(p) \right] - 1 \right\} \tag{2.39}$$

allow, by sequential functional integration, evaluation of the quantities v_n, namely the values of the potential and its derivatives at the origin. There is, however, one difficulty which is immediately apparent and which is discussed below.

The functional integration is accomplished through the formula[13,14]

$$v_n = \bar{v}_n + \frac{1}{(s - \bar{s})} \int_{\bar{s}}^{s} dt \int_{-\infty}^{+\infty} dk [\eta(k) - \bar{\eta}(k)] \times \frac{\delta v_n}{\delta \eta'(k)} \Bigg|_{\eta'(k) = \frac{(t-\bar{s})\eta(k) - (t-s)\bar{\eta}(k)}{(s-\bar{s})}}, \tag{2.40}$$

where $\bar{v}_n$ is the nth derivative (evaluated at the origin) of the potential $\bar{V}(r)$ corresponding to the phase shift $\bar{\eta}(k)$, just as v_n is the nth derivative of the potential $V(r)$ corresponding to the phase shift $\eta(k)$. As implied by the remark after Eq. (2.37), the bound states of the two potentials $V(r)$ and $\bar{V}(r)$ must be characterized by the same parameters.

However, in order for Eq. (2.40) to hold, it is required that *the asymptotic vanishing at large k of the difference* $\eta(k) - \bar{\eta}(k)$ *be sufficiently rapid.* We give now a heuristic justification of this fact, and thereby we specify exactly how fast the difference $\eta(k) - \bar{\eta}(k)$ must vanish asymptotically in order that Eq. (2.40) holds. Because the final results derived in this section will be reobtained in Sec. 4, we content ourselves with this heuristic analysis.

First of all we introduce the asymptotic expansion of the phase shift $\eta(k)$:

$$\eta(k) = \sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}). \tag{2.41}$$

An analogous formula holds for $\bar{\eta}(k)$. The validity of these asymptotic expansions (and, in particular, the absence of logarithmic terms) is a consequence of the assumed holomorphy of the potential and of the restriction to S-wave scattering.[7]

We then consider Eq. (2.1) in the asymptotic (large k) region. From Eq. (2.3b) and the fact that

$$f(-k, r) = e^{ikr}[1 + O(k^{-1})], \tag{2.42}$$

we get the equation, valid at large k,

$$\frac{\delta V(r)}{\delta \eta(k)} = \frac{4}{\pi} k \cos 2kr. \tag{2.43}$$

Differentiating n times, we also obtain

$$\frac{\delta V^{(n)}(r)}{\delta \eta(k)} = \frac{4}{\pi} 2^n k^{n+1} (-)^{n/2} \cos 2kr, \quad \text{for even } n, \tag{2.44a}$$

$$= \frac{4}{\pi} 2^n k^{n+1} (-)^{n+1/2} \sin 2kr, \quad \text{for odd } n, \tag{2.44b}$$

where the notation $V^{(n)}(r)$ indicates the nth derivative of $V(r)$. We may now read from these equations the condition on the difference $\eta(k) - \bar{\eta}(k)$ sufficient to allow the functional integration because it corresponds simply to the requirement that the integral over k of the right-hand side of Eqs. (2.44) be absolutely convergent after multiplication by the difference $\eta(k) - \bar{\eta}(k)$. We therefore infer that a sufficient condition for the validity of Eq. (2.40) is that

$$\eta(k) - \bar{\eta}(k) = O(k^{-n-3}). \tag{2.45}$$

[One might think that our estimate could be made more stringent for odd n because the right-hand side of Eq. (2.44b) vanishes with r. In fact this is not the case because the order of the two limits $r \to 0$ and $k \to \infty$ is certainly relevant since the argument of the sine function is $2kr$.] On the other hand, using Eq. (2.41) in addition to Eq. (2.45), we conclude that a

---

[13] V. Volterra, *Theory of Functionals and Integro-differential Equations* (Dover Publications, Inc., New York, 1959), p. 159.

[14] This functional integration formula applies, provided that the functional derivative satisfies the integrability condition

$$\delta^2 v_n / \delta\eta(k)\delta\eta(k') = \delta^2 v_n / \delta\eta(k')\delta\eta(k).$$

While the validity of these conditions might be explicitly verified in each case, this is not necessary because the consistency of the procedure of derivation of the expression for the functional derivative from the Gel'fand–Levitan formalism automatically implies that the integration must be allowed and (therefore) that the integrability condition must be satisfied.

sufficient condition for the validity of Eq. (2.40) is the following:

$$a_j = \bar{a}_j, \quad j = 0, 1, \cdots, \tfrac{1}{2}n, \quad (n \text{ even}), \tag{2.46a}$$

$$a_j = \bar{a}_j, \quad j = 0, 1, \cdots, \tfrac{1}{2}(n+1), \quad (n \text{ odd}). \tag{2.46b}$$

[It should be emphasized that these conditions are certainly sufficient to guarantee the convergence of the integration on the right-hand side of Eq. (2.40); but, as the discussion which we have given above implies, for odd n they are *more stringent* than those which would be required merely to secure convergence.]

We are finally in a position to begin the process of sequential integration. Because the final results will be reobtained in Sec. 4, we concentrate here on explaining the approach rather than on the details of the derivation. From Eq. (2.38) and using Eq. (2.40), we obtain

$$v_0 = \bar{v}_0 + \frac{4}{\pi}\int_{-\infty}^{+\infty} dk\, k[\eta(k) - \bar{\eta}(k)], \quad (a_0 = \bar{a}_0), \tag{2.47a}$$

or, equivalently,

$$v_0 - \frac{4}{\pi}\int_{-\infty}^{+\infty} dk\, k\left[\eta(k) - \frac{a_0}{k}\right] = \bar{v}_0 - \frac{4}{\pi}\int_{-\infty}^{+\infty} dk k\left[\bar{\eta}(k) - \frac{\bar{a}_0}{k}\right], \quad (a_0 = \bar{a}_0). \tag{2.47b}$$

This equation may also be written in the form

$$v_0 - \frac{4}{\pi}\int_{-\infty}^{+\infty} dk\, k\left[\eta(k) - \frac{a_0}{k}\right] = A_0(a_0; E_n; C_n), \tag{2.47c}$$

where the function A_0 is a universal function of the arguments shown. This function may be evaluated inserting the values corresponding to a known potential in the left-hand side of Eq. (2.47c), having associated an arbitrary number of bound states with arbitrary parameters and an arbitrary value of a_0. Using this procedure, Newton found[2]

$$A_0(a_0; E_n; C_n) = -4\sum_n E_n. \tag{2.48}$$

This equation together with Eqs. (2.47c) and (D5) implies Eq. (1.1).

We then integrate Eq. (2.39) and, using Eq. (2.40) and some labor, we get

$$v_1 = \bar{v}_1 - \frac{4}{\pi}\int_{-\infty}^{+\infty} dk\, k^2 \times [|f(k)|^{-2} + 2\Delta(k) - |\bar{f}(k)|^{-2} - 2\bar{\Delta}(k)], \quad (a_0 = \bar{a}_0, \quad a_1 = \bar{a}_1), \tag{2.49}$$

where

$$\Delta(k) = -\frac{1}{\pi} P\int_{-\infty}^{+\infty} dq\, \frac{q}{q^2 - k^2}\eta(k), \tag{2.50}$$

so that [see Eq. (2.29)]

$$|f(k)| = \left[\prod_n\left(1 - \frac{E_n}{k^2}\right)\right]\exp[\Delta(k)]. \tag{2.29'}$$

From this equation we obtain

$$v_1 + \frac{4}{\pi}\int_{-\infty}^{+\infty} dk\, k^2[|f(k)|^{-2} + 2\Delta(k) - 1] = A_1(a_0, a_1; E_n; C_n). \tag{2.51}$$

The universal function A_1 might again be evaluated using a solvable potential. However, we defer to Sec. 4 the derivation of the expression for v_1 which finally obtains.

The process of sequential functional integration obviously can be continued indefinitely. For instance, from Eqs. (2.37), (2.38), and (1.1), one obtains

$$\frac{\delta v_2}{\delta\eta(k)} = -\frac{16}{\pi}k\left[k^2 + \frac{4}{\pi}\int_{-\infty}^{+\infty} dq\, q\frac{d}{dq}[q\eta(q)] + 4\sum_n E_n\right]. \tag{2.52}$$

Applying the functional integration formula Eq. (2.40) to this equation, one easily obtains an expression for v_2, except, of course, for the ambiguity embodied in the universal function $A_2(a_0, a_1; E_n; C_n)$, which should again be evaluated using a known solvable potential.

Clearly the process we have just described allows, in principle, the derivation of as many derivatives of the potential at the origin as one likes, except for the practical difficulty of performing the functional integration of ever more complicated expressions and of determining each universal function

$$A_{2m}(a_0, a_1, \cdots, a_m; E_n; C_n)$$

or

$$A_{2m+1}(a_0, a_1, \cdots, a_{m+1}; E_n; C_n)$$

using known solvable potentials which must be sufficiently general to contain, in addition to an arbitrary number of bound states characterized by arbitrary parameters, at least $m + 1$ (or $m + 2$) free parameters to correspond to the $m + 1$ (or $m + 2$) asymptotic constants a_i.

A simplification of the procedure to evaluate the universal functions A_p obtains if, in addition to the functional derivative of the potential with respect to the phase shift, Eq. (2.1), its (partial) derivatives

with respect to the bound state parameters C_n and E_n are also introduced.[2,15] By applying to these expressions an analogous treatment to that given above, it is then possible to obtain a recursive system involving, in addition to the (functional) derivatives of the quantities v_n with respect to the phase shift, the (partial) derivatives of these quantities with respect to the parameters E_n and C_n. Sequential integration over both the (functional) dependence upon $\eta(k)$ and the (ordinary) dependence upon the variables C_n and E_n can then be performed. The advantage of such a procedure is to reduce the ambiguity in the knowledge of the v_m's which obtains as an outcome of the integration process to a universal function $\tilde{A}_p$ which, although still depending on the $m + 1$ (if $p = 2m$) or $m + 2$ (if $p = 2m + 1$) variables a_i, is independent of the bound state parameters E_n and C_n. Thus the determination of these universal functions requires the introduction of a much simpler comparison potential, one which need not have any bound state. However, this advantage is compensated by the added complication of the sequential integration process. But since, as we have repeatedly emphasized, the method of this section, although sufficient in principle to determine all the v_n's, is not the most convenient route to reach that end, we choose to close the discussion at this point, relegating to Appendix B a sketch of the formulas relevant to the generalization we have just mentioned.

3. METHOD OF OPERATOR TRACES

In this section we discuss the method of operator traces. Its mathematical treatment is mainly due to Dikiĭ[10]; its application to scattering theory, to Buslaev and Faddeev[4] and, independently, to Percival.[5] Although our own contribution to this approach is minimal, we believe it is useful, for completeness, to illustrate also this method and to report the results it yields.

The basic idea of this approach is best introduced, following Percival,[5] by considering the Hamiltonian H of a quantum-mechanical system which admits only a finite number N of linearly independent states. Such a Hamiltonian may be represented by an $N \times N$ matrix, and the invariance of the trace of a matrix under unitary transformation implies the well-known formula

$$\text{Tr}\,[H^p] = \sum_{n=1}^{N} E_n^p, \tag{3.1}$$

where E_n are the eigenvalues of H.

[15] R. Jost and W. Kohn, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. 27, No. 9 (1953).

Let us now return to our problem, which is characterized by the Hamiltonian

$$H = T + V \tag{3.1a}$$

$$= -(d^2/dr^2) + V(r). \tag{3.1b}$$

The eigenvalues of this Hamiltonian are all the positive values $E = k^2$, and possibly, in addition, a finite number of negative values E_n corresponding to (S-wave) bound states.

Thus, in contrast to the previous case, we now have an infinite (nondenumerable) number of eigenstates. Thus the formula corresponding to Eq. (3.1) contains an integral in addition to a sum

$$\text{Tr}\,[H^p] = \sum_n E_n^p + \int_0^\infty dE\, E^p n(E), \tag{3.2}$$

where $n(E)$ is the density of states in the continuum. This equation, however, has only a symbolic meaning because the integral on the right-hand side does not converge (nor, of course, does the trace on the left-hand side). On the other hand, because the density of states $n(E)$ is connected with the scattering phase shift $\eta(k)$ (see below), if it were possible to assign a definite meaning to this formula, one would obtain relationships between the energy moments of the scattering phase shift on the right-hand side and the interaction which is contained in the Hamiltonian operator on the left-hand side. This motivates the attempt to extract definite relationships from the symbolic equation (3.2).

The first trick to use is to subtract the same expression for the case without interaction from Eq. (3.2) obtaining

$$\text{Tr}\,[H^p - T^p] = \sum_n E_n^p + \int_0^\infty dE\, n_I(E) E^p, \tag{3.3}$$

where $n_I(E)$ is now the increment of the density of states due to the interaction. It is well known (see Appendix C) that

$$n_I(E) = \frac{1}{2\pi k} \frac{d}{dk} \eta(k), \tag{3.4}$$

where $\eta(k)$ is the scattering phase shift. Thus, in place of Eq. (3.3), we may write

$$\text{Tr}\,[H^p - T^p] = \sum_n E_n^p + \frac{1}{\pi} \int_0^\infty dk\, k^{2p} \frac{d}{dk} \eta(k). \tag{3.5}$$

It should be emphasized that this equation still has only a symbolic meaning because, in general, the integral on the right-hand side does not converge.

For $p = 0$, however, the equation is directly meaningful, and it implies the result

$$\eta(0) - \eta(\infty) = N\pi, \tag{3.6}$$

where N is the number of bound states. This is just Levinson's theorem.[1]

The task of extracting definite results from the symbolic equation (3.5) is closely connected to that of the regularization of divergent sums and integrals, a mathematical problem which recently has received much attention.[16] We sketch below the two different approaches used by Percival[5] and by Buslaev and Faddeev,[4] and we report their results.

We begin with the work of Percival,[5] who uses a more physical approach, borrowing certain techniques from the theory of statistical mechanics. In place of Eq. (3.3) he substitutes the equation

$$\mathrm{Tr}\,[e^{-\beta H} - e^{-\beta T}] = \sum_n e^{-\beta E_n} + \frac{1}{\pi}\int_0^\infty dk e^{-\beta k^2}\frac{d}{dk}\eta(k), \tag{3.7}$$

which is now certainly meaningful for positive β. He then notes that, while an expansion in powers of β of both sides of this equation is not allowed [as it would merely reproduce the symbolic equation (3.5)], one can expand in powers of $\beta^{\frac{1}{2}}$ and equate the coefficients of equal powers of $\beta^{\frac{1}{2}}$. In this manner he is able to derive the following moment relationships:

$$-\frac{1}{4\pi}\int_{-\infty}^{+\infty}dr\int_{-\infty}^{+\infty}dk e^{ikr}[(T+V)^p - T^p]e^{ikr} = \sum_n E_n^p + \frac{1}{\pi}\int_0^\infty dk\, k^{2p} \times\left[\frac{d\eta(k)}{dk} + \sum_{n=0}^{p-1}(2n+1)a_n k^{-2n-2}\right], \quad p = 1, 2, 3, \cdots \tag{3.8}$$

The coefficients a_n are those characterizing the asymptotic behavior of $\eta(k)$; i.e., they are the same coefficients which were previously introduced through Eq. (2.41). Percival shows that they could be obtained equating the coefficients of β^p in

$$\frac{\sqrt{\pi}}{(2\pi)^2}\sum_{s=1}^{\infty}\frac{(-)^s\beta^{s+\frac{1}{2}}}{s!}\int_{-\infty}^{+\infty}dr \times\int_{-\infty}^{+\infty}dk\left[-\frac{d^2}{dr^2} - 2ik\frac{d}{dr} + V(r)\right]^s e^{-\beta k^2} = \sum_{p=0}^{\infty}\frac{(-)^p 2^{2p}}{(2p)!}p!\,a_p\beta^p. \tag{3.9}$$

[16] See, for instance, I. M. Gel'fand and G. S. Šilov, *Generalized functions. Vol. I. Properties and operations* (Academic Press Inc., New York, 1964).

Thereby one would obtain explicit expressions of the coefficients a_p as an integral over powers of the potential and its derivatives. We are more interested in Eq. (3.8), which we rewrite, using the results of Appendices D and E, in the more compact form

$$-\tfrac{1}{4}\{1\cdot[V(r) - \tfrac{1}{4}(\vec{d} - \overleftarrow{d})^2]^p\cdot 1\}\big|_{r=0} = \sum_n E_n^p + \frac{p}{\pi(2p-1)!}\int_{-\infty}^{+\infty}dk D_k^{2p-1}\eta(k). \tag{3.10}$$

(The differential operators $\vec{d}$, $\overleftarrow{d}$, and D_k are defined in Appendices D and E.)

It should be emphasized at this point that the treatment of Percival applies only to even potentials, i.e., potentials such that

$$v_{2n+1} \equiv \frac{d^{2n+1}V(r)}{dr^{2n+1}}\bigg|_{r=0} = 0. \tag{3.11}$$

It is this condition (together with the assumed asymptotic vanishing of the potential) that guarantees the convergence of the r integration in Eqs. (3.8) and (3.9) at both ends. The validity of the results we have just described requires, in addition, just as in the previous section, that the potential be a holomorphic function of r.

It also follows from the discussion of Appendix E that Eq. (3.10) may be written in the form

$$v_{2p} = (-)^{p+1}\frac{2^{2p+2}}{p+1}\sum_n E_n^{p+1} + \frac{(-)^{p+1}2^{2p+2}}{\pi(2p+1)!} \times\int_{-\infty}^{+\infty}dk\, D_k^{2p+1}\eta(k) + F_p(v_0, v_2, \cdots, v_{2p-2}), \tag{3.12}$$

where F_p is defined by the equation

$$F_p(v_0, v_2, \cdots, v_{2p-2}) = \frac{(-)^{p+1}2^{2p}}{p+1}\left\{1\cdot[V(r) - \tfrac{1}{4}(\vec{d} - \overleftarrow{d})^2]^{p+1}\cdot 1\right\}\bigg|_{r=0} + v_{2p}, \tag{3.13}$$

and depends only on the v_n's whose order n does not exceed $2p - 2$. It is therefore seen from Eq. (3.12) that in principle it is possible to obtain sequentially all the even derivatives of the potential at the origin in terms of the generalized moments

$$L_p = \frac{2}{p}\sum_n E_n^p + \frac{2}{\pi(2p-1)!}\int_{-\infty}^{+\infty}dk D_k^{2p-1}\eta(k), \quad p = 1, 2, \cdots. \tag{3.14}$$

The derivative of order $2p$ depends only on the moments L_n whose order n does not exceed $p + 1$. The knowledge of all the even derivatives is, of course,

sufficient to reconstruct the whole potential, in view of its assumed property of being even. On the other hand, the conditions that the phase shift and the bound state parameters must satisfy in order that a corresponding even potential exist are far from trivial; moreover, their investigation is completely beyond the power of the approach we have just described. These conditions are instead easily obtained from the results of the following section; they are given explicitly in Sec. 5.

We turn now to a brief description of the method employed by Buslaev and Faddeev[4] and to a summary of their results. As will be seen, their approach, although motivated by the approach of operator traces discussed above, is, in fact, closer to the method based on the asymptotic expansion of the scattering parameters and on their analyticity properties (which will be discussed in the following section and in Appendices F, G, and H).

The basic equation of their method is written in our notation as follows:

$$\sin \pi z \int_0^\infty dk\, k^{2z-1}\eta(k) + \cos \pi z \int_0^\infty dk\, k^{2z-1} \ln |f(k)| + \frac{\pi}{2z}\sum_n (-E_n)^z = 0, \quad (0 < \operatorname{Re} z < \tfrac{1}{2}). \tag{3.15}$$

They establish this equation by contour integration of the function $k^{2z}(d/dk) \ln f(k)$, exploiting the holomorphy of $f(k)$ in the lower half-plane $\operatorname{Im} k < 0$ and the fact that its zeros there correspond to the bound states. The equation might be considered a realization of the symbolic equation (3.5); however, there is no explicit dependence on the interaction (see below). Of course, the restriction to the strip $0 < \operatorname{Re} z < \frac{1}{2}$ is essential. The moment relationships are then obtained by analytic continuation in z outside this strip. The continuation leads to a formula which, using the results of Appendix D, may be cast into the form:

$$a_p = \frac{2(-)^p}{2p+1}\sum_n (-E_n)^{p+\frac{1}{2}} - \frac{2}{\pi(2p)!}\int_0^\infty dk D_k^{2p} \ln |f(k)|, \tag{3.16}$$

$$d_p = \frac{2}{\pi(2p-1)!}\int_0^\infty dk D_k^{2p-1}\eta(k), \tag{3.17}$$

where a_p and d_p are the asymptotic coefficients of $\eta(k)$ [see Eq. (2.41)] and $\Delta(k)$:

$$\Delta(k) = \sum_{n=1}^{N} d_n k^{-2n} + O(k^{-2N-2}), \tag{3.18}$$

and the function $\Delta(k)$ is defined by Eq. (2.50).

Using Eq. (2.29′) and performing (with standard techniques) the integral over the bound state contributions, we may recast Eq. (3.16) in the form

$$a_p = -\frac{2}{\pi(2p)!}\int_0^\infty dk D_k^{2p}\Delta(k). \tag{3.19}$$

Let us point out (along with Roberts[6]) that the machinery employed by Buslaev and Faddeev to derive these equations is somewhat more complicated than is really necessary. For instance, it is easier (see Appendix G) to derive Eq. (3.17) performing an asymptotic expansion directly on Eq. (2.50). This remark is the starting point for the discussion of the next section.

The results obtained thus far do not provide an explicit connection between energy moments of the phase shift (or of the modulus of the Jost function), bound state energies, and interaction. Such a connection is introduced using the explicit asymptotic expressions of the phase shift and of the modulus of the Jost function in terms of the potential. Such expressions are derived and discussed in the following section and in Appendix F, using a technique due to Verde.[7] Similar results have been obtained with a different technique by Buslaev and Faddeev. Their results imply that

$$a_p = (-)^p 2^{-2p-1} Q_{2p+1}, \tag{3.20}$$

$$d_p = p^{-1}\sum_n E_n^p + (-)^{p+1}2^{-2p}Q_{2p}, \tag{3.21}$$

where the coefficients Q_n are obtained from the recursion relations[17]

$$Q_p = V_{p-1} + \sum_{j=1}^{p-1} \frac{j}{p} V_{p-j-1}Q_j, \tag{3.22}$$

[17] It is easily seen that these recursion relations are solved by the explicit formulas

$$Q_p = \sum (m-1)! \prod_{j=1}^{l} [V_{j-1}^{s_j}/s_j!],$$

$$V_{p-1} = -\sum (-)^m \prod_{j=1}^{l} [Q_j^{s_j}/s_j!],$$

where, in both sums, we have

$$m = \sum_{j=1}^{l} s_j,$$

and both sums extend over all positive integral values of l and s_l and nonnegative integral values of s_j, $(j = 1, 2, \cdots, l-1,)$ such that

$$\sum_{j=1}^{l} s_j = p.$$

These equations are obtained using the technique of Appendix G, Eqs. (G18) ff. This is applicable because the Q_j's are the coefficients of the asymptotic expansion of the logarithm of a function whose asymptotic expansion is characterized by the coefficients V_j.

where, by definition,

$$V_l = \lim_{r\to\infty} V_l(r), \tag{3.23}$$

and the functions $V_l(r)$ are obtained by recursion from

$$V_0(r) = -\int_0^r ds V(s), \tag{3.24}$$

and

$$V_l(r) = V^{(l-1)}(0) + \sum_{m=0}^{l-1} \binom{l-1}{m} \int_0^r ds V_m(s) V^{(l-1-m)}(s). \tag{3.25}$$

As usual, by $V^{(n)}(r)$ we indicate the nth derivative of the potential. From these relationships Buslaev and Faddeev find

$$Q_1 = -\int_0^\infty dr V(r), \tag{3.26}$$

$$Q_2 = V(0), \tag{3.27}$$

$$Q_3 = V'(0) + \int_0^\infty dr V^2(r), \tag{3.28}$$

$$Q_4 = V''(0) - 2V^2(0). \tag{3.29}$$

Inserting the expressions for Q_2 and Q_4 in Eqs. (3.21) and (3.17), they obtain explicit expressions for $V(0)$ and $V''(0)$. However, it is clear that this approach cannot produce an expression for the first derivative $V'(0)$ nor for the higher derivatives, except in the case of an even potential. (In fact, Q_6 is a linear combination of $V^{IV}(0)$, $V''(0)V(0)$, $[V'(0)]^2$ and $[V(0)]^3$; see Appendices F, G, and H.)

4. METHOD OF ASYMPTOTIC EXPANSION

Because the method which we treat in this section is mainly based on the relationship between the coefficients of the asymptotic expansion of the scattering parameters and the potential, we begin our discussion with an analysis of this problem. We follow the approach of Verde.[7]

Let us introduce the function

$$g(k, r) = ik + f'(k, r)/f(k, r). \tag{4.1}$$

The radial Schrödinger equation (2.4) satisfied by $f(k, r)$ implies for $g(k, r)$ the Riccati equation

$$g'(k, r) = V(r) + 2ikg(k, r) - g^2(k, r), \tag{4.2}$$

while the condition Eq. (2.5) implies that $g(k, r)$ vanishes asymptotically:

$$g(k, \infty) = 0. \tag{4.3}$$

On the other hand, at the origin

$$g(k, 0) \equiv g(k), \tag{4.4}$$

where the function $g(k)$ is the same as the one introduced in Sec. 2, Eq. (2.20).

We now introduce the asymptotic expansion in k:

$$g(k, r) = \sum_{n=0}^{N} g_n(r)(-2ik)^{-n-1} + O(k^{-N-2}). \tag{4.5}$$

Inserting this expansion in Eq. (4.2), we obtain the recursion relations:

$$g_n'(r) + g_{n+1}(r) + \sum_{m=0}^{n-1} g_m(r) g_{n-1-m}(r) = 0, \quad n \geq 1. \tag{4.6}$$

These equations, together with the starting conditions

$$g_0(r) = V(r), \quad g_1(r) = -V'(r) \tag{4.7}$$

and the asymptotic property

$$g_n(\infty) = 0, \tag{4.8}$$

are sufficient to determine by recursion all the functions $g_n(r)$. We also introduce the notation

$$g_n \equiv g_n(0), \tag{4.9}$$

which implies the asymptotic expansion

$$g(k) = \sum_{n=0}^{N} g_n(-2ik)^{-n-1} + O(k^{-N-2}). \tag{4.10}$$

It is important to notice that knowledge of the quantities g_m for all m up to n implies knowledge of the quantities

$$g_0^{(m)} \equiv \frac{d^m}{dr^m} g_0(r)\bigg|_{r=0} = v_m \equiv \frac{d^m}{dr^m} V(r)\bigg|_{r=0} \tag{4.11}$$

for all m up to n. This is a consequence of the recursion relations

$$\bar{g}_n^{(p)} = (-)^p g_n + (-)^p \sum_{m=0}^{p-1} (-)^m \sum_{t=0}^{m} \binom{m}{t} \times \sum_{u=t}^{n-m+t-2} \bar{g}_u^{(t)} \bar{g}_{n-2-u}^{(m-t)}, \quad (n \geq p), \tag{4.12}$$

where

$$\bar{g}_n^{(p)} = \frac{d^p}{dr^p} g_{n-p}(r)\bigg|_{r=0}, \quad (n \geq p). \tag{4.13}$$

These relations are derived from Eq. (4.6) in Appendix F. They explicitly imply that $\bar{g}_n^{(p)}$ is determined once the quantities $\bar{g}_l^{(m)}$ are given for all m up to $p-1$ and for all l up to n; therefore, by recursion, $\bar{g}_n^{(p)}$ is determined once the quantities $\bar{g}_l^{(0)}$ are given for all l up to n. But by definition,

$$\bar{g}_l^{(0)} = g_l. \tag{4.14}$$

Note, that in addition to having shown that the quantity $g_0^{(n)} = v_n$ is determined by the quantities g_l

for all l up to n, we may assert that the structure of the relationship between these quantities must be the following:

$$v_p = \sum (p, l \mid s_1, s_2, \cdots, s_l) \prod_{j=1}^{l} g_{s_j}, \quad (4.15a)$$

where the sum extends over all positive integers l and nonnegative integers s_j. The coefficients $(p, l \mid s_1, \cdots, s_l)$ depend only on the arguments shown, and vanish unless

$$\sum_{j=1}^{l} s_j + 2l = p + 2. \quad (4.15b)$$

This dimensional condition severely restricts the number of terms in the sum. In particular, the maximum value of l allowed is

$$l_{\mathrm{MAX}} = 1 + p/2, \quad \text{for even } p, \quad (4.15c)$$

$$= 1 + (p-1)/2, \quad \text{for odd } p. \quad (4.15d)$$

A detailed discussion of the sum of Eq. (4.15a) and of the coefficients entering into it is given in Appendix F.

It should be emphasized that what we have now shown implies that knowledge of the quantities g_m for all m is sufficient, in principle, for the reconstruction of the whole potential (which is, by assumption, a holomorphic function of r, and can therefore be obtained from all its derivatives at the origin).

But the quantities g_m are known once the phase shift and the bound state parameters are known. In fact, from Eq. (2.19) we infer that

$$g_{2p+1} = \tilde{g}_{2p+1} + 2^{2p+2} \sum_n C_n (-E_n)^p, \quad (4.16)$$

$$g_{2p} = \tilde{g}_{2p}, \quad (4.17)$$

where the quantities $\tilde{g}_n$ are the asymptotic coefficients of the expansion in inverse power of $(-2ik)$ of the function $\tilde{g}(k)$:

$$\tilde{g}(k) = \sum_{n=0}^{N} \tilde{g}_n (-2ik)^{-n-1} + O(k^{-N-2}). \quad (4.18)$$

The explicit expression of $\tilde{g}(k)$ in terms of the phase shift and the bound state energies is given by Eqs. (2.30). From these expressions it is possible to derive explicit formulas for the coefficients $\tilde{g}_n$. The details of this derivation, which is based on a method due to Roberts,[6] are given in Appendix G. The final formulas are Eq. (G40) for the odd coefficients $\tilde{g}_{2p+1}$ and Eqs. (G34), (G35), and (G17) for the even coefficients $\tilde{g}_{2p}$. [Explicit expressions for the even coefficients $\tilde{g}_{2p}$ up to $2p = 10$ are displayed in Eqs. (G37).] We remark that the even coefficients $\tilde{g}_{2p}$ are given by combinations of the generalized moments (of odd order) of the phase shift, while the odd coefficients $\tilde{g}_{2p+1}$ are directly proportional to the generalized moments (of even order) of the "spectral function" $|f(k)|^{-2}$. As we already noted, these coefficients $\tilde{g}_n$ depend on the scattering phase shift and on the energies of the bound states, but are independent of the bound state normalization coefficients C_n. It is only the odd coefficients g_{2p+1} which depend on these parameters, and in the simple manner displayed in Eq. (4.16).

We are finally in the position to write down explicit expressions for the values of the potential and its derivatives at the origin in terms of the phase shift and the bound-state parameters. The procedure is to express, by means of Verde's recursion relations, the quantities v_n in terms of the coefficients g_s, through Eq. (4.15) [or, more conveniently, through Eqs. (F8) and (4.9)] and then to express the coefficients g_s in terms of the phase shift and the bound-state parameters through Eqs. (4.16), (4.17), (G40), (2.29), (G34), (G35), and (G17). This procedure is in principle the same one employed by Roberts,[6] except for the fact that he concentrated on the asymptotic expansion of the function $h(k)$ of Eqs. (G8) rather than on $g(k)$ or $\tilde{g}(k)$. He therefore, obtained the results corresponding to the even coefficients g_{2p}, but not those corresponding to the odd coefficients g_{2p+1}. Thus his results were not sufficient to derive an expression for the first derivative of the potential in the origin, nor for the derivatives higher than the second, except for even potentials. (A more detailed comparison of Roberts' results with ours is presented in Appendix H.)

We end this section by writing the explicit formulas for the derivatives of the potential at the origin which obtain with the technique we have just illustrated. We condense all these equations in the formula

$$v_p = \sum [p, l \mid s_2, s_3, \cdots, s_{l+1}] \prod_{j=2}^{l+1} J_j^{s_j}, \quad (4.19)$$

where

$$J_{2j} = (-)^{j+1} 2^{2j+2} \left[\frac{1}{j} \sum_n E_n^j + \frac{2}{\pi(2j-1)!} \times \int_0^\infty dk\, D_k^{2j-1} \eta(k) \right] = (-)^{j+1} 2^{2j+1} L_j \quad (4.20)$$

and

$$J_{2j+1} = (-)^{j+1} 2^{2j} \left[\sum_n C_n E_n^{j-1} - \frac{2}{\pi(2j+1)!} \times \int_0^\infty dk\, D_k^{2j+1} (k^{-1} |f(k)|^{-2}) \right] = g_{2j-1}. \quad (4.21)$$

(Note that these integrals extend only over positive values of k.) The sum in Eq. (4.19) extends over all

TABLE I. The coefficients $[p, l \mid s_2, \cdots, s_{l+1}]$ of Eq. (4.19) for p up to 10. These coefficients have been evaluated by computer from Eqs. (F8), (4.9), (4.16), (4.17), (G34), (4.20), (4.21), and from the data of Table II. The coefficients up to $p = 6$ have been checked by hand. All the nonvanishing coefficients are printed.

$[1, 2 \mid 0, 1] = -1.$
$[2, 1 \mid 2] = \frac{1}{8}$; $[2, 3 \mid 0, 0, 1] = -\frac{1}{4}.$
$[3, 2 \mid 1, 1] = 1$; $[3, 4 \mid 0, 0, 0, 1] = -1.$
$[4, 1 \mid 3] = -\frac{1}{6}$; $[4, 2 \mid 0, 2] = 5$; $[4, 3 \mid 1, 0, 1] = \frac{1}{2}$; $[4, 5 \mid 0, 0, 0, 0, 1] = -\frac{1}{4}.$
$[5, 2 \mid 2, 1] = -\frac{13}{4}$; $[5, 3 \mid 0, 1, 1] = \frac{9}{2}$; $[5, 4 \mid 1, 0, 0, 1] = 2$; $[5, 6 \mid 0, 0, 0, 0, 0, 1] = -1.$
$[6, 1 \mid 4] = \frac{1}{8}$; $[6, 2 \mid 1, 2] = -28$; $[6, 3 \mid 0, 0, 2] = \frac{5}{4}$; $[6, 3 \mid 2, 0, 1] = -2$; $[6, 4 \mid 0, 1, 0, 1] = 28$;
$[6, 5 \mid 1, 0, 0, 0, 1] = \frac{3}{4}$; $[6, 7 \mid 0, 0, 0, 0, 0, 0, 1] = -\frac{1}{4}.$
$[7, 2 \mid 0, 3] = -140$; $[7, 2 \mid 3, 1] = \frac{97}{6}$; $[7, 3 \mid 1, 1, 1] = -37$; $[7, 4 \mid 0, 0, 1, 1] = 17$;
$[7, 4 \mid 2, 0, 0, 1] = -\frac{23}{2}$; $[7, 5 \mid 0, 1, 0, 0, 1] = 10$; $[7, 6 \mid 1, 0, 0, 0, 0, 1] = 3$;
$[7, 8 \mid 0, 0, 0, 0, 0, 0, 0, 1] = -1.$
$[8, 1 \mid 5] = -\frac{38}{15}$; $[8, 2 \mid 2, 2] = 258$; $[8, 3 \mid 1, 0, 2] = -14$; $[8, 3 \mid 3, 0, 1] = \frac{38}{3}$;
$[8, 4 \mid 0, 0, 0, 2] = 69$; $[8, 4 \mid 1, 1, 0, 1] = -246$; $[8, 5 \mid 0, 0, 1, 0, 1] = 7$; $[8, 5 \mid 2, 0, 0, 0, 1] = -5$;
$[8, 6 \mid 0, 1, 0, 0, 0, 1] = 54$; $[8, 7 \mid 1, 0, 0, 0, 0, 0, 1] = 1$; $[8, 9 \mid 0, 0, 0, 0, 0, 0, 0, 0, 1] = -\frac{1}{4}.$
$[9, 2 \mid 1, 3] = 2310$; $[9, 2 \mid 4, 1] = -\frac{6047}{48}$; $[9, 3 \mid 0, 1, 2] = -\frac{791}{4}$; $[9, 3 \mid 2, 1, 1] = \frac{1667}{4}$;
$[9, 4 \mid 1, 0, 1, 1] = -208$; $[9, 4 \mid 0, 2, 0, 1] = -2310$; $[9, 4 \mid 3, 0, 0, 1] = \frac{523}{6}$;
$[9, 5 \mid 0, 0, 0, 1, 1] = \frac{1265}{2}$; $[9, 5 \mid 1, 1, 0, 0, 1] = -115$; $[9, 6 \mid 0, 0, 1, 0, 0, 1] = \frac{83}{8}$;
$[9, 6 \mid 2, 0, 0, 0, 0, 1] = -\frac{1077}{4}$; $[9, 7 \mid 0, 1, 0, 0, 0, 0, 1] = \frac{365}{2}$; $[9, 8 \mid 1, 0, 0, 0, 0, 0, 0, 1] = 4$;
$[9, 10 \mid 0, 0, 0, 0, 0, 0, 0, 0, 0, 1] = -1.$
$[10, 1 \mid 6] = \frac{59}{3}$; $[10, 2 \mid 0, 4] = 11550$; $[10, 2 \mid 3, 2] = -\frac{91956}{3}$; $[10, 3 \mid 0, 0, 3] = -\frac{1654}{3}$;
$[10, 3 \mid 2, 0, 2] = 190$; $[10, 3 \mid 1, 2, 1] = 5852$; $[10, 4 \mid 1, 0, 0, 2] = -836$;
$[10, 4 \mid 0, 1, 1, 1] = -3652$; $[10, 4 \mid 2, 1, 0, 1] = 3344$; $[10, 5 \mid 0, 0, 0, 0, 2] = \frac{63}{4}$;
$[10, 5 \mid 1, 0, 1, 0, 1] = -108$; $[10, 5 \mid 0, 2, 0, 0, 1] = -1100$; $[10, 5 \mid 3, 0, 0, 0, 1] = 46$;
$[10, 6 \mid 0, 0, 0, 1, 0, 1] = 418$; $[10, 6 \mid 1, 1, 0, 0, 0, 1] = -682$; $[10, 7 \mid 0, 0, 1, 0, 0, 0, 1] = 15$,
$[10, 7 \mid 2, 0, 0, 0, 0, 0, 1] = -10$; $[10, 8 \mid 0, 1, 0, 0, 0, 0, 0, 1] = 88$;
$[10, 9 \mid 1, 0, 0, 0, 0, 0, 0, 0, 0, 1] = \frac{5}{4}$; $[10, 11 \mid 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1] = -\frac{1}{4}.$

positive integral values of l and s_{l+1} and all nonnegative integral values of s_j, $j = 2, 3, \cdots, l$, such that

$$\sum_{j=2}^{l+1} j s_j = p + 2. \tag{4.22}$$

(A simple way to check this condition is a dimensional count.) Thus the maximum value of l appearing in Eq. (4.19) is

$$l_{\max} = p + 1. \tag{4.23}$$

All the coefficients $[p, l \mid s_2, \cdots, s_{l+1}]$ for p up to 10 are given in Table I.

5. REMARKS ON THE INVERSE PROBLEM

As it is well known, given the scattering phase shift $\eta(k)$, the (negative) bound-state energies E_n and the (positive) bound-state normalization coefficients C_n, there exists one and only one corresponding potential $V(r)$.[18] On the other hand, there exists an N parameter family of "(S-wave) physically equivalent" potentials, which have associated with them the same (S-wave) scattering phase shift $\eta(k)$ and the same bound-state energies E_n without any restriction on the coefficients C_n; N is the number of (S-wave) bound states.[19] The relation between these physically equivalent[20] potentials has received much attention, especially from the point of view of their asymptotic behavior.[9,15,21] Here we give some results which instead concern their behavior at the origin. We then discuss some properties of the phase shift and bound-state parameters associated with even potentials.

Theorem 1: The value at the origin of all physically equivalent potentials is the same. This also applies to the value at the origin of their second derivatives.

This well-known[9,15] result follows trivially from the explicit expressions of $V(0)$ and $V''(0)$, Eqs. (1.1) and (6.6). It should be noted that these formulas were already derived by Buslaev and Faddeev,[4] Percival,[5] and Roberts.[6]

Theorem 2: Among a class of physically equivalent potentials there may exist only one which is even.

This theorem follows already, for instance, from the work of Percival,[5] reported in Sec. 3, because that work implies that an even potential depends only on

[18] This well-known result (see Ref. 9) is directly implied by the discussion of the preceding section.

[19] Throughout we exclude from consideration the marginal case of a bound state (or rather, resonance) occurring at zero energy.

[20] They are also called "Jost equivalent" because they have the same Jost function $f(k)$. Recall that the Jost function depends only on the phase shift and the bound-state energies; see Eqs (2.12) and (2.29).

[21] V. Bargmann, Phys. Rev. **75**, 301 (1949).

the phase shift and the bound state energies, and not on the C_n's. A more direct method to prove this result is to note that, from Eqs. (B2) and (B3), it follows that $\partial v_{2n}/\partial C_l$ vanishes for all (integral) values of n and l if all v_{2n+1} vanish.

Theorem 3: The necessary and sufficient conditions that the S-wave phase shift and bound state parameters must satisfy in order for the corresponding potential to be even are

$$\sum_n C_n E_n^{j-1} - \frac{2}{\pi(2j+1)!} \times \int_0^\infty dk\, k^{2j+1} \frac{d^{2j+1}}{dk^{2j+1}} [k^{2j} |f(k)|^{-2}] = 0, \quad j = 1, 2, \cdots. \tag{5.1}$$

An equivalent condition is

$$\lim_{k\to\infty} [k^m \operatorname{Re} g(k)] = 0, \quad (\operatorname{Im} k = 0), \tag{5.2}$$

where m is arbitrary

The functions $|f(k)|$ and $\operatorname{Re} g(k)$ in these formulas are those defined in Eqs. (2.29), (2.19), and (2.23).

The proof of this theorem is immediate. Note that Eq. (5.1) corresponds simply, through Eq. (4.21), to the condition that all J_j's with odd j vanish. This implies, through Eqs. (4.19) and (4.22), that all v_p with odd p vanish. [In fact, if all the j's in Eq. (4.22) are even, p is also necessarily even.] On the other hand, it is easily seen that Eqs. (4.19) and (4.22) imply

$$v_p = [p, p+1 \,|\, 0, 0, \cdots, 1]J_p + G_p[J_2, J_3, \cdots, J_{p-2}], \tag{5.3}$$

where G_p vanishes if p is odd and if all J_j vanish for odd $j < p$ [see the explicit Eqs. (6.5)–(6.8)]. Thus, if all v_p's with odd p vanish, also all J_p's with odd p must vanish. Finally, the equivalence of Eq. (5.2) to Eq. (5.1) follows from the fact that the J_j's (or, equivalently, the g_j's) with odd j are the coefficients of the asymptotic expansion of $\operatorname{Re} g(k)$, for real k, as implied by Eqs. (2.19), (2.25a), and (4.10).

Note that this theorem implies that, for an even potential, the energies of the bound states and the bound-state normalization constants may be expressed in terms of the phase shift (which, in addition, must satisfy an infinite number of additional equations). For instance, for an even potential which possesses only one (S-wave) bound state, we find, from Eqs. (5.1) and (2.29), the implicit relations

$$E_1 = \frac{1}{20} \frac{\int_0^\infty dk\, D_k^5[k^3(k^2 - E_1)^{-2} e^{-2\Delta(k)}]}{\int_0^\infty dk\, D_k^3[k^3(k^2 - E_1)^{-2} e^{-2\Delta(k)}]}, \tag{5.4}$$

$$C_1 = \frac{1}{3\pi} \int_0^\infty dk\, D_k^3[k^3(k^2 - E_1)^{-2} e^{-2\Delta(k)}], \tag{5.5}$$

the operator D_k being that introduced in Appendix D and $\Delta(k)$ being explicitly expressed in terms of the phase shift in Eq. (2.50). It should be emphasized that these results hold for a potential which is represented by a holomorphic function of r (and which vanishes sufficiently fast at infinity), and that by the term "even" we mean a potential such that all its derivatives of odd order vanish at the origin.

Some consequences of Theorem 3 which concerns the very special case of even Bargmann's potentials[9,21] are discussed in Appendix I.

Theorem 4: Two (S-wave) physically equivalent potentials are identical if the values at the origin of all their derivatives up to the order $2N - 1$ coincide, N being the number of (S-wave) bound states.

Proof: It follows from Verde's recursion relations, Eqs. (4.6) and (4.7), and the definition equation (4.9), that all quantities g_n for n up to p are determined if the values of all derivatives of the potential up to the order p are assigned. On the other hand, once the quantities g_n for n up to $2N - 1$ are assigned, the N equations (4.16) with $p = 0, 1, \cdots, N - 1$ are sufficient to determine the N coefficients C_n—provided, of course, that the quantities g_n and E_n are known. These equations in fact constitute a linear system of N equations for the N unknown coefficients C_n; and it is easily seen that the determinant of its coefficients cannot vanish, since all the bound-state energies must be different from one another.[22] Thus, if in addition to the phase shift and the bound-state energies, which determine all the g_n's, the values at the origin of the derivatives of the potential are assigned up to the order $2N - 1$, the bound-state normalization coefficients C_n are determined, and therefore, the corresponding potential is uniquely characterized. Q.E.D.

6. SUMMARY OF RESULTS

For the convenience of the reader who is more interested in the results than in their derivation we collect here the relevant final formulas.

[22] It is, in fact, a so-called Van der Monde determinant. See, for instance, M. Picone and G. Fichera, *Trattato di analisi matematica* (Tumminelli, Roma, 1954), p. 43.

This paper has focussed on the derivation of explicit expressions for the values of the (holomorphic) potential $V(r)$ and its derivatives at the origin in terms of the S-wave scattering phase shift $\eta(k)$, the (negative) S-wave bound-state energies E_n, and the bound-state normalization coefficients C_n [defined in Eq. (2.10) and (2.9)]. These expressions are condensed in the formula

$$V^{(p)}(0) = \sum [p, l \mid s_2, \cdots, s_{l+1}] \prod_{j=2}^{l+1} J_j^{s_j}, \quad (4.19)$$

the sum being extended over all positive integral values of l and s_{l+1} and over nonnegative integral values of s_j, $j = 2, 3, \cdots, l$, which are consistent with the dimensional condition

$$\sum_{j=2}^{l+1} j s_j = p + 2. \quad (4.22)$$

The universal coefficients $[p, l \mid s_2, \cdots, s_{l+1}]$ depend only on the indices shown and are tabulated up to $p = 10$ in Table I. The even quantities J_{2j} are essentially the generalized momentum moments (of odd order) of the phase shift; they depend on the phase shift and on the bound state energies, but are independent of the normalization coefficients:

$$J_{2j} = (-)^{j+1} 2^{2j+2} \left\{ \frac{2}{\pi(2j-1)!} \int_0^\infty dk\, k^{2j-1} \frac{d^{2j-1}}{dk^{2j-1}} \times [k^{2j-1}\eta(k)] + \frac{1}{j} \sum_n E_n^j \right\}, \quad (4.20)$$

$$= (-)^{j+1} 2^{2j+2} \left\{ -\frac{2}{\pi} \int_0^\infty dk\, k^{2j-1} \times \left[\eta(k) - \sum_{n=0}^{j-1} a_n k^{-2n-1} \right] + \frac{1}{j} \sum_n E_n^j \right\}. \quad (6.1)$$

The coefficients a_n entering in this equation are those characterizing the asymptotic behavior of $\eta(k)$,

$$\eta(k) = \sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}). \quad (2.41)$$

The odd quantities J_{2j+1} are essentially the generalized momentum moments (of even order) of the spectral function $|f(k)|^{-2}$; they depend explicitly also upon the normalization coefficients C_n:

$$J_{2j+1} = (-)^{j+1} 2^{2j} \left\{ \frac{-2}{\pi(2j+1)!} \int_0^\infty dk\, k^{2j+1} \frac{d^{2j+1}}{dk^{2j+1}} \times [k^{2j} |f(k)|^{-2}] + \sum_n C_n E_n^{j-1} \right\}, \quad (4.2)$$

$$= (-)^{j+1} 2^{2j} \left\{ \frac{2}{\pi} \int_0^\infty dk\, k^{2j} \times \left[|f(k)|^{-2} - 1 - \sum_{n=0}^{j-1} (-)^n \tilde{g}_{2n} k^{-2n-2} \right] + \sum_n C_n E_n^{j-1} \right\}. \quad (6.2)$$

The coefficients $\tilde{g}_{2n}$ entering in this equation have been introduced previously; they characterize the asymptotic behavior of $|f(k)|^{-2}$ through

$$|f(k)|^{-2} = 1 + \sum_{n=0}^{N} (-)^n \tilde{g}_{2n} k^{-2n-2} + O(k^{-2N-4}), \quad (6.3)$$

as implied by Eqs. (2.22) and (4.18). The "spectral function" $|f(k)|^{-2}$ is expressed in terms of the phase shift and bound state energies by the well-known formula [see Eq. (2.29)]

$$|f(k)|^{-2} = e^{-2\Delta(k)} \left[\prod_n \left(1 - \frac{E_n}{k^2} \right) \right]^{-2}, \quad (6.4)$$

with

$$\Delta(k) = -\frac{2}{\pi} P \int_0^\infty dq \frac{q}{q^2 - k^2} \eta(q). \quad (2.50)$$

Below we display the first few cases contained in Eq. (4.19) above:

$$V(0) = -\tfrac{1}{4} J_2, \quad (1.1)$$

$$V'(0) = -J_3, \quad (6.5)$$

$$V''(0) = \tfrac{1}{8} J_2^2 - \tfrac{1}{4} J_4, \quad (6.6)$$

$$V'''(0) = J_2 J_3 - J_5, \quad (6.7)$$

$$V^{\mathrm{iv}}(0) = -\tfrac{1}{6} J_2^3 + 5 J_3^2 + \tfrac{1}{2} J_2 J_4 - \tfrac{1}{4} J_6. \quad (6.8)$$

In the very special case of an even potential, all the odd quantities J_{2j+1} vanish; this implies conditions on the scattering phase shift (and on the bound-state energies and normalization coefficients, if bound states are present). These conditions, Eqs. (5.1) or (5.2), yield (implicit) expressions of the bound state parameters in terms of the phase shift. For instance, if only one bound state is present, one finds

$$E_1 = \frac{1}{20} \frac{\int_0^\infty dk\, k^5 \frac{d^5}{dk^5} [k^8 (k^2 - E_1)^{-2} e^{-2\Delta(k)}]}{\int_0^\infty dk\, k^3 \frac{d^3}{dk^3} [k^6 (k^2 - E_1)^{-2} e^{-2\Delta(k)}]}, \quad (5.4)$$

$$C_1 = \frac{1}{3\pi} \int_0^\infty dk\, k^3 \frac{d^3}{dk^3} [k^6 (k^2 - E_1)^{-2} e^{-2\Delta(k)}], \quad (5.5)$$

with $\Delta(k)$ given by Eq. (2.50) reported above. Obviously, in this case the (even) derivatives of the potential at the origin are expressed as polynomials of the even coefficients J_{2j}, i.e., only in terms of the generalized momentum moments of the phase shift. This conclusion was already obtained in the work of Percival[5] (see Sec. 3).

In addition to these results, it is possible to obtain expressions for certain integrals over the potential,

namely

$$\int_0^\infty dr\, V(r) = -2a_0, \tag{H10}$$

$$\int_0^\infty dr\, V^2(r) = -2^3 a_1 + 4\sum_n C_n + \frac{4}{3\pi}\int_0^\infty dk\, k^3 \frac{d^3}{dk^3}\left[k^2 |f(k)|^{-2}\right], \tag{6.9}$$

where the coefficients a_n are those characterizing the asymptotic behavior of the scattering phase shift as given by Eq. (2.41) above, and may therefore also be expressed [see Eq. (D7)] through the equations

$$a_n = \frac{1}{(2n)!}\int_0^\infty dk\, k^{2n} \frac{d^{2n+1}}{dk^{2n+1}}\left[k^{2n+1}\eta(k)\right]. \tag{6.10}$$

These expressions are obtained from Verde's results on the asymptotic behavior of the scattering phase shift; to obtain Eq. (6.9) one needs, in addition, the expression for $V'(0)$ in terms of the phase shift and the bound state parameters, Eq. (6.5), derived in this paper. As is emphasized at the end of Appendix H, it is not possible to obtain other simple expressions of this kind for the potential—for instance, an expression for the integral of $V^3(r)$. In fact, the next expression one would get by this method is one for the quantity

$$\int_0^\infty dr V^3(r) + \frac{1}{2}\int_0^\infty dr[V'(r)]^2, \tag{6.11}$$

and it is implied by Eq. (H12).

Finally, we report the exact expression, valid within the manifold of physically equivalent potentials, which is derived in Appendix B:

$$\int_0^\infty dr\, rV(r) - 2\sum_n \ln C_n = \int_0^\infty dr\, r\bar{V}(r) - 2\sum_n \ln \bar{C}_n. \tag{B13}$$

It is, on the other hand, easily seen that Eq. (B9) does not provide a new result, being directly implied by Eq. (6.9) above.

7. CONCLUSIONS

The principal aim of this paper has been to obtain explicit expressions for the values of a potential and of its derivatives at the origin in terms of its S-wave scattering phase shift and of the parameters of its S-wave bound states. The discussion has been restricted to potentials represented by holomorphic functions of r, and vanishing at infinity sufficiently fast for the usual treatment of scattering theory to apply. We have reviewed previous work on this subject and have discussed different routes to obtain this kind of result. Our attempt to explore all the available techniques was motivated by our intention to initiate with this paper a program aimed at the extension of this approach to more general cases.

The method which proved most convenient in the present case is that based on the asymptotic and analyticity properties of the scattering functions: the use of asymptotic techniques leads to explicit relationships between the features of the interaction at short distances and the parameters characterizing the asymptotic behavior of appropriate functions (or rather functionals) of the scattering phase shift; the analyticity properties characterizing the dependence of these functions upon the linear momentum k make it possible to express these asymptotic parameters as integrals extending over all physical values of k. These methods bear a certain resemblance to certain techniques recently employed in the theory of elementary particles, within the framework of dispersion relations and the so-called superconvergence sum rules.

As we have already emphasized, we hope that results of this kind will eventually be useful in phenomenological applications. We are, however, aware that the concentration on the properties of the potential in the neighborhood of the origin, which has been characteristic of this paper, does not appear, from this point of view, very promising. In fact, generally in the realm of elementary particle interactions and of nuclear physics, the potential model (even when it is adequate to describe the interaction at large distances) becomes inapplicable at short distances. Connected with this difficulty is the fact that the present approach concentrates upon the asymptotic behavior of the scattering parameters at large energies, namely in that very energy region where the potential model is phenomenologically inadequate; but this difficulty is, to a certain extent, bypassed in our results exploiting the analyticity properties of the scattering functions to obtain the parameters characterizing their asymptotic behavior from their over-all dependence upon the linear momentum k. Of course, the necessity to use the values of the scattering phase shift for all physical energies as input data remains an unavoidable difficulty.

We do not elaborate on the extension of these results to more general cases, except to note that the generalizations which are naturally suggested include nonholomorphic potentials, higher partial waves, multichannel problems and complex potentials, the scattering of Klein–Gordon and Dirac particles, and more general relativistic models. The principal hindrance to progress along these lines is connected with the (largely unexplored) complications which

are a feature of the asymptotic behavior of scattering phase shifts as soon as the restrictions to the case of holomorphic potentials and to S waves are dropped. Another promising direction of research, consistent with the philosophy advocated in this paper, but implying more radical modifications in the approach, takes as point of departure a kind of information derived from the scattering process different from that contained within a single scattering phase shift (or something analogous in more complicated cases) assumed known for all energies; for instance, the angular scattering amplitude at one energy, or the total cross section for all energies.

ACKNOWLEDGMENTS

The authors wish to acknowledge fruitful conversations on the asymptotic behavior of scattering phase shifts and on the recurrence relations of Appendix F with E. Presutti, and on the integration of functional equations with Professor G. Jona-Lasinio. We also wish to thank Dr. A. Morgana who wrote the Fortran program for the evaluation of the coefficients

$$\{n, l \,|\, s_1, \cdots, s_l\} \quad \text{and} \quad [n, l \,|\, s_2, \cdots, s_{l+1}].$$

APPENDIX A

In this Appendix we prove that the function $g(k)$ defined by Eq. (2.19) is holomorphic in the half-plane Im $k < 0$. As noted in Sec. 2, the meromorphy of this function is immediately implied by the well-known property of holomorphy for Im $k < 0$ of the Jost solution $f(k, r)$, and therefore also of the two functions $f'(k, 0)$ and $f(k)$. To complete the proof there remains to show that all the poles of the function $g(k)$, defined by Eq. (2.20), which occur at the zeros of $f(k)$ in the lower half-plane, are exactly cancelled by the poles explicitly contained in the sum on the right-hand side of Eq. (2.19). [That the only zeros of $f(k)$ in the lower half-plane occur at $k = -i(-E_n)^{\frac{1}{2}}$, where E_n indicate the energies of the bound states, is a well-known fact which we take for granted.] Thus it is required to be shown that

$$C_n = 2ip_n f'(-ip_n, 0)/\dot{f}(-ip_n), \tag{A1}$$

where we have used the notation

$$p_n = +(-E_n)^{\frac{1}{2}} \tag{A2}$$

and

$$\dot{f}(k) = df(k)/dk. \tag{A3}$$

This formula is well known[9]; we provide here a proof for completeness.

The proof is based on the differential equation satisfied by the function $\dot{f}(k, r) = \partial f(k, r)/\partial k$, which is obtained by straightforward differentiation from Eq. (2.4). We get

$$\ddot{f}(k, r) + [k^2 - V(r)]\dot{f}(k, r) = -2kf(k, r). \tag{A4}$$

Applying the Wronskian theorem, from this equation and Eq. (2.4) we obtain

$$[f'(k, r)\dot{f}(k, r) - \dot{f}'(k, r)f(k, r)]\Big|_{r=0}^{r=\infty} = 2k\int_0^\infty dr\, f^2(k, r). \tag{A5}$$

We use this equation for $k = -ip_n$, in which case the integral in the right-hand side is certainly convergent because

$$f(-ip_n, r) = f'(ip_n, 0)\varphi(E_n, r), \tag{A6}$$

where $\varphi(E_n, r)$ is the bound-state wavefunction. This relation is implied by Eq. (2.6), by the fact that

$$f(-ip_n) = 0, \tag{A7}$$

and by the normalization of $\varphi(E_n, r)$, Eq. (2.9).

Now inserting Eq. (A6) in Eq. (A5) and using the asymptotic vanishing of the bound-state wavefunction $\varphi(E_n, r)$ and Eq. (A7), we secure

$$\dot{f}(-ip_n, 0) \equiv \dot{f}(-ip_n) = 2ip_n f'(-ip_n, 0)\int_0^\infty dr\, \varphi^2(E_n, r). \tag{A8}$$

This equation together with Eq. (2.10) imply Eq. (A1). Q.E.D.

It should be emphasized that this proof used only properties of the Jost solution $f(k, r)$ in the lower half-k-plane, where this function (and, of course, its derivative) is holomorphic. On the other hand, from Eqs. (A6) and (2.6) one might infer the relation

$$f'(-ip_n) = -2p_n/\dot{f}(ip_n), \tag{A9}$$

which, inserted in Eq. (A1), yields the well-known formula[23]

$$C_n = -4ip_n^2/[\dot{f}(ip_n)\dot{f}(-ip_n)]. \tag{A10}$$

However, both Eqs. (A9) and (A10) hold only if the value $k = ip_n$ lies within the analyticity domain of the Jost function. This explains why to one Jost function $f(k)$ there need not correspond a unique set of normalization coefficients C_n.[23] On the other hand, Eq. (A1), which implies that

$$C_n = \lim_{k \to -ip_n} [(E_n - k^2)g(k)], \tag{A11}$$

shows that once the function $g(k)$ is assigned in the lower half-k-plane, where it is meromorphic, the

[23] See, for instance, the book by de Alfaro and Regge cited in Ref. 9.

coefficients C_n are determined. Thus, to a given Jost function $f(k)$ there may correspond more than one potential, while to a given $g(k)$, there corresponds a unique potential (see also Sec. 5 and Appendix H).

APPENDIX B

In this Appendix we outline the generalization of the approach of Sec. 2 which treats explicitly the dependence of the potential on the bound state parameters.

