JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 6, NUMBER 8 AUGUST 1965

On the Algebraic Structure of the Cluster Expansion in Statistical Mechanics

CAHİT ARF,* KAYA İMRE, AND ERCÜMENT ÖZİZMİR†
Çekmece Nuclear Research and Training Center, Istanbul, Turkey
(Received 25 February 1964; revised manuscript received 12 March 1965)

The structure of cluster expansion which is widely used in statistical mechanics is studied from an algebraic point of view. In doing this, a commutative algebra is constructed which is generated by partitions of a finite set by regarding them as operators which divide the set into disjoint parts. Physically, these operators correspond to operations which remove interaction among certain clusters of particles. It is shown that the cluster expansion stems from the relation between two basis sets of this algebra; the first set is the set of all partitions and the second is the set of pairwise orthogonal minimal idempotents. This property enables one to demonstrate the equivalence of the product versus cluster properties of the distribution and correlation functions, respectively, in general terms. This is done by constructing a simple representation space for the partition algebra corresponding to the distribution functions. A second application of the partition algebra is considered in the case when correlations are well-ordered with respect to the interaction strength λ, so that to a given order in λ the distribution functions are not all independent and can be expressed in terms of a finite irreducible set of functions involving smaller numbers of particles. The combinatorial problem of calculating the expansion coefficients is carried out explicitly using a graded representation space with respect to the order in λ. It is concluded that the partition algebra can be used as a mathematical tool in handling problems involving cluster expansion.

I. INTRODUCTION

ONE of the main problems in statistical mechanics is to express macroscopic physical quantities pertaining to interacting systems of large degrees of freedom in terms of the interaction potential. In equilibrium, all thermodynamic quantities can be expressed in terms of the partition function. The calculation of the partition function leads us to the evaluation of an integral over the state variables of a large number of particles. A method which involves the cluster development initiated by Ursell[1] and elaborated further by Mayer[2] and his collaborators provided a means of evaluating of this integral for imperfect gases. The method was later extended for application to quantum systems[3] and to systems with nonbinary potentials.[4] Recently, analogous methods were employed in treating nonequilibrium problems.[5]

The combinatorial problems involved in the calculation of the partition function, and also in reduction of the distribution functions in powers of a suitable parameter, can be handled by making use of the powerful techniques of the linear graph theory.[6] In the application of the graph theory to statistical mechanics, the connected graphs defined on a set of points play a special role. For example, the Ursell cluster functions $U_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ are defined as a sum over all connected graphs on the set of l points.[6] The relation between the probability distribution $W_N(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ which are defined as a sum over

* Present Address: Institute for Advanced Study, Princeton, New Jersey.
† Present Address: Ministry of Defence, Research and Development Center, Etimesğut, Ankara, Turkey.

[1] H. D. Ursell, Proc. Cambridge Phil. Soc. **23**, 685 (1927).
[2] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940).
[3] B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938).
[4] H. L. Friedman, *Ionic Solution Theory* (Interscience Publishers, New York, 1962).
[5] E. G. D. Cohen, Physica **28**, 1025, 1045, 1060 (1962).
[6] G. E. Uhlenbeck and G. W. Ford, in *Studies in Statistical Mechanics*, edited by J. De Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1961), Vol. I, Part B.

all graphs on the set of N points and the cluster functions U_l is given by

$$W_1(i) = U_1(i),$$
$$W_2(ij) = U_2(ij) + U_1(i)U_1(j),$$
$$W_3(ijk) = U_3(ijk) + [U_2(ij)U_1(k) + \text{cyclic terms}] + U_1(i)U_1(j)U_1(k), \quad (1)$$

and so on. In general, W_N is equal to the sum of all possible products of cluster functions U_l such that in every term the N particles are divided into disjoint groups (partitions). The inverse relation is also well known and is given by

$$U_1(i) = W_1(i),$$
$$U_2(ij) = W_2(ij) - W_1(i)W_1(j),$$
$$U_3(ijk) = W_3(ijk) - [W_2(ij)W_1(k) + \text{cyclic terms}] + 2W_1(i)W_1(j)W_1(k), \quad (2)$$

and so on. In general, the cluster function U_l is equal to the sum of all possible products of probability functions W_N as above except that each term is multiplied by a factor of $(-1)^{k-1}(k-1)!$ where k is the number of groups in that term. The relations (1) and (2) are referred to as Ursell development. In problems where this development is suitable, it does not seem necessary to go into the graph-theoretical details of the interaction picture, since W_N and U_l both involve certain collections of graphs. In such problems the relations (1) and (2) may be considered as a one-to-one correspondence between two sets of functions W_N and U_l, which are symmetric functions of their arguments. There is a further decomposition for U-functions used in statistical mechanics which is due to Husimi and is referred to as star expansion. This is not considered in the context of the present investigation.

The purpose of this paper is to study the structure of the cluster expansion using algebraic methods and demonstrate that the combinational problems involved in the application of this expansion can be handled conveniently through the use of the *partition algebra* which will be introduced in detail in the sequel. The generators of this algebra are *partitions* which are regarded as operators on finite sets which divide these sets into disjoint parts. Physically, partitions may be interpreted as operators which cut off interactions among separate groups of particles. Thus, the probability of finding a system in a certain state becomes a product when operated by such an operator. Repeated application of these operators induces more factorization. In this manner the composition rule of the partition operators will be defined. The mathematical structure of the partition algebra is a well-known type which has been studied extensively. Therefore, the use of it as a mathematical tool presents no algebraic novelty. Indeed, many of the properties demonstrated herein can be recognized by those acquainted with the semisimple structures.[7] Nevertheless, a self-contained description of this algebra will be given in Sec. II. The application of the algebra to a specific physical problem is made through the construction of an appropriate representation space upon which the elements of the algebra are defined as linear operators. In Sec. III we construct a simple representation space of the partition algebra which is particularly suited to the study of the implications of the defining relations (1) and (2) between the distribution and the cluster functions. It will be shown that the proof of equivalence of the product and cluster properties[6] of the distribution and correlation functions, respectively, is immediate once the partition algebra and the appropriate representation space are constructed. This will be given in Sec. IV. In the proof of equivalence no assumption will be made on the W- and U-functions other than the fact that they are related through the cluster series.

In Secs. V–VIII, we give another application of the method developed here to a combinatorial problem associated with the cluster development when the correlation functions are well ordered with respect to a parameter λ.

II. THE PARTITION ALGEBRA $C(M)$

Let M be a finite set of $|M|$ identical elements, and denote the subsets of this set by $\mathfrak{M}_i$, $i=1, 2, \cdots$.

A set of disjoint, nonempty, subsets, $\mathfrak{M}_1, \cdots, \mathfrak{M}_k$, is said to be a partition of M if the union of these subsets is M. Alternatively, partitions P of M can be regarded as operators on M which divide M into disjoint subsets whose union is M. In what follows partitions of M will only be used in the latter sense. The subsets $\mathfrak{M}_i$, $(i = 1, \cdots, k)$ will be referred to as the parts of P.

With every pair of partitions P and P' (of M), we associate another partition, the product PP' which is formed by the nonempty intersections of the parts of P and P'. It follows from this definition that the multiplication of two partitions is commuta-

[7] N. Bourbaki, *Eléments de Mathématique. Livre II: Algèbre* (Hermann & Cie., Paris, 1958). (See in particular Chap. 8).

tive, i.e., $PP' = P'P$, and also operators P are idempotents, i.e., $P^2 = PP = P$.

We also consider the collection of all formal linear combinations of operators P with coefficients taken from a field $\mathfrak{R}$ (say, e.g., the field of rational numbers). In this way a commutative algebra $C(M)$ is defined. The collection of all partitions of M contitutes a basis set for $C(M)$.

Two special partitions are of interest: (i) Identity partition E. This partition introduces no division on the set M. Therefore, one has $PE = P$, for all P of M. (ii) Null partition P_0. This partition divides M such that there is only one element in each of its $|M|$ parts. One clearly has $PP_0 = P_0$, for all P of M.

It is convenient to introduce a nonnegative integer $d(P)$ corresponding to each partition P, defined by

$$d(P) = |M| - k(P), \tag{3}$$

where $k(P)$ is the number of parts of P. We call $d(P)$ the defect of partition P. In particular, $d(E) = |M| - 1$, and $d(P_0) = 0$.

It follows from the composition rule given above that the number of parts of PP' cannot be less than either $k(P)$ or $k(P')$. Alternatively, this can be expressed as

$$d(PP') \leq \min\{d(P), d(P')\}. \tag{4}$$

When the equality sign holds, either $PP' = P$ or $PP' = P'$; if also $d(P) = d(P')$ then one must have $P = P'$. If $PP' = P'$, then the parts of P' must be contained in some of the parts of P. In this case we say that "P contains P'" (or "P' is contained in P") and write $P \supseteq P'$ (or $P' \subseteq P$). If $PP' \neq P'$, then we write $P \not\supseteq P'$.

Minimal Idempotents of $C(M)$

Corresponding to every partition P of M, we shall construct an element e_P in $C(M)$ which satisfies the following conditions:

$$\begin{aligned} &\text{(i)} \quad e_P e_P = e_P, \\ &\text{(ii)} \quad P'e_P = e_P \quad \text{if} \quad P' \supseteq P, \\ &\qquad\qquad = 0 \quad \text{if} \quad P' \not\supseteq P. \end{aligned} \tag{5}$$

It is noted that the above conditions define uniquely the elements e_P which are pairwise orthogonal and which form a basis for $C(M)$. To see the pairwise orthogonality, we note that the product $e_P e_{P'} = PP'e_P e_{P'}$ [cf. (ii)] is nonvanishing if and only if $P \supseteq P'$ and $P \subseteq P'$, i.e., $P = P'$. The set of all e_P forms a basis for $C(M)$ since they are linearly independent. To prove uniqueness, suppose that there exist two sets of elements say e_P and $\tilde{e}_P$ which satisfy (5). Because of completeness, $\tilde{e}_P = \sum_{P'} a_{P,P'} e_{P'}$, hence $\tilde{e}_P e_{P'} = a_{P,P'} e_{P'}$. Also, if $P \neq P'$, then $\tilde{e}_P e_{P'} = 0$, which implies $a_{P,P'} = 0$. Thus $\tilde{e}_P = a_{P,P} e_P$. Using (i) we get $a_{P,P} = a_{P,P}^2$. Since $\tilde{e}_P \neq 0$, $a_{P,P} = 1$, i.e., these sets are identical.

Consider the element $A(0) = E - P_0$ of the algebra $C(M)$. We see that $P_0 A(0) = 0$, and $A(0)A(0) = A(0)$. Similarly, we define a set of elements by the following recurrence formula:

$$A(d) = A(d-1)(E - \sum_{d(P)=d} P), \qquad (d = 1, \cdots, |M| - 1), \tag{6}$$

where the summation is over all partitions with defect d. These elements are idempotents, and, moreover, $A(d)$ annihilates all partitions with defects $\leq d$. These properties can be proved by induction, observing that the element

$$P(E - \sum_{d(P')=d(P)} P') = - \sum_{\substack{d(P')=d(P) \\ P' \neq P}} (PP')$$

contains only partitions with defects $< d(P)$. One gets

$$A(d) = \prod_{i=0}^{d} (E - \sum_{d(P')=i} P'). \tag{7}$$

The minimal idempotents e_P can be written in terms of the annihilators $A(d)$ as

$$\begin{aligned} e_{P_0} &= P_0, \\ e_P &= PA(d(P) - 1), \end{aligned} \tag{8}$$

since the defining conditions (5) are satisfied.

It follows from (6) by successive substitution that

$$A(d) = E - \sum_{d(P) \leq d} e_P. \tag{9}$$

Since annihilator of E, i.e., $A(|M| - 1)$, must be zero, one obtains

$$E = \sum_{P \in C(M)} e_P, \tag{10}$$

which is an expression of completeness of the set of minimal idempotents.

Subalgebras of $C(M)$

Let $\mathfrak{M}$ be a subset, and P be a partition of the set M. The set of all non-empty intersections of $\mathfrak{M}$ and the parts of P defines a partition on $\mathfrak{M}$. We denote this partition by $P(\mathfrak{M})$, and this correspondence by $T(\mathfrak{M})$. $C(M)$ is mapped by $T(\mathfrak{M})$ homomorphically onto $C(\mathfrak{M})$: that is $T(\mathfrak{M}): C(M) \to C(\mathfrak{M})$.

On the other hand, under the same transforma-

tion $T(\mathfrak{M})$, the subalgebra of $C(M)$ which is generated by the partitions having the subset $M - \mathfrak{M}$ as one of their parts, is mapped isomorphically onto $C(\mathfrak{M})$. We denote this subalgebra also by $C(\mathfrak{M})$, and its elements by the same symbols used for their corresponding elements in $C(\mathfrak{M})$.

Therefore, any partition $P = (\mathfrak{M}_1, \cdots, \mathfrak{M}_k)$of M can be written as the product of the unit elements $E(\mathfrak{M}_i) \in C(\mathfrak{M}_i)$, $(i = 1, \cdots, k)$. We shall now show that similar relationships also hold for e_P's, i.e.,

$$e_P = \prod_{i=1}^{k} e(\mathfrak{M}_i), \tag{11}$$

where we set, for simplicity, $e(\mathfrak{M}) = e_{E(\mathfrak{M})}(\mathfrak{M})$, i.e. the minimal idempotent in $C(\mathfrak{M})$ corresponding to $E(\mathfrak{M})$. To prove this we observe that

$$P \prod_{i=1}^{k} e(\mathfrak{M}_i) = \prod_{i=1}^{k} E(\mathfrak{M}_i)e(\mathfrak{M}_i) = \prod_{i=1}^{k} e(\mathfrak{M}_i),$$

and

$$P' \prod_{i=1}^{k} e(\mathfrak{M}_i) = \prod_{i=1}^{k} P'(\mathfrak{M}_i)e(\mathfrak{M}_i).$$

In the latter equation, if $P' \supseteq P$, then $P'(\mathfrak{M}_i) = E(\mathfrak{M}_i)$, hence the right-hand side reduces to the product of $e(\mathfrak{M}_i)$'s. But if $P' \not\supseteq P$, then at least for one value of i one has $P'(\mathfrak{M}_i) \neq E(\mathfrak{M}_i)$, hence the right-hand side vanishes. This completes the proof since $e(\mathfrak{M}_i)$'s are idempotents [cf. conditions (5)].

Another useful formula is

$$E = \sum_{i \in \mathfrak{M} \subseteq M} e(\mathfrak{M})E(M - \mathfrak{M}), \tag{12}$$

where the summation runs over all subsets of M which contain a fixed element i of M. This element merely plays the role of a counting device and the resulting sum does not depend on it.

To prove this formula, we consider Eq. (10). We can write $P = E(\mathfrak{M})P(M - \mathfrak{M})$, where $\mathfrak{M}$ is the part of P which contains the fixed element i. We first sum over all $P(M - \mathfrak{M}) \in C(M - \mathfrak{M})$, and then over all subsets $\mathfrak{M}$ of M which contain i. In this way every partition $P \in C(M)$ is counted once and only once. Hence

$$E = \sum_{i \in \mathfrak{M} \subseteq M} e(\mathfrak{M}) \sum_{P(M-\mathfrak{M}) \in C(M-\mathfrak{M})} e_{P(M-\mathfrak{M})},$$

which, by virtue of (10), is identical to (12).

Transformation Rules between the Basis Sets $\{P\}$ and $\{e_p\}$

In order to express P's in terms of e_P's, it suffices to multiply (10) by P, to get

$$P = \sum_{P' \subseteq P} e_{P'}, \tag{13}$$

where the summation is over all P' contained in P. The inversion formula is somewhat more involved. We shall only give the relation expressing $e(M)$ in terms of P's. To calculate e_P, (11) can be used.

We shall now give the proof of the formula

$$e(M) = \sum_{P \in C(M)} (-1)^{k(P)-1}(k(P) - 1)!\, P. \tag{14}$$

This formula clearly holds when $|M| = 1$. Let us assume, as induction hypothesis, that it also holds for all proper subsets of M. Then, by making use of (10), we can write

$$\begin{aligned} e(M) &= E - \sum_{i \in \mathfrak{M} \subset M} e(\mathfrak{M})E(M - \mathfrak{M}) \\ &= E + \sum_{i \in \mathfrak{M} \subset M} E(M - \mathfrak{M}) \\ &\quad \times \sum_{P(\mathfrak{M}) \in C(\mathfrak{M})} P(\mathfrak{M})(-1)^{k(P(\mathfrak{M}))}(k(P(\mathfrak{M})) - 1)!. \end{aligned}$$

Now, $E(M - \mathfrak{M})P(\mathfrak{M})$ is a partition of $C(M)$, hence the two summations in the latter term can be combined into one summation over the partitions $P \in C(M)$ which have at least two parts. One has $k(P) = k(P(\mathfrak{M})) + 1$. Moreover, each partition $P \in C(M)$ with $k(P) \geq 2$ is obtained exactly $k(P) - 1$ times as every one of the $k(P(\mathfrak{M}))$ parts of $P(\mathfrak{M})$ becomes the set $M - \mathfrak{M}$. Therefore,

$$e(M) = E + \sum_{\substack{P \in C(M) \\ k(P) \geq 2}} (-1)^{k(P)-1}(k(P) - 1)!\, P,$$

which completes the proof.

Ideals and Residue Classes in $C(M)$

Let us consider $C(M)$ for a moment as a vector space. Then the subspace $\mathfrak{I}$ is said to be an ideal of $C(M)$ if $X \in \mathfrak{I}$ implies that $XY \in \mathfrak{I}$ for all $Y \in C(M)$. Let $\mathfrak{I}$ be an ideal of $C(M)$. Then the congruence relation $Y \equiv Z \pmod{\mathfrak{I}}$ indicates that there is an element $X \in \mathfrak{I}$ such that $Y = Z + X$. Then we say that Y is congruent to Z modulo $\mathfrak{I}$. The subspace generated by such elements Z is called the residue class modulo $\mathfrak{I}$. The reason for introducing these classes is that in this way the equations under consideration are transformed into congruences modulo a suitable ideal so that a considerable simplification is obtained in carrying out the manipulations. We shall now consider some useful examples.

Given a subset $\mathfrak{M} \subseteq M$, let P be a partition of M having the property that $P(\mathfrak{M}) \neq E(\mathfrak{M})$, i.e., P divides $\mathfrak{M}$ at least into two parts. It is clear that, for any $P' \in C(M)$, the product PP' has the same

property. Therefore, the linear combinations (over the field $\mathfrak{R}$) of all such partitions form an ideal $\mathfrak{I}(\mathfrak{M})$ of $C(M)$. It follows from the relation $e_P = Pe_P$ that if P is in this ideal, then so is e_P. Hence, the set of all e_P's corresponding to the partitions in $\mathfrak{I}(\mathfrak{M})$ form a basis for this ideal.

Now we consider all partitions of M which do not divide the subset $\mathfrak{M}$, i.e., $P(\mathfrak{M}) = E(\mathfrak{M})$. Then, the linear combinations of such partitions form the residue class modulo $\mathfrak{I}(\mathfrak{M})$, viz. $C(M)$ [mod $\mathfrak{I}(\mathfrak{M})$]. It is readily seen that the set of all e_P's corresponding to such partitions form a basis set for this class. With respect to ideal $\mathfrak{I}(\mathfrak{M})$, Eq. (12) can be written as the following congruence:

$$E(M) \equiv \sum_{\mathfrak{M} \subseteq \mathfrak{M}' \subseteq M} e(\mathfrak{M}')E(M - \mathfrak{M}') \quad [\text{mod } \mathfrak{I}(\mathfrak{M})]. \tag{15}$$

III. THE REPRESENTATION SPACE $L(M)$

In order to make use of the algebra $C(M)$ in physical problems of interest it is necessary to construct vector spaces, appropriate to the problem at hand, in which the elements of $C(M)$ are defined as linear operators. We shall now construct such a vector space axiomatically, and later use it, as an application, to study some properties of the so-called cluster expansion of statistical mechanics.

We associate a symbol $Q_{\mathfrak{M}}$ with every subset $\mathfrak{M} \subseteq M$, and consider the formal products of these symbols. We write $Q_{P(\mathfrak{M})}$ to denote the product $Q_{\mathfrak{M}_1} \cdots Q_{\mathfrak{M}_k}$, where $P(\mathfrak{M}) = (\mathfrak{M}_1 \cdots \mathfrak{M}_k)$. The linear combinations of all such $Q_{P(\mathfrak{M})}$ over the field $\mathfrak{R}$ generate a vector space which will be denoted by $L(\mathfrak{M})$.

We adopt the following axioms: (i) $PQ_{\mathfrak{M}} = P(\mathfrak{M})Q_{\mathfrak{M}} = Q_{P(\mathfrak{M})}$, for all $P \in C(M)$; (ii) $P(\alpha A) = (\alpha P)A = \alpha(PA)$, $P(A + A') = PA + PA'$, $(P + P')A = PA + P'A$, for all $A, A' \in L(\mathfrak{M})$, $\alpha \in \mathfrak{R}$; (iii) $P(AB) = (PA)(PB)$ for all $A \in L(\mathfrak{M})$, $B \in L(\mathfrak{M}')$, where $\mathfrak{M}$ and $\mathfrak{M}'$ are any two subsets of M with empty intersection.

It is easy to see that if X and Y are two elements of $C(M)$, one has the property

$$X(YA) = (XY)A. \tag{16}$$

Therefore, the vector spaces $L(\mathfrak{M})$ are representation spaces for the algebra $C(M)$.

Now we consider the element $K_{\mathfrak{M}}$ of $L(\mathfrak{M})$ defined as $K_{\mathfrak{M}} = e(\mathfrak{M})Q_{\mathfrak{M}}$. By making use of the relations (10), (12), (14), and (16) we can readily write relations between the elements Q_P and $K_P = K_{\mathfrak{M}_1} \cdots K_{\mathfrak{M}_k}$. We have respectively,

$$Q_M = \sum_{P \in C(M)} K_P, \tag{17}$$

$$Q_M = \sum_{i \in \mathfrak{M} \subseteq M} K_{\mathfrak{M}} Q_{M-\mathfrak{M}}, \quad (Q_\phi = 1), \tag{18}$$

$$K_M = \sum_{P \in C(M)} (-1)^{k(P)-1}(K(P) - 1)!\, Q_P, \tag{19}$$

where in proving these relations we use (16), e.g.,

$$\sum_{P \in C(M)} K_P = \sum_{P \in C(M)} e_P Q_P = \Big(\sum_{P \in C(M)} e_P\Big) Q_M = E(M)Q_M = Q_M. \tag{20}$$

The representation space $L(M)$ developed in this section is still abstract in nature. In order to apply it to the physical problems of interest we should demonstrate an operator isomorphy between $L(M)$ and the space of functions of the physical problem in which the elements of $C(M)$ are defined as linear operators. In the next section, we shall consider a special problem, namely the cluster expansion of statistical mechanics, and demonstrate how some of the properties can be deduced in a straightforward manner by making use of the operator formalism developed here.

IV. APPLICATION TO THE CLUSTER EXPANSION OF STATISTICAL MECHANICS

Let us consider the $|M|$-particle reduced distribution function F_M, where each element of the set M denotes the collection of state variables of a particle of the system under consideration. Suppose there is a grouping in the particles of the system so that particles belonging to different groups do not interact with each other. Then the probabilities of such particles being in certain states are mutually exclusive, which implies that F_M, in this case, becomes a product of distribution functions corresponding to noninteracting groups.

In most cases, the interaction potential is a strongly vanishing function as the distance between the interacting particles increases. Then, the distribution function corresponding to a set of particles which is formed by groups of particles situated far apart in the configuration space has the tendency to become factorized into the product of the distribution functions corresponding these subgroups. This is the so-called "product property."

It is for this reason that in constructing the vector space of the distribution functions we take as generating elements the (ordinary) products $F_{\mathfrak{M}_1} \cdots F_{\mathfrak{M}_k}$, where $\mathfrak{M}_i$ are disjoint subsets of M. The linear combinations of all such elements over the field $\mathfrak{R}$ form this vector space which will be denoted by $L^F(M)$.

We now introduce linear operators in this space

which materialize this factorization process by cutting off the interaction between some groups of particles. Denoting this grouping by the partition $P = (\mathfrak{M}_1, \cdots, \mathfrak{M}_k)$, we write $PF_{\mathfrak{M}} = F_P = \prod_{i=1}^{k} F_{\mathfrak{M}_i}$. We also adopt the axioms introduced in the previous section with regard to operations by P.

One can readily verify that there is an operator isomorphy between $L^F(M)$ and the abstract representation space $L(M)$ of the last section, defined by the correspondence $Q_M \leftrightarrow F_M$. In view of Eqs. (17) to (19), we observe that the elements in $L^F(M)$ corresponding to $K_{\mathfrak{M}}$ of $L(M)$ are nothing but the correlation functions $G_{\mathfrak{M}}$. Indeed, these equations correspond to the so-called cluster expansion (and its inversion).

Product versus Cluster Properties

It is physically a well-known fact that, in a set of particles, if there are noninteracting groups, then the correlation function corresponding to this set vanishes (while the distribution function becomes a product). This is the so-called "cluster property." Conversely, if $G_{\mathfrak{M}}$ is zero, then $F_{\mathfrak{M}}$ has the product property. It has been stated that when two sets of functions are related only by means of relations (17) and (19), then there is no simple and direct proof of either statements.[6] We shall now demonstrate that the partition algebra provides these proofs.

If $\tilde{P}$ is a partition whose parts denote the noninteracting groups of particles, then the product property is expressed, in our language, as $\tilde{P}F_{\mathfrak{M}} = F_{\mathfrak{M}}$. When $F_{\mathfrak{M}}$ have the product property, then, by definition, $G_{\mathfrak{M}} = e(\mathfrak{M})F_{\mathfrak{M}} = e(\mathfrak{M})\tilde{P}(\mathfrak{M})F_{\mathfrak{M}} = 0$ if $\tilde{P}(\mathfrak{M}) \neq E(\mathfrak{M})$. Thus, we have shown that the product property implies the cluster property.

Conversely, let us suppose that $G_{\mathfrak{M}}$ has a cluster property. That is, there is a certain grouping of particles denoted by $\tilde{P}$ such that $G_{\mathfrak{M}} = 0$ for all $\mathfrak{M} \subseteq M$ whenever $\tilde{P}(\mathfrak{M}) \neq E(\mathfrak{M})$. In other words, G_P is nonzero only if $P \subseteq \tilde{P}$. Thus, $F_M = \sum_P G_P = \sum_{P \subseteq \tilde{P}} G_P = \sum_{P \subseteq \tilde{P}} \tilde{P}G_P = F_{\tilde{P}}$. This completes the proof of the inverse statement that the cluster property implies the product property.

V. PRELIMINARIES TO AN APPLICATION OF THE METHOD TO A COMBINATORIAL PROBLEM

In Secs. V–VIII, we consider an application of the present method to a combinatorial problem involving the cluster development in which distribution functions are expanded with respect to a small parameter.

The combinatorial problem under consideration arises as follows: Consider

$$G_M = \sum_{P \in C(M)} (-1)^{k(P)-1}(k(P) - 1)!\, F_P \tag{21}$$

and the expansions

$$F_M = \sum_{\nu=0}^{\infty} \lambda^\nu F_M^\nu \quad \text{and} \quad G_M = \sum_{\nu=0}^{\infty} \lambda^\nu G_M^\nu,$$

where λ is a small parameter. One obtains

$$G_M^\nu = \sum_{P \in C(M)} (-1)^{k(P)-1}(k(P) - 1)! \sum_{|\nu|=\nu} F_P^\nu, \tag{22}$$

where $\nu = (\nu_1, \cdots, \nu_k)$ denotes a set of k [$= k(P)$] integers, $|\nu| = \nu_1 + \cdots + \nu_k$, and $F_P^\nu = F_{\mathfrak{M}_1}^{\nu_1} \cdots F_{\mathfrak{M}_k}^{\nu_k}$, $\mathfrak{M}_i$ being the parts of P.

Suppose correlation functions G_M are of order $\lambda^{|M|-1}$, i.e., $G_M = 0$, whenever $\nu < |M| - 1$. This property is referred to as the *well-ordering* of correlations. (See Sec. VIII.) In this case, we have from (22) a set of relations

$$F_M^\nu = \sum_{P \subset E} (-1)^{k(P)}(k(P) - 1)! \sum_{|\nu|=\nu} F_P^\nu, \qquad (\nu < |M| - 1), \tag{23}$$

which shows that F_P^ν are not all linearly independent. Observe that there are some factors $F_{\mathfrak{M}_i}^{\nu_i}$ ($\mathfrak{M}_i \subset M$) in the right-hand side of (23) such that $\nu_i < |\mathfrak{M}_i| - 1$. Therefore, through repeated use of (23), one can express any F_M^ν with $\nu < |M| - 1$ in terms of a set of functions $\{F_{\mathfrak{M}}^\mu\}$ for which $\mu \geq |\mathfrak{M}| - 1$, $\mu \leq \nu$, $\mathfrak{M} \subset M$. This set will be referred to as the irreducible set. Our problem is to find the total number that a certain cluster of irreducible factors appear in the final expression.

The program of the remainder of this paper is as follows. In Sec. VI, we construct *abstract* representation spaces $\mathfrak{L}(\mathfrak{M})$, ($\mathfrak{M} \subseteq M$), of $C(M)$, so that the elements of $\mathfrak{L}(\mathfrak{M})$ correspond to the linear combinations of F_P^ν. A *grading* with respect to nonnegative integers ν (the degree) is introduced in these spaces so that $\mathfrak{L}(M)$ can be written as a direct sum of the graded spaces $L^\nu(M)$. The degree ν corresponds to the expansion power of λ. A subspace $L_0^\nu(M)$ is considered whose elements correspond to those G_M^ν which vanish under the well-ordering assumption. Thus, it suffices to consider the quotion space $L^\nu(M)$ [mod $L_0^\nu(M)$] so that to the equation which expresses F_M^ν in terms of the irreducible set there corresponds a congruence relation modulo $L_0^\nu(M)$. In calculating the coefficients in this congruence relation we use operators $\partial_{\mathfrak{M}}$ which map $L^\nu(M)$ into $L^{\nu-1}(M)$. In Sec. VII an application to the cluster series in which correlations are well-ordered is given.

A short discussion of the results obtained is presented in Sec. VIII.

VI. THE VECTOR SPACE $\mathfrak{L}(M)$

Let us associate with every nonempty subset $\mathfrak{M} \subseteq M$ the symbols $Q^{\nu}_{\mathfrak{M}}$ ($\nu = 0, 1, \cdots$) and consider the formal products of these symbols $Q^{\nu}_{P} = Q^{\nu_1}_{\mathfrak{M}_1} \cdots Q^{\nu_k}_{\mathfrak{M}_k}$, where $\mathfrak{M}_i$ ($i = 1, \cdots, k$) are the parts of the partition P of M and $\nu = \nu_1, \cdots, \nu_k$ is a set of nonnegative integers. All linear combinations of such formal products over the field $\mathfrak{R}$ form a vector space $\mathfrak{L}(M)$.

We adopt the following axioms by which the elements of $C(M)$ are defined as linear operators on $\mathfrak{L}(M)$. (i) $PQ^{\nu}_{\mathfrak{M}} = P(\mathfrak{M})Q^{\nu}_{\mathfrak{M}} = \sum_{|\nu|=\nu} Q^{\nu}P(\mathfrak{M})$; (ii) for $A, A' \in \mathfrak{L}(\mathfrak{M})$, $\alpha \in \mathfrak{R}$, $P(\alpha A) = (\alpha P)A = \alpha(PA)$, $P(A + A') = PA + PA'$, $(P + P')A = PA + P'A$; (iii) for $A \in \mathfrak{L}(\mathfrak{M})$, $B \in \mathfrak{L}(\mathfrak{M}')$ with $\mathfrak{M} \cap \mathfrak{M}' = \phi$, $P(AB) = (PA)(PB)$. It is readily verified that $\mathfrak{L}(\mathfrak{M})$ are representation spaces for $C(M)$ since (16) also holds for these spaces.

Grading in $\mathfrak{L}(M)$

We define the degree of a product Q^{ν}_{P} as the integer $\nu = |\nu| = \nu_1 + \cdots + \nu_k$. The degree of an arbitrary element of $\mathfrak{L}(M)$ is defined as the maximum of the degrees of all products with nonvanishing coefficients in this element. Zero-degree term is the constant term, and, for $A \in \mathfrak{L}(\mathfrak{M})$, PA is the constant term if $P(\mathfrak{M})$ is the null partition (zero defect).

If all terms in $A \in \mathfrak{L}(\mathfrak{M})$ have the same degree ν, then A is said to be a homogeneous element with degree ν. If $X \in C(M)$, and if $A \in \mathfrak{L}(M)$ is a homogeneous element with degree ν, then XA is either zero or also homogeneous with the same degree ν.

Let $L^{\nu}(M)$ denote the subspace of $\mathfrak{L}(M)$ which is spanned by the homogeneous elements with degree ν. It is clear that the sum of all these subspaces for $\nu = 0, 1, \cdots$, is equal to $\mathfrak{L}(M)$. Moreover, this is a direct sum since the zero element of $\mathfrak{L}(M)$ can only be obtained by the sum of the zero elements of $L^{\nu}(M)$. We thus write

$$\mathfrak{L}(M) = \bigoplus_{\nu=0}^{\infty} L^{\nu}(M). \tag{24}$$

It is easily verified that every one of the subspaces $L^{\nu}(M)$ is a representation space for $C(M)$.

A further decomposition of $L^{\nu}(M)$ into direct sums of its subspaces is obtained by virtue of (5) and (10), giving

$$L^{\nu}(M) = \bigoplus_{P \in C(M)} e_P L^{\nu}(M). \tag{25}$$

It is of interest to note that for any $A \in L^{\nu}(M)$, one has $X(e_P A) = (Xe_P)A = (\sum_{P' \geq P} \xi_{P'}) e_P A$, where $X = \sum_P \xi_P P \in C(M)$. Therefore, $e_P A$ can be regarded as the eigenvector of any element $X \in C(M)$.

Remark: Any subspace of $L^{\nu}(M)$ which is mapped onto itself by $C(M)$ can be decomposed uniquely by forming the direct sum over all partitions of certain subspaces of $e_P\mathfrak{L}(M)$. The converse statement is also true.

A basis set for the space $e_P L^{\nu}(M)$ is obtained by operating e_P on the basis elements $Q^{\nu'}_{P'}$ of $L^{\nu}(M)$. One has

$$\begin{aligned} e_P Q^{\nu'}_{P'} &= 0 \qquad \text{if} \quad P' \not\supseteq P \\ &= \prod_{i=1}^{k'} \sum_{|\nu_i|=\nu_{i'}} e_{P(\mathfrak{M}_i)}(\mathfrak{M}_i) Q^{\nu_i}_{P(\mathfrak{M}_i)} \\ &\qquad \text{if} \quad P' \supseteq P, \end{aligned} \tag{26}$$

where $P' = (\mathfrak{M}_1, \cdots, \mathfrak{M}_{k'})$. Note that each term in the right-hand side is of the form $e_P Q^{\nu}_{P}$, therefore, the set of elements $e_P Q^{\nu}_{P}$ with $|\nu| = \nu$ spans $e_P L^{\nu}(M)$ so that

$$e_P L^{\nu}(M) = \bigoplus_{|\nu|=\nu} \mathfrak{R} e_P Q^{\nu}_{P}, \tag{27}$$

where the symbol $\mathfrak{R}$ in the right-hand side denotes the linear combinations over the field $\mathfrak{R}$.

The Subspace $L^{\nu}_0(M)$

Let I_ν denote the set of Q^{ν}_{P} such that $|\nu| = \nu$ and $\nu_i \geq |\mathfrak{M}_i| - 1$, ($i = 1, \cdots, k$). Define the subspace $L^{\nu}_0(M)$ by

$$L^{\nu}_0(M) = \bigoplus_{Q^{\nu}_{P} \in I_\nu} \mathfrak{R} e_P Q^{\nu}_{P}, \tag{28}$$

where the sum is over all Q^{ν}_{P} ($|\nu| = \nu$) which are not in I_ν. $L^{\nu}_0(M)$ is mapped onto itself by $C(M)$, i.e., $C(M)L^{\nu}_0(M) = L^{\nu}_0(M)$. Therefore, the quotion space $L^{\nu}(M)$ [mod $L^{\nu}_0(M)$] is also a representation space of $C(M)$. One has from (25) the following:

$$L^{\nu}(M) = \bigoplus_{Q^{\nu}_{P} \in I_\nu} \mathfrak{R} e_P Q^{\nu}_{P} \oplus L^{\nu}_0(M), \tag{29}$$

therefore, a basis of the above quotion space is the residue classes $e_P Q^{\nu}_{P}$ (mod $L^{\nu}_0(M)$) with $Q^{\nu}_{P} \in I_\nu$.

Now, we show that, for $Q^{\nu}_{P} \in I_\nu$,

$$e_P Q^{\nu}_{P} \in \sum_{\substack{Q^{\nu'}_{P'} \in I_\nu \\ P' \subseteq P}} \mathfrak{R} Q^{\nu'}_{P'} + L^{\nu}_0(M). \tag{30}$$

We use induction with respect to the defect of P. The statement is obvious for $d(P) = 0$. On the other hand, $e_P Q^{\nu}_{P} = Q^{\nu}_{P} - \sum_{P' \subset P} e_{P'} Q^{\nu}_{P}$, and $e_{P'} Q^{\nu}_{P}$

is equal to a sum over ν' of some terms of the form $e_{P'}Q_{P'}^{\nu'}$ [cf. (26)]. Now, if $Q_{P'}^{\nu'} \notin I_\nu$ these terms are in $L_0^\nu(M)$. If, however, $Q_{P'}^{\nu'} \in I_\nu$, then by using the induction hypothesis the proof is complete, since $P' \subset P$ implies that $d(P') < d(P)$.

It follows from (30) that

$$L^\nu(M) \equiv \sum_{Q_{P'}^{\nu'} \in I_\nu} \Re Q_{P'}^{\nu'} \quad [\text{mod } L_0^\nu(M)]. \tag{31}$$

By virtue of (29) and the one-to-one correspondence between Q_P^ν and $e_P Q_P^\nu$, one has

$$L^\nu(M) = \bigoplus_{I_\nu} \Re Q_P^\nu \oplus L_0^\nu(M). \tag{32}$$

Therefore, any basis element Q_P^ν of $L^\nu(M)$ can be expressed uniquely as

$$Q_P^\nu \equiv \sum_{Q_{P'}^{\nu'} \in I_\nu} \sum_{P' \subseteq P} \alpha_{PP'}^{\nu\nu'} Q_{P'}^{\nu'} \quad [\text{mod } L_0^\nu(M)]. \tag{33}$$

Our problem is thus to calculate the coefficients $\alpha_{PP'}^{\nu\nu'}$. It suffices to calculate the coefficients $a_{EP}^{\nu\nu'}$ which appear in the expansion of Q_M^ν with $\nu < |M| - 1$, namely,

$$Q_M^\nu \equiv \sum_{\substack{d(P')+|\mathfrak{p}'|=\nu \\ \mathfrak{p}' \geq 0}} a_{EP'}^{p\mathfrak{p}'} Q_{P'}^{\mathfrak{d}(P')+\mathfrak{p}'} \quad [\text{mod } L_0^\nu(M)], \tag{34}$$

where $p = \nu - |M| + 1$ is a negative integer.

In order to facilitate the calculation of these coefficients we shall now introduce a set of operators $\partial_{\mathfrak{M}}$ for each $\mathfrak{M} \subseteq M$ which is somewhat analogous to differentation in power series. By a suitable successive application of these operators, one can single out the term in the right-hand side of (34) whose coefficient is sought.

The Operators $\partial_{\mathfrak{M}}$

Let us define

$$\begin{aligned} \partial_{\mathfrak{M}} Q_P^{\mathfrak{d}(P)+\mathfrak{p}} &= p_i Q_P^{\mathfrak{d}(P)+\mathfrak{p}-1_i} \quad \text{if} \quad P(\mathfrak{M}) = E(\mathfrak{M}) \\ &\qquad \text{and} \quad |\mathfrak{M}_i| - 1 + p_i \neq 0 \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{35}$$

where $\mathfrak{M}_i$ is the part of P which contains $\mathfrak{M}$ and 1_i is a vector whose jth component is $\delta_{i,j}$, i.e., $1_i = (0, \cdots, 1, \cdots, 0)$. We shall write $\partial_{\mathfrak{M}}\partial_{\mathfrak{M}} = \partial_{\mathfrak{M}}^2$.

The operator $\partial_{\mathfrak{M}}$ maps $L^\nu(M)$ into $L^{\nu-1}(M)$, but it does not map $L_0^\nu(M)$ into $L_0^{\nu-1}(M)$. Therefore, one has to be careful in operating by $\partial_{\mathfrak{M}}$ on the congruence relations in the quotion space $L^\nu(M)$ [mod $L_0^\nu(M)$]. We shall demonstrate that

$$\partial_{\mathfrak{M}} L_0^\nu(M) \subseteq L_0^{\nu-1}(M) + \mathfrak{I}(\mathfrak{M}) L^{\nu-1}(M), \tag{36}$$

where $\mathfrak{I}(\mathfrak{M})$ is the ideal spanned by the partitions which divide the subset $\mathfrak{M}$ at least into two parts (cf. Sec. II). If we apply $\partial_{\mathfrak{M}}$ again on (36), we see that the last term vanishes by definition, i.e., $\partial_{\mathfrak{M}}^2 L_0^\nu(M) \subseteq L_0^{\nu-2}(M) + \mathfrak{I}(\mathfrak{M}) L^{\nu-2}(M)$.

We shall now give some of the properties of this operator. It is easy to see that

$$(P\partial_{\mathfrak{M}} - \partial_{\mathfrak{M}} P)P = 0, \tag{37}$$

which can be verified by applying on any basis element $Q_{P'}^{\nu'}$. For an arbitrary element $X = \sum_P \xi_P P \in C(M)$, we then obtain

$$e(\mathfrak{N})\partial_{\mathfrak{M}} X = e(\mathfrak{N})\partial_{\mathfrak{M}} \sum_{\substack{P(\mathfrak{N})=E(\mathfrak{N}) \\ P(\mathfrak{M})=E(\mathfrak{M})}} \xi_P P,$$

where the latter sum is over all P's which do not divide the subsets $\mathfrak{N}$ and $\mathfrak{M}$, since those which divide $\mathfrak{N}$ are annihilated by $e(\mathfrak{N})$ by virtue of (37) and the others by $\partial_{\mathfrak{M}}$. In particular, we have

$$\begin{aligned} &e(M - \mathfrak{M})\partial_{\mathfrak{M}} e(M) \\ &= e(M - \mathfrak{M})\partial_{\mathfrak{M}}[E(M) - E(\mathfrak{M})E(M - \mathfrak{M})] \end{aligned} \tag{39}$$

where we have made use of (14).

Now, we shall prove another useful formula, namely

$$E(\mathfrak{M})\partial_{\mathfrak{M}} e(M) = e(M - \mathfrak{M})\partial_{\mathfrak{M}} e(M). \tag{40}$$

Since $e(M - \mathfrak{M})$ is of the form $E(M - \mathfrak{M}) + \sum_{P \subset E(M-\mathfrak{M})} \xi_P P$, it suffices to show that, for all subsets $\mathfrak{M}' \neq \phi$, $\mathfrak{M}' \cup \mathfrak{M} \subset M$, $\mathfrak{M}' \cap \mathfrak{M} = \phi$.

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}} e(M) = 0. \tag{41}$$

The latter holds obviously when $|M - \mathfrak{M}| = 1$, since there is no subset $\mathfrak{M}'$ in this case satisfying the above conditions. We proceed by induction with respect to the number of elements in $M - \mathfrak{M}$. Substituting (15) in (41) we obtain

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}}[E(M) - \sum_{\mathfrak{M} \subseteq \mathfrak{N} \subset M} e(\mathfrak{N})E(M - \mathfrak{N})],$$

Every term for which the subset $\mathfrak{N} - \mathfrak{M}$ contains elements from both of the sets $\mathfrak{M}'$ and $M - (\mathfrak{M} \cup \mathfrak{M}')$ must vanish due to the induction hypothesis. Thus, the only contributions come from the summations over $\mathfrak{M} \subseteq \mathfrak{N} \subseteq \mathfrak{M} \cup \mathfrak{M}'$ and $\mathfrak{M} \subset \mathfrak{N} \subseteq M - \mathfrak{M}'$. Moreover, in the former sum $\partial_{\mathfrak{M}}$ commutes with $E(M - (\mathfrak{M} \cup \mathfrak{M}'))$, whereas in the latter it commutes with $E(\mathfrak{M}')$, so that we can replace $E(M - \mathfrak{N})$ by $E(\mathfrak{M} \cup \mathfrak{M}' - \mathfrak{N})E(M)$ and $E(M - \mathfrak{M}' - \mathfrak{N})E(M)$, respectively. We thus obtain

$$\begin{aligned} E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}} E(M)[E(M) &- E(\mathfrak{M}' \cup \mathfrak{M}) \\ &- E(M - \mathfrak{M}') + E(\mathfrak{M})]. \end{aligned}$$

Since $e(M)Q_P^\nu = 0$ unless $P(M) = E(M)$, then it

suffices to verify that the latter quantity vanishes when operated on Q^{ν}_{M}. One can readily see that it is indeed so after a straightforward calculation. (See Appendix for details.)

Now we are in a position to prove the statement (36). Clearly it suffices to show that the basis element $e(M)Q^{\nu}_{M}$ ($\nu < |M| - 1$) is mapped in the space $L_0^{\nu-1}(M) + \mathfrak{J}(\mathfrak{M})L^{\nu-1}(M)$.

We have from (39) and (40)

$$\begin{aligned} & e(\mathfrak{M})\partial_{\mathfrak{M}}e(M)Q^{\nu}_{M} \\ & \quad = e(\mathfrak{M})e(M - \mathfrak{M})\partial_{\mathfrak{M}}[E(M) - E(\mathfrak{M})]Q^{\nu}_{M} \\ & \quad = \sum_{\nu_1+\nu_2=\nu-1} (\nu_2 - |M - \mathfrak{M}|) \\ & \qquad \times e(\mathfrak{M})Q^{\nu_1}_{\mathfrak{M}}e(M - \mathfrak{M})Q^{\nu_2}_{M-\mathfrak{M}}. \end{aligned} \tag{42}$$

We see that the last quantity is in $L_0^{\nu-1}(M)$. On the other hand, for any $\mathfrak{M}' \supseteq \mathfrak{M}$ we have, by virtue of (38),

$$e(\mathfrak{M}')\partial_{\mathfrak{M}}e(M)Q^{\nu}_{M} = e(\mathfrak{M}')\partial_{\mathfrak{M}'}e(M)Q^{\nu}_{M} \in L_0^{\nu-1}(M), \tag{43}$$

where we have used the fact that $\partial_{\mathfrak{M}}Q^{\nu}_{M} = \partial_{\mathfrak{M}'}Q^{\nu}_{M}$ if $\mathfrak{M} \subseteq \mathfrak{M}' \subseteq M$, which follows from the definition. Now, summing (43) over all subsets $\mathfrak{M}'$ such that $\mathfrak{M} \subseteq \mathfrak{M}' \subseteq M$, we complete the proof of (36) [by virtue of (13)].

We introduce a slight generalization of (36) by writing

$$\partial^{\mathbf{p}}_{P}L^{\nu}_{0}(M) \subseteq L_0^{\nu-|\mathbf{p}|}(M) + \mathfrak{J}(P)L^{\nu-|\mathbf{p}|}(M), \tag{44}$$

where $\partial^{\mathbf{p}}_{P} = \partial^{\mathbf{p}_1}_{\mathfrak{M}_1} \cdots \partial^{\mathbf{p}_k}_{\mathfrak{M}_k}$, and $\mathfrak{J}(P)$ is the ideal spanned by the partitions P' such that $PP' \neq P$.

Calculation of the Coefficients $a^{\mathbf{pp}'}_{\mathbf{PP}'}$.

Let us select a partition P and a vector $\mathbf{p}$ satisfying the condition $d(P) + |\mathbf{p}| = \nu$, and apply the corresponding operator $\partial^{\mathbf{p}}_{P}$ on the congruence (34). The left-hand side gives $p(p-1)\cdots(p - |\mathbf{p}| + 1)Q^{d(P)}_{M}$. In the right-hand side, only partitions which contain P contribute, and (34) transforms into a congruence relation modulo $L_0^{\nu-|\mathbf{p}|}(M) + \mathfrak{J}(P)L^{\nu-|\mathbf{p}|}(M)$.

Let P_i be the restriction of P on $\mathfrak{M}_i$, i.e. $P_i = P(\mathfrak{M}_i)$ where $\mathfrak{M}_i$ are the parts of P'. Also let $\mathbf{p}_i$ denote the part of $\mathbf{p}$ corresponding to this restriction. We then have

$$\begin{aligned} \partial^{\mathbf{p}}_{P}Q^{\mathbf{d}(P')+\mathbf{p}'}_{P'} = \prod_{i=1}^{k(P')} & p'_i(p'_i - 1)\cdots \\ & \times (p'_i - |\mathbf{p}_i| + 1)Q^{m'_i}_{\mathfrak{M}_i}, \end{aligned} \tag{45}$$

where $m'_i = |\mathfrak{M}_i| - 1 + p'_i - |\mathbf{p}_i|$. Therefore, the terms in the right-hand side of $\partial^{\mathbf{p}}_{P}Q^{\nu}_{M}$ which have nonvanishing contributions must satisfy the inequality $p'_i \geq |\mathbf{p}_i|$, in addition to the conditions imposed on the summation, namely $d(P') \geq d(P)$ and $|\mathbf{p}| + d(P) = |\mathbf{p}'| + d(P')$. The only combination which satisfies all these conditions is that $P = P'$ and $p'_i = |\mathbf{p}_i|$. Thus, recalling $p < 0$, we have

$$\begin{aligned} & (-1)^{|\mathbf{p}|}(|\mathbf{p}| - 1 - p)\cdots(-p+1)(-p)Q^{d(P)}_{M} \\ & \quad \equiv a^{\mathbf{pp}}_{EP}p_1!\,p_2!\cdots p_k!\,Q^{\mathbf{d}(P)}_{P} \\ & \qquad [\mathrm{mod}\ L_0^{\mu-|\mathbf{p}|}(M) + \mathfrak{J}(P)L^{\nu-|\mathbf{p}|}(M)]. \end{aligned}$$

By making use of the fact that $Q^{d(P)}_{M}$ is congruent to $Q^{\mathbf{d}(P)}_{P}$ modulo $\mathfrak{J}(P)L^{\nu-|\mathbf{p}|}(M)$, the sought coefficients are found as multinomials

$$a^{\mathbf{pp}}_{EP} = (-1)^{|\mathbf{p}|}\binom{|\mathbf{p}|-p-1}{\mathbf{p}}. \tag{46}$$

It follows from the defining relations that $|\mathbf{p}| - p - 1 = k(P) - 2$.

VII. APPLICATION TO THE CLUSTER SERIES

The combinatorial problem of expressing the νth order, $|M|$-particle distribution function under the well-ordering assumption in terms of the irreducible set was explained in Sec. V. In order to apply the results obtained in Sec. VI to this problem, we first construct a vector space generated by the clusters of distribution functions in which the elements of $C(M)$ are defined as linear operators. Then we demonstrate that between this space and the abstract representation space developed previously there is an isomorphy which is compatible with the operations of $C(M)$.

Let $L^{F}_{\nu}(M)$ denote the vector space spanned (over the field $\mathfrak{R}$) by the (ordinary) products $F^{\boldsymbol{\nu}}_{P} = F^{\nu_1}_{\mathfrak{M}_1}\cdots F^{\nu_k}_{\mathfrak{M}_k}$ of the distribution functions where $\nu_i \geq 0$. Define the operation $P = (\mathfrak{M}_1, \cdots, \mathfrak{M}_k)$ such that

$$PF^{\nu}_{M} = \sum_{|\mathbf{v}|=\nu} \prod_{i=1}^{k} F^{\nu_i}_{\mathfrak{M}_i}.$$

Also assume that this operation satisfies the axioms given in the previous section.

It is readily seen that this space is operator isomorphic to the representation space $L^{\nu}(M)$ of the previous section through the correspondence $Q^{\nu}_{\mathfrak{M}} \leftrightarrow F^{\nu}_{\mathfrak{M}}$. We also observe that under this correspondence the elements $e(\mathfrak{M})Q^{\nu}_{\mathfrak{M}}$ are mapped onto the correlation functions $G^{\nu}_{\mathfrak{M}}$, so that the space $e_PL^{\nu}(M)$ is isomorphic to the space generated by the clusters of correlation functions $G^{\boldsymbol{\nu}}_{P}$ of order ν ($= |\mathbf{v}|$).

The subspace of L^{F}_{ν} which corresponds to $L^{\nu}_{0}(M)$, on the other hand, is generated by such clusters $G^{\boldsymbol{\nu}}_{P}$ that there is at least one factor $G^{\nu_i}_{\mathfrak{M}_i}$ satisfying

the condition $\nu_i < |\mathfrak{M}_i| - 1$. Therefore, if correlations are well-ordered, this subspace contains only the zero element. This implies that under the well-ordering assumption the space $L^F_\nu(M)$ is operator isomorphic to the quotion space $L^\nu(M)[\text{mod } L^\nu_0(M)]$. Thus, the congruence relations previously given are valid as equations in $L^F_\nu(M)$.

The following expansion of F^ν_M in terms of the irreducible functions is then obtained:

$$F^\nu_M = \sum_{\substack{d(P)+|\mathfrak{p}|=\nu \\ \mathfrak{p}\geq 0}} (-1)^{|\mathfrak{p}|}\binom{k(P)-2}{\mathfrak{p}} F_P^{d(P)+\mathfrak{p}}, \qquad (\nu < |M| - 1).$$

VIII. DISCUSSION

In the present investigation, an attempt is made to study the algebraic foundations of the cluster expansion which is widely used in statistical mechanics. For this purpose, *the partition algebra, $C(M)$*, is introduced and studied to some extent. Generators of this algebra are partitions which are regarded as operators on finite sets M. The set of minimal idempotents of this algebra is constructed explicitly. In physical problems, the elements of the set M correspond to state variables of particles in an interacting system and the partitions are interpreted as operators which remove the interactions among certain clusters of particles. It turns out that the correlation functions of statistical mechanics correspond to the representation of the minimal idempotents of $C(M)$ in an appropriately constructed representation space of this algebra. This interesting result facilitates the demonstration of the equivalence of the cluster vs product properties under rather general conditions.[8]

It is expected that the techniques presented herein will be useful in handling the combinatorial problems associated with the use of the cluster expansion. To demonstrate this point, we have considered the cluster expansion in which the distribution and correlation functions are expanded in powers of a small parameter. When correlations are well-ordered, i.e., G_M is of order $\lambda^{|M|-1}$, the distribution functions to a given order are not all independent and can be expressed in terms of a finite irreducible set of distribution functions. As an application of the present technique the expansion coefficients are calculated. The well-ordering of correlations is a property of many physical systems. For example, in plasmas in equilibrium and also under Bogoliubov hypothesis correlations are well-ordered with respect to the interaction strength.[9] The same order property is usually assumed in the nonequilibrium initial-value problem associated with the BBGKY chain. Hence, the above result is applicable to a wide range of problems.

On the basis of above applications we hope that the algebraic techniques described presently will be useful in statistical mechanics.

[8] Since this paper was submitted, a paper by S. Sherman appeared in J. Math. Phys. 5, 1137 (1964) in which the equivalence of product property and cluster property is proved by a different method. Sherman uses the generalized Bohnenblust algebra and obtains the result in a simple fashion. However, we feel that the use of the partition algebra provides an insight to the understanding of the structure of the cluster relations and, moreover, it seems to have the possibility of other applications.

APPENDIX

In this appendix we shall demonstrate in detail that

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}}[E(M) - E(\mathfrak{M}\cup\mathfrak{M}') - E(M-\mathfrak{M}') + E(\mathfrak{M})]Q^\nu_M = 0.$$

The first term gives, using the definition of $\partial_{\mathfrak{M}}$,

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}}Q^\nu_M = (\nu - |M| + 1)\sum_{\nu_1+\nu_2+\nu_3=\nu-1} Q^{\nu_1}_{\mathfrak{M}}Q^{\nu_2}_{\mathfrak{M}'}Q^{\nu_3}_{M'},$$

where we set $M' = M - (\mathfrak{M}'\cup\mathfrak{M})$. The second term yields

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}}E(\mathfrak{M}\cup\mathfrak{M}')Q^\nu_M = E(\mathfrak{M})\sum_{\nu_1+\nu_2=\nu-1}(\nu_1 - |\mathfrak{M}'\cup\mathfrak{M}| + 2)Q^{\nu_1}_{\mathfrak{M}\cup\mathfrak{M}'}Q^{\nu_2}_{M'}$$

$$= \sum_{\nu_1+\nu_2+\nu_3=\nu-1}(\nu_1+\nu_2-|\mathfrak{M}|-|\mathfrak{M}'|+2)Q^{\nu_1}_{\mathfrak{M}}Q^{\nu_2}_{\mathfrak{M}'}Q^{\nu_3}_{M'}.$$

The third term yields

$$E(\mathfrak{M})E(\mathfrak{M}')\,\partial_{\mathfrak{M}}E(M-\mathfrak{M}')Q^\nu_M = E(\mathfrak{M})\sum_{\nu_2+\nu_3=\nu-1}(\nu_3 - |M| + |\mathfrak{M}'| + 2)Q^{\nu_2}_{\mathfrak{M}'}Q^{\nu_3}_{M-\mathfrak{M}'}$$

$$= \sum_{\nu_1+\nu_2+\nu_3=\nu-1}(\nu_1+\nu_3-|M|+|\mathfrak{M}'|+2)Q^{\nu_1}_{\mathfrak{M}}Q^{\nu_2}_{\mathfrak{M}'}Q^{\nu_3}_{M'},$$

and finally for the last term

$$E(\mathfrak{M})E(\mathfrak{M}')\partial_{\mathfrak{M}}E(\mathfrak{M})Q^\nu_M = E(\mathfrak{M}')\sum_{\nu_1+\nu_2=\nu-1}(\nu_1 - |M| + 2)Q^{\nu_1}_{\mathfrak{M}}Q^{\nu_2}_{M-\mathfrak{M}}$$

$$= \sum_{\nu_1+\nu_2+\nu_3=\nu-1}(\nu_1 - |\mathfrak{M}| + 2)Q^{\nu_1}_{\mathfrak{M}}Q^{\nu_2}_{\mathfrak{M}'}Q^{\nu_3}_{M'}.$$

Collecting the terms, and observing that $\nu_1 + \nu_2 + \nu_3 = \nu - 1$ and $|M| = |\mathfrak{M}| + |\mathfrak{M}'| + |M'|$ we readily arrive at the desired result.

[9] K. İmre and E. Özizmir, Nucl. Fusion 4, 1, 105, 179 (1964).

Cluster–Star Inversion and the Möbius Formula*

S. SHERMAN
Department of Mathematics, Indiana University, Bloomington, Indiana
(Received 30 November 1964)

The cluster–star equations are inverted by using the general Möbius inversion formula. Since Uhlenbeck and Ford have many instances of systems of equations to which general Möbius inversion can be applied, this paper first presents a general survey of Möbius inversion, then some graph terminology, and finally the cluster–star inversion.

I. INTRODUCTION

IN Ref. 1, Uhlenbeck and Ford raised the problem of inverting the relation they gave between the cluster functions $U_m(r_1, \cdots, r_m)$ and the star functions $V_m(r_1, \cdots, r_m)$. This we do here using the general Möbius inversion formula.[2] [UF] has many instances[3] of systems of equations to which the general Möbius inversion formula may be applied. In the case at hand, cluster–star inversion, the answer is of physical interest and the Möbius function is new, so we present the details here. First we give a general survey of Möbius inversion, then some graph terminology and the cluster–star inversion.

Before continuing we should remark that (1) the Möbius function has already been used as an aid in solving a statistical mechanical problem by Nettleton and Green[4] who obtained an expression for the entropy density of a statistical mechanical system in terms of molecular distribution functions and (2) an alternative method using generating functionals for cluster–star inversion due to Stell is given in an adjoining article together with additional physical and historical remarks.[5]

II. INCIDENCE ALGEBRA

Since the general Möbius inversion formula can be applied to physical problems other than those associated with the specific inversion presented here, we give a survey of the general theory and then follow it by some graph terminology and the cluster–star inversion.

Let P be a *partially ordered set* whose relation is written $\leq$ so that

$$x \leq y \quad \text{and} \quad y \leq z \quad \text{imply} \quad x \leq z$$

and

$$x \leq y \quad \text{and} \quad y \leq x \quad \text{imply} \quad x = y.$$

We write $x \not\leq y$ if it is not true that $x \leq y$.

Examples:

1. $P_1 = N = \{1, 2, 3, \cdots\}$, the natural numbers, with the ordinary order so that $2 \leq 5$ and $2 \leq 6$.

2. $P_2 = N = \{1, 2, 3, \cdots\}$, the natural numbers, with $x \leq y$ meaning x divides y so that $2 \leq 6$ but $2 \not\leq 5$.

3. $P_3 = 2^A$, the collection of subsets of A with $x \leq y$ meaning $x \subset y$. Thus, if $A = \{1, 2, 3, 4, 5\}$, then $\{1, 2, 3\} \leq A$ but $\{1, 2, 3\} \not\leq \{2, 3, 4\}$.

4. P_4 is the collection of all *partitions* of A. A partition of A is a collection of disjoint subsets (called *blocks*) of A which exhaust A. Let $x \leq y$ mean each block in x is a subset of some block in y. If $A = \{1, 2, 3, 4, 5\}$, then $\{\{1, 2\}, \{3, 4\}, \{5\}\} \leq \{\{1, 2, 3, 4\}, \{5\}\}$ and $\{\{1, 2\}, \{3, 4\}, \{5\}\} \not\leq \{\{1, 2, 3\}, \{4, 5\}\}$.

The *segment* $[x, y] = \{z: x \leq z \leq y\}$, the set of all z greater than or equal to x and less than or equal to y. A partially ordered set is *locally finite* if every segment is finite. P_1 and P_2 are locally finite. If A is finite, P_3 and P_4 are locally finite, otherwise not. The *product*, $P \times Q$, of partially ordered sets P and Q is the partially ordered set $\{(p, q): p \in P \text{ and } q \in Q\}$ where $(r, s) \leq (p, q)$ means $r \leq p$ in P and $s \leq q$ in Q.

Suppose P is a locally finite partially ordered set. We can now define the *incidence algebra* of P by first considering the set of all real-valued functions $h(x, y)$, where both $x, y \in P$ and such that $h(x, y) = 0$

* The work of S. Sherman was supported in part by the National Science Foundation, Grant No. G 24334 and GP 3941.

[1] G. E. Uhlenbeck and G. W. Ford, "Theory of Linear Graphs," in *Studies in Statistical Mechanics, I* edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1962). This reference will be referred to as [UF]. See pp. 158–9.