We begin with the derivative of the potential with respect to the bound-state normalization constants[2,15]

$$[\partial V(r)/\partial C_l] = -2(d/dr)[\varphi(E_l, r)]^2. \tag{B1}$$

The similarity of this equation to Eq. (2.1) implies immediately the recursive system of partial differential equations

$$\frac{\partial v_{n+2}}{\partial C_l} = -4E_l \frac{\partial v_n}{\partial C_l} + 2\sum_{m=0}^{n} \frac{n+m+2}{n+1} \binom{n+1}{m+1} v_{n-m} \frac{\partial v_m}{\partial C_l}, \tag{B2}$$

with boundary conditions

$$\partial v_0/\partial C_l = 0, \tag{B3}$$

$$\partial v_1/\partial C_l = -4. \tag{B4}$$

Here the quantities v_n are the same as were introduced by Eq. (2.31); namely, v_n is simply the nth derivative of the potential evaluated at the origin. Of course, this system of recursive differential equations should be integrated in conjunction with the system, Eq. (2.37). Note that Eqs. (B2) and (B3) also imply

$$\partial v_2/\partial C_l = 0. \tag{B5}$$

It is immediately seen that Eqs. (B3), (B4), and (B5) imply

$$A_0(a_0; E_n; C_n) = \bar{A}_0(a_0; E_n), \tag{B6}$$

$$A_1(a_0, a_1; E_n; C_n) = \bar{A}_1(a_0, a_1; E_n) - 4\sum_n C_n, \tag{B7}$$

$$A_2(a_0, a_1; E_n; C_n) = \bar{A}_2(a_0, a_1; E_n), \tag{B8}$$

where the quantities $\bar{A}$ are independent of the C_l's and the quantities A are those introduced in Sec. 2.

From Eq. (B1) it is easy to derive two simple exact relationships. The first is

$$\int_0^\infty dr\, V^2(r) - 4\sum_n C_n = \int_0^\infty dr\, \bar{V}^2(r) - 4\sum_n \bar{C}_n, \tag{B9}$$

where the potential $V(r)$ and $\bar{V}(r)$ must have the same phase shift and the same bound state energies. (They are thus physically equivalent; see Sec. 5.) This equation is proved by multiplying both sides of Eq. (B1) by $V(r)$ and integrating over r, exploiting the fact that the radial Schrödinger equation

$$\varphi''(E_l, r) + E_l\varphi(E_l, r) = V(r)\varphi(E_l, r) \tag{B10}$$

implies that the quantity $V(r)\varphi'(E_l, r)\varphi(E_l, r)$ is a perfect differential.

The second relationship is obtained after multiplying both sides of Eq. (B1) by r and integrating over r. Because by partial integration

$$\int_0^\infty dr\, r\varphi'(E_l, r)\varphi(E_l, r) = -\tfrac{1}{2}C_l^{-1}, \tag{B11}$$

we get

$$\frac{\partial}{\partial C_l}\int_0^\infty dr\, rV(r) = 2C_l^{-1}, \tag{B12}$$

which yields

$$\int_0^\infty dr\, rV(r) - 2\sum_n \ln C_n = \int_0^\infty dr\, r\bar{V}(r) - 2\sum_n \ln \bar{C}_n, \tag{B13}$$

where again the potentials $V(r)$ and $\bar{V}(r)$ must be physically equivalent.

The expression of the partial derivative of the potential with respect to the energies of the bound states is somewhat awkward[24]:

$$\frac{\partial V(r)}{\partial E_l} = -4\frac{d}{dr}\left\{\frac{1}{f(-ip_l)} f(-ip_l, r)\varphi(E_l, r) - \sum_n B_n\varphi^2(E_n, r)\right\}, \tag{B14}$$

with

$$B_n = C_n/(E_n - E_l), \quad \text{for} \quad l \neq n, \tag{B15a}$$

$$B_l = \frac{C_l}{4E_l}\left[1 - k\frac{d}{dk}\ln \dot{f}(k)\right]\Big|_{k=-ip_l}, \tag{B15b}$$

and

$$p_l^2 = -E_l, \quad p_l > 0. \tag{B16}$$

The dot indicates differentiation with respect to k.

In the light of the results of Sec. 4 it is easy to understand why these expressions are much more complicated than those for the derivative of the potential with respect to the normalization coefficients C_n; in fact, while the dependence of the quantities g_p, and therefore ultimately of the potential, on the parameters C_n is very simple [see Eqs. (4.16)–(4.17)], the dependence upon the parameters E_n is much more complicated for the odd quantities g_{2p+1} because, in addition to the explicit dependence in the sum, there

[24] This formula, first derived by Jost and Kohn (see Ref. 15) (who, however, do not write the explicit expressions of the quantities B_l), is misprinted in the book by de Alfaro and Regge (see Ref. 9), where it is written in the special case when only one bound state is present, altogether omitting the contribution of the sum on the right-hand side.

is a hidden dependence through the function $|f(k)|$ [see Eq. (G40)].

It is not possible to derive from Eq. (B14) recursion relations for the quantities $\partial v_n/\partial E_l$ as simple as in the preceding case. But of course one can derive from it explicit expressions by direct computation. For instance, one readily gets

$$\frac{\partial V(0)}{\partial E_l} = -4, \tag{B17}$$

which is consistent with Eq. (1.1), and (somewhat less readily)

$$\frac{\partial V'(0)}{\partial E_l} = 8\sum_n B_n - \frac{8}{f(-ip_l)} f'(-ip_l, 0) \tag{B18a}$$

$$= -8[\tilde{g}(-ip_l) - p_l], \tag{B18b}$$

whose consistency with Eq. (6.5) is easily verified.

We close the discussion at this point because, as was already emphasized, the approach based on the variational and partial derivatives of the potential with respect to the scattering and bound state parameters is, in any case, less convenient to use in order to establish the final results than that of Sec. 4.

APPENDIX C

In this Appendix for the sake of completeness we report an explicit elementary derivation of the well-known formula, Eq. (3.4), which connects the derivative of the scattering phase shift $\eta(k)$ due to the potential $V(r)$ to the increment $n_I(E)$ in the density of states due to the same potential.[25]

We enclose the system in a spherical box of radius R. In the absence of potential, the allowed states are characterized by the wavefunctions $\sin(k_n r)$, with

$$k_n R = n\pi. \tag{C1}$$

Thus the density of states in this case is

$$n_0(E) = \frac{\Delta n}{\Delta E} = \frac{1}{2k}\frac{\Delta n}{\Delta k} = \frac{R}{2\pi k}. \tag{C2}$$

On the other hand, if the potential is present and if R is large enough that the asymptotic expression $\sin[kR + \eta(k)]$ may be used in place of the radial wavefunction, then, in place of Eq. (C1), we have

$$k_n R + \eta(k_n) = n\pi. \tag{C3}$$

Here, of course, $\eta(k)$ is the phase shift due to the potential. Thus the density of states in this case is

$$n(E) = \frac{\Delta n}{\Delta E} = \frac{1}{2\pi k}\left(R + \frac{\Delta n}{\Delta k}\right). \tag{C4}$$

But, by definition,

$$n_I(E) = n(E) - n_0(E). \tag{C5}$$

Inserting expressions (C2) and (C4) in this equation, we get, in the limit of the continuum ($R = \infty$), Eq. (3.4). Q.E.D.

APPENDIX D

In this Appendix we collect some useful operator and integral formulas.

First of all we introduce the operator

$$D_k = k\frac{d}{dk}k, \tag{D1}$$

and we note the identity

$$D_k^n \equiv \left(k\frac{d}{dk}k\right)^n = k^n\frac{d^n}{dk^n}k^n, \tag{D2}$$

which is easily proved by induction.

A remarkable property of the operator D_k is expressed by the following formula:

$$D_k^n\sum_{m=1}^{N} f_m k^{-m} = 0, \quad \text{if} \quad n \geq N, \tag{D3a}$$

$$= (-)^n\sum_{m=n+1}^{N} f_m\frac{(m-1)!}{(m-n-1)!}k^{n-m}, \quad \text{if} \quad n < N. \tag{D3b}$$

We then consider a function $f(k)$ whose asymptotic expansion ($k \to \pm\infty$) is of the following type:

$$f(k) = k^{-p}\sum_{n=1-p}^{N} f_n k^{-2n} + O(k^{-p-2N-2}), \quad p = 0 \text{ or } 1. \tag{D4}$$

We also assume that $f(k)$ and all its derivatives are finite on the real axis, and that the asymptotic behaviors of the derivatives of $f(k)$ may be obtained by differentiating Eq. (D4).

For functions of this type we write the integral identity

$$\int_{-\infty}^{+\infty} dk\, k^{2n+p}\left[f(k) - k^{-p}\sum_{m=1-p}^{n} f_m k^{-2m}\right] = \frac{(-)^p}{(2n+p)!}\int_{-\infty}^{+\infty} dk\, D_k^{2n+p} f(k). \tag{D5}$$

[25] See, for instance, J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (John Wiley & Sons, Inc., New York, 1954).

The validity of this relation is easily proved by partial integration and the use of Eqs. (D3a) and (D4). Of course the value of p in Eqs. (D4) and (D5) must be the same.

Another useful formula, whose proof requires only one extra partial integration, is the following:

$$\int_{-\infty}^{+\infty} dk\, k^{2n+p+1} \frac{d}{dk}\left[f(k) - k^{-p} \sum_{m=1-p}^{n} f_m k^{-2m}\right] = (-)^{p+1} \frac{(2n+p+1)}{(2n+p)!} \int_{-\infty}^{+\infty} dk\, D_k^{2n+p} f(k). \quad \text{(D6)}$$

Again the value of p must be the same as in Eq. (D4).

It is easily seen that Eqs. (D5) and (D6) remain valid if the integral runs only over one semiaxis and if the function $f(k)$ has a definite parity, so that the integrands in Eqs. (D5) and (D6) are even.

Finally, we write integral expressions for the asymptotic coefficients f_n themselves:

$$f_n = \frac{(-)^{p+1}}{(2n+p-1)!} \int_0^{\infty} \frac{dk}{k} D_k^{2n+p} f(k). \quad \text{(D7)}$$

These are easily proved by partial integration and the use of Eq. (D3a). Again the value of p must be the same as in Eq. (D4). Note that the integrand on the right-hand side of this equation is an odd function of k if $f(k)$ has the same parity of its asymptotic expansion.

APPENDIX E

In this Appendix we investigate the expression

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr}[(T+V)^p - T^p]e^{ikr}, \quad \text{(E1)}$$

where

$$T = -(d^2/dr^2) \quad \text{(E2)}$$

and V is a holomorphic function of r which vanishes sufficiently fast at infinity to allow any number of integrations by parts with neglect of the contributions at both extremes.

We begin evaluating the integral

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr} f(r) T g(r) e^{ikr} = \frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk \times e^{2ikr} f(r)[-g''(r) - 2ikg'(r) + k^2 g(r)], \quad \text{(E3)}$$

where $f(r)$ and $g(r)$ are arbitrary functions which vanish asymptotically together with their derivatives. We use the identities

$$2ike^{2ikr} = \frac{d}{dr} e^{2ikr} \quad \text{(E4a)}$$

and

$$k^2 e^{2ikr} = -\frac{1}{4} \frac{d^2}{dr^2} e^{2ikr} \quad \text{(E4b)}$$

to eliminate the terms linear and quadratic in k on the right-hand side. We then integrate by parts to eliminate the derivatives acting on the exponential, and we thus secure

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr} f(r) T g(r) e^{ikr} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk \times e^{2ikr}[2f'(r)g'(r) - f''(r)g(r) - f(r)g''(r)]. \quad \text{(E5)}$$

Finally, using the expression

$$\delta(r) = \frac{1}{\pi} \int_{-\infty}^{+\infty} dk e^{2ikr}, \quad \text{(E6)}$$

where δ is the Dirac distribution, we obtain

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr} f(r) T g(r) e^{ikr} = \tfrac{1}{4}[2f'(r)g'(r) - f''(r)g(r) - f(r)g''(r)]\big|_{r=0}. \quad \text{(E7)}$$

At this point it is convenient to introduce the two differential operators $\vec{d}$ and $\overleftarrow{d}$, defined as follows:

$$f(r)\vec{d}g(r) = f(r)g'(r), \quad \text{(E8a)}$$

$$f(r)\overleftarrow{d}g(r) = f'(r)g(r). \quad \text{(E8b)}$$

We may then rewrite Eq. (E7) in the form

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr} f(r) T g(r) e^{ikr} = -\tfrac{1}{4} f(r)(\vec{d} - \overleftarrow{d})^2 g(r)\big|_{r=0}. \quad \text{(E9)}$$

On the other hand, we also have, for an arbitrary function $f(r)$,

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr} f(r) e^{ikr} = f(0). \quad \text{(E10)}$$

From this equation and Eq. (E9) we infer that

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} dr \int_{-\infty}^{+\infty} dk e^{ikr}[(T+V)^p - T^p]e^{ikr} = \{1 \cdot [V(r) - \tfrac{1}{4}(\vec{d} - \overleftarrow{d})^2]^p \cdot 1\}\big|_{r=0}. \quad \text{(E11)}$$

It is interesting to note a remarkable property of the operator $\vec{d} + \overleftarrow{d}$, namely that

$$f(r)(\vec{d} + \overleftarrow{d})g(r) = d/dr[f(r)g(r)]. \quad \text{(E12)}$$

Thus the operator $\vec{d} + \overleftarrow{d}$ commutes with a function of r; i.e., its position within a product is irrelevant:

$$(\vec{d} + \overleftarrow{d})f(r)g(r) = f(r)(\vec{d} + \overleftarrow{d})g(r) = f(r)g(r)(\vec{d} + \overleftarrow{d}). \quad \text{(E13)}$$

Unfortunately, this is not true for the case of the operator $\vec{d} - \overleftarrow{d}$.

We may use this property of the operator $\vec{d} + \overleftarrow{d}$ to express the right-hand side of Eq. (E11) in the form

$$\{1 \cdot [V(r) - \tfrac{1}{4}(\vec{d} - \overleftarrow{d})^2]^p \cdot 1\} = \sum_{l=0}^{p-1} \binom{p}{l} (-)^l 2^{-2l} \frac{d^{2l}}{dr^{2l}} \{1 \cdot [\overleftarrow{d} \cdot \vec{d} + V(r)]^{p-l} \cdot 1\}, \quad \text{(E14)}$$

which is obtained using the identity

$$(\vec{d} - \overleftarrow{d})^2 = (\vec{d} + \overleftarrow{d})^2 - 4\,\vec{d} \cdot \overleftarrow{d} \quad \text{(E15)}$$

and the binomial expansion, which is applicable thanks to the commutativity of $\vec{d} + \overleftarrow{d}$. This expansion may be useful because the operator $\vec{d} \cdot \overleftarrow{d}$ is such that

$$1 \cdot [\vec{d} \cdot \overleftarrow{d} + V(r)]^p \cdot 1 = V(r)[\overleftarrow{d} \cdot \vec{d} + V(r)]^{p-2} V(r), \quad p = 2, 3, \cdots. \quad \text{(E16)}$$

Moreover, this term contains derivatives only up to the order $p - 2$. On the other hand, the quantity on the left-hand side of Eq. (E14) contains derivatives up to the order $2p - 2$. Note that from Eqs. (E14) and (E16) one easily gets

$$1 \cdot [-\tfrac{1}{4}(\vec{d} - \overleftarrow{d})^2 + V(r)]^p \cdot 1 = (-)^{p+1} p 2^{2-2p} \frac{d^{2p-2}}{dr^{2p-2}} V(r) + \text{terms containing derivatives of order } 2p - 4 \text{ or less.} \quad \text{(E17)}$$

APPENDIX F

In this Appendix we discuss Verde's recursion relation, Eq. (4.6).

First of all we derive from them the more general recursion relations:

$$g_n(r) = (-)^p g^{(p)}_{n-p}(r) - \sum_{m=0}^{p-1} (-)^m \frac{d^m}{dr^m} \sum_{s=0}^{n-m-2} g_s(r) g_{n-m-2-s}(r), \quad (n \geq p). \quad \text{(F1)}$$

(We always adopt the convention that a sum vanishes if its upper limit is smaller than its lower limit.) This equation is true for $p = 1$ because, in this case, it is identical to Eq. (4.6). Let us then prove it by induction: Assuming that it is true for a certain value of p (and all values of n not smaller than p), we show its validity for the value $p + 1$. In fact, performing in the first term on the right-hand side, the substitution

$$g_{n-p}(r) = -g'_{n-p-1}(r) - \sum_{s=0}^{n-p-2} g_s(r) g_{n-p-2-s}(r), \quad (n > p), \quad \text{(F2)}$$

which follows from Eq. (4.6), we get

$$g_n(r) = (-)^{p+1} g^{(p+1)}_{n-p-1}(r) + (-)^{p+1} \frac{d^p}{dr^p} \sum_{s=0}^{n-p-2} g_s(r) g_{n-p-2-s}(r) - \sum_{m=1}^{p-1} (-)^m \frac{d^m}{dr^m} \sum_{s=0}^{n-m-2} g_s(r) g_{n-m-2-s}(r), \quad \text{(F3a)}$$

$$= (-)^{p+1} g^{(p+1)}_{n-p-1}(r) - \sum_{m=0}^{p} (-)^m \frac{d^m}{dr^m} \sum_{s=0}^{n-m-2} g_s(r) g_{n-m-2-s}(r), \quad (n > p). \quad \text{(F3b)}$$

But this equation is identical to Eq. (F1), except for the substitution of $p + 1$ in place of p. Q.E.D.

From Eq. (F1), using Leibniz formula for the differentiation of a product, setting $r = 0$, and using the definition equation (4.13), we obtain

$$\bar{g}^{(p)}_n = (-)^p g_n + (-)^p \sum_{m=0}^{p-1} (-)^m \sum_{t=0}^{m} \binom{m}{t} \times \sum_{s=0}^{n-m-2} \bar{g}^{(t)}_{s+t} \bar{g}^{(m-t)}_{n-2-s-t}, \quad (n \geq p), \quad \text{(F4)}$$

which becomes Eq. (4.12) if we substitute the index $u = s + t$ in the last sum for the index s.

Let us now introduce a procedure for the evaluation of the coefficients $(n, l \mid s_1, \cdots, s_l)$ of Eq. (4.15). We begin by recalling their definition:

$$V^{(n)}(r) = \sum (n, l \mid s_1, \cdots, s_l) \prod_{j=1}^{l} g_{s_j}(r), \quad \text{(F5)}$$

the sum extending over all positive values of l and non-negative values of s_j such that the dimensional restriction

$$\sum_{j=1}^{l} s_j + 2l = n + 2 \quad \text{(F6)}$$

holds. In the following we set to zero any coefficient $(n, l \mid s_1, \cdots, s_l)$ for which this condition is not satisfied.

It should be emphasized that the sum in Eq. (F5) contains, in general, several terms which are multiplied by the same (except for the ordering) monomial $\prod_{j=1}^{l} g_{s_j}$. Such a definition of the coefficients

$$(n, l \mid s_1, \cdots, s_l)$$

is more convenient to use in deriving recursion relations for them (see below); it ignores the fact that the ordering of the factors in the monomial $\prod_{j=1}^{l} g_{s_j}$ is unimportant.[26] A more compact formula for $V^{(n)}(r)$ is obtained by introducing the coefficients $\{n, l \mid s_1, \cdots, s_l\}$ through

$$\{n, l \mid s_1, \cdots, s_l\} = \sum (n, l \mid s_1, \cdots, s_l), \quad \text{(F7)}$$

the sum extending over all permutations over different indices s_j. As is well known, the number of these permutations is $l! \prod [m(s)!]^{-1}$, where $m(s)$ is the number of indices s_j which take the value s. Obviously the coefficients $\{n, l \mid s_1, \cdots, s_l\}$ are completely symmetrical in the indices s_j; they also vanish unless the dimensional condition (F6) holds. Using these coefficients, Eq. (F5) simplifies to

$$V^{(n)}(r) = \sum{}' \{n, l \mid s_1, \cdots, s_l\} \prod_{j=1}^{l} g_{s_j}(r), \quad \text{(F8)}$$

the sum now being extended only over different sets of indices s_j, a set being by definition independent from the order in which the indices s_j appear. The prime appended to the sum is a reminder of this fact.

We also note that it follows from Eq. (F6) that the maximum value of l entering in the sum on the right-hand side of Eqs. (F5) or (F8) is

$$l_{\mathrm{MAX}} = 1 + \tfrac{1}{2}n \quad \text{for even } n, \quad \text{(F9a)}$$

$$= 1 + \tfrac{1}{2}(n-1) \quad \text{for odd } n. \quad \text{(F9b)}$$

To establish a recursive equation for the evaluation of the coefficients $(n, l \mid s_1, \cdots, s_l)$, it is sufficient to differentiate Eq. (F5), and then use Eq. (4.6) in the right-hand side of Verde's recursion relations. In this manner one obtains

$$\begin{aligned}(n+1, l+1 \mid s_1, \cdots, s_{l+1}) &= -\sum_{j=1}^{l+1} (n, l+1 \mid s_1, \cdots, s_{j-1}, s_j - 1, s_{j+1}, \cdots, s_{l+1}) \\ &\quad - \sum_{j=1}^{l} (n, l \mid s_1, \cdots, s_{j-1}, s_j + s_{j+1} + 1, s_{j+2}, \cdots, s_{l+1}).\end{aligned} \quad \text{(F10)}$$

Note that the index s_i occurs in the coefficients appearing in the second sum on the right-hand side, for $i > j$, at the $(i-1)$th place. We use the convention that a coefficient vanishes if any one of its arguments s_j is negative.

The relationships of Eq. (F6) are recursive in the index n; i.e., they yield the values of the coefficients $(n+1, L \mid s_1, \cdots, s_L)$ for all values of L and $s_1, \cdots, s_L$, provided that the coefficients

$$(n, l \mid s_1, \cdots, s_l)$$

are known for all values of l and s_j. [The restriction to values consistent with Eq. (F6) is always implied.] On the other hand, it is easily seen that, for $n = 1$, there is only one coefficient, namely

$$(1, 1 \mid 1) = -1. \quad \text{(F11)}$$

From this one, using the recursive relations of Eq. (F10), it is possible to evaluate all the coefficients $(n, l \mid s_1, \cdots, s_l)$; from these, the symmetrical coefficients $\{n, l \mid s_1, \cdots, s_l\}$ are immediately obtained through Eq. (F7). These last coefficients are displayed, for n up to 10, in Table II.

A different, but less direct, procedure to evaluate these coefficients is based on recursive relations in the index l. To obtain such relations, however, it is necessary to introduce more general coefficients $(n, p, l \mid s_1, \cdots, s_l)$ through

$$\bar{g}_n^{(p)} = \sum (n, p, l \mid s_1, \cdots, s_l) \prod_{j=1}^{l} g_{s_j}, \quad (n \geq p), \quad \text{(F12)}$$

the sum again extending over all integral positive values of l and nonnegative values of s_j consistent with the dimensional equation (F6).

Inserting this equation in Eq. (F4) or (4.12), with some algebra we secure the recursion relations

$$(n, p, 1 \mid s) = (-)^p \delta_{ns}, \quad \text{(F13)}$$

$$\begin{aligned}(n, p, L \mid s_1, \cdots, s_L) &= (-)^p \sum_{m=0}^{p-1} (-)^m \\ &\quad \times \sum_{l=1}^{L-1} \sum_{t=T_1}^{T_2} \binom{m}{t} (\Sigma_l + 2l - 2, t, l \mid s_1, \cdots, s_l) \\ &\quad \times (n - \Sigma_l - 2l, m - t, L - l \mid s_{l+1}, \cdots, s_L),\end{aligned} \quad \text{(F14)}$$

with

$$\Sigma_l = \sum_{j=1}^{l} s_j \quad \text{(F15)}$$

and

$$T_1 = \max [0, \Sigma_l + 2l + m - n], \quad \text{(F16)}$$

$$T_2 = \min [m, \Sigma_l + 2l - 2]. \quad \text{(F17)}$$

Equation (F14) is clearly a recursion relation in the index l: inserting in its right-hand side the coefficients $(n, p, l \mid s_1, \cdots, s_l)$, assumed known for all n and p up to N and for all l up to L, we obtain these same coefficients for all n and p up to N and for $l = L + 1$.

On the other hand, Eq. (F13) provides an explicit formula for the coefficient with $L = 1$ (and arbitrary values of n and p). Therefore it is possible,

[26] It should, however, be emphasized that the coefficients $(n, l \mid s_1, \ldots, s_l)$ which are defined by the recursion relations (F10) are not completely symmetrical in the indices s_j.

TABLE II. The coefficients $\{n, l \mid s_1, \cdots, s_l\}$ of Eq. (F8) for n up to 10. These coefficients have been evaluated by computer from Eqs. (F7), and the recursion relations (F10) with (F11). The coefficients up to $n = 6$ have been checked by hand. All the nonvanishing coefficients are printed, except, of course, those obtainable by permutation of the indices s_j.

$\{1, 1 \mid 1\} = -1.$
$\{2, 1 \mid 2\} = 1$; $\{2, 2 \mid 0, 0\} = 1.$
$\{3, 1 \mid 3\} = -1$; $\{3, 2 \mid 1, 0\} = -4.$
$\{4, 1 \mid 4\} = 1$; $\{4, 2 \mid 2, 0\} = 6$, $\{4, 2 \mid 1, 1\} = 5$; $\{4, 3 \mid 0, 0, 0\} = 4.$
$\{5, 1 \mid 5\} = -1$; $\{5, 2 \mid 3, 0\} = -8$, $\{5, 2 \mid 2, 1\} = -18$; $\{5, 3 \mid 1, 0, 0\} = -34.$
$\{6, 1 \mid 6\} = 1$; $\{6, 2 \mid 4, 0\} = 10$, $\{6, 2 \mid 3, 1\} = 28$, $\{6, 2 \mid 2, 2\} = 19$; $\{6, 3 \mid 2, 0, 0\} = 68$,
$\{6, 3 \mid 1, 1, 0\} = 112$; $\{6, 4 \mid 0, 0, 0, 0\} = 34.$
$\{7, 1 \mid 7\} = -1$; $\{7, 2 \mid 5, 0\} = -12$, $\{7, 2 \mid 4, 1\} = -40$, $\{7, 2 \mid 3, 2\} = -68$;
$\{7, 3 \mid 3, 0, 0\} = -116$, $\{7, 3 \mid 2, 1, 0\} = -512$, $\{7, 3 \mid 1, 1, 1\} = -140$; $\{7, 4 \mid 1, 0, 0, 0\} = -496.$
$\{8, 1 \mid 8\} = 1$; $\{8, 2 \mid 6, 0\} = 14$, $\{8, 2 \mid 5, 1\} = 54$, $\{8, 2 \mid 4, 2\} = 110$, $\{8, 2 \mid 3, 3\} = 69$;
$\{8, 3 \mid 4, 0, 0\} = 180$, $\{8, 3 \mid 3, 1, 0\} = 984$, $\{8, 3 \mid 2, 2, 0\} = 660$, $\{8, 3 \mid 2, 1, 1\} = 1080$;
$\{8, 4 \mid 2, 0, 0, 0\} = 1240$, $\{8, 4 \mid 1, 1, 0, 0\} = 3048$; $\{8, 5 \mid 0, 0, 0, 0, 0\} = 496.$
$\{9, 1 \mid 9\} = -1$; $\{9, 2 \mid 7, 0\} = -16$, $\{9, 2 \mid 6, 1\} = -70$, $\{9, 2 \mid 5, 2\} = -166$, $\{9, 2 \mid 4, 3\} = -250$;
$\{9, 3 \mid 5, 0, 0\} = -262$, $\{9, 3 \mid 4, 1, 0\} = -1700$, $\{9, 3 \mid 3, 2, 0\} = -2828$,
$\{9, 3 \mid 3, 1, 1\} = -2310$, $\{9, 3 \mid 2, 2, 1\} = -3094$; $\{9, 4 \mid 3, 0, 0, 0\} = -2584$,
$\{9, 4 \mid 2, 1, 0, 0\} = -16944$, $\{9, 4 \mid 1, 1, 1, 0\} = -9240$; $\{9, 5 \mid 1, 0, 0, 0, 0\} = -11056.$
$\{10, 1 \mid 10\} = 1$, $\{10, 2 \mid 8, 0\} = 18$, $\{10, 2 \mid 7, 1\} = 88$, $\{10, 2 \mid 6, 2\} = 238$, $\{10, 2 \mid 5, 3\} = 418$,
$\{10, 2 \mid 4, 4\} = 251$; $\{10, 3 \mid 6, 0, 0\} = 364$, $\{10, 3 \mid 5, 1, 0\} = 2728$, $\{10, 3 \mid 4, 2, 0\} = 5392$,
$\{10, 3 \mid 3, 3, 0\} = 3344$, $\{10, 3 \mid 4, 1, 1\} = 4400$; $\{10, 3 \mid 3, 2, 1\} = 14608$; $\{10, 3 \mid 2, 2, 2\} = 3260$;
$\{10, 4 \mid 4, 0, 0, 0\} = 4808$, $\{10, 4 \mid 3, 1, 0, 0\} = 38896$, $\{10, 4 \mid 2, 2, 0, 0\} = 25956$,
$\{10, 4 \mid 2, 1, 1, 0\} = 84832$, $\{10, 4 \mid 1, 1, 1, 1\} = 11550$; $\{10, 5 \mid 2, 0, 0, 0, 0\} = 33168$,
$\{10, 5 \mid 1, 1, 0, 0, 0\} = 108416$; $\{10, 6 \mid 0, 0, 0, 0, 0, 0\} = 11056.$

starting from this equation and using Eq. (F14) recursively, to compute sequentially all the coefficients $(n, p, l \mid s_1, \cdots, s_l)$. Once these more general coefficients have been evaluated, we easily obtain the coefficients $(p, l \mid s_1 z \cdots, s_l)$ of Eqs. (4.15):

$$(p, l \mid s_1, \cdots, s_l) = (p, p, l \mid s_1, \cdots, s_l). \quad \text{(F18)}$$

For instance, in this manner one easily obtains

$$(n, 2 \mid n - s - 2, s) = (n, 2 \mid s, n - s - 2) = (-)^n \sum_{m=0}^{n-1} \sum_{t=T_1}^{T_2} \binom{m}{t}, \quad \text{(F19)}$$

with

$$T_1 = \max [0, s + 2 + m - n], \quad \text{(F20a)}$$

$$T_2 = \min [m, s]; \quad \text{(F20b)}$$

and, in particular, from this equation one gets

$$(n, 2 \mid 0, n - 2) = (n, 2 \mid n - 2, 0) = (-)^n (n - 1), \quad \text{(F21a)}$$

$$(n, 2 \mid 1, n - 3) = (n, 2 \mid n - 3, 1) = (-)^n \tfrac{1}{2}(n - 2)(n + 1). \quad \text{(F21b)}$$

APPENDIX G

In this Appendix we derive the coefficients of the asymptotic expansion of the function $\tilde{g}(k)$, defined by Eqs. (2.30). We use a technique employed by Roberts.[6]

We consider separately the expansion of the real and imaginary parts of $\tilde{g}(k)$, since we obviously have, from Eq. (4.18),

$$\text{Re } \tilde{g}(k) = \sum_{n=0}^{N} \tilde{g}_{2n+1} (-)^{n+1} (2k)^{-2n-2} + O(k^{-2N-4}), \quad (\text{Im } k = 0), \quad \text{(G1)}$$

$$\text{Im } \tilde{g}(k) = \sum_{n=0}^{N} \tilde{g}_{2n} (-)^n (2k)^{-2n-1} + O(k^{-2N-3}), \quad (\text{Im } k = 0). \quad \text{(G2)}$$

We treat first the imaginary part. Let us recall Eqs. (2.22) and (2.29′) writing

$$\text{Im } \tilde{g}(k) = k \left[1 - e^{-2\Delta(k)} \prod_n \left(1 - \frac{E_n}{k^2} \right)^{-2} \right], \quad (\text{Im } k = 0), \quad \text{(G3)}$$

where

$$\Delta(k) = -\frac{1}{\pi} p \int_{-\infty}^{+\infty} dq \, \frac{q}{q^2 - k^2} \eta(q). \quad (2.50)$$

It is convenient to discuss first the asymptotic expansion of $\Delta(k)$ setting

$$\Delta(k) = \sum_{n=1}^{N} d_n k^{-2n} + O(k^{-2N-2}). \quad (3.18)$$

We now prove that

$$d_p = -\frac{1}{\pi (2p - 1)!} \int_{-\infty}^{+\infty} dk \, D_k^{2p-1} \eta(k). \quad (3.17)$$

This formula from the paper of Buslaev and Faddeev[4] has been quoted in Sec. 3. The operator D_k is that defined in Appendix D.

Before providing a proof of this formula, we show how it may be obtained by a less rigorous but more straightforward approach. In fact, if we substitute in Eq. (2.50) the asymptotic expansion

$$(q^2 - k^2)^{-1} = -\sum_{p=1}^{\infty} q^{2p-2}k^{-2p}, \tag{G4}$$

we get formally

$$\Delta(k) = \sum_{p=1}^{\infty} k^{-2p}\,\frac{1}{\pi}\int_{-\infty}^{+\infty} dq\, q^{2p-1}\eta(q), \tag{G5}$$

which implies Eq. (3.17) because, by partial integration (see Appendix D),

$$-\frac{1}{(2p-1)!}\int_{-\infty}^{+\infty} dk\, D_k^{2p-1}\eta(k) = \int_{-\infty}^{+\infty} dk\, k^{2p-1}\eta(k). \tag{G6}$$

Of course the last two equations written above need not make sense because the integrals on the right-hand side generally do not converge. Both equations, however, become valid, provided that an appropriate part of its asymptotic expansion is subtracted from the function $\eta(k)$, so as to make the integrals convergent. Specifically, in place of $\eta(k)$ in both equations we must substitute

$$\eta(k) - \sum_{n=0}^{p-1} a_n k^{-2n-1}, \tag{G7}$$

where the coefficients a_n are those of Eq. (2.41). It has been proved in Appendix D that with this substitution, Eq. (G6) holds true. There remains now to show that with this substitution also Eq. (G5) becomes correct.

To prove this result we introduce, after Roberts,[6] the function

$$h(k) = \frac{d}{dk}\ln f(k) \equiv \frac{d}{dk} f(k)/f(k) \tag{G8a}$$

$$= \frac{d}{dk}\Delta(k) + \frac{d}{dk}\sum_n \ln\left(1 - \frac{E_n}{k^2}\right) + i\frac{d}{dk}\eta(k), \tag{G8b}$$

and its asymptotic expansion

$$h(k) = i\sum_{n=0}^{N}(-2ik)^{-n-2}h_n + O(k^{-N-3}). \tag{G9}$$

Equation (G8b) follows from Eqs. (2.12) and (2.29′). A comparison of this expansion with those for $\Delta(k)$ and $\eta(k)$, Eqs. (3.18) and (2.41), implies [through Eq. (G8b)] the relations

$$d_p = \frac{(-)^p}{p}\,2^{-2p-2}h_{2p-1} + \frac{1}{p}\sum_n E_n^p, \tag{G10}$$

$$a_p = \frac{(-)^p}{2p+1}\,2^{-2p-2}h_{2p}. \tag{G11}$$

On the other hand, integrating the function

$$F(k) = k^{2p}\left[h(k) - i\sum_{n=0}^{2p-1} h_n(-2ik)^{-n-2}\right] \tag{G12}$$

along a contour composed of a large semicircle in the lower half-k-plane and of the real axis indented at $k = 0$, and using the fact that the function $h(k)$ is meromorphic in the lower half-plane and that its poles there have unit residues and correspond to the bound states, one finds[27]

$$(-)^p 2^{-2p-2}h_{2p-1} = -\sum_n E_n^p - \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, k^{2p}\frac{d}{dk} \times \left[\eta(k) - \sum_{n=0}^{p-1} a_n k^{-2n-1}\right]. \tag{G13}$$

From this equation and Eq. (G10) we get

$$d_p = -\frac{1}{2\pi p}\int_{-\infty}^{+\infty} dk\, k^{2p}\frac{d}{dk}\left[\eta(k) - \sum_{n=0}^{p-1} a_n k^{-2n-1}\right]. \tag{G14}$$

This equation, by Eq. (D6), becomes Eq. (3.17). Q.E.D.

Let us now return to Eq. (G3). It is convenient to introduce the auxiliary function

$$L(k) = \ln\left[1 - k^{-1}\,\mathrm{Im}\,\tilde{g}(k)\right] \tag{G15a}$$

$$= -2\Delta(k) - 2\sum_n \ln\left(1 - \frac{E_n}{k^2}\right) \tag{G15b}$$

and its asymptotic expansion

$$L(k) = \sum_{n=1}^{N} L_n k^{-2n} + O(k^{-2N-2}). \tag{G16}$$

A comparison of this expansion with Eq. (3.18) implies [through Eqs. (G15b) and (3.17)]

$$L_p = \frac{2}{\pi(2p-1)!}\int_{-\infty}^{+\infty} dk\, D_k^{2p-1}\eta(k) + \frac{2}{p}\sum_n E_n^p. \tag{G17}$$

This is an explicit, simple expression for L_p in terms of the phase shift and the bound state energies.

There remains now to express the quantity $\tilde{g}_{2p}$ in terms of the coefficients L_n. To achieve this aim we must first solve the following algebraic problem: Let

$$z(x) = l^{y(x)}, \tag{G18}$$

$$y(x) = \sum_{n=1}^{\infty} y_n x^n, \tag{G19}$$

$$z(x) = 1 + \sum_{n=1}^{\infty} z_n x^n, \tag{G20}$$

and express z_p in terms of the y_n's.

[27] Note that we utilize the validity of the asymptotic expansion not only on the real axis but also in the lower half-plane.

The solution of this problem may be found by noting that

$$z_n = \frac{1}{n!} z^{(n)}(0) \tag{G21}$$

and

$$y_n = \frac{1}{n!} y^{(n)}(0), \tag{G22}$$

the parenthetical superscript indicating the order of differentiation, and using the formula for the differentiation of a composite function,[28] i.e.,

$$\frac{d^n}{dx^n} F[y(x)] = n! \sum \frac{d^m}{dy^m} F[y(x)] \prod_{j=1}^{l} \left\{ \frac{1}{s_j!} \left[\frac{1}{j!} y^{(j)}(x) \right]^{s_j} \right\}, \tag{G23}$$

where

$$m = \sum_{j=1}^{l} s_j \tag{G24}$$

and the sum is extended over all positive integral values of l and s_l and nonnegative integral values of s_j, $(j = 1, 2 \cdots, l-1)$, such that

$$\sum_{j=1}^{l} j s_j = n. \tag{G25}$$

In this manner we find

$$z_p = \sum \prod_{j=1}^{l} \left[\frac{1}{s_j!} y_j^{s_j} \right], \tag{G26}$$

this sum being now extended over all the positive integral values of l and s_l and the nonnegative values of s_j, $(j = 1, 2 \cdots, l-1)$, such that

$$\sum_{j=1}^{l} j s_j = p. \tag{G27}$$

But since Eq. (G15a) implies

$$-k^{-1} \operatorname{Im} \tilde{g}(k) = e^{L(k)} - 1, \tag{G28}$$

we may now set

$$x = k^{-2}, \tag{G29}$$

$$-k^{-1} \operatorname{Im} \tilde{g}(k) = z(x) - 1, \tag{G30}$$

$$L(k) = y(x). \tag{G31}$$

Then, from Eqs. (G2) and (G20) we find

$$\tilde{g}_{2p} = (-)^{p+1} 2^{2p+1} z_{p+1}, \tag{G32}$$

and from Eqs. (G16) and (G19)

$$y_n = L_n. \tag{G33}$$

Thus, in conclusion,

$$\tilde{g}_{2p} = (-)^{p+1} 2^{2p+1} \sum \prod_{j=1}^{l} \left[\frac{1}{s_j!} L_j^{s_j} \right], \tag{G34}$$

the sum being extended over all positive integral values of l and s_l and nonnegative integral values of s_j, $(j = 1, 2, \cdots, l-1)$, such that

$$\sum_{j=1}^{l} j s_j = p + 1. \tag{G35}$$

Note that this last equation implies that the maximum value of l in Eq. (G34) cannot exceed

$$l_{\mathrm{MAX}} = p + 1. \tag{G36}$$

In particular, from these equations we obtain:

$$\tilde{g}_0 = -2L_1, \tag{G37a}$$

$$\tilde{g}_2 = 2^3 \left\{ \frac{1}{2!} L_1^2 + L_2 \right\}, \tag{G37b}$$

$$\tilde{g}_4 = -2^5 \left\{ \frac{1}{3!} L_1^3 + L_1 L_2 + L_3 \right\}, \tag{G37c}$$

$$\tilde{g}_6 = 2^7 \left\{ \frac{1}{4!} L_1^4 + \frac{1}{2!} L_1^2 L_2 + \frac{1}{2!} L_2^2 + L_1 L_3 + L_4 \right\}, \tag{G37d}$$

$$\tilde{g}_8 = -2^9 \left\{ \frac{L_1^5}{5!} + \frac{L_1^3 L_2}{3!} + \frac{L_1 L_2^2}{2!} + \frac{L_1^2 L_3}{2!} + L_2 L_3 + L_1 L_4 + L_5 \right\}, \tag{G37e}$$

$$\tilde{g}_{10} = 2^{11} \left\{ \frac{L_1^6}{6!} + \frac{L_1^4 L_2}{4!} + \frac{L_1^2 L_2^2}{2!\,2!} + \frac{L_2^3}{3!} + L_1 L_2 L_3 + \frac{L_1^3 L_3}{3!} + \frac{L_2^2}{2!} + L_2 L_4 + \frac{L_1^2 L_4}{2!} + L_1 L_5 + L_6 \right\}. \tag{G37f}$$

Let us proceed now to discuss the odd coefficients $\tilde{g}_{2p+1}$, namely those appearing in the asymptotic expansion of $\operatorname{Re} \tilde{g}(k)$ for real k. We use again the technique which we used above, except for the fact that we now integrate the function[27]

$$F(k) = k^{2p+1} \left[\tilde{g}(k) - \sum_{n=0}^{2p+1} \tilde{g}_n (-2ik)^{-n-1} \right]. \tag{G38}$$

The contour is the same as in the previous case. Note that now $F(k)$ is holomorphic in the lower half-plane (see Appendix A). Considering only the imaginary part of the result, in this manner we secure

$$\tilde{g}_{2p+1} = \frac{(-)^{p+1}}{\pi} 2^{2p+2} \int_{-\infty}^{+\infty} dk\, k^{2p+1} \times \left[\operatorname{Im} \tilde{g}(k) - \sum_{n=0}^{p} (-)^n \tilde{g}_{2n} (2k)^{-2n-1} \right], \tag{G39}$$

which, using Eqs. (2.22) and (D5), may be written

[28] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products* (Academic Press Inc., New York, 1965), Eq. 0.430. Note, however, that there is a misprint in this formula: The restriction to "positive integral values" should be corrected to read "nonnegative integral values."

in the form[29]

$$\tilde{g}_{2p+1} = \frac{(-)^{p+1}2^{2p+2}}{\pi(2p+2)!}\int_{-\infty}^{+\infty} dk\, D_k^{2p+2}[1-|f(k)|^{-2}] \tag{G40a}$$

$$= \frac{(-)^{p+1}2^{2p+2}}{\pi(2p+3)!}\int_{-\infty}^{+\infty} dk\, D_k^{2p+3}[k^{-1}|f(k)|^{-2}]. \tag{G40b}$$

The second line of this equation is obtained from the first by an integration by parts.

APPENDIX H

In this Appendix we review the approach of Roberts,[6] and we compare his results with those obtained in Sec. 4.

Roberts concentrates on the function $h(k)$ of Eqs. (G8). We have already shown in Appendix G how it is possible to express the odd coefficients of its asymptotic behavior in terms of energy moments of the scattering phase shift. This connection is given explicitly by Eq. (G13), which we may rewrite as

$$h_{2p-1} = (-)^{p+1}p2^{2p+1}L_p. \tag{H1}$$

The coefficients L_p are defined in Eq. (G17).

On the other hand, via the Verde technique, it is possible to connect the coefficients h_{2p-1} to the potential and its derivatives at the origin. Roberts uses only the explicit expressions for h_1 and h_3 which are easily obtained from the explicit formulas of Verde. We indicate here a more systematic approach; the results are, of course, the same ones we have already obtained (more directly) in the previous two Appendices.

We begin noticing that from Eqs. (G8), (4.1), and (4.3) it follows that

$$h(k) = -\int_0^\infty dr\,\frac{d}{dk}\,g(k,r). \tag{H2}$$

This equation, through Eqs. (G9) and (4.5), implies that

$$h_p = -2(p+1)\int_0^\infty dr\, g_p(r). \tag{H3}$$

We now show that the odd coefficients $g_{2p+1}(r)$, which are connected to the coefficients h_{2p+1}, are in fact an exact differential of a combination of even coefficients $g_{2n}(r)$. It is therefore possible to express the quantities h_{2p+1} in terms of the quantities $g_{2n} \equiv g_{2n}(0)$.

To achieve this aim it is convenient to introduce the two auxiliary functions $g_\pm(k,r)$, connected to the function $g(k,r)$ of Sec. 4 through

$$g_\pm(k,r) = \tfrac{1}{2}[g(k,r) \pm g(-k,r)]. \tag{H4}$$

From this definition and the differential equation satisfied by $g(k,r)$, Eq. (4.2), we easily obtain

$$g'_-(k,r) = 2ikg_+(k,r) - 2g_+(k,r)g_-(k,r). \tag{H5}$$

This equation may be rewritten in the form

$$g_+(k,r) = -\frac{1}{2}\frac{d}{dr}\ln\left[1-\left(\frac{1}{ik}\right)g_-(k,r)\right]. \tag{H6}$$

If in both sides of this equation we insert the asymptotic expansions

$$g_+(k,r) = \sum_{n=0}^{N} g_{2n+1}(r)(-2ik)^{-2n-2} + O(k^{-2N-4}), \tag{H7a}$$

$$g_-(k,r) = \sum_{n=0}^{N} g_{2n}(r)(-2ik)^{-2n-1} + O(k^{-2N-3}), \tag{H7b}$$

which obviously follow from the asymptotic expansion of $g(k,r)$, Eq. (4.5), and the definitions of $g_+(k,r)$ and $g_-(k,r)$, Eq. (H4), we find

$$g_{2p+1}(r) = \frac{1}{2}\frac{d}{dr}\sum(-2)^m(m-1)!\prod_{j=1}^{l}\left[\frac{1}{s_j!}\,g_{2j-2}^{s_j}(r)\right], \tag{H8a}$$

the sum being extended over all sets of positive integral values of l and s_l and nonnegative integral values of s_j, $(j = 1, 2, \cdots, l-1)$, such that

$$\sum_{j=1}^{l} js_j = p+1, \tag{H8b}$$

and with

$$m = \sum_{j=1}^{l} s_j. \tag{H8c}$$

This equation is derived by a technique similar to that used to derive Eq. (G26).

Inserting this equation in Eq. (H3) and using Eq. (4.8), we secure

$$h_{2p+1} = 2(p+1)\sum(-2)^m(m-1)!\prod_{j=1}^{l}\left(\frac{1}{s_j!}\,g_{2j-2}^{s_j}\right), \tag{H9}$$

the sum being extended over the same values of l, p, and s_j as specified above.

This last equation expresses the energy moments of the phase shift through the values of the potential and its derivatives at the origin, which are connected to the coefficients g_n through Verde's recursion relations, as is discussed in Sec. 4 and Appendix F. If, on the other hand, it is our purpose to express the values of the potential and its derivatives at the origin in terms of the scattering parameters, then we should invert Eq. (H9)—i.e., derive from it an expression for each g_{2p} in terms of the h_{2n+1}'s. In this manner we would recover Eq. (G34). In fact, it is apparent that the procedure we have now described is the inverse of that of Appendix G, just as the logarithm is the inverse of the exponential.

[29] Of course this equation might have been inferred, in a straightforward but nonrigorous way, expanding directly under the integral sign in Eq. (2.23), as was done at the beginning of this Appendix to obtain the coefficients d_p.

However, it should be emphasized that with the approach we have just described it is impossible to obtain the odd coefficients $\tilde{g}_{2p+1}$, which are connected to the moments of the "spectral function" $|f(k)|^{-2}$ (see Appendix G). On the other hand, as is clear from the discussion of Appendix F (see, also, Sec. 5), only if we know all the coefficients g_n—and this requires the knowledge of all the coefficients $\tilde{g}_n$ and, in addition, of the bound-state normalization coefficients C_n [see Eqs. (4.16) and (4.17)]—only then can we obtain all the derivatives of the potential.

The reason why the approach of Sec. 4 is more powerful than that of Roberts[6] is that the function $g(k)$, on which we concentrate, contains more information than the function $h(k)$, or, for that matter, $f(k)$. This is easily recognized because the function $g(k)$ contains all the information which is required for the complete specification of the potential (see Appendix A and Sec. 5); in fact, its poles in the lower half-plane, through their positions and their residues, characterize both the energies and the normalization coefficients of the bound states. On the other hand, the Jost function $f(k)$, as is well known, characterizes the bound-state energies, but not the bound-state normalization coefficients (see Appendix A). It is an additional boon that the asymptotic behavior of the function $g(k)$ turns out to be more directly related (via the Verde technique) to the interaction.

Finally, let us note that, by considering the even quantities h_{2n} as well as the odd quantities h_{2n+1}, one does not get any additional information on the values of the potential and its derivatives at the origin because the even functions $g_{2n}(r)$ [which yield the even coefficients h_{2n} through Eq. (H3)] cannot be expressed as exact differentials. In fact, the consideration of the even coefficients h_{2n} would simply reproduce Verde's results for the asymptotic expansion of the scattering phase shift because the even coefficients h_{2n} are directly related to the asymptotic coefficients a_n of the phase shift through Eq. (G11). For instance, the first three expressions thus obtained read

$$a_0 = -\frac{1}{2}\int_0^\infty V(r)\,dr, \tag{H10}$$

$$a_1 = -\frac{1}{2^3}\left[V'(0) + \int_0^\infty V^2(r)\,dr\right], \tag{H11}$$

$$a_2 = \frac{1}{2^5}\left[V'''(0) - 3[V^2(0)]' - \int_0^\infty [V'(r)]^2\,dr - 2\int_0^\infty [V(r)]^3\,dr\right]. \tag{H12}$$

Of course, expressions of this kind also qualify as exact explicit relationships between the scattering parameters and the potential; but unfortunately the information on the potential they yield is quite "implicit" because the potential enters in these formulas in a complicated way, involving its powers and the powers of its derivatives not only at the origin but also as arguments of integrals extending over all (positive) values of r.

APPENDIX I

In this Appendix we discuss some consequences of Theorem 3 of Sec. 5 for the very special class of even Bargmann's potentials.[9,21] As in the rest of the paper, we consider only S waves.

By definition a Bargmann potential is a potential whose Jost function $f(k)$ is a rational function of k:

$$f(k) = \prod_{j=1}^{n}\left[\frac{k-\alpha_j}{k-\beta_j}\right]. \tag{I1}$$

The constants β_j must satisfy the condition

$$\text{Im}\,\beta_j > 0 \tag{I2}$$

in order that the corresponding potential vanish asymptotically sufficiently fast; the constants α_j must satisfy the same condition, i.e.,

$$\text{Im}\,\alpha_j > 0, \tag{I3}$$

unless they are purely imaginary, in which case they may also occur in the lower half-plane, corresponding to bound states with energy $E_j = \alpha_j^2$:

$$\text{Im}\,\alpha_j < 0, \quad \text{Re}\,\alpha_j = 0, \quad j = 1, 2, \cdots, N. \tag{I4}$$

The number $N\,(\leq n)$ is the number of bound states. Finally, the reality of the potential implies that both the α_j's and the β_j's either be purely imaginary or occur in real conjugate pairs:

$$\alpha_j = -\alpha_i^*, \quad \beta_j = -\beta_i^*. \tag{I5}$$

Theorem 1: The functions $g(k)$ and $\tilde{g}(k)$, corresponding to a Bargmann potential, are rational functions of k.

The validity of this theorem is directly implied by the definition of a Bargmann potential [Eq. (I1) above] and by Eqs. (2.22), (2.23), and (2.19).

Theorem 2: The function Re $g(k)$, corresponding to an even Bargmann potential, vanishes identically.

This theorem is a direct consequence of Theorem 1 above and of Theorem 3 of Sec. 5 [see Eq. (5.2)].

Corollary 2.1: There does not exist any even Bargmann potential without bound states.

This result is a direct consequence of the preceding theorem and of the equality of $g(k)$ to $\tilde{g}(k)$ if there are no bound states [see Eq. (2.19)] because the holomorphy of $\tilde{g}(k)$ in the lower half-plane implies that it must vanish identically if its real part vanishes on the real axis.

Corollary 2.2: There exists one and only one even Bargmann potential with given bound state energies E_n and normalization coefficients C_n.

In fact, from Theorem 2 above and Eq. (2.19) we get

$$\text{Re } \tilde{g}(k) = \sum_n \frac{C_n}{k^2 - E_n}, \quad \text{Im } k = 0. \tag{I6}$$

From this equation and the dispersion relation implied by the holomorphy of $\tilde{g}(k)$ in the lower half-plane we get

$$\text{Im } \tilde{g}(k) = -k \sum_n \frac{C_n |E_n|^{-\frac{1}{2}}}{k^2 - E_n}, \quad \text{Im } k = 0. \tag{I7}$$

Thus we see that once the bound state energies E_n and the normalization coefficients C_n are given, the functions $\tilde{g}(k)$ (and therefore also the phase shift and the potential) is uniquely determined.

For instance, the only even Bargmann potential having one bound state with binding energy E and normalization constant C is

$$V(r) = 4 \frac{C}{p} \frac{d}{dr} \times \left[\frac{\sinh (br) \cosh (pr)}{(p - b) \cosh (p + b)r - (p + b) \cosh (p - b)r} \right], \tag{I8}$$

and the corresponding phase shift is

$$\eta(k) = \arctan (p/k) + \arctan (b/k), \tag{I9}$$

with

$$p = + |E|^{\frac{1}{2}} \tag{I10}$$

and

$$b = (p^2 + C/p)^{\frac{1}{2}}. \tag{I11}$$

We emphasize that the requirement that the corresponding potential be of Bargmann type and even implies that the phase shift is determined once the bound-state parameters are assigned, and, in fact, one easily finds the general formula

$$\eta(k) = 2 \sum_{n=1}^{N} \arctan \left(\frac{|E_n|^{\frac{1}{2}}}{k} \right) - \frac{k}{2\pi} P \int_{-\infty}^{+\infty} dq \, \frac{1}{q^2 - k^2} \ln \left[1 + \sum_{n=1}^{N} \frac{C_n |E_n|^{-\frac{1}{2}}}{q^2 - E_n} \right]. \tag{I12}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Relativistic Kinetic Equations Including Radiation Effects.* I. Vlasov Approximation

RÉMI HAKIM†
Department of Physics, Temple University, Philadelphia, Pennsylvania
AND
ANDRÉ MANGENEY
Laboratoire d'Astronomie Spatiale, Observatoire de Meudon, France‡

(Received 20 June 1967)

In a preceding series of papers, we developed relativistic statistical mechanics of electromagnetically interacting particles. Here we derive a hierarchy of manifestly covariant equations for the reduced densities. This hierarchy is valid at order one in a parameter which is closely related to radiation phenomena. At the lowest order (absence of correlation between particles), we derive a kinetic equation which is a modification of the relativistic Vlasov equation. The added term introduces the effects of the emission of collective radiation. A dimensional analysis shows that this term is important when $mc^2 \sim kT$, i.e., when the plasma is relativistic. The modified Vlasov equation is shown to be irreversible: an H theorem is proved. Several applications are studied: dispersion relations, hydrodynamical equations. We have also discussed the Einstein–Ritz controversy on the connection between retarded actions and irreversibility.

1. INTRODUCTION

In a preceding series of papers,[1,2] we develop manifestly covariant statistical mechanics. In particular, we studied electromagnetic interactions and obtained a rigorous hierarchy of equations for suitable reduced densities. This formalism presents several peculiarities. First it is completely free of the usual self-energy divergences of classical field theory; in other words, it is renormalized. Second, the formalism includes all the effects of radiation phenomena through a *finite* number of additional variables.

Accordingly, the present series of papers is mainly concerned with the study of manifestly covariant

* This work was begun while one of the authors (R. H.) was at Orsay; at Temple University it was supported by the Aerospace Research Laboratories of the Office of Aerospace Research, USAF.
† Permanent address: Laboratoire de Physique Théorique et Hautes Energies, 91 Orsay, France.
‡ Postal address: Laboratoire d'Astronomie Spatiale, Observatoire de Meudon, 92 Meudon, France.
[1] R. Hakim, J. Math. Phys. **8**, 1315 (1967).
[2] R. Hakim, J. Math. Phys. **8**, 1379 (1967).

Corollary 2.2: There exists one and only one even Bargmann potential with given bound state energies E_n and normalization coefficients C_n.

In fact, from Theorem 2 above and Eq. (2.19) we get

$$\text{Re } \tilde{g}(k) = \sum_n \frac{C_n}{k^2 - E_n}, \quad \text{Im } k = 0. \tag{I6}$$

From this equation and the dispersion relation implied by the holomorphy of $\tilde{g}(k)$ in the lower half-plane we get

$$\text{Im } \tilde{g}(k) = -k \sum_n \frac{C_n |E_n|^{-\frac{1}{2}}}{k^2 - E_n}, \quad \text{Im } k = 0. \tag{I7}$$

Thus we see that once the bound state energies E_n and the normalization coefficients C_n are given, the functions $\tilde{g}(k)$ (and therefore also the phase shift and the potential) is uniquely determined.

For instance, the only even Bargmann potential having one bound state with binding energy E and normalization constant C is

$$V(r) = 4 \frac{C}{p} \frac{d}{dr} \times \left[\frac{\sinh (br) \cosh (pr)}{(p - b) \cosh (p + b)r - (p + b) \cosh (p - b)r} \right], \tag{I8}$$

and the corresponding phase shift is

$$\eta(k) = \arctan (p/k) + \arctan (b/k), \tag{I9}$$

with

$$p = + |E|^{\frac{1}{2}} \tag{I10}$$

and

$$b = (p^2 + C/p)^{\frac{1}{2}}. \tag{I11}$$

We emphasize that the requirement that the corresponding potential be of Bargmann type and even implies that the phase shift is determined once the bound-state parameters are assigned, and, in fact, one easily finds the general formula

$$\eta(k) = 2 \sum_{n=1}^{N} \arctan \left(\frac{|E_n|^{\frac{1}{2}}}{k} \right) - \frac{k}{2\pi} P \int_{-\infty}^{+\infty} dq \, \frac{1}{q^2 - k^2} \ln \left[1 + \sum_{n=1}^{N} \frac{C_n |E_n|^{-\frac{1}{2}}}{q^2 - E_n} \right]. \tag{I12}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Relativistic Kinetic Equations Including Radiation Effects.* I. Vlasov Approximation

RÉMI HAKIM†
Department of Physics, Temple University, Philadelphia, Pennsylvania
AND
ANDRÉ MANGENEY
Laboratoire d'Astronomie Spatiale, Observatoire de Meudon, France‡

(Received 20 June 1967)

In a preceding series of papers, we developed relativistic statistical mechanics of electromagnetically interacting particles. Here we derive a hierarchy of manifestly covariant equations for the reduced densities. This hierarchy is valid at order one in a parameter which is closely related to radiation phenomena. At the lowest order (absence of correlation between particles), we derive a kinetic equation which is a modification of the relativistic Vlasov equation. The added term introduces the effects of the emission of collective radiation. A dimensional analysis shows that this term is important when $mc^2 \sim kT$, i.e., when the plasma is relativistic. The modified Vlasov equation is shown to be irreversible: an H theorem is proved. Several applications are studied: dispersion relations, hydrodynamical equations. We have also discussed the Einstein–Ritz controversy on the connection between retarded actions and irreversibility.

1. INTRODUCTION

In a preceding series of papers,[1,2] we develop manifestly covariant statistical mechanics. In particular, we studied electromagnetic interactions and obtained a rigorous hierarchy of equations for suitable reduced densities. This formalism presents several peculiarities. First it is completely free of the usual self-energy divergences of classical field theory; in other words, it is renormalized. Second, the formalism includes all the effects of radiation phenomena through a *finite* number of additional variables.

Accordingly, the present series of papers is mainly concerned with the study of manifestly covariant

* This work was begun while one of the authors (R. H.) was at Orsay; at Temple University it was supported by the Aerospace Research Laboratories of the Office of Aerospace Research, USAF.
† Permanent address: Laboratoire de Physique Théorique et Hautes Energies, 91 Orsay, France.
‡ Postal address: Laboratoire d'Astronomie Spatiale, Observatoire de Meudon, 92 Meudon, France.
[1] R. Hakim, J. Math. Phys. **8**, 1315 (1967).
[2] R. Hakim, J. Math. Phys. **8**, 1379 (1967).

kinetic equations including radiation effects. In particular we want to investigate the statistical effects of radiation on the particle states.

Due to the fact that in the relativistic framework there exists a new universal constant c, the velocity of light in vacuum, there also exists two dimensionless expansion parameters,[3] one more than in the Newtonian case. One of these parameters is the usual plasma parameter[4] ($e^2 n^{\frac{1}{3}} \beta$) while *another one* may be[5] ($\tau_0 \omega_p$), where ω_p is the plasma frequency and

$$\tau_0 = \tfrac{2}{3} \frac{e^2}{mc^3} \quad (\tau_0 \sim 10^{-23} \text{ for an electron})$$

(τ_0 is sometimes called the "noncausality time" [6]) which is characteristic of radiation phenomena.[2,6]

In these papers, we consider kinetic equations valid at order $\sim\tau_0$, essentially because "it is empirically well known that only the first order in τ_0 is physically significant." [6] It should be noticed that neglecting terms in τ_0^2 amounts to neglecting (for instance) double Thomson scattering. In this paper (I), we derive a simple kinetic equation valid at order $\sim(ne^2\tau_0)$, in other words, we consider a self-consistent approximation (Vlasov approximation) which takes radiation phenomena into account. Next, we apply this kinetic equation to several problems: dispersion relations, hydrodynamical equations, etc.

Paper II is devoted to a "Landau approximation," i.e., to order $(ne^4\tau_0)$, weak collisions, etc., also including radiation effects.

In Sec. 2, we recall the basic equations of relativistic statistical mechanics. In Sec. 3 we derive the generating equation of the relativistic BBGKY hierarchy which is used in Sec. 4 to obtain a kinetic equation valid at order $\sim ne^2\tau_0$. Section 5 is devoted to the derivation of the dispersion relations while Sec. 6 deals with the covariant form of the conductivity tensor. In Sec. 7 we prove an H theorem. Finally, Sec. 8 is concerned with the hydrodynamical equations.

2. SUMMARY OF RELATIVISTIC STATISTICAL MECHANICS

In this section we give the basic concepts, definitions, and notations used throughout this paper. We do not justify or discuss them here, the details having been given in Refs. 1 and 2.

A. Equations of Motion

As is well known, the relativistic dynamics of electromagnetically interacting particles may be approached in two equivalent ways. They are (a) the action-at-a-distance formalism and (b) the field one. Both approaches, which have been discussed and compared by Havas,[7] lead to the *same* equations of motion in *a rigorous manner*.[8] These equations are the Lorentz–Dirac equations,[6,9]

$$m\gamma_i^\mu = \frac{e}{c} \overset{(i)}{F}{}^{\mu\nu}(x_{\rho i}) u_{\nu i} + m\tau_0 \Delta^{\mu\nu}(u_{\rho i}) \dot{\gamma}_{\nu i}, \quad i = 1 \cdots N, \tag{2.1}$$

where we have[10]

$$u_i^\mu = \frac{dx_i^\mu}{d\tau_i}, \quad \gamma_i^\mu = \frac{du_i^\mu}{d\tau_i}, \quad \dot{\gamma}_i^\mu = \frac{d\gamma_i^\mu}{d\tau_i}, \tag{2.2}$$

$$\Delta^{\mu\nu}(u_\rho) = g^{\mu\nu} - \frac{u^\mu u^\nu}{c^2}, \tag{2.3}$$

τ_i = proper time of the ith particle. In Eq. (2.1), m is the finite observable *renormalized mass* and $\overset{(i)}{F}{}^{\mu\nu}$ is the "electromagnetic field" due to all particles *but the* ith, the "self-fields" being subtracted and leading to the last term of the right-hand side of Eq. (2.1) (and to the renormalization of mass). $\overset{(i)}{F}{}^{\mu\nu}$ is given by

$$\begin{aligned} \overset{(i)}{F}{}^{\mu\nu}(x_\rho) = \frac{4\pi e}{c} \sum_{j \neq i} \iint d_4x' \cdot d_4u' \cdot d\tau' \\ \times \{u'^\nu \partial^\mu - u'^\mu \partial^\nu\} D_{\text{ret}}(x_\rho - x'_\rho) \\ \times \delta[x'_\rho - x_{\rho j}(\tau')] \otimes \delta[u'_\rho - u_{\rho j}(\tau')], \end{aligned} \tag{2.4}$$

where $D_{\text{ret}}(x_\delta)$ is the usual retarded elementary solution of $\Box D_{\text{ret}} = \delta$. We strongly emphasize that Eqs. (2.1) are *rigorous* and hence are admissible as a starting point of relativistic statistical mechanics. Furthermore, they contain all the effects of radiation through the term involving $\dot{\gamma}^\mu$.

---

[3] Such a situation also occurs in quantum statistical mechanics where Planck's constant leads to a new dimensionless expansion parameter.

[4] n is the numerical density and $\beta = (kT)^{-1}$, T being the temperature and k the Boltzmann constant.

[5] Such a new parameter involving c is of course not unique. For instance, we could also choose $mc^2(kT)^{-1}$.

[6] F. Rohrlich, *Classical Charged Particles* (Addison-Wesley Publishing Company, Reading, Mass., 1965).

[7] P. Havas, Phys. Rev. **74**, 456, (1948).

[8] Actually, these two formalisms are equivalent only when we assume the validity of the Wheeler–Feynman "absorber" conditions. However, we shall not consider a time symmetrical formalism, but rather, retarded actions. Therefore, the two formalisms are equivalent without extra conditions (see the article by P. Havas, Ref. 7) and not in an approximated way as is sometimes stated (see, e.g., Refs. 1 and 2, and references quoted therein).

[9] In all that follows, we limit ourselves to an assembly of N electrons embedded in a positive uniform neutralizing background. The extension to more complicated systems is straightforward.

[10] Throughout this paper we consider the following pseudometric tensor:

$$g^{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu; \quad g^{00} = -g^{ii} = +1, \quad \text{for} \quad i = 1, 2, 3.$$

Moreover, the Einstein summation convention is used; Greek indices run from zero to three while Latin ones run from one to three.

B. Phase Space—Relativistic Gibbs Ensembles

Contrary to the usual case where phase space may be identified with the space of initial data, the relativistic phase space is only suggested by the form of the equations of motion (2.1) and hence is chosen for convenience.[1,2] Here, phase space will be chosen as being the following $12N$-dimensional space:

$$\Gamma = \{x_\mu\}^N \times \{u_\mu\}^N \times \{\dot{\gamma}_\mu\}^N. \tag{2.5}$$

[In fact, relativistic phase space is only $10N$-dimensional because of the $2N$ constraints,

$$u_i^\mu u_{\mu i} = c^2, \quad i = 1 \cdots N, \tag{2.6}$$

$$\gamma_i^\mu u_{\mu i} = 0, \quad i = 1 \cdots N. \tag{2.7}$$

The latter conditions, obtained by differentiation of Eq. (2.6), imply N conditions on the γ_i^μ's through Eq. (2.1). However, the constraints (2.6) and (2.7) will be incorporated in the densities so that we can actually consider Eq. (2.5) as defining phase space.]

A relativistic Gibbs ensemble is now defined as: (a) the manifold of solutions of Eq. (2.1) and (b) a probability on this manifold. In other words, what is random is the data of the *entire trajectories* of the particles of the system. We assume the existence of an averaging operation, not as usual over initial data but rather over the trajectories of the particles. This last assumption is needed because both complete relativistic invariance and the precise nature of the initial data of the nonlinear integrodifferential equations (2.1) are yet unknown.[1,11]

C. Densities

In these papers we only need the one- and two-particle densities, so that we only give the definition of these later.

We first define the *random* one-particle density:

$$R_1(x_\nu, u_\nu, \dot{\gamma}_{\nu i}\tau) = \sum_{i=1}^{i=N} \delta[x_\nu - x_{\nu i}(\tau)] \otimes \delta[u_\nu - u_{\nu i}(\tau)] \otimes \delta[\dot{\gamma}_\nu - \dot{\gamma}_{\nu i}(\tau)] \tag{2.8}$$

which is random because the $x_{\nu i}(z)$'s are random. From R_1 we define the proper time-dependent density,

$$D_1(x_\nu, u_\nu, \dot{\gamma}_\nu; \tau) = N\langle R_1(x_\nu, u_\nu, \dot{\gamma}; \tau)\rangle, \tag{2.9}$$

where the brackets represent the averaging operation. It should be realized that D_1 has no direct physical meaning and is only a useful intermediary to obtain the relativistic one-particle density:

$$\mathcal{N}_1(x_\nu, u_\nu, \dot{\gamma}_\nu) = \int_{-\infty}^{+\infty} D_1(x_\nu, u_\nu, \dot{\gamma}_\nu)\, d\tau. \tag{2.10}$$

[11] See the general discussion by P. Havas, in "Some Basic Problems in the Formulation of a Relativistic Statistical Mechanics" *Statistical Mechanics of Equilibrium and Nonequilibrium*, J. Meixner, Ed. (North-Holland Publishing Company, Amsterdam, 1965).

D_1 is normalized through

$$\int D_1(x_\nu, u_\nu, \dot{\gamma}_\nu)\, d_4x_4 \cdot d_4u \cdot d_4\dot{\gamma} = 1, \tag{2.11}$$

while the normalization of $\mathcal{N}_1$ may be found in Ref. 1. In the same way, we define the random two-particle density,

$$R_2(X, \tau; X', \tau') = \sum_{i \neq j} \delta[X - X_i(\tau)] \otimes \delta[X' - X_j(\tau')] \tag{2.12}$$

with $X \equiv (x_\nu, u_\nu, \dot{\gamma}_\nu)$, and D_2 by

$$D_2(X, \tau; X', \tau') = N(N-1)\langle R_2(X, \tau; X', \tau')\rangle; \tag{2.13}$$

D_2 is also normalized to unity (in what follows we assume that $N \gg 1$).

Since we want to obtain a kinetic equation for the particles, we only need the "reduced" $\mathcal{N}_1(x_\nu, u_\nu)$ obtained from $\mathcal{N}_1(x_\nu, u_\nu, \dot{\gamma}_\nu)$ by an integration over $\dot{\gamma}$. $\mathcal{N}_1(x_\nu, u_\nu)$ is normalized through

$$\iint_\Sigma d_4u\, d\Sigma_\mu u^\mu \mathcal{N}_1 = 1, \quad \text{for every spacelike three-surface } \Sigma, \tag{2.14}$$

where $d\Sigma_\mu$ is the differential form "surface element."

It must be noted that the constraints (2.6) and (2.7) are actually included in the densities because of the occurrence of $x_{\nu i}(\tau)$, $u_{\nu i}(\tau)$, $\dot{\gamma}_{\nu i}(\tau)$ in their definitions.

D. The Relativistic Hierarchy

We now use the elegant methods of Klimontovich. In order to obtain a relativistic hierarchy, we first derive a generating equation, i.e., an equation satisfied by *random densities*. Starting from the continuity equation,[12]

$$\frac{\partial R_1}{\partial \tau} + u^\mu \partial_\mu R_1 + \frac{\partial}{\partial u^\mu}\{\gamma^\mu R_1\} + \frac{\partial}{\partial \dot{\gamma}^\mu}\{\gamma^\mu R_1\} = 0, \tag{2.15}$$

the generating equation is obtained after an integration over $\dot{\gamma}$ and taking Eq. (2.1) into account,

$$\frac{\partial}{\partial \tau} R_1(x_\nu, u_\nu; \tau) + u^\mu \partial_\mu R_1(x_\nu, u_\nu; \tau) + \frac{4\pi e^2}{mc^2} u_\nu \frac{\partial}{\partial u_\mu} \int d\tau' \cdot d_4x' \cdot d_4u' \times \{u'^\nu \partial^\mu - u'^\mu \partial^\nu\} D_{\text{ret}}(x_\rho - x'_\rho) R_2(X, \tau; X', \tau') = -\tau_0 \frac{\partial}{\partial u^\mu}\left\{\Delta^{\mu\nu}(u_\rho) \int d_4\dot{\gamma} \cdot \dot{\gamma}_\nu R_1(x_\nu, u_\nu, \dot{\gamma}_\nu; \tau)\right\}. \tag{2.16}$$

[12] Note that the independent variables are $(\tau, x_\nu, u_\nu, \dot{\gamma}_\nu)$. This equation has only a formal meaning. One must explicitly perform on R_1 the various derivations required.

The right-hand side of Eq. (2.16) is proportional to τ_0 and hence represents the effects of radiation on the motion of particles. This term couples the ordinary random densities for the particles to the generalized density including radiation effects. The left-hand side represents the generating renormalized Klimontovich equation.[1,2,13] By taking average values of products of $R_1^{\otimes k}$ $(k = 1 \cdots)$ by both sides of Eq. (2.16), a relativistic hierarchy is obtained.[2]

E. Remark

The equations of motion (2.1) actually do have some unphysical solutions, i.e., the so-called "runaway solutions," which have to be eliminated with asymptotic conditions. We could take these asymptotic conditions into account in this formalism. However, we are dealing with kinetic equations valid at order one in τ_0 and hence *we assume* that the distribution functions are analytic in τ_0. This analyticity assumption is sufficient to prevent the occurrence of such "runaway solutions" as shown by Lichnérowicz.[14]

In fact, this analyticity of the solution of the equation of motion eliminates the runaway solutions for a one-particle system embedded in an external force field. For interacting particles, such a property has never been proved so that we have to *assume* its validity. This assumption seems to be very plausible.

3. RELATIVISTIC HIERARCHY AT ORDER ONE IN τ_0

Before deriving the generating equation at "order one in radiation effects," we first discuss the order of magnitude of the different terms of Eq. (2.16) using dimensional arguments.

A. Dimensional Arguments

We rewrite Eq. (2.16) with dimensionless quantities. To this end we use the following substitutions:

$$\tau = \tau_1 s, \tag{3.1}$$

$$x_\mu = l_0 \hat{x}_\mu, \tag{3.2}$$

$$u_\mu = l_0 \tau_1^{-1} \hat{u}_\mu \equiv \epsilon \hat{u}_\mu, \tag{3.3}$$

$$R_1(x_\nu, u_\nu; \tau) = \epsilon^{-4} l_0^{-4} \hat{R}_1(\hat{x}_\nu, \hat{u}_\nu, s), \tag{3.4}$$

where l_0 and τ_1 are, respectively, some characteristic length and time of the electron plasma. It is important to note that substitution (3.3) *does not* preserve the pseudolength of the four-vector u^μ:

$$u^\mu u_\mu = c^2, \quad \hat{u}^\mu \hat{u}_\mu = \epsilon^{-2} c^2. \tag{3.5}$$

This means that changing *independently* the scales of length and of time amounts to changing the scale of velocity and hence the *numerical value* of the speed of light,

$$\hat{c} = c\epsilon^{-1}. \tag{3.6}$$

It may be readily verified that substitution (3.2) preserves the light cone, as is well known.

Now using substitutions (3.1)–(3.4), Eq. (2.16) may be rewritten as

$$\begin{aligned} &\frac{\partial \hat{R}_1}{\partial s}(\hat{x}_\nu, \hat{u}_\nu; s) + \hat{u}^\mu \hat{\partial}_\mu \hat{R}_1(\hat{x}_\nu, \hat{u}_\nu; s) + \frac{4\pi e^2}{mc} \tau_1 l_0^{-2} \\ &\times u_\nu \frac{\partial}{\partial u^\mu} \int ds' \cdot d_4\hat{x}' \cdot d_4\hat{u}' \{\hat{u}'^\nu \hat{\partial}^\mu - \hat{u}'^\mu \hat{\partial}^\nu\} D_{\text{ret}}(\hat{x}_\rho - \hat{x}'_\rho) \\ &\times \hat{R}_2(\hat{x}_\rho, \hat{u}_\rho, s; \hat{x}'_\rho, \hat{u}'_\rho, s') \\ &= -\tau_0 \tau_1^{-1} \frac{\partial}{\partial \hat{u}^\mu} \left\{ \Delta^{\mu\nu}(\hat{u}_\rho) \int d_4\hat{\gamma} \cdot \hat{\dot{\gamma}}_\nu \hat{R}_1(\hat{x}_\nu, \hat{u}_\nu, \hat{\dot{\gamma}}_\nu; s) \right\}, \end{aligned} \tag{3.7}$$

where we have set $\hat{\partial}_\mu = (\partial/\partial \hat{x}^\mu)$.

The last term of the left-hand side of Eq. (3.7) (i.e., the interaction term) will be of the order of magnitude of the right-hand side of Eq. (3.7) when the following condition is satisfied:

$$l_0 \tau_1^{-1} \sim c. \tag{3.8}$$

Now choosing the Debye length for l_0,

$$l_0 = e^{-1} n^{-\frac{1}{2}} \beta^{-\frac{1}{2}}, \tag{3.9}$$

and the plasma frequency for τ_1^{-1},

$$\tau_1^{-1} \sim \omega_p \sim e n^{\frac{1}{2}} m^{-\frac{1}{2}}, \tag{3.10}$$

condition (3.8) reduces to

$$kT \sim mc^2. \tag{3.11}$$

This condition signifies that the electron plasma should be relativistic. In other words, radiation effects are important for relativistic plasmas. This result is expected after the analysis of radiation by Schwinger.[15] Of course, dimensional arguments are not proofs. They simply give an approximate idea.

B. Generating Equation at Order One in τ_0

At order one in τ_0, the equations of motion (2.7) are to be replaced by the following:

$$m\gamma_i^\mu = \frac{e}{c} \overset{(i)}{F}{}^{\mu\nu} u_{\nu i} + m\tau_0 \Delta^{\mu\nu}(u_\rho) \dot{\gamma}_{\nu i}^{[0]}, \tag{3.12}$$

[13] In fact, the Klimontovich generating equation {Yu. L. Klimontovitch, Zh. Eksp. Teor. Fiz. 37, 735 (1959) [Sov. Phys.—JETP 10, 524 (1960)]} is not renormalized and instead of R_2, involves a term in $R_1 \otimes R_1$, which actually contains self-energy divergences (see Ref. 2).

[14] A. Lichnérowicz, private communication.

[15] J. Schwinger, Phys. Rev. 75, 1912 (1949).

where $\dot{\gamma}_{vi}^{[0]}$ is nothing but $\dot{\gamma}_{vi}$ calculated at order zero in τ_0, i.e.,

$$\dot{\gamma}_{vi}^{[0]} = \frac{d}{d\tau}\gamma_{vi}^{[0]} = \frac{d}{d\tau_i}\left\{\frac{e}{mc}\overset{(i)}{F}{}^{\mu\nu}u_{vi}\right\}. \tag{3.13}$$

In the same way, Eq. (2.16) should be approximated by taking the order zero in τ_0 of its right-hand side. Therefore, we have

$$\begin{aligned}&\frac{\partial}{\partial\tau}R_1(x_\nu, u_\nu;\tau) + u^\mu\partial_\mu R_1(x_\nu, u_\nu;\tau)\\ &\quad + \frac{4\pi e^2}{mc^2}u_\nu\frac{\partial}{\partial u^\mu}\left\{\int d\tau'\cdot d_4x'\cdot d_4u'\{u'^\nu\partial^\mu - u'^\mu\partial^\nu\}\right.\\ &\quad\left.\times D_{\text{ret}}(x_\rho - x'_\rho)R_2(x_\rho, u_\rho,\tau; x'_\rho, u'_\rho,\tau')\right\}\\ &\quad = -\tau_0\frac{\partial}{\partial u^\mu}\left\{\Delta^{\mu\nu}(u_\rho)\int d_4\dot{\gamma}\dot{\gamma}_\nu^{[0]}R_1^{[0]}(x_\nu, u_\nu, \dot{\gamma}_\nu;\tau)\right\}.\end{aligned} \tag{3.14}$$

In Eq. (3.14) $R_1^{[0]}$ represents the zeroth-order approximation of R_1, i.e.,

$$\begin{aligned}R_1^{[0]}(x_\nu, u_\nu, \dot{\gamma}_\nu,\tau) &= \sum_{i=1}^{i=N}\delta[x_\nu - x_{vi}^{[0]}(\tau)]\\ &\quad\otimes\delta[u_\nu - u_{vi}^{[0]}(\tau)]\otimes\delta[\dot{\gamma}_\nu - \gamma_{vi}^{[0]}(\tau)].\end{aligned} \tag{3.15}$$

It should be realized that Eq. (3.13) implies that $\dot{\gamma}_{vi}^{[0]}$ is now a function of $u_{vi}^{[0]}$ and $\gamma_{vi}^{[0]}$. From Eq. (3.13) we easily find

$$\begin{aligned}\dot{\gamma}_{vi}^{[0]} &= \frac{e}{mc}\left\{\overset{(i)}{F}{}_{\nu\alpha}\gamma_i^\alpha + u_i^\rho u_i^\alpha\partial_{\rho i}\overset{(i)}{F}{}_{\nu\alpha}\right\}^{[0]}\\ &= \frac{e}{mc}\left\{\frac{e}{mc}\overset{(i)}{F}{}_{\nu\alpha}\cdot\overset{(i)}{F}{}^{\alpha\beta}u_{\beta i} + u_i^\rho u_i^\alpha\partial_{\rho i}\overset{(i)}{F}{}_{\nu\alpha}\right\}.\end{aligned} \tag{3.16}$$

Using now expression (3.15) in Eq. (3.14), taking into account the various δ factors and Eq. (3.16), using the δ factors and integrating finally over $\dot{\gamma}$, we find

$$\begin{aligned}&\frac{\partial}{\partial\tau}R_1(x_\nu, u_\nu;\tau) + u^\mu\partial_\mu R_1(x_\nu, u_\nu;\tau)\\ &\quad + \frac{4\pi e^2}{mc^2}u_\nu\frac{\partial}{\partial u^\mu}\left\{\int d\tau'\cdot d_4x'\cdot d_4u'\{u'^\nu\partial^\mu - u'^\mu\partial^\nu\}\right.\\ &\quad\left.\times D_{\text{ret}}(x_\rho - x'_\rho)\cdot R_2(x_\nu, u_\nu,\tau; x'_\nu, u'_\nu,\tau')\right\}\\ &= -\tau_0\frac{\partial}{\partial u^\mu}\left\{\Delta^{\mu\nu}(u_\rho)\times\left[\left[\frac{4\pi e^2}{mc^2}u^\rho u^\alpha\int d\tau'\cdot d_4x'\cdot d_4u'\right.\right.\right.\\ &\quad\times\{u'_\alpha\partial_\nu - u'_\nu\partial_\alpha\}D_{\text{ret}}(x_\rho - x'_\rho)\\ &\quad\left.\times R_2^{[0]}(x_\nu, u_\nu,\tau; x'_\nu, u'_\nu,\tau')\right]\\ &\quad + \left[\left(\frac{4\pi e^2}{mc^2}\right)^2 u_\alpha\int\int d\tau'\cdot d_4x'\cdot d_4u'\cdot d\tau''\cdot d_4x''\cdot d_4u''\right.\\ &\quad\times\{u'_\nu\partial_\beta - u'_\beta\partial_\nu\}D_{\text{ret}}(x_\nu - x'_\nu)\\ &\quad\times\{u''^\beta\partial^\alpha - u''^\alpha\partial^\beta\}D_{\text{ret}}(x_\nu - x'_\nu)\\ &\quad\times[\![R_3^{[0]}(x_\nu, u_\nu,\tau; x'_\nu, u'_\nu,\tau; x''_\nu, u''_\nu,\tau'')\\ &\quad\left.\left.\left.+ W_{3\,\text{stoch}}^{2[0]}(x'_\nu, u'_\nu,\tau'; x''_\nu, u''_\nu,\tau''; \{x_\nu, u_\nu,\tau\})]\!]\right]\right]\right\},\end{aligned} \tag{3.17}$$

where we have designated by $W_{3\,\text{stoch}}^2(X', X'', \{X\})$ the random density

$$\begin{aligned}W_{3\,\text{stoch}}^2(X', X'', \{X\}) &= \sum_{i\neq j}\delta[X' - X_i(\tau')]\\ &\quad\otimes\delta[X'' - X_i(\tau'')]\otimes\delta[X - X_j(\tau)]\end{aligned} \tag{3.18}$$

which has already been discussed.[1,2]

Equation (3.17) is the generating equation for which we were looking. The hierarchy at order one in τ_0 is now obtained in a straightforward way by taking averages of products of R_1 by Eq. (3.17).

In view of the applications considered in this paper, we need only the first equation of the hierarchy. This equation is of the same form as Eq. (3.17) with the following replacements:

$$R_3 \to N^2D_3, \quad \text{and} \quad W_{3\,\text{stoch}}^2 \to NW_3^2.$$

Note also that the diagrams established in Ref. 2 may also be used to derive higher-order equations[16]; this requires only slight modifications.

4. A KINETIC EQUATION AT ORDER $(ne^2\tau_0)$

At order (ne^2), the two particle density $D_2(x_\nu, u_\nu,\tau; x'_\nu, u'_\nu,\tau')$ is usually assumed to factorize:

$$\begin{aligned}D_2(x_\nu, u_\nu,\tau; x'_\nu, u'_\nu,\tau') &= D_1(x_\nu, u_\nu;\tau)\\ &\quad\otimes D_1(x'_\nu, u'_\nu;\tau')\end{aligned} \tag{4.1}$$

(or equivalently, the two-particle correlation function vanishes). We adopt this common assumption here. Therefore, starting with the first equation of the relativistic hierarchy at order one in τ_0, derived from Eq. (3.17), taking into account the order considered (i.e., $ne^2\tau_0$), the following kinetic equation is obtained:

$$\begin{aligned}&\frac{\partial}{\partial\tau}D_1(x_\nu, u_\nu;\tau) + u^\mu\partial_\mu D_1(x_\nu, u_\nu;\tau)\\ &\quad + \frac{4\pi e^2N}{mc^2}u_\nu\frac{\partial}{\partial u^\mu}\left\{\int d\tau'\cdot d_4x'\cdot d_4u'[u'^\nu\partial^\mu - u'^\mu\partial^\nu]\right.\\ &\quad\left.\times D_{\text{ret}}(x_\rho - x'_\rho)\times D_1(x'_\nu, u'_\nu;\tau')\otimes D_1(x_\nu, u_\nu;\tau)\right\}\\ &= -\tau_0\frac{\partial}{\partial u^\mu}\left\{\Delta^{\mu\nu}(u_\rho)\left(\frac{4\pi e^2N}{mc^2}\right)u^\rho u^\alpha\int d\tau'\cdot d_4x'\cdot d_4u'\right.\\ &\quad\times[u'_\alpha\partial_\nu - u'_\nu\partial_\alpha]\partial_\rho D_{\text{ret}}(x_\rho - x'_\rho)D_1(x'_\nu, u'_\nu;\tau')\\ &\quad\left.\otimes D_1(x_\nu, u_\nu;\tau)\right\};\end{aligned} \tag{4.2}$$

[16] It may be of interest to note that, at order one in τ_0, two dotted lines are needed to generalize our previous diagrams (see Ref. 2).

it may be readily verified that the radiation term [i.e., the right-hand side of Eq. (4.2) actually vanishes when the system under study is *homogeneous*, as expected.] At this point it may be worth emphasizing that the right-hand side of Eq. (4.2) represents the action of the (coherent) collective radiation field of the system on the motion of the particles. Note also that in the right-hand side of Eq. (4.2), we have replaced $D_1^{[0]}$ by D_1, thereby adding negligibly small terms of order $O(\tau_0^2)$. The left-hand side of Eq. (4.2) is the usual covariant form of the relativistic Vlasov equation.

Henceforth, we use the more "physical" [1] density $\mathcal{N}_1(x_\nu, u_\nu)$ rather than $D_1(x_\nu, u_\nu; \tau)$. Integrating Eq. (4.2) over τ and taking into account the property,[1]

$$\lim_{\tau \to \pm\infty} D_1(x_\nu, u_\nu; \tau) = 0,$$

we find[17]

$$u^\mu \partial_\mu \mathcal{N} + \left(\frac{4\pi e^2 N}{m}\right) u_\nu \int d_4x' \cdot d_4u' [u'^\nu \partial^\mu - u'^\mu \partial^\nu] \times D_{\text{ret}}(x_\rho - x'_\rho) \mathcal{N}(x'_\rho, u'_\rho) \frac{\partial}{\partial u^\mu} \mathcal{N}(x_\rho, u_\rho)$$
$$= -\frac{4\pi e^2 N \tau_0}{m} \cdot \frac{\partial}{\partial u^\mu} \left\{ \Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \int d_4x' \cdot d_4u' \times [u'_\alpha \partial_\nu - u'_\nu \partial_\alpha] \partial_\rho D_{\text{ret}}(x_\rho - x'_\rho) \times \mathcal{N}(x'_\rho, u'_\rho) \otimes \mathcal{N}(x_\rho, u_\rho) \right\}. \quad (4.3)$$

This equation will be referred to as the "modified Vlasov equation."

Linearization of the Modified Vlasov Equation

We now assume that the plasma under consideration is in an homogeneous equilibrium state described by the Jüttner–Synge distribution function[18]:

$$\mathcal{N}_{\text{eq}}(u_\rho) = \frac{n_0 m \xi}{4\pi K_2(m\xi)} \exp\{-m\xi^\mu u_\mu\} \cdot 2\theta(u^0)\delta(u^\mu u_\mu - 1), \quad (4.4)$$

i.e., Eq. (4.4) represents an equilibrium distribution function *at order* 0 *in* e^2. In expression (4.4), n_0 is the invariant (constant) world density of the electron gas, K_2 is a Kelvin function of order 2, and

$$\xi^\mu = \xi \bar{u}^\mu; \quad \xi = (kT)^{-1}; \quad \bar{u}^\mu \cdot \bar{u}_\mu = 1, \quad (4.5)$$

$\bar{u}^\mu$ being the average four-velocity. Now we study the motion of a "small disturbance" within this plasma. To this end, we linearize the modified Vlasov equation by setting

$$\mathcal{N}(x_\nu, u_\nu) = \mathcal{N}_{\text{eq}}(u_\nu) + \mathcal{N}_{(1)}(x_\nu, u_\nu) \quad (4.6)$$

and assuming that the part of the distribution function which describes the disturbance is "small," i.e.,

$$\{\mathcal{N}_{(1)}\}^2 \ll \mathcal{N}_{(1)}. \quad (4.7)$$

Substituting Eq. (4.6) into Eq. (4.3) and taking condition (4.7) into account, we get[19]

$$u^\mu \partial_\mu \mathcal{N}_{(1)} + \frac{4\pi e^2}{m} u_\nu \frac{\partial}{\partial u^\mu} \mathcal{N}_{\text{eq}}(u_\rho) \times \int d_4x' \cdot d_4u' \{u'^\nu \partial^\mu - u'^\mu \partial^\nu\} \times D_{\text{ret}}(x_\rho - x'_\rho) \mathcal{N}_{(1)}(x'_\rho, u'_\rho)$$
$$= -\left(\frac{4\pi e^2 \tau_0}{m}\right) \frac{\partial}{\partial u^\mu} \left\{ \Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \int d_4x' \cdot d_4u' \times \{u'_\alpha \partial_\nu - u'_\nu \partial_\alpha\} \partial_\rho D_{\text{ret}}(x_\rho - x'_\rho) \times \mathcal{N}_{\text{eq}}(u_\rho) \mathcal{N}_{(1)}(x'_\rho, u'_\rho) \right\}. \quad (4.8)$$

It is important to realize that $\mathcal{N}_{\text{eq}}$ could as well have been chosen to be any other four-velocity-dependent distribution and not only the one given in Eq. (4.4).

5. DERIVATION OF DISPERSION RELATIONS

Let us first rewrite the above linearized modified Vlasov equation in terms of the Fourier transform $\hat{\mathcal{N}}_{(1)}$ of $\mathcal{N}_{(1)}$:

$$\mathcal{N}_{(1)}(x_\nu, u_\nu) = \frac{1}{(2\pi)^2} \int d_4k \exp[ik^\mu x_\mu] \hat{\mathcal{N}}_{(1)}(k_\nu, u_\nu),$$
$$\hat{\mathcal{N}}_{(1)}(k_\nu, u_\nu) = \frac{1}{(2\pi)^2} \int d_4x \exp[-ik^\mu x_\mu] \mathcal{N}_{(1)}(x_\nu, u_\nu). \quad (5.1)$$

From Eq. (5.1) and Eq. (4.8) we immediately find

$$ik^\mu u_\mu \hat{\mathcal{N}}_{(1)} + i\left(\frac{4\pi e^2}{m}\right) u_\nu \frac{\partial}{\partial u^\mu} \mathcal{N}_{\text{eq}}(u_\rho) \times \int d_4u \frac{k^\mu u'^\nu - k^\nu u'^\mu}{-k^\lambda k_\lambda} \hat{\mathcal{N}}_{(1)}(k_\rho, u'_\rho)$$
$$= \left(\frac{4\pi e^2 \tau_0}{m}\right) \frac{\partial}{\partial u^\mu} \{\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \mathcal{N}_{\text{eq}}(u_\rho)\} \times \int d_4u' \frac{u'_\alpha k_\nu - u'_\nu k_\alpha}{-k^\lambda k_\lambda} k_\rho \hat{\mathcal{N}}_{(1)}(k_\rho, u'_\rho), \quad (5.2)$$

[17] In all that follows, we suppress the index 1 of $\mathcal{N}_1$ and take $c = 1$.

[18] J. L. Synge, *The Relativistic Gas* (North-Holland Publishing Company, Amsterdam, 1957).

[19] Note that, as usual, we have used the assumption of a neutralizing positive background.

from which we get at once

$$\hat{\mathcal{N}}_{(1)}(k_\nu, u_\nu) = \frac{\omega_p^2}{n_0} u_\nu \frac{\partial}{\partial u^\mu} \mathcal{N}_{eq}(u_\rho) \times \frac{1}{k^\lambda u_\lambda} \int d_4u' \frac{k^\mu u'^\nu - k^\nu u'^\mu}{k^\lambda k_\lambda} \hat{\mathcal{N}}_{(1)}(k_\rho, u_\rho) + i\left(\frac{\tau_0 \omega_p^2}{n_0}\right) \frac{\partial}{\partial u^\mu} \{\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \mathcal{N}_{eq}(u_\rho)\} \times \frac{1}{k^\lambda u_\lambda} \int d_4u' \frac{u'_\alpha k_\nu - u'_\nu k_\alpha}{k^\lambda k_\lambda} k_\rho \hat{\mathcal{N}}_{(1)}(k_\rho, u'_\rho), \quad (5.3)$$

where ω_p is the plasma frequency. In Eq. (5.3) it is *intended* that

$$\frac{1}{k^\lambda u_\lambda} \equiv i\pi\delta^-(k^\lambda u_\lambda) = i\pi \left\{\delta(k^\lambda u_\lambda) - \frac{i}{\pi} P \frac{1}{k^\lambda u_\lambda}\right\}, \quad (5.4)$$

where P denotes the Cauchy principal value.

The Fourier transform of the four-current due to the "small disturbance" $\mathcal{N}_{(1)}$, is

$$J^\beta(k_\nu) = e \int d_4u\, u^\beta \hat{\mathcal{N}}_{(1)}(k_\nu, u_\nu). \quad (5.5)$$

Now multiplying Eq. (5.3) by eu^β and integrating over four-velocity u^μ, we get the homogeneous linear system

$$J^\beta(k_\nu) = \frac{\omega_p^2}{k^\alpha k_\alpha} \{-I_\lambda k^\beta - k^\alpha k_\alpha I^\beta_\lambda + k^\alpha I_\alpha g^\beta_\lambda + i\tau_0(k^\alpha k_\alpha J^\beta_\lambda + \xi^\alpha k_\alpha \xi^{-1} g^\beta_\lambda - k^\beta \xi_\lambda \xi^{-1})\} \cdot J^\lambda(k_\nu). \quad (5.6)$$

In Eq. (5.6) we have used the following notations:

$$I_\nu(k_\rho, \xi_\rho) = n_0^{-1} \int d_4u \frac{u_\nu}{k^\alpha u_\alpha} \mathcal{N}_{eq}(u_\rho), \quad (5.7)$$

$$I^\nu_\mu(k_\rho, \xi_\rho) = -n_0^{-1} \int d_4u \frac{u_\mu u^\nu}{(k^\alpha u_\alpha)^2} \mathcal{N}_{eq}(u_\rho), \quad (5.8)$$

$$J^\nu_\mu(k_\rho, \xi_\rho) = n_0^{-1} \int d_4u \frac{u_\mu u^\nu}{k^\alpha u_\alpha} \mathcal{N}_{eq}(u_\rho). \quad (5.9)$$

These integrals are studied in the Appendix.

The dispersion relation is now obtained by setting the determinant of the linear system (5.6) equal to zero:

$$\det \left\{ g^\beta_\lambda \left(1 - \omega_p^2 \frac{k^\alpha I_\alpha + i\tau_0 \xi^\alpha k_\alpha \xi^{-1}}{k^\alpha k_\alpha}\right) + \omega_p^2 (I^\beta_\lambda - i\tau_0 J^\beta_\lambda) + \omega_p^2 k^\beta \frac{I_\lambda + i\tau_0 \xi_\lambda \xi^{-1}}{k^\alpha k_\alpha} \right\} = 0. \quad (5.10)$$

Finally, Eq. (5.10) leads to the following dispersion relations [in a frame of reference where ξ_μ reduces to $(\xi, 0)$ and choosing the third axis as the direction of propagation, i.e., k_μ reduces to

$$(k^0, 0, 0, k^3) \equiv (\omega, 0, 0, k)]: \quad (5.11)$$

$$\left\{1 - \frac{\omega_p^2}{k^\alpha k_\alpha} G\right\} + \omega_p^2 \{I^1_1 + i\tau_0 J^1_1\} = 0 \quad (5.12)$$

(for transverse modes),

$$\left\{\left[1 - \frac{\omega_p^2}{k^\alpha k_\alpha} G\right] + \omega_p^2 (I^0_0 - i\tau_0 J^0_0) + \frac{\omega_p^2 k^0}{k^\alpha k_\alpha} (I_0 + i\tau_0)\right\} \times \left\{\left[1 - \frac{\omega_p^2}{k^\alpha k_\alpha} G\right] + \omega_p^2 (I^3_3 - i\tau_0 J^3_3) + \frac{\omega_p^2 k^3}{k^\alpha k_\alpha} (I_3)\right\} = \omega_p^4 \left\{\frac{k^3}{k^\alpha k_\alpha} (I_0 + i\tau_0) + I^3_0 - i\tau_0 J^3_0\right\} \times \left\{I^0_3 - i\tau_0 J^0_3 + \frac{k^0 I_3}{k^\alpha k_\alpha}\right\} \quad (5.13)$$

(for longitudinal modes).

In Eqs. (5.12) and (5.13) we have set

$$G = k^\nu I_\nu + i\tau_0 k^\nu \xi_\nu \cdot \xi^{-1}. \quad (5.14)$$

The method used above to derive the dispersion relations has already been employed by Kurşunoğlu[20] in the case where radiation emission is neglected. However, the dispersion relations obtained by this author are incorrect, due to the fact that his $\mathcal{N}_{eq}$ *does not transform* as a scalar but rather as the zeroth component of a four-vector.[21]

In the dispersion relations (5.12) and (5.13), the radiation terms appear in two different ways. First, they appear in front of integrals J^ν_μ; these may be important under certain circumstances. Second, they appear in G (for instance); these are negligibly small since in order to be appreciable they would require frequencies of the order of 10^{23} cps! Therefore, we drop these terms.

Note also that all the various integrals involved in in Eqs. (5.12)–(5.14) can be expressed in terms of

$$I_0, \quad \frac{\partial I_0}{\partial \omega} \quad \text{and} \quad \frac{\partial}{\partial(m\xi)} I_0$$

(see Appendix). After expressing the dispersion relations for the longitudinal modes in terms of I_0 and its derivative with respect to ω, neglecting the radiation terms and also terms in ω_p^4, we recover *formally* the formula given by Kurşunoğlu.[20,21]

[20] B. Kurşunoğlu, Nucl. Fusion **1**, 213 (1961).

[21] In particular, the property $I^\nu k_\nu = 1$ is no longer true.

A. The Dispersion Relation at Zero Temperature

At zero temperature, the Jüttner–Synge distribution function becomes

$$\mathcal{N}_{eq}(u_\rho) = n_0\delta(u_\rho - \bar{u}_\rho)2\theta(\bar{u}^0)\delta(\bar{u}^\rho\bar{u}_\rho - 1). \quad (5.15)$$

In what follows we drop the second δ factor and only write

$$\mathcal{N}_{eq}(u_\rho) = n_0\delta(\bar{u}_\rho - u_\rho), \quad (5.16)$$

and always keep in mind that $\bar{u}^\rho\bar{u}_\rho = 1$, $\bar{u}^0 > 0$.

At zero temperature the various integrals involved in the dispersion relations are easily calculated and hence exact results may be obtained.

Putting Eq. (5.16) into the definitions (5.7), (5.8), and (5.9), we get immediately

$$I_\nu = \frac{\bar{u}_\nu}{k^\lambda\bar{u}_\lambda}, \quad (5.17)$$

$$I_\mu^\nu = -\frac{\bar{u}_\mu\bar{u}^\nu}{(k^\lambda\bar{u}_\lambda)^2}, \quad (5.18)$$

$$J_\mu^\nu = \frac{\bar{u}_\mu\bar{u}^\nu}{k^\lambda\bar{u}_\lambda}. \quad (5.19)$$

The dispersion relation for transverse waves now reads

$$k^\alpha k_\alpha = \omega_p^2, \quad (5.20)$$

or

$$\omega^2 = k^2 + \omega_p^2, \quad (5.21)$$

which is the usual expected result (in the reference frame already considered and for wave propagation along the third axis).

For the longitudinal oscillations, we obtain

$$\begin{aligned}\omega_1^2 &= k^2 + \omega_p^2,\\ \omega_2^2 &= \omega_p^2.\end{aligned} \quad (5.22)$$

The second solution ω_2 is the usual expected non-propagating longitudinal wave at zero temperature. The first solution seems to be typically relativistic since it disappears at the Newtonian limit, as can easily be verified by taking $c = \infty$ in the longitudinal dispersion relation[22,23]:

$$\omega^4 - \omega^2(k^2c^2 + \omega_p^2) + \omega_p^2c^2k^2 = 0. \quad (5.23)$$

In Eqs. (5.22) and (5.23) we have neglected the radiation terms. Had we considered these terms, we should have obtained

$$\omega^4 - \omega^2(k^2c^2 + \omega_p^2) + i\tau_0\omega_p^2\omega c^2k^2 + \omega_p^2c^2k^2 = 0 \quad (5.24)$$

for the longitudinal dispersion relation instead of Eq. (5.23). In fact, the radiation terms are probably irrelevant since our previous dimensional analysis seems to indicate that they are important only at high temperatures.

B. The Dispersion Relations at Infinite Temperatures

Using the expressions at the extreme relativistic limit of the various integrals involved in the dispersion relations (5.12) and (5.13), it can be seen that they reduce to

$$k^\mu k_\mu = 0. \quad (5.25)$$

Therefore, it seems that the extreme relativistic plasma behaves like the vacuum for both longitudinal and transverse electromagnetic waves. However, the situation is not so simple. Indeed, when Eq. (5.25) is verified $\omega = k$, and hence the various integrals J_μ^ν diverge (see Appendix). This means that electromagnetic waves are so strongly damped that they cannot propagate within an extreme relativistic plasma. This feature was of course expected since radiation emission becomes more and more effective in damping waves as the temperature increases. This result substantially disagrees with the one obtained by Kurşunoğlu. This is not surprising because of the reasons already mentioned.

Finally, we want to emphasize that the extreme relativistic approximation of our equations is meaningless, especially because, at super high temperatures our approximation is itself meaningless; we expect that order τ_0^2, etc., should be significant.

6. THE CONDUCTIVITY TENSOR

We now look for a tensor $\sigma_{\mu\nu}^\lambda$ such that

$$\hat{J}^\lambda = \sigma_{\mu\nu}^\lambda\hat{F}^{\mu\nu}, \quad (6.1)$$

where $\hat{F}^{\mu\nu}$ is the Fourier transform of the electromagnetic field generated by the disturbance. We have

$$\hat{F}^{\mu\nu}(k_\rho) = -4\pi i\frac{k^\mu\hat{J}^\nu - k^\nu\hat{J}^\mu}{k^\alpha k_\alpha}. \quad (6.2)$$

Using Eq. (5.3) written in the form

$$\begin{aligned}&\hat{\mathcal{N}}_{(1)}(k_\rho, u_\rho)\\ &= \left\{\frac{1}{k^\alpha u_\alpha}\left[\frac{\omega_p^2}{4\pi ien_0}u_\nu\frac{\partial}{\partial u^\mu}\mathcal{N}_{eq}(u_\rho)\right.\right.\\ &\quad\left.\left.+\frac{\tau_0\omega_p^2}{4\pi n_0 e}\cdot\frac{\partial}{\partial u^\lambda}\{\Delta_\nu^\lambda(u_\rho)k^\alpha u_\alpha u_\mu\mathcal{N}_{eq}(u_\rho)\}\right]\right\}\cdot\hat{F}^{\mu\nu}(k_\rho),\end{aligned} \quad (6.3)$$

[22] When dealing with electromagnetic phenomena, it is not easy to consider a Newtonian limit. Indeed, to obtain the Newtonian limit of Maxwell's equations amounts to suppressing the displacement current term and hence, we no longer have electromagnetic waves! Furthermore, if we conserve the displacement current, the resulting nonrelativistic theory is neither invariant under Galilei transformations nor under Lorentz transformations and we get a theory which is not consistent.

[23] We have re-established the factors c in Eqs. (5.23)–(5.24).

and multiplying both sides of this equation by eu^λ, integrating over the four velocity, and using definition (6.1), we get

$$\sigma^\lambda_{\mu\nu}(k_\rho) = \frac{\omega_p^2}{8\pi i}\{(g^\lambda_\nu I_\mu - g^\lambda_\mu I_\nu) + (k_\nu I^\lambda_\mu - k_\mu I^\lambda_\nu)\} + \frac{\tau_0\omega_p^2}{8\pi}\{\xi^{-1}(g^\lambda_\mu\xi_\nu - g^\lambda_\nu\xi_\mu) + (k_\nu J^\lambda_\mu - k_\mu J^\lambda_\nu)\}. \tag{6.4}$$

When we neglect the radiation terms in τ_0 in Eq. (6.4), we recover the formula (up to unessential numerical factors) already given by Kurşunoğlu[20]; as we have pointed out in preceding sections, the difference lies in the different value of the various integrals involved therein. Note also that in Eq. (6.4), the first term in τ_0 is in general negligible and hence will be dropped.

We now calculate the conductivity tensor at zero temperature. Using Eqs. (5.16) and (5.18) we get

$$\sigma^\lambda_{\mu\nu}(k_\rho) = \frac{\omega_p^2}{8\pi i}\left\{\frac{(g^\lambda_\mu\bar{u}_\nu - g^\lambda_\nu\bar{u}_\mu)}{k^\alpha\bar{u}_\alpha} - \frac{(k_\nu\bar{u}_\mu\bar{u}^\lambda - k_\mu\bar{u}_\nu\bar{u}^\lambda)}{(k^\alpha\bar{u}_\alpha)^2}\right\} + \frac{\tau_0\omega_p^2}{8\pi}\left\{\frac{k_\nu\bar{u}_\mu\bar{u}^\lambda - k_\mu\bar{u}_\nu\bar{u}^\lambda)}{k^\alpha\bar{u}_\alpha}\right\}, \tag{6.5}$$

which reduces to the usual form[24] in the frame of reference where $\bar{u}_\nu = (1, \mathbf{0})$ (and after neglecting the irrelevant radiation terms).

At infinite temperature, it may be shown that the radiation terms become infinite while the usual terms vanish.

7. *H*-THEOREM

The modified relativistic Vlasov equation contains, as we have seen, all the effects of the emission of radiation by the electrons of the plasma. This emisson of radiation is by essence *irreversible*. Therefore, we expect an irreversible behavior of the plasma itself and hence an H theorem to be valid. Let us now show that we actually have such a theorem.

To this end, we multiply both sides of Eq. (4.3) by

$$-k\log\mathcal{N}(x_\nu, u_\nu) \quad (k = \text{Boltzmann's const}) \tag{7.1}$$

and integrate the resulting equation over the four-velocities. We get

$$\partial_\mu S^\mu(x_\rho) + \frac{ke}{m}F^{\mu\nu}(x_\rho) \times \int d_4u\, u_\nu \frac{\partial}{\partial u^\mu}\log\mathcal{N}(x_\rho, u_\rho)\cdot\mathcal{N}(x_\rho, u_\rho) = \frac{\tau_0 ek}{m}\left\{\int d_4u \log\mathcal{N}(x_\rho, u_\rho) \times \frac{\partial}{\partial u^\mu}\{\Delta^{\mu\nu}(u_\rho)u^\rho u^\alpha\mathcal{N}(x_\rho, u_\rho)\} \times \partial_\rho F_{\nu\alpha}(x_\rho), \tag{7.2}$$

where the entropy four-current density $S^\mu(x_\rho)$ is given by[1]

$$S^\mu(x_\rho) = -k\int d_4u\cdot u^\mu\mathcal{N}(x_\rho, u_\rho)\log\mathcal{N}(x_\rho, u_\rho), \tag{7.3}$$

while the total entropy of the fluid is

$$S(\Sigma) = \int_\Sigma d\Sigma_\mu\cdot S^\mu(x_\rho),$$

where Σ is an arbitrary spacelike three-surface. The second term of the right-hand side of Eq. (7.2) vanishes as can easily be seen by first integrating by parts, and then taking into account the antisymmetry of $F^{\mu\nu}$. Finally, we obtain (after integrating by parts twice)

$$\partial_\mu S^\mu(x_\rho) = \frac{\tau_0 ek}{m}\left\{\int d_4u\mathcal{N}(x_\rho, u_\rho)\frac{\partial}{\partial u^\mu}[g^{\mu\nu}u^\rho u^\alpha]\right\}\partial_\rho F_{\nu\alpha}, \tag{7.4}$$

$$= \frac{\tau_0 ek}{m}\left\{\int d_4u\mathcal{N}(x_\rho, u_\rho)[g^{\rho\nu}u^\alpha + g^{\nu\alpha}u^\rho]\right\}\partial_\rho F_{\nu\alpha},$$

$$= \frac{\tau_0 k}{m}\{J^\alpha\partial_\nu F^\nu_\alpha + J^\rho\partial_\rho F^\alpha_\alpha\}, \tag{7.5}$$

or

$$\partial_\mu S^\mu(x_\rho) = \frac{\tau_0 k}{m}J^\alpha\cdot J_\alpha, \tag{7.6}$$

where we have used Maxwell's equations,

$$\partial_\mu F^{\mu\nu}(x_\rho) = J^\nu(x_\rho). \tag{7.7}$$

Since the four-current is necessarily a timelike four-vector, it follows from Eq. (7.6) that

$$\partial_\mu S^\mu(x_\rho) > 0 \tag{7.8}$$

which inequality expresses our H theorem in a covariant way. As is well known, Eq. (7.8) leads immediately to[25]

$$S(\Sigma_0) < S(\Sigma_1), \tag{7.9}$$

where the arbitrary spacelike three-surface Σ_1 is entirely *situated in the future* of Σ_0. Equation (7.8) or Eq. (7.9) expresses the increase of entropy in the course of "time."

Connection with the Einstein–Ritz Controversy

At the beginning of the century, an interesting controversy took place between Einstein and Ritz[26] on the relation between the use of retarded actions in physics and the Carnot principle. Essentially, Ritz's point of view was that the basis of Carnot principle was due to retarded actions only.

[24] See, e.g., S. Garthenhaus, *Elements of Plasma Physics* (Holt, Rinehart and Winston, Inc., New York, 1964).

[25] See, e.g., W. Israel, J. Math. Phys. **4**, 1163 (1963). Equation (7.9) may easily be obtained through an integration of Eq. (7.8) over a four-dimensional domain bounded by Σ_0, Σ_1, and a surface generated by timelike lines $S^\mu(x_\rho)$, which we make tend to infinity.

[26] A. Einstein and W. Ritz, Z. Physik **10**, 323 (1909).

This point of view has also been discussed by Costa de Beauregard.[27] Einstein's view on this problem was that in nature, actions should be time symmetrical and therefore the Carnot principle (and irreversibility) should follow mainly from probabilistic considerations.

In this paper we deal with retarded actions only and therefore it may be worthwhile examining more closely the connection of Einstein's and Ritz's views with our irreversible equations. Let us first notice that had we used *advanced actions* instead of retarded ones, we would have found the same equations as those derived in this paper, *except* that the terms in τ_0 would have been changed to terms in *minus* τ_0. Consequently, instead of Eq. (7.8) we would have found

$$\partial_\mu S^\mu(x_\rho) < 0, \tag{7.10}$$

which belongs to the equations governing the paradoxial universe of Flamarion[27]: entropy would always decrease! This remark seems to imply that it is only retarded actions which lead to the increase of entropy. Unfortunately, the situation is by no means so simple. Indeed, the irreversibility of our kinetic equation could perfectly well be a result of our approximations as is the case in Newtonian physics. (For instance, Boltzmann's equation is derived through approximations and assumptions which do introduce irreversibility.) However, we *believe* (but are as yet unable to prove) that the complete hierarchy at order one in τ_0 is irreversible as a result of *retarded interactions*. Let us specify this point more precisely. In a preceding paper,[2] we find a kinetic equation for the noninteracting particle embedded in an *external* electromagnetic force field $F^{\mu\nu}$. It was shown that this equation was reversible [i.e., $S(\Sigma) = \text{const}$, $\forall\Sigma$]. However, if we examine more closely the proof of this statement,[2] we can see that it rests on the assumption that the external field $F^{\mu\nu}$ was a *free field* $[\partial_\mu F^{\mu\nu}(x_\rho) = 0]$. When this hypothesis is relaxed, we again have an irreversible equation. From these remarks we can draw two conclusions: (a) The use of retarded actions does not necessarily imply irreversibility.[2,28] (b) Irreversibility seems to come from interactions, no matter whether they are symmetrical or not,[29] *and* the irreversible emission of radiation.

Finally, it seems to us that Ritz's views could be restated in the following form which is considerably weaker.[30]

Conjecture: The irreversible emission of radiation joined to interactions lead to H theorems for all "reasonable" kinetic equations (i.e., lead to a weak irreversibility[31]).

Of course, this "weak" irreversibility has nothing to do with the usual irreversibility occurring through coarse graining, for instance.

8. HYDRODYNAMICAL EQUATIONS

In order to get the various hydrodynamical equations governing a plasma obeying the modified Vlasov equation (4.3), it is sufficient to derive the various possible (first) moment equations. However, if we want to have a deeper insight of their physical meaning, we have to discuss several notions such as local averages etc.

A. Local Averages—Physical Quantities

When dealing with hydrodynamics we are mainly concerned with x_ν dependent quantities which are local averages.[1] Given a physical quantity $A^{\cdots}(x_\nu, u_\nu)$, we first define the "current of quantity $A^{\cdots}(x_\nu, u_\nu)$" as being

$$J^\mu(\{A^{\cdots}\}, x_\nu) = \int d_4u A^{\cdots}(x_\nu, u_\nu) u^\mu \mathcal{N}(x_\nu, u_\nu), \tag{8.1}$$

so that the *total average* of $A^{\cdots}(x_\nu, u_\nu)$ is

$$\langle A^{\cdots}\rangle_\Sigma = \int_\Sigma d\Sigma_\mu J^\mu(\{A^{\cdots}\}, x_\nu), \tag{8.2}$$

this total average not depending on Σ only when the following integrability condition is fulfilled:

$$\partial_\mu J^\mu(\{A^{\cdots}\}, x_\nu) = 0. \tag{8.3}$$

Let us now define *local averages*. To this end we need a conditional distribution function (x_ν given). Unfortunately, such a conditional distribution function *is in general not* unique. Indeed, given a field of timelike four-vectors $\alpha^\mu(x_\nu)$ (with $\alpha^\mu \cdot \alpha_\mu = 1$), all the possible conditional distribution functions are

$$\mathcal{N}(x_\nu, \{\alpha^\mu\} \mid u_\nu) = \frac{\mathcal{N}(x_\nu, u_\nu) u^\mu \alpha_\mu(x_\nu)}{J^\mu(x_\nu) \cdot \alpha_\mu(x_\nu)}. \tag{8.4}$$

The quantity $J^\mu \cdot \alpha_\mu$ represents the density of particles at point x_ν in a local frame of reference whose time axis is parallel to α_μ.

---

[27] O. Costa De Beauregard, *la théorie synthétique de la relativité et des quanta* (Gauthiers-Villars, Paris, 1957).

[28] Even the case of only one particle is too simple to be really significant.

[29] Indeed, if we start with the kinetic equation derived in Ref. 2, and if we impose that the external field is a self-consistent field (i.e., $\partial_\mu F^{\mu\nu} = J^\nu$ with J^ν = current due to the system), we recover the kinetic equation obtained in this paper. Therefore, it is the fact that we consider interactions (no matter what their symmetries) which introduces irreversibility (because of interaction $F^{\mu\nu}$ is no longer a free field and hence we have an H theorem) *joined* to the fact that *irreversible radiation emission* is taken into account.

[30] The preceding theory is of no help in a discussion (or a reformulation) of the first converse statement (see Ref. 27).

[31] Let us notice that this irreversibility is extremely "weak" since the right-hand side of Eq. (7.6) is of order $\sim e^2\tau_0$. However, we expect that at high temperatures it might become important. In the nonrelativistic case, this "weak" irreversibility is not important and we have to face coarse graining, etc., to deal with this problem.

One easily verifies that

$$\int d_4u\mathcal{N}(x_\nu, \{\alpha^\mu\} \mid u_\nu) = 1, \tag{8.5}$$

as expected. As a consequence, the local averages of a physical quantity $A^{\cdots}(x_\nu, u_\nu)$ will be defined as

$$\tilde{A}^{\cdots}(\{\alpha^\mu\}) = \int d_4u\mathcal{N}(x_\nu, \{\alpha^\mu\} \mid u_\nu)A^{\cdots}(x_\nu, u_\nu), \tag{8.6}$$

$$= \frac{J^\mu(\{A^{\cdots}\}, x_\nu)\alpha_\mu(x_\nu)}{J^\mu(x_\nu)\alpha_\mu(x_\nu)}, \tag{8.7}$$

[where J^μ designates the usual four-current

$$J^\mu(x_\nu) \equiv J^\mu(\{1\}, x_\nu)]. \tag{8.8}$$

In some important cases we know only one vector field $\alpha^\mu(x_\nu)$ so that the above local averages are uniquely defined. For instance, this is the case of a fluid at local equilibrium whose distribution function depends on only one vector field [e.g., from the Jüttner–Synge distribution we can get only one vector field, say $\bar{u}^\mu(x_\nu)$]. However, in the general case, we can construct an infinity of vector fields by contracting the infinite number of moments together. This feature of relativistic local averages is closely related to the absence of a unique notion of simultaneity.

The main physical quantities $A^{\cdots}$ are: (a) the mass m (or the electric charge, etc.), (b) the momentum–energy mu^μ, and (c) the entropy, which has already been considered in Sec. 7. To these quantities we may add $mu^\mu u^\nu$, which is related to the heat flux (when existing) within the gas.

The currents of these quantities are the usual current, the momentum–energy tensor,

$$T^{\mu\nu}(x_\rho) = \int d_4u m u^\mu u^\nu \mathcal{N}(x_\rho, u_\rho), \tag{8.9}$$

and the entropy four-current already considered.

B. Thermodynamical Quantities

Once the transformation law of entropy is fixed (information theory imposes that it is a scalar), the transformation law of all other thermodynamical quantities is completely determined when the one of temperature is given. However, the relativistic notion of temperature is far from being devoid of ambiguity. For instance, if temperature is *defined* through the Lagrange multipliers of energy–momentum (when looking for the equilibrium distribution function), then it should behave like $(T^{-1})^\mu$. On the other hand, temperature may also be *defined* as the local average of the kinetic energy of the fluid at point x_ν. In such a case it should behave like $(T)^\mu$. Of course, these two definitions mutually agree in the nonrelativistic case. Unfortunately, they are nearly completely disconnected in the relativistic framework (even in the case of the relativistic perfect gas). Furthermore, the relativistic notion of kinetic momentum energy itself is not unique! Indeed, kinetic momentum energy may be defined only with respect to a given vector field,

$$K^\mu = m(u^\mu - \alpha^\mu(x_\nu)) \tag{8.10}$$

so that the "current of kinetic momentum–energy" is

$$J^\mu(\{K^\mu\}, x_\nu) = m\int d_4u\mathcal{N}(x_\rho, u_\rho)u^\mu[u^u - \alpha^\nu(x_\nu)] \tag{8.11}$$

$$= T^{\mu\nu}(x_\rho) - mJ^\mu(x_\rho)\alpha^\nu(x_\rho). \tag{8.12}$$

Besides the arbitrariness in the choice of the vector field $\alpha^\mu(x_\rho)$, there exists an ambiguity in the mass term to be subtracted (*either* the rest mass *or* the rest mass plus the contribution of internal energy per particle).