[2] Historically significant references are Weisner, L., Trans. Am. Math. Soc., 38, 474 (1935); P. Hall, 7, Quart. J. Math. 1, 134 (1936); and M. Ward, Duke Math. J. 5, 357 (1939). The author learned about this theory from G.-C. Rota, Z. Wahrscheinlichkeits-theorie Verw. 2, 340 (1964). Section II is a summary of relevant parts of Rota. The numbering of propositions is that used by Rota.

[3] See Eqs. II (10), (11), (24), (28), and (39) and III (9), (30), and (33) in [UF].

[4] R. E. Nettleton and M. S. Green, J. Chem. Phys. 29, 1565 (1958).

[5] G. Stell, J. Math. Phys. 6, 1193 (1965) (following article).

if $x \not\leq y$. Generalizations to other ranges are possible but not considered here. If g and h are two such functions the sum of g and h defined by $(g+h)(x, y) = g(x, y) + h(x, y)$ is such a function. Similarly a scalar multiple of such a function is such a function. We define the *product* $k = gh$ by

$$k(x, y) = \sum_{x \leq z \leq y} g(x, z)h(z, y).$$

Since P is locally finite, the sum on the right is over a finite set and no convergence problem arise. Note that if $x \not\leq y$, then $k(x, y) = 0$. The collection of functions considered here with operations of addition, multiplication and multiplication by a scalar constitutes a real associative algebra, the *incidence algebra* [written $\mathfrak{A}(P)$ or $\mathfrak{A}$ if P is clear]. In general $\mathfrak{A}$ is not commutative. There is an obvious parallel between matrix operations and those in $\mathfrak{A}(P)$. The identity element of $\mathfrak{A}$ is δ, the Kronecker delta, defined as usual by

$$\delta(x, y) = \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases}$$

Thus $h\delta = \delta h = h$. We define the *zeta function* $\zeta \in \mathfrak{A}(P)$ by

$$\zeta(x, y) = \begin{cases} 1, & x \leq y, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 1. The zeta function of a locally finite partially ordered set is invertible in the incidence algebra.

Proof. Let

$$\mu(x, y) = \begin{cases} 1, & x = y, \\ -\sum_{x \leq z < y} \mu(x, z), & x < y, \\ 0, & \text{otherwise.} \end{cases}$$

This constitutes a recursive definition. Note that $\mu\zeta(x, x) = 1.\ 1 = 1$ and if $x < y$, then

$$\mu\zeta(x, y) = \sum_{x \leq z \leq y} \mu(x, z)\ \zeta(z, y)$$
$$= \left(\sum_{x \leq z < y} \mu(x, z)\right) + \mu(x, y) = 0.$$

Noting the parallel with matrices we also have $\zeta\mu = \delta$ and μ, which is called the *Möbius function*, is the unique inverse to ζ.

Examples:

(1) In P_1

$$\mu(m, n) = \begin{cases} 1, & n = m, \\ -1, & n = m + 1, \\ 0, & \text{otherwise.} \end{cases}$$

(2) In P_2

$$\mu(d, n) = \begin{cases} 0, & d \nmid n, \\ \tilde{\mu}(n/d), & d \mid n. \end{cases}$$

Here $\tilde{\mu}(m)$ is the classical Möbius function; $\tilde{\mu}(m) = (-1)^k$ if m is square-free and k is the number of prime factors in m, while $\mu(m) = 0$, otherwise.

(3) In P_3, with A finite,

$$\mu(x, y) = (-1)^{n(y)-n(x)},$$

where $x \leq y$ and $n(x)$ is the number of elements in x.

(4) In P_4, the Möbius function[6] is more complicated. If $A = \{1, 2, 3, \cdots, a\}$ with $a < \infty$ then

$$\mu(\{\{1\}, \{2\}, \cdots, \{a\}\}, \{\{1, 2, 3, \cdots, a\}\}) = (-1)^{a-1}(a - 1)!.$$

Proposition 2 (Möbius inversion formula). Let f be a real-valued function such that, for each x in locally finite P, f vanishes at all except a finite number of $y \leq x$. Suppose

$$F(x) = \sum_{y \leq x} f(y).$$

Then

$$f(x) = \sum_{y \leq x} F(y)\mu(y, x).$$

Proof.

$$F(x) = \sum_{y} f(y)\zeta(y, x).$$

Thus

$$\sum_{x} F(x)\mu(x, z) = \sum_{x} \sum_{y} f(y)\zeta(y, x)\mu(x, z)$$
$$= \sum_{y} f(y)\delta(y, z) = f(z).$$

Examples:

(1) In P_1, the Möbius inversion formula asserts $F(n) = \sum_{j=1}^{n} f(j)$ implies $f(n) = g(n) - g(n-1) = \Delta p(n)$, the first difference of p at n. Thus in this case the Möbius inversion formula is a discrete version of the fundamental theorem of the calculus.

(2) In P_2, the Möbius inversion formula asserts

$$F(n) = \sum_{d \mid n} f(d)$$

and implies

$$f(n) = \sum_{d \mid n} F(d)\mu\left(\frac{n}{d}\right).$$

[6] Schützenberger, M. P., *Contributions aux Applications Statistiques de la Théorie de L'Information*, Pub. Inst. Stat. Univ. Paris, 3, 1 (1954), especially p. 25, and R. Frucht and G. C. Rota, "A Möbius Type Inversion Formula for Partitions," Notices Am. Math. Soc. [Abstract 63 T-250 10 (August, 1963), p. 495].

(3) In P_3, the Möbius inversion formula asserts that if $A = \{1, 2, \cdots, a\}$ and $F(x) = \sum_{y \le x} f(y)$, then $f(x) = \sum_{y \le x} (-1)^{n(x-y)} g(y)$.

Proposition 5. Let $P \times Q$ be the direct product of locally finite partially ordered sets P and Q. The Möbius functions of $P \times Q$ is given by

$$\mu((x, y), (u, v)) = \mu(x, u)\mu(y, v).$$

Note the use of a single letter to denote the Möbius functions of three partially ordered sets.

Proof. It is trivial that $\mu((x, y), (x, y)) = 1 = \mu(x, x)\mu(y, y)$. Suppose for each $(z_1, z_2) < (u, v)$, $\mu((x, y), (z_1, z_2)) = \mu(x, z_1)\mu(y, z_2)$. Then

$$\begin{aligned}
\mu((x, y),(u, v)) &= -\sum_{(x,y) \le (z_1,z_2) < (u,v)} \mu((x, y),(z_1, z_2)) \\
&= -\sum_{x \le z_1 < u} \mu(x, z_1)\mu(y, v) \\
&\quad -\sum_{y \le z_2 < v} \mu(x, u)\mu(y, z_2) \\
&\quad -\sum_{x \le z_1 < u, y \le z_2 < v} \mu(x, z_1)\mu(y, z_2) \\
&= \mu(x, u)\mu(y, v) + \mu(x, u)\mu(y, v) \\
&\quad -\mu(x, u)\mu(y, v) \\
&= \mu(x, u)\mu(y, v).
\end{aligned}$$

III. GRAPH TERMINOLOGY

While the terminology of [UF] will be followed, some additional comments useful for this problem are made here. Among the labeled connected graphs are the labeled complete stars. These are labeled graphs such that each pair of different points or vertices is joined by a single line. In graph G, complete star K is *maximal* if K is not a subgraph of a larger complete star in G. A *complete-star tree* (c-s tree) is a graph such that each line is part of a unique maximal complete star and if points P_1 and P_2 (P_1 may be equal to P_2) are in the same maximal complete star K, then P_1 may not be connected to P_2 by a sequence of different lines not in S. A labeled c-s tree can be regarded as a collection of labeled complete stars hung together at articulation points.

Suppose we consider the collection of labeled c-s trees with points labeled $1, 2, \cdots, n$. This collection of c-s trees can be regarded as a partially ordered set if $x \le y$ means each line of x is a line of y. K_n the labeled complete star with n points labeled $1, 2, \cdots, n$ satisfies $x \le K_n$ for each x.

Let $\alpha(x)$ be the *number of articulation points* of x and let $A(x) = \{[1; x], \cdots, [\alpha(x); x]\}$ be the set of articulation points of x. If $x = K_n$, then $\alpha(x) = 0$. Let $\sigma(x)$ be the *number of maximal complete stars* in x and let $K(x) = \{K(1; x), \cdots, K(\sigma; x)\}$ be the *set of maximal complete stars* in x. If $x = K_n$, then $\sigma(x) = 1$ and $K(x) = \{x\}$. Let $\delta([i; x])$ be the *number of maximal complete stars* containing $[i; x]$ and let $\mathfrak{B}([i; x]) = \{K(j; x) : [i ; x] \in K(j ; x)\}$ be the set of those maximal complete stars, $1 \le i \le \alpha(x)$, which contain $[i; x]$. Of course, $n(\mathfrak{B}([i; x])) = \delta([i; x])$, where $n(A)$ means the numbers of elements in A. While $\bigcup_{i=1}^{\alpha(x)} \mathfrak{B}([i; x]) = K(x)$, if $\alpha(x) \ge 2$, then the $\mathfrak{B}([i; x])$ may fail to be disjoint. Thus $\{\mathfrak{B}([i; x]) : 1 \le i \le \alpha([i; x])\}$ in general fails to be a partition of $K(x)$. A relation between $\sigma(x)$ and the $\delta([i; x])$ is

$$1 + \sum_{i=1}^{\alpha(x)} (\delta([i; x]) - 1) = \sigma(x).$$

This may be proved by starting with a complete star, successively hanging complete stars at articulation points, and observing the change at each stage on both sides of the equation. Any c-s tree can be generated by so hanging complete stars on a complete star.

We associate with labeled c-s tree x having n points two functions $\mathfrak{U}_x$ and $\mathfrak{V}_x$ of n variables $r_1, r_2, \cdots, r_n$ satisfying the following requirements:

(1) If x is a labeled complete star, then $\mathfrak{U}_x$ and $\mathfrak{V}_x$ are symmetric functions of the r's whose subscripts are labels of points of x.

(2) If x has the set of maximal complete stars $K(x) = \{K(j; x) : 1 \le j \le \sigma(x)\}$, then

$$\mathfrak{U}_x = \prod_{j=1}^{\sigma(x)} \mathfrak{U}_{K(j;x)}$$

and

$$\mathfrak{V}_x = \prod_{j=1}^{\sigma(x)} \mathfrak{V}_{K(j;x)}.$$

IV. CLUSTER-STAR INVERSION

Equation (III,30) of [UF] is equivalent to the following: for each $n \ge 2$ in the partially ordered set of labeled c-s trees with n points

$$\mathfrak{U}_x = \sum_{y \le x} \mathfrak{V}_y.$$

The translation between the [UF] notation and that used here is as follows: Suppose x is a labeled c-s tree with n points. Then $\mathfrak{U}_x$ as used here corresponds to $\mathfrak{U}_n$ in [UF]. Note that there are more equations here since x may fail to be a complete star. The new equations introduced here follow trivially from the definitions here and Equation (III,30) in [UF]. It should be mentioned that the set up here is more general than that in [UF] since we allow $\mathfrak{U}_x \ne \mathfrak{U}_y$

FIG. 1.

even though x and y are complete stars with the same number of points, e.g., if x and y have three points and x is labeled by 1, 2, and 3, while y is labeled by 2, 3, and 4, there is nothing in our considerations that forces us to assume $\mathfrak{U}_x = \mathfrak{U}_y$.

The question raised in ([UF], p. 159) is the inversion of the equations and this is given by

$$\mathfrak{V}_x = \sum_{y \leq x} \mathfrak{U}_y \mu(y, x),$$

where $\mu(y, x)$ is the Möbius function for the partially ordered set of c-s trees with n points. Since the solution for $x = K_n$ immediately yields the general solution, we consider only the case $x = K_n$. This is also the only case considered in [UF].

4. EVALUATION OF MÖBIUS FUNCTION

The evaluation of the Möbius function $\mu(y, x)$ is the next order of business. To that end we examine the partially ordered set of labeled c-s trees with n points.

Suppose $y_1 \leq y_2$, then $A(y_2) \subset A(y_1)$. Thus for each $[i; y_2]$ there is a unique $[i_2; y_1]$. If $\mathcal{K} = \{K^1, K^2, \cdots, K^m\}$ is a set of labeled complete stars let $\mathcal{K}^*$ be the labeled complete star containing all the points of $K^1 \cup K^2 \cup \cdots \cup K^m$. Each $[j; y_1]$ in the complement of $A(y_2)$ with respect to $A(y_1)$ is a pseudoarticulation point of c-s degree 1 of y_1 in the sense that all the lines of y_1 incident with $[j; y_1]$ lie in a minimal complete star "$K([j; y_1]; y_2)$" of y_2 which may fail to be a maximal complete star of y_1. This is illustrated in Fig. 1 where $y_1 \leq y_2$, $A(y_1) = \{2, 3, 4\}$, $A(y_2) = \{2\}$, "$K([4; y_1], y_2)$" is the complete star associated with the set of points $\{3, 4, 5\}$ and "$K([4; y_1], y_2)$" is not a maximal complete star of y_2.

In the original definitions each nonarticulation point could have been called a pseudoarticulation of c-s degree 1. With that convention the relation between $\sigma(x)$ and c-s degrees would be maintained where now the summation could be over all points of c-s tree.

To continue with the analysis of $y_1 \leq y_2$, we note that each $K(j_1; y_1)$ is a subgraph of some unique $K(j_2; y_2)$. From the above remarks we see also that the segment $[y_1, K_n]$ as a partially ordered set is isomorphic to

$$\bigtimes_{i=1}^{\alpha(y_1)} [\{\{1\}, \{2\}, \cdots, \{\delta[i; y_1]\}\}, \times \{\{1, 2, \cdots, \delta[i; y_1]\}\}],$$

i.e., to a product of initial segments in the partially ordered set of partitions. From Proposition 5 and Example 3 after Proposition 1 we conclude that[7]

$$\mu(y_1, K_n) = \prod_{i=1}^{\alpha(y_1)} (-1)^{\delta([i;y_1])-1}(\delta([i; y_1]) - 1)!$$

$$= (-1)^{1+\sigma(y_1)} \prod_{i=1}^{\alpha(y_1)} (\delta([i; y_1]) - 1)!.$$

Thus

$$\mathfrak{V}_{K_n} = \sum_{y \leq K_n} \mathfrak{U}_y (-1)^{1+\sigma(y)} \prod_{i=1}^{\alpha(y)} (\delta([i; y]) - 1)!.$$

[7] The partially ordered set of labeled c-s trees with n points is what Rota calls a *geometric lattice or matroid* which he studies in Sec. 7 of the reference cited in (2). The rank of x is $n - \sigma(x)$ which explains the dependence of the sign of $\mu(x, K_n)$ on $\sigma(x)$.

Cluster–Star Inversion by Means of Generating Functionals*

G. STELL†
Courant Institute of Mathematical Sciences, Division of Electromagnetic Research, New York University, New York, New York
(Received 14 December 1964)

The cluster–star equations are inverted. These equations, Eq. (2.1), define a set of cluster functions U_m in terms of star functions V_n for $2 \leq n \leq m$. The inversion, Eq. (2.2) gives the V_n in terms of the U_m; $2 \leq m \leq n$.

The cluster–star inversion yields the coefficients β_n that appear in the number-density expansion of the logarithm of the grand partition function Ξ of a statistical mechanical system in terms of the coefficients b_n that appear in its fugacity expansion, even when the system is in the presence of an external field.

In the course of obtaining the inversion, an expression for the work necessary to bring a particle from infinity to a point r inside a classical system at equilibrium is obtained in terms of the V_m and the one-particle distribution function $\rho(r)$, and also in terms of the U_m and ρ. The expression in terms of the V_m and ρ has been previously derived by others with restrictions on the form of the potential energy; the alternative expression in terms of the U_m seems to be new.

Operators with essentially the same algebraic structure as the V_m appear in a recently considered asymptotic expansion of the distribution functions of a nonequilibrium system. The cluster–star inversion facilitates the determination of these operators in terms of the solution operators of the n-body Liouville equation, $n \geq 2$.

I. INTRODUCTION

In their monograph, *The Theory of Linear Graphs*, Uhlenbeck and Ford[1] pointed out that the problem of inverting the relationship that defines the cluster or Ursell functions $U_m(r_1, \cdots, r_m)$ in terms of the star or Husimi functions $V_m(r_1, \cdots, r_m)$ had not yet been solved at the time the monograph was written.

Integrals of the U_m and the V_m appear as coefficients in the fugacity and number-density expansions, respectively, of the logarithm of the grand partition function of a statistical mechanical system. It is a simple matter to characterize the U_m as a sum of products of the V_n, $n \leq m$, by representing the latter as linear graphs [see Eq. (2.1)] or by using some equivalent notational device. Also it is clear that any V_m can be written as a linear combination of products of U_n, $n \leq m$. The inversion problem is to find the rule for writing the coefficients in this linear combination. This rule is given in Eq. (2.2) in language appropriate to the graphical notation we have used to write the equations.

Our method of obtaining the inversion is an application of the formalism that has been independently developed and used by a number of workers[2] in the past few years to obtain the general terms of a variety of cluster expansions. This formalism is based on functional differentiation and a simple procedure for resumming series. An alternative method of obtaining the cluster–star inversion, due to Sherman, is given in a separate article.[3]

An equation that is closely related to our result was first obtained by Mayer,[4] and later by Kilpatrick.[5] It involves the integrals of $U_m(r_1, \cdots, r_m)$ and $V_m(r_1, \cdots, r_m)$,

$$b_m = \frac{1}{m!} \int U_m(r_1, \cdots, r_n)\, dr_2 \cdots dr_m,$$

$$\beta_{m-1} = \frac{1}{(m-1)!} \int V_m(r_1, \cdots, r_m)\, dr_2 \cdots dr_m.$$

Here we assume that each integration is over all r_i such that $|r_i| \leq \infty$. The equation is

$$\beta_k = \sum \left\{ (-1)^{\sum_i m_i - 1} [(k - 1 + \sum_i m_i)!/k!] \times \prod_j \frac{(jb_j)^{m_j}}{m_j!} \right\},$$

where j runs between 2 and $k + 1$ and the first sum is over all sets of nonnegative integers $\{m_2, m_3 \cdots\}$, such that $\sum_i (j - 1)m_i = k$. The derivations of this equation given by Mayer and Kilpatrick avoid the problem of inverting the cluster–star equations themselves. However the equation is only correct when the integral of each product of the $V_m(r_1, \cdots, r_m)$

* This work was supported by the National Science Foundation under Grant No. GP 2003.

† Present address: The Belfer Graduate School of Science, Yeshiva University, New York, New York.

[1] G. E. Uhlenbeck and G. W. Ford, "Theory of Linear Graphs," in *Studies in Statistical Mechanics. I*, edited by J. de Boer and G. E. Uhlenbeck. This reference will be referred to as UF.

[2] See, for example, T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) 25, 531 (1961) or C. De Dominicis, J. Math. Phys. 3, 983 (1962).

[3] S. Sherman, J. Math. Phys. 6, 1189 (1965) (preceding article).

[4] J. E. Mayer, J. Chem. Phys. 10, 629 (1942).

[5] J. E. Kilpatrick, J. Chem. Phys. 21, 274 (1953).

that appears in the cluster–star expression for $U_m(r_1, \cdots r_m)$ can be factored into a product of the integrals. This happens to be the case when these functions are associated in the usual way[1] with a physical system that is free of external fields and hence is translation invariant. Our expression for each V_m in terms of U_n, $n \leq m$ immediately yields the general relationship between the b_m and the β_m, which reduces to Mayer's equation for the special case in which the integrals of the products of V_m's can be factored into products of integrals. Our result cannot be recovered from Mayer's equation, however, since information is irretrievably lost when the integrals are factored. Recently Wu[6] obtained a result that is equivalent to the general expression for β_m in terms of b_n, $n \leq m$, and hence is equivalent to our result. His method bears no obvious resemblance to the procedures used by either Sherman[3] or the author.

Although the equation giving the β_m in terms of the b_n was first considered in connection with a classical statistical mechanical system at equilibrium, it is possible to express the V_n directly in terms of the interaction potential and temperature[7] for such a system. Hence it is unnecessary to find the inversion in order to obtain cluster expansions for thermodynamical quantities and the n-particle distribution functions in terms of the one-particle distribution function, interaction potential, and temperature. On the other hand, the U_n can themselves be expressed simply in terms of the potential and temperature[1] and hence the inversion leads to a second representation of the V_n as functions of these quantities. For most purposes the first representation, which can be expressed in terms of irreducible clusters of Mayer f-functions[7] (and suitable generalizations of f-functions if n-body forces are present for $n > 2$), is probably more useful.

Recently, asymptotic expansions have been derived for the n-particle distribution functions of a nonequilibrium system. One of these expansions[8] involves operators $\mathfrak{V}_m$ with the same algebraic structure as the V_m (except for a certain lack of commutitivity.) The cluster–star inversion enables one to give a representation of these operators in terms of the solution operators of the n-body Liouville equation, $n = 2, 3, \cdots$. The author has investigated an alternative representation, analogous to the f-function representation, and intends to make a comparison of the utility of the two representations in a future publication.

When used to obtain the cluster–star inversion, the generating-functional method yields as a bonus an expression for the work necessary to bring a particle from infinity to a point inside a classical system at equilibrium in terms of the one-particle distribution function $\rho(r)$ and the star functions V_m. The expression (from which the $\ln \Xi$ in terms of ρ and the V_m can be easily obtained) is valid for a system with n-body terms with no restriction on n, i.e., for a nonuniform system with a nonadditive interaction potential. Similar derivations of this expression have previously been given for uniform systems[9] (for which $n \geq 2$) and systems with additive potentials (for which $n \leq 2$).[2] Our alternative representation in terms of ρ and the U_m, rather than the V_m, seems to be new.

As the above paragraph suggests, when the generating functional formalism is applied to a problem associated with a physical system, it is apt to be rich in objects that have physical significance. This seems to us an attractive feature, and we have made no attempt to bypass any of the structure of the method in order to save a few steps in getting out the inversion prescription as quickly as possible. Sherman's use of the Möbius inversion formula, on the other hand, is ideally suited to obtaining the inversion with a minimum of apparatus.

II. SOLUTION BY MEANS OF GENERATING FUNCTIONALS

1. Terminology and Lemmas

In this section we follow the notation of UF closely but for convenience we make a slight conceptual generalization. We represent products of functions and integrals of such products in a way that is a straightforward extension of the use of linear graphs to represent products of functions of two vector variables and their integrals. Just as *linear graphs* are collections of vertices, between certain pairs of which are lines, our *graphs* are collections of vertices, between certain pairs of which are lines (one-dimensional faces), among certain triplets of which are triangles (two-dimensional

[6] F. Y. Wu, J. Math. Phys. 4, 1438 (1963).

[7] See for example Ref. 1, p. 150 for V_m in terms of the Mayer f-function where the potential consists only of two-body terms. A discussion of the more general non-additive case can be found in Friedman's book, Ref. 13.

[8] E. G. D. Cohen, Physica 28, 1025 (1962). A similar expansion that does not involve the introduction of the V_m has been given by M. S. Green and R. A. Piccerelli, Phys. Rev. 132, 1388 (1963). In order to directly compare the two expansions, the cluster–star inversion must be used.

[9] See for example, K. Husimi, J. Chem. Phys. 18, 682 (1950).

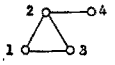
faces), and, in general, among certain n-tuples of which are $(n-1)$-dimensional faces.[10]

A vertex is called an *articulation point* if upon its removal, the graph of which it is a part separates into two or more pieces. Here the removal of a vertex is meant to imply the removal of a neighborhood of the vertex so that all connections at the vertex are broken. A graph containing no articulation points is an *irreducible* graph.

We call a graph *simple* if any two faces share at most a single vertex. Thus

and


are not simple, while

and


are simple. *Simply connected* will have its usual meaning. Thus the third graph shown above is simply connected and the fourth is not. In dealing with nonsimple graphs one can think of connectivity either in terms of the point set associated with the rectilinear realization of a graph or in terms of faces that are considered to be connected only at vertices,

e.g., 

can be considered as a simply connected traingular point set or as the doubly connected graph 

Since we shall consider only simple graphs in this paper, the question of which convention is most convenient to use does not arise here.

We refer to a vertex that is used to represent the function $v(r_1)$ as a v-vertex; a 1-vertex is the vertex that represents the function unity. We call a face that represents the function $F_n(r_1, r_2, \cdots, r_n)$ an F_n-face. Here $v(r_1)$ is a function of a single variable r_1 and $F_n(r_1, r_2, \cdots, r_n)$ is a symmetric function of the variables $r_1, r_2, \cdots, r_n$.

In the terminology of UF, a *complete star* is a linear graph such that between every pair of vertices there is a line. An F_n-face can be identified with a complete star containing n 1-vertices and $\frac{1}{2}[n(n-1)]$ f-lines by setting

$$F_n(r_1, r_2, \cdots, r_n) = \prod_{1\le i<j\le n} f(r_i, r_j). \qquad (1,1)$$

When (1.1) is used, a *complete star tree* of UF is identical with the simply connected simple graph of this section.

In our correspondence between graphs and integrals a black vertex is not labeled and corresponds to a function of a dummy variable over which integration is performed, while a white vertex is labeled by the variable with which it is associated. Thus a white ρ-vertex labeled r_2 is a vertex that represents the function $\rho(r_2)$. To each integral there corresponds not just a graph, but a graph times σ where σ is the symmetry number of the graph. This is the order of the group of permutations of the black vertices that leave all of the connections of the circles invariant. An example of the correspondence between integrals and graphs is shown below.

$$= \frac{1}{6}\int dr_2\, dr_3\, dr_4\, g(r_1)f(r_2)f(r_3) \times g(r_4)F_4(r_1, r_2, r_3, r_4).$$

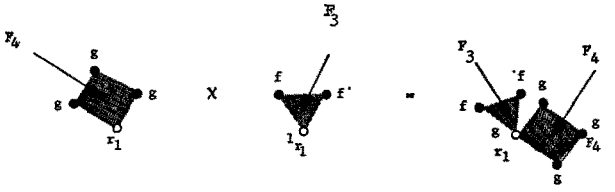
In this example $\sigma = 3! = 6$ since any permutation of dummy labels affixed to the black vertices would leave such a labeled graph unchanged.[11] We note that such dummy labels do not actually appear on the graph, which is labeled only by the coordinates that have not been integrated over.

For the purposes of this section, only the algebraic properties of integrals are of interest, not their function-theoretic significance. When we speak of the addition and multiplication of integrals, we refer to the manipulation of integrals as algebraic objects, using the ordinary rules of algebra. Thus

$$\frac{1}{6}\int dr_2\, dr_3\, dr_4\, g(r_1)g(r_2)g(r_3)g(r_4)F_4(r_1, r_2, r_3, r_4) \times \frac{1}{2}\int dr_2\, dr_3\, f(r_2)f(r_3)F_3(r_1, r_2, r_3)$$

$$= \frac{1}{12}\int dr_1\, dr_2\, dr_3\, dr_4\, dr_5\, dr_6\, g(r_1)g(r_2)g(r_3)g(r_4) \times f(r_5)f(r_6)F_4(r_1r_2r_3r_4)F_3(r_1r_5r_6).$$

This can be represented graphically as

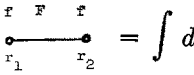


[10] H. L. Friedman and E. G. D. Cohen have also considered the use of such graphs. See for example *Ionic Solution Theory* by H. L. Friedman (John Wiley & Sons, Inc., New York, 1962).

[11] Another way of putting this is to say that any permutation of the variables of the fully labeled graph that corresponds to the *integrand* leaves that graph invariant. This is consistent with the fact that the integrand itself remains invariant, since F_4 is symmetric.

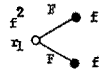
We shall use the concept of functional differentiation in the same spirit. The operation of functional or variational differentiation has a well known[12] definition in terms of a limit, but only the algebraic structure associated with this operation is of interest to us here.

We define addition, multiplication, and exponentiation of graphs in terms of the sum, product, and exponent of the integrals which the graphs represent. There is then the following difficulty with products: the product of two graphs is not necessarily a graph. That is, if Γ_1 is a graph corresponding to the integral I_1 divided by the symmetry numbers σ_1 and Γ_2 corresponds to I_2/σ_2, then $\Gamma_1\Gamma_2 = I_1I_2/\sigma_1\sigma_2$, but if we draw the graph associated with $I_3 = I_1I_2$, its symmetry number σ_3 will not in general be $\sigma_1\sigma_2$. For example, consider

$\Gamma_1 =$  $= \int dr_2\, f(r_1)f(r_2)F(r_1, r_2); \quad \sigma_1 = 1,$

$\Gamma_2 = \Gamma_1,$

$\Gamma_3 = \Gamma_1\Gamma_2 = \int dr_2\, dr_3\, f^2(r_1)f(r_2)f(r_3)F(r_1, r_2)F(r_1r_3).$

Here the σ_3 of  is not one but two.

For convenience, we define the *star product* of two simple graphs each containing one white circle labeled r_1. Using the notation of the last few sentences, we set

$$\Gamma_1 * \Gamma_2 = \Gamma_1\Gamma_2(\sigma_1\sigma_2/\sigma_3) = I_1I_2/\sigma_3.$$

The star product of more than two such graphs is defined in an obvious way. It is always a graph.

In our discussion, we shall use the following lemmas. Since these results, with minor differences, have been proved before in the literature[2] we shall not reproduce the proofs here.

Lemma I. Let $\mathcal{G}$ be a set of distinct connected graphs $\Gamma_1, \Gamma_2, \cdots, \Gamma_i, \cdots$, each consisting of some or no black vertices, one white 1-vertex labeled r, and some or no faces. It is assumed that no star products of members of $\mathcal{G}$ are in $\mathcal{G}$. Then if $\mathfrak{F}$ is the set consisting of all graphs in $\mathcal{G}$ and all star products of graphs in $\mathcal{G}$, the sum of all graphs in $\mathfrak{F} = \exp\{\text{the sum of all graphs in } \mathcal{G}\} - 1$. The exponentiation here is relative to the ordinary product, rather than the star product.

[12] See, for example, V. Volterra *Theory of Functionals* (Dover Publications, Inc., New York, 1959).

Lemma II. Let Γ be a graph consisting of faces and black vertices such that certain of the black vertices are f-vertices. Then

$$\frac{\delta\Gamma}{\delta f(r)} = \left[\begin{array}{l}\text{The sum of all distinct graphs that are obtained from } \Gamma \text{ by} \\ \text{changing a black } f\text{-vertex into a white 1-vertex labeled r}\end{array}\right].$$

We note that the statement of Lemma II could be used as the algebraic *definition* of functional differentiation. Then, for the purposes of this paper, the only lemma related to differentiation that we would have to prove is the statement that if two sums of graphs are equal, their functional derivatives are equal.

Lemma III. Given $\mathcal{G}$, a set of all distinct connected graphs, $\Gamma_1(r_1)$, $\Gamma_2(r_1)$ $\cdots$, each of which consists of one white vertex labeled r_1, some or no black vertices, and faces; let $f(r_1)$ be the function defined by the sum of the graphs in $\mathcal{G}$ and let G be any connected irreducible graph containing one white vertex labeled with any label, say r. Suppose that $\mathfrak{F}$ is the set of all distinct graphs $\Phi_1, \Phi_2, \cdots$, each of which is obtained by replacing each black vertex in G by some Γ_α. The replacement is to be made so that when Γ_α is attached to G, it is attached by its white vertex (which is first stripped of its label and blackened). Then

The sum of graphs in $\mathfrak{F}$ = the graph obtained from G by letting all the black vertices be f-vertices.

The following points are worth noting.

(i) It is unnecessary to specify anything about the functions that correspond to the vertices and faces of the Γ_α and G.

(ii) In forming a particular Φ_1 from G, the same Γ_α can be used in replacing several, or all, of the black vertices of G.