In conclusion, we suggest adopting the idea of the existence of several notions of temperatures, the one defined through the Jüttner–Synge distribution function being referred to as the *equilibrium temperature*, another one defined as the local average of momentum–energy[32]:

$$3kT^\mu(x_\rho) = \frac{T^{\mu\nu}(x_\rho)\alpha_\nu(x_\rho)}{J^\mu(x_\rho)\alpha_\mu(x_\rho)} \tag{8.13}$$

and referred to as the *kinetic temperature*. The choice of $\alpha^\mu = J^\mu$ seems to be appropriate, although for instance we could readily choose α^μ to be the timelike eigenvector of $T^{\mu\nu}$.

In the case of the perfect gas[18,25] we find

$$3kT^\mu = \left\{m\frac{K_1(m\xi)}{K_2(m\xi)} + \frac{3}{\xi}\right\}\bar{u}^\mu,$$

from which we obtain

$$T_K = T_E + \frac{m}{k}\frac{K_1(m/kT_E)}{K_2(m/kT_E)},$$

where T_E and T_K are the equilibrium and the kinetic temperatures, respectively. Obviously, their nonequality reflects the equivalence of mass and energy. Indeed, at the nonrelativistic approximation, we have $T_E = T_K$ up to an irrelevant mass term $\sim mc^2$.

We shall return to these questions in another paper.

C. Local Equilibrium

Let us now return to our problem and examine whether or not a plasma obeying the kinetic equation (4.3) may be in a state of local equilibrium, i.e., whether or not its distribution function could be the Jüttner–Synge distribution (4.4) with ξ^μ and n_0 now depending on x_ν (although in a sufficiently "smooth"

[32] This temperature is related to internal energy.

way). To this end, we introduce Eq. (4.4) into Eq. (4.3) (with ξ^μ and n_0 depending on x_ν) and we find

$$u^\mu\left\{\partial_\mu A + eF_{\mu\nu}\xi^\nu A + \frac{\tau_0 e}{m} A\partial^\nu F_{\nu\mu}\right\} - u^\mu u^\lambda \times \{m\partial_\mu \xi_\lambda \cdot A + \tau_0 e A\xi^\nu \partial_\mu F_{\nu\lambda}\} = 0, \quad (8.14)$$

where we have set

$$A \equiv \frac{n_0(x_\nu)m\xi(x_\nu)}{4\pi K_2[m\xi(x_\nu)]}. \quad (8.15)$$

Equation (14) should be valid for any u^μ, so that it yields the following conditions:

$$\partial_\mu A + \left(eF_{\mu\nu}\xi^\nu + \frac{\tau_0 e^2}{m} n_0(x_\nu)\xi_\mu \xi^{-1}\right)A = 0, \quad (8.16)$$

$$m\partial_\mu \xi_\lambda + \tau_0 e\partial_\mu F_{\nu\lambda} \cdot \xi^\lambda = 0, \quad (8.17)$$

[to obtain Eq. (8.16) we used Maxwell's equations]. Equations (8.16) and (8.17) represent twenty equations for only five unknowns. Therefore, it is not certain that they have solutions. However, these equations seem to be compatible. In what follows we assume that they have either solutions or "quasi-solutions" [i.e., solutions up to terms in $O(\tau_0^2)$ or $O(e^4)$, etc.].

D. Moment Equations

1. Let us first integrate Eq. (4.3) over four-velocity space. We immediately get the continuity equation

$$\partial_\mu J^\mu(x_\rho) = 0, \quad (8.18)$$

where we have assumed a sufficiently rapidly decreasing behavior of $\mathcal{N}(x_\rho, u_\rho)$ at infinity in four-velocity space. This assumption will be made in what follows.

2. In order to get the equations satisfied by the second-order moments, we multiply Eq. (4.3) by mu^λ and integrate over four-velocity space. We obtain

$$\partial_\mu T^{\mu\lambda} - eF^{\lambda\mu}J_\mu = \tau_0 e g^\lambda_\mu \times \left\{\int d_4u \cdot \Delta^{\mu\nu}(u_\rho)u^\rho u^\alpha \mathcal{N}(x_\rho, u_\rho)\right\}\partial_\rho F_{\nu\alpha} \quad (8.19)$$

or

$$\partial_\mu T^{\mu\lambda} - eF^{\lambda\mu}J_\mu = \frac{\tau_0 e}{m} T^{\rho\alpha}\partial_\rho F^\lambda_\alpha. \quad (8.20)$$

This last equation expresses the conservation of momentum and energy. The first term on the left-hand side of Eq. (8.20) refers to the particles momentum energy, while the second term refers to the momentum energy of the electromagnetic field of the plasma. The right-hand side of Eq. (8.20) represents the contribution to momentum energy balance due to the radiation emission. When the plasma is homogenous, the electromagnetic field is constant and hence this term vanishes as expected. The radiation effects are of order $\sim\tau_0 e^2$ (or e^4) since $F^{\mu\nu}$ is of order e. Equation (8.20) could also be put in the following form:

$$\partial_\mu\{T^{\mu\nu}_{\text{particles}} + T^{\mu\nu}_{\text{e.m. field}} + T^{\mu\nu}_{\text{radiation field}}\} = 0. \quad (8.21)$$

3. The equations verified by the third-order moments are obtained by multiplying Eq. (4.3) by $mu^\lambda u^\gamma$ and integrating over four-velocity space, we obtain

$$\partial_\mu Q^{\mu\lambda\gamma} - e\{F^{\lambda\nu}T^\gamma_\nu + F^{\gamma\nu}T^\lambda_\nu\} = \frac{\tau_0 e}{m}\{g^{\lambda\nu}Q^{\gamma\rho\alpha} + g^{\gamma\nu}Q^{\lambda\rho\alpha}\}\partial_\rho F_{\nu\alpha}. \quad (8.22)$$

The tensor $Q^{\mu\lambda\gamma}$ is sometimes referred to as the "*heat flux tensor*." [25] The right-hand side of Eq. (8.22) is, once again, the contribution of radiation emission and vanishes when the system is homogeneous.

4. More general moment equations than Eqs. (8.18), (8.20), and (8.22) may be easily obtained. However, for practical purposes such as the erection of a relativistic magnetohydrodynamics, they are sufficient. Of course, we have to add the specific forms of the tensors J^μ, $T^{\mu\nu}$, $Q^{\mu\lambda\gamma}$, and possibly an equation of state or the knowledge of the conductivity tensor, etc. Furthermore, $F^{\mu\nu}$ and J^μ are connected through Maxwell's equations. Note also the following relations:

$$T^\mu_\mu(x_\rho) = m\int d_4u \cdot \mathcal{N}(x_\rho, u_\rho), \quad (8.23)$$

$$Q^{\lambda\mu}_\mu(x_\rho) = mJ^\lambda(x_\rho). \quad (8.24)$$

E. Hydrodynamics of a Small Disturbance

Linearizing Eq. (4.3) (around an *equilibrium* solution), we get Eq. (4.8). This equation immediately yields the linearized hydrodynamical equations

$$\partial_\mu T^{\mu\nu}_{(1)} - eF^{\mu\nu}_{(1)}J_{\mu(0)} = \frac{\tau_0 e}{m} T^{\rho\alpha}_{(0)}\partial_\rho F^\nu_{\alpha(1)} \quad (8.25)$$

and

$$\partial_\mu Q^{\mu\lambda\gamma}_{(1)} - e\{F^{\lambda\nu}_{(1)}T^\gamma_{\nu(0)} + F^{\gamma\nu}_{(1)}T^\lambda_{\nu(0)}\} = \frac{\tau_0 e}{m}\{g^{\lambda\nu}Q^{\gamma\rho\alpha}_{(0)} + g^{\gamma\nu}Q^{\lambda\rho\alpha}_{(0)}\}\partial_\rho F_{\nu\alpha(1)} \quad (8.26)$$

which could just as well have been derived from Eqs. (8.20) and (8.22). In Eqs. (8.25) and (8.26), indices (0) or (1) denote the zeroth- or first-order terms, respectively.

The tensors $T^{\mu\nu}_{(0)}$ and $Q^{\mu\nu\gamma}_{(0)}$ are given by [18,25]

$$T^{\mu\nu}_{(0)} = m\frac{K_3(m\xi)}{K_2(m\xi)} n_0 \bar{u}^\mu \bar{u}^\nu - \frac{n_0}{\xi} g^{\mu\nu}, \quad (8.27)$$

$$Q^{\mu\nu\lambda}_{(0)} = -\frac{\partial}{\partial \xi_\lambda} T^{\mu\nu}_{(0)}\left\{\frac{4\pi K_2(m\xi)}{m\xi}\right\}. \quad (8.28)$$

F. Calculation of $T^{\lambda\gamma}_{(1)}$

We start with Eq. (4.8) which we rewrite in the form

$$u^\mu \partial_\mu \mathcal{N}_{(1)} + \frac{e}{m} F^{\mu\nu}_{(1)} u_\nu \frac{\partial}{\partial u_\mu} \mathcal{N}_{(0)} = -\frac{\tau_0 e}{m} \cdot \frac{\partial}{\partial u^\mu} \times \{\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \mathcal{N}_{(0)}\} \partial_\rho F_{\nu\alpha(1)}. \quad (8.29)$$

Furthermore, we provisionally assume that $\mathcal{N}_{(0)}$ is a *local equilibrium* distribution function. Now taking the Fourier transform of this last equation, we get

$$\hat{\mathcal{N}}_{(1)}(k_\rho, u_\rho) = \frac{1}{ik^\mu u_\mu} \Big\{ -\frac{e}{m} \hat{F}^{\mu\nu}_{(1)}(k_\rho) u_\nu \frac{\partial}{\partial u^\mu} \hat{\mathcal{N}}_{(0)}(k_\rho, u_\rho) - \frac{\tau_0 e}{m} \frac{\partial}{\partial u^\mu} [\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \hat{\mathcal{N}}_{(0)}(k_\rho, u_\rho)] * [ik_\rho \hat{F}_{\nu\alpha(1)}(k_\rho)] \Big\}, \quad (8.30)$$

where the convolution $*$ concerns the k_α variables. Multiplying Eq. (8.30) by $mu^\lambda u^\gamma$, we get, after on integration over four velocities,

$$\hat{T}^{\gamma\lambda}_{(1)} = \frac{1}{i} \int d_4k' \{ -e\hat{F}^{\mu\nu}_{(1)}(k_\rho - k'_\rho) \cdot \int d_4u \frac{u_\nu u^\lambda u^\gamma}{k^\mu u_\mu} \frac{\partial}{\partial u^\mu} \hat{\mathcal{N}}_{(0)}(k'_\rho, u_\rho) - \frac{\tau_0 e}{i} \int d_4k' \Big\{ (k_\rho - k'_\rho) \hat{F}_{\nu\alpha(1)}(k_\rho - k'_\rho) \cdot \int d_4u \cdot \frac{u^\lambda u^\gamma}{k^\mu u_\mu} \frac{\partial}{\partial u^\mu} [\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \hat{\mathcal{N}}_{(0)}(k_\rho, u_\rho)]. \quad (8.31)$$

Equation (8.31) allows us to obtain $T^{\lambda\gamma}_{(1)}$ in terms of an integral expression involving $F^{\mu\nu}_{(1)}$ and terms which may be calculated at least in principle. Equation (8.31) may be simplified further when dealing with the problem of a small disturbance occurring in a *uniform and homogeneous* plasma. In such a case, Eq. (8.31) becomes

$$i\hat{T}^{\lambda\gamma}_{(1)}(k_\rho) = +\tau_0 e k_\rho \hat{F}_{\nu\alpha(1)}(k_\rho) \int d_4u \frac{u^\lambda u^\gamma}{k^\mu u_\mu} \frac{\partial}{\partial u^\mu} \times [\Delta^{\mu\nu}(u_\rho) u^\rho u^\alpha \mathcal{N}_{\text{eq}}(u_\rho)] - eF^{\mu\nu}_{(1)}(k_\rho) \int d_4u \times \frac{u_\nu u^\lambda u^\gamma}{k^\mu u_\mu} \frac{\partial}{\partial u^\mu} \mathcal{N}_{\text{eq}}(u_\rho). \quad (8.32)$$

Equation (8.32) may be rewritten as

$$\hat{T}^{\gamma\lambda}_{(1)}(k_\rho) = \frac{e}{i} \hat{F}^{\nu\mu}_{(1)}(k_\rho) \int d_4u \mathcal{N}_{\text{eq}}(u_\rho) u_\nu \times \left[\frac{g^\lambda_\mu u^\gamma + g^\gamma_\mu u^\lambda}{k^\alpha u_\alpha} - \frac{k_\mu u^\gamma u^\lambda}{(k^\alpha u_\alpha)^2} \right] + \tau_0 e k_\rho \hat{F}_{\nu\alpha(1)}(k_\rho) \times \int d_4u \mathcal{N}_{\text{eq}}(u_\rho) u^\rho u^\alpha \left[\frac{g^{\lambda\nu} u^\gamma + g^{\gamma\nu} u^\lambda}{k^\alpha u_\alpha} - \frac{k^\nu u^\lambda u^\gamma}{(k^\alpha u_\alpha)^2} \right]. \quad (8.33)$$

Using the integrals defined in Sec. 5, this last equation reads (K is defined in the Appendix):

$$\hat{T}^{\gamma\lambda}_{(1)}(k_\rho) = \frac{e}{i} n_0 \hat{F}^{\mu\nu}_{(1)}(k_\rho) \{ -I^\gamma_\nu g^\lambda_\mu - I^\lambda_\nu g^\gamma_\mu + k_\mu \frac{\partial}{\partial k_\nu} J^{\gamma\lambda} \} + \tau_0 e \hat{F}_{\nu\alpha(1)}(k_\rho) \Big\{ g^{\lambda\nu} T^{\alpha\gamma}_{(0)} + g^{\gamma\nu} T^{\alpha\lambda}_{(0)} - \frac{\partial [K^{-1} n_0 I^{\gamma\lambda}]}{\partial (m\xi_\alpha)} \Big\}. \quad (8.34)$$

The hydronamical equations (8.25) now read

$$k_\lambda \hat{T}^{\gamma\lambda}_{(1)} - e^2 n_0 \xi^{-1} F^{\gamma\lambda}_{(1)} \xi_\lambda = \frac{i\tau_0 e}{m} T^{\rho\alpha}_{(0)} k_\rho \hat{F}^\gamma \alpha_{(1)}, \quad (8.35)$$

and more explicit expressions may be obtained by using Eqs. (34) and (35). Of course, more complicated equations are obtained for $Q^{\lambda\mu\nu}_{(1)}$.

ACKNOWLEDGMENT

One of the authors (R. H.) is indebted to Professor Peter Havas for his kind hospitality at Temple University and for interesting comments and discussions.

APPENDIX: CALCULATION OF SOME INTEGRALS

In the derivation of the dispersion relations and also in the expression of the conductivity tensor, the following integrals occur:

$$I_\mu = K \int \frac{d_3u}{u^0} \cdot \frac{u_\mu}{k^\lambda u_\lambda} \exp\{-m\xi^\alpha u_\alpha\}, \quad (A1)$$

$$I_{\mu\nu} = -K \int \frac{d_3u}{u^0} \cdot \frac{u_\mu u_\nu}{(k^\lambda u_\lambda)^2} \exp\{-m\xi^\alpha u_\alpha\}, \quad (A2)$$

$$J_{\mu\nu} = K \int \frac{d_3u}{u^0} \cdot \frac{u_\mu u_\nu}{k^\lambda u_\lambda} \exp\{-m\xi^\alpha u_\alpha\}, \quad (A3)$$

where K is given by

$$K = \frac{m\xi}{4\pi K_2(m\xi)}. \quad (A4)$$

In all what follows, we set

$$m\xi = \alpha. \quad (A5)$$

The preceding integrals (A1), (A2), (A3) are interrelated in the following way[33]:

$$I_{\mu\nu} = \frac{\partial}{\partial k^\nu} I_\mu = \frac{\partial}{\partial k^\mu} I_\nu, \quad (A6)$$

$$J_{\mu\nu} = -\frac{\partial}{\partial (m\xi^\nu)} I_\mu = -\frac{\partial}{\partial (m\xi^\mu)} I_\nu, \quad (A7)$$

so that these integrals may be (at least in principle) obtained from the knowledge of I_μ only. Note also the useful relations,

$$I_{\mu\nu} k^\nu = -I_\mu, \quad (A8)$$

$$J_{\mu\nu} k^\nu = \int \frac{d_3u}{u^0} u_\mu K \exp\{-\alpha \bar{u}^\lambda u_\lambda\} = \bar{u}_\mu \equiv \xi_\mu \xi^{-1}; \quad (A9)$$

we first calculate I_μ.

[33] In Eq. (A7), the derivations with respect to ξ^μ *do not* apply to the factor K involved in I_μ which is to be regarded as a constant in ξ.

A. Calculation of I_μ (Generalities)

Since I_μ is a four-vector, it can be expressed as a linear combination of the only four-vectors which are at our disposal, i.e., ξ_μ and k_μ. Therefore, we can set

$$I_\mu = c_1\xi_\mu + c_2 k_\mu, \tag{A10}$$

where c_1 and c_2 are scalar functions of the following invariants:

$$\xi^\mu\xi_\mu, \quad \xi^\mu k_\mu, \quad k^\mu k_\mu. \tag{A11}$$

Multiplying Eq. (A10) successively by k_μ and ξ_μ, we get the following linear system for c_1 and c_2:

$$\begin{aligned} k^\mu I_\mu &= c_1 k^\mu\xi_\mu + c_2 k^\mu k_\mu, \\ \xi^\mu I_\mu &= c_1\xi^\mu\xi_\mu + c_2\xi^\mu k_\mu, \end{aligned} \tag{A12}$$

from which we at once obtain

$$c_1 = -\frac{k^\mu\xi_\mu\varphi - k^\mu k_\mu\psi}{(k^\mu k_\mu)(\xi^\mu\xi_\mu) - (k^\mu\xi_\mu)^2}, \tag{A13}$$

$$c_2 = \frac{\xi^\mu\xi_\mu\varphi - k^\mu\xi_\mu\psi}{(k^\mu k_\mu)(\xi^\mu\xi_\mu) - (k^\mu\xi_\mu)^2}, \tag{A14}$$

where we have set

$$\varphi \equiv k^\mu I_\mu = \int \frac{d_3u}{u^0} K \exp\{-\alpha\bar{u}^\mu \cdot u_\mu\} \tag{A15}$$

$$= K_1(\alpha)/K_2(\alpha), \tag{A16}$$

$$\psi \equiv \xi^\mu I_\mu = K\int \frac{d_3u}{u^0}\cdot\frac{\xi^\mu u_\mu}{k^\mu u_\mu}\exp\{-\alpha\bar{u}^\mu\cdot u_\mu\}. \tag{A17}$$

We first rewrite ψ in a Lorentz frame where ξ^μ reduces to $(\xi, \mathbf{0})$ and studying propagation along the third axis,

$$k^\mu = (k^0, 0, 0, k^3) \underset{\text{def}}{\equiv} (\omega, 0, 0, k).$$

Note that in this Lorentz frame, one has

$$\psi = \xi^0 I_0 = \xi I_0. \tag{A18}$$

Now using relativistic polar coordinates,

$$\begin{aligned} u^0 &= \cosh\chi, \\ u^1 &= \sinh\chi\sin\theta\cos\phi, \\ u^2 &= \sinh\chi\sin\theta\sin\phi, \\ u^3 &= \sinh\chi\cos\theta, \end{aligned} \tag{A19}$$

Eq. (A17) may be written as

$$\psi = K\xi\int_0^\infty\int_{-\pi/2}^{+\pi/2}\int_0^{2\pi} d\chi\, d\theta\, d\phi \times \frac{\sinh^2\chi\cdot\cosh\chi\cdot\sin\theta\exp[-\alpha\cosh\chi]}{\omega\cosh\chi - k\sinh\chi\cdot\cos\theta} \tag{A20}$$

$$= \frac{2\pi K\xi}{k}\int_0^\infty\int_{-\pi/2}^{+\pi/2} d\chi\, d(\cos\theta) \times \frac{\sinh^2\chi\cosh\chi}{\mu\cosh\chi - \sinh\chi\cos\theta}\exp[-\alpha\cosh\chi], \tag{A21}$$

where we have set

$$\mu = \frac{\omega}{k} = \text{phase velocity}; \quad \alpha = m\xi. \tag{A22}$$

It is important to bear in mind that the expression

$$(\mu\cosh\chi - \sinh\chi\cos\theta)^{-1} \tag{A23}$$

appearing in the integral (A21) has the following significance:

$$\{\mu\cosh\chi - \sinh\chi\cos\theta\}^{-1} \equiv i\pi\delta^-\{\mu\cosh\chi - \sinh\chi\cos\theta\} = i\pi\delta\{\mu\cosh\chi - \sinh\chi\cos\theta\} + P\{\mu\cosh\chi - \sinh\chi\cos\theta\}^{-1}, \tag{A24}$$

so that the calculation of ψ actually reduces to the calculation of the two integrals

$$\operatorname{Im}\psi = \frac{2\pi^2 K\xi}{k}\iint d\chi\, d(\cos\theta) \times \delta\{\mu\cosh\chi - \sinh\chi\cos\theta\}\sinh^2\chi\cosh\chi \times \exp[-\alpha\cosh\chi], \tag{A25}$$

$$\operatorname{Re}\psi = \frac{2\pi K\xi}{k}\iint d\chi\, d(\cos\theta) \times \frac{\sinh^2\chi\cosh\chi}{\mu\cosh\chi - \sinh\chi\cos\theta} \times \exp[-\alpha\cosh\chi]. \tag{A26}$$

B. Calculation of Im ψ

The integral over θ is easily computed and yields

$$\operatorname{Im}\psi = \frac{2\pi^2 K\xi}{k}\int_0^\infty d(\cosh\chi)\cdot\cosh\chi\exp[-\alpha\cosh\chi] \tag{A27}$$

with the condition

$$\mu\coth\chi \leq 1. \tag{A28}$$

Finally we obtain

$$\operatorname{Im}\psi = \frac{\pi}{2k}\cdot\frac{\xi}{K_2(\alpha)}\exp\left\{\frac{-\alpha}{(1-\mu^2)^{\frac{1}{2}}}\right\}\cdot\left[\frac{1}{\alpha} + \frac{1}{(1-\mu^2)^{\frac{1}{2}}}\right]. \tag{A29}$$

C. Calculation of Re ψ

Performing the θ integration, Eq. (A26) becomes

$$\operatorname{Re}\psi = \frac{2\pi K\xi}{k}\int_0^\infty d(\cosh\chi)\cosh\chi\exp[-\alpha\cosh\chi] \times \log\left|\frac{1+\mu\coth\chi}{1-\mu\coth x}\right|. \tag{A30}$$

Integrating by part this last expression, we get

$$\operatorname{Re}\psi = -\frac{\xi\mu\alpha}{kK_2(\alpha)}\cdot\frac{d}{d\alpha} \times \left\{\alpha^{-1}\int_0^{+\infty}\frac{dx\exp\{-\alpha\cosh\chi\}}{(\mu^2-1)\cosh^2\chi + 1}\right\}. \tag{A31}$$

Taking into account the fact that

$$K_2(\alpha) \sim \alpha^{-2}, \tag{A32}$$

when α tends to zero, a brief inspection of Eq. (A29) and (A31) shows immediately that ψ tends to zero at the extreme relativistic limit ($\alpha = 0$). The integral (A31) may now be evaluated approximately for various physical situations.

D. Calculation of I_3, I_0^0, I_1^1, I_3^3, I_3^0

These integrals are the only nonvanishing ones among the I_μ^ν, I_μ.

Starting from Eq. (A16), we get

$$k^\mu I_\mu = K_1(\alpha)/K_2(\alpha), \\ = \omega I_0 + k I_3, \tag{A33}$$

and hence,

$$I_3 = \left\{\frac{K_1(\alpha)}{K_2(\alpha)} - \omega I_0\right\} \cdot k^{-1}. \tag{A34}$$

Inspection of Eqs. (A21) and (A18) shows immediately that

$$I_0^0 = \frac{\partial}{\partial \omega} I_0. \tag{A35}$$

To calculate I_1^1 it is sufficient to use Eqs. (A6), (A10), and (A17). It turns out that

$$I_1^1 = c_2 = \left\{\omega I_0 - \frac{K_1(\alpha)}{K_2(\alpha)}\right\} \cdot k^{-2}. \tag{A36}$$

To calculate I_3^3 it is sufficient to use relation (A8),

$$I_\mu^\nu k^\mu = -I^\nu \\ = I_0^\nu \omega + I_3^\nu k \tag{A37}$$

or

$$+I_3 = -I^3 = I_0^3 \omega + I_3^3 k, \tag{A38}$$

so that we actually need I_0^3 in order to calculate I_3^3. We have in a straightforward manner,

$$I_3^0 = -I_0^3 = \frac{\partial}{\partial \omega} I_3, \tag{A39}$$

or

$$I_3^0 = -k^{-1}\left\{\omega \frac{\partial}{\partial \omega} I_0 + I_0\right\}. \tag{A40}$$

Substituting Eq. (A40) into Eq. (A38), we get

$$I_3^3 = k^{-2}\left\{\frac{K_1(\alpha)}{K_2(\alpha)} - \frac{\partial}{\partial \omega}(\omega^2 I_0)\right\}. \tag{A41}$$

E. Calculation of J_0^0, J_1^1, J_3^3, J_3^0

In the same way as above we, have

$$J_0^0 = -K \frac{\partial}{\partial \alpha}(K^{-1} I_0), \\ J_1^1 = -m^{-1} c_1, \tag{A42}$$

and more explicitly,

$$J_1^1 = m^{-1} k^{-2}\left\{\frac{K_1(\alpha)}{K_2(\alpha)} - \omega I_0\right\}; \tag{A43}$$

we also have

$$J_3^0 = -K \frac{\partial}{\partial \alpha}(K^{-1} I_3), \\ = -K \frac{\partial}{\partial \alpha}\left\{\frac{4\pi K_2(\alpha)}{\alpha k}\left[\frac{K_1(\alpha)}{K_2(\alpha)} - \omega I_0\right]\right\}, \tag{A44}$$

and finally, [using Eq. (A9)],

$$J_3^3 = +\mu J_3^0$$

or

$$J_3^3 = \mu K \frac{\partial}{\partial \alpha}\left\{\frac{4\pi K_2(\alpha)}{\alpha k}\left[\frac{K_1(\alpha)}{K_2(\alpha)} - \omega I_0\right]\right\}. \tag{A45}$$

In Eqs. (A42)–(A45), K is given by Eq. (A4).

F. Evaluation of Integrals at the Extreme Relativistic Approximation ($\alpha \sim 0$)

At the extreme relativistic approximation, most particles have an energy which is much greater than the rest-mass energy. Therefore, at this approximation we have practically massless particles. Hence, instead of integrating the various above integrals on the hyperboloid $u^\mu u_\mu = 1$, we can evaluate them on the null cone $u^\mu u_\mu = 0$ (with $m\xi \sim 0$). At this approximation, the normalization factor given in Eq. (A4) becomes[18]

$$K = \frac{\alpha^3}{8\pi}. \tag{A46}$$

As an example, we now calculate the real part of I^0. We have

$$\text{Re } I^0 = \frac{\alpha^3}{8\pi} \int R^2 \, dR \, d\theta \, d\phi \, \frac{\sin\theta \exp[-\alpha R]}{R\omega - Rk\cos\theta}, \tag{A47}$$

and finally

$$\text{Re } I^0 = \frac{\alpha}{4k} \log\left|\frac{1-\mu}{1+\mu}\right|, \tag{A48}$$

which tends to zero with α. Using Eq. (A29) and

$$K_2(\alpha) \simeq \frac{2}{\alpha^2} \quad (\text{with } \alpha \ll 1), \tag{A49}$$

we find

$$\text{Im } I^0 = \frac{\pi}{8k} \alpha^2 \ll \text{Re } I^0. \tag{A50}$$

The extreme relativistic behavior of the other integrals may be found in the same way. Essentially, the integrals I^μ and I_ν^μ tends to zero while the integrals J_ν^μ remain finite except when $\omega = k$, in which case they diverge.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Cluster Expansions in Many-Fermion Theory. I. "Factor-Cluster" Formalisms*

JOHN W. CLARK†
Compton Laboratory of Physics, Washington University, St. Louis, Missouri
AND
PAUL WESTHAUS
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

(Received 5 June 1967)

Cluster development may furnish a powerful device for the calculation of the expectation values of the observables of a many-fermion system with respect to dynamically correlated state vectors. The generalized normalization integral, a generating function for the required expectation values, is defined, and four of the many possible decompositions of this function into cluster integrals are explored. Two of these decompositions are slight extensions of the conventional ones of Iwamoto and Yamada and Aviles, Hartogh, and Tolhoek. The other two are product decompositions, leading to new "factor-cluster" formalisms. A factor-cluster expansion is applied to the evaluation of the n-particle spatial distribution function.

1. INTRODUCTION

In this paper we explore the formal aspects of cluster expansions as a tool for the systematic evaluation of expectation values of the observables of a system of N identical fermions with respect to dynamically correlated state vectors. The techniques to be developed may be useful in the calculation of properties of the bound states of such finite systems as nuclei and the electronic subsystems of atoms and molecules in the fixed nuclei approximation, and of such infinite systems as quantum fluids (including liquid He^3 and nuclear matter) and quantum solids (solid He^3).

In treating a many-fermion system, one usually starts with some intelligently chosen independent-particle model and then corrects this model for its most disconcerting inadequacies. There will in general be important correlation effects, whose description is by definition outside the scope of the input independent-particle model. We devote the major portion of this introduction to a discussion of how correlation effects may be built into the assumed form for the many-particle wavefunction.

The correlation structure of the exact stationary-state wavefunction has been investigated by many authors.[1–11] Their discoveries may be conveniently reviewed in terms of the highly formal expression

$$\Psi_{\mathbf{m}}(x_1 \cdots x_N) = (N!)^{-\frac{1}{2}} \langle 0| \psi(x_N) \cdots \psi(x_1) \mathcal{F}_{\mathbf{m}} a^\dagger_{m_N} \cdots a^\dagger_{m_1} |0\rangle \quad (1)$$

for the exact wavefunction. Here $a^\dagger_\kappa$ and a_κ are the usual fermion creation and annihilation operators associated with a complete set of single-particle states. The one-particle field operator $\psi(x_i)$ is a linear combination

$$\psi(x_i) = \sum_\kappa \varphi_\kappa(x_i) a_\kappa \quad (2)$$

of the annihilation operators a_κ, the coefficients $\varphi_\kappa(x_i)$ being the configuration-space representatives of the single-particle states κ. The argument x_i stands for all the coordinates—space ($\mathbf{r}_i$), spin (s_{z_i}), and, when appropriate, isospin (t_{z_i})—of the ith particle. $|0\rangle$ is the zero-particle state, the vacuum. The specific ket

$$|\Phi_{\mathbf{m}}\rangle = a^\dagger_{m_N} \cdots a^\dagger_{m_1} |0\rangle,$$

a (N-particle) basis ket of the occupation number representation generated by the $a^\dagger_\kappa$'s, represents the (input) independent-particle approximation to the exact N-fermion state of interest. We use $\mathbf{m}$ to denote the collection of labels $m_1, m_2, \cdots, m_N$, and, to be definite, take $m_1 < m_2 < \cdots < m_N$. The basis $\{|\varphi_\kappa\rangle\}$ of single-particle kets, or orbitals, may, for example, be chosen according to the Hartree–Fock scheme, the $N \times N$ determinant of a particular set of N such orbitals yielding a "self-consistent solution" to the N-body problem. The ket $|\Psi_{\mathbf{m}}\rangle$ corresponding to the

* Supported in part by the National Science Foundation, under Grant Nos. GP-3211 and GP-6210, a National Aeronautics and Space Administration Grant to Washington University, and National Aeronautics and Space Administration Grant NsG-275-62.

† Alfred P. Sloan Fellow.

[1] H. Primas, in *Modern Quantum Chemistry*, O. Sinanoğlu, Ed. (Academic Press Inc., New York and London, 1965), Vol. II, p. 45.

[2] W. Brenig, Nucl. Phys. **4**, 363 (1957).

[3] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1950).

[4] H. Kümmel, in *Lectures on the Many-Body Problem*, E. R. Caianiello, Ed. (Academic Press Inc., New York, 1962), p. 265ff.

[5] O. Sinanoğlu, in *Advances in Chemical Physics*, I. Prigogine, Ed. (Interscience Publishers, Inc., New York, 1963), Vol. VI.

[6] O. Sinanoğlu, Proc. Natl. Acad. Sci. U.S. **47**, 1217 (1961).

[7] O. Sinanoğlu, Rev. Mod. Phys. **35**, 517 (1963).

[8] O. Sinanoğlu, J. Chem. Phys. **36**, 706 (1962).

[9] O. Sinanoğlu, J. Chem. Phys. **36**, 3198 (1962).

[10] J. da Providencia, Nucl. Phys. **44**, 572 (1963).

[11] J. da Providencia, Nucl. Phys. **46**, 401 (1963).

wavefunction $\Psi_{\mathbf{m}}(x_1 \cdots x_N)$ as given in (1) is supposed to have unit overlap with $|\Phi_{\mathbf{m}}\rangle$. $\mathcal{F}_{\mathbf{m}}$, *the correlation operator*, serves to convert $|\Phi_{\mathbf{m}}\rangle$ into $|\Psi_{\mathbf{m}}\rangle$.

The familiar configuration interaction (CI) expansion for the wavefunction arises out of the following expansion for the correlation operator:

$$\mathcal{F}_{\mathbf{m}} = 1 + \sum_i \gamma^{\mathbf{m}}_{m_i} + \sum_{i<j} \gamma^{\mathbf{m}}_{m_i m_j} + \sum_{i<j<k} \gamma^{\mathbf{m}}_{m_i m_j m_k} + \cdots + \gamma^{\mathbf{m}}_{m_1 \cdots m_N} \quad (3)$$

in which the n-body excitation operator $\gamma^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}$ is defined by

$$\gamma^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}} \equiv \sum_{\kappa_1 < \cdots < \kappa_n \notin \mathbf{m}} C_{m_N \cdots \kappa_n \cdots \kappa_1 \cdots m_1; m_1 \cdots m_N} \times a^\dagger_{\kappa_n} \cdots a^\dagger_{\kappa_1} a_{m_{j(1)}} \cdots a_{m_{j(n)}}. \quad (4)$$

The CI coefficient

$$C_{m_N \cdots m_{j(n)+1} \kappa_n m_{j(n)-1} \cdots m_{j(1)+1} \kappa_1 m_{j(1)-1} \cdots m_1; m_1 \cdots m_{j(1)} \cdots m_{j(n)} \cdots m_N}$$

is just the amplitude, in the usual CI expansion for the wavefunction, of the independent-particle wavefunction which results when the orbitals $\varphi_{m_{j(1)}}, \cdots, \varphi_{m_{j(n)}}$ appearing in

$$\Phi_{\mathbf{m}} = (N!)^{-\frac{1}{2}} \langle 0| \psi(x_N) \cdots \psi(x_1) |\Phi_{\mathbf{m}}\rangle$$

are replaced, respectively, by orbitals $\varphi_{\kappa_1}, \cdots, \varphi_{\kappa_n}$, $\kappa_1 < \cdots < \kappa_n \notin \mathbf{m}$.

A form equivalent to (1) with (3) inserted has been introduced by Sinanoğlu[8]:

$$\begin{aligned}\Psi_{\mathbf{m}}(x_1 \cdots x_N) &= \mathcal{A}(N)\Big[\prod_{i=1}^N \varphi_{m_i}(x_i) + \sum_p \prod_{i \neq p} \varphi_{m_i}(x_i) U^{\mathbf{m}}_{m_p}(x_p) \\ &\quad + (2!)^{-\frac{1}{2}} \sum_{p<q} \prod_{i \neq p,q} \varphi_{m_i}(x_i) U^{\mathbf{m}}_{m_p m_q}(x_p, x_q) \\ &\quad + (3!)^{-\frac{1}{2}} \sum_{p<q<r} \prod_{i \neq p,q,r} \varphi_{m_i}(x_i) U^{\mathbf{m}}_{m_p m_q m_r}(x_p, x_q, x_r) \\ &\quad + \cdots + (N!)^{-\frac{1}{2}} U^{\mathbf{m}}_{m_1 \cdots m_N}(x_1 \cdots x_N)\Big]. \end{aligned} \quad (5)$$

Here $\mathcal{A}(N)$ is the N-particle antisymmetrizer

$$\mathcal{A}(N) = (N!)^{-\frac{1}{2}} \sum_\nu (-1)^\nu P_\nu, \quad (6)$$

the sum containing, with their signatures attached, all permutation operators on the set of N particle labels. Note $\mathcal{A}(N)$ is normalized such that $[\mathcal{A}(N)]^2 = (N!)^{\frac{1}{2}}\mathcal{A}(N)$. The identification

$$\begin{aligned} U^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}) &= (n!)^{-\frac{1}{2}} \langle 0| \psi(x_{j(n)}) \cdots \psi(x_{j(1)}) \gamma^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}} \\ &\quad \times a^\dagger_{m_{j(n)}} \cdots a^\dagger_{m_{j(1)}} |0\rangle \end{aligned} \quad (7)$$

is to be made. This form is studied in detail in the Appendix.

If we contemplate successive incorporation of one-body, two-body, $\cdots$, n-body, $\cdots$ correlation effects, the form (1) with (3) inserted, or, equivalently, Sinanoğlu's form (5), has the disturbing feature that specific n-body correlation effects are present not only in the terms involving n-indexed $\gamma \ldots$'s (or $U \ldots$'s), but also in all succeeding terms. One says that "unlinked clusters" are present. As N grows very large these expansions for the N-body wavefunction cease to have any meaning, since an equivalently large number of terms are required for a description of even two-body correlation effects.

The occurrence of "unlinked clusters" may be traced to the reducibility of the γ operators (or the U functions).[1] Upon asserting a cluster decomposition

$$\begin{aligned} \gamma^{\mathbf{m}}_{m_{j(1)}} &= \beta^{\mathbf{m}}_{m_{j(1)}} \\ \gamma^{\mathbf{m}}_{m_{j(1)} m_{j(1)}} &= \beta^{\mathbf{m}}_{m_{j(1)}} \beta^{\mathbf{m}}_{m_{j(2)}} + \beta^{\mathbf{m}}_{m_{j(1)} m_{j(2)}}, \\ &\cdot \\ &\cdot \\ &\cdot \\ \gamma^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}} &= \sum_{k=1}^n \Big\{ \sum_{\substack{\text{all partitions} \\ \text{of } m_{j(1)} \ldots m_{j(n)} \\ \text{among } n-k+1\ \beta\text{'s}}} \beta^{\mathbf{m}}_{\ldots} \beta^{\mathbf{m}}_{\ldots} \cdots \beta^{\mathbf{m}}_{\ldots} \Big\} \end{aligned} \quad (8)$$

of the excitation operators γ in terms of the *irreducible cluster operators* β, and invoking the (essentially obvious) facts that any two $\beta \ldots$'s commute and the product of two $\beta \ldots$'s with overlapping indices is zero, $\mathcal{F}_{\mathbf{m}}$ may readily be cast into the product form

$$\begin{aligned} \mathcal{F}_{\mathbf{m}} = \Big\{\prod_{i=1}^N (1 + \beta^{\mathbf{m}}_{m_i})\Big\}\Big\{\prod_{i<j} (1 + \beta^{\mathbf{m}}_{m_i m_j})\Big\} \\ \times \Big\{\prod_{i<j<k} (1 + \beta^{\mathbf{m}}_{m_i m_j m_k})\Big\} \cdots \{1 + \beta^{\mathbf{m}}_{m_1 \cdots m_N}\}. \end{aligned} \quad (9)$$

As opposed to the aforementioned expansions for $\Psi_{\mathbf{m}}$, Eq. (1) with (9) inserted displays a very clean separation of one-body, two-body, $\cdots$, n-body, $\cdots$ correlation effects. A single correlation factor operator $1 + \beta^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}$ appears for each of the $N!/[(N-n)!\,n!]$ distinct clustering possibilities for particles in n orbitals selected from the set $\varphi_{m_1} \cdots \varphi_{m_N}$, there being $2^N - 1$ factors in all. Now, if we want an approximation for $\Psi_{\mathbf{m}}$ which includes all correlation effects involving groups of, say, n or less particles, we simply put all $\beta^{\mathbf{m}}_{m_{l(1)} \cdots m_{l(p)}}$'s with $p > n$ equal to zero.

It is interesting to note that a cluster-decomposition law plays an important role even at this most elemental stage in the development of a many-body theory.

These formal manipulations, however, bring us no closer to a solution of the many-body problem since the C's, the basic ingredients of the γ's (or U's) and

therefore the basic ingredients of the β's, still remain to be determined, for example, by a perturbative or variational calculation. Obviously, direct determination of the β's would be highly desirable.[2-4]

Another form of correlated wavefunction that we consider is a generalization of the Jastrow wavefunction[12,13]:

$$\tilde{\Psi}_{\mathbf{m}}(x_1 \cdots x_N) = \mathcal{A}(N)[F_{\mathbf{m}}(x_1 \cdots x_N)\varphi_{m_1}(x_1) \cdots \varphi_{m_N}(x_N)], \quad (10)$$

where

$$F_{\mathbf{m}}(x_1 \cdots x_N) = \prod_{n=1}^{N} \prod_{\langle j(1) \cdots j(n) \mid 1 \cdots N \rangle} f^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}). \quad (11)$$

We adopt here and henceforth the notation

$$\langle j(1) \cdots j(k) \mid l(1) \cdots l(p) \rangle,$$

$k \leq p$, for a particular combination of k indices, these selected from the set of indices $\{l(1) \cdots l(p)\}$. The special case $\langle j(1) \cdots j(k) \mid 1 \cdots N \rangle$ will be abbreviated to $\langle j(1) \cdots j(k) \rangle$. In Eq. (11) the correlation function $f^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)})$—which one might hope to determine variationally—describes correlations of n particles which, in the independent-particle approximation, occupy single-particle states $m_{j(1)} \cdots m_{j(n)}$. [Observe that in the primitive function, i.e., the bracketed expression in (10), a 1–1 correspondence exists between the orbital and particle labels; application of the antisymmetrizer eradicates this correspondence and restores the indistinguishability of the particles.] If we choose all $N!/[(N-n)!\,n!]$ of the n-particle correlation functions to be the same symmetric function of their n coordinate arguments, (10) and (11) simplify to

$$\tilde{\Psi}_{\mathbf{m}}(x_1 \cdots x_N) = F_{\mathbf{m}}\Phi_{\mathbf{m}}(x_1 \cdots x_N), \quad (12)$$

where

$$\Phi_{\mathbf{m}}(x_1 \cdots x_N) = (N!)^{-\frac{1}{2}} \begin{vmatrix} \varphi_{m_1}(x_1) & \cdots & \varphi_{m_N}(x_1) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi_{m_1}(x_N) & \cdots & \varphi_{m_N}(x_N) \end{vmatrix}, \quad (13)$$

and

$$F_{\mathbf{m}} = \prod_{n=1}^{N} \prod_{\langle j(1) \cdots j(n) \rangle} f^{\mathbf{m}}_n(x_{j(1)} \cdots x_{j(n)}). \quad (14)$$

Upon setting $f^{\mathbf{m}}_n = 1$, $n \neq 2$, Eqs. (12)–(14) collapse to the definition of the well-known Jastrow wavefunction.

The product form (10) bears a superficial resemblance to the exact stationary-state wavefunction written in terms of the product form for $\mathcal{F}_{\mathbf{m}}$. However, it is really of quite different character, since, unlike $\prod_{\langle j(1) \cdots j(n) \rangle}(1 + \beta^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}})$, which produces only the effects of all specifically n-body correlations, $\prod_{\langle j(1) \cdots j(n) \rangle} f^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}$ incorporates not only irreducible n-body correlation effects but also correlation effects involving all larger groups of particles, these being incorporated in terms of *products* of n-body correlation functions. Thus, as in those cases of the expansion arising when (3) is inserted into (1) and the expansion (5) of Sinanoğlu, there is, in (10) or (12), no clean separation of correlation effects, but the mixing, in these latter forms for the wavefunction, is not intrinsically of such a nature as to vitiate their use in practical calculations and may even be beneficial.

Since we are thinking in terms of stationary states, we may regard $\Psi_{\mathbf{m}}$ of (1)–(3) or (5)–(7) as the exact wavefunction for our problem (or, with undetermined C's, as a trial wavefunction) and $\tilde{\Psi}_{\mathbf{m}}$ of (10) and (11), (12)–(14) as trial wavefunctions.

We remind the reader that any of the above wavefunctions may alternatively be viewed as the element $\mathbf{m}$ of a *basis of correlated functions* which, depending on how the $f^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(k)}}$'s, $f^{\mathbf{m}}_k$'s, or $C \cdots$'s are chosen, may provide a highly advantageous starting point for exact description of the N-fermion system, a far more appropriate starting point then the input independent-particle basis.[13-15]

Now let us turn to the task at hand. Working henceforth entirely in the configuration-space representation, our attention is centered on integrals like

$$(\Psi, S\Psi) \equiv \int \prod_{b=1}^{N} dx_b \Psi^*(x_1 \cdots x_N) S(1 \cdots N) \Psi(x_1 \cdots x_N) \quad (15)$$

and

$$(\Psi, \Psi) \equiv \int \prod_{b=1}^{N} dx_b \, |\Psi(x_1 \cdots x_N)|^2, \quad (16)$$

where $S(1 \cdots N)$ is a permutation-symmetric Hermitian operator and $\int \prod_{b=1}^{N} dx_b$ implies integrations over all continuous coordinates and summations over all discrete ones. [When written as an argument of an operator, i stands for $\mathbf{r}_i$, $-i\hbar\nabla_i$, and the spin and (if appropriate) isospin operators $\boldsymbol{\sigma}_i$, $\boldsymbol{\tau}_i$.] The methods to be developed for the evaluation of the ratio (15)/(16), the expectation value

$$\langle S \rangle = \frac{(\Psi, S\Psi)}{(\Psi, \Psi)}, \quad (17)$$

[12] R. Jastrow, Phys. Rev. **98**, 1479 (1955).
[13] J. W. Clark and P. Westhaus, Phys. Rev. **141**, 833 (1966).
[14] E. Feenberg and C. W. Woo, Phys. Rev. **137**, A391 (1965).
[15] See C. M. Shakin, Y. R. Waghmare, and M. H. Hull, Jr., Phys. Rev. **161**, 1006 (1967).

are sufficiently general that Ψ may be any of the forms we have considered, in fact, any N-particle wavefunction. We have dropped the **m** label, since we deal from now on with a particular state. (This allows a welcome simplification of the notation as regards single-particle state labels—we are able to write $j(1)\cdots j(n)$ in place of $m_{j(1)}\cdots m_{j(n)}$.)

In the absence of dynamical correlations (all β's except possibly the one-particle ones set zero, giving an independent-particle approximation), the evaluation of (17) is usually trivial. For example, if we take S to be the ordinary Hamiltonian, a symmetric sum of one-body operators plus a symmetric sum of two-body operators, the required expectation value reduces quickly to a sum of one-body integrals plus a sum of two-body integrals. But in general it is a practical impossibility to conclude the operations indicated in (17). Thus one is prompted to express (17) as a sum of one-body terms, plus a sum of two-body terms, . . . , plus a sum of N-body terms, *in such way* that truncation of the series after a manageable number of terms involving only calculable (few-body) integrals furnishes (hopefully) a useful approximation. Expansions of this type, called *cluster expansions*, first saw application in the classical statistical mechanics of imperfect gases, where they have long been employed to approximate the partition function.[16] Two cluster expansions, the one associated with Iwamoto and Yamada[17] (IY) and the other with Aviles[18] and Hartogh and Tolhoek[19] (AHT), have frequently been used in quantum-mechanical many-body calculations. In Sec. 2 the generalized normalization integral—a quantity from which the required expectation values may be extracted—is defined, and decomposed into cluster integrals according to schemes closely allied with those of IY and AHT. Then two radically different ("factor-cluster") decompositions, bearing the same relation to the Van Kampen[20] classical cluster development as that borne by the IY and AHT decompositions to the classical Ursell[16] development, are studied in detail. To conclude Sec. 2, general formulas for the expectation value of an operator in terms of the two new sets of cluster integrals are derived. An application of these formulas to the evaluation of the n-particle spatial distribution function is presented in Sec. 3. The final section is devoted to a general comparison of the four formalisms here studied, in the context of their practical application for the systems of interest.

In the second paper of this series we shift the emphasis from an investigation of the new factor-cluster formalisms *per se* to an exploitation of their properties and their relationships with the IY and AHT formalisms, with the aim of extending the applicability of these latter formalisms to finite N.

2. CLUSTER EXPANSIONS

A symmetric Hermitian operator $S(1\cdots N)$ may in general be resolved into a symmetric sum of one-body operators, plus a symmetric sum of two-body operators, . . . , plus a symmetric sum of N-body operators:

$$S(1\cdots N)=\sum_{i=1}^{N}S_i(1\cdots N) \tag{18}$$

with

$$S_p(1\cdots N)=\sum_{\langle j(1)\cdots j(p)\rangle}\bar{s}_p(j(1)\cdots j(p)), \tag{19}$$

where $\bar{s}_p(j(1)\cdots j(p))$ operates in the subspace of the particles labeled $j(1)\cdots j(p)$. There are in general many possible ways of resolving $S(1\cdots N)$. There will always be one particular resolution in which none of the $\bar{s}_p(j(1)\cdots j(p))$ may be further decomposed into the sum of terms which individually depend on a proper subset of particle labels $j(1),\cdots,j(p)$. Such a resolution will be said to be irreducible.

For example, the Hamiltonian of a system of N identical, nonrelativistic particles of mass M, located in an external field $V(i)$ and interacting via two-body potentials $v(ij)$,

$$H(1\cdots N)=\sum_{i=1}^{N}\frac{-\hbar^2}{2M}\nabla_i^2+\sum_{i=1}^{N}V(i)+\sum_{1\le i<j\le N}v(i,j), \tag{20}$$

may, as suggested above, be irreducibly resolved by taking

$$\begin{aligned}\bar{s}_1(i)&=\frac{-\hbar^2}{2M}\nabla_i^2+V(i),\\ \bar{s}_2(ij)&=v(ij),\\ \bar{s}_n(j(1)\cdots j(n))&=0,\quad 3\le n\le N.\end{aligned} \tag{21}$$

As a second illustration, consider the operator

$$\mathbf{P}^{(n)}(\mathbf{r}_1\cdots\mathbf{r}_n;\mathbf{r}_1'\cdots\mathbf{r}_n')$$

$$=\sum_{\langle j(1)\cdots j(n)\rangle}\left\{\sum_{\substack{\text{all permutations}\\ \text{of } l(1,\cdots l(n)\\ \text{over } 1\cdots n}}\prod_{i=1}^{n}\delta(\mathbf{r}_{l(i)}-\mathbf{r}'_{j(i)})\right\}, \tag{22}$$

whose expectation value yields the n-particle spatial

[16] H. D. Ursell, Proc. Cambridge Phil. Soc. **23**, 685 (1927).
[17] F. Iwamoto and M. Yamada, Progr. Theoret. Phys. (Kyoto) **17**, 543 (1957),
[18] J. B. Aviles, Jr., Ann. Phys. (N.Y.) **5**, 251 (1958).
[19] C. D. Hartogh and H. A. Tolhoek, Physica **24**, 721, 875, 896 (1958).
[20] G. N. Van Kampen, Physica **27**, 783 (1961).

distribution function

$$\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = \frac{N!}{(N-n)!} \times \frac{\sum_{\substack{\text{all spin} \\ \text{coordinates}}} \int \prod_{i=n+1}^{N} d\mathbf{r}_i \Psi^*(x_1 \cdots x_N)\Psi(x_1 \cdots x_N)}{\int \prod_{i=1}^{N} dx_i \Psi^*(x_1 \cdots x_N)\Psi(x_1 \cdots x_N)}. \quad (23)$$

In this case the obvious choice of $\bar{s}_p$'s is

$$\bar{s}_n^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n; \mathbf{r}'_{j(1)} \cdots \mathbf{r}'_{j(n)}) = \left\{ \sum_{\substack{\text{all permutations} \\ \text{of } l(1)\cdots l(n) \\ \text{over } 1\cdots n}} \prod_{i=1}^{n} \delta(\mathbf{r}_{l(i)} - \mathbf{r}'_{j(i)}) \right\}, \quad (24)$$

$$\bar{s}_p^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n; \mathbf{r}'_{j(1)} \cdots \mathbf{r}'_{j(p)}) = 0, \quad p \neq n.$$

Again the resolution we have chosen is irreducible.

Following the lead of Iwamoto and Yamada,[17] we define the (generalized) normalization integral

$$I_{\mathbf{m}}(\alpha) = \int \prod_{b=1}^{N} dx_b \Psi^*_{\mathbf{m}}(x_1 \cdots x_N) e^{\alpha S(1\cdots N)} \Psi_{\mathbf{m}}(x_1 \cdots x_N),$$

$$= \int \prod_{b=1}^{N} dx_b \Psi^*_{\mathbf{m}} \left[1 + \alpha S + \frac{\alpha^2}{2!} S^2 + \cdots \right] \Psi_{\mathbf{m}}, \quad (25)$$

from which the required expectation value may be obtained by means of the following differentiation:

$$\langle S \rangle = \frac{(d/d\alpha) I_{\mathbf{m}}(\alpha)}{I_{\mathbf{m}}(\alpha)} \bigg|_{\alpha=0} = \frac{d}{d\alpha} \ln I_{\mathbf{m}}(\alpha) \bigg|_{\alpha=0}. \quad (26)$$

Our goal is to calculate $I_{\mathbf{m}}(\alpha)$, or alternatively $\ln I_{\mathbf{m}}(\alpha)$, by the technique of cluster expansion. Although we are presently interested only in the expectation value of the operator $S(1 \cdots N)$, it might be pointed out that a knowledge of $I_{\mathbf{m}}(\alpha)$ yields a wealth of additional information. In particular, the variance of $S(1 \cdots N)$ may be computed as the second derivative of $\ln I_{\mathbf{m}}(\alpha)$ evaluated at $\alpha = 0$.[21] Finally, let us note that the basic idea embodied in (25) and (26) has recently been extended by Clark and Westhaus[13] to the evaluation of nondiagonal matrix elements in the representation defined by a set of dynamically correlated basis functions.

Our procedures for calculating $I_{\mathbf{m}}(\alpha)$ begin with the recognition that, given $\Psi_{\mathbf{m}}$ and $S(1 \cdots N)$, a set of related functions and operators may be defined for n-particle subspaces, $n \leq N$, of the N-particle Hilbert space. First, we consider a subset $m_{j(1)} \cdots m_{j(n)}$ of $\mathbf{m}$ and construct an n-particle "wavefunction." In particular, according as form (1) with (3) inserted, form (5), form (10), or form (12) has been chosen for $\Psi_{\mathbf{m}}$, we define the n-particle "wavefunction," respectively, as

$$\psi_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \ldots x_{j(n)}) = (n!)^{-\frac{1}{2}} \langle 0 | \psi(x_{j(n)}) \cdots \psi(x_{j(1)}) \left\{ \prod_{p=1}^{n} (1 + \beta_{m_{j(p)}}) \right\} \cdot \left\{ \prod_{p<q}^{n} (1 + \beta_{m_{j(p)} m_{j(q)}}) \right\} \cdots \{1 + \beta_{m_{j(1)} \cdots m_{j(n)}}\} \times a^\dagger_{m_{j(n)}} \cdots a^\dagger_{m_{j(1)}} |0\rangle, \quad (27a)$$

$$\psi_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}) = \mathcal{A}(n) \left[\prod_{i=1}^{n} \varphi_{m_{j(i)}}(x_{j(1)}) + \sum_{p=1}^{n} \prod_{i \neq p}^{n} \varphi_{m_{j(i)}}(x_{j(i)}) U^{\mathbf{m}}_{m_{j(p)}}(x_{j(p)}) + (2!)^{-\frac{1}{2}} \sum_{p<q}^{n} \prod_{i \neq p,q}^{n} \times \varphi_{m_{j(i)}}(x_{j(i)}) U^{\mathbf{m}}_{m_{j(p)} m_{j(q)}}(x_{j(p)} x_{j(q)}) + \cdots + (n!)^{-\frac{1}{2}} U^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}) \right], \quad (27b)$$

$$\psi_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}) = \mathcal{A}(n) \times \left[\prod_{p=1}^{n} \left\{ \prod_{\langle l(1) \cdots l(p) \mid j(1) \cdots j(n) \rangle} f^{\mathbf{m}}_{m_{l(1)} \cdots m_{l(p)}}(x_{l(1)} \cdots x_{l(p)}) \right\} \times \prod_{q=1}^{n} \varphi_{m_{j(q)}}(x_{j(q)}) \right], \quad (27c)$$

or

$$\psi_{m_{j(1)} \cdots m_{j(n)}}(x_{j(1)} \cdots x_{j(n)}) = \prod_{p=1}^{n} \left\{ \prod_{\langle l(1) \cdots l(p) \mid j(1) \cdots j(n) \rangle} f^{\mathbf{m}}_{p}(x_{l(1)} \cdots x_{l(p)}) \right\} \times (n!)^{-\frac{1}{2}} \begin{vmatrix} \varphi_{m_{j(1)}}(x_{j(1)}) & \cdots & \varphi_{m_{j(n)}}(x_{j(1)}) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi_{m_{j(1)}}(x_{j(n)}) & \cdots & \varphi_{m_{j(n)}}(x_{j(n)}) \end{vmatrix}. \quad (27d)$$

In the second and third definitions $\mathcal{A}(n)$ is the antisymmetrizer for the appropriate set of n coordinate labels, normalized such that $(\mathcal{A}(n))^2 = (n!)^{\frac{1}{2}} \mathcal{A}(n)$. Although we have written the $\psi_{m_{j(1)} \cdots m_{j(n)}}$ as functions of the coordinates $x_{j(1)}, \cdots, x_{j(n)}$, they can, of course, be written in terms of any set of n coordinates, i.e., there is no necessary correspondence between the place indices of the selected orbitals $m_{j(1)}$; $m_{j(2)}, \cdots, m_{j(n)}$ and the coordinate labels [see Eq. (29)]. Henceforth, the orbital indices $m_{j(1)}, \cdots, m_{j(n)}$ will, for brevity, be written simply as $j(1), \cdots, j(n)$; moreover, the same symbol $\psi_{j(1) \ldots j(n)}$ will denote, as appropriate, any of the four forms of the n-particle

[21] R. Norgard (private communication).

"wavefunction." Next, from the elements into which $S(1\cdots N)$ is resolved via (18) and (19), we construct the n-particle operator

$$S(1\cdots n)=\sum_{p=1}^{n}\sum_{\langle l(1)\cdots l(p)\,|\,1\cdots n\rangle}\bar{s}_p(l(1)\cdots l(p)). \quad (28)$$

Then we define in terms of a particular n-particle wavefunction and n-particle operator a subnormalization integral indexed with the same orbital labels as designate the corresponding n-particle wavefunction:

$$I_{j(1)\cdots j(n)}(\alpha)=\int\prod_{b=1}^{n}dx_b\psi^{*}_{j(1)\cdots j(n)}(x_1\cdots x_n)\times e^{\alpha S(1\cdots n)}\psi_{j(1)\cdots j(n)}(x_1\cdots x_n). \quad (29)$$

Counting $I_{1\cdots N}(\alpha)=I_{\mathrm{m}}(\alpha)\equiv I(\alpha)$ as a particular subnormalization integral, we see there are 2^N-1 such quantities, each indexed by a particular subset of orbital labels, $j(1)\cdots j(n)$.

Before proceeding to general techniques for the evaluation of $I(\alpha)$ let us imagine for a moment that all correlations among the particles have ceased to exist (e.g., all β operators vanish) and that all except the one-body terms in the operator $S(1\cdots N)$ are zero. Under these conditions we have

$$\begin{aligned}I_{j(1)\cdots j(n)}(\alpha)&\to\tilde{I}_{j(1)\cdots j(n)}(\alpha)\\&=(n!)^{-1}\int\prod_{b=1}^{n}dx_b\begin{vmatrix}\varphi_{j(1)}(x_1)&\cdots&\varphi_{j(n)}(x_1)\\ \cdot&\cdots&\cdot\\ \cdot&\cdots&\cdot\\ \cdot&\cdots&\cdot\\ \varphi_{j(1)}(x_n)&\cdots&\varphi_{j(n)}(x_n)\end{vmatrix}^{*}\\&\quad\times\exp\left\{\alpha\sum_{i=1}^{n}\bar{s}_1(i)\right\}\begin{vmatrix}\varphi_{j(1)}(x_1)&\cdots&\varphi_{j(n)}(x_1)\\ \cdot&\cdots&\cdot\\ \cdot&\cdots&\cdot\\ \cdot&\cdots&\cdot\\ \varphi_{j(1)}(x_n)&\cdots&\varphi_{j(n)}(x_n)\end{vmatrix}\\&=\prod_{i=1}^{n}\left[\int dx_1\varphi^{*}_{j(i)}(x_1)e^{\alpha\bar{s}_1(1)}\varphi_{j(i)}(x_1)\right]+O(\alpha_2)\\&=\prod_{i=1}^{n}\tilde{I}_{j(i)}(\alpha)+O(\alpha^2). \quad (30)\end{aligned}$$

In particular, if the operator $\bar{s}_1(1)=h_0(1)$ is a one-body Hamiltonian containing some *ad hoc* single-particle potential and $\{\varphi_1,\varphi_2,\cdots,\varphi_N\}$ a set of eigenfunctions of h_0, then the equalities in (30) hold with the addends $O(\alpha^2)$ omitted. Frequently, such an independent-particle model provides both a convenient image of the physical system and a framework in which those observables corresponding to single-particle operators may be accurately calculated. In any event, as long as we are only interested in computing $\langle S\rangle$, terms $O(\alpha^2)$ need never be considered.

We are now prepared to set forth the essential structure and state the underlying philosophy of a wide class of cluster formalisms applicable to the evaluation of $I(\alpha)$. Each of the above subnormalization integrals $I_{j(1)\cdots j(n)}$ or, more generally, some linear combination of those with the given number n of indices, is to be built up from the $\tilde{I}_{j(i)}$'s and a (finite) number of *cluster integrals* involving anywhere from one- to n-fold integrations. Alternatively, we say that the subnormalization integral, or the corresponding linear combination of subnormalization integrals, is decomposed into the $\tilde{I}_{j(i)}$'s and the cluster integrals. With the $\tilde{I}_{j(i)}$'s and the $I_{j(1)\cdots j(n)}$'s already defined by (30) and (29), respectively, such a decomposition equation provides the definition of the last-appearing n-body, cluster integral. Proceeding to $n=N$, an expansion for $I(\alpha)$ in terms of the cluster integrals is achieved. We see that for some decompositions (among these the most familiar) it is necessary to rearrange this *primitive cluster expansion* for $I(\alpha)$ in order to obtain a useful approximation scheme for large N. All of the primitive cluster expansions to be considered will contain a finite number of addends for finite N.

Different modes of decomposition of the basic subnormalization integrals, or appropriate linear combinations of them, lead to different definitions of the cluster integrals. But an important feature of general cluster theory as herein circumscribed will pervade all the following considerations for finite N: no matter what mode of decomposition is chosen, the explicit elimination of all the cluster integrals from the decomposition equation for, say, $I_{j(1)\cdots j(n)}$ must lead to an identity. In particular, it is the definition of the *last* cluster integral which ensures that the (primitive) cluster expansion for $I(\alpha)$, if completely summed, will regenerate this normalization integral. This feature is trivially obvious, but, as we shall see, formally useful.

Later discussions will be facilitated if we adopt a definite convention for what we mean by a *term* contributing to a cluster expansion of some quantity. Any *individual contribution* to a given cluster expansion will always be a product of the corresponding cluster integrals (perhaps with only one such factor), supplemented by some numerical factor. These cluster integrals may or may not be indexed with single-particle labels. In the former case, the *term* corresponding to the contribution in question will consist of this contribution summed over all combinations of the allowed single-particle labels, the original labels having been replaced by dummies. In the latter case, the *term* corresponding to the contribution is that

contribution multiplied by the number of times it occurs in the expansion. In both cases, if the contribution enters with a minus sign, that also is to be attached. The given cluster expansion is, of course, the sum of all such *unique* terms. The meaning of this convention will become clearer as we develop concrete examples of cluster expansions.

In the above, we have not specified any essential properties that the cluster integrals must possess, but have merely regarded them as the elements of the postulated decomposition—which of course determines their properties completely. It is not our immediate concern whether or not a given n-body cluster integral (involving integrations over n sets x_b) so determined is irreducible in the sense that this integral can be in turn decomposed into a sum of products of independent integrals over fewer than n sets of coordinates. Of course, a careful study of the structure of the cluster integrals, in particular with respect to their possible reducibility, is vital in analyzing the behavior of the terms in a cluster expansion as N grows large.[13,14,22,23]

To conclude these very general remarks: Although, strictly speaking, every cluster expansion is a tautology, the value of a given cluster formalism, when employed in its main role of generating approximations for $\langle S\rangle$, depends on the wisdom of our decomposition. We always have in mind systems for which irreducible correlations involving many particles are less important than those involving few. Thus we always seek a cluster expansion for $\langle S\rangle$ which, perhaps after suitable rearrangement, will reflect this situation in rapid convergence, allowing us to approximate $\langle S\rangle$ satisfactorily in terms of the first few cluster integrals, therefore in terms of the first few subnormalization integrals. Naturally the usefulness of a given cluster expansion of the expectation value of the observable S will depend crucially on the correlated wavefunction chosen; in the final analysis the merit of the approximation schemes suggested by this paper must be tested by detailed numerical calculations.

We now study in detail four modes of defining cluster integrals. The first is a straightforward extension of the procedure adopted by Iwamoto and Yamada.[17] Initially we approximate $I_{j(1)\cdots j(n)}(\alpha)$ by the product $\prod_{i=1}^{n}\tilde{I}_{j(i)}(\alpha)$ and then "build up" the subnormalization integral by replacing each combination of zero, one, two, ... n factors in $\prod_{i=1}^{n}\tilde{I}_{j(i)}(\alpha)$ with the sum of all possible products of one-body, two-body, ..., n-body cluster integrals involving, with no repetitions, the same set of indices as the replaced factors. Beginning with the one-indexed quantities, a hierarchy of equations is generated:

$$
\begin{aligned}
I_{j(1)} &= \tilde{I}_{j(1)} + X_{j(1)},\\
I_{j(1)j(2)} &= \tilde{I}_{j(1)}\tilde{I}_{j(2)} + X_{j(1)j(2)} + \tilde{I}_{j(1)}X_{j(2)}\\
&\quad + X_{j(1)}X_{j(2)} + X_{j(1)j(2)},\\
I_{j(1)j(2)j(3)} &= \prod_{i=1}^{3}\tilde{I}_{j(i)}\\
&\quad + \sum_{i=1}^{3}\left[X_{j(i)}\cdot\prod_{\substack{p=1\\p\neq i}}^{3}\tilde{I}_{j(p)} + \tilde{I}_{j(i)}\cdot\prod_{\substack{p=1\\p\neq i}}^{3}X_{j(p)}\right]\\
&\quad + \sum_{i=1}^{3}\left[\sum_{\substack{p<q\\p,q\neq i}}(\tilde{I}_{j(i)} + X_{j(i)})X_{j(p)j(q)}\right]\\
&\quad + \prod_{i=1}^{3}X_{j(i)} + X_{j(1)j(2)j(3)},\\
&\ \ \vdots\\
I_{12\cdots N} &= \prod_{i=1}^{N}\tilde{I}_i + \sum_{i=1}^{N}X_i\prod_{p\neq i}^{N}\tilde{I}_p + \cdots + X_{12\cdots N}.
\end{aligned}
\tag{31}
$$

Introducing the normalized cluster integrals

$$
x_{j(1)\cdots j(n)} = X_{j(1)\cdots j(n)}\Big/\prod_{i=1}^{n}\tilde{I}_{j(i)}, \tag{32}
$$

we can rewrite these decomposition equations as

$$
\begin{aligned}
I_{j(1)} &= \tilde{I}_{j(1)}\cdot[1 + x_{j(1)}],\\
I_{j(1)j(2)} &= \tilde{I}_{j(1)}\cdot\tilde{I}_{j(2)}\\
&\quad\times[1 + x_{j(1)} + x_{j(2)} + x_{j(1)}x_{j(2)}\\
&\qquad + x_{j(1)j(2)}],\\
I_{j(1)j(2)j(3)} &= \prod_{i=1}^{3}\tilde{I}_{j(i)}\\
&\quad\times\left[1 + \sum_{p=1}^{3}x_{j(p)} + \sum_{1\le p<q\le 3}(x_{j(p)}\cdot x_{j(q)} + x_{j(p)j(q)})\right.\\
&\quad + x_{j(1)}x_{j(2)}x_{j(3)} + x_{j(1)}\cdot x_{j(2)j(3)} + x_{j(2)}x_{j(1)j(3)}\\
&\quad\left. + x_{j(3)}x_{j(1)j(2)} + x_{j(1)j(2)j(3)}\right],\\
&\ \ \vdots\\
I_{12\cdots N} &= I = \prod_{i=1}^{N}\tilde{I}_i\\
&\quad\times\left[1 + \sum_{p}x_p + \sum_{p<q}(x_p\cdot x_q + x_{pq}) + \cdots x_{12\cdots N}\right],
\end{aligned}
\tag{33}
$$

where, in general,

$$
I_{j(1)\cdots j(n)} = \prod_{i=1}^{n}\tilde{I}_{j(i)}B_{j(1)\cdots j(n)} \tag{34}
$$

with

$$
\begin{aligned}
B_{j(1)\cdots j(n)} &= \left[1 + \sum_{p=1}^{n}\sum_{\langle l(1)\cdots l(p)\,|\,j(1)\cdots j(n)\rangle}\right.\\
&\quad\left.\times\left\{\sum_{q=1}^{p}\sum_{\substack{\text{all partitions}\\ \text{of } l(1)\cdots l(p)\\ \text{among } p-q+1\\ \text{factors}}}x\ldots x\ldots\cdots x\ldots\right\}\right].
\end{aligned}
\tag{35}
$$

[22] J. W. Clark, D. Chakkalakal, and P. Westhaus, Progr. Theoret. Phys. (Kyoto) **34**, 726 (1965).

[23] C. W. Woo, Phys. Rev. **151**, 138 (1966).

To the right of 1 in each $B_{j(1)\cdots j(n)}$, there stands the sum of all normalized cluster integrals with indices comprising a subset of $\{j(1)\cdots j(n)\}$, plus the sum of all possible *completely unlinked* products of these normalized cluster integrals. A *completely unlinked product* is one in which *no* pair of factors has an index in common. More generally, an *unlinked product* of cluster integrals is one in which at least one factor has no index in common with any other factor. On the other hand if each factor has at least one index in common with one or more other factors, the product will be called *linked.*

Now, although termination of the series (35) for $B_{12\cdots N}$ at one-index, two-index, ..., n-index, ... terms may lead to acceptable approximations for $B_{12\cdots N}$ when N is small, the nonuniform asymptotic N dependence of successive truncations indicates that, when N is large, some other form for $B_{12\cdots N}$ must be found for approximating $I(\alpha)$. This nonuniform N dependence results from the presence of the completely unlinked products.[13,17,22] Dropping terms which are of no consequence in the many-body limit (i.e., the limit in which $N\to\infty$ while ρ, the number density, is held constant), Iwamoto and Yamada[17] and Wu and Feenberg[24] have derived an expression for $B_{12\cdots N}$ in which all products of x's which appear are linked. Thus, upon writing

$$B_{12\cdots N} = \exp G_{\mathrm{IY}}, \tag{36}$$

they find in the many-body limit that

$$G_{\mathrm{IY}} = \sum_{i=1}^{N} x_i + \sum_{i<j}^{N} x_{ij} + \sum_{i<j<k}^{N} \times [x_{ijk} - x_{ij}x_{jk} - x_{ij}x_{ik} - x_{ik}x_{jk}] + \cdots, \tag{37}$$

an expansion which, upon successive truncations, displays a uniform asymptotic N dependence. In the second paper of this series, we offer another method of rearranging $B_{12\cdots N}$ to generate such an *exponential formula*, a method which demonstrates, and utilizes, the (quite general) linked character of G_{IY}. Our procedure, unlike former ones, is applicable to finite as well as infinite systems.

By the *order* of a given term in (37) or in any cluster expansion based on the IY subnormalization integrals (29), we mean the number of distinct single-particle indices involved in any individual contribution to that term. This is a classification according to "number of bodies"—there will be one-body terms, two-body terms, three-body terms, ..., or, speaking more loosely, one-body clusters, two-body clusters, three-body clusters, ...—and is quite different from the ordering prescription of Clark and Westhaus[13] based on a "smallness parameter." In Eq. (37) the successive addends

$$\textstyle\sum_i x_i,\ \sum_{i<j} x_{ij},\ -\sum_{i<j<k} x_{ij}x_{jk},\ \cdots$$

are, respectively, first-order, second-order, third-order, ... terms in the sense just defined.

In order to motivate the introduction of a well-known alternative to the cluster formalism just discussed, let us suppose that a correlated wave-function (12) with state-independent correlation factors has been selected. Now consider a unitary transformation in the one-body vector space spanned by the set of single-particle functions $\varphi_1, \cdots, \varphi_N$:

$$\varphi_i(x) \to \varphi_i'(x) = \sum_j U_{ij}\varphi_j(x), \quad i,j = 1, \cdots, N, \tag{38}$$

where

$$\sum_k U_{ki}^* U_{kj} = \sum_k U_{ik}U_{jk}^* = \delta_{ij}.$$

Since the replacement $\varphi_i \to \varphi_i'$, $i = 1, \cdots, N$, changes Φ by at most a phase factor, the expectation value $\langle S\rangle$ is invariant under this replacement. On the other hand, as one may readily verify, the IY expansion for $\langle S\rangle$ does not possess term-by-term nor indeed order-by-order invariance under the substitution (38). Of course, if summed to all orders, the expression $\prod_{i=1}^{N} I_i e^{G_{\mathrm{IY}}}$ must yield $I(\alpha)$ identically and thus lead to expectation values with the above invariance. But practical application of this cluster expansion theory demands that G_{IY} be terminated in low orders, whence the approximate expectation values lack a very desirable feature of the exact expectation value.[13,14]

A cluster expansion which is invariant term by term under such a unitary transformation of the orbitals has been investigated by Aviles[18] and by Hartogh and Tolhoek.[19] The quantities which are decomposed into cluster integrals are certain linear combinations, or averages, of the subnormalization integrals defined in (29):

$$\begin{aligned}
\mathfrak{I}_1 &= \frac{1}{N}\sum_{i=1}^{N} I_i, \\
\mathfrak{I}_2 &= \frac{2}{N(N-1)}\sum_{i<j} I_{ij}, \\
&\vdots \\
\mathfrak{I}_n &= \frac{n!\,(N-n)!}{N!}\sum_{\langle j(1)\cdots j(n)\rangle} I_{j(1)\cdots j(n)}, \\
&\vdots \\
\mathfrak{I}_N &= I_{12\cdots N} = I(\alpha).
\end{aligned} \tag{39}$$

These linear combinations are themselves invariant under the transformation (38) and so in turn will be the cluster integrals which they define. Having

[24] F. Wu and E. Feenberg, Phys. Rev. **28**, 943 (1962).

motivated the introduction of these new subnormalization integrals we may drop our restriction to the wavefunction (12). Upon so doing the above statements concerning invariance may no longer apply; we emphasize this fact by putting the word *invariant* within quotation marks from now on.

We proceed in analogy with the foregoing development, again first considering the situation in which no dynamical correlations exist among the particles and in which we need only be concerned with the one-body component of $S(1\cdots n)$. Here, in complete correspondence with (30), we have

$$\mathfrak{I}_n \to \tilde{\mathfrak{I}}_n = \frac{n!(N-n)!}{N!} \sum_{\langle j(1)\cdots j(n)\rangle} \frac{1}{n!} \int \prod_{b=1}^{n} dx_b \times \begin{vmatrix} \varphi_{j(1)}(x_1) & \cdots & \varphi_{j(n)}(x_1) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi_{j(1)}(x_n) & \cdots & \varphi_{j(n)}(x_n) \end{vmatrix}^* \times \exp\left[\alpha \sum_{i=1}^{n} \tilde{s}_1(i)\right] \times \begin{vmatrix} \varphi_{j(1)}(x_1) & \cdots & \varphi_{j(n)}(x_1) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi_{j(1)}(x_n) & \cdots & \varphi_{j(n)}(x_n) \end{vmatrix} = (\tilde{\mathfrak{J}}_1)^n + O(\alpha^2), \tag{40}$$

where

$$\tilde{\mathfrak{J}}_1 = \frac{1}{N} \sum_{i=1}^{N} \int dx_1 \varphi_i^*(x_1) \exp[\alpha \tilde{s}_1(1)] \varphi_i(x_1) = \frac{1}{N} \sum_{i=1}^{N} \tilde{I}_i. \tag{41}$$

Returning to the realistic situation, we "build up" $\mathfrak{I}_n$ by replacing successive numbers of $\tilde{\mathfrak{J}}_1$ factors in $(\tilde{\mathfrak{J}}_1)^n$ with appropriate sums of products of cluster integrals. This process may be viewed alternatively as a decomposition of $\mathfrak{I}_n$ into cluster integrals. From either point of view, of course, with $\tilde{\mathfrak{J}}_1$, $\mathfrak{I}_n$, and cluster integrals $\mathfrak{X}_p$ $(p < n)$ having already been defined, it is the cluster integral $\mathfrak{X}_n$ which is defined by this decomposition of $\mathfrak{I}_n$. Beginning the process with $\mathfrak{I}_1$, we write

$$\begin{aligned} \mathfrak{I}_1 &= \tilde{\mathfrak{J}}_1 + \mathfrak{X}_1, \\ \mathfrak{I}_2 &= \tilde{\mathfrak{J}}_1^2 + 2\tilde{\mathfrak{J}}_1\mathfrak{X}_1 + \mathfrak{X}_1^2 + \mathfrak{X}_2, \\ \mathfrak{I}_3 &= \tilde{\mathfrak{J}}_1^3 + 3\tilde{\mathfrak{J}}_1^2\mathfrak{X}_1 + 3\tilde{\mathfrak{J}}_1^2\mathfrak{X}_1 + 3\tilde{\mathfrak{J}}_1\mathfrak{X}_2 + \mathfrak{X}_1^3 \\ &\quad + 3\mathfrak{X}_2\mathfrak{X}_1 + \mathfrak{X}_3, \\ &\vdots \\ \mathfrak{I}_N &= \tilde{\mathfrak{J}}_1^N + N\tilde{\mathfrak{J}}_1^{N-1}\mathfrak{X}_1 + \cdots + \mathfrak{X}_N. \end{aligned} \tag{42}$$

Then, continuing the development in analogy with the decomposition of $I_{j(1)\cdots j(n)}$, we define

$$\mathfrak{X}_n = \mathfrak{X}_n/(\tilde{\mathfrak{J}}_1)^n \tag{43}$$

and rewrite the general equation in the above set as

$$\tilde{\mathfrak{J}}_n = (\tilde{\mathfrak{J}}_1)^n \mathfrak{B}_n \tag{44}$$

with

$$\mathfrak{B}_n = \left\{ 1 + \sum_{p=1}^{n} {}_pC_n \sum_{\substack{\text{all partitions} \\ \text{such that} \\ \sum_{b=1}^{p} b\nu_b = p}} \frac{p!\,\mathfrak{X}_1^{\nu_1}\cdots\mathfrak{X}_b^{\nu_b}\cdots\mathfrak{X}_p^{\nu_p}}{\prod_{b=1}^{p}(b!)^{\nu_b}\nu_b!} \right\} \tag{45}$$

and ${}_pC_n = n!/[(n-p)!\,p!]$.

Once more hoping to approximate $\mathfrak{B}_N$ in terms of the few-body cluster integrals, say $\mathfrak{X}_1$, $\mathfrak{X}_2$, $\mathfrak{X}_3$, and $\mathfrak{X}_4$, we discover that in the case of large N the series (45) must first be rearranged in order to express $\mathfrak{B}_N$ in terms of a series which, for large N, is uniform in N; and again the required rearrangement results in an exponential formula

$$\mathfrak{B}_N = \exp \mathfrak{S}_{\text{AHT}}, \tag{46}$$

with $\mathfrak{S}_{\text{AHT}}$ given, as the many-body limit is approached, by the "uniform" cluster expansion

$$\mathfrak{S}_{\text{AHT}} = N\mathfrak{X}_1 + \frac{N^2}{2!}\mathfrak{X}_2 + \frac{N^3}{3!}(\mathfrak{X}_3 - 3\mathfrak{X}_2^2) + \cdots. \tag{47}$$

A derivation of (46), together with a generalization of (47) which includes those terms which must also be accounted for when N is not large, will be given in the second paper of this series, where the properties of the "factor-cluster" expansions introduced below are more fully exploited.

By the *order* of a given term in (47) or in any cluster expansion based on the AHT subnormalization integrals (39) we shall mean the following: in the case of the asymptotically leading terms, simply the power of N that appears explicitly; in the case of terms down from these by $O(1/N)$, the minimum number of sets of particle coordinates which must be introduced to carry out the integrations involved. This definition is not as arbitrary as it appears, since it will be seen to maintain a close correspondence with the "number-of-bodies" ordering prescription set up for expansions based on the IY subnormalization integrals. In (47) the successive addends $N\mathfrak{X}_1$, $\frac{1}{2}N^2\mathfrak{X}_2$, $-\frac{1}{2}N^3\mathfrak{X}_2^2, \cdots$ are, respectively, first-order, second-order, third-order, ... terms in the sense just defined, while the asymptotically negligible term $\frac{1}{2}N\mathfrak{X}_1^2$, to be incorporated later, is of second order.

The necessity of rearranging the series for $B_{12\cdots N}$ and $\mathfrak{B}_N$ has prompted us to seek new and "more

natural" modes of decomposition to replace those of IY and AHT, modes which require no rearrangement to make sense for large N. The two alternative approaches presented here are closely related to a cluster expansion proposed by Van Kampen[20] in the imperfect-gas problem. They have the virtue of immediate applicability to finite as well as infinite many-body problems. The first, a "noninvariant" formalism, is based on the "noninvariant" subnormalization integrals of IY; the second, an "invariant" formalism, on the "invariant" subnormalization integrals of AHT.

Instead of constructing the $I_{j(1)\cdots j(n)}$ by replacing the $\tilde{I}_{j(i)}$'s in the product $\prod_{i=1}^{n}\tilde{I}_{j(i)}$ as we originally did in (31), we now propose that additional factors—the new cluster integrals—simply be attached to this product according to the following prescription:

$$
\begin{aligned}
I_{j(1)} &= \tilde{I}_{j(1)}Y_{j(1)},\\
I_{j(1)j(2)} &= \tilde{I}_{j(1)}\cdot\tilde{I}_{j(2)}\cdot Y_{j(1)}\cdot Y_{j(2)}\cdot Y_{j(1)j(2)},\\
I_{j(1)j(2)j(3)} &= \left[\prod_{i=1}^{3}\tilde{I}_{j(i)}\right]\cdot\left[\prod_{l=1}^{3}Y_{j(l)}\right]\\
&\quad\cdot\left[\prod_{1\le p<q\le 3}Y_{j(p)j(q)}\right]\cdot Y_{j(1)j(2)j(3)},\\
&\;\;\vdots\\
I_{12\cdots N} &= \left[\prod_{i=1}^{N}\tilde{I}_{i}\right]\cdot\left[\prod_{l=1}^{N}Y_{l}\right]\\
&\quad\cdot\left[\prod_{1\le p<q\le N}Y_{pq}\right]\cdots Y_{12\cdots N}. \qquad (48)
\end{aligned}
$$

Employing (26) together with the final equation of this set, we arrive at an extremely simple formula for the expectation value:

$$
\langle S\rangle = \sum_{i=1}^{N}\frac{\dfrac{d}{d\alpha}\tilde{I}_i}{\tilde{I}_i}\bigg|_{\alpha=0} + \sum_{n=1}^{N}\sum_{\langle j(1)\cdots j(n)\rangle}\frac{(d/d\alpha)Y_{j(1)\cdots j(n)}}{Y_{j(1)\cdots j(n)}}\bigg|_{\alpha=0}. \qquad (49)
$$

Our general statements concerning c'uster theories may be applied to this new mode of decomposing $I(\alpha)$. Specifically, the last cluster integral in (48) is defined in terms of the generalized normalization integral itself. But, of course, the hope is to approximate $\langle S\rangle$ by terminating the series in (49) after the first few sums, and thus avoid having to compute the many-indexed Y's. The rapidity with which (49) converges depends ultimately upon the problem at hand; but that the series does depend uniformly on N in the many-body limit will become clear in the following paper where the relationship between this cluster expansion and that of IY is explored.

Implicit definitions of the quotients appearing in (49) may be obtained immediately upon differentiating the logarithm of each equation in the set (48) with respect to α. Defining

$$
\begin{aligned}
\tilde{J}_{j(i)} &= \frac{(d/d\alpha)\tilde{I}_{j(i)}}{\tilde{I}_{j(i)}}\bigg|_{\alpha=0},\\
J_{j(1)\cdots j(n)} &= \frac{(d/d\alpha)I_{j(1)\cdots j(n)}}{I_{j(1)\cdots j(n)}}\bigg|_{\alpha=0}, \qquad (50)
\end{aligned}
$$

and

$$
Z_{j(1)\cdots j(n)} = \frac{(d/d\alpha)Y_{j(1)\cdots j(n)}}{Y_{j(1)\cdots j(n)}}\bigg|_{\alpha=0}, \qquad (51)
$$

we have

$$
J_{j(1)\cdots j(n)} = \sum_{i=1}^{n}\tilde{J}_{j(i)} + \sum_{q=1}^{n}\sum_{\langle l(1)\cdots l(q)\,|\,j(1)\cdots j(n)\rangle} \times Z_{l(1)\cdots l(q)}, \quad n = 1,\cdots,N. \qquad (52)
$$

The key to the inversion of this set of equations is the realization that the substitution for the $Z_{l(1)\cdots l(q)}$'s of their definitions in terms of the basic quantities $\tilde{J}_{l(i)}$ and $J_{l(1)\cdots l(p)}$ must transform these equations into identities. Hence we conclude that

$$
\begin{aligned}
Z_{j(i)} &= J_{j(i)} - \tilde{J}_{j(i)},\\
Z_{j(1)\cdots j(n)} &= \sum_{q=1}^{n}(-1)^{n-q}\\
&\quad\times\left\{\sum_{\langle l(1)\cdots l(q)\,|\,j(1)\cdots j(n)\rangle}J_{l(1)\cdots l(q)}\right\}, \quad n\ge 2. \qquad (53)
\end{aligned}
$$

For upon the insertion of these relations into

$$
J_{k(1)\cdots k(f)} = \sum_{i=1}^{f}\tilde{J}_{k(i)} + \sum_{p=1}^{f}\sum_{\langle h(1)\cdots h(p)\,|\,k(1)\cdots k(f)\rangle} Z_{h(1)\cdots h(p)},
$$

the coefficient of each $\tilde{J}_{k(i)}$ vanishes, while the coefficient of $J_{l(1)\cdots l(q)}\{l(1)\cdots l(q)\}\subset\{k(1)\cdots k(f)\}$, is given by

$$
\sum_{s=q}^{f}(-1)^{s-q}{}_{s-q}C_{f-q} = \sum_{t=0}^{f-q}(-1)^{t}{}_{t}C_{f-q} = \begin{cases}0, & q<f,\\ 1, & q=f,\end{cases}
$$

as required. Note that the $\tilde{J}_{j(i)}$ enter only into the one-indexed cluster integrals. The results of this new "factor-cluster" formalism may be succinctly stated as follows:

$$
\langle S\rangle = \sum_{i=1}^{N}\tilde{J}_i + \sum_{n=1}^{N}\sum_{\langle j(1)\cdots j(n)\rangle}Z_{j(1)\cdots j(n)}, \qquad (49)
$$

where

$$
\begin{aligned}
Z_{j(i)} &= J_{j(i)} - \tilde{J}_{j(i)},\\
Z_{j(1)\cdots j(n)} &= \sum_{q=1}^{n}(-1)^{n-q}\left\{\sum_{\langle l(1)\cdots l(q)\,|\,j(1)\cdots j(n)\rangle}J_{l(1)\cdots l(q)}\right\},\\
&\qquad n\ge 2. \qquad (53)
\end{aligned}
$$

The quotients

$$J_{j(i)} = \frac{\dfrac{d}{d\alpha}\tilde{I}_{j(i)}}{\tilde{I}_{j(i)}}\bigg|_{\alpha=0} = \frac{\int dx_1 \varphi^*_{j(i)}(x_1)\bar{s}_1(1)\varphi_{j(i)}(x_1)}{\int dx_1 \varphi^*_{j(i)}(x_1)\varphi_{j(i)}(x_1)}$$

and

$$J_{l(1)\cdots l(q)} = \frac{\dfrac{d}{d\alpha} I_{l(1)\cdots l(q)}}{I_{l(1)\cdots l(q)}}\bigg|_{\alpha=0}$$

$$= \frac{\int \prod_{b=1}^{q} dx_b \psi^*_{l(1)\cdots l(q)}(x_1 \cdots x_q) \times S(1\cdots q)\psi_{l(1)\cdots l(q)}(x_1\cdots x_q)}{\int \prod_{b=1}^{q} dx_b \psi^*_{l(1)\cdots l(q)}(x_1\cdots x_q)\psi_{l(1)\cdots l(q)}(x_1\cdots x_q)}$$

are not susceptible to further reductions and must therefore be evaluated as the basic ingredients of the expansion.

Let us now consider a factor decomposition alternative to (42). Again $\mathfrak{I}_n$ is "built up" from $(\tilde{\mathfrak{I}}_1)^n$ by attaching cluster integrals $\mathfrak{Y}_p$ $(p \leq n)$—and thus defining $\mathfrak{Y}_n$—in the following manner:

$$\begin{aligned}
\mathfrak{I}_1 &= \tilde{\mathfrak{I}}_1 \cdot \mathfrak{Y}_1, \\
\mathfrak{I}_2 &= \tilde{\mathfrak{I}}_1^2 \cdot \mathfrak{Y}_1^2 \cdot \mathfrak{Y}_2, \\
\mathfrak{I}_3 &= \tilde{\mathfrak{I}}_1^3 \cdot \mathfrak{Y}_1^3 \cdot \mathfrak{Y}_2^3 \cdot \mathfrak{Y}_3, \\
&\cdots \\
\mathfrak{I}_n &= (\tilde{\mathfrak{I}}_1)^n \cdot \left[\prod_{k=1}^{n} (\mathfrak{Y}_k)^{{}_kC_n}\right], \qquad (54) \\
&\cdots \\
\mathfrak{I}_N &= I = (\tilde{\mathfrak{I}}_1)^N \cdot \left[\prod_{k=1}^{N} (\mathfrak{Y}_k)^{{}_kC_N}\right].
\end{aligned}$$

Formulas completely analogous to (48) through (53) emerge. Indeed, we find that

$$\mathfrak{Z}_n = n\tilde{\mathfrak{Z}}_1 + \sum_{b=1}^{n} {}_kC_n \mathfrak{Z}_k \qquad (55)$$

and, in particular,

$$\langle S\rangle = \mathfrak{Z}_N = N\tilde{\mathfrak{Z}}_1 + \sum_{k=1}^{N} {}_kC_N \mathfrak{Z}_k, \qquad (56)$$

where we have defined

$$\begin{aligned}
\tilde{\mathfrak{Z}}_1 &\equiv \frac{(d/d\alpha)\tilde{\mathfrak{I}}_1}{\tilde{\mathfrak{I}}_1}\bigg|_{\alpha=0} \\
\mathfrak{Z}_n &\equiv \frac{(d/d\alpha)\mathfrak{I}_n}{\mathfrak{I}_n}\bigg|_{\alpha=0}, \qquad (57)
\end{aligned}$$

and

$$\mathfrak{Z}_n \equiv \frac{(d/d\alpha)\mathfrak{Y}_n}{\mathfrak{Y}_n}\bigg|_{\alpha=0}. \qquad (58)$$

The inversion of (55) parallels that of (52). Here the required identity is ensured by taking

$$\begin{aligned}
\mathfrak{Z}_1 &= \mathfrak{Z}_1 - \tilde{\mathfrak{Z}}_1, \\
\mathfrak{Z}_n &= \sum_{q=1}^{n} (-1)^{n-q} \frac{n!}{q!\,(n-q)!} \mathfrak{Z}_q, \quad n \geq 2.
\end{aligned} \qquad (59)$$

Equation (56) constitutes our newly proposed "invariant"-factor-cluster expansion, with successive terms given by (59). The computation of the $\mathfrak{Z}_n$ must be carried out in terms of the basic ingredients, the "invariant" $\tilde{\mathfrak{Z}}_1$, $\mathfrak{Z}_n$'s of (57).

To keep the terminology and the various relationships straight, we list the four cluster formalisms just analyzed:

(1). The conventional one of IY, a "noninvariant" formalism in which a sum-of-factors decomposition law is postulated for the IY subnormalization integrals.

(2). The conventional one of AHT, an "invariant" formalism in which a sum-of-factors decomposition law is postulated for the AHT subnormalization integrals.

(3). A new "noninvariant-factor-cluster" formalism (henceforth, the FIY formalism) in which a product decomposition law is postulated for the IY subnormalization integrals.

(4). A new "invariant-factor-cluster" formalism (henceforth, the FAHT formalism) in which a product decomposition law is postulated for the AHT subnormalization integrals.

There is one important feature of general cluster-expansion theory that we have not explicitly pointed out: the arbitrariness of the n-body subnormalization integrals with $n < N$. Since, in any complete cluster expansion of $I(\alpha)$, the *net* coefficient of every subnormalization integral except the last one, $I_{1\ldots N} = I(\alpha) = \mathfrak{I}_N$, is strictly zero, the subnormalization integrals with $n < N$ are entirely at our disposal. The choices (29), (39) are only two of an infinite number of possibilities. One should take advantage of this arbitrariness in tailoring the cluster formalism to the problem to be solved. An illuminating example of how this may be done is provided by the work of Feenberg and his collaborators on liquid He³.[14,23]

3. APPLICATION OF THE NEW "INVARIANT-CLUSTER" FORMALISM

We now apply the "invariant-factor-cluster" or FAHT formalism developed in the preceding section

to the evaluation of the n-particle spatial distribution function (23). The resolution of the corresponding operator given in (24) leads to the following expressions for the basic invariants of the expansion:

$$\tilde{\mathfrak{I}}_1^{(n)}(r_1) = \frac{1}{N}\sum_{i=1}^{N}\sum_{s_{z_1}}\sum_{t_{z_1}} \varphi_i^*(\mathbf{r}_1, s_{z_1}, t_{z_1})\varphi_i(\mathbf{r}_1, s_{z_1}, t_{z_1}), \quad n = 1,$$
$$= 0, \quad n \geq 2,$$

$$\mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = \sum_{\langle l(1)\cdots l(q)\rangle} \frac{q!}{(q-n)!}\int \prod_{b=1}^{q} dx_b' \psi^*_{l(1)\cdots l(q)} \times (x_1' \cdots x_q') \prod_{i=1}^{n} \delta(\mathbf{r}_i - \mathbf{r}_i')\psi_{l(1)\cdots l(q)}(x_1' \cdots x_q')$$
$$\left[\sum_{\langle l(1)\cdots l(q)\rangle}\int \prod_{b=1}^{q} dx_b' \times \psi^*_{l(1)\cdots l(q)}(x_1' \cdots x_q')\psi_{l(1)\cdots l(q)}(x_1' \cdots x_q')\right]^{-1} \quad q \geq n,$$
$$= 0, \quad q < n. \tag{60}$$

Thus, in accordance with formulas (56)–(59) the expansion for $\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$ takes the form

$$\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = N\tilde{\mathfrak{I}}_1^{(n)} + \sum_{p=1}^{N} {}_kC_N \mathfrak{I}_k^{(n)} \tag{61}$$

with

$$\mathfrak{I}_1^{(n)}(\mathbf{r}_1) = 0, \quad n > 1,$$
$$= \frac{1}{N}\sum_{s_{z_1}}\sum_{t_{z_1}}\sum_{i=1}^{N}\left\{\frac{\varphi_i^*(\mathbf{r}_1 s_{z_1} t_{z_1})\varphi_i(\mathbf{r}_1 s_{z_1} t_{z_1})}{\left[\sum_{j=1}^{N}\frac{1}{N}\int dx \psi_j^*(x)\psi_j(x)\right]} - \varphi_i^*(\mathbf{r}_1 s_{z_1} t_{z_1})\varphi_i(\mathbf{r}_1 s_{z_1} t_{z_1})\right\}, \quad n = 1,$$

$$\mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = 0, \quad q < n, \tag{62}$$
$$= \sum_{k=n}^{q} \frac{(-1)^{q-k}q!}{k!\,(q-k)!}\mathfrak{I}_k^{(n)}$$
$$= \sum_{k=n}^{q} \frac{(-1)^{q-k}q!}{(q-k)!\,(k-n)!} \times \frac{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k') \prod_{i=1}^{n} \delta(\mathbf{r}_i - r_i')}{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k')}, \quad q \geq n.$$

In arriving at the last line of (62), we have interchanged the integrations and summations which occur in both the numerator and denominator of the $\mathfrak{I}_k^{(n)}$'s and have defined

$$D_k(x_1' \cdots x_k') = \sum_{\langle l(1)\cdots l(k)\rangle} \psi^*_{l(1)\cdots l(k)}(x_1' \cdots x_k')\psi_{l(1)\cdots l(k)}(x_1' \cdots x_k'). \tag{63}$$

Some intuitively satisfying features of this new cluster expansion deserve special study. First of all, since

$$\int d\mathbf{r}_1 \cdots d\mathbf{r}_n \times \left\{\frac{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k') \prod_{i=1}^{n} \delta(\mathbf{r}_i - \mathbf{r}_i')}{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k')}\right\} = 1, \quad k \geq n, \tag{64}$$

it follows that

$$\int d\mathbf{r}_1 \mathfrak{I}_1^{(n)}(\mathbf{r}_1) = 0, \quad \text{all } n; \quad \int d\mathbf{r}_1 \tilde{\mathfrak{I}}_1^{(n)}(\mathbf{r}_1) = 1, \quad n = 1, \quad = 0, \quad n > 1;$$
$$\int d\mathbf{r}_1 \cdots d\mathbf{r}_n \mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) \equiv 0, \quad q < n; \tag{65}$$
$$\int d\mathbf{r}_1 \cdots d\mathbf{r}_n \mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = \sum_{k=n}^{q} \frac{(-1)^{q-k}q!}{(q-k)!\,(k-n)!} = n!, \quad q = n, \quad n > 1, \quad = 0, \quad q > n.$$

Thus the integral of the *first* nonvanishing term in the cluster expansion (61) provides the *complete* normalization, $N!/[(N-n)!]$, of $\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$. There are no contributions to the normalization from higher orders. Moreover, to the extent that the ratio

$$\mathcal{R}_k(\mathbf{r}_1 \cdots \mathbf{r}_n) = \frac{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k') \prod_{i=1}^{n} \delta(\mathbf{r}_i - \mathbf{r}_i')}{\int \prod_{b=1}^{k} dx_b' D_k(x_1' \cdots x_k')}, \quad k \geq n,$$

is independent of k, the quantities $\mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$, $q > n$, vanish. For if we assume this ratio is independent of k and denote it by $\bar{\mathcal{R}}(\mathbf{r}_1 \cdots \mathbf{r}_n)$, we obtain directly from (62) that

$$\mathfrak{I}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = \bar{\mathcal{R}}(\mathbf{r}_1 \cdots \mathbf{r}_n) \cdot \sum_{k=n}^{q} \frac{(-1)^{q-k}q!}{(q-k)!\,(k-n)!}$$
$$= n!\,\bar{\mathcal{R}}(\mathbf{r}_1 \cdots \mathbf{r}_n), \quad q = n,$$
$$= 0, \quad q > n. \tag{66}$$

Finally, employing two relations from combinatorial

analysis, we are able to perform partial summations of the addends in (61) resulting in a cogent expression for $\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$ in terms of the fundamental "invariants," the $\mathfrak{Z}_p$'s. We need only recall that

$$\sum_{k=n}^{\mu}\sum_{q=n}^{k} F(q,k) = \sum_{q=n}^{\mu}\sum_{k=q}^{\mu} F(q,k) \tag{67}$$

and

$$\sum_{k=q}^{\mu}\frac{(-1)^{k-q}(N-q)!}{(N-k)!(k-q)!} = \sum_{j=0}^{\mu-q}\frac{(-1)^{j}(N-q)!}{(N-q-j)!\,j!} = \frac{(-1)^{\mu-q}(N-q-1)!}{(N-\mu-1)!(\mu-q)!}, \tag{68}$$

in order to obtain for the sum of the first $\mu+1$ ($\mu \leq N$) terms

$$\begin{aligned}
&N\tilde{\mathfrak{Z}}_1^{(n)} + \sum_{k=1}^{\mu} {}_kC_N \mathfrak{Z}_k^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n) \\
&= N\tilde{\mathfrak{Z}}_1^{(n)} + \sum_{k=n}^{\mu} {}_kC_N \mathfrak{Z}_k^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n), \\
&= \sum_{k=n}^{\mu}\sum_{q=n}^{k} \frac{N!}{k!(N-k)!}\cdot\frac{(-1)^{k-q}k!}{q!(k-q)!}\,\mathfrak{Z}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n), \\
&= \sum_{q=n}^{\mu}\left[\sum_{k=q}^{\mu}\frac{(-1)^{k-q}N!}{(N-k)!(k-q)!}\right]\frac{\mathfrak{Z}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)}{q!}, \\
&= \sum_{q=n}^{\mu}\left[\frac{(-1)^{\mu-q}N!(N-q-1)!}{(N-q)!(N-\mu-1)!(\mu-q)!}\,\frac{\mathfrak{Z}_q^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)}{q!}\right].
\end{aligned} \tag{69}$$

This "invariant-factor-cluster" expansion of the n-particle spatial distribution function, unlike the corresponding AHT expansion generated via (46), (47), is applicable to finite as well as to infinite many-fermion systems. A reading of the following paper in this series will leave no doubt that, order by order, the terms by which the two expansions differ become negligible in the limit $N \to \infty$. Hence, in treating infinitely extended fermion systems, the "invariant-factor-cluster" and the AHT formalisms yield the same spatial distribution functions order by order. However, this statement *cannot* be made about the spatial integral $\int d\mathbf{r}_1 \cdots d\mathbf{r}_n$ of the two expansions; the AHT expansion does not possess the desirable feature that only its first nonvanishing term contributes to the normalization of $\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$.

It might be noted that relations analogous to Eqs. (64)–(69) serve to establish completely analogous features of the FIY expansion for the n-particle spatial distribution function. And we might also note, explicitly, that the usual IY expansion for $\rho^{(n)}(\mathbf{r}_1 \cdots \mathbf{r}_n)$ does not possess the very important feature of the FIY expansion that the only contribution to the normalization comes from the first nonvanishing term of the expansion.

The AHT expansion for the two-particle spatial distribution function has been studied in detail by Aviles[18] and by Hartogh and Tolhoek.[19] A numerical comparison of AHT and IY expansions for $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ was carried out by Woo in the framework of the liquid-He^3 problem.[14]

We conclude this section with a study of $D_k(x_1 \cdots x_k)$ for the special case that Ψ is given by (12)–(14). Since the correlation function in the corresponding k-particle wavefunctions [cf. (27d)] is independent of orbital labels, the structure of $D_k(x_1 \cdots x_k)$ simplifies to

$$D_k(x_1 \cdots x_k) = \left[\prod_{j=1}^{k}\prod_{\langle l(1)\cdots l(j)\,|\,1\cdots k\rangle} f_j^2(x_{l(1)} \cdots x_{l(j)})\right] \cdot \left\{\sum_{\langle q(1)\cdots q(k)\rangle} \Phi^*_{q(1)\cdots q(k)}(x_1 \cdots x_k)\Phi_{q(1)\cdots q(k)}(x_1 \cdots x_k)\right\}. \tag{70}$$

Explicitly, the independent-particle factor may be written

$$\Phi_{q(1)\cdots q(k)}(x_1 \cdots x_k) = (k!)^{-\frac{1}{2}} \det \Lambda_{q(1)\cdots q(k)}(x_1 \cdots x_k), \tag{71}$$

where

$$[\Lambda_{q(1)\cdots q(k)}]_{ij} = \varphi_{q(i)}(x_j), \quad i,j = 1, \cdots k. \tag{72}$$

Noting that

$$\begin{aligned}
&\Phi^*_{q(1)\cdots q(k)}(x_1 \cdots x_k)\Phi_{q(1)\cdots q(k)}(x_1 \cdots x_k) \\
&\quad = (1/k!)\det \Lambda^*_{q(1)\cdots q(k)} \cdot \det \Lambda_{q(1)\cdots q(k)} \\
&\quad = (1/k!)\det [\Lambda^\dagger_{q(1)\cdots q(k)}\Lambda_{q(1)\cdots q(k)}],
\end{aligned} \tag{73}$$

Eq. (70) becomes

$$\begin{aligned}
&D_k(x_1 \cdots x_k) \\
&\quad = \left[\prod_{j=1}^{k}\prod_{\langle l(1)\cdots l(j)\,|\,1\cdots k\rangle} f_j^2(x_{l(1)} \cdots x_{l(j)})\right] \\
&\qquad \cdot \left\{\frac{1}{k!}\sum_{\langle q(1)\cdots q(k)\rangle} \det Q_{q(1)\cdots q(k)}(x_1 \cdots x_k)\right\},
\end{aligned} \tag{74}$$

the elements of the matrix

$$Q_{q(1)\cdots q(k)} = \Lambda^\dagger_{q(1)\cdots q(k)}\Lambda_{q(1)\cdots q(k)} \tag{75}$$

being just

$$[Q_{q(1)\cdots q(k)}]_{ij} = \sum_{p=1}^{k} \varphi^*_{q(p)}(x_i)\varphi_{q(p)}(x_j). \tag{76}$$

To carry the analysis further, we must address ourselves to the crucial evaluation of the sum of determinants $\sum_{\langle q(1)\cdots q(k)\rangle} \det Q_{q(1)\cdots q(k)}$.

Each element of $Q_{q(1)\cdots q(k)}$ is the sum of the same k functions; the arguments of these functions vary from one element to another and thus serve as the row–column indices. Successive application of a well-known theorem governing the expansion of the determinant of

a matrix in which each row of the ith column is the sum of two terms allows us to express det $Q_{q(1)\cdots q(k)}$ as the sum of k^k determinants:

$$\sum_{\langle q(1)\cdots q(k)\rangle} \det Q_{q(1)\cdots q(k)} = \sum_{\langle q(1)\cdots q(k)\rangle} \left\{ \sum_{j(1)=q(1)}^{q(k)} \cdots \sum_{j(k)=q(1)}^{q(k)} \times \begin{vmatrix} \varphi^*_{j(1)}(x_1)\varphi_{j(1)}(x_1) & \cdots & \varphi^*_{j(k)}(x_1)\varphi_{j(k)}(x_k) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi^*_{j(1)}(x_k)\varphi_{j(1)}(x_1) & \cdots & \varphi^*_{j(k)}(x_k)\varphi_{j(k)}(x_k) \end{vmatrix} \right\}. \tag{77}$$

Notice, however, that any term with $j(i) = j(l)$, $i \neq l$ is zero, for the determinant comprising that addend has two columns proportional. This observation may be repeated as we take into account the sum over the $N!/[(N-k)!\,k!]$ choices for the orbitals from among $\{\varphi_1, \cdots, \varphi_N\}$ with the result that

$$\sum_{\langle q(1)\cdots q(k)\rangle} \det Q_{q(1)\cdots q(k)} = \sum_{j(1)=1}^{N} \cdots \sum_{j(k)=1}^{N} \times \begin{vmatrix} \varphi^*_{j(1)}(x_1)\varphi_{j(1)}(x_1) & \cdots & \varphi^*_{j(k)}(x_1)\varphi_{j(k)}(x_k) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \varphi^*_{j(1)}(x_k)\varphi_{j(1)}(x_1) & \cdots & \varphi^*_{j(k)}(x_k)\varphi_{j(k)}(x_k) \end{vmatrix}. \tag{78}$$

It now behooves us to reverse the procedure by which each det $Q_{q(1)\cdots q(k)}$ was written as the sum of k^k determinants and thus express the right-hand side of (78) as a single determinant. In this way we arrive at

$$\sum_{\langle q(1)\cdots q(k)\rangle} \det Q_{q(1)\cdots q(k)} = \begin{vmatrix} \sum_{i=1}^{N} \varphi^*_i(x_1)\varphi_i(x_1) & \cdots & \sum_{i=1}^{N} \varphi^*_i(x_1)\varphi_i(x_k) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \sum_{i=1}^{N} \varphi^*_i(x_k)\varphi_i(x_1) & \cdots & \sum_{i=1}^{N} \varphi^*_i(x_k)\varphi_i(x_k) \end{vmatrix} \tag{79}$$

and thus, returning to (74), conclude that

$$D_k(x_1 \cdots x_k) = \left[\prod_{j=1}^{k} \prod_{\langle l(1)\cdots l(j) \mid 1\cdots k\rangle} f^2_j(x_{l(1)} \cdots x_{l(j)}) \right] W_k(x_1 \cdots x_k), \tag{80}$$

where

$$W_k(x_1 \cdots x_k) = \frac{1}{k!} \begin{vmatrix} \Delta(x_1, x_1) & \cdots & \Delta(x_1, x_k) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \Delta(x_k, x_1) & \cdots & \Delta(x_k, x_k) \end{vmatrix} \tag{81}$$

and

$$\Delta(x_i, x_j) = \sum_{l=1}^{N} \varphi^*_l(x_i)\varphi_l(x_j). \tag{82}$$

Thus the sum of products of determinants (70) has been reduced to a single determinant derived from a matrix whose elements, defined by (82), are easily shown to be invariant under the unitary transformation (38).

4. COMPARISON OF THE FOUR CLUSTER FORMALISMS

The factor-cluster formalisms—beside being directly useful for practical computation—provide a crutch for the extension of the IY and AHT formalisms to finite systems. In previous derivations, terms down from the leading terms in G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$ by $O(1/N)$ have not been tracked down. These terms may now be generated as follows (we outline the procedure for the IY case): The final equations in (33) and (48) along with (34) and (36) imply that we must have, for all N,

$$G_{\mathrm{IY}} = \sum_{n=1}^{N} \sum_{\langle j(1)\cdots j(n)\rangle} \ln Y_{j(1)\cdots j(n)}. \tag{83}$$

The equations which serve to decompose the

$$\ln I_{i(1)\cdots i(p)}$$

in terms of the

$$\ln Y_{j(1)\cdots j(n)}, \quad \{j(1)\cdots j(n)\} \subset \{i(1)\cdots i(p)\},$$

are easily inverted (as usual one looks for an expression for $\ln Y_{j(1)\cdots j(n)}$ in terms of the $\ln I_{k(1)\cdots k(q)}$, $\{k(1)\cdots k(q)\} \subset \{j(1)\cdots j(n)\}$, which yields a set of identities). The $I_{k(1)\cdots k(q)}$ are then decomposed in the IY fashion in terms of the $\tilde{I}_{k(l)}$, $r = 1\cdots q$ and the $x_{l(1)\cdots l(s)}$, $\{l(1)\cdots l(s)\} \subset \{k(1)\cdots k(q)\}$. Upon expanding the logarithms, there results for G_{IY} an expansion in the IY normalized cluster integrals which is valid for all N. Indeed, in this the IY case a linked cluster theorem may be proven: G_{IY} is the sum of all distinct $x\ldots$'s and, with appropriate coefficients, all possible linked products of these cluster integrals. Corresponding to (83) we have, in the AHT case,

$$\mathfrak{G}_{\mathrm{AHT}} = \sum_{n=1}^{N} {}_nC_N \ln \mathfrak{Y}_n, \tag{84}$$

and everything goes through in precisely the same way except that there is no question of a linked cluster

theorem here since *linked product* is undefined in this formalism. It will be our task in the following paper to carry out the program sketched in this paragraph; in particular we prove the linked cluster theorem for the IY expansion and contrive a rule for determining the aforementioned appropriate coefficients.

So, finally, one has four cluster expansions for $\langle S\rangle$, all valid for finite N as well as in the many-body limit. The "invariant" expansions (AHT and FAHT) clearly differ from the "noninvariant" expansions (IY and FIY) order by order as well as term by term. IY and FIY expansions, and on the other hand AHT and FAHT expansions, coincide order by order but not, of course, term by term.

The question naturally arises: Which is the best of these four formalisms to use in a given calculation? Assuredly, the answer to this question will depend on the details of the problem at hand. Nevertheless, some very general observations can be made, based on the fact that the IY (AHT) expansion for $\langle S\rangle$ may be derived from the FIY (FAHT) expansion by employing the procedure just described. As a result of the logarithmic expansions, there are, "in general" (i.e., there exist special exceptions as we see in a moment), an infinite number of nonnegligible terms in each order of the IY and AHT expansions, even for finite N. On the other hand, the factorized expansions are characterized by a finite number of terms in each order, therefore a finite number of terms for finite N. (In fact there is just one term of each order in the FIY case as well as in the FAHT case if we obey our convention and regard each summand of $\sum_{n=1}^{N}$ as a single term.) Surely the only effect of the extra expansions involved in the transition FIY $\rightarrow$ IY or FAHT $\rightarrow$ AHT is to lessen the rate of convergence of IY relative to FIY, AHT relative to FAHT. Clearly the order-by-order identity of IY and FIY (AHT and FAHT) expansions is of no *practical* significance if a given order in the IY (AHT) expansion contains an infinite number of terms.

For *finite* systems we may remove the qualification "in general" of the preceding paragraph. We expect a preference of FIY over IY, FAHT over AHT. But there is as yet no numerical evidence favoring one expansion over any of the others, even in a single particular case, since there has only been one cluster-method calculation on a finite system, that one using the IY procedure in the approximate evaluation of $\langle H\rangle$ for the O^{16} nucleus assuming a Jastrow trial wavefunction.[25,26]

Let us now consider *extended* systems with short range forces. We keep N finite but suppose it to be large enough that contributions $O(N^0)$ may be discarded compared with contributions $O(N)$. Then only a finite number of terms survive in a given order of the AHT expansion. Consequently the order-by-order identity of AHT and FAHT does, in this case, have practical significance. On the other hand the IY expansion does not necessarily so collapse. Let there exist for the observable S in question an irreducible resolution (18)–(19) such that, for some q, $S(1\cdots q)$ is negligible for all configurations except those in which particles $j(1)\cdots j(p)$, $\{j(1)\cdots j(p)\}\subset\{1\cdots q\}$, are within a volume $O(1/N)\Omega$, and is not identically zero. (The Hamiltonian H fits this requirement; the operator $\mathrm{P}^{(n)}(\mathbf{r}_1\cdots\mathbf{r}_n;\mathbf{r}'_1\cdots\mathbf{r}'_N)$ whose expectation value yields the n-particle spatial distribution function does not.) If nonlocalized orbitals (with plane-wave spatial factors) are used, then all except a finite number of terms disappear from each order of the IY expansion. But if localized orbitals are used, there is no such simplification and there are still an infinite number of terms in each order. Proof of these statements rests upon the detailed results of the next paper; however, their plausibility may be enhanced at this point by a consideration of the following two terms—typical addends of the infinite sequences in question—contributing respectively to $(d/d\alpha)G_{\mathrm{IY}}\big|_{\alpha=0}$ and $(d/d\alpha)\mathfrak{G}_{\mathrm{AHT}}\big|_{\alpha=0}$:

$$\frac{d}{d\alpha}\sum_{i<j}x_{ij}^2\bigg|_{\alpha=0},$$

$$\frac{d}{d\alpha}\tfrac{1}{2}N(N-1)\,\mathfrak{X}_2^2\bigg|_{\alpha=0}$$

Thus in practical n-body cluster calculations of $\langle H\rangle$ for *extended systems* with short-range forces, it makes no difference whether we use the AHT or the FAHT expansion. For *quantum liquids* there will be no difference in practical IY and FIY calculations of this nature. For *quantum solids* there will be; indeed the numerical work of Nosanow and his collaborators[27] has revealed the expected superiority of the FIY formalism over the IY formalism in a variational calculation of the cohesive energy of solid He^3 through three-body clusters assuming a Jastrow trial wavefunction.

There exists one last piece of information bearing on the comparison of the four formalisms in practical application. Through third order, IY and AHT results for the two-particle spatial distribution function in

[25] J. Dabrowski, Proc. Phys. Soc. (London) **71**, 685 (1958).
[26] J. Dabrowski, Proc. Phys. Soc. (London) **72**, 499 (1958).
[27] L. H. Nosanow, Phys. Rev. **154**, 175 (1967).

liquid He³ assuming a Bose correlation factor are hardly distinguishable.[14,23]

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the helpful discussions we had with Professor E. Feenberg. One of the authors (J. W. C.) would like to thank Professor Aage Bohr for extending to him the hospitality of the Niels Bohr Institute, where part of this work was carried out. The other author (P. W.) gratefully acknowledges the assistance of the Theoretical Chemistry Institute where much of this work was done.

APPENDIX

The approximate evaluation of expectation values via the cluster expansion techniques of this paper may be carried through with any wavefunction once a method of defining the n-particle wavefunctions has been proposed. Among the possibilities considered in Sec. 1 is the exact N-particle wavefunction $\Psi_{\mathbf{m}}$, written in the manner of Sinanoğlu:

$$\Psi_{\mathbf{m}} = \mathcal{A}(N)\Big[\prod_{i=1}^{N}\varphi_{m_i}(x_i) + \sum_{p}\prod_{i\neq p}\varphi_{m_i}(x_i)U^{\mathbf{m}}_{m_p}(x_p) + (2!)^{-\frac{1}{2}}\sum_{p<q}\prod_{i\neq p,q}\varphi_{m_i}(x_i)U^{\mathbf{m}}_{m_p m_q}(x_p x_q) + \cdots + (N!)^{-\frac{1}{2}}U^{\mathbf{m}}_{m_1\cdots m_N}(x_1\cdots x_N)\Big]. \quad \text{(A1)}$$

This Appendix is devoted to an examination of Sinanoğlu's expression for $\Psi_{\mathbf{m}}$ with the intention of illuminating its structure and that of the corresponding n-particle wavefunctions.