2. The Problem and a Sketch of its Solution

We are given a set of symmetric functions $U_n(r_1, \cdots, r_n)$, $n \geq 2$, defined in terms of the set $V_n(r_1, \cdots, r_n)$, $n \geq 2$, by the following relationship.[13]

$$U_n(r_1, \cdots, r_n) = \left[\begin{array}{l}\text{The sum of all distinct simply connected} \\ \text{simple graphs consisting of } n \text{ white 1-} \\ \text{vertices labeled } r_1 \text{ through } r_n, \text{ respectively, and } V_m\text{-faces, } 2 \leq m \leq n\end{array}\right]. \tag{2.1}$$

We prove that

[13] Ref. 1, p. 158.

$$V_n(r_1, \cdots, r_n) = \left[\begin{array}{l}\text{The weighted sum of all distinct simply-connected simple graphs consisting of } n \text{ white 1-vertices labeled } r_1 \text{ through } r_n, \text{ respectively, and } U_m\text{-faces, where } 2 \leq m \leq n. \text{ The weight of each graph is} \\ \prod_{1 \leq \alpha \leq m} (-1)^{(k_\alpha - 1)}(k_\alpha - 1)!\end{array}\right]. \tag{2.2}$$

Here the product is taken over all vertices and k_α is the articulation *degree* or *multiplicity* of the vertex, defined as the number of pieces into which the graph separates if the vertex is removed.

We give a sketch here of the way things go and relegate the details to an appendix. We define a functional $L[U_n, z]$ of the functions U_n and the test function $z(r)$. We next introduce both a new test function

$$\rho(r) = \delta L/\delta \ln z(r) \tag{2.3}$$

and a new functional

$$W = \ln [\delta L/\delta z(r)]. \tag{2.4}$$

We then compare the expression for W written as a functional of the V_n and ρ with the expression for W written as a functional of the U_n and ρ. We shall denote the first function as $\tilde{W}[V_n, \rho]$ and the second as $W[U_n, \rho]$. The result of the comparison is Eq. (2.2).

Although the idea is simple enough, getting the two different representations of W involves a number of steps. The first step in obtaining W as a functional of the V_n and ρ entails rewriting L as a functional of the V_n and z. We shall denote this functional as $\tilde{L}[V_n, z]$. The rewriting can be done immediately, using Eq. (2.1). Lemma II then yields both $\delta\tilde{L}[V_n, z]/\delta z(r)$ and $\rho(r)$, which is $\delta\tilde{L}[V_n, z]/\delta \ln z(r)$. From $\delta\tilde{L}[V_n, z]/\delta z(r)$ and Lemma I, we obtain $\ln \delta\tilde{L}[V_n, z]/\delta z(r)$. It is then easy to guess what form $\ln (\delta\tilde{L}/\delta z)$ has as a functional of V_n and ρ. We could confirm the guess by using our expression for $\rho(r) = \delta\tilde{L}[V_n, z]/\delta \ln z(r)$ in the assumed expression for $\tilde{W}[V_n, \rho]$ and verifying that it gives $\ln\{\delta\tilde{L}[V_n, z]/\delta z(r)\}$. However, more formally and more directly, we simply apply Lemma III to $\delta\tilde{L}[V_n, z]/\delta z(r)$ and $\delta\tilde{L}[V_n, z]/\delta \ln z(r)$ and get $\tilde{W}[V_n, \rho]$, without guessing anything in the first place.

It is still necessary to find $\ln (\delta L/\delta z)$ as the functional $W[U_n, \rho]$. To find it, we start with $L[U_n, z]$ and apply Lemma II to get $\delta L[U_n, z]/\delta z(r)$, which we write as $1 + A[U_n, z]$. By noting that $\ln [1 + A] = A - \frac{1}{2}A^2 + \cdots$ we obtain an expansion of $\ln \{\delta L[U_n, z]/\delta z(r)\}$ and by noting that $1/(1 + A) = 1 - A + A^2 - \cdots$ we obtain an expansion of $\{\delta L[U_n, z]/\delta z(r)\}^{-1}$. From this latter expression and Lemma III we obtain both $[\delta L/\delta z(r)]^{-1}$ and $z(r)$ {which is just $\rho(r)\ [\delta L/\delta z(r)]^{-1}$} expressed as functionals of the U_n and ρ. We use Lemma III backwards here, noting that the U_n-face, ρ-vertex expressions for $[\delta L/\delta z(r)]^{-1}$ and $z(r)$ imply each other and that Lemma III applied to both of them yields our initial expression for $\{\delta L[U_n, z]/\delta z(r)\}^{-1}$. Finally we obtain $\ln (\delta L/\delta z)$ written as $W[U_n, \rho]$ by substituting our expression for $z(r)$ in terms of the U_n and $\rho(r)$ into our expression for $\ln \{\delta L[U_n, z]/\delta z(r)\}$. We use Lemma III to verify that the resulting expression for $W[U_n, \rho]$ contains the correct combinatorial factors. The whole development is shown schematically in Fig. 1. As indicated there, the final step of obtaining the V_n in terms of the U_n by comparing $\tilde{W}[V_n, \rho]$ and $W[U_n, \rho]$ is facilitated by

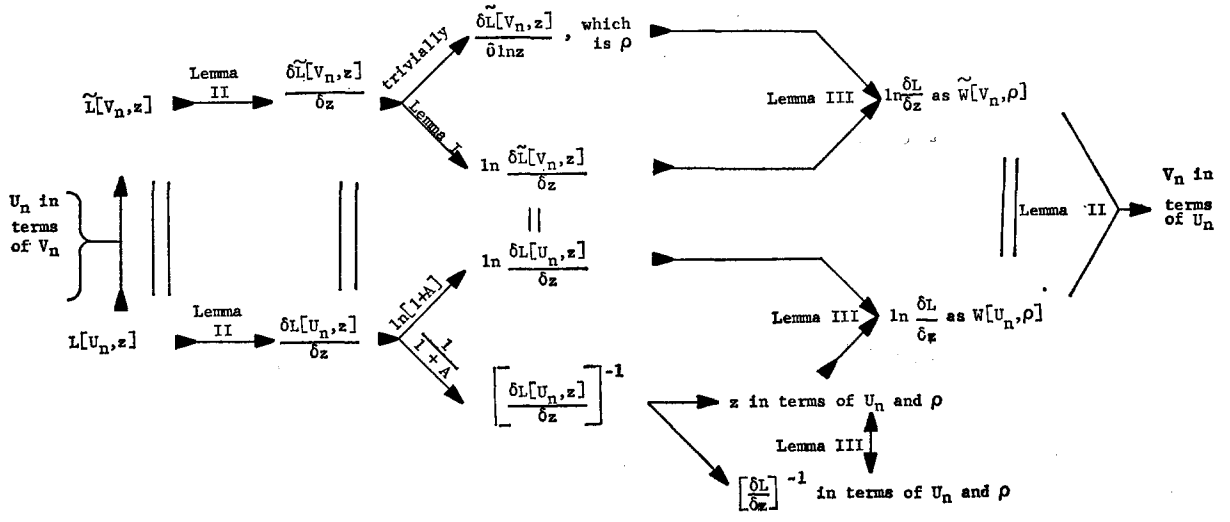


FIG. 1. "Flow diagram" for Sec. II. Double bars are drawn between functionals that are equal, such as $\tilde{L}[V_n, z]$ and $L[U_n, z]$.

the use of Lemma II, since V_n is simply the $(n-1)$th functional derivative of $\tilde{W}$ with respect to ρ, evaluated at $\rho = 0$.

No combinatorial factors beyond those that already appear as part of the symmetry numbers arise in obtaining the graphical representation of $\tilde{W}[V_n, \rho]$ [Eq. (A6) in Appendix]. The counting problem that must be solved to get the representation of $W[U_n, \rho]$ [Eq. (A7)] turns out to be trivial and yields the extra factor of $\prod_\alpha (-1)^{k_\alpha - 1}(k_\alpha - 1)!$ that appears in the final inversion rule. Details are given in the Appendix.

We note that the functional L can be identified with the logarithm of the grand partition function of a classical system, $\rho(r)$ with the one-particle distribution function, and $z(r)$ with the fugacity times exp {−external field/kT} where k is Boltzmann's constant and T is the absolute temperature. The expression $\tilde{W}[V_n, \rho]$ can then be identified with the {external field/kT} minus $1/kT$ times the work necessary to bring a particle from outside the system to position r in terms of $\rho(r)$ and the star functions, V_n. If the external field is zero, a single integration with respect to the number density ρ yields the number-density expansion of the equation of state. If further, the only interparticle forces are pair forces, then the star functions can easily be reexpressed in terms of Mayer's f-function[7] and the resulting expansion reduces to the familiar ρ-vertex, f-bond cluster expansion of the equation of state.

ACKNOWLEDGMENT

The author is indebted to Professor G. E. Uhlenbeck for bringing this problem to his attention and for helpful discussions. He also wishes to thank Professor S. Sherman for information concerning his own independent method of solving the inversion problem and for a critical appraisal of an early version of this manuscript.

APPENDIX. THE DETAILS OF THE SOLUTION

The functional $L[U_n, z]$ is defined as

$$L[U_n, z] = \text{the sum of all distinct graphs each of which consists of black } z\text{-vertices and a } U_n\text{-face, } n \geq 1, \text{ where } U_1 = 1,$$

$$= \int dr_1 z(r_1) U_1(r_1)$$

$$+ \frac{1}{2} dr_1\, dr_2\, z(r_1) z(r_2) U_2(r_1, r_2)$$

$$+ \frac{1}{6} \int \prod_{i=1}^{3} [dr_i\, z(r_i)] U_3(r_1, r_2, r_3) + \cdots, \quad \text{(A1)}$$

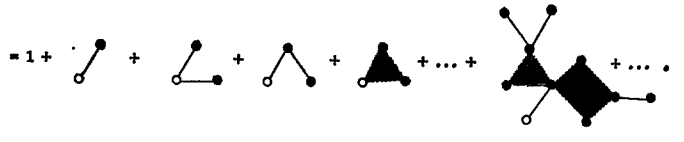
In illustrating the graphs in (A1) and in the expansions to follow we omit all labels designating the functions to which the vertices and faces correspond as well as the coordinates associated with the white vertices. These functions and coordinates will be obvious from the descriptions immediately preceding the illustrations.

We shall first find $\tilde{W}[V_n, \rho]$. We begin by inserting into the right-hand side of (A1) the expression given in (2.1) for the U_n in terms of the V_m, $m < n$. The result is the following expression for $\tilde{L}[V_n, z]$.

$$\tilde{L}[V_n, z] = \left[\begin{array}{l}\text{The sum of all distinct simply connected simple}\\ \text{graphs consisting of black } z\text{-vertices and } V_n\text{-faces}\end{array}\right] \quad \text{(A2)}$$

If we differentiate (A2) with respect to $z(r)$ to get $\delta\tilde{L}/\delta z(r)$, we obtain, from Lemma II,

$$\delta\Gamma/\delta z(r) = \left[\begin{array}{l}1 + \text{the sum of all distinct simply connected}\\ \text{simple graphs consisting of a white 1-vertex}\\ \text{labeled } r, \text{ at least one black } z\text{-vertex, and } V_n\text{-faces}\end{array}\right] \quad \text{(A3)}$$



Two expressions immediately follow from (A3). First, by multiplying both sides of (A3) by $z(r)$ and using the definition $\rho(r) = z(r)\delta\tilde{L}/\delta z(r)$ to identify the left-hand side of the result, we have

$$\rho(r) = \left[\begin{array}{l}z(r) + \text{the sum of all distinct simply connected simple}\\ \text{graphs consisting of a white } z\text{-vertex labeled } r, \text{ at least}\\ \text{one black } z\text{-vertex, and } V_n\text{-faces}\end{array}\right]. \quad \text{(A4)}$$

Second, applying Lemma I to Eq. (A4) we obtain an expression for $\ln \{\delta\Gamma[V_n, z]/\delta z(r)\}$:

$$\ln \frac{\delta\Gamma[V_n, z]}{\delta z(r)} = \left[\begin{array}{l}\text{The sum of all distinct simply-connected}\\ \text{simple graphs consisting of a white 1-vertex}\\ \text{labeled } r, \text{ at least one black } z\text{-vertex, and}\\ V_n\text{-faces, such that the white vertex is not}\\ \text{an articulation point}\end{array}\right]. \tag{A5}$$

We now apply Lemma III. We let the $f(r)$ of that lemma be $\rho(r)$ given by Eq. (A4) and the G be the maximal rooted irreducible subgraph of any graph in the right-hand side of (A5). By maximal rooted irreducible subgraph we mean the largest subgraph of the graph that contains the white vertex (the root) but does not contain any articulation point. An arbitrary graph need not contain a unique maximal irreducible subgraph, but the condition in (A5), stating that the white vertex is not an articulation point is enough to ensure the existence of such a subgraph for every graph in (A5).

We observe that any graph on the right-hand side of (A5) can be obtained from its maximal rooted irreducible subgraph G by the replacement of the black vertices of G by graphs in the expansion of ρ. (The replacement is to be made in the manner indicated in Lemma III.) Furthermore, only graphs that can be obtained by such a replacement are found in Eq. (A5). As a result, the right-hand side of Eq. (A5) can be replaced by a sum of all distinct maximal rooted irreducible subgraphs if all z-vertices are replaced by ρ-vertices. This result can be written as

$$\tilde{W}[V_n, \rho] = \left[\begin{array}{l}\text{The sum of all distinct graphs that consist of a}\\ \text{white 1-vertex labeled } r, n \text{ black } \rho\text{-circles, and a}\\ V_n\text{-face, } n \geq 2\end{array}\right]. \tag{A6}$$

Thus we have obtained $\ln [\delta\tilde{L}/\delta z(r)]$ as $\tilde{W}[V_n, \rho]$, a functional of V_n and ρ.

We turn next to the task of finding $W[U_n, \rho]$. Our claim is that

$$W[U_n, \rho] = \left[\begin{array}{l}\text{The weighted sum of all simply connected}\\ \text{simple graphs that consist of one white 1-vertex}\\ \text{labeled } r, \text{ one or more black } \rho\text{-vertices, and}\\ U_n\text{-faces, } n \geq 2. \text{ The weight of each graph is}\\ \displaystyle\prod_{1 \leq \alpha \leq n} (-1)^{(k_\alpha - 1)}(k_\alpha - 1)!\\ \text{where this factor has the same meaning as in}\\ \text{Eq. (2.2)}\end{array}\right]. \tag{A7}$$

The weight of each graph can be absorbed into the definition of the graph if the αth vertex is associated not with $\rho(r_\alpha)$ but with $w_\alpha\rho(r_\alpha)$, where

$$w_\alpha = (-1)^{(k_\alpha - 1)}(k_\alpha - 1)!. \tag{A8}$$

The label α here is not to be confused with the labeling used in designating the vector associated with each white vertex. Here α is used simply to identify each circle, so that an articulation degree can be associated with it. For uniformity, the α's of a graph will be enumerated always starting with the white vertex, so that the w_α for it will always be denoted by w_1. Otherwise, no particular ordering will be assumed. Absorbing the weight of each circle into the definition of the function associated with the circle, we rewrite (A7) in the form in which we shall actually obtain it:

$$W[U_n, \rho] = \left[\begin{array}{l}\text{the sum of all distinct simply-connected simple}\\ \text{graphs that consist of one white } w_1\text{-vertex}\\ \text{labeled } r, \text{ one or more black } (w_\alpha\rho)\text{-vertices,}\\ \text{and } U_n\text{-faces, } n \geq 2\end{array}\right]. \tag{A9}$$

Our first step in proving (A7) is to apply Lemma II to the graphs of Eq. (A1). We obtain for $\delta L/\delta z(r)$ the characterization

$$\frac{\delta L}{\delta z(r)} = \left[\begin{array}{l}1 + \text{the sum of all distinct graphs each of which}\\ \text{consists of one white 1-vertex labeled } r, (n-1)\\ \text{black } z\text{-vertices, and a } U_n\text{-face, } n \geq 2\end{array}\right]. \tag{A10}$$

Our next step is to obtain $\ln \{\delta L[U_n, z]/\delta z\}$. We do this by letting the sum in Eq. (A10) be denoted by A so that we can write

$$\delta L/\delta z(r) = 1 + A. \tag{A11}$$

Therefore

$$\ln \frac{\delta L}{\delta z} = A - \frac{A^2}{2} + \frac{A^3}{3} - \cdots (-1)^{k-1}\frac{A^k}{k} \pm \cdots. \tag{A12}$$

We claim that the application of this identity yields the prescription

$$\ln \frac{\delta L}{\delta z} = \left[\begin{array}{l}\text{the sum of all distinct simply connected simple}\\ \text{graphs that consist of one white } w_1\text{-vertex labeled } r,\\ \text{one or more black } z\text{-vertices, and } U_n\text{-faces, } n \geq 2,\\ \text{such that the only articulation point is the white}\\ \text{vertex}\end{array}\right]. \tag{A13}$$

In order to verify this, it is necessary only to consider the genesis of the factor w_1 in (A13) since the rest of the characterization of the sum follows immediately from (A12) and the characterization of A given by (A11). Thus we can take (A13) as given if we consider w_1 to be still undetermined, and our task is to evaluate w_1.

Comparing (A12) and (A13), we see that for a given k, $(-1)^{k-1}A^k/k$ is a sum of graphs, each of whose white vertex is a w_1-vertex which has k as an articulation degree. Let us examine one such graph, which we denote by Γ_t. The k pieces that

would separate if the white vertex were removed define k subgraphs that share the same white vertex. Among them, there will, in general, be n_1 identical graphs having a symmetry number of σ_1, n_2 identical graphs having a symmetry number of σ_2, $\cdots$, and n_r identical graphs having a symmetry number of σ_r. Thus we can write $k = \sum_{1\leq\alpha\leq r} n_\alpha$, and according to our general rule of correspondence between graphs and integrals, our typical graph Γ_t stands for an integral times (the symmetry number of $\Gamma_t)^{-1}$ where the integral has an integrand consisting of w_1 times a product of $z(r_i)$'s and $U_n(r_1, \cdots, r_n)$'s. If the integral over the product of z's and U_n's is denoted as J we have $I = w_1 J$. We further note that the symmetry number of Γ_t can be written in terms of the σ_α and n_α, so

$$\Gamma_t = w_1 J / \prod_\alpha [\sigma_\alpha^{n_\alpha}(n_\alpha!)]. \tag{A14}$$

On the other hand, A is itself a sum of graphs, each of which represents an integral I_β times σ_β^{-1}, where σ_β is the symmetry number of I_β. Hence Γ_t is equal to $(-1)^{k-1}/k$ times a sum of identical terms, every one of which is a product of graphs. Each such product is a term of the form $\prod_\beta [I_\beta/\sigma_\beta]$. This term can be re-expressed as $J/\prod_\beta \sigma_\beta$ or, changing to an index that sums over distinct graphs, as $J/\prod_\alpha (\sigma_\alpha^{n_\alpha})$. The question is: How many such identical terms are there? The number of such terms is the number of ways in which k objects can be partitioned into r groups of which the first contains n_1 objects, the second n_2 objects, etc., and hence this number is simply the multinomial coefficient

$$\frac{k!}{\prod_\alpha (n_\alpha!)}. \tag{A15}$$

Thus we have

$$\Gamma_t = \frac{(-1)^{k-1}k!J_t}{k \prod_\alpha (n_\alpha!) \prod_\alpha (\sigma_\alpha^{n_\alpha})} = \frac{w_1 J_t}{\prod_\alpha [\sigma_\alpha^{n_\alpha}(n_\alpha!)]} \tag{A16}$$

or

$$w_1 = (-1)^{k-1}k!/k = (-1)^{k-1}(k-1)!. \tag{A17}$$

We have verified Eq. (A13). Our next objective is to find an expression for z in terms of U_n and ρ. This can be accomplished in two steps. The first involves following exactly the same argument we have just used to obtain (A13) from (A10). Instead of looking at

$$\ln \frac{\delta L}{\delta z} = \sum_{k\geq 1} \frac{(-1)^{k-1}A^k}{k},$$

however, we look at

$$\left[\frac{\delta L}{\delta z}\right]^{-1} = \sum_{k\geq 0} (-1)^k A^k.$$

We find that

$$\left[\frac{\delta L}{\delta z}\right]^{-1} = \left[\begin{array}{l}1 + \text{the sum of all distinct simply connected simple graphs that consist of one white } [(-1)^{k_1}k_1!]\text{-vertex labeled } r, \text{ one or more black } z\text{-vertices and } U_n\text{-faces, } n \geq 2, \\ \text{such that the only articulation point is the white vertex}\end{array}\right]. \tag{A18}$$

Since $[\delta L/\delta z]^{-1}$ equals $z(r)/\rho(r)$, multiplying both sides of (3.18) by $\rho(r)$ gives us an expression for $z(r)$:

$$z(r) = \left[\begin{array}{l}\rho(r) + \text{the sum of all distinct simply connected simple graphs that consist of one white } ([(-1)^{k_1}k_1!]\rho)\text{-vertex labeled } r, \text{ one or more black } z\text{-vertices, and } U_n\text{-faces, } n \geq 2, \text{ such that the only articulation point is a white vertex}\end{array}\right]. \tag{A19}$$

We can now replace the z's that appear in the sum of (A18) by the right-hand side of (A19). The resulting expression still contains z's, since the right-hand side of (A19) does. They can be replaced by the right-hand side of (A19) again, etc. Taking the limit of this procedure, we can eliminate the z's that occur in (A18) in favor of ρ's. The result is

$$\left[\frac{\delta L}{\delta z}\right]^{-1} = \left[\begin{array}{l}1 + \text{the sum of all distinct simply connected graphs that consist of one white } [(-1)^{k_1}k_1!]\text{-vertex labeled } r, \text{ one or more black } w_\alpha\rho\text{-vertices, and } U_n\text{-faces, } n \geq 2.\end{array}\right]. \tag{A20}$$

To verify that (A20) is indeed the limiting result of our procedure of successive replacement, we can use Lemma III, taking as $f(r)$ the $z(r)$ obtained by multiplying (A20) through by $\rho(r)$. We identify G with the maximal rooted irreducible subgraph of any graph in the right-hand side of (A20) and the Φ_1, Φ_2, etc. with typical graphs in (A20) having the same maximal rooted irreducible subgraph G. The lemma establishes that (A20) reduces to (A18) when only the graphs free of black articulation points are retained and all black vertices are made z-vertices.

For convenience, we exhibit Eq. (A20) after it is multiplied through by ρ, since the result is the desired expression for z in terms of ρ and the U_n:

$$z(r) = \left[\begin{array}{l}\rho(r) + \text{the sum of all distinct simply connected simple graphs that consist of one white } [(-1)^{k_1}k_1!\rho]\text{-vertex labeled } r, \text{ one or more black } (w_\alpha\rho)\text{-vertices, and } U_n\text{-faces, } n \geq 2.\end{array}\right]. \tag{A21}$$

We can now use (A21) to eliminate z from the expression for $\ln(\rho/z)$ given by (A13). The result is (A9).

Decomposition of the Lorentz Transformation Matrix into Skew-Symmetric Tensors*

STANISLAW L. BAZANSKI†
Department of Physics, Lehigh University, Bethlehem, Pennsylvania
(Received 19 October 1964)

Any matrix describing a finite proper orthochronous Lorentz transformation of the null tetrad in Minkowski space may be written as a polynomial of the second order in skew-symmetric tensors. From a geometrical point of view these tensors describe two-dimensional planes which are mapped by the Lorentz transformation into themselves.

1. INTRODUCTION

THE infinitesimal proper orthochronous Lorentz transformation, is described as is well known, by the matrix[1]

$$L_\alpha{}^\beta = \delta_\alpha{}^\beta + \omega_\alpha{}^\beta, \tag{1.1}$$

where $\omega_{\alpha\beta} = \eta_{\beta\gamma}\omega_\alpha{}^\gamma$ are infinitesmial constants, skew-symmetric in their indices ($\omega_{\alpha\beta} = -\omega_{\beta\alpha}$). The six arbitrary coefficients $\omega_{\alpha\beta}$ are related to the six parameters of the infinitesimal Lorentz group. The very elementary fact expressed by (1.1) suggests that the matrices of the finite proper orthochronous group itself can be written in the form of a series[2]

$$L_\alpha{}^\beta = \delta_\alpha{}^\beta + \omega_\alpha{}^\beta + \sum_{i=1}^{\infty} \alpha_{(i)}\omega_\alpha{}^{\lambda_1}\omega_{\lambda_1}{}^{\lambda_2}\cdots\omega_{\lambda_i}{}^\beta, \tag{1.2}$$

where $\alpha_{(i)}$ are some numerical coefficients. Equation (1.1) might then be considered as the linear approximation of (1.2). Let us observe that if (1.2) is true, there will always be a simplification possible. As higher powers of any matrix in a four-dimensional space can always be reduced to powers not exceeding the fourth, due to its Hamilton–Cayley equation, the series (1.2) should really be a polynomial in $\omega_{\alpha\beta}$ of order not higher than fourth. This note aims to show that any finite proper orthochronous Lorentz transformation may indeed be represented in a polynomial form, and the geometrical properties of the skew-symmetric tensors involved are exhibited.

2. LORENTZ TRANSFORMATION IN TERMS OF NULL TETRADS

In this section the null-tetrad formalism will be briefly reviewed.[3]

Let us assume that in Minkowski space an orthonormal tetrad of vectors ${}_\mu e^\alpha$ is given ($\alpha\beta\gamma\cdots\lambda$-tensor indices; $\mu\nu\pi\rho\cdots$-tetrad indices). Then we have

$${}_\mu e^\alpha{}_\nu e_\alpha = \eta_{\mu\nu}, \qquad \eta^{\mu\nu}{}_\mu e_{\alpha\nu} e_\beta = \eta_{\alpha\beta}. \tag{2.1}$$

Following the standard procedure we introduce the basic vectors of a null tetrad

$$\begin{aligned} k_\alpha &= (1/\sqrt{2})({}_0e_\alpha + {}_1e_\alpha), \\ l_\alpha &= (1/\sqrt{2})({}_0e_\alpha - {}_1e_\alpha), \\ t_\alpha &= (1/\sqrt{2})({}_2e_\alpha - i\,{}_3e_\alpha). \end{aligned} \tag{2.2}$$

The vectors k_α and l_α are real, and t_α is complex; we denote its complex conjugate by $\bar{t}_\alpha$. The first of relations (2.1) then becomes

$$\begin{aligned} k_\alpha l^\alpha &= -t^\alpha \bar{t}_\alpha = 1, \\ k_\alpha k^\alpha &= l_\alpha l^\alpha = t_\alpha l^\alpha = t_\alpha k^\alpha = t_\alpha t^\alpha = 0, \end{aligned} \tag{2.3}$$

and the second

$$\eta_{\alpha\beta} = 2[k_{(\alpha}l_{\beta)} - t_{(\alpha}\bar{t}_{\beta)}]. \tag{2.4}$$

We now consider linear transformations of null-tetrad basic vectors which leave the form (2.4) of $\eta_{\alpha\beta}$ invariant [or equivalently, which preserve all the scalar products (2.3), and their complex conjugates], and which do not change the direction of one of the real null vectors, say of k_α. Besides that we demand that our transformations do not reverse the sense of time, and do not lead to space reflections.

[The Lorentz transformation is defined as an iso-

* Research supported by the Aerospace Research Laboratories of the Office of Aerospace Research, U. S. Air Force.
† On leave of absence from Institute of Physics, Polish Academy of Sciences, Warsaw, Poland.
[1] We consider Minkowski space with coordinates x^α and a metric tensor $\eta_{\alpha\beta}$ (with signature -2). All tensor indices are raised or lowered by means of this tensor. $\delta_\alpha{}^\beta$ is the Kronecker tensor. For any two vectors a_α, b_β, we define $a_{(\alpha}b_{\beta)} = \frac{1}{2}(a_\alpha b_\beta + a_\beta b_\alpha)$; $a_{[\alpha}b_{\beta]} = \frac{1}{2}(a_\alpha b_\beta - a_\beta b_\alpha)$.
[2] From the theory of continuous groups it is well known that this series is an exponential in the generator $\omega_{\alpha\beta}$. This information is, however, irrelevant for the considerations presented here.
[3] Null tetrads have recently been widely used in general relativity. A more detailed description of this formalism may be found, e. g., in R. K. Sachs, Proc. Roy. Soc. (London) 264, 309 (1961) and in the literature quoted there.

metry in Minkowski space ($\eta'_{\alpha\beta} = \eta_{\alpha\beta}$). Thus the demanded condition of the invariance of the direction of k_α could be regarded as an additional restriction. It is, however, well known that every *proper orthochronous Lorentz transformation leaves at least one null-direction invariant.* It may also be shown directly as follows:

Every future-pointing null vector can be written in the form: $k^0 = \tau$, $k^1 = \tau \sin\varphi \sin\vartheta$, $k^2 = \tau \cos\varphi \sin\vartheta$, $k^3 = \tau \cos\vartheta$; where $0 < \tau < \infty$, $0 \leq \varphi < 2\pi$, $0 \leq \vartheta \leq \pi$. Hence, the null directions are determined by the angles φ, ϑ on the unit sphere. The proper orthochronous Lorentz transformation maps the future light cone onto itself, and thus generates a continuous mapping of this sphere onto itself (which does not change the orientation of closed curves). From the well-known Brouwer's fixed-point theorem there is always at least one point (φ, ϑ) that remains fixed under such a mapping, and the null direction which corresponds to this point is left invariant by the Lorentz transformation.]

It may be easily shown that these conditions define the null-tetrad up to the transformations

$$\begin{aligned} k'_\alpha &= e^{\psi} k_\alpha, \\ l'_\alpha &= e^{-\psi}(l_\alpha + \gamma\bar{\gamma}k_\alpha - \gamma t_\alpha - \bar{\gamma}\bar{t}_\alpha), \\ t'_\alpha &= e^{i\varphi}(t_\alpha - \bar{\gamma}k_\alpha). \end{aligned} \tag{2.5}$$

Here the parameters ψ and φ are real numbers, and γ is complex. For $\gamma = 0$ (2.5) describes a special Lorentz transformation in the plane defined by k_α and l_α (i.e., in the plane ${}_0e^\alpha$, ${}_1e^\alpha$) and a rotation in the plane defined by t_α and $\bar{t}_\alpha$ (i.e., in the plane ${}_2e^\alpha$, ${}_3e^\alpha$):

$$\begin{aligned} k'_\alpha &= e^{\psi} k_\alpha, \\ l'_\alpha &= e^{-\psi} l_\alpha, \\ t'_\alpha &= e^{i\varphi} t_\alpha. \end{aligned} \tag{2.6}$$

For $\psi = 0$, $\varphi = 0$, however, transformation (2.5) goes over into a null rotation[4] about the vector

$$\begin{aligned} k''_\alpha &= k_\alpha, \\ l''_\alpha &= l_\alpha + \gamma\bar{\gamma}k_\alpha - \gamma t_\alpha - \bar{\gamma}\bar{t}_\alpha, \\ t''_\alpha &= t_\alpha - \bar{\gamma}k_\alpha. \end{aligned} \tag{2.7}$$

The general transformation (2.5) is a product of (2.6) and (2.7); this product is not commutative.