We note that in the language of a configuration interaction (CI) treatment,

$$U^{\mathbf{m}}_{m_{j(1)}\cdots m_{j(q)}}(x_1\cdots x_q) = \sum_{\kappa_1<\cdots\kappa_q\notin\mathbf{m}} C_{m_N\cdots\kappa_q\cdots\kappa_1\cdots m_1;m_1\cdots m_N}\mathcal{A}(q)[\varphi_{\kappa_1}(x_1)\cdots\varphi_{\kappa_q}(x_q)], \quad \text{(A2)}$$

where in practice the coefficients must be determined by a perturbative or variational procedure, but are formally given by

$$C_{m_N\cdots\kappa_q\cdots\kappa_1\cdots m_1;m_1\cdots m_{j(1)}\cdots m_{j(q)}\cdots m_N} = \int\prod_{b=1}^{N}dx'_b\{\mathcal{A}(N)[\varphi_{m_1}(x'_1)\cdots\varphi_{\kappa_1}(x'_{j(1)})\cdots\varphi_{\kappa_q}(x'_{j(q)})\cdots\varphi_{m_N}(x'_N)]\}^*\Psi_{\mathbf{m}}(x'_1\cdots x'_N). \quad \text{(A3)}$$

In (A2) the sum is over all combinations of q orbitals excluded from the set $\{m_1\cdots m_N\}$. The expression for the correlation function $U^{\mathbf{m}}_{m_{j(1)}\cdots m_{j(q)}}$ may be recast into a compact, but revealing, form. Upon inserting (A3) into (A2) and interchanging the sums and integrations which occur, we find that $U^{\mathbf{m}}_{m_{j(1)}\cdots m_{j(q)}}$ may be written

$$U^{\mathbf{m}}_{m_{j(1)}\cdots m_{j(q)}} = (N!\,q!)^{-\frac{1}{2}}\int\prod_{b=1}^{N}dx'_b\sum_{\kappa_1<\cdots<\kappa_q\notin\mathbf{m}}\begin{vmatrix}\varphi^*_{m_1}(x'_1) & \cdots & \varphi^*_{\kappa_1}(x'_1) & \cdots & \varphi^*_{\kappa_q}(x'_1) & \cdots & \varphi^*_{m_N}(x'_1)\\ \vdots & & \vdots & & \vdots & & \vdots\\ \varphi^*_{m_1}(x'_N) & \cdots & \varphi^*_{\kappa_1}(x'_N) & \cdots & \varphi^*_{\kappa_q}(x'_N) & \cdots & \varphi^*_{m_N}(x'_N)\end{vmatrix}$$

$$\times\begin{vmatrix}1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0\\ 0 & 1 & & & & & & \\ \vdots & & \ddots & & & & & \vdots\\ & & & \varphi_{\kappa_1}(x_1) & \cdots & \varphi_{\kappa_1}(x_q) & & \\ & & & \vdots & & \vdots & & \\ & & & \varphi_{\kappa_q}(x_1) & \cdots & \varphi_{\kappa_q}(x_q) & & \\ & & & \vdots & & \vdots & 1 & \\ 0 & & & \vdots & & \vdots & & 0\\ \vdots & & & \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1\end{vmatrix}\Psi_{\mathbf{m}}(x'_1\cdots x'_N). \quad \text{(A4)}$$

The integrand may be simplified according to the following scheme: For a given set of q orbital labels $\kappa_1 \cdots \kappa_q$ the product of two determinants found in (A4) can be written as a single determinant,

$$\det \Theta^{\kappa_1 \cdots \kappa_q}_{m_{j(1)} \cdots m_{j(q)}} = \begin{vmatrix} \varphi^*_{m_1}(x'_1) & \cdots & \sum_{i=1}^{q} \varphi^*_{\kappa_i}(x'_1)\varphi_{\kappa_i}(x_1) & \cdots & \sum_{i=1}^{q} \varphi^*_{\kappa_i}(x'_1)\varphi_{\kappa_i}(x_q) & \cdots & \varphi^*_{m_N}(x'_1) \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \varphi^*_{m_1}(x'_N) & \cdots & \sum_{i=1}^{q} \varphi^*_{\kappa_i}(x'_N)\varphi_{\kappa_i}(x_1) & \cdots & \sum_{i=1}^{q} \varphi^*_{\kappa_i}(x'_N)\varphi_{\kappa_i}(x_q) & \cdots & \varphi^*_{m_N}(x'_N) \end{vmatrix},$$

that is, a determinant in which the sth row of the tth column—occupied in $\Phi^*_{\mathbf{m}}(x'_1 \cdots x'_N)$ by $\varphi^*_{m_t}(x'_s)$—is now given by

$$[\Theta^{\kappa_1 \cdots \kappa_q}_{m_{j(1)} \cdots m_{j(q)}}]_{st} = \begin{cases} \sum_{i=1}^{q} \varphi^*_{\kappa_i}(x'_s)\varphi_{\kappa_i}(x_l), & \text{if} \quad t = j(l), \quad j(l) \in \{j(1) \cdots j(q)\}, \\ \varphi^*_{m_t}(x'_s), & \text{if} \quad t \notin \{j(1) \cdots j(q)\}. \end{cases} \tag{A5}$$

Moreover, by a procedure like that which leads from (73) to (79), the sum of such determinants may be reduced to a single determinant so that

$$\sum_{\kappa_1 < \cdots < \kappa_q \notin \mathbf{m}} \det \Theta^{\kappa_1 \cdots \kappa_q}_{m_{j(1)} \cdots m_{j(q)}} = \begin{vmatrix} \varphi^*_{m_1}(x'_1) & \cdots & \sum_{\kappa \notin \mathbf{m}} \varphi^*_{\kappa}(x'_1)\varphi_{\kappa}(x_1) & \cdots & \sum_{\kappa \notin \mathbf{m}} \varphi^*_{\kappa}(x'_1)\varphi_{\kappa}(x_q) & \cdots & \varphi^*_{m_N}(x'_1) \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \varphi^*_{m_1}(x'_N) & \cdots & \sum_{\kappa \notin \mathbf{m}} \varphi^*_{\kappa}(x'_N)\varphi_{\kappa}(x_1) & \cdots & \sum_{\kappa \notin \mathbf{m}} \varphi^*_{\kappa}(x'_N)\varphi_{\kappa}(x_q) & \cdots & \varphi^*_{m_N}(x'_N) \end{vmatrix}. \tag{A6}$$

Each of the sums appearing in (A6) is over all orbital labels $\kappa \notin \{m_1 \cdots m_N\}$. Since the single-particle states form an orthonormal basis in the one-particle Hilbert space, we can write

$$\sum_{\kappa \notin \mathbf{m}} \varphi^*_{\kappa}(x')\varphi_{\kappa}(x) = \delta(x', x) - \sum_{i=1}^{N} \varphi^*_{m_i}(x')\varphi_{m_i}(x). \tag{A7}$$

Upon substituting this expression into (A6) we discover that, in addition to $\delta(x', x)$, only the terms involving $\varphi^*_{m_l}(x')\varphi_{m_l}(x)$ with $l \in \{j(1), \cdots j(q)\}$ give rise to nonvanishing contributions when the resulting determinant is expanded. Thus, we define

$$\mathfrak{D}_{m_{j(1)} \cdots m_{j(q)}}(x_s, x_t) = \delta(x_s, x_t) - \sum_{i=1}^{q} \varphi^*_{m_{j(i)}}(x_s)\varphi_{m_{j(i)}}(x_t) \tag{A8}$$

and arrive at the result

$$\begin{aligned} &U^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(q)}}(x_1 \cdots x_q) \\ &= (N!\, q!)^{-\frac{1}{2}} \int \prod_{b=1}^{N} dx'_b \begin{vmatrix} \varphi^*_{m_1}(x'_1) & \cdots & \mathfrak{D}_{m_{j(1)} \cdots m_{j(q)}}(x'_1, x_1) & \cdots & \mathfrak{D}_{m_{j(i)} \cdots m_{j(q)}}(x'_1, x_q) & \cdots & \varphi^*_{m_N}(x'_1) \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \varphi^*_{m_1}(x'_N) & & \mathfrak{D}_{m_{j(i)} \cdots m_{j(q)}}(x'_N, x_1) & \cdots & \mathfrak{D}_{m_{j(i)} \cdots m_{j(i-1)}}(x'_N, x_q) & \cdots & \varphi^*_{m_N}(x'_N) \end{vmatrix} \\ &\quad \times \Psi_{\mathbf{m}}(x'_1 \cdots x'_N) \\ &= \sqrt{\frac{N!}{q!}} \int \prod_{b=1}^{N} dx'_b [\varphi^*_{m_1}(x'_1) \cdots \mathfrak{D}_{m_{j(1)} \cdots m_{j(q)}}(x'_{j(1)}, x_1) \cdots \mathfrak{D}_{m_{j(1)} \cdots m_{j(q)}}(x'_{j(q)}, x_q) \cdots \varphi_{m_N}(x'_N)] \Psi_{\mathbf{m}}(x'_1 \cdots x'_N). \end{aligned} \tag{A9}$$

The second expression, in which the determinant has been replaced by a simple product, is obtained by recognizing that $\Psi_{\mathbf{m}}(x_1' \cdots x_N')$ must be antisymmetric under the interchange of any two particle labels.

Although we have assumed $\Psi_{\mathbf{m}}(x_1' \cdots x_N')$ to be the exact N-particle wavefunction, the above derivation of (A9) may be bypassed and this formula viewed as a starting point for the analysis of trial wavefunctions.[5] Thus given a trial wavefunction $\tilde{\Psi}_{\mathbf{m}}(x_1' \cdots x_N')$ (which may or may not turn out to be the sought-for eigenfunction) the corresponding correlation functions $\tilde{U}^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(q)}}$ from which to construct the "n-particle wavefunctions" may be found via (A9). If, as in accordance with (27b), we define

$$\tilde{\psi}_{m_{j(1)} \cdots m_{j(n)}}(x_1 \cdots x_n) = \left(\frac{n!}{N!}\right)^{\frac{1}{2}} \mathcal{A}(n)\Big[\prod_{i=1}^{n} \varphi_{m_{j(i)}}(x_i) + \sum_{p} \prod_{i=1, i\neq p}^{n} \varphi_{m_{j(i)}}(x_i)\tilde{U}^{\mathbf{m}}_{m_{j(p)}}(x_p) + (2!)^{-\frac{1}{2}} \sum_{p<q} \prod_{i\neq p,q}^{n} \varphi_{m_{j(i)}}(x_j)\tilde{U}^{\mathbf{m}}_{m_{j(p)}m_{j(q)}}(x_p, x_q) + \cdots + (n!)^{-\frac{1}{2}} U^{\mathbf{m}}_{m_{j(1)} \cdots m_{j(n)}}(x_1 \cdots x_n)\Big],$$

then the insertion of (A9) for the $\tilde{U}^{\mathbf{m}}_{m_{l(1)} \cdots m_{l(p)}}$ yields the very natural expression

$$\tilde{\psi}_{m_{j(1)} \cdots m_{j(n)}}(x_1 \cdots x_n) = \int \prod_{b=1}^{N} dx_b' \Big[\prod_{i=1}^{n} \delta(x_{j(i)}', x_i) \prod_{p\notin\{j(1)\cdots j(n)\}}^{N} \varphi^*_{m_p}(x_p')\Big]\tilde{\Psi}_{\mathbf{m}}(x_1' \cdots x_N'). \tag{A10}$$

We have developed an explicit means of constructing the n-particle wavefunction from a given form for $\tilde{\Psi}_{\mathbf{m}}(x_1 \cdots x_N)$. Certainly our prescription is not unique. Indeed, since completely summing any cluster expansion for $\langle S \rangle$ leaves a final expression involving only the N-body wavefunction $\tilde{\Psi}_{\mathbf{m}}(x_1 \ldots x_N)$, we must admit that from a formal standpoint $\psi_{m_{j(1)} \cdots m_{j(n)}}$ (with $n < N$) may be defined quite arbitrarily. Nevertheless, we anticipate that the method outlined here, resulting in the intuitively natural form for $\tilde{\psi}_{m_{j(1)} \cdots m_{j(n)}}$ as evidenced in (A10), will prove most advantageous. Moreover, when the problem at hand allows the particles to be divided into subgroups such that the members of one subgroup are only weakly correlated with those of another, $\tilde{\psi}_{m_{j(1)} \cdots m_{j(n)}}(x_1 \cdots x_n)$ as defined by (A10) approximates the probability amplitude[28] for such a subgroup of n particles, these particles occupying orbitals $\varphi_{m_{j(1)}} \cdots \varphi_{m_{j(n)}}$ in the independent-particle picture. Finally, we observe that the n-particle wavefunctions defined in (27c) and (27d) are not equivalent to the function (A10) but involve the further step of replacing some of the correlation functions within the integrand of (A10) by unity.

[28] P. Westhaus, WIS-TCI Report No. 234 (unpublished).

Cluster Expansions in Many-Fermion Theory. II. Rearrangements of Primitive Decomposition Equations*

PAUL WESTHAUS
University of Wisconsin Theoretical Chemistry Institute, Madison, Wisconsin

AND

JOHN W. CLARK†
Department of Physics, Washington University, St. Louis, Missouri

(Received 5 June 1967)

The "factor-cluster" formalisms introduced in the preceding paper are used as a tool for a further development of the cluster theories proposed by Iwamoto and Yamada and by Aviles, Hartogh, and Tolhoek. The primitive decomposition characterizing each of the older formalisms is rearranged into the exponential of a series with a uniform N dependence in the limit of large N. In addition a "linked cluster" theorem is proven for the series comprising the exponent in the Iwamoto–Yamada formalism. Our derivations, unlike those of earlier authors, are valid for all N.

I. INTRODUCTION

In the first paper[1] (hereafter referred to as I) of this series, we reviewed the structure of correlated wavefunctions and discussed cluster expansion techniques, similar to those employed in statistical mechanics, as a means for the evaluation of expectation values with respect to such wavefunctions. In particular, two radically different, "factorized" cluster expansions were introduced to supplant the more familiar cluster developments of Iwamoto and Yamada[2] (IY) and of Aviles,[3] and Hartogh and Tolhoek[4] (AHT).

As presented in the literature, the final expansions of the IY and AHT cluster formalisms are suitable for use only in the many-body limit (N, the number of particles, and Ω, the volume of the system, approach infinity, while the density $\rho = N/\Omega$ is held constant). Using the notation of I, the generalized normalization integral (the generating function for the expectation value) is given in the IY formalism by

$$I(\alpha) = \left[\prod_{i=1}^{N} \tilde{I}_i(\alpha)\right] \cdot B_{12\ldots N}(\alpha), \tag{1}$$

where $B_{12\ldots N}$ has the decomposition, or primitive expansion,

$$B_{12\ldots N} = 1 + \sum_{n=1}^{N} \sum_{\langle j(1)\cdots j(n)\rangle} \sum_{l=1}^{n} \sum_{\substack{\text{all partitons of } j(1)\cdots j(n) \\ \text{among } n-l+1 \\ \text{factors}}} x_{\ldots} x_{\ldots} \cdots x_{\ldots} \tag{2}$$

in terms of the normalized cluster integrals $x_{j(p)\cdots j(q)}$, and in the AHT formalism by

$$I(\alpha) = (\tilde{\mathfrak{I}}_1(\alpha))^N \mathfrak{B}_N(\alpha), \tag{3}$$

where $\mathfrak{B}_N$ has the decomposition or primitive expansion

$$\mathfrak{B}_N = 1 + \sum_{n=1}^{N} {}_nC_N \sum_{\substack{\text{all partitions} \\ \text{such that} \\ \sum_{b=1}^{n} b\nu_b = n}} \frac{n!\, \mathfrak{X}_1^{\nu_1} \cdots \mathfrak{X}_b^{\nu_b} \cdots \mathfrak{X}_n^{\nu_n}}{\prod_{b=1}^{n} (b!)^{\nu_b} \nu_b!} \tag{4}$$

in terms of the "invariant" normalized cluster integrals $\mathfrak{X}_b$. Successive truncation of the above series for $B_{12\ldots N}$ and $\mathfrak{B}_N$ yields approximants possessing nonuniform N dependence; rearrangements of Eqs. (2) and (4) are necessary before useful approximations of the expectation value of the general operator S can be achieved. The formula which generates

$$\begin{aligned}\langle S\rangle &= \frac{d}{d\alpha} \ln I(\alpha)\Big|_{\alpha=0} \\ &= \left\{\sum_{i=1}^{N} \frac{d}{d\alpha} \ln \tilde{I}_i(\alpha) + \frac{d}{d\alpha} \ln B_{12\ldots N}\right\}\Big|_{\alpha=0} \\ &= \left\{N \frac{d}{d\alpha} \ln \tilde{\mathfrak{I}}_1(\alpha) + \frac{d}{d\alpha} \ln \mathfrak{B}_N\right\}\Big|_{\alpha=0}\end{aligned} \tag{5}$$

* Supported in part by the National Science Foundation, under Grant Nos. GP-3211 and GP-6210, a National Aeronautics and Space Administration Grant to Washington University, and National Aeronautics and Space Administration Grant NsG-275-62.

† Alfred P. Sloan Fellow.

[1] J. Clark and P. Westhaus, J. Math. Phys. **9**, 131 (1968) (preceding paper).

[2] F. Iwamoto and M. Yamada, Progr. Theoret. Phys. (Kyoto) **17**, 543 (1957).

[3] J. B. Aviles Jr., Ann. Phys. (N.Y.) **5**, 251 (1958).

[4] C. D. Hartogh and H. A. Tolhoek, Physica **24**, 721, 875, 896 1957).

suggests that we try to express $B_{12\ldots N}$ and $\mathfrak{B}_N$ in the form

$$B_{12\ldots N} = \exp G_{\mathrm{IY}} \tag{6}$$

and

$$\mathfrak{B}_N = \exp \mathfrak{G}_{\mathrm{AHT}}. \tag{7}$$

Since previous methods[2-5] for deriving G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$ are limited to the case in which $N \to \infty$, it has, in general, been impossible to apply the IY and AHT formalisms to problems in which N is reasonably small.

In this paper we exploit the relationships between the "factorized" cluster formalisms developed in I—which are immediately useful for all N—and the more traditional IY and AHT formalisms, using the former as crutches to extend the latter to finite systems. This goal is obtained in Sec. II, where expressions for G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$ which are valid for all N are generated. In Sec. III we focus our attention on the IY expansion, showing that G_{IY} is the sum of "linked products" of the IY normalized cluster integrals.

II. DERIVATION OF G_{IY} AND $\mathfrak{G}_{\mathrm{AHT}}$

The essential identity of the derivations of G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$ through the intermediacy of the two factorized cluster formalisms allows us to present them simultaneously. The explicit definitions of the subnormalization integrals $I_{j(1)\ldots j(n)}$ and $\tilde{I}_{j(i)}$ or their "unitarily invariant" counterparts $\mathfrak{I}_n$ and $\tilde{\mathfrak{I}}_1$ are of no consequence in the following manipulations. Indeed, only the relationships of the former with the cluster integrals $x_{j(p)\ldots j(q)}$ and $Y_{j(p)\ldots j(q)}$ and the latter with $\mathfrak{X}_p$ and $\mathfrak{Y}_p$ are needed in the forthcoming derivations of G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$. Let us review these relationships which were established in I. Each $I_{j(1)\ldots j(n)}$ may be written either as

$$I_{j(1)\ldots j(n)} = \left(\prod_{i=1}^{n} \tilde{I}_{j(i)}\right) B_{j(1)\ldots j(n)} \tag{8a}$$

or

$$I_{j(1)\ldots j(n)} = \left(\prod_{i=1}^{n} \tilde{I}_{j(i)}\right)\left(\prod_{p=1}^{n} Y_{j(p)}\right) \times \left(\prod_{1\le p<q\le n} Y_{j(p)j(q)}\right)\cdots Y_{j(1)\ldots j(n)}, \tag{8b}$$

where, in (8a),

$$B_{j(1)\ldots j(n)} = 1 + \sum_{p=1}^{n} \sum_{\langle l(1)\cdots l(p)\,|\,j(1)\cdots j(n)\rangle} \sum_{q=1}^{p} \sum_{\substack{\text{all partitions of } l(1)\cdots l(p) \\ \text{among } p-q+1 \text{ factors}}} x\ldots x\ldots\cdots x\ldots. \tag{9}$$

Similarly, the "invariant" $\mathfrak{I}_n$ may be decomposed either according to the prescription

$$\mathfrak{I}_n = (\tilde{\mathfrak{I}}_1)^n \mathfrak{B}_n \tag{10a}$$

or

$$\mathfrak{I}_n = (\tilde{\mathfrak{I}}_1)^n \prod_{k=1}^{n} (\mathfrak{Y}_k)^{kC_n}, \tag{10b}$$

where, in Eq. (10a),

$$\mathfrak{B}_n = 1 + \sum_{p=1}^{n} {}_pC_n \sum_{\substack{\text{all partitions} \\ \text{such that} \\ \sum_{b=1}^{p} b\nu_b = p}} \frac{p!\,\mathfrak{X}_1^{\nu_1}\cdots\mathfrak{X}_b^{\nu_b}\cdots\mathfrak{X}_p^{\nu_p}}{\prod_{b=1}^{p}(b!)^{\nu_b}\nu_b!}. \tag{11}$$

Clearly, Eqs. (1) and (3) are just particular cases of Eqs. (8a) and (10a), respectively, as are

$$I = I_{12\ldots N} = \left(\prod_{i=1}^{N} \tilde{I}_i\right)\left(\prod_{p=1}^{N} Y_p\right)\left(\prod_{1\le p<q\le N} Y_{pq}\right)\cdots Y_{12\ldots N} \tag{12}$$

and

$$I = \mathfrak{I}_N = (\tilde{\mathfrak{I}}_1)^N \prod_{k=1}^{N} (\mathfrak{Y}_k)^{kC_N}; \tag{13}$$

just particular cases of (8b) and (10b).

From Eqs. (1), (6), and (12) we see that

$$G_{\mathrm{IY}} = \ln B_{12\ldots N} = \sum_{p=1}^{N} \ln Y_p + \sum_{1\le p<q\le N} \ln Y_{pq} + \cdots + \ln Y_{12\ldots N}, \tag{14}$$

while Eqs. (3), (7), and (13) yield

$$\mathfrak{G}_{\mathrm{AHT}} = \ln \mathfrak{B}_N = \sum_{k=1}^{N} {}_kC_N \ln \mathfrak{Y}_k. \tag{15}$$

Our aim is, of course, to express G_{IY} and $\mathfrak{G}_{\mathrm{AHT}}$, respectively, in a series of the $x_{j(1)\ldots j(n)}$'s and a series of the $\mathfrak{X}_n$'s. $B_{12\ldots N}$, and thus $\ln B_{12\ldots N}$, is already known in terms of the normalized cluster integrals via (2). However, it is precisely this series which is unmanageable, and therefore all manipulations within the first line of (14) are questionable. On the other hand, the expansion comprising the second line of (14) appears to be of the intuitively desired form, but it remains to relate the $x_{j(1)\ldots j(n)}$ and the $Y_{k(1)\ldots k(p)}$. Completely analogous considerations apply to (15). Thus, our next task is to express each $Y_{k(1)\ldots k(p)}$ in terms of the $x_{j(1)\ldots j(n)}$ and each $\mathfrak{Y}_p$ in terms of the $\mathfrak{X}_n$.

We begin the attack upon this problem with the observation that equations similar to (14) and (15) hold among *all* the quantities defined in (8) or (10).

[5] F. Wu and E. Feenberg, Phys. Rev. **28**, 943 (1962).

Equating the logarithms of (8a) and (8b), we obtain

$$\ln B_{j(1)\cdots j(n)} = \sum_{\langle l(1)\,|\,j(1)\cdots j(n)\rangle} \ln Y_{l(1)} + \sum_{\langle l(1)l(2)\,|\,j(1)\cdots j(n)\rangle} \ln Y_{l(1)l(2)} + \cdots + \ln Y_{j(1)\cdots j(n)}; \tag{16}$$

a similar operation involving (10a) and (10b) yields

$$\ln \mathfrak{B}_n = \sum_{k=1}^{n} {}_kC_n \ln \mathfrak{Y}_k. \tag{17}$$

The low indexed $B_{j(1)\cdots j(n)}$ and $\mathfrak{B}_n$ contain a manageable number of terms even when the total number of particles is large. We shall thus express the $\ln Y_{j(1)\cdots j(p)}$ in (14) and the $\ln \mathfrak{Y}_p$ in Eq. (15) in terms of the $\ln B_{l(1)\cdots l(q)}$ and $\ln \mathfrak{B}_q$, respectively, and then, progressing from the low-indexed quantities, finally arrive at useful expressions for G_{IY} and $\mathfrak{S}_{\mathrm{AHT}}$ in terms of the $x_{j(1)\cdots j(n)}$ and the $\mathfrak{X}_n$ via Eqs. (9) and (11).

Recalling from I that the $\ln B_{l(1)\cdots l(q)}$'s are independent, we know there exists a unique inversion of Eq. (16). The correct formula for $\ln Y_{j(1)\cdots j(n)}$ in terms of the $\ln B_{l(1)\cdots l(p)}$, when used to eliminate the $\ln Y_{j(1)\cdots j(n)}$ from (16), must yield an identity. This criterion is satisfied by

$$\ln Y_{j(1)\cdots j(n)} = \sum_{p=1}^{n} (-1)^{n-p} \sum_{\langle l(1)\cdots l(p)\,|\,j(1)\cdots j(n)\rangle} \ln B_{l(1)\cdots l(p)}. \tag{18}$$

Similarly, we find that

$$\ln \mathfrak{Y}_n = \sum_{p=1}^{n} (-1)^{n-p} {}_pC_n \ln \mathfrak{B}_p. \tag{19}$$

Thus, by way of Eqs. (9) and (18) we have expressed $\ln Y_{j(1)\cdots j(n)}$ in terms of the x's, while Eqs. (11) and (19) yield $\ln \mathfrak{Y}_n$ in terms of the $\mathfrak{X}$'s.

We may elucidate the structure of Eqs. (18) and (19) by defining

$$B_0 = \mathfrak{B}_0 \equiv 1;$$

$$A^E_{j(1)\cdots j(n)} \equiv \left\{\prod_{p=1}^{\frac{1}{2}n}\left[\prod_{\langle l(1)\cdots l(2p)\,|\,j(1)\cdots j(n)\rangle} B_{l(1)\cdots l(2p)}\right]\right\} B_0, \quad \text{even } n,$$

$$\equiv \left\{\prod_{p=1}^{\frac{1}{2}(n-1)}\left[\prod_{\langle l(1)\cdots l(2p)\,|\,j(1)\cdots j(n)\rangle} B_{l(1)\cdots l(2p)}\right]\right\} B_0, \quad \text{odd } n,$$

$$\equiv \left\{\begin{array}{l}\text{product of all the } B\text{'s labeled by an even}\\ \text{number of indices which are subsets of}\\ \{j(1)\cdots j(n)\}\end{array}\right\} B_0$$

$$\equiv 1 + R^E_{j(1)\cdots j(n)};$$

$$A^O_{j(1)\cdots j(n)} \equiv \left\{\prod_{p=1}^{\frac{1}{2}n}\left[\prod_{\langle l(1)\cdots l(2p-1)\,|\,j(1)\cdots j(n)\rangle} B_{l(1)\cdots l(2p-1)}\right]\right\}, \quad \text{even } n,$$

$$\equiv \left\{\prod_{p=1}^{\frac{1}{2}(n+1)}\left[\prod_{\langle l(1)\cdots l(2p-1)\,|\,j(1)\cdots j(n)\rangle} B_{l(1)\cdots l(2p-1)}\right]\right\}, \quad \text{odd } n,$$

$$\equiv \left\{\begin{array}{l}\text{product of all the } B\text{'s labeled by an odd}\\ \text{number of indices which are subsets of}\\ \{j(1)\cdots j(n)\}\end{array}\right\}$$

$$\equiv 1 + R^O_{j(1)\cdots j(n)}; \tag{20a}$$

$$\mathcal{A}^E_n \equiv \left\{\prod_{p=1}^{\frac{1}{2}n} [\mathfrak{B}_{2p}]^{2pC_n}\right\} \mathfrak{B}_0, \quad \text{even } n,$$

$$\equiv \left\{\prod_{p=1}^{\frac{1}{2}(n-1)} [\mathfrak{B}_{2p}]^{2pC_n}\right\} \mathfrak{B}_0, \quad \text{odd } n,$$

$$\equiv \left\{\begin{array}{l}\text{product of all the even-indexed}\\ \mathfrak{B}_q\text{'s with } q \le n\end{array}\right\} \mathfrak{B}_0 \equiv 1 + \mathfrak{R}^E_n;$$

and

$$\mathcal{A}^O_n \equiv \left\{\prod_{p=1}^{\frac{1}{2}n} [\mathfrak{B}_{2p-1}]^{(2p-1)C_n}\right\}, \quad \text{even } n,$$

$$\equiv \left\{\prod_{p=1}^{\frac{1}{2}(n+1)} [\mathfrak{B}_{2p-1}]^{(2p-1)C_n}\right\}, \quad \text{odd } n,$$

$$\equiv \left\{\begin{array}{l}\text{product of all the odd-indexed}\\ \mathfrak{B}_q\text{'s with } q \le n\end{array}\right\} \equiv 1 + \mathfrak{R}^O_n. \tag{20b}$$

Now we can write Eq. (18) as

$$\ln Y_{j(1)\cdots j(n)} = (-1)^n\{\ln A^E_{j(1)\cdots j(n)} - \ln A^O_{j(1)\cdots j(n)}\},$$
$$= (-1)^n\{\ln (1 + R^E_{j(1)\cdots j(n)}) - \ln (1 + R^O_{j(1)\cdots j(n)})\} \tag{21}$$

and Eq. (19) as

$$\ln \mathfrak{Y}_n = (-1)^n\{\ln \mathcal{A}^E_n - \ln \mathcal{A}^O_n\},$$
$$= (-1)^n\{\ln (1 + \mathfrak{R}^E_n) - \ln (1 + \mathfrak{R}^O_n)\}. \tag{22}$$

Finally, anticipating that each $x_{j(1)\cdots j(n)} = O(1/N)$ so that $R^E_{j(1)\cdots j(n)}$ and $R^O_{j(1)\cdots j(n)}$ are both smaller than 1 [see Eq. (27)], we expand the logarithms appearing in Eq. (21) to find

$$\ln Y_{j(1)\cdots j(n)} = (-1)^n\left\{\sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} [(R^E_{j(1)\cdots j(n)})^p - (R^O_{j(1)\cdots j(n)})^p]\right\}.$$

Factoring each term in square brackets and removing

the common factor, we find

$$\begin{aligned}
\ln Y_{j(1)\cdots j(n)} &= (-1)^n\Big\{[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}] \\
&\quad\times\sum_{r=0}^{\infty}\frac{(-1)^r}{(r+1)}\sum_{q=0}^{r}(R^E_{j(1)\cdots j(n)})^q(R^O_{j(1)\cdots j(n)})^{r-q}\Big\}, \\
&= (-1)^n\Big\{[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}] \\
&\quad\times\sum_{q=0}^{\infty}\sum_{r=q}^{\infty}\frac{(-1)^r}{(r+1)}(R^E_{j(1)\cdots j(n)})^q(R^O_{j(1)\cdots j(n)})^{r-q}\Big\}, \\
&= (-1)^n[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}] \\
&\quad\times\Big\{\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{(-1)^{s+t}}{(s+t+1)}(R^E_{j(1)\cdots j(n)})^s(R^O_{j(1)\cdots j(n)})^t\Big\}.
\end{aligned} \tag{23}$$

The same manipulations performed on Eq. (22) give

$$\ln \mathcal{Y}_n = (-1)^n[\mathcal{R}^E_n - \mathcal{R}^O_n]\times\Big\{\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{(-1)^{s+t}}{(s+t+1)}(\mathcal{R}^E_n)^s(\mathcal{R}^O_n)^t\Big\}. \tag{24}$$

With the idea that the R's and $\mathcal{R}$'s are to be expressed in terms of the normalized cluster integrals, we thus propose for G_{IY} and $\mathcal{G}_{\mathrm{AHT}}$ the expansions

$$\begin{aligned}
G_{\mathrm{IY}} = \sum_{n=1}^{N}(-1)^n\sum_{\langle j(1)\cdots j(n)\rangle}\Big\{&[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}] \\
&\times\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{(-1)^{s+t}}{(s+t+1)}(R^E_{j(1)\cdots j(n)})^s(R^O_{j(1)\cdots j(n)})^t\Big\};
\end{aligned} \tag{25}$$

$$\mathcal{G}_{\mathrm{AHT}} = \sum_{n=1}^{N}{}_nC_N(-1)^n\Big\{[\mathcal{R}^E_n - \mathcal{R}^O_n]\times\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{(-1)^{s+t}}{(s+t+1)}(\mathcal{R}^E_n)^s(\mathcal{R}^O_n)^t\Big\}. \tag{26}$$

Truncation of these series at a low value of n is expected to yield valid approximations for the expectation value (5) in terms of the few-body cluster integrals. Moreover, for a given value of n, since the R's and the $\mathcal{R}$'s are smaller than 1, only the first few addends in the double infinite sums need to be kept. Indeed, in the many-body limit only the $s = t = 0$ term contributes to the expectation value of the Hamiltonian, the remaining terms being at least an order of magnitude greater in $1/N$. The expressions previously derived for G_{IY} and $\mathcal{G}_{\mathrm{AHT}}$ are thus recovered by taking only the $s = t = 0$ term for each value of n. We conclude this section with the explicit expressions for some of the low-indexed R's and $\mathcal{R}$'s in terms of the corresponding normalized cluster integrals:

$$\begin{aligned}
R^E_{j(1)} &= 0, \\
R^O_{j(1)} &= x_{j(1)}, \\
R^E_{j(1)j(2)} &= x_{j(1)} + x_{j(2)} + x_{j(1)}x_{j(2)} + x_{j(1)j(2)}, \\
R^O_{j(1)j(2)} &= x_{j(1)} + x_{j(2)} + x_{j(1)}x_{j(2)}, \\
&\vdots \\
\mathcal{R}^E_1 &= 0, \\
\mathcal{R}^O_1 &= \mathfrak{X}_1, \\
\mathcal{R}^E_2 &= \mathfrak{X}_1^2 + 2\mathfrak{X}_1 + \mathfrak{X}_2, \\
\mathcal{R}^O_2 &= \mathfrak{X}_1^2 + 2\mathfrak{X}_1. \\
&\vdots
\end{aligned} \tag{27}$$

III. LINKED CLUSTER EXPANSION OF G_{IY}

To facilitate the proof of the linked character of the addends contributing to the cluster expansion of G_{IY} we introduce the following definitions. The normalized cluster integrals $x_{j(1)\cdots j(n)}$ will be called *elements*. The addends comprising the expansion of the $R^E_{l(1)\cdots l(p)}$ or the $R^O_{l(1)\cdots l(p)}$ after Eq. (9) has been substituted for the B's will be called *configurations*. A configuration is, then, either a single element or the product of two or more elements. (In addition, the expansion of $B_{l(1)\cdots l(p)}$ includes the number 1.) Two elements are said to be *linked* if they have one or more indices in common. If a configuration of two or more elements can be factored into two proper subsets of elements such that no element of one subset is linked with any element of the other, the configuration is said to be *unlinked*. If no such factorization is possible, the configuration is said to be *linked* and will be called a *chain*. If, on the other extreme, no two elements are linked, the configuration is said to be *completely unlinked* and will be termed a *cell*. In certain instances, the terms chain and cell will be applied to configurations of one element. The extension of these terms to apply to a one-element configuration—which, in one sense, is not linked, but in another sense is not unlinked—is made for semantic reasons and should generate no real ambiguities. Finally, two chains are said to be *independent* if they have no index in common. Therefore, elements belonging to two independent chains must be unlinked.

We now examine the expansion (25) for G_{IY} in terms of the normalized cluster integrals, with a view to establishing the theorem that it is the sum of all

distinct single elements and, with appropriate numerical coefficients, all possible *linked configurations* of these elements. Consider a particular nth-order contribution:

$$[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}] \times \sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\frac{(-1)^{s+t}}{(s+t+1)}(R^E_{j(1)\cdots j(n)})^s(R^O_{j(1)\cdots j(n)})^t.$$

Our demonstration of the advertised "linked cluster theorem" rests on proving the assertion that the difference $[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}]$ contains only chains (linked configurations plus $x_{j(1)\cdots j(n)}$) having all n indices $j(1)\ldots j(n)$. Clearly, once this statement can be made for the factor $[R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}]$, it also applies to the entire nth-order contribution.

To prove the above assertion, we first note that

$$R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)} = A^E_{j(1)\cdots j(n)} - A^O_{j(1)\cdots j(n)} \tag{28}$$

and then set out to show the equality of the coefficients of a given unlinked configuration or a given chain with less than n indices arising from the product of B's in $A^E_{j(1)\cdots j(n)}$ and in $A^O_{j(1)\cdots j(n)}$. Let us define the *multiplicity* of a chain with respect to an $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$ as the total number of ways in which that chain may be formed from the x's in expanding the B's comprising that $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$. In addition to the term 1, each $B_{l(1)\cdots l(p)}$ consists of the sum of single elements whose indices form subsets of

$$\{l(1)\cdots l(p)\},$$

plus all possible completely unlinked products of these elements. Denoting the number 1 which occurs in $B_{l(1)\cdots l(p)}$ as a "trivial cell" and now extending the notion of "cell" to include one element configurations, we can describe $B_{l(1)\cdots l(p)}$ as the sum of all possible cells which include none, one, two, $\cdots$, all of the p indices $l(1)\cdots l(p)$. With this structure of each $B\ldots$ in mind, we observe that the coefficient of a given configuration (arising in the expansion of either $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$) is the *product* of the multiplicities (with respect to either $A^E_{j(1)\cdots (j)n}$ or $A^O_{j(1)\cdots j(n)}$) of all the *independent* chains comprising that configuration. Thus to find the coefficient of a given configuration we must first determine the multiplicity of each independent chain which composes it.

Since there are 2^{n-1} $B\ldots$ factors in both $A^E_{j(1)\cdots j(n)}$ and $A^O_{j(1)\cdots j(n)}$, a given chain may be said to arise as the product of 2^{n-1} cells (in the most inclusive sense), one from each of the 2^{n-1} factors. (Note that the trivial cell—i.e., the factor 1—must occur at least once in a chain arising from $A^E_{j(1)\cdots j(n)}$ since $B_0\ (=1)$ contains only this cell. Thus, there can never be more than $2^{n-1}-1$ nontrivial cells in the product forming such a chain.) Hence, to find the multiplicity of a given chain with respect to either $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$, we must first factor it into all possible combinations of cells, determine how many ways the given chain can arise for each such factorization, and then "sum" over all possible factorizations.

Consider some factorization of a given chain into ν nontrivial cells and $2^{n-1}-\nu$ factors of 1, the trivial cell. With columns labeled by the 2^{n-1} sets of indices on the $B\ldots$'s in $A^E_{j(n)\cdots j(n)}$ or $A^O_{j(n)\cdots j(n)}$ and rows labeled by the indices of the 2^{n-1} cells in this particular factorization (the symbol ϕ is to be associated with B_0 and the trivial cell), construct the square matrices $\Omega^E_{j(1)\cdots j(n)}$ and $\Omega^O_{j(1)\cdots j(n)}$ of order 2^{n-1} with elements prescribed by

$$\begin{aligned}\omega_{k(1)\cdots k(p);\phi l(1)\cdots l(q)} &= 1 \quad \text{if} \quad \{k(1)\cdots k(p)\} \subset \{\phi l(1)\cdots l(q)\},\\ &= 0 \quad \text{if} \quad \{k(1)\cdots k(p)\} \not\subset \{\phi l(1)\cdots l(q)\}.\end{aligned} \tag{29}$$

Here, the first set of (unlinked) indices corresponds with those of the cells labeling the rows and the second set with the $B\ldots$ indices labeling the columns; ϕ appears in each right-hand set of indices to denote the presence of the trivial cell in every $B\ldots$. It is not difficult to see what this scheme proposes: an element of the matrix so defined is either one or zero according to whether or not the cell corresponding to the set of row indices is to be found in the $B\ldots$ corresponding to the set of column indices. The permanent of the matrix $\Omega_{j(1)\cdots j(n)}$ (i.e., $\Omega^E_{j(1)\cdots j(n)}$ or $\Omega^O_{j(1)\cdots j(n)}$) divided by the factorials of the number of times a specific cell factor appears is the number of ways the given chain arises (in expanding $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$) for this particular factorization. Division by the factorials avoids counting contributions corresponding to permutations of identical cells among themselves.

We next prove that any chain with less than n different indices has the same multiplicity with respect to $A^E_{j(1)\cdots j(n)}$ as it does with respect to $A^O_{j(1)\cdots j(n)}$.

We already know that there can never be more than $2^{n-1}-1$ nontrivial cells in a factorization of any chain arising from the expansion of $A^E_{j(1)\cdots j(n)}$; we now assert that the same is true for any given chain with less than n different indices when it arises from the expansion of $A^O_{j(1)\cdots j(n)}$. Indeed, if there were 2^{n-1} nontrivial cells in the latter case, then all the one-indexed columns of $\Omega^O_{j(1)\cdots j(n)}$ denoted by indices absent from the chain would have zeros in every row. Thus the permanent would vanish, implying that the indicated factorization was, in fact, impossible.

Let us consider a particular factorization of a given chain with less than n indices into ν ($\nu < 2^{n-1}$) nontrivial cells and examine the matrices $\Omega^E_{j(1)\cdots j(n)}$ and $\Omega^O_{j(1)\cdots j(n)}$ which are associated with this particular factorization. There is clearly a one-to-one correspondence between the row indices of $\Omega^E_{j(1)\cdots j(n)}$ and those of $\Omega^O_{j(1)\cdots j(n)}$ since in both cases they are associated with the (nontrivial and trivial) cells into which the given chain has been factored. It is also possible to establish a one-to-one correspondence between the column indices of $\Omega^E_{j(1)\cdots j(n)}$ and the column indices of $\Omega^O_{j(1)\cdots j(n)}$ in the following manner: Choose one of the indices, say $j(a)$, which, although a member of the set $\{j(1)\cdots j(n)\}$, is not present on the chain under consideration. For every even-indexed $B_{\cdots}$ factor in $A^E_{j(1)\cdots j(n)}$, for instance $B_{l(1)\cdots l(2f)}$, the set of indices $\{l(1)\cdots l(2f)\}$ either contains the chosen index $j(a)$ or it does not. If $j(a) \in \{l(1)\cdots l(2f)\}$, associate with this even-indexed $B_{l(1)\cdots l(2f)}$ the odd-indexed $B_{\cdots}$ in $A^O_{j(1)\cdots j(n)}$ with the set of $2f-1$ indices formed by *omitting* $j(a)$ from $\{l(1)\cdots l(2f)\}$. On the other hand, if $j(a) \notin \{l(1)\cdots l(2f)\}$, then associate with the even-indexed $B_{l(1)\cdots l(2f)}$ the odd-indexed $B_{\cdots}$ with the set of $2f+1$ indices formed by *adding* $j(a)$ to the set $\{l(1)\cdots l(2f)\}$. Once $j(a)$ is chosen, the association between the factors in $A^E_{j(1)\cdots j(n)}$ and those in $A^O_{j(1)\cdots j(n)}$ is unique, and, hence, so is the correspondence between the column indices of $\Omega^E_{j(1)\cdots j(n)}$ and the column indices of $\Omega^O_{j(1)\cdots j(n)}$. Now, we may rearrange the columns of $\Omega^O_{j(1)\cdots j(n)}$ to form a new matrix $\Omega^{O'}_{j(1)\cdots j(n)}$ whose kth set of column indices ($k = 1, 2, \cdots, 2^{n-1}$) corresponds with the kth set of column indices of $\Omega^E_{j(1)\cdots j(n)}$. Finally, we note that the permanent of $\Omega^{O'}_{j(1)\cdots j(n)}$ is equal to the permanent of $\Omega^O_{j(1)\cdots j(n)}$ since the former is obtained from the latter just by column permutation; thus the prescription for enumerating the multiplicity of a given chain with less than n indices is unchanged upon substituting $\Omega^{O'}_{j(1)\cdots j(n)}$ for $\Omega^O_{j(1)\cdots j(n)}$.

Now we assert that the matrices $\Omega^E_{j(1)\cdots j(n)}$ and $\Omega^{O'}_{j(1)\cdots j(n)}$ are identical:

$$\omega^E_{r(1)\cdots r(s);\phi l(1)\cdots l(2f)} = \omega^{O'}_{r(1)\cdots r(s);\phi m(1)\cdots m(2g-1)}. \tag{30}$$

Here, the even set $\{l(1)\cdots l(2f)\}$ indexes the kth column of $\Omega^E_{j(1)\cdots j(n)}$, while the odd set

$$\{m(1)\cdots m(2g-1)\}$$

indexes the kth column of $\Omega^{O'}_{j(1)\cdots j(n)}$. To verify Eq. (30) observe that, since $j(a)$ cannot be a member of the set $\{r(1)\cdots r(s)\}$, if $\{r(1)\cdots r(s)\} \subset \{\phi, l(1)\cdots l(2f)\}$, then $\{r(1)\cdots r(s)\} \subset \{\phi m(1)\cdots m(2g-1)\}$; while if $\{r(1)\cdots r(s)\} \not\subset \{\phi l(1)\cdots l(2f)\}$, then $\{r(1)\cdots r(s)\} \not\subset \{\phi m(1)\cdots m(2g-1)\}$. The prescription for determining the matrix elements $\omega^E_{\cdots;\cdots}$ and $\omega^{O'}_{\cdots;\cdots}$ coupled with these observations establishes (30) and thereby the identity of $\Omega^E_{j(1)\cdots j(n)}$ and $\Omega^{O'}_{j(1)\cdots j(n)}$.

We see then, that, with respect to a particular factorization, a given chain with less than n indices arises in the same number of ways from $A^E_{j(1)\cdots j(n)}$ as it does from $A^O_{j(1)\cdots j(n)}$. Since, of course, all factorizations of such a chain occurring with respect to $A^E_{j(1)\cdots j(n)}$ also occur with respect to $A^O_{j(1)\cdots j(n)}$ and vice versa, the multiplicity of a chain with less than n different indices is the same with respect to $A^E_{j(1)\cdots(n)}$ as it is with respect to $A^O_{j(1)\cdots j(n)}$. Finally, recalling that the coefficient of a configuration arising in expanding $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$ is the product of the multiplicities of the independent chains which compose it, we find that only configurations with all n indices $j(1)\cdots j(n)$ present on a single chain remain upon forming the difference $A^E_{j(1)\cdots j(n)} - A^O_{j(1)\cdots j(n)}$. All other configurations—i.e., those having less than n indices and those composed of two or more independent chains, each with less than n indices—drop out in forming the difference, since their coefficients, upon arising from either $A^E_{j(1)\cdots j(n)}$ or $A^O_{j(1)\cdots j(n)}$, are the same.

Thus we have proven that $R^E_{j(1)\cdots j(n)} - R^O_{j(1)\cdots j(n)}$ is the sum of $x_{j(1)\cdots j(n)}$ and linked configurations with all n indices present on a single chain. Consequently, no unlinked configurations occur in the expansion (25) for G_{IY}.

ACKNOWLEDGMENT

One of the authors (J. W. C.) would like to thank Professor Aage Bohr for extending to him the hospitality of the Niels Bohr Institute, where part of this work was carried out.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Electromagnetic Scattering from an Inhomogeneous, Collisionless Plasma Cylinder

JOSEPH L. NEURINGER AND JAMES J. STEKERT
AVCO Space Systems Division, Research and Technology Laboratories, Wilmington, Massachusetts

(Received 28 March 1967)

An analytical investigation is presented of the coherent electromagnetic scattering at normal incidence from a quiescent, infinitely long, radially inhomogeneous, collisionless plasma cylinder. The inhomogeneity is characterized by an index of refraction $n = [1 - (\alpha/r)]^{\frac{1}{2}}$ such as would result from a line plasma source issuing plasma at a constant flow rate. Exact scattering solutions are obtained for both transverse electric and transverse magnetic polarizations of an incident plane wave. With k_0 and a representing the free-space wavenumber and cylinder radius, respectively, the partial radial wavefunctions and the scattered field are shown to depend markedly on the two nondimensional parameters β and η, where $\beta = k_0 a$ and $\eta = k_0 \alpha$. Numerical calculations of the differential scattering cross section for the TM mode show that forward scattering increases sharply with either increasing β or η although the increase is most sensitive to variations in β. The total scattering cross section is also shown to increase monotonically in the same fashion. Finally, it is observed that the scattering efficiency factor asymptotically approaches 2 in the limit as β and η both approach infinity, which is the same as that obtained for the homogeneous cylinder in the limit as β approaches infinity.

I. INTRODUCTION

It is the intent of this paper to present exact solutions to the coherent scattering of a normally incident plane electromagnetic wave by an infinitely long, radially inhomogeneous lossless plasma cylinder with a symmetric $1/r$-dependent particle density distribution. The plasma will be assumed to be quiescent, and density fluctuations about the mean distribution will be taken as unimportant. Two wave polarizations will be considered. The first assumes the electric (**E**) field to be parallel to the cylinder axis and gives rise to a transverse magnetic (TM) scattering system, while in the second, the magnetic (**H**) field is paraxial and gives rise to a transverse electric (TE) system. In the TM case, the solutions will prove to be exact and will be expressed as an expansion of partial waves in terms of known transcendental functions, the coefficients of which are the cylindrical analogies of those found by Mie for the homogeneous sphere. Solutions for the TE polarization will also be exact, but the partial waves will be expressed as power series instead of the more familiar functions of the TM expressions. Asymptotic and small argument expressions will be derived, and physical implications will be discussed. Computer calculations for the TM polarizations over a range of parameters will further exhibit the scattering properties.

II. THE WAVE EQUATIONS

Assuming all fields to be proportional to $e^{-i\omega t}$ and that no conduction currents, free charges or magnetic materials are present, Maxwell's equations can be written

$$\nabla \times \mathbf{E} = i\omega\mu_0\mathbf{H}, \tag{1a}$$

$$\nabla \times \mathbf{H} = -i\omega\epsilon\mathbf{E}, \tag{1b}$$

$$\nabla \cdot (\epsilon\mathbf{E}) = 0, \tag{1c}$$

$$\nabla \cdot \mathbf{H} = 0. \tag{1d}$$

Here, $\epsilon = \epsilon_0 n^2(\mathbf{r})$ is a function of position with $n(\mathbf{r})$ being the index of refraction of the propagation medium. Taking the curl of Eq. (1a) and using Eqs. (1b) and (1c) give the following wave equation for **E**:

$$(\nabla^2 + k_0^2 n^2)\mathbf{E} = -\nabla(\mathbf{E} \cdot \nabla n^2/n^2). \tag{2}$$

Taking the curl of Eq. (1b) and using Eqs. (1a) and (1d) results in the following wave equation for **H**:

$$(\nabla^2 + k_0^2 n^2)\mathbf{H} = -(\nabla n^2/n^2) \times (\nabla \times \mathbf{H}), \tag{3}$$

where $k_0 = [\omega^2\epsilon_0\mu_0]^{\frac{1}{2}}$ is the free-space wavenumber.

In the present case, the plasma of the cylinder is considered to be both collisionless and free from magnetic fields. Assuming neither the plasma ions nor the neutrals to have a direct effect on propagation, the index of refraction will be chosen so that

$$n^2 = 1 - Ne^2/\epsilon_0 m\omega^2 = 1 - (\alpha/r), \tag{4}$$

where N = electron number density, e = electron charge, and m = electron mass. Such a spatial dependence results from a line source which symmetrically issues plasma at a constant flow rate. The $1/r$ dependence will then be a property of N. It will be assumed that the plasma cylinder possesses a finite radius a and is illuminated from a great distance by a plane wave traveling in the $\hat{x}$ direction normal to the cylinder axis which is chosen to be the $\hat{z}$ axis as shown

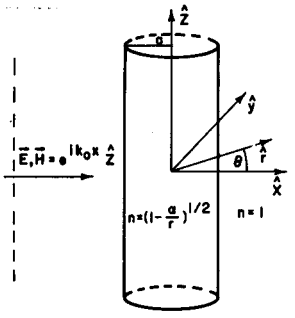

FIG. 1. Schematic diagram of plasma cylinder irradiated at normal incidence by a plane wave polarized in the axial direction.

in Fig. 1. Circular cylindrical coordinates will be used. The plane wave will first be given TM polarization and the electric field expanded into Bessel functions. Solutions will be found for Eq. (2), both inside and outside the cylinder, and tangential field components will be matched on the boundary. A similar procedure will be used with Eq. (3) for the TE polarizations.

III. TM POLARIZATION

Assume the incident plane wave to be of the form

$$\mathbf{E}_{z0} = e^{i(k_0 x - \omega t)}\hat{z}. \tag{5}$$

Equation (5) can be expanded in a Fourier–Bessel series[1] giving

$$\mathbf{E}_{z0} = \left[J_0(k_0 r) + 2\sum_{p=1}^{\infty} i^p J_p(k_0 r)\cos p\theta\right]\hat{z}, \tag{6}$$

where Eq. (6) follows from spatial symmetry and $e^{-i\omega t}$ has been surpressed.

Since $\mathbf{E}$ is everywhere normal to $\hat{r}$, $\mathbf{E}\cdot\nabla n^2 = 0$. Thus, Eq. (2) becomes

$$[\nabla^2 + k_0^2 n^2(r)]E_z = 0, \tag{7}$$

where $n = [1 - \alpha/r]^{\frac{1}{2}}$ for $r < a$ and $n = 1$ for $r > a$. Since there is no z dependence, one assumes that

$$E_z = R(r)\Theta(\theta). \tag{8}$$

Substituting Eq. (8) into Eq. (7) and using the method of separation of variables with requirements of single valuedness and symmetry in θ yields

$$E_z = \begin{pmatrix} c_0 \\ d_0 \end{pmatrix} R_0(r) + 2\sum_{p=1}^{\infty}\begin{pmatrix} c_p \\ d_p \end{pmatrix} i^p R_p(r)\cos p\theta, \tag{9}$$

where the coefficients c_p and d_p apply for $r < a$ and $r > a$, respectively. The functions $R_p(r)$ obey the equation

$$r^2\frac{d^2R_p}{dr^2} + r\frac{dR_p}{dr} + (k_0^2 r^2 - k_0^2\alpha r - p^2)R_p = 0. \tag{10}$$

When $r > a$, $\alpha = 0$ and Eq. (10) becomes Bessel's equation. Since a description of scattering is sought, solutions representing outgoing cylindrical waves are desired in the outer region. For the phase convention used here, such are provided by Hankel functions of the first kind. Thus, the scattered field can be written

$$E_{zs} = d_0 H_0^{(1)}(k_0 r) + 2\sum_{p=1}^{\infty} d_p i^p H_p^{(1)}(k_0 r)\cos p\theta, \quad r > a. \tag{11}$$

To find the radial solutions within the plasma cylinder, Eq. (10) can be transformed by means of the substitutions

$$R_p(r) = r^{-\frac{1}{2}}\Phi_p(r), \tag{12}$$

and

$$r = (1/i2k_0)\xi, \tag{13}$$

which results in

$$4\xi^2(d^2\Phi_p/d\xi^2) + (1 - 4p^2 + 2ik_0\alpha\xi - \xi^2)\Phi_p = 0. \tag{14}$$

Equation (14) is Whittaker's equation. Choosing only the solution which is regular at the origin yields[2]

$$\Phi_p(\xi) = (\text{const})\xi^{p+\frac{1}{2}}e^{-\frac{1}{2}\xi}M(\tfrac{1}{2} + p - \tfrac{1}{2}ik_0\alpha, 1 + 2p, \xi), \tag{15}$$

where $M(\gamma, \delta, z)$ is the confluent hypergeometric function defined by

$$M(\gamma, \delta, z) = 1 + \sum_{m=1}^{\infty}\frac{\gamma(\gamma+1)(\gamma+2)\cdots(\gamma+m-1)}{\delta(\delta+1)(\delta+2)\cdots(\delta+m-1)}\cdot\frac{z^m}{m!}. \tag{16}$$

Using Eqs. (12) and (13) and with a particular choice of the constant factor, Eq. (15) becomes

$$R_p(r) = (k_0 r)^p e^{-ik_0 r}M(\tfrac{1}{2} + p - \tfrac{1}{2}ik_0\alpha, 1 + 2p, i2k_0 r). \tag{17}$$

Using the identity[3]

$$M(p + \tfrac{1}{2}, 2p + 1, 2iz) = \Gamma(p + 1)e^{iz}(\tfrac{1}{2}z)^{-p}J_p(z),$$

it is seen that when $\alpha = 0$, $R_p(r)$ reduces to the Bessel function of the first kind of order p, as it must.

Thus, the field inside the plasma cylinder is

$$E_{zi} = c_0 e^{-ik_0 r}M(\tfrac{1}{2} - \tfrac{1}{2}ik_0\alpha, 1, i2k_0 r) + 2e^{-ik_0 r}\sum_{p=1}^{\infty} c_p i^p (k_0 r)^p \times M(\tfrac{1}{2} + p - \tfrac{1}{2}ik_0\alpha, 1 + 2p, i2k_0 r)\cos p\theta, \quad r < a. \tag{18}$$

The fields represented by Eqs. (6), (11), and (18) must obey boundary conditions for tangential fields

[1] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), Chap. 6, pp. 371, 372.

[2] E. D. Rainville, *Intermediate Differential Equations* (The Macmillan Company, New York, 1964), Chap. 11.

[3] *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, Eds. (U.S. Dept. of Commerce, Natl. Bur. Std., Appl. Math. Ser. 55, 1964), Chap. 13, p. 509, Formula 13.6.1.

at the cylindrical plasma surface. Thus at $r = a$,

$$J_p(\beta) + d_p H_p^{(1)}(\beta) = c_p R_p(\beta), \tag{19}$$

$$\left(\frac{dJ_p(k_0 r)}{d(k_0 r)}\right)_\beta + d_p\left(\frac{dH_p^{(1)}(k_0 r)}{d(k_0 r)}\right)_\beta = c_p\left(\frac{dR_p(k_0 r)}{d(k_0 r)}\right)_\beta, \tag{20}$$

where $\beta = k_0 a$. Defining $\eta = k_0\alpha$, $\gamma = \frac{1}{2} + p - \frac{1}{2}(i\eta)$, $\delta = 1 + 2p$ and using the recurrence relations

$$\frac{dJ_p(z)}{dz} = \frac{p}{z} J_p(z) - J_{p+1}(z), \tag{21}$$

$$\frac{dH_p^{(1)}(z)}{dz} = \frac{p}{z} H_p^{(1)}(z) - H_{p+1}^{(1)}(z), \tag{22}$$

and[4]

$$\frac{dM(\gamma, \delta, z)}{dz} = \frac{(\delta - 1)}{z}[M(\gamma, \delta - 1, z) - M(\gamma, \delta, z)], \tag{23}$$

the coefficients c_p and d_p from Eqs. (19) and (20) become

$$c_p = \frac{\beta^{1-p} e^{i\beta}[J_p(\beta)H_{p+1}^{(1)}(\beta) - J_{p+1}(\beta)H_p^{(1)}(\beta)]}{[\beta H_{p+1}^{(1)}(\beta) - (2p + i\beta)H_p^{(1)}(\beta)]M(\gamma, \delta, i2\beta) - 2pH_p^{(1)}(\beta)M(\gamma, \delta - 1, i2\beta)}, \tag{24}$$

$$d_p = \frac{2pJ_p(\beta)M(\gamma, \delta - 1, i2\beta) + [\beta J_{p+1}(\beta) - (2p + i\beta)J_p(\beta)]M(\gamma, \delta, i2\beta)}{[\beta H_{p+1}^{(1)}(\beta) - (2p + i\beta)H_p^{(1)}(\beta)]M(\gamma, \delta, i2\beta) - 2pH_p^{(1)}(\beta)M(\gamma, \delta - 1, i2\beta)}. \tag{25}$$

Equations (24) and (25) are the exact TM coefficients for the inhomogeneous cylinder, corresponding to the Mie coefficients for the homogeneous sphere. They are, however, rather cumbersome, and in the limits of large or small β, asymptotic and small argument approximations of simpler form can be found. When β is large,

$$J_p(\beta) \xrightarrow[\beta\to\infty]{} \left(\frac{2}{\pi\beta}\right)^{\frac{1}{2}} \cos\left(\beta - \frac{2p+1}{4}\pi\right), \tag{26}$$

$$H_p^{(1)}(\beta) \xrightarrow[\beta\to\infty]{} \left(\frac{2}{\pi\beta}\right)^{\frac{1}{2}} \exp\left[i\left(\beta - \frac{2p+1}{4}\pi\right)\right], \tag{27}$$

and[5]

$$M(\gamma, \delta, i2\beta) \xrightarrow[\beta\to\infty]{} (i2\beta)^{-\gamma}\left[\frac{\Gamma(\delta)}{\Gamma(\delta - \gamma)} e^{i\gamma\pi} + \frac{\Gamma(\delta)}{\Gamma(\gamma)} e^{i2\beta}(i2\beta)^{2\gamma-\delta}\right], \tag{28}$$

so that

$$c_p \xrightarrow[\beta\to\infty]{} \frac{2^p\Gamma(\gamma)}{\pi^{\frac{1}{2}} e^{(\eta/4)\pi} e^{i\frac{1}{2}[\eta\ln(2\beta) - 2(p+\frac{1}{2})\pi]}\{[\Gamma(\delta)\Gamma(\gamma)/\Gamma(\delta - \gamma)]e^{i(p+\frac{1}{2})\pi} + [\Gamma(\delta) - 2p\Gamma(\delta - 1)]e^{i[2\beta - \eta\ln(2\beta)]}\}}, \tag{29}$$

$$d_p \xrightarrow[\beta\to\infty]{} \frac{p\Gamma(\delta - 1)[e^{i[2\beta - \eta\ln(2\beta)]} - e^{-i[\eta\ln(2\beta) - (p-\frac{1}{2})\pi]}]}{\{[\Gamma(\delta)\Gamma(\gamma)/\Gamma(\delta - \gamma)]e^{i(p+\frac{1}{2})\pi} + [\Gamma(\delta) - 2p\Gamma(\delta - 1)]e^{i[2\beta - \eta\ln(2\beta)]}\}}$$
$$+ \frac{\frac{1}{2}\{[\Gamma(\delta)\Gamma(\gamma)/\Gamma(\delta - \gamma)]e^{-i[2\beta + (p+\frac{1}{2})\pi]} + \Gamma(\delta)e^{-i[\eta\ln(2\beta) - (p-\frac{1}{2})\pi]}\}}{\{[\Gamma(\delta)\Gamma(\gamma)/\Gamma(\delta - \gamma)]e^{i(p+\frac{1}{2})\pi} + [\Gamma(\delta) - 2p\Gamma(\delta - 1)]e^{i[2\beta - \eta\ln(2\beta)]}\}}. \tag{30}$$

Likewise, when β is small

$$J_p(\beta) \xrightarrow[\beta\to 0]{} \frac{1}{p!}(\tfrac{1}{2}\beta)^p, \tag{31}$$

$$H_0^{(1)}(\beta) \xrightarrow[\beta\to 0]{} i\frac{2}{\pi}\ln(\tfrac{1}{2}\beta), \tag{32}$$

$$H_p^{(1)}(\beta) \xrightarrow[\beta\to 0]{} -i\frac{(p-1)!}{\pi}(\tfrac{1}{2}\beta)^{-p} \quad (p \neq 0), \tag{33}$$

$$M(\gamma, \delta, i2\beta) \xrightarrow[\beta\to 0]{} 1, \tag{34}$$

which gives

$$c_p \xrightarrow[\beta\to 0]{} \frac{1}{2^p p!} e^{i\beta}, \tag{35}$$

$$d_p \xrightarrow[\beta\to 0]{} \frac{\pi\beta^{2p+1}}{2(2^p p!)^2}. \tag{36}$$

[4] Reference 3, Chap. 13, p. 507, Formula 13.4.13.
[5] Reference 3, Chap. 13, p. 508, Formula 13.5.1.