The matrices representing (2.6) and (2.7) may be very easily expressed as dyadic products of the vectors k_α, l_α, t_α. In the case of (2.6) we look for a matrix $L_\alpha{}^\beta$ which would allow us to write (2.6) in the form $k'_\alpha = L_\alpha{}^\beta k_\beta$, etc. Obviously

$$\begin{aligned} L_\alpha{}^\beta &= k'_\alpha l^\beta + l'_\alpha k^\beta - t'_\alpha \bar{l}^\beta - \bar{l}'_\alpha t^\beta \\ &= e^{\psi} k_\alpha l^\beta + e^{\psi} l_\alpha k^\beta - e^{i\varphi} t_\alpha \bar{l}^\beta - e^{-i\varphi} \bar{l}_\alpha t^\beta. \end{aligned} \tag{2.8}$$

Similarly the transformation (2.7) is described by a matrix

$$\begin{aligned} N_\alpha{}^\beta &= k''_\alpha l^\beta + l''_\alpha k^\beta - t''_\alpha \bar{l}^\beta - \bar{l}''_\alpha t^\beta \\ &= \delta_\alpha{}^\beta + \gamma(k_\alpha t^\beta - t_\alpha k^\beta) \\ &\quad + \bar{\gamma}(k_\alpha \bar{l}^\beta - \bar{l}_\alpha k^\beta) + \gamma\bar{\gamma} k_\alpha k_\beta. \end{aligned} \tag{2.9}$$

Formulas (2.8) and (2.9) may be considered as a kind of canonical form of the Lorentz matrices in Minkowski space, analogous to the spectral decomposition of a matrix in Euclidean space.

3. DECOMPOSITION OF LORENTZ TRANSFORMATION INTO SKEW-SYMMETRIC TENSORS

The matrix $N_\alpha{}^\beta$ (2.9) for the null transformation may be written as

$$\begin{aligned} N_\alpha{}^\beta &= \delta_\alpha{}^\beta + F_\alpha{}^\beta + \bar{F}_\alpha{}^\beta + F_{\alpha\lambda}\bar{F}^{\lambda\beta} \\ &= (\eta_{\alpha\lambda} + F_{\alpha\lambda})(\eta^{\lambda\beta} + \bar{F}^{\lambda\beta}), \end{aligned} \tag{3.1}$$

where

$$F_{\alpha\beta} = 2\gamma k_{[\alpha} t_{\beta]} \tag{3.2}$$

is a null antisymmetric tensor (i.e., $F_{\alpha\beta}F^{\alpha\beta} = 0$). It is immaterial in definition (3.2) which of the null-tetrads (2.7) (k_α, l_α, t_α or k''_α, l''_α, t''_α) was used. If γ is an infinitesmial quantity, (3.1) goes over into (1.1) with $\omega_{\alpha\beta} = F_{\alpha\beta} + \bar{F}_{\alpha\beta}$. In terms of $\omega_{\alpha\beta}$ (3.1) may also be written as

$$N_\alpha{}^\beta = \delta_\alpha{}^\beta + \omega_\alpha{}^\beta + \omega_\alpha{}^\lambda \omega_\lambda{}^\beta. \tag{3.3}$$

Thus Eqs. (3.1) or (3.3) are just expressions of the form (2.2) for null rotations.

A similar decomposition may be easily obtained for the matrix $L_\alpha{}^\beta$ (2.8). Let us observe that the vectors k_α and l_α in (2.8) may be replaced, without affecting $L_\alpha{}^\beta$, by any other pair of transformed null vectors $k'_\alpha = e^{\psi} k_\alpha$, $l'_\alpha = e^{-\psi} l_\alpha$. All these null vectors span the same bivector $f_{\alpha\beta} = 2k_{[\alpha} l_{\beta]}$. Thus we may try to represent $L_{\alpha\beta}$ in terms of $f_{\alpha\beta}$ and its dual $f^*_{\alpha\beta} = 2it_{[\alpha}\bar{l}_{\beta]}$.[5]

[4] Null rotations have been discussed by T. Shibata, J. Sci. Hiroshima Univ. **19**, 101 (1951) (and in his earlier papers quoted there), and were used in general relativity by H. Bondi, F. A. E. Pirani, and I. Robinson, Proc. Roy. Soc. (London) **A251**, 519 (1959); see also R. K. Sachs, Ref. 2, and H. M. Schwartz, Am. J. Phys. **31**, 864 (1963).

[5] In a null tetrad the Levi-Civita tensor density $\epsilon_{\alpha\beta\gamma\delta}$ equals $4!ik_{[\alpha}l_\beta\bar{l}_\gamma t_{\delta]}$. It is a real tensor, hence $f^*_{\alpha\beta}$ is also real.

Since our bivectors fullfil the relations[6]

$$f_\alpha{}^\gamma f_\gamma{}^\delta f_\delta{}^\beta = f_\alpha{}^\beta, \qquad f^*_\alpha{}^\gamma f^*_\gamma{}^\delta f^*_\delta{}^\beta = -f^*_\alpha{}^\beta, \tag{3.4}$$

all odd powers in $f_{\alpha\beta}$ and $f^*_{\alpha\beta}$ in an expansion like (1.2) can be reduced to $f_{\alpha\beta}$ and its dual $f^*_{\alpha\beta}$, and all even powers to $f_\alpha{}^\gamma f_\gamma{}^\beta$ and $f^*_\alpha{}^\gamma f^*_\gamma{}^\beta$. Thus without any loss of generality we may replace (1.2) by

$$L_\alpha{}^\beta = \delta_\alpha{}^\beta + \alpha_1 f_\alpha{}^\beta + \alpha_2 f_\alpha{}^\gamma f_\gamma{}^\beta + \beta_1 f^*_\alpha{}^\beta + \beta_2 f^*_\alpha{}^\gamma f^*_\gamma{}^\beta. \tag{3.5}$$

The vectors k_α and l_α are eigenvectors of $f_{\alpha\beta}$ belonging respectively to eigenvalues ± 1, and t_α and $\bar{t}_\alpha$ are eigenvectors of $f^*_{\alpha\beta}$ with corresponding eigenvalues $\mp i$. Thus

$$\begin{aligned} L_\alpha{}^\beta k_\beta &= e^{\psi} k_\alpha = (1 + \alpha_1 + \alpha_2) k_\alpha, \\ L_\alpha{}^\beta l_\beta &= e^{-\psi} L_\alpha = (1 - \alpha_1 + \alpha_2) L_\alpha, \\ L_\alpha{}^\beta t_\beta &= e^{i\varphi} t_\alpha = (1 - i\beta_1 - \beta_2) t_\alpha, \\ L_\alpha{}^\beta \bar{t}_\beta &= e^{-i\varphi} \bar{t}_\alpha = (1 + i\beta_1 - \beta_2) \bar{t}_\alpha. \end{aligned}$$

Using (2.3) we obtain

$$\begin{aligned} \alpha_1 &= \sinh\psi, \qquad \alpha_2 = \cosh\psi - 1, \\ \beta_1 &= -\sin\varphi, \qquad \beta_2 = 1 - \cos\varphi, \end{aligned} \tag{3.6}$$

and

$$L_\alpha{}^\beta = \delta_\alpha{}^\beta + \sinh\psi\, f_\alpha{}^\beta + (\cosh\psi - 1) f_\alpha{}^\gamma f_\gamma{}^\beta - \sin\varphi\, f^*_\alpha{}^\beta + (1 - \cos\varphi) f^*_\alpha{}^\gamma f^*_\gamma{}^\beta. \tag{3.7}$$

This formula[7] represents a decomposition of the Lorentz matrix $L_\alpha{}^\beta$ in terms of a skew-symmetric tensor $f_\alpha{}^\beta$.

4. CLASSIFICATION OF LORENTZ MATRICES

Formulas (3.3) and (3.7) allow us to write the Lorentz matrix $L_{\alpha\beta}$, if we know the "angles" φ and ψ, and the planes which are mapped by the transformation into themselves. (These planes are here described by simple bivectors of any pair of orthonormal vectors in them.)

Now we ask the opposite question: How can we find these invariant characteristics, if we are only given the components of a finite Lorentz matrix $L_{\alpha\beta}$?

Let us therefore assume that a general, proper orthochronous Lorentz matrix $L_{\alpha\beta}$ is given. Its skew-symmetric part $\omega_{\alpha\beta} = \frac{1}{2}L_{[\alpha\beta]}$ has two invariants $I_1 = \omega_{\alpha\beta}\omega^{\alpha\beta}$ and $I_2 = \omega_{\alpha\beta}\omega^{*\alpha\beta}$. If both of them differ from zero, our Lorentz transformation is of the type (3.7), i.e., it is a special Lorentz transformation and a rotation. We may see this easily if we remember that any skew-symmetric tensor $\omega_{\alpha\beta}$ (for which $I_1 \neq 0$ and $I_2 \neq 0$) may be represented as a duality rotation of two mutually dual simple bivectors.[8] To show it, it is sufficient to take

$$f_{\alpha\beta} = \omega_{\alpha\beta}\cos\alpha + \omega^*_{\alpha\beta}\sin\alpha$$

and define α as $\tan 2\alpha = I_2/I_1$. Then $f_{\alpha\beta}$ is simple ($f_{\alpha\beta}f^{*\alpha\beta} = 0$), and $\omega_{\alpha\beta}$ is a linear combination of $f_{\alpha\beta}$ and $f^*{}_{\alpha\beta}$:

$$\omega_{\alpha\beta} = f_{\alpha\beta}\cos\alpha - f^*_{\alpha\beta}\sin\alpha, \tag{4.1}$$

and the coefficients $\sin\alpha$ and $\cos\alpha$ in this decomposition are defined in terms of I_1 and I_2. If we normalize $f_{\alpha\beta}$ so that $f_{\alpha\beta}f^{\alpha\beta} = -2$ and compare (4.1) and the skew-symmetric part of (3.7), we can find the parameters ψ and φ of our Lorentz transformation in terms of I_1 and I_2. The simple bivectors $f_{\alpha\beta}$ and $f^*_{\alpha\beta}$ define then, respectively, two 2-dimensional, completely orthogonal planes in which the special Lorentz transformation characterized by ψ and the rotation by an angle φ are performed.

In the case $I_2 = 0$ and $I_1 \neq 0$, $\omega_{\alpha\beta}$ is already simple. Then we have to distinguish two cases: $I_1 > 0$ and $I_1 < 0$. In the first case $L_{\alpha\beta}$ represents a pure rotation and in the second a special Lorentz transformation.

If both $I_2 = I_1 = 0$, the transformation $L_{\alpha\beta}$ is a Lorentz null rotation.

From this classification we see that any proper orthochronous Lorentz transformation, generated by askew-symmetric tensor $\omega_{\alpha\beta}$, always can be written in either the form (3.7) or (3.3), i.e., it is a transformation (2.5). Thus it again demonstrates that any proper orthochronous Lorentz transformation leaves at least one null direction invariant. It is not difficult to see that this direction is one of the null eigendirections of the generator $\omega_{\alpha\beta}$.

ACKNOWLEDGMENTS

The author is indebted to Professor Peter Havas for his kind hospitality at the Department of Physics, Lehigh University, and to Professor Joshua Goldberg for helpful discussions.

The topic of this note arose during stimulating discussions with Professor Jerzy Plebański.

[6] These relations may be shown either directly, using the explicit form of $f_{\alpha\beta}$ and $f^*{}_{\alpha\beta}$, or as consequences of the Hamilton–Cayley equation for antisymmetric matrices $f_{\alpha\beta}$ and $f^*{}_{\alpha\beta}$; see, e.g., J. Plebański, Bull Acad. Polon. Sci. Cl. III, 9, 587 (1961).

[7] Equation (3.7) may also be written as

$$L_\alpha{}^\beta = \sinh\psi\, f_\alpha{}^\beta + \cosh\psi\, f_\alpha{}^\gamma f_\gamma{}^\beta - \sin\varphi\, f^*{}_{\alpha\beta} - \cos\varphi\, f_\alpha{}^{*\gamma} f_\gamma{}^\beta.$$

[8] A bivector (i.e., skew-symmetric tensor of the second rank) is called simple if it is the skew-symmetrized part of an external product of two vectors, i.e., if $\omega_{\alpha\beta} = p_{[\alpha}q_{\beta]}$. It is easy to show that $\omega_{\alpha\beta}$ is simple if and only if $\omega_{\alpha\beta}\omega^{*\alpha\beta} = 0$.

Atomic Terms for Equivalent Electrons

N. KARAYIANIS
Harry Diamond Laboratories, Washington, D. C.
(Received 18 August 1964; final manuscript received 1 October 1964)

The problem of determining the number of times, $b(l\ N\ L\ S)$, a given term occurs in an atomic configuration of the form $(nl)^N$ is considered by the methods of group theory. The b are shown to be related to the classical problem of partitioning numbers. Several expressions relating b to other, simpler partitions are obtained, which result in a recursion relationship for the b. A table for $b(4\ N\ L\ S)$ is included to complement existing tables for the range $1 \leq l \leq 3$, which are found in several places in the literature.

I. INTRODUCTION

THE restriction of atomic wavefunctions to those that are antisymmetric on the interchange of any two electrons limits the possible Russell–Saunders (LS) states for a given atomic configuration. The limitation is especially pronounced for configurations of the form $(nl)^N$, i.e., for equivalent electrons, because the full burden of antisymmetrization must be absorbed by the angular momentum variables—the only way in which the single-electron states may differ in these configurations. For two equivalent electrons, the allowed terms are those for which $L + S =$ even, and each such term is found to occur just once. For an arbitrary number of equivalent electrons, the manner in which antisymmetrization manifests itself cannot, however, be so simply expressed. One finds that the number of times, $b(l\ N\ L\ S)$ [hereafter written $b(lNLS)$ for conciseness], a given LS term appears increases rapidly as l (the arbital angular momentum of each electron) and N (the number of electrons) increase, and moreover, for fixed l, N, and S, the number of terms as L varies, fluctuates irregularly. The $b(lNLS)$, on the other hand, are not dependent on the principal quantum number n, provided n is greater than l.

The problem of determining the numbers $b(lNLS)$ dates back to Breit[1] in 1926. The following year, Russell[2] and Gibbs, Wilber, and White[3] simultaneously published results extending Breit's algorithm to obtain the b for all possible configurations involving p, d, and f electrons. The method developed by Breit relies on a manipulation of the angular momentum projections (the so-called "magnetic quantum numbers") of single electron states and is very laborious in spite of the improvements by the other authors. A good summary of the method is given by White.[4]

Recently, Curl and Kilpatrick[5] treated the same problem by the more elegant methods of group theory. They obtained a generating function $J_S(x)$ that generates the desired quantities in the form

$$J_S(x) = \sum_L b(lNLS)(x^L - x^{-L-1}). \qquad (1)$$

The contribution of Curl and Kilpatrick represents a significant improvement over Breit's method for obtaining the b even though, in practice, the process for finding a given b by expanding $J_S(x)$ in powers of x may still involve much labor.

A consideration of the known methods for determining the b indicates that these quantities are related to the classical problem of partitioning numbers, i.e., the problem of determining the number of unique ways an integer can be expressed as a summation over a set of other integers restricted by certain conditions. The method of Breit, particularly, gives a strong indication of this relationship. One may, therefore, reasonably expect the b to exhibit properties generally shared by the various partitions that have been studied.[6] Among these properties are (1) each partition is associated with a relatively simple generating function, (2) each satisfies a recursion relationship, and (3) there exist interconnecting formulas relating the various partitions. Only one of these properties, the first, has been found previously in connection with the b's. Owing to the facility with which tables for a given partition can be obtained using the recursion rela-

[1] G. Breit, Phys. Rev. **28**, 334 (1926).
[2] H. N. Russell, Phys. Rev. **29**, 782 (1927).
[3] R. C. Gibbs, D. T. Wilber, and H. E. White, Phys. Rev. **29**, 790 (1927).
[4] H. E. White, *Introduction to Atomic Spectra* (McGraw-Hill Book Company, Inc., New York, 1934), 1st ed., pp. 235, 293, and 437.
[5] R. F. Curl, Jr., and J. E. Kilpatrick, Am. J. Phys. **28**, 357 (1960).
[6] H. Gupta, *Royal Society Mathematical Tables, Partitions* 4 (Cambridge University Press, Cambridge, England, 1958).

TABLE I. Values for $b(4NLS)$ in the range $2 \leq N \leq 5$, $0 \leq L$, and $0 \leq S$. All b for arguments in this range not given in the table are zero.

L	$N=2$ $S=1$	$S=0$	$N=3$ $S=\frac{3}{2}$	$S=\frac{1}{2}$	$N=4$ $S=2$	$S=1$	$S=0$	$N=5$ $S=5/2$	$S=3/2$	$S=1/2$
0	0	1	0	0	1	0	3	1	1	3
1	1	0	1	1	0	4	0	0	4	5
2	0	1	0	2	2	3	5	2	6	10
3	1	0	2	2	1	6	2	1	8	11
4	0	1	1	3	2	5	6	2	9	15
5	1	0	1	3	1	7	3	1	9	14
6	0	1	1	2	2	5	6	2	9	15
7	1	0	1	2	1	6	3	1	8	13
8	0	1	0	2	1	4	5	1	7	13
9			1	1	0	4	2	0	6	10
10			0	1	1	2	3	1	4	9
11			0	1	0	2	1	0	3	6
12					0	1	2	0	2	5
13					0	1	0	0	1	3
14					0	0	1	0	1	2
15								0	0	1
16								0	0	1

tionship it satisfies, it is desirable to find such a relationship for b.

This paper reports the results of a successful investigation towards this end. In addition to obtaining a recursion relationship for the b, several new formulas are obtained relating this quantity to other, simpler partitions. Finally, the generating function of Curl and Kilpatrick and a new generating function related to it are derived for completeness.

The recursion relationship for b, given by (22) in Sec. III of this paper, lends itself well to computer calculations. A table for $b(lNLS)$ for all values of N, L, and S, and for l in the range $1 \leq l \leq 6$ was obtained in this manner and is available upon request.[7] The values for $l = 4$ are included in this paper in Tables I, II, and III to complement existing tables found in the literature.[3-5]

II. A REDUCIBLE REPRESENTATION

Consider the finite-dimensional linear vector space Γ comprising the antisymmetric wavefunctions of N equivalent electrons, each with orbital angular momentum l. This space is invariant with respect to the three-dimensional rotation operators in coordinate space, $R_L = \exp(-i\mathbf{n}\cdot\mathbf{L}\theta)$, and in spin space, $R_S = \exp(-i\mathbf{n}'\cdot\mathbf{S}\varphi)$, where $\mathbf{L} = \sum_1^N \mathbf{l}_i$ and $\mathbf{S} = \sum_1^N \mathbf{s}_i$. As a result of this invariance, Γ is a carrier of a representation for the product group formed

TABLE II. Values for $b(4NLS)$ in the range $6 \leq N \leq 7$, $0 \leq L$, and $0 \leq S$. All b for arguments in this range not given in the table are zero.

L	$N=6$ $S=3$	$S=2$	$S=1$	$S=0$	$N=7$ $S=7/2$	$S=5/2$	$S=3/2$	$S=1/2$
0	0	2	1	7	0	1	2	5
1	1	3	13	3	1	3	13	15
2	0	8	13	13	0	5	16	24
3	2	7	23	11	1	6	25	31
4	1	10	21	19	0	7	25	37
5	1	9	28	13	1	7	30	39
6	1	10	23	21	0	7	27	40
7	1	8	26	14	1	6	28	39
8	0	8	20	17	0	5	23	36
9	1	5	20	12	0	4	22	32
10	0	5	14	13	0	3	16	27
11	0	3	13	7	0	2	14	22
12	0	2	8	9	0	1	9	17
13	0	1	7	4	0	1	7	13
14	0	1	3	4	0	0	4	9
15	0	0	3	2	0	0	3	6
16	0	0	1	2	0	0	1	4
17	0	0	1	0	0	0	1	2
18	0	0	0	1	0	0	0	1
19					0	0	0	1

[7] N. Karayianis and Arthur Hausner, HDL Report No. R-RCB-64-1, June 1964.

TABLE III. Values for $b(4NLS)$ in the range $8 \leq N \leq 9$, $0 \leq L$, and $0 \leq S$. All b for arguments in this range not given in the table are zero.

L	$N=8$ $S=4$	$S=3$	$S=2$	$S=1$	$S=0$	$N=9$ $S=9/2$	$S=7/2$	$S=5/2$	$S=3/2$	$S=1/2$
0	0	0	5	3	10	1	0	3	6	8
1	0	2	7	23	7	0	1	3	16	19
2	0	2	16	27	24	0	1	9	24	35
3	0	4	16	43	20	0	1	8	34	40
4	1	3	22	43	34	0	1	12	38	52
5	0	4	20	53	28	0	1	10	40	54
6	0	3	23	48	37	0	1	12	42	56
7	0	3	19	52	28	0	1	9	39	53
8	0	2	19	43	34	0	1	9	35	53
9	0	2	14	42	24	0	0	6	32	44
10	0	1	13	32	26	0	0	6	26	40
11	0	1	8	29	17	0	0	3	20	32
12	0	0	7	20	18	0	0	3	16	26
13	0	0	4	17	10	0	0	1	11	19
14	0	0	3	10	10	0	0	1	7	15
15	0	0	1	8	5	0	0	0	5	9
16	0	0	1	4	5	0	0	0	3	7
17	0	0	0	3	2	0	0	0	1	4
18	0	0	0	1	2	0	0	0	1	2
19	0	0	0	1	0	0	0	0	0	1
20	0	0	0	0	1	0	0	0	0	1

by the operators $R_L R_S$. The reduction of this representation into its irreducible components provides the number of states, $b(lNLS)$, which belong simultaneously to the $(2L+1)$-dimensional irreducible representation of the rotation group in coordinate space and the $(2S+1)$-dimensional irreducible representation of the rotation group in spin space.

Let the matrix representation for $R_L R_S$ in terms of some basis in Γ be given by $D(R_L R_S)$. Further, let U represent the unitary transformation that effects the reduction of D as follows,

$$UD(R_L R_S)U^{-1} = \sum_{L'S'} b(lNL'S') D^{2L'+1}(R_L) \times D^{2S'+1}(R_S), \quad (2)$$

where the D matrices on the right-hand side are irreducible representations of the respective three-dimensional rotation groups. Taking traces of both sides, one has

$$\mathrm{Tr}\,[D(R_L R_S)] = \sum_{L'S'} b(lNL'S') \chi^{2L'+1}(R_L) \chi^{2S'+1}(R_S), \quad (3)$$

where the χ are characters of the respective rotations in the given irreducible representations.

To evaluate (3), it is sufficient to choose the rotations

$$R_L = e^{-iL_z\theta}, \qquad R_S = e^{-iS_z\varphi}, \quad (4)$$

for which the characters are given by the well-known expressions

$$\chi^{2L'+1}(R_L) = \sin(L'+\tfrac{1}{2})\theta / \sin\tfrac{1}{2}\theta,$$
$$\chi^{2S'+1}(R_S) = \sin(S'+\tfrac{1}{2})\varphi / \sin\tfrac{1}{2}\varphi. \quad (5)$$

The trace of $D(R_L R_S)$ may be evaluated by choosing any convenient orthonormal basis in Γ. We choose the set of determinantal wavefunctions u_ν that are known to span Γ and to be diagonal in $\mathbf{L}_z$ and $\mathbf{S}_z$. The u_ν are antisymmetrized product wavefunctions of the N single-electron states. A typical matrix element for $R_L R_S$ in this basis is

$$u_\nu^\dagger R_L R_S u_\lambda = D_{\nu\lambda}(e^{-iL_z\theta} e^{-iS_z\varphi}) = \delta_{\nu\lambda} e^{-iM_L(\nu)\theta} e^{-iM_S(\nu)\varphi}, \quad (6)$$

where $M_L(\nu) = \sum_1^N m_i$ and $M_S(\nu) = \sum_1^N \mu_i$, and the m_i and μ_i are respectively the individual z projections of the orbital angular momentum and spin for the single-electron states. The trace of $D(R_L R_S)$ is then given by

$$\mathrm{Tr}\,[D(R_L R_S)] = \sum_{M_L M_S} A(M_L M_S) e^{-iM_L\theta} e^{-iM_S\varphi}, \quad (7)$$

where $A(M_L M_S)$ represents the number of determinantal wavefunctions for a given M_L and M_S. Equating (3) and (7), and substituting the character values given by (5), one obtains

$$\sum_{L'S'} b(lNL'S') \sin(L'+\tfrac{1}{2})\theta \sin(S'+\tfrac{1}{2})\varphi = \sum_{M_L M_S} A(M_L M_S) \sin\tfrac{1}{2}\theta \sin\tfrac{1}{2}\varphi e^{-iM_L\theta} e^{-iM_S\varphi}. \quad (8)$$

The summation on the left is restricted to $L' \geq 0$, $S' \geq 0$, whereas M_L and M_S assume positive and

negative values. Owing to the fact A is even in both M_L and M_S, operating on (8) with

$$(\pi)^{-2}\int_0^{2\pi} d\theta \int_0^{2\pi} d\varphi \sin(L+\tfrac{1}{2})\theta \sin(S+\tfrac{1}{2})\varphi \quad (9)$$

gives the result

$$b(lNLS) = \sum_{\lambda\sigma} (-)^{\lambda+\sigma}\binom{1}{\lambda}\binom{1}{\sigma} A(L+\lambda, S+\sigma). \quad (10)$$

Considering the original definition of $A(M_L M_S)$ given by (7), this function is readily seen to be a partition. In particular, it represents the number of unique ways in which N integers m_i, restricted by $-l \leq m_i \leq l$, can add to give M_L *and* the number of ways N related quantities $\mu_i = \pm\frac{1}{2}$ can add to give M_S with the added restriction that the simultaneous equalities $m_i = m_j$, $\mu_i = \mu_j$ are disallowed for all $i \neq j$. To relate A to a simpler partition from which a recursion relationship can be obtained, we separate the sets $\{m_i, \mu_i\}$ according to the number of pairs of m_i that are equal (three or more m_i may not be equal).

A. $m_1 > \cdots > m_N$

The number of states for each given set of m_i satisfying these conditions is the number of unique ways the μ_i can be distributed consistent with $\sum \mu_i = M_S$, which is

$$\binom{N}{\frac{1}{2}N + M_S}. \quad (11)$$

To obtain the total number of states in this category, (11) must be multiplied by the number of unique sets of m_i which satisfy the conditions above, and whose sum is M_L. This number is represented by

$$\sum_{m_1>\cdots>m_N} \delta(\textstyle\sum m_i, M_L). \quad (12)$$

B. $m_1 = m_2 \neq \{m_3 > \cdots > m_N\}$

For a given set of m_i satisfying these conditions, one spin-up and one spin-down are necessarily associated with m_1 and m_2, hence the number of states for each given set of m_i is

$$\binom{N-2}{\frac{1}{2}N - 1 + M_S}. \quad (13)$$

The total number of states of this kind is therefore the product of (13) with

$$\sum_{m_1=m_2\neq\{m_3>\cdots>m_N\}} \delta(\textstyle\sum m_i, M_L). \quad (14)$$

C. $\{m_1 = m_2 > m_3 = m_4 > \cdots > m_{2p-1} = m_{2p}\}$
$\neq \{m_{2p+1} > \cdots > m_N\}$

For this general case, the total number of states is

$$\binom{N-2p}{\frac{1}{2}N - p + M_S} \sum_{\{p\}} \delta(\textstyle\sum m_i, M_L), \quad (15)$$

where $\{p\}$ represents the lengthy restriction on the m_i given above by C.

The partition A is equal to the summation over all the possibilities for p which is

$$A(M_L M_S) = \sum_{p=0}^{\frac{1}{2}N-|M_S|} \binom{N-2p}{\frac{1}{2}N - p + M_S} F(lNM_L p), \quad (16)$$

where F is the partition defined for arbitrary L by,

$$F(lNLp) = \sum_{\{p\}} \delta(\textstyle\sum m_i, L). \quad (17)$$

This method for expressing A in terms of a new partition F is by no means unique. For example, a given set $\{m_i, \mu_i\}$ can be split into two sets of m_i, those belonging to $\mu_i = +\frac{1}{2}$ and those belonging to $\mu_i = -\frac{1}{2}$. The physical requirements on the m_i now are that no two m_i in the same set are equal, and no reference is made to the equality or inequality of m_i in different sets. A consideration of A from this point of view leads to a new expression for A that can be obtained directly, however this expression will be obtained indirectly but more profitably later in the paper.

III. THE RECURSION RELATIONSHIP

The partition $F(lNLp)$ defined by (17) may be defined alternatively as the number of ways one can choose $N - p$ integers m_i' with the properties that $m_1' > \cdots > m_{N-p}'$, $-l \leq m_i' \leq l$, and that p of these integers are doubled when performing the summation to give L. When viewed in this manner, F readily can be related to itself through a recursion formula. Consider the sets $\{m_i'\}$ that satisfy the conditions for belonging to $F(lNLp)$. These sets can be separated into four categories determined by the two independent possibilities: m_1' equals or does not equal l, and m_{N-p}' equals or does not equal $-l$. By this explicit consideration of the m_i' that can attain the extremum values $\pm l$, the number of sets of m_i' in each category can be related to F of the form $F(l-1\ N'\ L'\ p')$. The results are

$$m_1' \neq l,\ m_{N-p}' \neq -l\text{:}\quad F(l-1\ N\ L\ p), \quad (18)$$

$$m_1' = l,\ m_{N-p}' \neq -l\text{:}\quad F(l-1\ N-1\ L-l\ p) + F(l-1\ N-2\ L-2l\ p-1), \quad (19)$$

$$m_1' \neq l,\ m_{N-p}' = -l:\ F(l-1\ N-1\ L+l\ p) + F(l-1\ N-2\ L+2l\ p-1), \quad (20)$$

$$m_1' = l,\ m_{N-p}' = -l:\ F(l-1\ N-2\ L\ p) + F(l-1\ N-3\ L-l\ p-1) + F(l-1\ N-3\ L+l\ p-1) + F(l-1\ N-4\ L\ p-2). \quad (21)$$

The quantity $F(lNLp)$ is therefore equal to the summation of the nine terms in (18) through (21), since there are no other possibilities for m_1' and m_{N-p}'.

The recursion relationship for b is obtained by relating F to b through (10) and (16), and using the above recursion formula for F. The result can be expressed in the following symmetric form:

$$b(lNLS) = \sum_{\lambda_1 \cdots \lambda_4} \prod_i \binom{1}{\lambda_i} b[l-1,\ N-2-\lambda_1-\lambda_2+\lambda_3+\lambda_4,\ L+l(\lambda_1-\lambda_2+\lambda_3-\lambda_4),\ S+\tfrac{1}{2}(-\lambda_1+\lambda_2+\lambda_3-\lambda_4)]. \quad (22)$$

The physically significant values for b are generated by this recursion formula starting with the following boundary conditions for $l = 0$,

$$b(0000) = 1,\ b(010\tfrac{1}{2}) = 1, \quad (23)$$

and the symmetries

$$b(l\ N\ -L\ S) = -b(l\ N\ L-1\ S), \quad (24)$$

$$b(l\ N\ L\ -S) = -b(l\ N\ L\ S-1), \quad (25)$$

$$b(l\ N\ L\ S) = b(l\ 2(2l+1)-N\ L\ S), \quad (26)$$

which can be obtained from (10). It can be further shown that the symmetries will be propagated to all b by the recursion relationship once incorporated into the boundary conditions. Consequently, it is convenient to combine the set of conditions (23) through (26) into the following compact form:

$$b[0\ \alpha+\beta\ -\lambda\ \tfrac{1}{2}(\alpha-\beta)] = (-)^{\lambda}\binom{1}{\lambda} \sum_{\sigma} (-)^{\sigma}\binom{1}{\sigma}\binom{1}{\alpha+\sigma}\binom{1}{\beta-\sigma}. \quad (27)$$

Expressions (22) and (27) are sufficient to generate all the physically significant b and to ensure the symmetry conditions.

By repeated application of the recursion relationship (22), one can obtain a general solution for $b(lNLS)$ in terms of the boundary values $b(0N'L'S')$. If, then, the solution is specialized to the $b(0N'L'S')$ given by (27), a new, symmetric expression for b will result which can be recognized as the partition described at the end of Sec. II.

To simplify the calculation, it is convenient first to define two functions, $B_\lambda^k(x)$ and $g_\lambda^k(\mu)$ as follows:

$$\prod_{r=0}^{k} (1 + yx^r) = \sum_{\lambda} B_\lambda^k(x) y^\lambda \quad (28)$$

$$= \sum_{\lambda\mu} g_\lambda^k(\mu) y^\lambda x^\mu. \quad (29)$$

The difference equations satisfied by these functions are, respectively,

$$B_\lambda^k(x) = B_\lambda^{k-1}(x) + x^k B_{\lambda-1}^{k-1}(x) \quad (30)$$

and

$$g_\lambda^k(\mu) = g_\lambda^{k-1}(\mu) + g_{\lambda-1}^{k-1}(\mu), \quad (31)$$

which have the following particular solutions satisfying (28) and (29):

$$B_\lambda^k(x) = x^{\frac{1}{2}\lambda(\lambda-1)} \prod_{r=1}^{\lambda} (1 - x^{k+2-r})(1 - x^r)^{-1} \quad (32)$$

and

$$g_\lambda^k(\mu) = \sum_{n_1=0}^{k} \sum_{n_2=0}^{n_1-1} \cdots \sum_{n_\lambda=0}^{n_{\lambda-1}-1} \delta(\sum n_i,\ \mu) \quad (33)$$

with

$$B_0^k(x) = 1 \quad (34)$$

and

$$g_0^k(\mu) = \delta_{\mu 0}. \quad (35)$$

The function $g_\lambda^k(\mu)$ has, in addition, the following properties that are useful for this problem, and which are easily derived from (29):

$$g_\lambda^k(\mu) = g_\lambda^k(k\lambda - \mu), \quad (36)$$

$$g_\lambda^k(\mu) = g_{k+1-\lambda}^k(\tfrac{1}{2}k(k+1) - \mu). \quad (37)$$

The iterated application of the recursion relationship (22) results in the expression

$$b(lNLS) = \sum_{\lambda_i,\mu_i} \prod_i g(i) b[0,\ N-2-\lambda_1-\lambda_2+\lambda_3+\lambda_4,\ L+\mu_1-\mu_2+\mu_3-\mu_4,\ S+\tfrac{1}{2}(-\lambda_1+\lambda_2+\lambda_3-\lambda_4)], \quad (38)$$

where the temporary designation $g(i)$ has been given to the expression

$$g(i) = \sum_{\lambda_{i1} \cdots \lambda_{il}} \delta(\sum_j \lambda_{ij},\ \lambda_i)\, \delta(\sum_j (l+1-j)\lambda_{ij},\ \mu_i) \times \prod_{j=1}^{l} \binom{1}{\lambda_{ij}}. \quad (39)$$

Expressing the Kronecker deltas in (39) by appropriate contour integrals, one obtains

$$g(i) = (2\pi i)^{-2} \oint dy \oint dx\, y^{-\lambda_i-1} x^{-\mu_i-1} \times \prod_{\nu=0}^{l-1} (1 + yxx^{\nu}), \quad (40)$$

which, from (29), is

$$g(i) = g_{\lambda_i}^{l-1}(\mu_i - \lambda_i). \quad (41)$$

The boundary conditions given by (27) can be enforced in the general solution (38) by appropriate Kronecker deltas to give the following special solution:

$$\begin{aligned} b(lNLS) = {} & \sum_{\alpha\beta\lambda\sigma} (-)^{\lambda+\sigma} \binom{1}{\alpha}\binom{1}{\beta}\binom{1}{\lambda}\binom{1}{\sigma} \\ & \times \sum_{\lambda_i,\mu_i} \prod_i g_{\lambda_i}^{l-1}(\mu_i - \lambda_i) \\ & \times \delta(\tfrac{1}{2}N+S+\sigma-l-\alpha-\lambda_1+\lambda_3, 0) \\ & \times \delta(\tfrac{1}{2}N-S-\sigma-l-\beta-\lambda_2+\lambda_4, 0) \\ & \times \delta(L+\lambda+\mu_1-\mu_2+\mu_3-\mu_4, 0). \end{aligned} \quad (42)$$

Owing to the restriction on the arguments of the g, the Kronecker deltas together with these functions can be expressed as three contour integrals in the manner of (40). One obtains the following result:

$$\begin{aligned} b(lNLS) = {} & \sum_{\lambda\sigma} (-)^{\lambda+\sigma} \binom{1}{\lambda}\binom{1}{\sigma} (2\pi i)^{-3} \\ & \times \oint dx \oint dy \oint dz\, x^{L+\lambda-1} \\ & \times y^{\frac{1}{2}N+S+\sigma-l-2} z^{\frac{1}{2}N-S-\sigma-l-2} (1 + y)(1 + z) \\ & \times \prod_{\nu=0}^{l-1} (1 + y^{-1}x^{\nu+1})(1 + yx^{\nu+1}) \\ & \times (1 + z^{-1}x^{-\nu-1})(1 + zx^{-\nu-1}), \end{aligned} \quad (43)$$

where the summations over α and β have introduced the factors $y^{-1}(1 + y)$ $z^{-1}(1 + z)$. These are just the factors required to combine the pi products involving y and z separately to obtain

$$(yz)^{-l} \prod_{\nu=0}^{2l} (1 + yx^{-l}x^{\nu})(1 + zx^{l}x^{-\nu}). \quad (44)$$

The contours are then evaluated using (29), and the arguments rearranged according to (36) and (37) to obtain finally,

$$\begin{aligned} b(lNLS) = {} & \sum_{\lambda\sigma} (-)^{\lambda+\sigma} \binom{1}{\lambda}\binom{1}{\sigma} \\ & \times \sum_{\mu} g_{\frac{1}{2}N+S+\sigma}^{2l}(\mu) g_{\frac{1}{2}N-S-\sigma}^{2l}(Nl+L+\lambda-\mu). \end{aligned} \quad (45)$$

Considering the definition of g given by (33), and the alternative method for expressing A according to the discussion at the end of Sec. II, it is clear that the summation over μ in (45) is related to A by

$$A(M_L M_S) = \sum_{\mu} g_{\frac{1}{2}N+M_S}^{2l}(\mu) g_{\frac{1}{2}N-M_S}^{2l}(Nl+M_L-\mu). \quad (46)$$

The somewhat lengthy method employed in the derivation of (45) clarifies the method of obtaining a generating function for the b. By operating on (45) with $\sum_L x^L$ and letting $L = \mu' - Nl - \lambda + \mu$ on the right-hand side, then according to (28) and (29), one has

$$\begin{aligned} \sum_L b(lNLS)x^L = {} & \sum_{\lambda\sigma} (-)^{\lambda+\sigma} \binom{1}{\lambda}\binom{1}{\sigma} \\ & \times x^{-Nl-\lambda} B_{\frac{1}{2}N+S+\sigma}^{2l}(x) B_{\frac{1}{2}N-S-\sigma}^{2l}(x). \end{aligned} \quad (47)$$

The right-hand side of (47) is the generating function of Curl and Kilpatrick, where the explicit values for the B are given by (32).

A more compact generating function is obtained by introducing two new variables by operating on (47) with $\sum_{NS} y^{\frac{1}{2}N+S} z^{\frac{1}{2}N-S}$. Then we obtain

$$\sum_{L\alpha\beta} b[l\ \alpha+\beta\ L\ \tfrac{1}{2}(\alpha-\beta)] x^L y^{\alpha} z^{\beta} = (1 - x^{-1})(1 - zy^{-1}) \times \prod_{\nu=-l}^{l} (1 + yx^{\nu})(1 + zx^{\nu}), \quad (48)$$

which also can be deduced directly from (43).

With this result, the aims of this paper have been fulfilled. The recursion relationship for the b given by (22) has been most useful for obtaining the values of b in the range $1 \leq l \leq 6$, and these are available on request.[7] In Tables I, II, and III of this paper the values for $l = 4$ are included to supplement existing tables for $1 \leq l \leq 3$ that appear in several places in the literature.

ACKNOWLEDGMENTS

The author wishes to acknowledge with many thanks the assistance of C. A. Morrison in obtaining and refining several of the results of this paper. He also thanks Arthur Hausner for programming the problem for the HDL computer and for his general interest.

Trace Formalism for Quantum Mechanical Expectation Values

JOHN LANGERHOLC
University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 2 November 1964)

The trace of a positive bounded operator in a Hilbert space is defined and shown to be independent of the basis used in the definition. The trace of the product WA is defined for bounded A and positive bounded W with unit trace and is also shown to be basis-invariant. An integral representation is derived and used to define $\mathrm{Tr}(WA)$ for unbounded A. Linearity, isotony, and continuity with respect to uniform covergence are proved for $\mathrm{Tr}(WA)$ as a function of (bounded) A. Several formulas used in statistical mechanics are derived, and it is proved that if A and B are commuting positive operators, then $\mathrm{Tr}(W(A + B)) = \mathrm{Tr}(WA) + \mathrm{Tr}(WB)$. A counter example is given which shows that the commonly used definition of $\mathrm{Tr}(WA)$ in terms of an orthonormal basis is not invariant under permutation of the basis vectors even in the case of a very simple unbounded operator A.

An expectation-value function M which associates the expectation value $M(A)$ to the operator A is assumed given and subject to certain restrictions. For bounded operators, these include positivity, normalization, additivity, and a condition which may be considered as a requirement of regularity for the probability function induced by M. Modifications of these requirements are imposed for unbounded operators, and von Neumann's statistical formula is proved: there is a unique bounded positive operator W with unit trace for which $M(A) = \mathrm{Tr}(WA)$. The requirements placed on M are weaker than those of von Neumann, and are in fact satisfied by $\mathrm{Tr}(WA)$ as a function of A.

1. INTRODUCTION

IN the quantum theory, the operators in infinite-dimensional spaces whose traces are of interest are usually of the form WA where W is a positive bounded operator of unit trace and A is a (possibly unbounded) self-adjoint operator corresponding to a dynamical variable. Because of this and the extremely pathological character of Tr (A) as compared to Tr (WA) as a function of A, the concept of Tr (A) is treated rather briefly and attention is concentrated on Tr (WA). The treatment is valid for unbounded self-adjoint operators; and expressions such as $AW\psi$, which restrict the region of validity of the usual treatments to bounded operators, do not appear except when W is a projector commuting with A. In this case $AW\psi = WA\psi$ has meaning if $\psi \in \mathrm{dom}\ (A)$. It is extremely important to be able to discuss unbounded operators, since many of the commonly used operators (position, momentum, and energy operators) are unbounded, and statistical mechanics would be impossible if the Hamiltonian of a system were bounded unless the Hilbert space were finite dimensional. With the exception of a collection of spacially fixed particles with spins, however, most interesting physical systems are associated with infinite-dimensional spaces.

Although the definition of Tr (WA) for bounded A is a natural extension of the definition in finite-dimensional spaces, it is not capable of basis invariant extension to the case where A is unbounded as the counterexample in Appendix I shows. This leaves the integral formula (7) in Sec. 3 as possibly the most natural extension of the definition.

Following von Neumann[1], Chap. IV, an expectation value function M is assumed defined which associates to each operator the expectation value of its corresponding observable. M is assumed defined for all bounded self-adjoint operators and all unbounded self-adjoint operators for which either $M(A_+)$ or $M(A_-)$ is finite. Here the decomposition $A = A_+ - A_-$ where $A_+ = \int_0^\infty \lambda\, dE_\lambda^A$ and $A_- = \int_{-\infty}^0 -\lambda\, dE_\lambda^A$ are positive self-adjoint operators is used. Several restrictions are placed on the function M in order that it be suitable to represent the expectation values of physical quantities. One complete set of such restrictions is weaker than von Neumann's set of axioms, and uses two noncommuting operators simultaneously only in the additivity requirement $M(A + B) = M(A) + M(B)$ for bounded self-adjoint operators A and B. In the rest of the axioms, all operators simultaneously considered commute and are therefore simultaneously measurable. This increases the plausibility of these axioms, since to such operators and their corresponding measurements, classical logic and the classical theory of probability are applicable. From these restrictions, the statistical formula is proved: there is a unique positive bounded operator W with unit trace for which $M(A) = \mathrm{Tr}\ (WA)$.

2. THE TRACE OF AN OPERATOR

If W is a positive operator in a separable Hilbert space $\mathfrak{H}$ and $\Psi = \{\psi_n\}_{n=1}^\infty$ is an orthonormal basis of $\mathfrak{H}$ contained in dom (W), then we define $\mathrm{Tr}_\Psi(W)$

[1] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1936).

to be $\sum_{n=1}^{\infty} (\psi_n, W\psi_n)$. This is a sum of positive real numbers which therefore either converges or diverges to $+\infty$. It is clearly linear as a function of W: If $W \leq V$, then $V - W \geq 0$ and $\mathrm{Tr}_\Psi(V) = \mathrm{Tr}_\Psi(W) + \mathrm{Tr}_\Psi(V - W) \geq \mathrm{Tr}_\Psi(W)$—the trace is isotonic as a function of W for a given basis Ψ. Of special importance are operators with finite trace. If $\{E_\lambda^W\}_{\lambda \in R}$ is the spectral resolution of W: $W = \int_0^\infty \lambda\, dE_\lambda^W$, and if we set $F_\lambda^W = I - E_\lambda^W$, $\mathfrak{N}_\lambda = F_\lambda^W \mathfrak{H}$, then $\mathfrak{N}_\lambda$ is finite dimensional for every $\lambda > 0$ provided $\mathrm{Tr}_\Psi(W) < +\infty$ for some basis Ψ. To prove this, we recall that E_λ^W commutes with W and that the product of two commuting positive operators is positive; hence $WE_\lambda^W \geq 0$. From this it follows that

$$W \geq W - WE_\lambda^W = F_\lambda^W \int_0^\infty \mu\, dE_\mu^W = \int_\lambda^\infty \mu\, dE_\mu^W \geq \lambda \int_\lambda^\infty dE_\mu^W = \lambda F_\lambda^W$$

and consequently,

$$\mathrm{Tr}_\Psi(W) \geq \mathrm{Tr}_\Psi(\lambda F_\lambda^W) = \lambda\, \mathrm{Tr}_\Psi(F_\lambda^W).$$

$\mathrm{Tr}_\Psi(F_\lambda^W)$ is thus finite for every $\lambda > 0$. It remains to be shown that if $\mathfrak{N}_\lambda$ has dimension D_λ and $\mathrm{Tr}_\Psi(F_\lambda^W)$ is finite, then $D_\lambda = \mathrm{Tr}_\Psi(F_\lambda^W)$. To do this, we take an orthonormal basis $\Phi = \{\phi_m\}_{m=1}^{D_\lambda}$ of $\mathfrak{N}_\lambda$ and write $||F_\lambda^W \psi_n||^2$ in terms of Φ; this is possible since $F_\lambda^W \psi_n \in \mathfrak{N}_\lambda$, and $\mathfrak{N}_\lambda$ is itself a Hilbert space.

$$\mathrm{Tr}_\Psi(F_\lambda^W) = \sum_{n=1}^{\infty} (\psi_n, F_\lambda^W \psi_n) = \sum_{n=1}^{\infty} ||F_\lambda^W \psi_n||^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{D_\lambda} |(\phi_m, F_\lambda^W \psi_n)|^2.$$

Because F_λ^W is self-adjoint and $F_\lambda^W \phi_m = \phi_m$, F_λ^W may be removed in the last sum. Since the sum converges and all the terms are positive, it is absolutely convergent and may be resummed in any order. In particular, we may interchange the summations and apply Parseval's formula again,

$$\mathrm{Tr}_\Psi(F_\lambda^W) = \sum_{m=1}^{D_\lambda} \sum_{n=1}^{\infty} |(\phi_m, \psi_n)|^2 = \sum_{m=1}^{D_\lambda} ||\psi_m||^2 = D_\lambda.$$

As shown above, $\mathrm{Tr}_\Psi(W) \geq \lambda\, \mathrm{Tr}_\Psi(F_\lambda^W) = \lambda D_\lambda$ consequently $D_\lambda \to 0$ as $\lambda \to \infty$. Since D_λ is an integer, it must vanish for all $\lambda \geq \mu$, for some μ. Thus $F_\lambda^W = 0$ if $\lambda \geq \mu$, and W is bounded. The spectral theorem then reads

$$W = \int_0^\mu \lambda\, dE_\lambda^W = \int_0^\epsilon \lambda\, dE_\lambda^W + \int_\epsilon^\mu \lambda\, dE_\lambda^W \quad \text{if} \quad 0 < \epsilon < \mu.$$

The first integral converges uniformly to 0 since

$$\left|\left| \int_0^\epsilon \lambda\, dE_\lambda^W \phi \right|\right|^2 = \int_0^\epsilon \lambda^2\, d\, ||E_\lambda^W \phi||^2 \leq \epsilon^2 \int_0^\mu d\, ||E_\lambda^W \phi||^2 = \epsilon^2\, ||\phi||^2$$

and consequently,

$$\left|\left| \int_0^\epsilon \lambda\, dE_\lambda^W \right|\right| \leq \epsilon.$$

We thus have

$$W(\epsilon) = \int_\epsilon^\mu \lambda\, dE_\lambda^W \rightrightarrows W \quad \text{as} \quad \epsilon \to 0.$$

The null space of $W(\epsilon)$ is $\mathfrak{M}_\epsilon^W = E_\epsilon^W \mathfrak{H}$ whose orthogonal complement $\mathfrak{N}_\epsilon^W = F_\epsilon^W \mathfrak{H}$ is finite dimensional. Thus $W(\epsilon)$ can be written as a "diagonal matrix" $W(\epsilon) = \sum_{m=1}^{N(\epsilon)} w_m P_{\phi_m}$ with orthonormal $\phi_m \in \mathfrak{N}_\epsilon^W$. Because of this, $(W - W(\epsilon))\phi_m = 0$ and $W\phi_m = W(\epsilon)\phi_m = w_m \phi_m$. If we specify the choice of the ϕ_m at each degeneracy, then W will be represented as an infinite series

$$W = \sum_{m=1}^{\infty} w_m P_{\phi_m}$$

in which the convergence is uniform and $W\phi_m = w_m\phi_m$. If we include in $\Phi = \{\phi_m\}_{m=1}^{\infty}$ a basis of $\mathfrak{M}_0^W$ and allow the corresponding w_m to be 0, all the results above remain valid, and Φ is a basis of $\mathfrak{H}$.[2] We now turn to the problem of invariance of $\mathrm{Tr}_\Psi(W)$ with respect to Ψ which will appear as a special case of the considerations of the next paragraph when A is set equal to I.

If A is a bounded operator, the quantity $\mathrm{Tr}_\Psi(W, A)$ is defined as the series $\sum_{n=1}^{\infty} (\psi_n, WA\psi_n)$ provided this series converges. This reduces in the case $A = I$ to $\mathrm{Tr}_\Psi(W)$. If this quantity is finite for some basis Ψ_0, then the series does converge, and

$$\mathrm{Tr}_\Psi(W, A) = \mathrm{Tr}_\Phi(W, A) = \mathrm{Tr}_\Phi(A, W), \qquad (1)$$

where Φ is the basis of eigenvectors of W as dis-

[2] This means that W is not only bounded, but also completely continuous, cf. R. Schatten, R. Schatten, *Norm Ideals of Completely Continuous Operators* [Erg. Math. Wiss. 27 (Springer-Verlag, Berlin, 1960)]. From, the summability of the eigenvalues of W which is proved later, it follows that $W^{\frac{1}{2}}$ belongs to the Schmidt class, the set of all completely continuous operators for which the series of eigenvalues is square summable. Thus $W = W^{\frac{1}{2}} W^{\frac{1}{2}}$ belongs to the trace class, the set of all products of two operators from the Schmidt class. In the counter-example of Appendix I, the operator A is such that AW is bounded but not in the trace class, and this permits convergence without absolute convergence of the series. Some of the following discussion can also be found in Schatten's book.

cussed above. For the proof, we apply Parseval's formula in the following computation:

$$\begin{aligned}\mathrm{Tr}_\Psi(W, A) &= \sum_{n=1}^{\infty} (\psi_n, WA\psi_n) = \sum_{n=1}^{\infty} (W\psi_n, A\psi_n) \\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} (W\psi_n, \phi_m)(\phi_m, A\psi_n) \\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} w_m(A^*\phi_m, \psi_n)(\psi_n, \phi_m) \\ &= \sum_{m=1}^{\infty} w_m \sum_{n=1}^{\infty} (A^*\phi_m, \psi_n)(\psi_n, \phi_m). \qquad (2)\end{aligned}$$

If the absolute convergence of the last series is demonstrated, the interchange of the summations is justified, and the steps may be reversed in order to show the convergence of the defining series for $\mathrm{Tr}_\Psi(W, A)$. For this purpose, we apply the inequality $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$, remembering that $||A^*|| = ||A||$,

$$\begin{aligned}&\sum_{m=1}^{\infty} w_m \sum_{n=1}^{\infty} |(A^*\phi_m, \psi_n)(\psi_n, \phi_m)| \\ &\quad\leq \sum_{m=1}^{\infty} w_m \sum_{n=1}^{\infty} \tfrac{1}{2}(|(\psi_n, A^*\phi_m)|^2 + |(\psi_n, \phi_m)|^2) \\ &\quad= \sum_{m=1}^{\infty} \tfrac{1}{2} w_m(||A^*\phi_m||^2 + ||\phi_m||^2) \\ &\quad\leq \sum_{m=1}^{\infty} \tfrac{1}{2} w_m(||A||^2 + 1).\end{aligned}$$

The convergence of the last series above is a result of the isotony of the trace and the fact that for any n,

$$W \leq \sum_{m=1}^{n} w_m P_{\phi m}.$$

$$\mathrm{Tr}_{\Psi_0}(W) \geq \sum_{m=1}^{n} w_m \mathrm{Tr}(P_{\phi m}) = \sum_{m=1}^{n} w_m$$

for any n, and consequently

$$\sum_{m=1}^{\infty} w_m < \mathrm{Tr}_{\Psi_0}(W).$$

Returning to (2), we may immediately apply Parseval's formula and get

$$\begin{aligned}\mathrm{Tr}_\Psi(W, A) &= \sum_{m=1}^{\infty} w_m(A^*\phi_m, \phi_m) = \sum_{m=1}^{\infty} (w_m\phi_m, A\phi_m) \\ &= \sum_{m=1}^{\infty} (\phi_m, WA\phi_m) = \mathrm{Tr}_\Phi(W, A) \\ &= \sum_{m=1}^{\infty} (\phi_m, Aw_m\phi_m) = \mathrm{Tr}_\Phi(A, W) \qquad (3)\end{aligned}$$

which proves (1).

From (1) we may draw the following conclusions:

(a) If $\mathrm{Tr}_{\Psi_0}(W)$ is finite, then $\mathrm{Tr}_\Psi(W)$ is finite for any Ψ and equal to $\mathrm{Tr}_{\Psi_0}(W)$. The trace of a positive bounded operator is thus a unitary invariant. We designate the common value of $\mathrm{Tr}_\Psi(W)$ for all Ψ by $\mathrm{Tr}(W)$. By direct computation, it is seen that $\mathrm{Tr}_\Phi(W) = \sum_{m=1}^{\infty} w_m = \sum_{m=1}^{\infty} (\phi_m, W\phi_m)$.

(b) If $\mathrm{Tr}(W)$ is finite, then $\mathrm{Tr}_\Psi(W, A) = \mathrm{Tr}_\Phi(W, A)$ for any Ψ. Their common value is designated by $\mathrm{Tr}(WA)$.

(c) From (3), $\mathrm{Tr}(WA) = \sum_{m=1}^{\infty} w_m(\phi_m, A\phi_m)$ is positive if A is. Since $\mathrm{Tr}(WA)$ is linear in A, it follows that

$$A \leq B \Rightarrow \mathrm{Tr}(WA) \leq \mathrm{Tr}(WB). \qquad (4)$$

If P is any projector and Ψ is an orthonormal basis of eigenvectors of P, then direct computation shows that $\mathrm{Tr}_\Psi(WP) = \mathrm{Tr}_\Psi(PW) = \mathrm{Tr}_\Psi(PWP) \leq \mathrm{Tr}(W)$.

Finally, we prove that if $A_n \to A$, then $\mathrm{Tr}(WA_n) \to \mathrm{Tr}(WA)$. From the strong convergence of the operators A_n, it follows that there is an upper bound for their norms,[3] C. By linearity of the trace, A may be taken as 0. Then since

$$\begin{aligned}|\mathrm{Tr}(WA_n)| &\leq \sum_{m=1}^{\infty} w_m |(\psi_m, A_n\psi_m)| \\ &\leq \sum_{m=1}^{\infty} w_m C \leq \mathrm{Tr}(W)C,\end{aligned}$$

the terms of this series which all approach 0 are uniformly bounded, and by the criterion of Weierstrass, the series itself converges to 0.

3. INTEGRAL REPRESENTATION OF THE TRACE

If A is a bounded self adjoint operator, then we define $\Phi(A) = \mathrm{Tr}(WA)$. Φ is an isotonic linear functional with $\Phi(I) = 1$. With the notation of Nagy[4] p. 24, the spectral theorem may be expressed

$$\sum_i \lambda_i E^A(\delta_i) \leq A \leq \sum_i \lambda_{i+1} E^A(\delta_i),$$

where $E^A(\delta_i) = E_{\lambda_{i+1}}{}^A - E_{\lambda_i}{}^A$ for a given partition $Z = \{\lambda_i\}_i$ of the interval $[-||A||, +||A||]$. We define $||Z|| = \max_i(\lambda_{i+1} - \lambda_i)$ and from the isotony and linearity of Φ obtain

$$\sum_i \lambda_i \Phi(E^A(\delta_i)) \leq \Phi(A) \leq \sum_i \lambda_{i+1} \Phi(E^A(\delta_i)),$$

where the difference between the extreme members is

$$\begin{aligned}&\sum_i (\lambda_{i+1} - \lambda_i)\Phi(E^A(\delta_i)) \\ &\qquad\leq \sum_i ||Z||\, \Phi(E^A(\delta_i)) = ||Z||\, \Phi(I).\end{aligned}$$

[3] S. Banach, *Théorie des opérationes linéaires* (Monografije Matematyczne, Warsaw, 1932), p. 80.

[4] B. v. Sz.- Nagy, *Spektraldarstellung Linearer Transformationen des Hilbertschen Raumes* (Springer-Verlag, Berlin 1942).

As $||Z_n|| \to 0$, the two integral sums approach $\Phi(A)$ whatever the choice of the sequence of partitions Z_n, and since $\Phi(E^A(\delta_i)) = \Phi(E_{\lambda_{i+1}}{}^A) - \Phi(E_{\lambda_i}{}^A)$, we have

$$\Phi(A) = \int_{-\infty}^{\infty} \lambda \, d\Phi(E_\lambda{}^A) \tag{5}$$

and, consequently, a simple upper bound

$$|\Phi(A)| \leq \int_{-||A||}^{||A||} |\lambda| \, d\Phi(E_\lambda{}^A)$$
$$\leq ||A|| \int_{-\infty}^{\infty} d\Phi(E_\lambda{}^A) = ||A|| \, \Phi(I) \tag{6}$$

which shows that Φ preserves uniform convergence. With $\Phi(A) = \mathrm{Tr}\,(WA)$, (7) reads

$$\mathrm{Tr}\,(WA) = \int_{-\infty}^{\infty} \lambda \, d\,\mathrm{Tr}\,(WE_\lambda{}^A). \tag{7}$$

From the monotony of $\lambda \to E_\lambda{}^A$ and isotony of Φ follows the monotony of $\lambda \to \Phi(E_\lambda{}^A)$. This suggests the use of (7) as a definition of $\mathrm{Tr}\,(WA)$ for unbounded self-adjoint operators A as long as one of the integrals over $(-\infty, 0]$, $[0, +\infty)$ is finite.

From this definition, several commonly used formulas may be derived. If A has a discrete spectrum: $A = \sum_{n=1}^{\infty} \alpha_n P_n$ where the α_n are the eigenvalues of A and the $P_n = E_{\alpha_n}{}^A - E_{\alpha_n -}{}^A$ are the projections on the corresponding eigenspaces, then the Stieltjes integral above reduces to the infinite series

$$\mathrm{Tr}\,(WA) = \sum_{n=1}^{\infty} \alpha_n \,\mathrm{Tr}\,(WP_n)$$

If $\{\psi_m^{(n)}\}_m$ is a basis of $P_n\mathfrak{H}$, then $\mathrm{Tr}\,(WP_n) = \sum_m (\psi_m^{(n)}, W\psi_m^{(n)})$ and

$$\mathrm{Tr}\,(WA) = \sum_{n=1}^{\infty} \sum_m \alpha_m^{(n)} (\psi_m^{(n)}, W\psi_m^{(n)}),$$

where $\alpha_m^{(n)} = \alpha_n$ is the eigenvalue of $\psi_m^{(n)}$ under A. From the assumption that either $\mathrm{Tr}\,(WA_+)$ or $\mathrm{Tr}\,(WA_-)$ is finite, it follows that either the above series is absolutely convergent or definitely divergent and therefore the series may be rearranged in any way. If we relabel the eigenvectors $\{\psi_r\}_{r=1}^{\infty}$ and the corresponding eigenvalues α_r and rearrange the series, we obtain

$$\mathrm{Tr}\,(WA) = \sum_{r=1}^{\infty} \alpha_r (\psi_r, W\psi_r). \tag{8}$$

If A is positive, then

$$\mathrm{Tr}\,(WA) = \lim_{\mu\to\infty} \int_0^{\mu} \lambda \, d\,\mathrm{Tr}\,(WE_\lambda^A) = \lim_{\mu\to\infty} \mathrm{Tr}\,(WAE_\mu^A).$$

Since $AE_\mu{}^A$ is a bounded operator, we may apply the formula in Sec. 2(c) to obtain

$$\mathrm{Tr}\,(WAE_\mu{}^A) = \sum_{m=1}^{\infty} w_m(\phi_m, AE_\mu{}^A\phi_m)$$
$$= \sum_{m=1}^{\infty} w_m \int_0^{\mu} \lambda \, d(\phi_m, E_\lambda{}^A\phi_m)$$
$$= \sum_{m=1}^{\infty} w_m \int_0^{\mu} \lambda \, d\, ||E_\lambda{}^A\phi_m||^2.$$

Since the terms of the sum are positive, increasing functions of μ, we may take the limit inside and find

$$\mathrm{Tr}\,(WA) = \sum_{m=1}^{\infty} w_m \int_0^{\infty} \lambda \, d\, ||E_\lambda{}^A\phi_m||^2. \tag{9}$$

The individual terms are finite iff $\phi_m \in \mathrm{dom}\,(A^{\frac{1}{2}})$ as seen from the substitution $\mu = \lambda^{\frac{1}{2}}$:

$$\int_0^{\infty} \lambda \, d\, ||E_\lambda{}^A\phi_m||^2 = \int_0^{\infty} \mu^2 \, d\, ||E_\lambda{}^{A^{\frac{1}{2}}}\phi_m||^2.$$

If, in addition, $\phi_m \in \mathrm{dom}\,(A)$ for all m, then we have

$$\mathrm{Tr}\,(WA) = \sum_{m=1}^{\infty} w_m \, ||A^{\frac{1}{2}}\phi_m||^2 = \sum_{m=1}^{\infty} w_m(\phi_m, A\phi_m) \tag{10}$$

from which linearity in A is obvious. It is proved in Appendix II that if A commutes with B, then $\mathrm{Tr}\,(WA) + \mathrm{Tr}\,(WB) = \mathrm{Tr}\,(W(A + B))$. However, no general results are obtained about additivity if A and B do not commute.

4. THE STATISTICAL FORMULA

The considerations of this section will be restricted to bounded operators; the formula will be extended in the next section. The first two axioms are of an algebraic nature (not requiring the concept of limit) and may be stated as follows: for any pair of bounded self adjoint operators A, B,

(I) $M(A^2) \geq 0$, $M(I) = 1$;

(II) $M(A + B) = M(A) + M(B)$.