IV. TE POLARIZATION

As in the TM case, seeking a separable solution of the form

$$H_z = \bar{c}_0\bar{R}_0(r) + 2\sum_{p=1}^{\infty}\bar{c}_p i^p \bar{R}_p(r)\cos p\theta \tag{37}$$

and substituting into Eq. (3) yields the following ordinary differential equation for the pth radial wavefunction $\bar{R}_p(r)$:

$$r^2(r - \alpha)\frac{d^2\bar{R}_p}{dr^2} + (r^2 - 2\alpha r)\frac{d\bar{R}_p}{dr} + [k_0^2 r(r - \alpha)^2 - p^2(r - \alpha)]\bar{R}_p = 0. \tag{38}$$

Introducing the new independent variable $u = (\beta/a)r$ and remembering that $\alpha = \eta/k_0$ and $\beta = k_0 a$,

Eq. (38) becomes

$$u^2(u-\eta)\frac{d^2\bar{R}_p}{du^2}+u(u-2\eta)\frac{d\bar{R}_p}{du}+[u^3-2u^2\eta+(\eta^2-p^2)u+p^2\eta]\bar{R}_p=0. \quad (39)$$

Unlike the TM case, we have so far been unsuccessful in attempting to transform Eq. (39) into a differential equation whose solutions correspond to any of the well-known transcendental functions. Hence, we proceed to generate appropriate convergent power-series solutions corresponding to the two regular singular points of the differential equation. Equation (39) has regular singular points at $u=0$ and $u=\eta$ corresponding to $r=0$ and $r/a=\eta/\beta$, respectively. If our plasma is such that $\eta>\beta$, then the power-series solution generated about $u=0$, which is regular at $u=0$, will converge everywhere in the region $0\leq r\leq a$. Physically, the condition $\eta>\beta$ or $\alpha>a$ corresponds to an overdense plasma whose index of refraction is purely imaginary everywhere. If $\eta<\beta$, the above solution does not converge in the region beyond the next singular point $u=\eta$, and this solution cannot be evaluated at the plasma edge to satisfy boundary conditions.

For the case $\eta<\beta$, the method of attack is as follows: We first seek that power-series solution about $u=\eta$ which is regular at $u=\eta$. This solution involves one arbitrary constant and its radius of convergence is $|u|=\eta$. If the boundary, $u=\beta$, is such that $\eta<\beta<2\eta$, the solution so obtained may be used to match boundary conditions. If $\beta>2\eta$, a convergent solution valid at $u=\beta$ and involving one arbitrary constant is obtained by seeking a second power-series solution about the ordinary point $u=\beta$. The power series so obtained has a radius of convergence $|\beta-\eta|$ and involves two arbitrary constants (since $u=\beta$ is an ordinary point). The two expansions obtained above have the common region of convergence $\eta<u<2\eta$. Choosing a u_1 anywhere inside this region, we match the functions and their first derivatives at that point, leading to two algebraic equations for the three constants. One arbitrary constant, therefore, is involved in the power-series solution about $u=\beta$ and this solution is then used to match the solutions in the free-space region. Thus, it is observed that for all arbitrary values of η and β it requires, in general, three power series to generate exact solutions throughout the plasma cloud. Note that physically, the condition $\beta>\eta$ or $a>\alpha$ corresponds to a plasma whose index of refraction is purely imaginary in the region $0\leq r\leq\alpha$ and is positive real in the region $\alpha\leq r\leq a$.

A. Solution for $\eta>\beta$

To illustrate the form of the solution, we will fully treat the $\eta>\beta$ case only. The more complicated $\eta<\beta$ case can be handled similarly using the prescription outlined above. Expanding about $u=0$, we seek a solution of the form

$$\bar{R}_p=u^c\sum_{m=0}^{\infty}a_{mp}u^m, \quad (40)$$

and substituting into Eq. (39) leads to the following indicial equation:

$$p^2-2c-c(c-1)=0.$$

We choose the root $c=\frac{1}{2}[-1+(1+4p^2)^{\frac{1}{2}}]$ which guarantees that $\bar{R}_p$, for all orders p, is regular at the origin.

Setting the coefficients of like powers of u equal to zero yields

$$a_{0p}=a_{0p}, \quad (41a)$$

$$a_{1p}=\frac{1}{\eta}\left[\frac{a_{0p}[c^2+(\eta^2-p^2)]}{(c+1)(c+2)-p^2}\right], \quad (41b)$$

$$a_{2p}=\frac{1}{\eta}\left[\frac{2a_{0p}-a_{1p}[(\eta^2-p^2)+(c+1)^2]}{p^2-(c+2)(c+3)}\right], \quad (41c)$$

$$a_{3p}=\frac{1}{\eta}\left[\frac{a_{0p}-2\eta a_{1p}+a_{2p}[(\eta^2-p^2)+(c+2)^2]}{(c+3)(c+4)-p^2}\right], \quad (41d)$$

and, thereafter, the four-term recurrence relation

$$a_{mp}=\frac{1}{\eta}\left[\frac{a_{m-3p}-2a_{m-2p}+a_{m-1p}[(\eta^2-p^2)+(c+m-1)^2]}{(c+m)(c+m-1)-p^2}\right]. \quad (41e)$$

Hence, the fundamental solution of the radial wave equation (38) which is regular at the origin is

$$\bar{R}_p(r)=\left(\frac{\beta}{a}r\right)^{\frac{1}{2}[(1+4p^2)^{\frac{1}{2}}-1]}\sum_{m=0}^{\infty}a_{mp}\left(\frac{\beta}{a}r\right)^m, \quad (42)$$

where the a_{mp} are given by Eqs. (41a) through (41e).

B. The Partial Wave Coefficients $\bar{c}_p$ and $\bar{d}_p$

The incident and scattered fields H_{z0} and H_{zs} are given by Eqs. (6) and (11), respectively, where we replace d_p by $\bar{d}_p$. The continuity of H_z and E_θ at the plasma edge, $r = a$, using Eqs. (6), (11), and (37) yields

$$J_p(\beta) + \bar{d}_p H_p^{(1)}(\beta) = \bar{c}_p \bar{R}_p(\beta)$$

$$\left(\frac{dJ_p(k_0 r)}{d(k_0 r)}\right)_\beta + \bar{d}_p \left(\frac{dH_p^{(1)}(k_0 r)}{d(k_0 r)}\right)_\beta = \frac{\beta}{\beta - \eta} \bar{c}_p \left(\frac{d\bar{R}_p(k_0 r)}{d(k_0 r)}\right)_\beta,$$

where $E_{\theta i}$ in the medium, using Eq. (1b), is given by

$$E_{\theta i}(r) = \frac{1}{i\omega\epsilon_0} \frac{r}{r - \alpha} \frac{\partial H_{zi}}{\partial r}.$$

Solving for $\bar{c}_p$ and $\bar{d}_p$, yields

$$\bar{c}_p = \frac{(\beta - \eta)[J_p(\beta)(dH_p^{(1)}(k_0 r)/d(k_0 r))_\beta - H_p^{(1)}(\beta)(dJ_p(k_0 r)/d(k_0 r))_\beta]}{(\beta - \eta)\bar{R}_p(\beta)(dH_p^{(1)}(k_0 r)/d(k_0 r))_\beta - \beta H_p^{(1)}(\beta)(d\bar{R}_p(k_0 r)/d(k_0 r))_\beta}, \tag{43}$$

$$\bar{d}_p = \frac{\beta J_p(\beta)(d\bar{R}_p(k_0 r)/d(k_0 r))_\beta - (\beta - \eta)\bar{R}_p(\beta)(dJ_p(k_0 r)/d(k_0 r))_\beta}{(\beta - \eta)\bar{R}_p(\beta)(dH_p^{(1)}(k_0 r)/d(k_0 r))_\beta - \beta H_p^{(1)}(\beta)(d\bar{R}_p(k_0 r)/d(k_0 r))_\beta}. \tag{44}$$

The small-argument approximation is given by the leading term of Eq. (42):

$$\bar{R}_p(r) \underset{r\to 0}{\longrightarrow} \left(\frac{\beta}{a} r\right)^{\frac{1}{2}[(1+4p^2)^{\frac{1}{2}}-1]} \tag{45}$$

Using Eq. (45) together with the small-argument forms of Eqs. (31) to (34), Eqs. (42) and (44), in the small-argument limit, become

$$\bar{c}_p \underset{\beta\to 0}{\longrightarrow} \frac{\beta^{p-\frac{1}{2}[(1+4p^2)^{\frac{1}{2}}-1]}}{2^{p-1}p!}, \quad (p \neq 0), \tag{46}$$

$$\bar{c}_0 \underset{\beta\to 0}{\longrightarrow} 1, \tag{47}$$

$$\bar{d}_p \underset{\beta\to 0}{\longrightarrow} \frac{i\pi\beta^{2p}}{p[2^p(p-1)!]^2}, \quad (p \neq 0), \tag{48}$$

$$\bar{d}_0 \underset{\beta\to 0}{\longrightarrow} -\frac{i\pi}{4}\beta^2. \tag{49}$$

As a further point of interest for the case $\beta > \eta$, it can be shown, using the method outlined in Ref. 6, that the form of the asymptotic behavior of the solution of Eq. (39) is

$$\bar{R}_p(u) \underset{u\to\infty}{\longrightarrow} C_p \frac{(u-\eta)^{\frac{1}{2}}}{u} \times \cos[(u-\eta) - \tfrac{1}{2}\eta \ln(u-\eta) + D_p], \tag{50}$$

where the amplitude and phase C_p and D_p are constant parameters, independent of u and may depend in a complicated fashion on η and p.

[6] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1955), Chap. 3, pp. 91, 92.

V. DISCUSSION

Several observations can be made on the characteristics of the solutions $R_p(r)$ and $\bar{R}_p(r)$ presented in the last two sections. The solutions, as they must be, are real. Proof is given in Appendix A. Two forms are to be expected within the plasma cylinder. In the case of a homogeneous cylindrical plasma, the internal partial-wave solutions are the Bessel functions[7] $J_p(k_0 n r)$. The wavefunction is then sinusoidal in nature when $n^2 > 0$, indicating standing waves in the bounded region. When $n^2 < 0$, the solution is monotonically decreasing toward the center like $I_p(k_0 |n| r)$, indicating radial evanescence or a nonpropagating condition. Such is also to be expected here, modified by the inhomogeneity with the solution changing form at $r = \alpha$. It would also be supposed, inasmuch as $|n|$ becomes unbounded at the origin, that all fields disappear at $r = 0$. While this does occur for all partial waves with $p \geq 1$, $R_0(r)$ and $\bar{R}_0(r)$, strangely enough, are seen from Eqs. (17), (34), and (42) to be unity at the cylinder axis. Evidently, the effect of the symmetric wave and cylindrical geometry is sufficient to overcome the damping effect of the unbounded imaginary index of refraction.

As would be expected physically, the solutions do not exhibit any infinite scattering resonances at finite frequency (in the sense that the denominator of c_p and d_p vanishes for real positive values of η and β). The mathematical proof is given in Appendix B. Similarly, it is noticed that the small-argument forms

[7] H. C. Van de Hulst, *Light Scattering by Small Particles* (John Wiley & Sons, Inc., New York, 1957), Chap. 15.

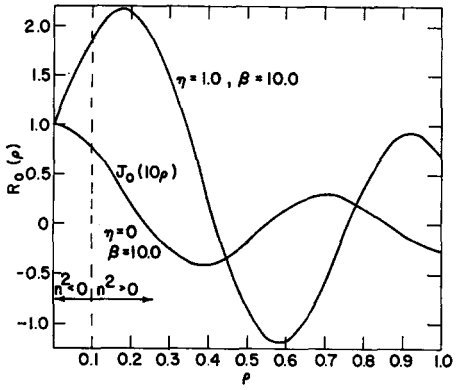

FIG. 2. Variation with dimensionless radial distance of zeroth-order radial wavefunction inside plasma for $\eta = 1.0$, $\beta = 10.0$, and comparison with $\eta = 0$, $\beta = 10.0$.

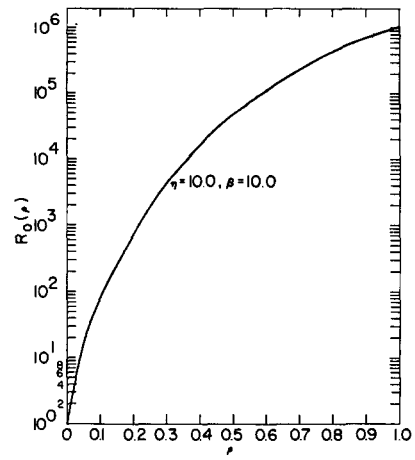

FIG. 3. Variation with dimensionless radial distance of zeroth-order radial wavefunction inside plasma for $\eta = 10.0$, $\beta = 10.0$.

of c_p and d_p are not functions of the plasma density parameter η and decrease rapidly with increasing partial-wave order p, while the asymptotic forms are just phase factors and functions of η at every order as found in more familiar cases.[8]

Finally, it should be observed that we have arbitrarily and somewhat unphysically truncated the plasma cylinder at $r = a$, leaving an abrupt discontinuity which adds to the scattering. The possibly unwanted effects of this calculational convenience can be made arbitrarily small simply by making the radius sufficiently large so that n approaches unity at the plasma edge.

VI. NUMERICAL RESULTS

By way of illustrating the nature of the scattering solutions, numerical calculations were made for the TM case. One is particularly interested in describing the behavior of partial radial wavefunctions inside the plasma cylinder as well as the behavior of the differential and total scattering cross sections as functions of the two plasma parameters η and β.

Letting $\rho = r/a$, $0 \leq \rho \leq 1$, and remembering that $\beta = k_0 a$ and $\eta = k_0 \alpha$, the radial-wave solutions of Eq. (17) become

$$R_p(\rho) = (\beta\rho)^p e^{-i\beta\rho} M(\tfrac{1}{2} + p - \tfrac{1}{2}i\eta,\ 1 + 2p,\ i2\beta\rho). \tag{51}$$

The behavior of the zeroth-order wavefunction $R_0(\rho)$ inside the plasma for two different (η, β) combinations is shown in Figs. 2 and 3. Since $n^2 = 1 - (\eta/\beta\rho)$, propagation cutoff occurs at $\rho = \eta/\beta$. Thus, if $\eta/\beta < 1$, sinusoidal-like wave behavior is to be expected in the region $\eta/\beta \leq \rho \leq 1$. If, however, $\eta/\beta \geq 1$, the index of refraction is everywhere purely imaginary inside the plasma cylinder and so we should expect a monotonically decreasing behavior from the plasma edge to the center for all $R_p(\rho)$ (see discussion in Sec. V). In Fig. 2, with $\eta = 1.0$ and $\beta = 10.0$, cutoff occurs at $\rho = 0.1$ and so wavelike behavior is exhibited in the region $0.1 < \rho < 1.0$. In Fig. 3, cutoff occurs at the plasma edge $\eta/\beta = 1$, so that $R_0(\rho)$ decreases monotonically with distance from the edge toward the center.

By way of comparison, with no plasma present, the $\eta = 0$, $\beta = 10.0$ solution, i.e., $J_0(10\rho)$, is shown in Fig. 2. It is seen that in the region $\eta/\beta \leq \rho \leq 1$, the solution $R_0(\rho)$ is similar in behavior to the corresponding Bessel function. In fact, for the (η, β) combination chosen, the (variable) wavelengths are almost of equal magnitude although the zeros of $R_0(\rho)$ are shifted to the right. Finally, from Eq. (53), it is to be noted that $R_0(\rho)$ like $J_0(\beta\rho)$ is the only partial wave that is different from zero at the origin.

The behavior of the first- and second-order partial wavefunctions with dimensionless radial distance are shown in Figs. 4 and 5, respectively. Again to be

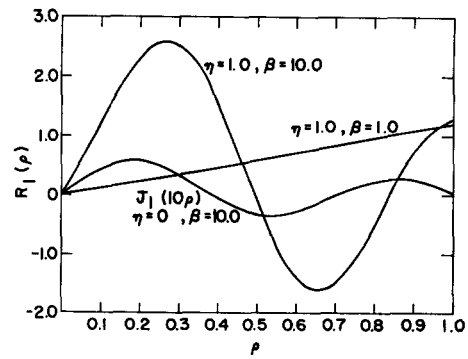

FIG. 4. Variation with dimensionless radial distance of first-order radial wavefunction inside plasma for the two cases $\beta = 1.0$ and 10.0 with $\eta = 1.0$ and comparison with $\eta = 0$, $\beta = 10.0$.

[8] Reference 1, Chap. 9, pp. 570–572.

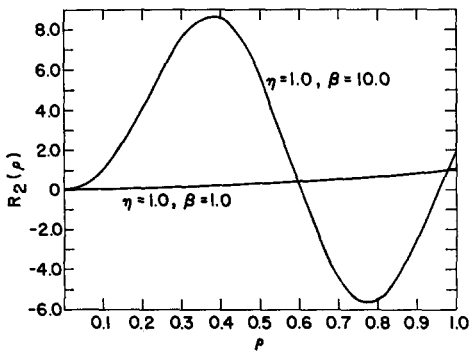

FIG. 5. Variation with dimensionless radial distance of second-order radial wavefunction inside plasma for the two cases $\beta = 1.0$ and 10.0 with $\eta = 1.0$.

noted is the marked difference in behavior depending on the magnitude of the ratio η/β. Again, for comparison purposes, the solution for $\eta = 0$, $\beta = 10.0$, i.e., $J_1(10\rho)$, is shown in Fig. 4.

Next, we examine numerically the behavior of the differential and total scattering cross sections of the plasma cylinder as functions of the plasma parameters η and β. The differential scattering cross section $\sigma(\theta)$ and the total scattering cross section Σ are given, respectively, by the following formulas[9]:

$$\sigma(\theta) = \frac{4}{k_0}\left| d_0 + 2\sum_{p=1}^{\infty} d_p \cos p\theta \right|^2, \tag{52}$$

$$\Sigma = \frac{2}{k_0}\left[2\,|d_0|^2 + 4\sum_{p=1}^{\infty}|d_p|^2\right], \tag{53}$$

where d_p is the pth partial wave scattering coefficient given by Eq. (25).

Polar diagrams showing the variation of the dimensionless differential scattering cross section $k_0\sigma(\theta)$ with θ for various (η, β) combinations is shown in Figs. 6–9. Figures 6 and 7 taken together compare the variation of $k_0\sigma(\theta)$ for $\beta = 1.0$ and 10.0, with fixed $\eta = 1.0$, while Figs. 8 and 9 taken together compare $k_0\sigma(\theta)$ for the same two values of β and for $\eta = 10.0$. It is seen that increasing β with η fixed causes a sharp increase in the forward scattering of the incident energy. If we compare figure groupings (6) and (8) with (7) and (9), [i.e., figures comparing the variation of $k_0\sigma(\theta)$ for two different values of η with β fixed] we see that while also intensifying the forward scattering, variations in η do not have as marked an effect on the pattern as do changes in β.

Figure 10 is a plot of the variation of the dimensionless total scattering cross section $k_0\Sigma$ with β for the three values of $\eta = 0.1$, 1.0, and 10.0. With η fixed, it is seen that the total scattering cross section increases monotonically as β increases with slight undulations evident in the curves for the two smaller values of η. The conventional efficiency factor for scattering is defined as

$$Q_{\text{sca}} = k_0\Sigma/2\beta. \tag{54}$$

It is then seen from Fig. 10 that as both η and β become large, Q_{sca} approaches the limit 2, which is just the value obtained for a homogeneous cylinder[10] in the limit as $\beta \to \infty$.

Figure 11 is a plot of the variation of the dimensionless total scattering cross section $k_0\Sigma$ with η for the two different values $\beta = 1.0$ and 10.0. It is seen that the total scattering cross section also increases monotonically with increasing η. Note again, for $\beta = 10.0$, that the efficiency factor asymptotically approaches 2 as η becomes large.

ACKNOWLEDGMENT

The authors wish to express their appreciation to D. Gootkind for his help in carrying out the numerical calculations.

APPENDIX A: REALITY OF THE RADIAL WAVEFUNCTIONS

The radial wavefunctions

$$R_p(r) = (k_0r)^p e^{-ik_0r} M(\tfrac{1}{2} + p - \tfrac{1}{2}ik_0\alpha,\ 1 + 2p,\ i2k_0r) \tag{A1}$$

are real.

Proof: The first factor of Eq. (A1), i.e., $(k_0r)^p$, is real. Hence, we need concern ourselves only with the product of the last two factors.

We make use of the following integral representations for the confluent hypergeometric function[11]:

$$M(\gamma, \delta, z) = \frac{\Gamma(\delta)}{\Gamma(\delta - \gamma)\Gamma(\gamma)}\int_0^1 e^{zt}t^{\gamma-1}(1-t)^{\delta-\gamma-1}\,dt, \tag{A2}$$

valid for $R_e(\delta) > R_e(\gamma) > 0$. In our case, $R_e(\delta) = 1 + 2p, R_e(\gamma) = p + \frac{1}{2}$ and since $p = 0, 1, 2, \cdots$ the above condition is satisfied. Thus, Eq. (A1) becomes

$$\frac{R_p(r)}{(k_0r)^p} = \frac{\Gamma(1+2p)}{\Gamma(\frac{1}{2} + p + \frac{1}{2}ik_0\alpha)\Gamma(\frac{1}{2} + p - \frac{1}{2}ik_0\alpha)} \times \int_0^1 e^{ik_0r[2t-1]}[t(1-t)]^{p-\frac{1}{2}}\left(\frac{t}{1-t}\right)^{-\frac{1}{2}ik_0\alpha} dt. \tag{A3}$$

Changing the variable of integration: $u = 1 - t$; $du = -dt$, we can write

$$\frac{R_p(r)}{(k_0r)^p} = \frac{\Gamma(1+2p)}{\Gamma(\frac{1}{2} + p + \frac{1}{2}ik_0\alpha)\Gamma(\frac{1}{2} + p - \frac{1}{2}ik_0\alpha)} \times \int_0^1 e^{-ik_0r[2u-1]}[u(1-u)]^{p-\frac{1}{2}}\left(\frac{u}{1-u}\right)^{\frac{1}{2}ik_0\alpha} du. \tag{A4}$$

[9] Reference 7, Chap. 15, pp. 301–303.

[10] Reference 7, Chap. 15, pp. 313–315.

[11] Reference 3, Chap. 13, p. 505, Formula 13.2.1.

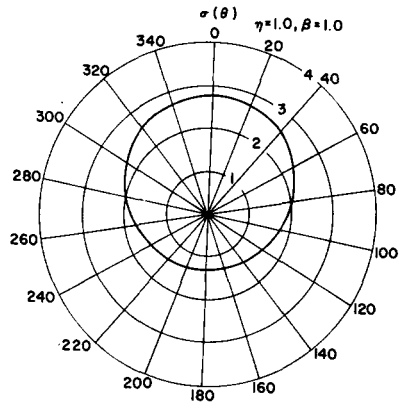

FIG. 6. Polar diagram of the dimensionless differential scattering cross section for $\eta = 1.0$, $\beta = 1.0$.

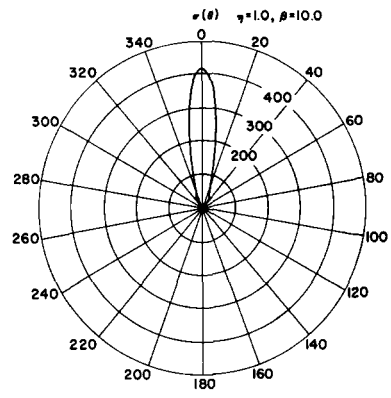

FIG. 7. Polar diagram of the dimensionless differential scattering cross section for $\eta = 1.0$, $\beta = 10.0$.

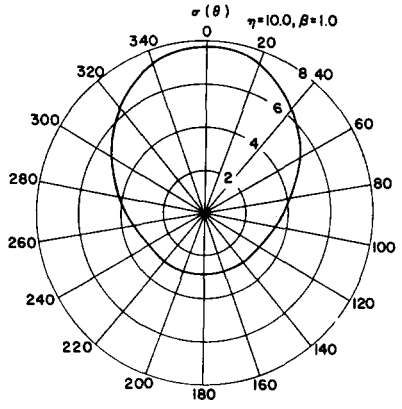

FIG. 8. Polar diagram of the dimensionless differential scattering cross section for $\eta = 10.0$, $\beta = 1.0$.

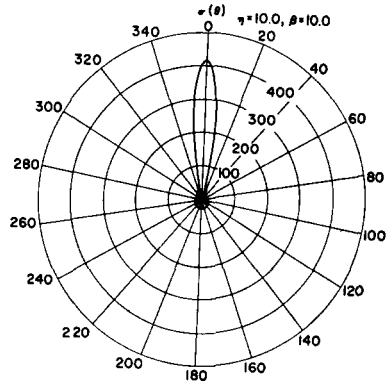

FIG. 9. Polar diagram of the dimensionless differential scattering cross section for $\eta = 10.0$, $\beta = 10.0$.

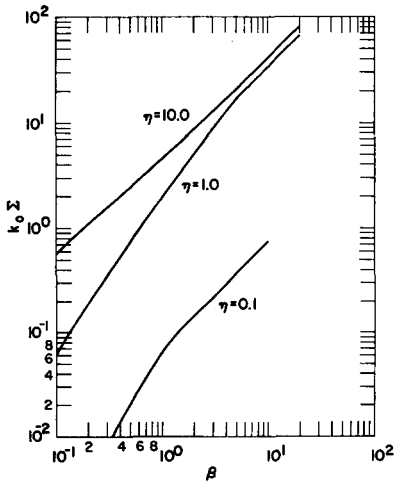

FIG. 10. Variation of the dimensionless total scattering cross section with β for the three η values 0.1, 1.0, and 10.0.

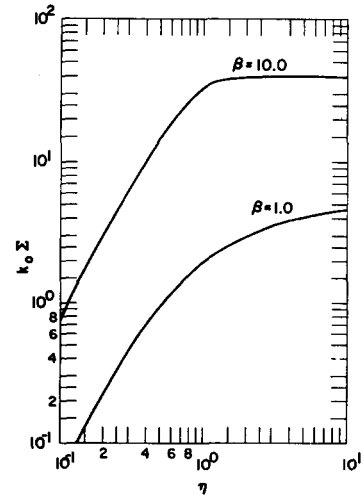

FIG. 11. Variation of the dimensionless total scattering cross section with η, for the two β values 1.0 and 10.0.

Comparing the right-hand sides of Eqs. (A3) and (A4), it is seen that they are complex conjugates of each other. Hence, R_p is real. Q.E.D.

APPENDIX B: ABSENCE OF INFINITE RESONANCES

The condition for resonance requires that there exist real positive values of η and β for which the denominators of the coefficients c_p and d_p go to zero. The denominator for both coefficients is given by

$$H_p^{(1)}(\beta)[dR_p(k_0r)/d(k_0r)]_\beta - H_p^{(1)\prime}(\beta)R_p(\beta) = 0, \quad \text{(B1)}$$

where the prime represents the derivative with respect to the argument.

We proved in Appendix A that the function $R_p(r)$ is real. Hence, using the definition of the Hankel function,

$$H_p^{(1)}(\beta) = J_p(\beta) + iY_p(\beta), \quad \text{(B2)}$$

substituting into Eq. (B1), and setting the real and imaginary parts equal to zero yields

$$J_p(\beta)\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta - J_p'(\beta)R_p(\beta) = 0, \quad \text{(B3)}$$

$$Y_p(\beta)\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta - Y_p'(\beta)R_p(\beta) = 0. \quad \text{(B4)}$$

Eliminating

$$R_p(\beta)\Big/\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta$$

from Eqs. (B3) and (B4) gives

$$J_p(\beta)Y_p'(\beta) - J_p'(\beta)Y_p(\beta) = 0 \quad \text{(B5)}$$

as the requirement for infinite resonance. However, the left-hand side of Eq. (B5) is the Wronskian[12] having the value of $2/\pi k_0 a$ which can never be zero for finite values of $k_0 a$.

[12] Reference 3, Chap. 9, p. 360, Formula 9.1.16.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Scattering of Massless Scalar Waves by a Schwarzschild "Singularity"*

RICHARD A. MATZNER†

Department of Physics and Astronomy, University of Maryland, College Park, Maryland,

and

Department of Applied Mathematics and Theoretical Physics, University of Cambridge

(Received 13 March 1967)

This paper investigates the scattering and absorption of scalar waves satisfying the equation $\phi^{;\mu}_{;\mu} = 0$ in the Schwarzschild metric. This problem has been previously considered by Hildreth. We find, for a Schwarzschild mass m, the following cross sections in the zero-frequency limit for s-waves: σ(absorption) $= 0$, $d\sigma/d\Omega \simeq [c + \frac{1}{2}(2m) \ln (2m\omega)]^2$, where c is a constant of order m. These results disagree with the previous calculation. We exhibit a method of solution for the equation. Its limiting (Newtonian) form, with suitable identification of the coefficients, is the problem of Coulomb scattering in nonrelativistic quantum mechanics. By demanding coordinate conditions which for large l allow the usual Coulomb results in a partial-wave expansion, we are able to define a partial-wave cross section. The (summed) differential cross section for small frequencies inherits the logarithmic behavior of the s-wave part, which is the only contribution explicitly calculated. (The $l \neq 0$ contributions and the behavior of the cross sections for $\omega \neq 0$ are qualitatively indicated.) Cosmological considerations are given which cut off this divergence.

I. INTRODUCTION AND SUMMARY

At one time the Schwarschild "singularity" at[1] $r = 2m$ was dismissed as unphysical, since it does not occur in the gravitational fields of normal stars. But as unexpected observations have more recently made speculations about possible peculiar astrophysical objects interesting, there has been an increased appreciation of the work of Oppenheimer and Snyder[2] which showed, not only that one could tolerate the actual existence of the Schwarzschild $r = 2m$ surface, but that there were even reasonable physical processes by which it could be produced.

* Supported in part by NASA Grant NSG-436, and by the Aerospace Research Laboratory, Office of Aerospace Research, U.S. Air Force.

† National Science Foundation Pre-doctoral Fellow 1963–1967.

[1] We take $G = c = 1$.

[2] J. R. Oppenheimer and H. Snyder, Phys. Rev. 56, 455 (1939).

Comparing the right-hand sides of Eqs. (A3) and (A4), it is seen that they are complex conjugates of each other. Hence, R_p is real. Q.E.D.

APPENDIX B: ABSENCE OF INFINITE RESONANCES

The condition for resonance requires that there exist real positive values of η and β for which the denominators of the coefficients c_p and d_p go to zero. The denominator for both coefficients is given by

$$H_p^{(1)}(\beta)[dR_p(k_0r)/d(k_0r)]_\beta - H_p^{(1)\prime}(\beta)R_p(\beta) = 0, \quad \text{(B1)}$$

where the prime represents the derivative with respect to the argument.

We proved in Appendix A that the function $R_p(r)$ is real. Hence, using the definition of the Hankel function,

$$H_p^{(1)}(\beta) = J_p(\beta) + iY_p(\beta), \quad \text{(B2)}$$

substituting into Eq. (B1), and setting the real and imaginary parts equal to zero yields

$$J_p(\beta)\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta - J_p'(\beta)R_p(\beta) = 0, \quad \text{(B3)}$$

$$Y_p(\beta)\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta - Y_p'(\beta)R_p(\beta) = 0. \quad \text{(B4)}$$

Eliminating

$$R_p(\beta)\Big/\left(\frac{dR_p(k_0r)}{d(k_0r)}\right)_\beta$$

from Eqs. (B3) and (B4) gives

$$J_p(\beta)Y_p'(\beta) - J_p'(\beta)Y_p(\beta) = 0 \quad \text{(B5)}$$

as the requirement for infinite resonance. However, the left-hand side of Eq. (B5) is the Wronskian[12] having the value of $2/\pi k_0 a$ which can never be zero for finite values of $k_0 a$.

[12] Reference 3, Chap. 9, p. 360, Formula 9.1.16.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Scattering of Massless Scalar Waves by a Schwarzschild "Singularity"*

RICHARD A. MATZNER†

Department of Physics and Astronomy, University of Maryland, College Park, Maryland,

and

Department of Applied Mathematics and Theoretical Physics, University of Cambridge

(Received 13 March 1967)

This paper investigates the scattering and absorption of scalar waves satisfying the equation $\phi^{;\mu}_{;\mu} = 0$ in the Schwarzschild metric. This problem has been previously considered by Hildreth. We find, for a Schwarzschild mass m, the following cross sections in the zero-frequency limit for s-waves: σ(absorption) $= 0$, $d\sigma/d\Omega \simeq [c + \frac{1}{2}(2m) \ln (2m\omega)]^2$, where c is a constant of order m. These results disagree with the previous calculation. We exhibit a method of solution for the equation. Its limiting (Newtonian) form, with suitable identification of the coefficients, is the problem of Coulomb scattering in nonrelativistic quantum mechanics. By demanding coordinate conditions which for large l allow the usual Coulomb results in a partial-wave expansion, we are able to define a partial-wave cross section. The (summed) differential cross section for small frequencies inherits the logarithmic behavior of the s-wave part, which is the only contribution explicitly calculated. (The $l \neq 0$ contributions and the behavior of the cross sections for $\omega \neq 0$ are qualitatively indicated.) Cosmological considerations are given which cut off this divergence.

I. INTRODUCTION AND SUMMARY

At one time the Schwarschild "singularity" at[1] $r = 2m$ was dismissed as unphysical, since it does not occur in the gravitational fields of normal stars. But as unexpected observations have more recently made speculations about possible peculiar astrophysical objects interesting, there has been an increased appreciation of the work of Oppenheimer and Snyder[2] which showed, not only that one could tolerate the actual existence of the Schwarzschild $r = 2m$ surface, but that there were even reasonable physical processes by which it could be produced.

* Supported in part by NASA Grant NSG-436, and by the Aerospace Research Laboratory, Office of Aerospace Research, U.S. Air Force.

† National Science Foundation Pre-doctoral Fellow 1963–1967.

[1] We take $G = c = 1$.

[2] J. R. Oppenheimer and H. Snyder, Phys. Rev. 56, 455 (1939).

(For some recent discussions see May and White[3] and Chap. 8 of Thorne's lectures.[4])

We refer to the remnant of such a process as an Oppenheimer–Snyder (OS) star, when some connotation of physical reality is not out of place, or as a Schwarzschild horizon, when it is more appropriate to suggest the mathematical idealizations we invoke in studying it. These idealizations include spherical symmetry, ignoring any angular momentum of the original matter, and the neglect of any matter expelled in the formation process, as well as of any orbiting debris with entrained magnetic fields, etc.

In this paper we illustrate by an example how an OS star is, like more normal final equilibrium states such as cooled white dwarfs and neutron stars, an object whose structure and whose response to various probes is independent of the detailed dynamical processes by which it is formed. In fact, we require no description whatsoever of the region $r < 2m$ beyond the assumption that all the matter responsible for the gravitational field has, some time in the past, fallen in beyond this $r = 2m$ Schwarzschild "horizon" leaving a spherically symmetric field. This assumption will be translated into a boundary condition at $r = 2m$.

With these idealizations, the exterior to a collapsed star is the Schwarzschild solution, and so the star's mass can be measured by observing planetary orbits or by doing bending of light experiments. In the latter case, Darwin[5] has shown that we cannot expect to probe more closely than $3m$ by means of high-frequency light, as light rays from infinity which reach smaller radii are trapped and never return to $r = \infty$. But with this limitation, we cannot tell with certainty whether the mass we observe is an OS star or, for instance, a neutron star with radius $2.5m$., A probe which gets closer in would be desirable, and we should investigate longer wavelength scattering. This will replace the corpuscular picture of Einstein's deflection-of-light calculation by a wave packet picture. We would expect the maximum amount of detail to be found from wave packets whose typical wavelength is approximately m. Lower frequencies would lose resolution; higher frequencies would be trapped if they probed close to $r = 2$m.

The problem that we actually discuss is a simplified version of that stated above. Instead of light, we calculate for a scalar field ϕ. And, although we state the problem for finite frequencies ω and for any angular momentum quantum number l, we will present a solution only for the $\omega \to 0$, s-wave limit. We seek cross sections for the scattering and absorption of scalar waves ϕ by a background Schwarzschild field.

The equation which determines the scattering is a generalization of the flat-space wave equation, namely

$$\phi^{;\mu}_{;\mu} = (-g)^{\frac{1}{2}} \frac{\partial}{\partial x^\mu} \left[(-g)^{\frac{1}{2}} g^{\mu\nu} \frac{\partial}{\partial x^\nu} \phi \right] = 0. \quad (1)$$

The metric entering this equation is taken to be the Schwarzschild metric, and it is its differing from the flat Lorentz metric which causes the scattering. By using the Schwarzschild metric, we have of course neglected any small gravitational field which the packet of ϕ waves itself might produce.

This same scattering problem has been considered previously by Hildreth[6] who thought of ϕ as an additional component of the gravitational field according to the Brans–Dicke[7] theory. The present restudy of this scattering problem obtains solutions of the basic differential equations somewhat more directly and leads to results which differ from Hildreth's

The radial equation is cast into the form of an effective potential equation by using the Regge–Wheeler[8] coordinate r^*. The boundary conditions at $r = 2m+$ are found. The difficulties of long-range forces, which are familiar from the problem of Coulomb scattering, are discussed. Although they have been extensively investigated in the literature, they pose the most difficult problems to be solved in the cross section calculation, particularly in terms of the coordinate invariance of the result. The separation into a "distorted plane wave" and outgoing waves is discussed. An ambiguity arises in the definition of the partial-wave scattering phase shifts. We fix the ambiguity in such a way that the usual "Coulomb" result is obtained for large l, so the large l partial cross sections contribute the usual large amount to small angle scattering (only).

We quote our results for the s wave $\omega \to 0$ limit:

$$\sigma(\text{absorption}) = 0,$$
$$d\sigma/d\Omega \simeq \left[c + \frac{2m}{3} \ln (2m\omega) \right]^2. \quad (2)$$

Here c is a constant of order m.

[3] M. M. May and R. H. White, Phys. Rev. **141**, 1232 (1966).
[4] K. S. Thorne, in *Les Houches Summer School Proceedings, 1966* (Gordon and Breach Science Publishers, London, 1967).
[5] Sir Charles Darwin, Proc. Roy. Soc. **A249**, 180 (1958).
[6] W. W. Hildreth, Ph.D. thesis, Princeton University (1964).
[7] R. H. Dicke, Phys. Rev. **125**, 2163 (1962), and earlier references there.
[8] T. Regge and J. A. Wheeler, Phys. Rev. **108**, 1063 (1957).

These disagree with Hildreth's results,[9] which are

$$\sigma(\text{absorption}) \simeq 12.5\,(2m)^2,$$

$$\frac{d\sigma}{d\Omega} - \frac{d\sigma}{d\Omega}\,\text{newtonian} \simeq 1.96\,(2m)^2.$$

The logarithmic divergence in Eq. (2) is cut off by the requirement that the observer be outside the near field zone, which requires $\omega \gtrsim 1/R_0$ where R_0 is, as a maximum, the Hubble radius. The cross section is then limited to $\sim(50\,m)^2$ for scatterers the mass of the sun. (In our $G = c = 1$ units, $m_{\text{sun}} = 1.5 \times 10^5$ cm.) However, this cross section is small enough so that it would contribute negligibly to the behavior of the scalar wave ϕ in a universe which contains many real stars whose radii are $\sim 10^{11}$ cm.

II. GENERAL CONSIDERATIONS FOR SCALAR SCATTERING

A. Radial Equation

With the usual Schwarzschild metric

$$ds^2 = dr^2\left(1 - \frac{2m}{r}\right)^{-1} + r^2\,d\Omega^2 - \left(1 - \frac{2m}{r}\right)dt^2$$

(where m is the mass of the scatterer), Eq. (1) is

$$r^{-2}\left(1 - \frac{2m}{r}\right)\frac{d}{dr}\left[r^2\left(1 - \frac{2m}{r}\right)\frac{dR}{dr}\right] - \left(1 - \frac{2m}{r}\right)\frac{l(l+1)}{r^2}R + \omega^2 R = 0, \quad (3)$$

where separation of variables has given

$$\phi = e^{-i\omega t}R(r)P_l(\theta). \quad (4)$$

If

$$R = u/r, \quad (5)$$

we find

$$\left(1 - \frac{2m}{r}\right)\frac{d}{dr}\left[\left(1 - \frac{2m}{r}\right)\frac{du}{dr}\right] + \left[\omega^2 - \left(1 - \frac{2m}{r}\right)\left(\frac{2m}{r^3} + \frac{l(l+1)}{r^2}\right)\right]u = 0. \quad (6)$$

Equation (6) is still not free from first derivatives du/dr.

However, the substitution[8]

$$dr^* = dr\left(1 - \frac{2m}{r}\right)^{-1}, \quad (7)$$

$$r^* = r + 2m\ln(r/2m - 1) + \text{const}, \quad (8)$$

puts Eq. (6) in the form

$$\left[\frac{d^2}{dr^{*2}} + \omega^2 - \left(1 - \frac{2m}{r}\right)\left(\frac{2m}{r^3} + \frac{l(l+1)}{r^2}\right)\right]u = 0. \quad (9)$$

[9] W. W. Hildreth, Ref. 6, p. 86. His B is $\frac{1}{2}m$, and we have factored out $(2m)^2$.

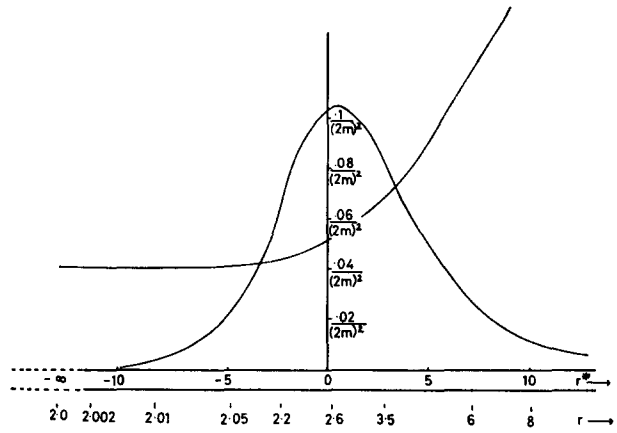

FIG. 1. The effective potential $V(r^*)$ plotted against a scale linear in r^*. (r and r^* are in units of m.) The constant in Eq. (8) has been set equal to zero. Also shown is the curve $u = r$ (not to scale), which is the solution to the s-wave radial equation for $\omega = 0$.

The problem has been reduced to a one-dimensional Schrödinger equation with independent variable r^*. This Schrödinger equation has an effective potential

$$V(r^*) = \left(1 - \frac{2m}{r}\right)\frac{2m}{r^3}, \quad (10)$$

with r considered a function of r^*.

The effective potential has been plotted in Fig. 1, taking the integration constant in r^* to be zero. Note that $V(r^*)$ has a rather low peak:

$$V_{\text{peak}} = 3\left(\frac{3}{16}\frac{1}{2m}\right)^2 \quad \text{at} \quad r = \tfrac{4}{3}(2m).$$

Furthermore, $V(r^*)$ is a finite range potential: it vanishes as $(r^*)^{-3}$ for $r^* \to +\infty$ and as $\exp(r^*/2m)$ for $r^* \to -\infty$ (corresponding to $r - 2m \to 0^+$). Therefore solutions of Eq. (9) can be obtained by familiar approximations.

B. Boundary Condition at $r = 2m$

We can write Eq. (9) in a dimensionless form by multiplying it by $(2m)^2$:

$$\left[\frac{d^2}{dy^{*2}} + q^2 - \left(1 - \frac{1}{y}\right)\frac{1}{y^3}\right]u = 0. \quad (11)$$

Here $y^* = r^*/2m$, $y = r/2m$, and $q = 2m\omega$. We set $l = 0$ since that is the case we investigate in Sec. III. (The boundary condition for $l \neq 0$ can be done in exactly the same way.) We set the integration constant in Eq. (8) equal to zero for convenience, and so

$$y^* = y + \ln(y - 1). \quad (12)$$

Now consider a well-defined pulse of scalar radiation of small but nonzero frequency, incident from $y^* = +\infty$, where $V(y^*) = 0$. This pulse will be partly reflected by the potential, so for $y^* \to +\infty$, we expect both ingoing and outgoing waves. Because the

potential disappears so rapidly for $y^* \to -\infty$, there is no backscatter from even small negative values of y^*. The pulse will be entirely ingoing for $y^* \to -\infty$, i.e., at the Schwarzschild horizon $r = 2m+$. We then take as the condition at $y^* = -\infty$ $(r - 2m = +0)$ that the waves be pure ingoing.

If we suppose the source of the field to be an OS star, this boundary condition means that there can be no scalar radiation from the star.

Since at $t \to \infty$ the surface of an OS star will approach arbitrarily close to $r = 2m$, any radiation from its surface will be red-shifted, so that its intensity goes to zero.[10] We saw above that the active region for scattering of the ϕ waves is $r \gtrsim \frac{4}{3}(2m)$, well away from the infinite red-shift surface. So the OS star, when "completely" collapsed, becomes a cold sink for radiation and satisfies our idealized boundary condition.

C. Comparison Coulomb Problem

Equation (9) has permitted us to find boundary conditions for $r - 2m = 0+$; but since r^* differs by a logarithm from r, it is not a good coordinate to use to interpret results at $r \to \infty$. We can give a statement of the problem which makes the asymptotic properties more transparent by making a different substitution in (3):

$$R = \frac{\bar{u}}{r\left(1 - \frac{2m}{r}\right)^{\frac{1}{2}}}, \tag{13}$$

which gives

$$\frac{d^2\bar{u}}{dr^2} + \left[\frac{\omega^2}{\left(1 - \frac{2m}{r}\right)^2} + \left(\frac{m}{r}\right)^2 \frac{1}{r^2\left(1 - \frac{2m}{r}\right)} - \frac{l(l+1)}{r^2}\frac{1}{\left(1 - \frac{2m}{r}\right)}\right]\bar{u} = 0. \tag{14}$$

As $r \to \infty$, Eq. (14) becomes

$$\frac{d^2\bar{u}}{dr^2} + \left[\omega^2 + \frac{4m\omega^2}{r} + \frac{12m^2\omega^2}{r^2} - \frac{l(l+1)}{r^2} + O(r^{-3})\right]\bar{u} = 0. \tag{15}$$

There is an attractive Newtonian coupling between the "energy density" $4\omega^2$ and the mass of the OS star.[11]

For large l, we expect the wavefunction solving Eqs. (9) and (14) to have small amplitude for a radius less than some distance of closest approach $r_\perp \sim [l(l+1)]^{\frac{1}{2}}/\omega$. The solution depends essentially only on the region $r > r_\perp$. Thus, for any ω, there is an $l_{\min}(\omega)$ such that $l \geq l_{\min}$ means the term $12m^2\omega^2/r^2$ and the terms $O(r^{-3})$ are negligible in Eq. (14). So for large enough l, we need only consider the solutions to

$$\frac{d^2\bar{u}}{dr^2} + \left[\omega^2 + \frac{4m\omega^2}{r} - \frac{l(l+1)}{r^2}\right]\bar{u} = 0, \qquad R = \frac{\bar{u}}{r}, \tag{14'}$$

which are satisfactory approximations to those of Eq. (9) or Eq. (14). The second of Eq. (14') follows from the assumption that the neglected terms in the first Eq. (14') are small.

We call Eq. (14') the "comparison Coulomb problem." We should investigate solutions to Eq. (9) for $l < l_{\min}$; but for $l \geq l_{\min}$, we may take the solutions given by Eq. (14'). This means considering the standard nonrelativistic wave mechanics problem of scattering of (negative) electrons from a nucleus of charge $+Ze$ (and infinite mass) with the substitutions

$$\hbar^2/2m_e \to 1, \quad Ze^2 \to 4m\omega^2.$$

We may immediately compute the "Newtonian" scattering [given by Eq. (14')] by making these substitutions into the Coulomb formula. In particular, we find the Coulomb "phase shift," $-\Omega_l(\omega)$:

$$-\Omega_l = -\arg \Gamma(l + 1 + 2im\omega).$$

For small ω and $l = 0$

$$\Omega_l \simeq 2im\omega\psi(1).$$

$\psi(x)$ is the logarithmic derivative of $\Gamma(x)$; $\psi(1) = 0.577\cdots$ (Euler's constant). The $l = 0$, $\omega \to 0$ "partial cross section" is then

$$\frac{d\sigma}{d\Omega} \cong \omega^{-2}\,|2im\omega\psi(1)|^2 \cong \tfrac{1}{3}(2m)^2.$$

The summed cross section is given by the Rutherford formula[12,13]

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{\sin^4 \theta/2}. \tag{16}$$

[10] C. W. Misner, January 1965 Relativity Meeting, Stevens Institute of Technology, Hoboken, N.J. (unpublished).

[11] This same $1/r$ coupling would be found by reducing Eq. (1) to Schroedinger form—free from first derivatives—in any coordinate system r' which goes as $r' = r + \text{const} + O(1/r)$. The $1/r^2$ terms in Eq. (14) are not invariant under such coordinate changes, but it is shown in Secs. IIE and IIIB below that the cross sections are invariant.

[12] H. Bethe and E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms* (Academic Press Inc., New York 1957), p. 31, Eq. (6.24).

[13] It is consistent to use the Rutherford formula for the geometrical optics case of deflection by the sun, where $2m \ll r_\perp$, since the dominant terms in the partial-wave expansion will come from the l values such that $l \sim r_\perp\omega$. Thus the term $l(l+1)/r^2$ completely dominates $12m^2\omega^2/r^2$ and $O(1/r^3)$ in Eq. (14), so Eq. (14') gives a correct result.

We note in passing that the Einstein deflection $\Delta\theta \simeq 4m/r_\perp$ gives a scattering cross section $d\sigma/d\Omega \simeq m^2(\theta/2)^{-4}$, in satisfactory agreement with Eq. (16) [we expect scalar and vector (light) waves to give the same answer in the geometrical optics limit] and justifying Eq. (14).

We have used the terms "phase shift" and "partial cross section" in the conventional sense, but it should be emphasized that for long-range forces, like the Newtonian gravitational force, these quantities are to a large extent really conventional, although the summed (over l) cross sections are well defined.

For short-range forces in ordinary quantum mechanics, the separation into partial waves and the separation into "plane incident" and scattered outgoing spherical waves is unambiguous. The boundary condition (usually square integrability at the origin) is also easily handled. Although the boundary condition in the ϕ-scattering problem was readily found in Sec. IIB above, the separation into "plane" and "scattered" waves and the application of the method of partial waves are more difficult. The difficulty arises partly because of the long-range force which distorts "plane" waves, even asymptotically, but also because of a need in relativity for invariance under coordinate transformation. These problems are dealt with in the next sections.

D. Form for the Asymptotic "Plane Wave"

A solution which satisfies Eq. (9), if we ignore terms in the equation which are $O(r^{*-3} \ln r^*)$ or smaller, is $r^* j_l(\omega r^*)$ where j_l is a spherical Bessel function with asymptotic behavior $j_l(\rho) \to \rho^{-1} \sin(\rho - \frac{1}{2}l\pi)$.

It also solves Eq. (14'), if *all* $O(r^{-2})$ terms are ignored. These solutions are significantly different from the solutions for short-range potentials, because the term $\ln r$ appears in the argument of j_l. In fact, it is well known[14] that for the comparison Coulomb Eq. (14'), a partial-wave solution has the asymptotic form

$$R_l(\text{Coulomb}) = \frac{\bar{u}}{r} \to r^{-1} \sin[\omega r + 2m\omega \ln(2\omega r) - \Omega_l(\omega) - l\pi/2]. \tag{17}$$

If we were presented with only the partial-wave expansion of this Coulomb solution, it would be impossible to apply straightforward partial-wave scattering theory because of the logarithmic term in Eq. (17), which apparently precludes an unambiguous asymptotic phase-shift determination. But the logarithm can be ignored, and it is the $\Omega_l(\omega)$ which determines the scattering. We can see this by noting the identity[14,15]

$$\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) = 2\delta(1-\cos\theta).$$

We use this identity in the definition of the scattering amplitude $f(\theta)$:

$$f(\theta) = (2i\omega)^{-1}\sum_{l=0}^{\infty}(2l+1)(e^{2i\delta_l}-1)P_l(\cos\theta), \tag{18}$$

to obtain, for $\theta \neq 0$:

$$f(\theta \neq 0) = (2i\omega)^{-1}\sum_{l=0}^{\infty}(2l+1)e^{2i\delta_l}P_l(\cos\theta). \tag{19}$$

Note that this holds only for the *summed* scattering amplitude. But in Eq. (19), it is clear that if δ_l is real, then additive terms in δ_l which are independent of l in the asymptotic region—even if they involve $\ln r$—change the value of $f(\theta \neq 0)$ by only a phase. They do not affect $|f(\theta \neq 0)|^2$, and we have, for instance,

$$\frac{d\sigma}{d\Omega}(\theta \neq 0) = f^*(\theta \neq 0)f(\theta \neq 0),$$
$$= (2\omega)^{-2}\,|\sum(2l+1)e^{-2i\Omega_l(\omega)}P_l(\cos\theta)|^2.$$

the scattering cross section for the equivalent Coulomb problem, Eq. (14').

We are dealing with a potential $\sim r^{-1}$ at large distances. [It is explicitly $1/r$ in Eq. (14) and the long-range character is included in the definition of r^*, and so appears when solutions of Eq. (9) are expressed as a function of r.] Both the forward-scattering amplitude and the total-scattering cross section (which diverge for r^{-n}, $n \ll 3$ and $n \ll 2$, respectively,[14]) are infinite. The only relevant quantity remaining to calculate or measure is the differential-scattering cross section. We thus lose nothing by modifying the forward behavior of $f(\theta)$, as was done in going from Eq. (18) to Eq. (19) above.

Ambiguities still exist in the definition of each particular phase shift, since we must decide which zero point to take for δ_l. (We must clearly take the *same* reference for every δ_l.) For short-range forces the phase shifts are also required to give the correct total cross section, which effectively fixes them so that $\lim_{l\to\infty}\delta_l = 0$, but this normalization does not hold for Coulomb quantum scattering.

[14] E. Landau and L. Lifschitz, *Quantum Mechanics, Non Relativistic Theory* (Pergamon Press, Inc., London 1958), para. 105, 106, and 112.

[15] This equation is true (the left side converges uniformly) except for $\cos\theta = -1$, where the convergence is only "in the mean." We point out that this lack of convergence in the backward direction is shared by the expansion for a plane wave, $\exp(ikr\cos\theta) = \sum i^l(2l+1)j_l(kr)P_l(\cos\theta)$.

We fix the constant in δ_l by demanding that the $\delta_l(\omega)$ approach $-\Omega_l(\omega)$, the phase shifts for the equivalent Coulomb problem (14′), for large l. (For fixed ω these are the partial-wave components which always stay far from the scattering center.) We do this by taking the integration constant in Eq. (8) to be $2m \ln (4m\omega)$, so

$$r^* = r + 2m \ln [2\omega(r - 2m)]. \tag{20}$$

As $r \to \infty$, $r^* \to r_c + O(m/r)$, where

$$r_c = r + 2m \ln (2\omega r)$$

is the combination appearing in the Coulomb partial-wave terms, Eq. (17). Thus, by the argument in Sec. IIC above, the phase shifts determined by Eq. (9), $\delta_l(\omega)$, tend to $-\Omega_l(\omega)$ as $l \to \infty$.

The method to use is (14′) and the known Coulomb phase shifts for the large l partial waves, and to apply Eq. (9) with the constant fixed by Eq. (20) to find the few small l phase shifts which will be significantly different, because of the effective potential $V(r^*)$, from the Coulomb ones. [We actually compute only one partial wave from Eq. (9), the $l = 0$, $\omega \to 0$ case.]

It is easy to show that this choice of constant (20) and this type of partial-wave manipulation is equivalent to splitting the total solution ϕ into

$$\phi_{\text{"plane"}} = \frac{r^*}{r} e^{i\omega r^* \cos\theta} \to (\omega r)^{-1} \sum i^l (2l + 1) P_l(\cos\theta) \times \sin(\omega r^* - l\pi/2), \tag{21a}$$

$$\phi_{\text{"out"}} = \phi - \phi_{\text{"plane"}} \to (2i\omega)^{-1} r^{-1} e^{i\omega r^*} \times \sum (2l + 1)(e^{2i\delta_l} - 1) P_l(\cos\theta), \tag{21b}$$

where the $\delta_l(\omega)$ are obtained by the usual partial-wave manipulation of the radial equation, treating r^* as an ordinary radial variable at infinity. Note that each term in the sums of Eq. (21) is an asymptotic solution to Eq. (9) and Eq. (14).[16]

E. Coordinate Invariance of the Scattering Cross Section

The total (incident + outgoing) solution to Eq. (9) or (14) for each partial wave has an asymptotic form like

$$u_l \to \sin(\omega r^* - l\pi/2 + \delta_l). \tag{22}$$

Since adding a constant to r^* amounts to adding a constant to each of the phase shifts, the summed differential cross section is independent of the choice of the constant in r^*. Similarly, coordinate changes $r' = r + \text{const} + O(r^{-1})$ add only an l-independent constant to the phase shift in the asymptotic region, and so the summed differential cross section is invariant under them. Also, the asymptotic value ($|\omega|$ in the z direction) of the current given by the "plane" wave part Eq. (21a) of the total solution is invariant both under $r' = r + \text{const} + O(r^{-1})$ and under additive constants in r^*, as can be verified by direct calculation.

We emphasize that the constant in Eq. (20) is irrelevant for determining a summed-scattering cross section, but the choice of (20) allows us to compute only a few phase shifts and then rely on tabulated Coulomb results for the higher l values.

III. CALCULATION OF CROSS SECTIONS

A. Exact solution for $\omega = 0$, $l = 0$

We write Eq. (9) for $l = 0$, $\omega = 0$:

$$\left(1 - \frac{2m}{r}\right)\left[\frac{d}{dr}\left(1 - \frac{2m}{r}\right)\frac{d}{dr} - \frac{2m}{r^3}\right]u = 0.$$

This equation has the solution $u = r$, as may be verified by direct substitution. We make use of this solution in the following sections as a starting point for our discussion of small ω behavior. It is especially useful since, as shown below, in the limit $\omega = 0$ this solution is exactly the scattering solution that we require (i.e., it solves the equation and fits the boundary condition.)

B. Calculation at Low Energy

It is shown in the Appendix that the scattering length approximation is valid for the $l = 0$, $\omega \to 0$ case which we now consider.

The ingoing wave condition at $y^* = -\infty$ corresponds (in the scattering length approximation) to

$$u = 1 - iqy^* + O(q^2 y^{*2}), \qquad \frac{du}{dy^*} = -iq + O(q^2 y^{*2}), \tag{23}$$

for $y^* \to -\infty$ with $qy^* \ll 1$. [The potential falls off exponentially in this region, so this ingoing condition is well satisfied for y^* even slightly to the inside of the peak of $V(y^*)$.] As $q \to 0$, these become

$$u \simeq 1, \quad du/dy^* \simeq 0.$$

The $q = 0$ wavefunction, $u(y^*) = y$, fits smoothly onto the ingoing wave Eq. (23) with errors of order q (see Appendix). We thus take ingoing free waves in the region inside the peak in $V(y^*)$, match them smoothly to this solution $u = y$, which is valid in the region where the potential dominates, and then match (by a scattering-length approximation) to a combination of

[16] An objection might be raised because the "plane wave," Eq. (21a), does not have a divergenceless current, even in the asymptotic region where $r^* \to r_c$. This characteristic is shared, however, by the "plane" part of the total Coulomb solution (see Ref. 14, p. 419). The current has the numerical value $|\omega|$ in the $z = r\cos\theta$ direction at spatial infinity.

ingoing and outgoing free waves on the outside of the peak in $V(y^*)$. In the $q = 0$ limit, we may approximate the free waves outside by a straight line, just as we approximated those on the inside by a straight line (of zero slope).