From (I) and the existence of a self-adjoint square root $A^{\frac{1}{2}} = \int_0^{\infty} (\lambda)^{\frac{1}{2}} \, dE_\lambda{}^A$ of any positive operator A, it follows that if $A \geq 0$, then $M(A) = M((A^{\frac{1}{2}})^2) \geq 0$. It can now be shown that M is linear. The proof follows.

(a) $M(nA) = nM(A)$ (by induction) for any integer $n \geq 0$;

(b) $mM(A) = M(mA)$

$$= M\left(n \frac{m}{n} A\right)$$
$$= nM\left(\frac{m}{n} A\right),$$
$$M\left(\frac{m}{n} A\right) = \frac{m}{n} M(A).$$

(c) $M(I) = M(I + 0)$
$= M(I) + M(0),$
$M(0) = 0.$

(d) $M(A) + M(-A)$
$= M(A - A)$
$= M(0) = 0,$
$M(-A) = -M(A).$

(e) If $A \leq B$, then $B - A \geq 0$ and $M(B - A) \geq 0$. Since $M(B) = M(A) + M(B - A) \geq M(A)$, we have $A \leq B \Rightarrow M(A) \leq M(B)$.

(f) If $A \geq 0$ and $\alpha > 0$, then for any pair of rational numbers q, r with $q < \alpha < r$, we have $qA \leq \alpha A \leq rA$ so that from (b) and (e), we get $qM(A) = M(qA) \leq M(\alpha A) \leq M(rA) = rM(A)$. Also $qM(A) \leq \alpha M(A) \leq rM(A)$ so that $|M(\alpha A) - \alpha M(A)| \leq (r - q)M(A)$. Since $r - q$ may be arbitrarily small, $M(\alpha A) = \alpha M(A)$ which is true also if $\alpha = 0$.

(g) If $A \geq 0$ and $\alpha < 0$, then $|\alpha| = -\alpha > 0$ and $M(\alpha A) = M(-|\alpha| A) = -M(|\alpha| A) = -|\alpha| M(A) = \alpha M(A)$.

(h) If $\alpha \geq 0$ and $A = A_+ - A_-$, then $M(\alpha A) = M(\alpha A_+ - \alpha A_-) = M(\alpha A_+) - M(\alpha A_-) = \alpha M(A_+) - \alpha M(A_-) = \alpha M(A_+ - A_-) = \alpha M(A)$. If $\alpha < 0$, we may consider $|\alpha| (-A)$.

The function M is linear, and from (e), isotonic. $M(I)$ is finite so that the integral formula (5) of Sec. 3 applies to M. If A is a bounded self-adjoint operator, then

$$M(A) = \int_{-\infty}^{\infty} \lambda \, dM(E_\lambda^A). \tag{11}$$

To make the following considerations easier (to avoid rearrangement of series and complicated, unnatural definitions), we extend the domain of definition of M to include all bounded operators by means of the canonical decomposition $A = A^R + iA^I$, $A^R = \frac{1}{2}(A + A^*)$, $A^I = \frac{1}{2}i(A^* - A)$ as follows:

$$\hat{M}(A) = M(A^R) + iM(A^I). \tag{12}$$

The function $\hat{M}$ is linear (preserves linear combinations with *complex* coefficients); and if $A^* = A$, then $A^R = A$, $A^I = 0$, $\hat{M}(A) = M(A)$. $\hat{M}$ also preserves uniform convergence since

$$\begin{aligned}|\hat{M}(A)|^2 &= |M(\tfrac{1}{2}(A + A^*))|^2 \\ &+ |M(\tfrac{1}{2}i(A^* - A))|^2 \leq ||\tfrac{1}{2}(A + A^*)||^2 \\ &+ ||\tfrac{1}{2}i(A^* - A)||^2 \leq \tfrac{1}{4}(||A|| + ||A^*||)^2 \\ &+ \tfrac{1}{4}(||A^*|| + ||A||)^2 = 2\,||A||.\end{aligned} \tag{13}$$

It is clear from a twofold application of Parseval's theorem that we have the following strong convergence:

$$P_\psi = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (\psi, \psi_n)\, |\psi_m \times \psi_n|\, (\psi_m, \psi) \tag{14}$$

if $||\psi|| = 1$ and $\{\psi_n\}_{n=1}^{\infty}$ is an orthonormal basis. The operator $|\psi \times \phi|$ is defined by $|\psi \times \phi|\, \chi = \psi(\phi, \chi)$ for $\chi \in \mathfrak{H}$. To show that the convergence is uniform, we recall the definition of

$$||A|| = \sup_{||\phi||=1} ||A\phi||$$

and note that

$$\begin{aligned}&\left|\left|\sum_{n=N}^{\infty} (\psi, \psi_n)\, |\psi_m \times \psi_n|\right|\right|^2 \\ &\quad = \sup_{||\phi||=1} \left|\left|\sum_{n=N}^{\infty} (\psi, \psi_n)\psi_m(\psi_n, \phi)\right|\right|^2 \\ &\quad = \sup_{||\phi||=1} \left|\left(\sum_{n=N}^{\infty} \psi_n(\psi_n, \psi), \phi\right)\right|^2 \\ &\quad \leq \left|\left|\sum_{n=N}^{\infty} \psi_n(\psi_n, \psi)\right|\right|^2 \\ &\quad = \sum_{n=N}^{\infty} |(\psi_n, \psi)|^2 \to 0 \quad \text{as} \quad N \to \infty\end{aligned}$$

since it is the remainder of a convergent series. For the summation over m, we have similarly

$$\begin{aligned}&\left|\left|\sum_{m=M}^{\infty} \sum_{n=1}^{\infty} (\psi, \psi_n)\, |\psi_m \times \psi_n|\, (\psi_m, \psi)\right|\right|^2 \\ &\quad = \sup_{||\phi||=1} \left|\left|\sum_{m=M}^{\infty} (\psi_m, \psi)\psi_m \sum_{n=1}^{\infty} (\psi, \psi_n)(\psi_n, \phi)\right|\right|^2 \\ &\quad = \sup_{||\phi||=1} \left|\left|\sum_{m=M}^{\infty} (\psi_m, \psi)\psi_m\right|\right|^2 |(\psi, \phi)|^2 \\ &\quad = \sum_{m=M}^{\infty} |(\psi_m, \psi)|^2 \to 0 \quad \text{as} \quad M \to \infty.\end{aligned}$$

Since both series converge uniformly, we may apply the result (13) and linearity to obtain the formula

$$\begin{aligned}M(P_\psi) &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (\psi, \psi_n)\hat{M}(|\psi_m \times \psi_n|)(\psi_m, \psi) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (\psi, \psi_n) w_{nm}(\psi_m, \psi),\end{aligned} \tag{15}$$

where w_{nm} is defined as $\hat{M}(|\psi_m \times \psi_n|)$. To show that there is an operator W for which $w_{nm} = (\psi_n, W\psi_m)$, we consider the bilinear form

$$f(\phi, \psi) = \sum_{m=1}^{\infty} (\phi, \psi_n) w_{nm}(\psi_n, \psi).$$

The quadratic form induced by f, defined by $\hat{f}(\phi) = f(\phi, \phi)$, is clearly equal to $M(P_\phi)$. Since $0 \leq P_\phi \leq I$,

$0 \leq M(P_\phi) \leq M(I) = 1$, f is bounded. This implies[5] that f is bounded and that there is a unique bounded self-adjoint [since $w_{mn} = (w_{nm})^*$] operator W for which $w_{nm} = (\psi_n, W\psi_m)$. W is positive since $(\phi, W\phi) = f(\phi, \phi) = M(P_\phi) \geq 0$. By applying Parseval's formula in the opposite direction, we find

$$M(P_\psi) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (\psi, \psi_n)(\psi_n, W\psi_m)(\psi_m, \psi)$$
$$= (\psi, W\psi) = \mathrm{Tr}\,(WP_\psi). \qquad (16)$$

This formula extends immediately to projectors P on finite-dimensional subspaces $\mathfrak{M}$ if an orthonormal basis of $\mathfrak{M}$, $\{\psi_m\}_{m=1}^n$ is taken and P is expressed as $\sum_{m=1}^n P_{\psi_m}$. Then

$$M(P) = \sum_{m=1}^{n} M(P_{\psi_m})$$
$$= \sum_{m=1}^{n} \mathrm{Tr}\,(WP_{\psi_m}) = \mathrm{Tr}\,(WP). \qquad (17)$$

The third axiom requires for its interpretation a consideration of the probability function induced by M[6] on the set of projection operators. In the correspondence between self-adjoint operators and observables, the projectors correspond to observables having only 0 and 1 as measurable values. These in turn correspond to "yes–no" propositions; the proposition $\pi(P)$ is verified when the measured value of the observable corresponding to P is 1. The probability that π is verified is then the expectation value of P, i.e., if π is the proposition corresponding to P, then

$$w(\pi) = M(P) \qquad (18)$$

is the probability of verification of π when the state of the system is described by the function M.

For the concepts and results in this paragraph, we refer to Nikodým.[7] If $\{P_n\}_{n=1}^\infty$ is a decreasing sequence of projectors, then the corresponding propositions π_n will also decrease in the sense that $\pi_{n+1} \Rightarrow \pi_n$. The projectors commute and thus correspond to simultaneously measurable operators. The π_n will thus be compatible, and from the isotony of M, it follows that $w(\pi_n) \geq w(\pi_{n+1})$. We may define the infimum of the projectors $\wedge_{n=1}^\infty P_n$ as the greatest projector which is less than or equal to P_n for all n. This exists for any family of projectors; and in particular, if $\mathfrak{M}_n = P_n\mathfrak{H}$, $\mathfrak{M} = \bigcap_{n=1} \mathfrak{M}_n$, then $\wedge_{n=1}^\infty P_n = \mathrm{proj}\,(\mathfrak{M})$. The conjunction of the propositions is then $\pi\,(\mathrm{proj}\,(\mathfrak{M})) = \wedge_{n=1}^\infty \pi_n$; this is the proposition which is verified iff π_n is verified for each n. If $\wedge_{n=1}^\infty P_n = 0$, then $w(\wedge_{n=1}^\infty \pi_n) = 0$; this proposition is impossible to verify. As the third axiom, then, we require that

(III) If $\{P_n\}_{n=1}^\infty$ is a decreasing sequence of projectors and if $\wedge_{n=1}^\infty P_n = 0$, then $w(\pi_n) = w(\pi(P_n)) \to 0$ as $n \to \infty$.

We first verify that if $P_n \to 0$ (strong convergence) under the assumption that the sequence is decreasing, then $\wedge_{n=1}^\infty P_n = 0$. For if $\psi \in \mathfrak{M} \subseteq \mathfrak{M}_n$ for all n, then $P_n\psi = \psi$, and from $P_n\psi \to 0$ follows $\psi = 0$. Thus $\mathfrak{M} = \{0\}$ and $\wedge_{n=1}^\infty P_n = 0$. If $\{P_n\}_{n=1}^\infty$ is an increasing sequence of projectors for which $P_n \to P$, then $\wedge_{n=1}^\infty (P - P_n) = 0$ because $\{P - P_n\}_{n=1}^\infty$ is a sequence of projectors decreasing strongly to 0. But from axiom (III), we have $M(P) = M(P_n) + M(P - P_n)$ and $M(P - P_n) = w(\pi(P - P_n)) \to 0$ so that $M(P) = \lim_{n\to\infty} M(P_n)$. It is easily checked that given any projector P and any basis of $\mathfrak{M} = P\mathfrak{H}$, it is the strong limit of the increasing sequence $\{P_n\}_{n=1}^\infty$ where P_n is the projection on the subspace generated by the first n vectors in the basis (We assume them indexed by integers which is possible since $\mathfrak{H}$, and therefore $\mathfrak{M}$, is separable). This, with (17) yields

$$M(P) = \lim_{n\to\infty} \mathrm{Tr}\,(WP_n) = \mathrm{Tr}\,(WP),$$

where the last equality results from the fact that $\mathrm{Tr}\,(WP) - \mathrm{Tr}\,(WP_n)$, when computed with respect to the basis discussed above is the remainder of a convergent series.

Thus we may conclude, using the integral representations for M and the trace and the fact that $M(E_\lambda^A) = \mathrm{Tr}\,(WE_\lambda^A)$, that for any bounded self-adjoint A,

$$M(A) = \int_{-\infty}^{\infty} \lambda\, dM(E_\lambda^A)$$
$$= \int_{-\infty}^{\infty} \lambda\, d\,\mathrm{Tr}\,(WE_\lambda^A) = \mathrm{Tr}\,(WA). \qquad (19)$$

5. THE FORMULA FOR UNBOUNDED OPERATORS

Although two unbounded self-adjoint operators may not have a common domain which is dense in $\mathfrak{H}$, this is always the case for commuting operators. Two operators $A = \int_\infty^\infty \lambda\, dE_\lambda^A$ and $B =$

[5] P. R. Halmos, *Introduction to Hilbert Space* Chelsea, New York (1957) pp. 32, 39.

[6] G. Ludwig, *Die Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1954), p. 51. Here the statistical formula is used in some proofs before the discussion of the probabilities is completed, but all of these proofs can be easily modified so as not to depend on this specific characterization of the expectation value function, and circularity of reasoning is thus avoided.

[7] O. M. Nikodým, *Une nouvel appareil mathématique pour la théorie des quanta.* Ann. Inst. H. Poincaré (1947). It should be pointed out that if two propositions are not compatible, then their formal conjunction does not possess the intuitively necessary properties of a conjunction, and that moreover, there is no other proposition which qualifies as a substitute.

$\int_{-\infty}^{\infty} \lambda\, dE_\lambda^B$ are said to commute if E_λ^A commutes with E_λ^B for every $\lambda, \mu \in R$. In this case

$$\mathfrak{D} = \bigcup_{\lambda,\mu\in R} (\mathfrak{M}_\lambda^A \ominus \mathfrak{M}_{-\lambda}^A) \cap (\mathfrak{M}_\mu^B \ominus \mathfrak{M}_{-\mu}^B)$$

is included in dom $(A) \cap$ dom (B) and is dense in $\mathfrak{H}$. For this reason, and because commuting operators are also simultaneously measurable, we restrict our attention to commuting self adjoint operators. This will be seen to be sufficient, however, for the extension of the formula. Axioms (I) and (II) of Sec. 4 may be immediately generalized. For any pair of commuting self-adjoint operators A, B,

(I$^\infty$) $M(A^2) \geq 0$;

(II$^\infty$) $M(A + B) = M(A) + M(B)$ unless $M(A) = -M(B) = \pm\infty$.

With $F_\mu^A = I - E_\mu^A$, A may be decomposed into a sum of commuting operators $AE_\mu^A + AF_\mu^A$ to which (II$^\infty$) may be applied to yield $M(A) = M(AE_\mu^A) + M(AF_\mu^A)$ in case $M(A)$ exists. For $\mu = 0$, this reads $M(A) = M(-A_-) + M(A_+) = M(A_+) - M(A_-)$ [$M(-A) = -M(A)$ is proved as in Sec. 4(d)] so that only positive operators A need be considered. If this is the case, then AE_μ^A is bounded; and $M(AE_\mu^A) = \mathrm{Tr}\,(WAE_\mu^A) = \int_0^\mu \lambda\, d\,\mathrm{Tr}\,(WE_\lambda^A) \to \int_0^\infty \lambda\, d\,\mathrm{Tr}\,(WE_\lambda^A) = \mathrm{Tr}\,(WA)$ as $\mu \to \infty$. Because the integral above is an increasing function of μ, $M(AF_\mu^A)$ is decreasing and is positive since $AF_\mu^A \geq 0$. Hence, the limit $\alpha(A) = \lim_{\mu\to\infty} M(AF_\mu^A)$ exists and is nonnegative. $M(A) = \mathrm{Tr}\,(WA) + \alpha(A) \geq \mathrm{Tr}\,(WA)$.

That equality above cannot be proved from the axioms is seen from the following counter example.

$$M(A) = \begin{cases} +\infty & \text{if } A \geq 0 \text{ is unbounded,} \\ \mathrm{Tr}\,(WA) & \text{if } A \text{ is bounded,} \\ -\infty & \text{if } A \leq 0 \text{ is unbounded.} \end{cases}$$

To eliminate such possibilities, a regularity condition is needed which permits the calculation of the expectation value of an unbounded operator from the expectation values of certain bounded operators. One possibility is a much weakened form of von Neumann's axiom B',

(III$_a^\infty$) if $\{A_n\}_{n=1}^\infty$ is a family of commuting bounded positive operators and $A = \sum_{n=1}^\infty A_n$ (strong convergence), then $M(A) = \sum_{n=1}^\infty M(A_n)$,

or the equivalent

(III$_b^\infty$) if $\{A_n\}_{n=1}^\infty$ is an increasing sequence of bounded operators such that $A_n \to A$, then $M(A_n) \to M(A)$.

The statistical formula follows from this if A_n is set equal to AE_n^A. Then $AE_n^A \to A$ so that $M(A_nE_n^A) \to M(A)$, and as has been seen above, $M(A_n) = \mathrm{Tr}\,(WA_n) \to \mathrm{Tr}\,(WA)$.

An alternate approach to this axiom, which may be more appealing, relies on the connection with classical probability theory[8] in which the proposition corresponding to E_λ^A is that of measuring a value of A less than or equal to λ. If we set $w(\lambda) = M(E_\lambda^A)$, then w is a probability distribution since $w(-\infty) = 0$, $w(+\infty) = 1$, and is right continuous and nondecreasing. Also, from the integral representation (11),

$$M(A) = \int_{-\infty}^{\infty} \lambda\, dM(E_\lambda^A) = \int_{-\infty}^{\infty} \lambda\, dw(\lambda) \qquad (19)$$

for any bounded self-adjoint operator A. This suggests the requirement that the formula remain valid for unbounded A, i.e.,

(III$_c^\infty$) $M(A) = \int_{-\infty}^{\infty} \lambda\, dw(\lambda)$ if $w(\lambda) = M(E_\lambda^A)$.

From the observation that $w(\lambda) = \mathrm{Tr}\,(WE_\lambda^A)$ and the definition of Tr (WA) for unbounded A, the statistical formula follows.

As a final point, we remark that the possibility $M(A) = +\infty$ is not restricted to this theory, but even in the pure-state theory where it is assumed that $W = P_\psi$ for a certain ψ, one has $M(A) \geq \mathrm{Tr}\,(P_\psi A) = \int_{-\infty}^{\infty} \lambda\, d\, ||E_\lambda^A\psi||^2$. If $A \geq 0$ and $\psi \notin$ dom $(A^{\frac{1}{2}})$, then this integral is infinite. It is not clear even if the integral is finite that $M(A)$ is a linear function of A. If $\psi \in$ dom (A), then from the spectral theorem, $M(A) = (\psi, A\psi)$, but it is possible that $\psi \notin$ dom (A) and $\psi \in$ dom $(A^{\frac{1}{2}})$. In this case, the integral is finite, but $(\psi, A\psi)$ is not defined.

ACKNOWLEDGMENT

The last result of Sec. 2 was suggested to me by Dr. T. F. Jordan.

APPENDIX I

One might be tempted to use the definition $\mathrm{Tr}_\Psi\,(WA) = \sum_{n=1}^\infty (\psi_n, WA\psi_n)$ when A is unbounded provided the basis $\{\psi_n\}_{n=1}^\infty \subseteq$ dom (A) so that the expression is well defined with the possible exception of the convergence of the series. The following example shows that the definition is not independent of the basis. We choose a basis and separate it into

[8] G. Ludwig Ref. 6, Chap. II 2; G. Birkhoff and J. von Neumann Ann. Math. 37, 823 (1936).

two infinite subsets $\Psi = \{\psi_n\}_{n=1}^{\infty} \cup \{\phi_n\}_{n=1}^{\infty}$. ψ_n and ϕ_n span a two-dimensional subspace of $\mathfrak{H}$ for each n, and $\mathfrak{H}$ is the direct sum of these subspaces. Now let

$$W = \sum_{n=1}^{\infty} W_n; \qquad W_n = 2^{-n}\,\text{proj}\,(2^{-\frac{1}{2}}(\psi_n + \phi_n)),$$

$$A = \sum_{n=1}^{\infty} A_n; \qquad A_n = 2^{n}\,\text{proj}\,(\alpha_n\phi_n + \beta_n\psi_n),$$

where α_n, β_n are real and $\alpha_n^2 + \beta_n^2 = 1$.[9] If ψ_n and ϕ_n are grouped and the summation is performed with this grouping, then we obtain

$$\text{Tr}_{\Psi_0}(WA) = \sum_{n=1}^{\infty} \text{Tr}_{\{\psi_n,\phi_n\}}(W_nA_n),$$

where the terms of the series are traces in the two-dimensional spaces. Now

$$\text{Tr}_{\{\psi,\phi\}}(P_{(\psi+\phi)/\sqrt{2}}P_{(\alpha\psi+\beta\phi)})$$
$$= \tfrac{1}{2}(\psi,\ \psi+\phi)(\psi+\phi,\ \alpha\psi+\beta\phi)(\alpha\psi+\beta\phi,\ \psi) + \cdots$$
$$= \tfrac{1}{2}(\alpha+\beta)\alpha + \tfrac{1}{2}(\alpha+\beta)\beta,$$

where the terms arising from ψ and ϕ are still separate. The series becomes

$$\text{Tr}_{\Psi_0}(WA) = \sum_{n=1}^{\infty} \tfrac{1}{2}((\alpha_n+\beta_n)\alpha_n + (\alpha_n+\beta_n)\beta_n)$$
$$= \sum_{n=1}^{\infty} \tfrac{1}{2}(\alpha_n+\beta_n)^2.$$

Conclusions

(1) If α_n, β_n are chosen so that $\alpha_n + \beta_n \geq 0$, and $\sum_{n=1}^{\infty}(\alpha_n+\beta_n)^2$ converges, then $\alpha_n + \beta_n \to 0$ as $n \to \infty$. Since $|\alpha_n| \leq 1$, $|\beta_n| \leq 1$, the series also converges if the parentheses are removed. If $\alpha_n \geq 0$ then $\alpha_n, -\beta_n \to 2^{-\frac{1}{2}}$

(2) If at the same time $\sum_{n=1}^{\infty}(\alpha_n+\beta_n)$ does not converge, then $\sum_{n=1}^{\infty}(\alpha_n+\beta_n)\alpha_n$ will not converge; and therefore the series $\sum_{n=1}^{\infty}(\psi_n, WA\psi_n)$ does not converge absolutely, and

(3) by a rearrangement theorem of Riemann, the vectors in the basis may be rearranged so that the series has any value desired.

(4) An example is provided by $\alpha_n = (\frac{1}{2}+n^{-1})^{\frac{1}{2}}$ $\beta_n = -|\frac{1}{2}-n^{-1}|^{\frac{1}{2}}$. For $n > 1$, $\alpha_n+\beta_n = (\frac{1}{2}+n^{-1})^{\frac{1}{2}} = [1 + s(n^{-1})]/\sqrt{2}n$ where s is a convergent power series without a constant term.

[9] That A is a self-adjoint operator follows from Lemma of Sec. 120 in F. Riesz and B. v. Sz.- Nagy, *Functional Analysis*, (Fredrick Ungar Publishing Company, New York).

APPENDIX II

The linearity of $\Phi(A) = \text{Tr}\,(WA)$ as a function of A is trivaial for bounded A because of the series representation, but this series can no longer be used if A is an arbitrary self adjoint operator because of the result of Appendix I. However, if A and B commute, then it is possible to prove that $\Phi(A+B) = \Phi(A) + \Phi(B)$ and apply the considerations in Sec. 4, (a)–(h) to prove linearity as long as there is not a subtraction of infinities. We assume that $A \geq 0$, $B \geq 0$ and thus $A + B \geq 0$.

The proof begins with the observation that as the product of commuting positive operators, $E_\alpha{}^A E_\beta{}^B E_\mu{}^{A+B} \geq 0$ and is increasing as a function of α, β, μ; thus $\Phi(E_\alpha{}^A E_\beta{}^B E_\mu{}^{A+B})$ is increasing in α, β, and μ. We have also that

$$\lim_{\alpha,\beta\to\infty}(E_\alpha{}^A E_\beta{}^B E_\mu{}^{A+B}) = \Phi(E_\mu{}^{A+B}) \tag{20}$$

since $E_\mu{}^{A+B}(I - E_\alpha{}^A E_\beta{}^B)$ is decreasing and from (10),

$$\Phi(E_\mu{}^{A+B}(I - E_\alpha{}^A E_\beta{}^B)) = \sum_{m=1}^{\infty} w_m\,||E_\mu{}^{A+B}(I - E_\alpha{}^A E_\beta{}^B)\phi_m||^2.$$

Since $(I - E_\alpha{}^A E_\beta{}^B)\phi \to 0$ for any ϕ and the series is convergent for any α, β, the criterion of Weierstrass may be applied to prove that the sum approaches 0 as $\alpha, \beta \to \infty$.

Using (9) and Lebesgue's bounded convergence criterion after an integration by parts, we have

$$\Phi(A+B) = \lim_{\mu\to\infty}\int_0^{\mu} \lambda\, d\Phi(E_\lambda{}^{A+B})$$
$$= \lim_{\mu\to\infty}\left[\mu\Phi(E_\mu{}^{A+B}) - \int_0^{\mu}\Phi(E_\lambda{}^{A+B})\,d\lambda\right]$$
$$= \lim_{\mu\to\infty}\lim_{\alpha,\beta\to\infty}\left[\mu\Phi(E_\alpha{}^A E_\beta{}^B E_\mu{}^{A+B}) - \int\Phi(E_\alpha{}^A E_\beta{}^B E_\lambda{}^{A+B})\,d\lambda\right]$$
$$= \lim_{\mu\to\infty}\lim_{\alpha,\beta\to\infty}\int_0^{\mu}\lambda\, d\Phi(E_\lambda{}^{A+B}E_\alpha{}^A E_\beta{}^B)$$
$$= \lim_{\mu\to\infty}\lim_{\alpha,\beta\to\infty}\Phi((A+B)E_\mu{}^{A+B}E_\alpha{}^A E_\beta{}^B).$$

The last written quantity is increasing in α, β, and μ so that the limits may be interchanged (The limits may be replaced by suprema. These are completely associative and commutative, i.e., $\sup_\alpha \sup_\beta x_{\alpha\beta} =$

$\sup_\beta \sup_\alpha x_{\alpha\beta}$; and thus the limits commute). As a result,

$$\Phi(A+B) = \lim_{\alpha,\beta,\mu\to\infty} \Phi((A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}),$$

$$\Phi(A) = \lim_{\alpha,\beta,\mu\to\infty} \Phi(AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}),$$

$$\Phi(B) = \lim_{\alpha,\beta,\mu\to\infty} \Phi(BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}). \quad (21)$$

If ψ is an arbitrary vector, then $E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi \in$ dom $(A) \cap$ dom $(B) \subseteq$ dom $(A+B)$. Thus $(A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi = AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi + BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi$ and we have

$$\Phi((A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}) = \Phi(AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}) + \Phi(BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B})$$

and consequently, from (21),

$$\Phi(A+B) = \Phi(A) + \Phi(B). \quad (22)$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Can the Position Variable be a Canonical Coordinate in a Relativistic Many-Particle Theory?

EDWARD H. KERNER
Department of Physics, University of Delaware, Newark, Delaware
(Received 30 November 1964)

The time-symmetrical interaction of charges (Wheeler–Feynman electrodynamics) is shown in principle to yield second-order Newtonian-type equations of motion under the restriction that the motions be analytic extensions of free-particle motions. The means for explicitly generating the electrodynamic equations of motion describing invariant world-lines are given. These physically relevant equations do not fit into Dirac's (1949) formulation of Hamiltonian relativistic particle dynamics, where either world-line invariance is given up, or only trivially straight world lines can be described ("zero-interaction theorem" of Currie, Jordan, Sudarshan, 1963). The misfit is due to the requirement in Dirac's scheme that position x be canonical. Under the Lie–Königs theorem, however, Hamiltonian statements of dynamics with invariant world lines remain possible when suitable $Q(x, \dot{x})$ are introduced instead of x as canonical position variables. A group of necessary conditions on the structure of any dynamics that permits x to be canonical are worked out to indicate how stringent is this permission in general.

I. INTRODUCTION

FOLLOWING Dirac's formulation[1] of the general conditions for an Hamiltonian statement of relativistic particle dynamics, it was recognized by Thomas[2] that the dynamics will not in general admit a description of particle motions in terms of invariant world lines.

Thomas[2] and later Foldy[3] have shown explicitly how to construct Dirac's ten fundamental dynamical variables that generate infinitesimal time and space translations, space rotations, and pure Lorentz transformations in accordance with the required symmetry of the dynamics under the full group of inhomogeneous Lorentz transformations; and Currie, Jordan, and Sudarshan have underscored Thomas' point by proving a "zero-interaction theorem" to the effect that (at least in the case of two or three particles) world-line invariance, plus Hamiltonian description of motion, demands free-particle motion only.[4]

This species of misfit between relativity requirements and Hamiltonian requirements must be reckoned a significant one, because on the one hand world-line invariance is so simple and compelling a feature of the relativistic view of nature, saying merely that one and the same particle orbit should indeed be an orbit no matter which inertial observer considers it, and on the other hand because an Hamiltonian rendition of motion stands at the root of the quantal understanding of motion; the denial of world-line invariance may well be a denial of an

[1] P. A. M. Dirac, Rev. Mod. Phys. 21, 392, (1949).
[2] L. H. Thomas, Phys. Rev. 85, 868 (1952); B. Bakamjian and L. H. Thomas, *ibid.* 92, 1300 (1953); B. Bakamjian, *ibid.* 121, 1849 (1961).
[3] L. L. Foldy, Phys. Rev. 122, 275 (1961).
[4] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963); D. G. Currie, J. Math. Phys. 4, 1470 (1963); J. T. Cannon and T. F. Jordan, *ibid.* 5, 299 (1964). For n particles, it appears that the zero-interaction theorem has been proven by H. Leutwyler, to be published in Helv. Phys. Acta. I thank the referee for bringing the latter to my notice.

$\sup_\beta \sup_\alpha x_{\alpha\beta}$; and thus the limits commute). As a result,

$$\Phi(A+B) = \lim_{\alpha,\beta,\mu\to\infty} \Phi((A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}),$$

$$\Phi(A) = \lim_{\alpha,\beta,\mu\to\infty} \Phi(AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}),$$

$$\Phi(B) = \lim_{\alpha,\beta,\mu\to\infty} \Phi(BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}). \quad (21)$$

If ψ is an arbitrary vector, then $E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi \in$ dom $(A) \cap$ dom $(B) \subseteq$ dom $(A+B)$. Thus $(A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi = AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi + BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}\psi$ and we have

$$\Phi((A+B)E_\mu^{A+B}E_\alpha^{A}E_\beta^{B}) = \Phi(AE_\mu^{A+B}E_\alpha^{A}E_\beta^{B}) + \Phi(BE_\mu^{A+B}E_\alpha^{A}E_\beta^{B})$$

and consequently, from (21),

$$\Phi(A+B) = \Phi(A) + \Phi(B). \quad (22)$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Can the Position Variable be a Canonical Coordinate in a Relativistic Many-Particle Theory?