We make the outside match at some matching point. The wavefunction must have continuous amplitude and slope.

$y_0 \equiv$ (Solution for potential region at match)

$=$ (solution for free region at match) $\equiv A(y_0^* + b)$. (24)

A and b are constants to be determined; the subscript zero means "at the joining point." For $y^* \gg 1$, $V(y^*) \sim y^{*-3}$, and the joining point will be the point where $V(y_0) \simeq h^{-3}q^2$ for some constant h of order unity. We may write

$$y_0 \equiv hy_{+TP} \simeq hq^{-\frac{2}{3}}. \tag{25}$$

Here y_{+TP} is the outside turning point:

$$q^2 = y_{+TP}^{-3}(1 - y_{+TP}^{-1}).$$

Matching slopes, we have

$$A \equiv \frac{dy}{dy^*}\bigg|_0 = 1 - y_0^{-1}, \tag{26}$$

$$b = y_0(1 - y_0^{-1})^{-1} - y_0^*. \tag{27}$$

The wavefunction u may be written (for large y_0)

$$u \simeq (1 - y_0^{-1})(y^* + y_0 + 1 - y_0^*) \tag{24a}$$

and is clearly independent of additive constants in y^*.

The explicit form for y^* which was chosen in Eq. (20) is

$$y^* = y + \ln [2q(y - 1)].$$

Substituting this into Eq. (27), we find (in the limit $q \to 0$ so $y_0 \gg 1$)

$$b \simeq 1 - \ln [2q(y_0 - 1)].$$

From Eq. (25) above, we have

$$b \simeq 1 - \ln [2q(hq^{-\frac{2}{3}})] = 1 - \ln 2h - \tfrac{1}{3} \ln q. \tag{28}$$

We note that Eq. (28) contains a term proportional to $\ln q$. This term cannot be subtracted out of the scattering length b, as constants in y^* can be, because it is a specifically small l term. As mentioned in Sec. IIID below, the higher angular momentum components are kept from feeling the effective potential $V(y^*)$ by the centrifugal terms $l(l + 1)/r$. Only the small l solutions can feel the y^{*-3} potential which gives the logarithmic shift. [In the "starred" coordinate, the wavefunction in the potential region is suffering a logarithmic phase shift compared to the "free" wave $\sim \sin qy^*$. The difference between the inner solution and the comparison free wave increases as the turning point moves out ($y_{+TP} \simeq hq^{-\frac{2}{3}}$) so the value of this logarithmic shift increases.]

Since we argue that only the $l = 0$ term will be significantly different from the Coulomb result, we conclude there must be a logarithmic term in the summed differential cross section.

The absorption cross section is particularly easy to determine. On the inside of the potential bump, the wavefunction u has unit amplitude and its imaginary part $-i \sin qy^*$ is approximately zero. The wavefunction on the outside, however, is

$$u \simeq y^* \simeq q^{-1} \sin qy^*.$$

We see that the ratio of the amplitudes, outside to in, is q^{-1}; so if the incident wave (outside) were normalized, we would have no flux on the inside of the potential barrier as $q \to 0$. We may immediately conclude that *there is no absorption in the s wave, $\omega \to 0$ limit.* From the way this result was obtained, it is clearly coordinate invariant.

An observer who is monitoring extremely long wavelength scattering may, in fact, be within the effective potential region. This observer will see the solution $u = r$. This corresponds to $R = r/r = \text{const}$, i.e., $\phi(S \text{ wave}) = \text{constant}$ in space. (Of course there will be the slow time variation also.) The observer then sees no reason to prefer the origin over any other point. He would say that the scattering cross section is zero, since otherwise the location of the massive scatterer would have been distinguished by the behavior of the field.

As an example, if the orbit of the earth ($\sim 1.5 \times 10^8$ km; $y_E = r_E/2m. \sim 0.5 \times 10^8$) is within the region of the effective potential due to the sun, we must have

$$y_E^{\frac{3}{2}} \sim q^{-1} \sim \text{wavelength}/2m. \sim 0.3 \times 10^{12},$$

so

$$\text{wavelength} \sim 10^{12} \text{ km},$$

$$\text{period} \sim 3 \times 10^6 \text{ sec},$$

$$\sim \text{month}.$$

Periods of this order or longer are of cosmological interest.

C. Cutoff for the Divergent Zero-Energy Cross Section

The last paragraphs of Sec. IIIB have shown (a) that an observer in the free wave zone would note a logarithmically diverging scattering length, but (b) an observer within the potential-dominated region would not be able to see any scattering at all.

We have a scattering cross section

$$\frac{d\sigma}{d\Omega} = b^2 \simeq \left[2am + \frac{2m}{3} \ln (2\omega m)\right]^2,$$

where a is a constant of order unity. We see that as

$\omega \to 0$, the scattering length becomes infinite. However, the size of the universe gives a cutoff for this infinity, and the cutoff is rather small.

Consider the case of an OS star of one solar mass. Since $\omega_{\min} \sim 1/R_H$ where R_H = "Radius of the universe" $\sim 10^{28}$ cm (the Hubble radius),

$$b \simeq 2m[a + \tfrac{1}{3} \ln (10^5 \times 10^{-28})],$$
$$\simeq 2m[a - 23] \sim 50 \text{ km}.$$

In this case there is nothing like an infinity. The longest scattering length occurs for $m \sim R_H$, in which case the scattering length $b \sim R_H$, also. The scattering length decreases monotonically for decreasing mass of the scatterers.

D. Higher Energy and Angular Momentum

The results given above hold only for the S-wave case in the zero-frequency limit. If S waves with ω^2 greater than

$$V_{\text{peak}} = 3\left(\frac{3}{16} \cdot \frac{1}{2m}\right)^2 = \frac{0.105}{(2m)^2}$$

are incident, they will be little affected by the bump, and so the absorption cross section would go up and the scattering becomes more "Newtonian" for s waves.

Since the effective potential is a close-in effect ($r_{\text{peak}} = \frac{4}{3}2m$), the higher angular momenta will be prevented by the repulsive centrifugal potential from feeling it. Recall the definition of a "distance of closest approach," $r_\perp = [l(l+1)]^{\frac{1}{2}}/\omega$. For any frequency low enough to feel the potential ($\omega \ll \frac{1}{2}m$), we have $r_\perp \geq 2m[l(l+1)]^{\frac{1}{2}}$. For $l = 1$ this is already outside, though near, the peak of the effective repulsive potential. Lower-frequency waves (for $l \neq 0$) will remain farther away and won't be affected by the effective potential, while higher-frequency waves, though they can get in to the peak of the effective potential, are not much affected by it. Thus we conclude, as Hildreth did by a similar argument, that the higher angular momentum components suffer essentially Newtonian scattering. However, there is a range ($\omega \sim \frac{1}{2}m$) where the $l = 1$ waves are somewhat affected by the non-Newtonian aspects of the scattering.

ACKNOWLEDGMENT

The author would like to thank Professor C. W. Misner for suggesting this problem and for his invaluable suggestions and very patient advice throughout this work.

APPENDIX: MATCHING BETWEEN FREE AND POTENTIAL REGIONS

Because $V(y^*) \sim \exp(y^*)$ for $y^* \to -\infty$, the potential falls off very quickly inside the peak, and so the linear approximation can be used ($e^{-iqy^*} \simeq 1 - iqy^*$). The turning point y^*_{-TP}, where the potential becomes negligible, occurs at a negative y^* which is much smaller than a wavelength ($\sim q^{-1}$).

Similarly, on the outside of the potential bump, the wavelength is $\sim q^{-1}$, while the turning point goes as $q^{-\frac{2}{3}}$. Thus the linear approximation through the turning point region is valid at both sides of the potential region.

Now consider the match at the inside of the potential. We assume the matching is done exactly at y^*_{-TP}, and note that the interior wave has $u = y$ and

$$y_{-TP} \simeq 1 + q^2,$$
$$\left.\frac{dy}{dy^*}\right|_{-TP} = (1 - y^{-1}_{-TP}) \simeq q^2.$$

The ingoing free wave has $u = \exp(-iqy^*)$:

$$u(y^*_{-TP}) \simeq 1 - iqy^*_{-TP},$$
$$\left.\frac{du}{dy^*}\right|_{-TP} \equiv u'_{-TP} \simeq -iq.$$

For sufficiently small q, the errors in joining these functions become small with q. Thus the solution $u = y$ is the correct continuation, in the potential region, of the ingoing wave determined by the boundary condition. We can see that errors in the match described here do not propagate and accumulate to cause large errors in the wavefunction for $y^* \to +\infty$. For we can write the relevant integral equations:

$$u'(x^*) = \int_{y^*-TP}^{x^*} Vu\, dy^* - \int_{y^*-TP}^{x^*} uq^2\, dy^* + u'(y^*_{-TP}), \tag{A1}$$

$$u(x^*) = \int_{y^*-TP}^{x^*} \left[u'(y^*_{-TP}) + \int_{y^*-TP}^{z^*} (V - q^2)u\, dy^* \right] dz^* + u(y^*_{-TP}). \tag{A2}$$

Consider (A1). Since $u(y^*)$ is of order unity near the potential peak, errors of order q in $u(y^*)$ will lead only to errors of order q in the first integral. The kinetic term will be of order q^2, so the second integral goes to zero as q^2. The error in u' at the inside matching point is also of order q, and so the last term $u(y^*_{-TP})$ is of that order (since it is of at least that smallness in the two solutions matched at y^*_{-TP}). Thus errors of order q in u can make only errors of order q in u'. Consideration of (A2) in the same way then shows that the system is stable; errors of order q in u' induce only errors of order q in u.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Pair Distribution Function of a One-Dimensional Hard Rod Gas

MICHAEL FLICKER
Case Western Reserve University,
Department of Physics, Cleveland, Ohio

(Received 12 June 1967)

The pair distribution function $D_2(x_1, x_2)$ for an infinite system of one-dimensional hard rods, each of length d, is derived directly from the expression for a finite system. In addition, for a finite system of N rods, it is shown that $D_2(x_1, x_2)$ is constant in the region of translational invariance if $|x_2 - x_1| \geq (N-1)d$. A relationship between $D_2(x_1, x_2)$ and $D_1(x)$ is also discussed.

I. INTRODUCTION

In a recent paper, Leff and Coopersmith[1] have discussed the translational invariance properties of a finite one-dimensional hard core fluid. In particular, they have discussed a system of N identical hard rods, each of length d, in the interval $[0, L]$. They have shown that the canonical two-particle distribution function, $D_2(x_1, x_2)$, is translationally invariant if $(N-2)d \leq x_1, x_2 \leq L - (N-2)d$.

The purpose of this paper is to present the result that, for a finite system, $D_2(x_1, x_2)$ is *constant* in the region of translational invariance if $|x_2 - x_1| \geq (N-1)d$. In addition, we show how the well-known expression for the two-particle distribution function in an infinite system follows directly from the result for the finite system.[2]

Section II of this paper contains the essentials of the calculation. We borrow notation and results quite liberally from Ref. 1.

In Sec. III we discuss the results in the light of a connection between $D_2(x_1, x_2)$ and $D_1(x)$.

II. D_2 IN THE REGION OF TRANSLATIONAL INVARIANCE

The ordered two-particle distribution function in the canonical ensemble, where by ordered is meant $x_1 \leq x_2$ as given in Ref. 1, is

$$D_2^0(x_1, x_2) = \frac{(M+1)!}{2\xi^{M+1}} \sum{}'' \frac{(x_1 - nd)^n}{n!} \times \frac{(L - x_2 - md)^m}{m!} \frac{[x_2 - x_1 - (l+1)d]^l}{l!} \theta_1\theta_2\theta_3, \tag{1}$$

where

$$M = N - 1, \xi = L - Md, \theta(s) = \begin{cases} 1, & \text{for } s \geq 0 \\ 0, & \text{for } s < 0 \end{cases}$$

and

$$\theta_1 = \theta(x_1 - nd), \quad \theta_2 = \theta(L - x_2 - md), \quad \theta_3 = \theta(x_2 - x_1 - ld - d). \tag{2}$$

The doubly primed summation means n, m, and l run from zero to $M-1$ subject to the condition that $l + n + m = M - 1$.

Since we are in the region of translational invariance, that is $(M-1)d \leq x_1, x_2 \leq L - (M-1)d$, $\theta_1 = \theta_2 = 1$. If we define

$$y \equiv \frac{x_2 - x_1 - d}{d} \tag{3}$$

and

$$C \equiv L - (x_2 - x_1) - (M - l - 1)d = \left(\frac{\xi}{d} + l - y\right)d \tag{4}$$

then, after some manipulation, we have

$$D_2^0(x_1, x_2) = \frac{(M+1)!}{2\xi^{M+1}} \sum_{l=0}^{M-1} \frac{d^l(y-l)^l}{l!} \theta(y-l) \times \sum_{n=0}^{M-1-l} \frac{(x_1 - nd)^n}{n!} \frac{(C - x_1 + nd)^{M-1-l-n}}{(M-1-l-n!)}. \tag{5}$$

(Note: In Ref. 1, C is referred to as $\xi_l^{(2)}$.) But from Theorem 1 in Ref. 1, we have the rather amazing result that the second summation is independent of x_1. Furthermore, it is given by

$$\sum_{n=0}^{P} \frac{(x - nd)^n}{n!\,(P-n)!} [C - (x - nd)]^{P-n} = \sum_{n=0}^{P} \frac{(-d)^{P-n}}{n!} C^n, \tag{6}$$

where $P = M - 1 - l$. Thus, (5) can be simplified to read

$$D_2^0(x_1, x_2) = \frac{(M+1)!}{2\xi^{M+1}} (-d)^{M-1} \sum_{l=0}^{M-1} \sum_{n=0}^{M-1-l} \frac{(-1)^{l+n}}{l!\,n!} \times (y-l)^l \left(\frac{\xi}{d} + l - y\right)^n \theta(y-l), \tag{7}$$

where we have put back the explicit form of C. If the expression $(\xi/d + l - y)^n$ in (7) is expanded by the binomial theorem, we obtain

$$D_2^0(x_1, x_2) = \frac{(M+1)!\,(-d)^{M-1}}{2\xi^{M+1}} \times \sum_{l=0}^{M-1} \sum_{n=0}^{M-1-l} \sum_{k=0}^{n} \frac{(-1)^l (y-l)^{l+n-k}}{l!\,k!\,(n-k)!} \theta(y-l) \left(-\frac{\xi}{d}\right)^k. \tag{8}$$

[1] H. S. Leff and M. H. Coopersmith, J. Math. Phys. **8**, 306 (1967).
[2] Z. W. Salsburg, R. W. Zwanzig, and J. G. Kirkwood, J. Chem. Phys. **21**, 1098 (1953).

We now observe that

$$\sum_{l=0}^{M-1}\sum_{n=0}^{M-1-l}\sum_{k=0}^{n}=\sum_{n=0}^{M-1}\sum_{k=0}^{n}\sum_{l=0}^{M-1-n}=\sum_{k=0}^{M-1}\sum_{n=k}^{M-1}\sum_{l=0}^{M-1-n}.$$

If we change the summation variables such that

$$\lambda = n - k, \quad s = M - 1 - k,$$

then

$$\sum_{k=0}^{M-1}\sum_{n=k}^{M-1}\sum_{l=0}^{M-1-n}=\sum_{s=0}^{M-1}\sum_{\lambda=0}^{s}\sum_{l=0}^{s-\lambda}.$$

One final variable change is now necessary to obtain the desired form, that is

$$p = l + \lambda.$$

This implies

$$\sum_{\lambda=0}^{s}\sum_{l=0}^{s-\lambda}=\sum_{p=0}^{s}\sum_{l=0}^{p}.$$

$D_2^0(x_1, x_2)$ can now be written, after some rearranging, as

$$\frac{(M+1)!}{2\xi^2}\sum_{s=0}^{M-1}\frac{(-d/\xi)^s}{(M-1-s)!}\sum_{p=0}^{s}\sum_{l=0}^{p}\frac{(-1)^l}{l!\,(p-l)!}\times(y-l)^p\theta(y-l). \quad (9)$$

Let us consider the special situation where $y \geq M - 1$. This is equivalent to requiring $x_2 - x_1 \geq Md$. If y is subject to this restriction it is clear that

$$\theta(y - l) = 1$$

for all possible l. The l summation that appears in (9) is now a special case of the sum considered in (6). In particular, it corresponds to the situation where $C = 0$. The ordered two-particle distribution function, subject to this condition, can be simplified to

$$D_2^0(x_1, x_2) = \frac{(M+1)!}{2\xi^2}\sum_{s=0}^{M-1}\left(\frac{-d}{\xi}\right)^s\frac{(s+1)}{(M-1-s)!} = \text{a constant}. \quad (10)$$

Thus, $D_2^0(x_1, x_2)$ is a constant in the region of translational invariance whenever $x_2 - x_1 \geq Md$. A sufficient condition that such values of x_1, x_2 exist is that the number density be less than one third of the close-packing density.

In order to obtain $D_2^0(x_1, x_2)$ in the thermodynamic limit, for values of x_1 and x_2 in the region of translational invariance it is convenient to rewrite Eq. (9) in a slightly different form. Define

$$\zeta \equiv Md/\xi$$

and

$$A(p, y) \equiv \sum_{l=0}^{p}\frac{(-1)^l}{l!\,(p-l)!}(y-l)^p\theta(y-l).$$

$D_2^0(x_1, x_2)$ can now be written as

$$D_2^0(x_1, x_2) = \frac{(M+1)(M)}{2\xi^2}\times\sum_{s=0}^{M-1}\sum_{p=0}^{s}A(p, y)\frac{(-\zeta)^s(M-1)!}{(M-1-s)!\,M^s}. \quad (11)$$

Changing the order of summations this becomes

$$D_2^0(x_1, x_2) = \frac{(M+1)(M)}{2\xi^2}\sum_{p=0}^{M-1}A(p, y)\sum_{s=p}^{M-1}(-\zeta)^s\times\frac{(M-1)!}{(M-1-s)!\,M^s}. \quad (12)$$

We are interested in obtaining $D_2^0(x_1, x_2)$ in the limit as $N, L \to \infty$ such that $D = N/L =$ constant. The requirement that there exist a region of translational invariance in the thermodynamic limit is that the number density be less than one-half the close-packing density. As a consequence, the region of translational invariance is infinite and $0 < \zeta < 1$. In what follows we do not present an exact derivation of the limiting form of $D_2^0(x_1, x_2)$ but give a convincing plausibility argument. The result can be proved rigorously, but the proof is rather tedious and will not be presented. For finite s,

$$\frac{(M-1)!}{(M-1-s)!\,M^s} \to 1 \quad \text{as} \quad M \to \infty.$$

Thus, in the thermodynamic limit, one would expect Eq. (12) to become

$$D_2^0(x_1, x_2) = \frac{D^2}{2(1-Dd)^2}\sum_{p=0}^{\infty}A(p, y)\sum_{s=p}^{\infty}(-\zeta)^s. \quad (13)$$

Since $0 < \zeta < 1$,

$$\sum_{s=p}^{\infty}(-\zeta)^s = \frac{(-\zeta)^p}{1+\zeta}. \quad (14)$$

But $1/(1 + \zeta) = 1 - Dd$. Hence, from the definition of $A(p, y)$ we have

$$D_2^0(x_1, x_2) = \frac{D^2}{2(1-Dd)}\times\sum_{p=0}^{\infty}\sum_{l=0}^{p}\frac{(-\zeta)^p(-1)^l(y-l)^p\theta(y-l)}{l!\,(p-l)!}. \quad (15)$$

If we interchange the order of summation and shift the p index by l, we obtain the final form

$$D_2^0(x_1, x_2) = \frac{D^2}{2(1-Dd)}\sum_{l=0}^{\infty}\frac{[\zeta(y-l)]^l}{l!}\times e^{-\zeta(y-l)}\theta(y-l). \quad (16)$$

To see that this is equivalent to the result of Salsburg,

Zwanzig, and Kirkwood,[2] define

$$x \equiv \frac{|x_2 - x_1|}{d} \quad \text{and} \quad l = \frac{1}{Dd}.$$

The two-particle distribution function then becomes

$$D_2(x_1, x_2) = \frac{1}{ld^2(l-1)} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \left(\frac{x-k}{l-1}\right)^{k-1} \times \exp\left[-\left(\frac{x-k}{l-1}\right)\right] \theta(x-k). \quad (17)$$

Note that this is the expression for $D_2(x_1, x_2)$ rather than $D_2^0(x_1, x_2)$; the factor $\frac{1}{2}$ that appears in (16) does not appear in (17) because the variables are no longer ordered.

III. DISCUSSION

The purpose of this paper has been to derive, from the result for a finite system, the canonical two-particle distribution function $D_2(x_1, x_2)$ in the thermodynamic limit for values of x_1, x_2 in the region of translational invariance. In addition, we proved that, for a finite system, $D_2(x_1, x_2)$ is constant in the region of translational invariance if $|x_2 - x_1| \geq (N-1)d$.

It is interesting to note that, for the infinite system, there is a relationship between the two-particle distribution function and the one-particle distribution function. In Ref. 1, it was proved that, in the thermodynamic limit,

$$D_1(x) = \frac{D}{1-dD} \sum_{n=0}^{k} \frac{1}{n!} \left[\frac{D(x-nd)}{1-dD}\right]^n \times \exp\left[-\frac{(x-nd)D}{1-dD}\right],$$

where $kd \leq x \leq (k+1)d$. In our notation this can be written as

$$D_1(x) = \frac{D}{1-dD} \sum_{n=0}^{\infty} \frac{[\zeta(y'-n)]^n}{n!} e^{-\zeta(y'-n)} \theta(y'-n),$$

where $y' = x/d$. $D_1(x)$ is the density at a finite distance x from the wall. The similarity between this relation and the result obtained in Sec. II for the two-particle distribution, i.e.,

$$D_2(x_1, x_2) = \frac{D^2}{1-dD} \sum_{n=0}^{\infty} \frac{[\zeta(y-n)]^n}{n!} e^{-\zeta(y-n)} \theta(y-n),$$

where $y = (x_2 - x_1 - d)/d$, is striking. Intuitively, however, it seems quite reasonable that these functions should be so connected. Recall that the condition that there is a hard wall at $x = 0$ is equivalent to having a fixed particle at $x = -d$. Therefore, $D_1(x)$ is related to the probability of finding a particle at x knowing there is a particle at $-d$. Hence, in essence, it is a two-particle distribution function. There is a similar connection between the fact that $D_1(x)$ is constant in the region of translational invariance and the fact that $D_2(x_1, x_2)$ is constant in the region of translational invariance if $|x_2 - x_1| \geq (N-1)d$.

We conclude this section with a question. Is the connection between $D_2(x_1, x_2)$, where x_1, x_2 are in the interior of the fluid, and $D_1(x)$, where x is near a wall, peculiar to one-dimensional hard rods or is it more general?

ACKNOWLEDGMENT

The author wishes to acknowledge helpful discussions with Professor H. S. Leff.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 1 JANUARY 1968

Further Generalized Bose Condensation, Anisotropic ODLRO and Thin ^{4}He Films

GIDEON CARMI
Belfer Graduate School of Science, Yeshiva University, New York, New York

(Received 6 June 1967)

This paper contains a detailed analysis of the Bose condensation of an ideal Bose gas, finite in one dimension and infinite in the other two. Limiting processes, which were treated questionably in previous work, are here re-examined, and some seemingly paradoxical conclusions of the past work are explained. Two new concepts emerge—those of "further generalized Bose condensation" (FGBC) and "anisotropic off-diagonal long-range order" (AODLRO). The two concepts are shown to be connected to each other, and both are shown to exist at $T = 0$ for the model discussed.

1. INTRODUCTION

This paper is part of an investigation about off-diagonal long-range order (ODLRO) in systems in which at least one dimension is finite or zero. The results we aim at have close bearing on a kind of ODLRO which can be said to exist in a laser beam which is reflected between two parallel mirrors. They are also related to the properties of a one-dimensional system of "electrons" and "ions" on a circle, interacting with a potential whose Fourier transform is $1/k^2$, which are developed in detail in another paper.[1]

In the present paper, we restrict ourselves to a re-examination of Ziman's discussion[2] of thin ^{4}He films.

In 1953 Ziman came to the seemingly paradoxical conclusion that a three-dimensional system of free Bosons with energies $\sim k^2$, two dimensions of which are infinite ($L^2 = \infty^2$) while the third is finite ($D < \infty$), cannot undergo a Bose condensation (i.e., $T_c = 0$) *no matter how large D may be*. More recent papers[3] have corroborated the computational parts of Ziman's results without, however, challenging his basic method. We arrived at the conclusion that although his method contains two questionable steps (the unjustified interchange of two limiting processes and the unjustified but vital assumption of the smallness of a parameter), his results are nonetheless valid. We show this by giving two alternative versions of his calculation: one version tries to follow his as closely as possible; the other uses a more general and much shorter calculation. From this case we draw two conclusions. First, the introduction of interactions, however weak, whether short- or long-range, makes a drastic change in the properties of a Bose system and may compensate for the finiteness of a dimension. This again is closely connected with Schafroth's finding[4] that the wavefunction in a finite box changes from sinusoidal to a practically flat shape due to the self-consistent field set up by the interactions: i.e., the wavefunction "forgets" the existence of the boundary, up to very small distances from the latter. The second conclusion stems from the question whether all semblance to a Bose condensation (BC) is indeed lost in Ziman's case or whether, at least, some kind of generalized Bose condensation (GBC)[5] and ODLRO will still hold. In trying to answer this question, two concepts seem to suggest themselves:

(1) A "further generalized Bose condensation" (FGBC), which can loosely be defined as a situation in which a finite fraction of the particles condenses into a subset of measure zero of the set of states (see Sec. 3 for a more accurate definition). This subset is not necessarily "infinitesimally close to the $k = 0$ state" as in Girardeau's GBC, and we show that indeed the Ziman model exhibits such a situation at $T = 0$.

(2) "Anisotropic off-diagonal long-range order" (AODLRO), e.g., in ρ_1, meaning that

$$\langle xyz|\ \rho_1\ |xyz'\rangle \to \rho_c \neq 0$$

when $|z - z'| \to \infty$ (or when $|z - z'| \to D$, where D is the length of the box in the z-direction), whereas $\langle \rho_1 \rangle \to 0$ when $|x - x'| \to \infty$ or $|y - y'| \to \infty$. This concept derives its importance from the anisotropic type of ODLRO which may be said to exist in a laser (coherent radiation field between two mirrors). The interesting aspect of Ziman's model from this point of view is, however, that it provides another case of AODLRO, as we show in Sec. 3, where we also prove a theorem on the equivalence of FGBC and AODLRO, analogous to Girardeau's theorem[5] on the equivalence of GBC and ODLRO.

Models with FGBC and AODLRO at $T \neq 0$ will

[1] G. Carmi, "A Circular One-Dimensional Model for the Electron–Ion System" (to be published).
[2] J. M. Ziman, Phil. Mag. **44**, 548 (1953).
[3] See, e.g., D. F. Goble and L. E. H. Trainor, Can. J. Phys. **44**, 27 (1966) and the references cited therein.
[4] M. R. Schafroth, Phys. Rev., **100**, 463 (1955).
[5] M. D. Girardeau, J. Math. Phys. **6**, 1083 (1965).

be exhibited elsewhere. Here we merely wish to introduce these concepts and exhibit them in a simple case.

2. ZIMAN'S CALCULATIONS FOR THIN BOSE FILMS: AN ALTERNATIVE PROCEDURE

In 1953, J. M. Ziman[2] intended to show that a film of large but finite thickness (e.g., 1 cm) of ideal Bose particles would exhibit ordinary Bose–Einstein condensation (BC) whose transition temperature gradually reduces with the thickness of the film (as suggested by experiments on thin liquid ^{4}He films and by earlier theoretical work by Osborne[6]). But he arrived at quite a different and rather paradoxical conclusion: The condensation temperature of a Bose–Einstein gas of finite thickness, however large, tends to zero as $1/\ln N$, where N, the number of particles, tends to infinity together with the other two dimensions (while the density is kept constant). Thus a film of finite thickness always behaves as if it were two-dimensional. As for the two-dimensional case, Osborne[6] had previously observed that, although it does not show a Bose condensation, it undergoes a rapid "accumulation" into the ground state at a more or less well-defined "accumulation temperature" T_a. As Goble and Trainor (in particular) have shown in great detail,[3] a similar T_a exists in the case of three finite dimensions, although there are at least five different ways of defining T_a, each giving slightly different results. In both cases (the two- and the three-dimensional) $T_a \to 0$ as two dimensions are made large, although in the three-dimensional case T_a tends to the ordinary BC temperature T_λ if the third dimension is increased together with the other two.

To escape this paradoxical situation (according to which a finite bulk helium sample should, if the ideal Bose–Einstein gas is to have any value at all as a model, exhibit a λ temperature which depends on the size and shape of the bulk, however large its dimensions) and at the same time retain the relevance of the ideal gas as a model, Ziman suggested that one should consider a three-dimensional Bose–Einstein gas (finite or infinite) as made up of small, cubic "cells," each of side length $L_0 = \sim 7 \cdot 10^{-6}$ cm, which are assumed to be uncorrelated to each other. Therefore any finite system would exhibit the (nonzero) accumulation temperature belonging to its individual cells, with a value independent of the size of the system. While a system of thickness larger than L_0 exhibits about the same degree of agreement between its T_a and the λ point of a real liquid He of corresponding density as the three-infinite-dimensional ideal Bose gas, the size of the cells is so chosen that systems of width $D < L_0$ will show (according to Zimans' calculations) a suitably decreasing T_a. By fitting the single parameter L_0, Ziman obtains a very good agreement with experiment over the calculated D range of 4–10 atomic layers. This agreement has recently been checked by Goble and Trainor[3] by means of detailed computer calculations for various definitions of T_a. We refer the reader to their paper for additional literature on the subject, all of which was concerned with essentially the same type of questions—namely, that of checking Ziman's results against experiment; the question of various alternative definitions of T_a and the question of boundary conditions in the D direction (which are found[3] to have negligible effects on the results). As far as we know, the basic assumptions (e.g., approximation procedures) of Ziman's calculation have not been challenged since the appearance of his article thirteen years ago.

Since Ziman's procedure seems questionable in several points, and also since we wish to arrive at some more general expressions [see Eqs. (2.13) and (3.3) below] which will lead to the suggestion of the concepts of FGBC and of AODLRO, we shall, after a short review of Ziman's procedure, present two alternative derivations of his results. Then we proceed to interpret these derivations.

We consider N noninteracting Bose–Einstein particles confined to a box of dimensions $L \times L \times D$. The thermodynamics limit will be defined by $L, N \to \infty$, so that $\rho = N/L^2D$ remains finite and constant. The energy levels are given by

$$E_{l,m,n} = \frac{\pi^2\hbar^2}{2M}\left(\frac{l^2+m^2}{L^2} + \frac{n^2}{D^2}\right) = [\alpha(l^2+m^2) + \beta n^2]\kappa T, \tag{2.1}$$

where we are using Ziman's notation.[2] In equilibrium at temperature T the average occupation of the level l, m, n is

$$N_{lmn} = \{\exp[\alpha(l^2+m^2) + \beta n^2 - \mu/\kappa T] - 1\}^{-1}. \tag{2.2}$$

μ is the chemical potential, which must fulfill the condition that $E_{111} - \mu \geq 0$ in order to give $N_{lmn} \geq 0$ for all l, m, n, i.e.,

$$\epsilon \equiv 2\alpha + \beta - \mu/\kappa T \geq 0. \tag{2.3}$$

From (2.2), the total number of particles can be written

$$N = \rho L^2 D = \sum_{l,m,n=1}^{\infty} \sum_{j=1}^{\infty} \exp[j(\mu - E_{lmn})/\kappa T] = \sum_{j=1}^{\infty} e^{j\epsilon}\left[\sum_{l=1}^{\infty} e^{-jl^2\alpha}\right]^2 \sum_{n=1}^{\infty} e^{-jn^2\beta}. \tag{2.4}$$

[6] M. F. M. Osborne, Phys. Rev. 76, 396 (1949).

In order to approximate the sums appearing in (2.4), Ziman applies Jacobi's transformation[7]:

$$f(x) = \sum_{n=1}^{\infty} \exp(-xs^2) = \frac{1}{2}\left(\frac{\pi}{x}\right)^{\frac{1}{2}}\left[1 + 2\sum_{n=1}^{\infty}\exp(-\pi^2 s^2/x)\right] - \frac{1}{2}. \quad (2.5)$$

If $x \leq \pi/2$, the error of approximating the right-hand side by $\frac{1}{2}(\pi/x)^{\frac{1}{2}} - \frac{1}{2}$ is at most one part in 500; if $x > \pi/2$, the error of approximating the left-hand side by its first term, $\exp(-x)$, is, again, one part in a few hundred. At $x \simeq 1.68$ the two approximations agree exactly, deviating from (2.5) by the same order of magnitude (one part in several hundred).

Ziman therefore divides the summation over j into three regions (using $D < L$, i.e., $\beta > \alpha$):

(a) $1 \leq j \leq \pi/2\beta$;

(b) $\pi/2\beta < j \leq \pi/2\alpha$;

(c) $\pi/2\alpha < j$.

Taking $x = j\beta$ in (2.5), $\sum_n \exp(-jn^2\beta)$ can be approximated by $\frac{1}{2}(\pi/j\beta)^{\frac{1}{2}} - \frac{1}{2}$ in region (a) (where $x \leq \pi/2$), and by $\exp(-j\beta)$ in regions (b) and (c) (where $x > \pi/2$). Taking $x = j\alpha$ in (2.5),

$$\sum_l \exp(-jl^2\alpha)$$

can be approximated by $\frac{1}{2}(\pi/j\alpha)^{\frac{1}{2}} - \frac{1}{2}$ in regions (a) and (b) (where $x \leq \pi/2$) and by $\exp(-j\alpha)$ in region (c), giving

$$N \approx \frac{1}{8}\sum_{j=1}^{\pi/2\beta} e^{\mu j/\kappa T}\left[\left(\frac{\pi}{j\alpha}\right)^{\frac{1}{2}} - 1\right]^2\left[\left(\frac{\pi}{j\beta}\right)^{\frac{1}{2}} - 1\right] + \frac{1}{4}\sum_{j=\pi/2\beta+1}^{\pi/2\alpha} e^{j[(\mu/\kappa T)-\beta]}\left[\left(\frac{\pi}{j\alpha}\right)^{\frac{1}{2}} - 1\right]^2 + \sum_{j=\pi/2\alpha+1}^{\infty} \exp j\left(\frac{\mu}{\kappa T} - 2\alpha - \beta\right). \quad (2.6)$$

Here $\pi/2\beta$ and $\pi/2\alpha$ in the summation limits stand for $\lceil \pi/2\beta \rceil$ and $\lceil \pi/2\alpha \rceil$, where $\lceil x \rceil$ means the integer nearest to x. The last sum in (2.6) can be summed as a geometrical series; this gives

$$\exp\left(-\left\lceil \frac{\pi}{2\alpha} + 1 \right\rceil \epsilon\right)(1 - e^{-\epsilon})^{-1}, \quad (2.7)$$

which has a singularity when $\epsilon \to 0$. (2.6) is to be taken as an implicit definition of μ. Ziman now argues essentially in the following way.

When N is increased (at given L, D, and T), μ increases so that ϵ approaches $0+$. However, the first two sums cannot exceed the value N_0 which they assume at $\epsilon = 0$. This value is practically attained when ϵ reaches a value (very near to 0) for which the third sum is of the order of N_0. Any further increase in N is accompanied by practically no change in both ϵ and in the first two terms; but there is an increase in the third term. This presents an "accumulation" process into the lowest level, which is reached when the density N/L^2D is increased (at fixed T, L^2D) to a point where μ remains practically constant. Since N is always finite, one cannot expect true discontinuities (e.g., in the gradient of μ); but, at best, there are steep changes.

Ziman then defines the transition point T_a by (a) omitting the third term in (2.6), (b) setting ϵ equal to 0 and (c) equating the remainder to N_0; or, after multiplication by $8\alpha\beta^{\frac{1}{2}}/\pi^{\frac{3}{2}}$,

$$\frac{8\rho}{(\pi\kappa T_a)^{\frac{3}{2}}} = \frac{8N_0\alpha\beta^{\frac{1}{2}}}{\pi^{\frac{3}{2}}} = \sum_{j=1}^{\pi/2\beta} e^{j(2\alpha+\beta)}\left[\left(\frac{1}{j}\right)^{\frac{1}{2}} - \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}}\right]^2\left[\left(\frac{1}{j}\right)^{\frac{1}{2}} - \left(\frac{\beta}{\pi}\right)^{\frac{1}{2}}\right] + 2\left(\frac{\beta}{\pi}\right)^{\frac{1}{2}}\sum_{j=\pi/2\beta+1}^{\pi/2\alpha} e^{2j\alpha}\left[\left(\frac{1}{j}\right)^{\frac{1}{2}} - \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}}\right]^2. \quad (2.8)$$

The right-hand side is then expanded as functions of α and β, which are taken to be small. The result can be written in terms of T_λ, the "lambda point" of a three-infinite-dimensional BE system of the same density:

$$\left(\frac{T_\lambda}{T_a}\right)^{\frac{3}{2}} = 1 - \frac{1}{D}\left(\frac{T_\lambda}{T_a}\right)^{\frac{1}{2}}\left(0.15 - 0.504\ln\frac{T_\lambda}{T_a} + 1.512\ln D - 1.008\ln L\right) - \frac{1.008}{L}\left(\frac{T_\lambda}{T_a}\right)^{\frac{1}{2}}\ln D. \quad (2.9)$$

If now $L \to \infty$ while D, ρ remain constant, one obtains from (2.9)

$$T_a \sim \frac{1}{\ln L}. \quad (2.10)$$

This means that, in the thermodynamic limit (in which T_a goes over into a true condensation temperature), one obtains $T_c = 0$, no matter how large D may be.

The first question arising about this procedure is the "Ansatz" (2.8). First, while it is true that the third term of (2.6) arises solely from the terms $l = m = n = 1$ in (2.4), i.e., from the lowest energy level, it is by no means true that the third term of (2.6) exhausts *all* contributions from the lowest level—at least not as long as $\alpha \neq 0$. As a matter of fact, quite a large contribution from the lowest level is contained in the first two terms of (2.6) before the thermodynamic limit is taken, as can be seen by inspection. Thus the

[7] E. T. Whittaker and G. M. Watson, *A Course on Modern Analysis* (Cambridge University Press, Cambridge, 1966).

prescription (2.6) is not equivalent to the usual prescription[8] to obtain the λ point, where *first* the contributions of the lowest level are dropped (actually one drops an arbitrary, but finite, number of lowest levels[8]) and *then* setting the chemical potential equal to the lowest energy level. [It is obviously crucial *first* to carry out the thermodynamic limit $L^2D \to \infty$ ($\alpha \to 0$) and *then* the limit $\mu \to E_{111}$ as one sees, e.g., from (2.4); otherwise one would get

$$N = \sum_{j=1}^{\infty} \sum_{l^2+m^2+n^2>1} \exp\left[-j(l^2+m^2) - jn^2\beta\right] + \sum_{j=1}^{\infty} 1,$$

from which $N \to \infty$ will never recover the result obtained by taking the limits in the reverse order.]

Thus we have to make sure that Ziman's prescription is indeed equivalent to the way which first takes $L \to \infty$ (i.e., $\alpha \to 0$) and then sets $\epsilon \to 0$ in the result. Moreover, we have to make sure that setting $\epsilon = 0$ in the expressions which contains it (which stand after $\sum_j$ and $\sum_s$ signs) is the right way of taking $\lim_{\epsilon\to 0} N\alpha\beta^{\frac{1}{2}}$; i.e., we have to make sure that $\lim_{\epsilon\to 0}$ commutes with $\sum_j$ and $\sum_s$ [as the above example has shown, $\lim_{\epsilon\to 0}$ certainly does *not* commute, for example, with $\sum_{l,m,n}$ in (2.4)].

Secondly, it is true that, as $\epsilon \to 0$, the first two terms will continue to accommodate (beyond a certain point) only a negligible part of the particles which are being added to the system. However, this result holds in the process of increasing N at constant volume and temperature (i.e., while increasing ρ), and it is by no means clear *a priori* that we may conclude from this that, in another process, namely that of changing N and the volume at constant ρ, the contribution to the first two terms is negligible beyond a certain term. In view of the paradoxial character of Ziman's result and the delicacy of the mathematical process, it seems worthwhile to check these questions closely; it also seems desirable to modify the method in such a way that the assumption that β is small, which is made by Ziman in order to arrive at Eq. (2.9), is no longer necessary. (The epsilontics of the order of limits requires some tricks and will be described in Appendix A.) Here we start again from Eq. (2.4) and introduce into it the *full* transformation (2.5), this time for $\sum_{l,m}$ only, *without* making any approximation. Equation (2.4) thus goes over into the *exact* equation

$$N\alpha\beta^{\frac{1}{2}} = \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \sum_{j=1}^{\infty} e^{j(2\alpha+\beta-\epsilon)} \times \left[j^{-2}\left(1 + 2\sum_{s=1}^{\infty} e^{-\pi^2 s^2/j\alpha}\right) - (\alpha)^{\frac{1}{2}}\right]^2 \sum_n e^{-j\beta n^2}. \tag{2.11}$$

[8] P. T. Landsberg, Proc. Cambridge Phil. Soc. **50**, Pt. 1, 65 (1953).

We next take the thermodynamic limit $L \to \infty$, i.e., $\alpha \to 0$, of (2.11). In Appendix A we show that this can be done simply by replacing α with 0 whereever it occurs; i.e., the operation $\lim_{\alpha\to 0}$ can be interchanged with the summations in (11) (which is by no means obvious). This gives

$$\lim_{\substack{N\to\infty \\ (\alpha\to 0)}} N\alpha\beta^{\frac{1}{2}} = \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \sum_{j=1}^{\infty} \frac{e^{j(\beta-\epsilon)}}{j} \sum_{n=1}^{\infty} e^{-jn^2\beta}. \tag{2.12}$$

Using the same type of epsilontics which led to the previous interchange of limits, one can show that the sums over j and n in (2.12) are interchangeable, as long as $\epsilon \neq 0$. If we use $\int$ to denote the indefinite integral [keeping (2.2) in mind] this gives

$$\begin{aligned}\lim_{\substack{N\to\infty \\ \alpha\to 0}} N\alpha\beta^{\frac{1}{2}} &= \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \sum_{n=1}^{\infty}\left[-\int d\epsilon \sum_{j=1}^{\infty} \exp\left(\mu' - n^2\beta\right)\right] \\ &= \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \int_{\mu_0'}^{\mu'} d\mu \sum_{n=1}^{\infty} \left[\exp\left(n^2\beta - \mu\right) - 1\right]^{-1} \\ &= -\tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \sum_{n=1}^{\infty} \ln \frac{e^{n^2\beta} - e^{\mu'}}{e^{n^2\beta} - e^{\mu_0'}}. \qquad \left(\mu' \equiv \frac{\mu}{\kappa T}\right).\end{aligned}$$

Comparison with (2.12) shows that the integration constant μ_0' has to be chosen equal to $-\infty$, which gives

$$\begin{aligned}\lim_{\substack{N\to\infty \\ \alpha\to 0}} N\alpha\beta^{\frac{1}{2}} &= \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \sum_{n=1}^{\infty} \ln \frac{1}{1 - e^{\mu' - n^2\beta}} \\ &= \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}}\left[\ln \frac{1}{1 - e^{\mu'-\beta}} + \sum_{n=2}^{\infty} \ln \frac{1}{1 - e^{\mu'-n^2\beta}}\right].\end{aligned} \tag{2.13}$$

Here we have split off the first term which becomes singular when $\epsilon \to 0$ (i.e., $\mu \to 2\alpha + \beta = \beta$), i.e., that term which gives a finite fraction of particles condensing into the $n = 1$ state. It is to be noted, though, that this is a state with $n = 1$ and $l, m =$ arbitrary, whereas the behavior of μ (or ϵ), which gives the singularity of the first term of (2.13), is determined by its implicit equation

$$\frac{N}{\Omega} = \frac{1}{\Omega} \frac{1}{e^{E_{111}-\mu} - 1} + \frac{1}{\Omega} \sum_{i^2+m^2+n^2>1} N_{lmn} \tag{2.14}$$

in which the first term ($l = m = n = 1$) is dropped first by the customary procedure of passing to the thermodynamic limit before allowing $\mu \to E_{111}$. This is because the behavior of μ is dictated by the fact that we wish to determine the ordinary Bose condensation temperature T_c. Thus the first term of (2.13), which already contains a sum over all possible values of l and m, is quite different in meaning from the first term of (2.14). As we shall see, it reflects the existence of a certain type of GBC and ODLRO.

To show that (2.13) implies Ziman's result $T_c = 0$ or, more specifically,

$$T_c = \lim_{N\to\infty} \frac{\gamma}{\ln N}, \tag{2.15}$$

where γ is a constant, we use the fact that

$$\mu' - \beta = \mu' - \lim_{\alpha\to 0} \frac{E_{111}}{\kappa T} = \lim_{N\to\infty} O\left(\frac{1}{N}\right),$$

the proof of which we postpone. (See the corresponding stage of the alternative procedure given in the next section.)

Thus we find from (2.13) that the right-hand side of

$$\frac{1}{T_c^{\frac{3}{2}}} \frac{1}{L^2 D} \sum_{\substack{n=1 \\ l,m \text{ arbitrary}}} N_{lmn} = \tfrac{1}{2}(\pi\beta)^{\frac{1}{2}} \ln \frac{1}{1 - e^{\mu'-\beta}}$$

$$= \lim_{N\to\infty} O\left(\frac{\ln N}{(\kappa T_c)^{\frac{1}{2}}}\right)$$

diverges as $\ln N$, whereas the factor

$$\sum_{n=1}^{\infty} N_{lmn}/L^2 D \leq \rho < \infty$$

is finite. Therefore T_c must be zero, to the order $\sim 1/\ln N$.

3. SECOND ALTERNATIVE TO ZIMAN'S CALCULATION; FGBC AND AODLRO IN THIN BOSE FILMS

We now present an alternative way of obtaining (2.13) and the result $T_c = 0$, a way which has the advantage of compactness and of using a type of calculation which is more familiar. We start out from the fact that for an ideal Bose–Einstein gas of two infinite dimensions ($L^2 \to \infty$) in the x and y directions and one finite dimension ($D < \infty$) in the z direction, the average occupation number $N(k_x, k_y, n)$ of momentae

$$\mathbf{k} = \left(k_x, k_y, \frac{2\pi\hbar}{D} n\right),$$

where n = integer, k_x, k_y = real, is given by

$$N(\mathbf{k}) = \{\exp[\beta n^2 + \hbar^2(k_x^2 + k_y^2)/2m\kappa T - \mu'] - 1\}^{-1} \tag{3.1}$$

and the density is given by

$$\rho = \lim \frac{N_{\text{tot}}}{\Omega}$$

$$= \frac{1}{(2\pi)^2 D} \times \sum_{n=1}^{\infty} \mathrm{P} \int \frac{dk_x\, dk_y}{\exp[\beta n^2 + \hbar^2(k_x^2 + k_y^2)/2m\kappa T - \mu'] - 1} + \lim_{\Omega\to\infty} \frac{1}{\Omega} \frac{1}{e^{\beta-\mu'} - 1} \tag{3.2}$$

where P means principal value and the last term describes a possible contribution from the lowest level $k_x = k_y = 0$, $n = 1$ below T_c. The nth integral can now be evaluated:

$$\mathrm{P} \int \frac{dk_x\, dk_y}{\exp[\beta n^2 + \hbar^2(k_x^2 + k_y^2)/2m\kappa T - \mu'] - 1} = \frac{m\kappa T}{\hbar^2} \int_0^{\infty} \frac{dx}{e^{x+\beta n - \mu'} - 1},$$

$$= \frac{m\kappa T}{\hbar^2} \ln \frac{1}{1 - \exp[-\epsilon - (n^2 - 1)\beta]},$$

where $\epsilon = \beta - \mu'$, i.e.,

$$\rho = \frac{m\kappa T}{\hbar^2}\left[\ln \frac{1}{1 - e^{-\epsilon}} + \sum_{n=2}^{\infty} \ln \frac{1}{1 - \exp[-\epsilon - (n^2 - 1)\beta]}\right],$$

$$= \frac{mKT}{\hbar^2}\left[\ln \frac{1}{1 - e^{-\epsilon}} + \sum_{n=2}^{\infty}\right]. \tag{3.3}$$

This is identical with our previous result (2.13). Here we have again separated the term $n = 1$ and dropped the last term of (3.2), since the only use to be made of (3.3) was to calculate T_c, which, as usual, is done by approaching T_c from above, i.e., in the range where the condensate is excluded.

We may read off the result $T_c = 0$ immediately from (3.3), by noting that (3.2) may be taken to describe a countable number of two-dimensional systems of "effective" chemical potentials

$$[-\epsilon - (n^2 - 1)\beta]\kappa T, \quad (n = 1, 2, \cdots,),$$

where $\epsilon \to 0$ when $T = 0$: i.e., the first system has a condensation point at $T = 0$, while the others do not have any condensation points. Indeed[8]

$$\frac{m\kappa T}{\hbar^2} \ln \frac{1}{1 - \exp(\mu' - E_0/\kappa T)}$$

is the density of a two-dimensional system of free bosons of chemical potential μ and lowest energy level E_0.

Alternatively, we may argue that, as long as T_c is defined by the provision that for $T < T_c$ a finite fraction of the particles condenses into the lowest level (i.e., the last term of 17 is $\neq 0$), ϵ must tend to zero when $T \to T_c$. Thus (3.2) shows that T_c cannot be > 0, since $\ln 1/1 - e^{-\epsilon}$ and $\sum_2$ are both positive (but $\ln 1/1 - e^{-\epsilon} \approx \ln 1/\epsilon$ diverges as $T \to T_c$), while ρ remains finite. Moreover, as $\sum_2$ remains finite, we must have $\epsilon^{-1} = O(N)$ as $\Omega \to \infty$ and $T_c = O(1/\ln N)$.

The behavior of the term $n = 1$ of (3.2) suggests strongly that the subset $\{n = 1; k_x k_y = \text{arbitrary}\}$ of all states plays a role similar to that of the state $k_x = k_y = k_z = 0$ in ordinary Bose condensation. Although here this subset is not one out of a continuum, the fact that $\ln 1/1 - e^{-\epsilon}$ tends to infinity as

$T \to 0$ while the other terms remain finite suggests that in the limit $T \to 0$ *the system may be considered as "condensing" into that subset in a way which might be called "Further Generalized Bose Condensation"* (FGBC), since the subset of states is more general than the one considered by Girardeau in defining GBC.

For the purposes of the present paper, we shall tentatively define FGBC as *a condensation into one of a countable set of subsets of states, such that the ratio of occupation of this particular subset to the occupation of any other subset becomes infinite whenever T is below or equal to a certain FGBC temperature T_{FC}, and at the same time the ratio to the occupation numbers of all other subsets combined is at least $O(1)$.* In our case, $T_{FC} = 0$; but according to this definition, a FGBC definitely occurs at $T = 0$. Systems with $T_{FC} > 0$ require interactive forces and will be reported elsewhere.

A decisive fact in the occurrence of FGBC in our case is the discreteness of the levels belonging to the z direction which result from taking a finite width D and which are physically equivalent to the case of attractive forces which produce bound states with a discrete spectrum. Such attractive forces are assumed by Girardeau[8] in his example for the occurrence of his type of GBC. Therefore it will not be surprising to find that, while our system undergoes a FGBC into the subset $n = 1$ $\{k_x k_y = \text{arbitrary}\}$ at $T = 0$, the system also undergoes (ordinary) GBC into "infinitesimal vicinity" of the level $k_x = k_y = 0$ and $k_z = 2\pi\hbar/D$ at $T = 0$. In fact, if $0 < k_0 < 4\pi\hbar/D$, and if we denote $\beta - \mu/\kappa T = \beta - \mu' = \epsilon$,

$$\begin{aligned}\rho_c &= \lim_{k_0\to 0}\lim_{N\to\infty}\frac{1}{\Omega}\sum_{k<k_0} n_k = \lim_{k\to 0}\frac{1}{2\pi^2 D}\\ &\quad\times\int_{k_x^2+k_y^2<k_0^2}\frac{dk_x\,dk_y}{\exp\left[\epsilon+\hbar^2(k_x^2+k_y^2)/2m\kappa T\right]}\\ &= \lim_{k_0\to 0}\frac{m\kappa T}{(2\pi)^2 D\hbar^2}\int_0^{k_c}\frac{dx}{e^{x+\epsilon}-1}\\ &= \lim_{k_0\to 0}\frac{m\kappa T}{4\pi^2 D^2\hbar^2}\ln\frac{1-e^{-k_0-\epsilon}}{1-e^{-\epsilon}}. \qquad (3.4)\end{aligned}$$

If we perform $\lim_{k_0\to 0}$ *before* $\epsilon, T \to 0$, we obtain $\rho_c = 0$; i.e., there is no GBC at $T \neq 0$. But if we take $\epsilon, T \to 0$ before taking $\lim_{k_0\to 0}$, then the product $\kappa T \ln\dfrac{1-e^{-k_0-\epsilon}}{1-e^{-\epsilon}}$ gives, according to our previous results, the nonzero number ρ which is independent of k_0 [while the term $\kappa T \ln(1-e^{-k_0-\epsilon}) \to 0$] and remains so when $k_0 \to 0$. In other words, we have GBC at $T = 0$.

The question now naturally arises, whether (as in the case of Girardeau's GBC) FGBC will be associated with some kind of ODLRO. If so, it is natural to expect, in the systems discussed here, that this is a type of ODLRO associated with a particular direction in space (the z direction, in our case). Furthermore, it has to be associated with a finite range (since in the thermodynamic limit the system remains contained within a finite width D).

We are thus suggesting to look into a specific kind of ODLRO—which we might call "anisotropic ODLRO" (AODLRO)—for example, in ρ_1, (the first reduced density matrix), defined by

$$\lim_{|z-z'|\to D}\lim \text{therm}\,\langle xyz|\,\rho_1\,|xyz'\rangle = \rho_z > 0. \qquad (3.5)$$

Using periodic boundary conditions,

$$\langle xyz|\,\rho_1\,|x'y'z'\rangle = \frac{1}{\Omega}\sum_{\mathbf{k}} n_k e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}.$$

In our case this will reduce to

$$\lim \text{therm}\,\langle xyz|\,\rho_1\,|xyz'\rangle = \frac{1}{D}\sum_n N_n e^{i(2\pi/D)n(z-z')}, \qquad (3.6)$$

where, by (3.2) and (3.3),

$$\begin{aligned}N_n &= \frac{1}{(2\pi)^2}\\ &\quad\times\int\frac{dk_x\,dk_y}{\exp\left[(n^2-1)\beta+\hbar^2(k_x^2+k_y^2)/2m\kappa T-\epsilon\right]-1}\\ &= \frac{\kappa T}{4\pi^2\hbar^2}\ln\frac{1}{1-\exp\left[-\epsilon-(n^2-1)\beta\right]}. \qquad (3.7)\end{aligned}$$

Hence, for $T = 0$, $\epsilon = 0$ (the case of FGBC),

$$\begin{aligned}&\lim_{|z-z'|\to D}\lim \text{therm}\,\langle xyz|\,\rho_1\,|xyz'\rangle\\ &\qquad= \left[\frac{\kappa T}{4\pi^2\hbar^2 D}\ln\frac{1}{1-e^{-\epsilon}}\right]_{\substack{T=0\\ \epsilon=0}} = \rho > 0,\end{aligned}$$

where we have used (3.3) in the last step and the fact that in the limit $T = 0$, $\epsilon = 0$, the terms $n \geq 2$ drop out in (3.3). We thus obtain ODLRO in the sense defined above (AODLRO).

This proof is easily generalized for the case of FGBC (at $T_{FC} \geq 0$) to a theorem analogous to Girardeau's theorem[8] concerning the equivalence of GBC and ODLRO:

Theorem: A system having FGBC will have AODLRO and *vice versa* (under certain mild conditions on the functions n_k involved).

The proof follows the same lines as the above calculation and Girardeau's proof; the $n \geq 2$ terms again drop out because of their vanishing ratios to the $n = 1$ term, as required in the definition of FGBC.

APPENDIX. INTERCHANGE OF LIMITS APPEARING IN SEC. 2

In (2.4) we have to calculate

$$L = \lim_{\alpha \to 0} \frac{\alpha}{2} \sum_j e^{j(\mu-\beta)} \sum_{l,m,n=1}^{\infty} \exp\{-j(l^2 + m^2)\alpha - jn^2\beta\} \tag{A1}$$

We obviously cannot exchange $\lim_{\alpha\to 0}$ with $\sum_{l,m}$, ($\sum_{l,m} \lim_{\alpha\to 0}$ is indefinite). The crux of the matter is that, by applying (2.5) on the $\sum_{l,m}$, one has, *de facto*, carried out the limiting process which defines the infinite sum $\sum_{l,m}$. One may then carry out $\lim_{\alpha\to 0}$ simply by putting $\alpha = 0$ wherever α occurs. To see this, it is sufficient to show that the rest term,

$$R_{s_0} = \sum_{s=s_0}^{\infty} e^{-\pi s^2/j\alpha}$$

is as small as we please, for given j, if either s_0 is given and is sufficiently small or α is given and s_0 is sufficiently large. Clearly,

$$R_{s_0} < \sum_{r=0}^{\infty} e^{-\pi(s_0^2+r)/j\alpha} = \frac{e^{-\pi s_0^2/j\alpha}}{1 - e^{-\pi/j\alpha}},$$

which is as small as we please under these conditions; on the other hand, the rest term

$$R_{l_0} = \sum_{l=l_0}^{\infty} e^{-jl^2\alpha}$$

is, for any given l_0, as large as we please, provided that α is taken small enough.

To see this, we can write

$$R_{l_0} > \sum_{r=1}^{k} \exp\{-j[l_0^2 + r(2l_0k + 1)]\alpha$$
$$= e^{-jl_0^2} \frac{1 - \exp[-j(k+1)(2l_0 + 2k - 1)\alpha]}{1 - \exp[-j(2l_0 + 2k - 1)\alpha]} \tag{A2}$$

Here we have replaced the first k term of R_{l_0} by k smaller terms, making a geometric series in which the difference $2l_0k + 1$ of two consecutive powers is equal to the *largest* difference of the powers of any two consecutive terms of

$$R_{l_0}^k = \sum_{l=l_0}^{l_0+k} e^{-jl^2\alpha}.$$

k is chosen so large that the terms $r > k$ contribute much less than the difference of the right-hand side from R_{l_0} above. This can always be done, as is easily seen. Now, for each given M one can find an α and a k such that the right-hand side of (A2) will be larger than M. One first chooses k, such that $k + 1 > M$, and then α, such that for this k the exponents on the right-hand side of (A2) can be replaced by the terms (linear in α) of their Taylor expansions in α (with error amounting to less than 1), from which the result follows.

What we have shown is that in

$$\lim_{\alpha\to 0} \sum_j e^{j(\mu-\beta)} \times \left\{ j^{-\frac{1}{2}} \left[1 + 2\sum_{s=1}^{\infty} e^{-\pi s^2/j\alpha} \right] - (\alpha)^{\frac{1}{2}} \right\}^2 \sum_n e^{-j\beta n^2} \tag{A3}$$

[which is the result of applying (2.5) to (A1)], one may cut the sum over s at s_0, with an error ϵ which can be made as small as we please for any given α (by choosing s_0 large enough) and which will not grow if α is chosen smaller:

$$L = \lim_{\alpha\to 0} \sum_j e^{j(\mu-\beta)} \times \left\{ j^{-\frac{1}{2}} \left[1 + \frac{\epsilon}{2} + \sum_{j=1}^{s_0} e^{-\pi s^2/j\alpha} \right] - (\alpha)^{\frac{1}{2}} \right\}^2 \sum_n e^{-j\beta n},$$

where $s_0 = s_0(\epsilon, \alpha, j)$. We now may interchange $\lim_{\alpha\to 0}$ with $\sum_{s=1}^{s_0}$. But in order to do that we first have to interchange $\lim_{\alpha\to 0}$ with the $\sum_j$. To show that this is permissible it is sufficient to show that $\sum_j$ converges uniformly in α. But if $\sum_j$ converges for any α, it will clearly converge even better for a smaller α, as one sees by inspection; i.e., it will converge uniformly for all smaller α. In fact, $\sum_j$ converges for any $\alpha \neq 0$.

ACKNOWLEDGMENTS

The author wishes to thank Drs. Moshe Fibich and Gerald Horwitz for stimulating discussions.

This research was supported by the National Science Foundation (under contract GP-6134) and by the Air Force Office of Scientific Research (under contract AF 19(628)-4381).