EDWARD H. KERNER
Department of Physics, University of Delaware, Newark, Delaware
(Received 30 November 1964)

The time-symmetrical interaction of charges (Wheeler–Feynman electrodynamics) is shown in principle to yield second-order Newtonian-type equations of motion under the restriction that the motions be analytic extensions of free-particle motions. The means for explicitly generating the electrodynamic equations of motion describing invariant world-lines are given. These physically relevant equations do not fit into Dirac's (1949) formulation of Hamiltonian relativistic particle dynamics, where either world-line invariance is given up, or only trivially straight world lines can be described ("zero-interaction theorem" of Currie, Jordan, Sudarshan, 1963). The misfit is due to the requirement in Dirac's scheme that position x be canonical. Under the Lie–Königs theorem, however, Hamiltonian statements of dynamics with invariant world lines remain possible when suitable $Q(x, \dot{x})$ are introduced instead of x as canonical position variables. A group of necessary conditions on the structure of any dynamics that permits x to be canonical are worked out to indicate how stringent is this permission in general.

I. INTRODUCTION

FOLLOWING Dirac's formulation[1] of the general conditions for an Hamiltonian statement of relativistic particle dynamics, it was recognized by Thomas[2] that the dynamics will not in general admit a description of particle motions in terms of invariant world lines.

Thomas[2] and later Foldy[3] have shown explicitly how to construct Dirac's ten fundamental dynamical variables that generate infinitesimal time and space translations, space rotations, and pure Lorentz transformations in accordance with the required symmetry of the dynamics under the full group of inhomogeneous Lorentz transformations; and Currie, Jordan, and Sudarshan have underscored Thomas' point by proving a "zero-interaction theorem" to the effect that (at least in the case of two or three particles) world-line invariance, plus Hamiltonian description of motion, demands free-particle motion only.[4]

This species of misfit between relativity requirements and Hamiltonian requirements must be reckoned a significant one, because on the one hand world-line invariance is so simple and compelling a feature of the relativistic view of nature, saying merely that one and the same particle orbit should indeed be an orbit no matter which inertial observer considers it, and on the other hand because an Hamiltonian rendition of motion stands at the root of the quantal understanding of motion; the denial of world-line invariance may well be a denial of an

[1] P. A. M. Dirac, Rev. Mod. Phys. 21, 392, (1949).
[2] L. H. Thomas, Phys. Rev. 85, 868 (1952); B. Bakamjian and L. H. Thomas, *ibid.* 92, 1300 (1953); B. Bakamjian, *ibid.* 121, 1849 (1961).
[3] L. L. Foldy, Phys. Rev. 122, 275 (1961).
[4] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963); D. G. Currie, J. Math. Phys. 4, 1470 (1963); J. T. Cannon and T. F. Jordan, *ibid.* 5, 299 (1964). For n particles, it appears that the zero-interaction theorem has been proven by H. Leutwyler, to be published in Helv. Phys. Acta. I thank the referee for bringing the latter to my notice.

understanding of the most physically consequential particle interactions. To the field theorist the problem has perhaps no standing at all, since the premise of the problem is that interacting particles are described purely in terms of particle variables essentially on a direct action-at-a-distance basis. But in at least one case (electrodynamics) this premise is completely tenable and moreover makes a long step in obviating the inconsistencies and incompletenesses that have marked field theory since its inception; so that inquiring into the Lorentzian–Hamiltonian incompatibility is physically relevant, albeit by a circuitous route, even for field theory.

The object of the present note is to show that the incompatibility can be traced to the assumption that particle positions may be taken to be canonical coordinates in an Hamiltonian scheme for the particle motions; in other words, responding to the query in the title, *that both world-line invariance and an Hamiltonian description of motion of interacting particles are quite compatible so long as an a priori identification of positions as canonical coordinates is eschewed.* That one may not count position among the canonical variables of course raises appreciable new problems in any program of quantization, but the knowledge of the existence of *some* canonical formalism may at least open the way to dealing with the problems on familiar ground. It is perhaps not too great a price to pay for keeping invariant world-lines.

The march of the argument is as follows: (a) We shall work first of all within the frame of an action-at-a-distance theory—Wheeler–Feynman[5] electrodynamics—where at the outset the theoretical structure embraces a description of motion in terms of particle variables only, and thereby has the right footing physically for entertaining the question of an Hamiltonian scheme; and where Lorentz covariance of the theory, including world-line invariance, is built in and can, so to say, be left to shift for itself. (b) Next we show that the primitive equations of motion of interacting charges are formally reducible to Newtonian order (second-order equations for each Cartesian coordinate of each particle), and shall produce in an approximation of the latter, an example of ordinary Newtonian equations generating invariant world-lines. (c) We then invoke the Lie–Königs[6] theorem, giving a simple direct proof of it, to tell us that the Newtonian-order equations have Hamiltonian representations. (d) Finally we use the zero-interaction theorem to tell that among these representations there is none in which particle positions are canonical variables—for the motions being described are invariant, but *curved,* world-lines. This indirect proof, using the zero-interaction theorem, will be substantiated by a more direct plausibility argument, depending upon severe integrability conditions for constructing an Hamiltonian containing positions as canonical variables.

[5] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).

[6] E. T. Whittaker, *Analytical Dynamics* (Cambridge University Press, Cambridge, England, 1937).

II. EQUATIONS OF MOTION FOR INTERACTING CHARGES

It will suffice to consider just two charges interacting via the half-retarded + half-advanced fields produced by each at the position of the other, as in our earlier discussions[7] of Wheeler–Feynman theory. Self-interaction is explicitly ruled out, and the time-symmetrical interactions of the charges not only with each other but with the rest of the electrodynamic universe (viz. the Wheeler–Feynman absorber) are then comprehending the conventional scheme of purely retarded interactions of charges together with the Lorentz–Dirac force of radiation damping for each.[8]

Using a Taylor expansion about present time t of the Lienard–Wiechert potentials due to e_2 at $\mathbf{r}_2(t)$ at the position $\mathbf{r}_1(t)$ of e_1, we can write the Lagrangian[7]

$$L_1(\mathbf{r}_1, \mathbf{v}_1, t) = -m_1c^2(1 - \mathbf{v}_1^2/c^2)^{\frac{1}{2}} - e_1e_2 \sum_{p=0}^{\infty} \frac{D_2^{2p}}{2p!\, c^{2p}} \left(1 - \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{c^2}\right) r^{2p-1} \qquad (1)$$

for e_1's motion, that of e_2 being thought of as prescribed. (c is not placed equal to unity because it will be used as an ordering parameter later). Here $r = |\mathbf{r}_1 - \mathbf{r}_2|$ and the operator D_2 is a time differentiation of particle-2 variables only:

$$D_2^{2p}\left(1 - \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{c^2}\right) r^{2p-1} = \left\{\left(\frac{d}{dt'}\right)^{2p}\left(1 - \frac{\mathbf{v}_1 \cdot \mathbf{v}_2(t')}{c^2}\right) |\mathbf{r}_1 - \mathbf{r}_2(t')|^{2p-1}\right\}_{t'=t}.$$

This merely translates the fields at $\mathbf{r}_1(t)$, coming from the right advanced and retarded world points on e_2's trajectory, into the dynamical language of e_2's

[7] E. H. Kerner, J. Math. Phys. **3**, 35 (1962); Phys. Rev. **125**, 2184 (1962).

[8] It should be pointed out that for what follows we could just as well use the conventional scheme, insofar as the object is simply to obtain Newtonian-order equations of motion giving curved invariant world lines. This is achievable on the conventional basis of electrodynamics, but is simpler in the Wheeler–Feynman scheme which is any way the more complete and consistent one, and admits conservation laws.

present motion; it is assumed for the moment that the motion is regular enough to permit this. The mark of the time symmetry is in the evenness of the powers of D_2. The reason for concentrating on the single present time is of course that we must give the time a Newtonian role in aiming toward an Hamiltonian theory (Dirac's instant form of the theory) that looks to the evolution of the particle motions differentially in any inertial observer's frame; the two separate proper times for the particles have to be set aside, though we shall have an auxiliary use for them.

For the motion of e_2 in e_1's fields we find similarly a Lagrangian $L_2(\mathbf{r}_2, \mathbf{v}_2, t)$ by reversing the roles of 1 and 2 above. A joint Lagrangian comprising both motions simultaneously, which formed the basis for our earlier discussions,[7] can also be constructed by observing that the operator of total differention $D = d/dt$ is identically $D_1 + D_2$ so that

$$D_2^{2p} = D_2^p(D - D_1)^p$$
$$= (-D_2D_1)^p + \text{exact derivatives.}$$

The exact derivatives may be dropped in Eq. (1) leaving a symmetric interaction term (D_1 and D_2 commute) whereupon, after symmetrizing also the first term in L_1, a joint Lagrangian is ($\epsilon \equiv e_1e_2$)

$$L = -m_1c^2\left(1 - \frac{\mathbf{v}_1^2}{c^2}\right)^{\frac{1}{2}} - m_2c^2\left(1 - \frac{\mathbf{v}_2^2}{c^2}\right)^{\frac{1}{2}}$$
$$- \epsilon \sum \frac{(-D_1D_2)^2}{2p!\,c^{2p}}\left(1 - \frac{\mathbf{v}_1\cdot\mathbf{v}_2}{c^2}\right)r^{2p-1}. \quad (2)$$

The equations of motion from the private Lagrangians L_i,

$$\partial L_i/\partial \mathbf{r}_i - D\,\partial L_i/\partial \mathbf{v}_i = 0, \quad (3)$$

are not different, except for a rearrangement, from those coming from L,

$$\frac{\partial L}{\partial \mathbf{r}_i} - D\frac{\partial L}{\partial \mathbf{v}_i} + D^2\frac{\partial L}{\partial \dot{\mathbf{v}}_i} - \cdots = 0. \quad (4)$$

These infinite-order equations of motion are plainly Lorentz-invariant and are describing invariant world-lines. For, from the private-Lagrangian point of view we have the individual action principles

$$\delta\int\left\{-m_ic^2\left(1 - \frac{\mathbf{v}_i^2}{c^2}\right)^{\frac{1}{2}} + \frac{e_i}{c}\,\mathbf{v}_i\cdot\mathbf{A}^i(\mathbf{r}_i, [\mathbf{r}_j(t)])\right.$$
$$\left. - e_i\phi^i(\mathbf{r}_i, [\mathbf{r}_j(t)])\right\}dt = 0, \quad (5)$$

($i = 1, 2$; $j = 2, 1$) where the potentials with argument $[\mathbf{r}_j(t)]$ are indicating that they are functionals of the motion of e_j whatever this motion may be. Therefore, in this view point, e_i's motion can be discussed covariantly through introduction of the proper time $d\tau_i = dt(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}}$ and writing

$$\delta\int\left\{\tfrac{1}{2}m_i\mu_\nu^i\mu_\nu^i + \frac{e_i}{c}\,\mu_\nu^iA_\nu^i(\mathbf{r}_i, [\mathbf{r}_j(t(\tau_i))])\right\}d\tau_i = 0 \quad (6)$$

($i = 1, 2$; $j = 2, 1$; summation over ν only; μ_ν^i = 4-velocity of e_i; A_ν^i = 4-potential at e_i due to e_j).

Though formally the joint Lagrangian can be used to build up an Hamiltonian theory (Ostrogradsky's method[6]), comprising an infinite set of canonical coordinates and momenta to cope with the infinite-order of the equations of motion, this is not relevant to the question whether there can be an Hamiltonian scheme for interacting particles employing exactly the Newtonian quota of *three* coordinates and *three* momenta per particle. If there is such a scheme, it is implied that the primitive equations of motion are of Newtonian-(second-) order; and one suspects that there might therefore be some kind of reduced, finite form of the electrodynamic joint Lagrangian L. This possibility however,[9] on a direct attempt at order reduction, is at best fraught with difficulties.

Alternatively it suggests itself to examine directly the primitive equations of motion and to ask in what sense they might be reducible to finite-order equations. Writing Eq. (3) in extended form, we have

$$D\frac{m_i\mathbf{v}_i}{(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}}} = m_i\dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \mathbf{v}_i}\frac{\mathbf{v}_i}{(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}}}$$
$$= \left\{\frac{(1 - \mathbf{v}_i^2/c^2)I + \mathbf{v}_i\mathbf{v}_i/c^2}{(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}}}\right\}\cdot m_i\dot{\mathbf{v}}_i$$
$$= -\epsilon\left(\frac{\partial}{\partial \mathbf{r}_i} - D\frac{\partial}{\partial \mathbf{v}_i}\right)\sum\frac{D_j^{2p}}{2p!c^{2p}}(1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2)\mathbf{r}^{2p-1}.$$

The inverse of the dyadic $\{\ \}$ is $(I - \mathbf{v}_i\mathbf{v}_i/c^2)(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}} \equiv \Lambda_i$ so we obtain

---

[9] In effect we attempted to investigate this in earlier work,[7] constructing from $L(\mathbf{r}_i, \dot{\mathbf{r}}_i, \ddot{\mathbf{r}}_i, \cdots)$ a unique energy integral $E(\mathbf{r}_i, \dot{\mathbf{r}}_i, \ddot{\mathbf{r}}_i, \cdots)$ and proposing the algorithm $H(\mathbf{r}_i, \mathbf{p}_i) = E = E(\mathbf{r}_i, (\mathbf{r}_i, H), ((\mathbf{r}_i, H), H), \cdots)$ for the computation of H in powers of ϵ and of $1/c^2$. The algorithm seems clearly a *necessary* condition for H, but the proof of its *sufficiency* (i. e., of its power to provide H that generates motions in agreement with the primitive equations of motion) has remained undemonstrable. A somewhat similar attempt, working only up to a maximum of $1/c^4$ terms in a joint Lagrangian based on purely retarded interactions and reducing the Lagrangian directly, was made by Golubenkov and Smorodinski [reported in L. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1951)]; but a direct calculation shows that the primitive equations of motion are not properly comprehended.

$$\dot{\mathbf{v}}_1 = \frac{\epsilon}{m_1}\Lambda_1\left(D\frac{\partial}{\partial \mathbf{v}_1} - \frac{\partial}{\partial \mathbf{r}_1}\right) \times \sum \frac{D_2^{2p}}{2p!c^{2p}}(1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2)r^{2p-1} \equiv \frac{\epsilon}{m_1}\mathbf{F}_1,$$
$$\dot{\mathbf{v}}_2 = \frac{\epsilon}{m_2}\Lambda_2\left(D\frac{\partial}{\partial \mathbf{v}_2} - \frac{\partial}{\partial \mathbf{r}_2}\right) \times \sum \frac{D_1^{2p}}{2p!c^{2p}}(1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2)r^{2p-1} \equiv \frac{\epsilon}{m_2}\mathbf{F}_2, \tag{7}$$

as the complete formal dynamics of a pair of inter-acting charges.

Continuing formally we can reduce Eq. (7) to Newtonian order by observing that the higher-order derivatives in $\mathbf{F}_i$ can be displayed as

$$\frac{\partial}{\partial \mathbf{r}_j}D_i^{2p} = \left(\mathbf{v}_i^{(2p)}\cdot\frac{\partial}{\partial \mathbf{v}_i^{(2p-1)}} + \cdots + \dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \mathbf{v}_i} + \mathbf{v}_i\cdot\frac{\partial}{\partial \mathbf{r}_i}\right)\cdots\left(\ddot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \dot{\mathbf{v}}_i} + \dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \mathbf{v}_i} + \mathbf{v}_i\cdot\frac{\partial}{\partial \mathbf{r}_i}\right)\left(\dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \mathbf{v}_i} + \mathbf{v}_i\cdot\frac{\partial}{\partial \mathbf{r}_i}\right)\frac{\partial}{\partial \mathbf{r}_j},$$
$$D\frac{\partial}{\partial \mathbf{v}_j}D_i^{2p} = \left(\mathbf{v}_i^{(2p+1)}\cdot\frac{\partial}{\partial \mathbf{v}_i^{(2p)}} + \cdots + \dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial \mathbf{v}_i} + \mathbf{v}_i\cdot\frac{\partial}{\partial \mathbf{r}_i} + \mathbf{v}_j\cdot\frac{\partial}{\partial \mathbf{r}_j}\right)D_i^{2p}\frac{\partial}{\partial \mathbf{v}_j},$$

and by then introducing into $\dot{\mathbf{v}}_i = (\epsilon/m_i)\mathbf{F}_i(\dot{\mathbf{v}}_i, \ddot{\mathbf{v}}_i, \cdots)$ the derivatives $\ddot{\mathbf{v}}_i = (\epsilon/m_i)\mathbf{F}_i$, $\dddot{\mathbf{v}}_i = D(\epsilon/m_i)\mathbf{F}_i, \cdots$ and ordering terms according to powers of ϵ.

Let us see in the opening terms how this proceeds. In $(\epsilon/m_i)\mathbf{F}_i$ to first order in ϵ we must retain only the part $(\mathbf{v}_i\cdot\partial/\partial\mathbf{r}_i)^{2p}$ of D_i^{2p} because terms in $\dot{\mathbf{v}}_i$, $\ddot{\mathbf{v}}_i$, $\cdots$ will be proportional to ϵ, owing to $\dot{\mathbf{v}}_i = (\epsilon/m_i)\mathbf{F}_i$. It is easy to show that

$$(\mathbf{v}_i\cdot\partial/\partial\mathbf{r}_i)^{2p}r^{2p-1}$$
$$= (2p-1)^2(2p-3)^2\cdots 1^2(\mathbf{v}_i\times\hat{r})^{2p}/r, \qquad (\hat{r}\equiv\mathbf{r}/r),$$

and thence

$$\sum_p \frac{(\mathbf{v}_i\cdot\partial/\partial\mathbf{r}_i)^{2p}}{2p!c^{2p}}(1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2)r^{2p-1} = \frac{1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2}{r(1 - (\mathbf{v}_i\times\hat{r})^2/c^2)^{\frac{1}{2}}},$$

whereupon in first approximation

$$\dot{\mathbf{v}}_1 = \frac{\epsilon}{m_1}(I - \mathbf{v}_1\mathbf{v}_1/c^2)(1 - \mathbf{v}_1^2/c^2)^{\frac{1}{2}}\left(D\frac{\partial}{\partial \mathbf{v}_1} - \frac{\partial}{\partial \mathbf{r}_1}\right) \times \frac{1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2}{r(1 - (\mathbf{v}_2\times\hat{r})^2/c^2)^{\frac{1}{2}}} \equiv \frac{\epsilon}{m_1}\mathbf{F}_{10}(\mathbf{r}_k, \mathbf{v}_k),$$
$$\dot{\mathbf{v}}_2 = \frac{\epsilon}{m_2}\left(I - \frac{\mathbf{v}_2\mathbf{v}_2}{c^2}\right)\left(1 - \frac{\mathbf{v}_2^2}{c^2}\right)^{\frac{1}{2}}\left(D\frac{\partial}{\partial \mathbf{v}_2} - \frac{\partial}{\partial \mathbf{r}_2}\right) \times \frac{1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2}{r(1 - (\mathbf{v}_1\times\hat{r})^2/c^2)^{\frac{1}{2}}} \equiv \frac{\epsilon}{m_2}\mathbf{F}_{20}(\mathbf{r}_k, \mathbf{v}_k). \tag{8}$$

The physical meaning is transparent: we have merely reckoned the 4-potential $(\mathbf{A}^i, \Phi)$ in Eq. (3) for e_i's motion due to fields from e_j as

$$(\mathbf{a}^i, \phi) = \left(\frac{e_j\mathbf{v}_j/c}{r[1 - (\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}}, \frac{e_j}{r[1 - (\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}}\right), \tag{9}$$

viz., as fields from e_j in *uniform* motion,[10] $\dot{\mathbf{v}}_j = 0$.

[10] See for instance M. Abraham and R. Becker, *Theorie der Elektrizität II* (B. G. Teubner, Leipzig and Berlin, 1933).

The forces in Eq. (8) are simply Lorentz forces on each charge due to the other charge's uniform motion,

$$\epsilon\left(D\frac{\partial}{\partial \mathbf{v}_i} - \frac{\partial}{\partial \mathbf{r}_i}\right)\frac{1 - \mathbf{v}_1\cdot\mathbf{v}_2/c^2}{r[1 - (\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}} = e_i\left\{\mathbf{E}\ (\text{at } i, \text{ from } j) + \frac{\mathbf{v}_i}{c}\times\mathbf{H}\ (\text{at } i, \text{ from } j)\right\},$$

where D is $\mathbf{v}_1\cdot\partial/\partial\mathbf{r}_1 + \mathbf{v}_2\cdot\partial/\partial\mathbf{r}_2$ in the present order of approximation.

To go to the next order in ϵ^2 we need D_i^{2p}, D_i^{2p+1} up to terms linear in $\dot{\mathbf{v}}_i$, $\ddot{\mathbf{v}}_i$, $\cdots$; a computation by induction gives, abbreviating $d_i \equiv \mathbf{v}_i\cdot\partial/\partial\mathbf{r}_i$,

$$D_i^n = d_i^n + \sum_{l=0}^{n-1}\mathbf{v}_i^{(l+1)}\cdot\mathbf{J}_l^n(d_i)$$
$$\mathbf{J}_l^n(d_i) = \binom{n-1}{l}d_i^{n-l-1}\frac{\partial}{\partial \mathbf{v}_i} + \binom{n-2}{l}d_i^{n-l-2}\frac{\partial}{\partial \mathbf{v}_i}d_i + \cdots + \frac{\partial}{\partial \mathbf{v}_i}d_i^{n-l-1}.$$

Then the equations of motion become

$$\dot{\mathbf{v}}_1 = \frac{\epsilon}{m_1}\mathbf{F}_{10}(\mathbf{r}_k, \mathbf{v}_k) + \frac{\epsilon^2}{m_1m_2}\Lambda_1\left(D\frac{\partial}{\partial \mathbf{v}_1} - \frac{\partial}{\partial \mathbf{r}_1}\right) \times \sum_{p=0}^{\infty}\sum_{l=0}^{2p-1}(d^{l+1}\mathbf{F}_{20})\cdot J_l^{2p}(d_2)\left(1 - \frac{\mathbf{v}_1\cdot\mathbf{v}_2}{c^2}\right)r^{2p-1},$$
$$\dot{\mathbf{v}}_2 = \frac{\epsilon}{m_2}\mathbf{F}_{20}(\mathbf{r}_k, \mathbf{v}_k) = \frac{\epsilon^2}{m_1m_2}\Lambda_2\left(D\frac{\partial}{\partial \mathbf{v}_2} - \frac{\partial}{\partial \mathbf{r}_2}\right) \times \sum_{p=0}^{\infty}\sum_{l=0}^{2p-1}(d^{l+1}\mathbf{F}_{10})\cdot J_l^{2p}(d_1)\left(1 - \frac{\mathbf{v}_1\cdot\mathbf{v}_2}{c^2}\right)r^{2p-1},$$

in which the right-hand members depend only on $\mathbf{r}_k$, $\mathbf{v}_k$, and $d \equiv d_1 + d_2$. The reduction process restates the high-order differentiation in terms of just the elementary d_i. Continuing in this cumbersome but stepwise well-defined fashion, we come formally to express electrodynamics as a Newtonian system

$$\dot{\mathbf{v}}_i = \sum_m \epsilon^m\mathbf{f}_{im}(\mathbf{r}_k, \mathbf{v}_k),$$

where already in the opening term curved invariant world-lines $\dot{\mathbf{v}}_i = (\epsilon/m_i)\mathbf{F}_{i0}$ are being generated out of the initially straight ones $\dot{\mathbf{v}}_i = 0$.

This "solving" of the higher-order equations of motion by expansion in powers of a parameter attached to the higher derivatives amounts to seeking out of the large set of motions, parametrized by a large number of initial conditions $\mathbf{r}_k(0)$, $\dot{\mathbf{r}}_k(0)$, $\ddot{\mathbf{r}}_k(0)$, $\cdots$, some very particular small subset parametrized by just $\mathbf{r}_k(0)$, $\dot{\mathbf{r}}_k(0)$, whose distinguishing feature is their analytic contiguity to free-particle motions ($\epsilon \to 0$). This contiguity seems physically compelling; if it is adopted as a physical requirement, then electrodynamics on its original field-theoretical basis [or its action-at-a-distance translation via Eq. (4)] must be counted grossly overcomplete. The high derivatives are being taken to signify "corrections" to lower-order motions, instead of being allowed the power to generate new classes of motions. The situation is not unlike that occurring for ordinary radiatively damped motion, e.g., $-\frac{2}{3}(e^2/c^3)\dddot{x} + m\ddot{x} = -\partial V/\partial x$, where only two-thirds of the motions are admissable and one-third of them are improper, being marked by essential singular behavior as $e^2 \to 0$; not until a rule is given for the right motion and placed in force theoretically can such an equation of motion be considered satisfactory.

Nothing can be said at this point about convergence of the procedure we have sketched. The primary aim has been just to show that it is not ruled out that physically significant particle interactions, perhaps the only significant ones, may be fully comprehensible in covariant Newtonian-order equations of motion, giving invariant world-lines, despite the traditional bias of field physics that an infinite number of degrees of freedom is inherent in the problem; and particularly to show in a concrete case, Eqs. (8), how such equations of motion get developed for purposes of discussing in how far an Hamiltonian statement of them is possible—it being immaterial to this purpose that Eqs. (8) are only approximative to the whole electrodynamics.

Before turning to the main question of Hamiltonians, a few sidepoints may be noticed.

(i) Another formal way to develop second-order equations of motion out of Eqs. (7) is to order the right-hand members by powers of $1/c^2$,

$$\dot{\mathbf{v}}_i = \sum_{n=0} \frac{1}{c^{2n}} \mathbf{g}_{in}(\mathbf{r}_i, \mathbf{v}_i; \mathbf{r}_j, \dot{\mathbf{r}}_j, \cdots, \mathbf{r}_j^{(2n)});$$

$$\mathbf{g}_{i0} = \frac{\epsilon}{m_i} \frac{\partial}{\partial \mathbf{r}_i}\left(\frac{1}{r}\right),$$

and to expand $\dot{\mathbf{v}}_i$ as $\sum \gamma_{in}(\mathbf{r}_k, \mathbf{v}_k)/c^{2n}$, beginning with γ_{i0} = the Coulomb interaction, $\mathbf{g}_{i0}$, and computing the higher derivatives needed in $\mathbf{g}_{in}$ from derivatives of the already determined $\dot{\mathbf{v}}_i = \sum_0^{n-1} \gamma_{is}(\mathbf{r}_k, \mathbf{v}_k)/c^{2s}$. The opening terms are, for instance

$$\gamma_{i0} = \frac{\epsilon}{m_i} \frac{(\mathbf{r}_i - \mathbf{r}_j)}{r^3},$$

$$\gamma_{i1} = \frac{\epsilon^2}{m_1 m_2} \frac{(\mathbf{r}_i - \mathbf{r}_j)}{r^4}$$

$$- \frac{\epsilon}{m_i}\left\{\frac{\mathbf{v}_i \cdot (\mathbf{r}_i - \mathbf{r}_j)}{r^3}(\mathbf{v}_i - \mathbf{v}_j) + \frac{1}{2}\frac{(\mathbf{v}_i^2 - \mathbf{v}_j^2)}{r^3}(\mathbf{r}_i - \mathbf{r}_j)\right.$$

$$\left. + \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{r^3}(\mathbf{r}_i - \mathbf{r}_j) + \frac{3}{2}\frac{(\mathbf{v}_j \cdot \mathbf{r})^2}{r^5}(\mathbf{r}_i - \mathbf{r}_j)\right\}.$$

It is now the motions that are contiguous, for $c \to \infty$, to the "unperturbed" Kepler motions, instead of to free-particle motions, that are being built up. Lorentz covariance is sacrificed in the individual terms, but this is compensated to some extent by their relative ease of computation and by their containing all orders of the interaction strength ϵ (γ_{in} = polynomial in ϵ). The expansion plainly exhibits in fact the explicit entry of the charge radii $e_i^2/m_i c^2$ into the dynamics, hinting at how convergence of the series may be controlled, in contrast to the position in field theory where this fundamental length is quickly seen to exist and then sits uneasily in the background playing some important but unknown role. It is interesting in this connection to note the distinction that can be made between the *non-relativistic limit* ($c \to \infty$) of electrodynamics, governed completely by the Coulomb interaction, and the *static limit* ($\mathbf{v}_i \to 0$), governed by

$$\dot{\mathbf{v}}_i = \frac{\epsilon(\mathbf{r}_i - \mathbf{r}_j)}{m_i r^3} + \frac{\epsilon^2}{m_1 m_2 c^2}\frac{(\mathbf{r}_i - \mathbf{r}_j)}{r^4} + \cdots.$$

(ii) It is visible from Eqs. (8) that the occurrence of both $(1 - \mathbf{v}_i^2/c^2)^{\frac{1}{2}}$ and $(1 - (\mathbf{v}_j \times \hat{r})^2/c^2)^{\frac{1}{2}}$ in $\mathbf{F}_{i0}$ will not be altered in the development of further terms in the ϵ-expansion, since these grow out of the rational operation of differentiation. There is therefore no reason of principle to discountenance orbits for which both $|\mathbf{v}_1| > c$ and $|\mathbf{v}_2| > c$. Such orbits are plainly disjoint, classically, from orbits with $|\mathbf{v}_i| < c$; if initially $|\mathbf{v}_i(0)| < c$, then apparently at no later time can $|\mathbf{v}_i(t)| > c$ and conversely—the light barrier $|\mathbf{v}_i| = c$ is seemingly impassable. In quantum theory the situation can be possibly expected to be otherwise. For quantally, an entire ensemble of classical motions must always be sur-

veyed, owing to the finiteness of Planck's constant h, which defines an irreducible volume in phase space whose boundaries may not in general be set ahead of time to exclude a classically forbidden region; this is at any rate what lies back of the so characteristically quantum phenomenon of barrier penetration in ordinary nonrelativistic potential problems, where classically disjoint motions necessarily have to be looked at together and be allowed to overlap. It is just the intimacy of the codetermination of motions of only the sources of the field afforded in the action-at-a-distance view, that brings into evidence the extended class of motions hidden in field theory behind the mathematically neat but physically clumsy conceptualization of sources emitting and absorbing fields. The meaning, if any, of the extended set of motions is difficult to assess before a program of Hamiltonization and quantization is formulated, but already at the outset there appears from Eq. (2) a curious and probably essential ambiguity of any fundamental time-development operator as to its Hermiticity; to span all possible motions the operator would seem to have to be capable of sometimes being Hermitian and sometimes anti-Hermitian, or of possessing in some fashion an eigenspectrum with both pure-real and pure-imaginary sectors.

(iii) We note finally now a very simple and natural way to fit gravitation and electrodynamics together, similar to that previously suggested[7] in the context to Eq. (2). This amounts to the recognition that the physically relevant motions left over when $\epsilon \to 0$ are not the simple unaccelerated motions used in Eq. (8), but the motions of gravitating mass points in the general relativistic sense of Einstein, Infeld, and Hoffman (EIH).[11] To thus accredit gravitation as the fundament of all other motions it suffices, as a physical hypothesis, to place

$$\dot{\mathbf{v}}_1 = \mathfrak{F}_1(\mathbf{r}_k, \mathbf{v}_k) + \frac{\epsilon}{m_1}\mathbf{F}_1,$$

$$\dot{\mathbf{v}}_2 = \mathfrak{F}_2(r_k, v_k) + \frac{\epsilon}{m_2}\mathbf{F}_2,$$

where (a) $\mathfrak{F}_i$ are the gravitational accelerations computed from the EIH action-at-a-distance Lagrangian

$$\mathfrak{L} = \tfrac{1}{2}m_1\mathbf{v}_1^2 + \tfrac{1}{8}m_1\frac{\mathbf{v}_1^4}{c^2} + \cdots + \tfrac{1}{2}m_2\mathbf{v}_2^2 + \tfrac{1}{8}m_2\frac{\mathbf{v}_2^4}{c^2} + \cdots + \frac{\gamma m_1 m_2}{r} + \frac{\gamma m_1 m_2}{2c^2 r}\left(3\mathbf{v}_1^2 + 3\mathbf{v}_2^2 - 7\mathbf{v}_1\cdot\mathbf{v}_2 - \frac{\mathbf{v}_1\cdot\mathbf{r}\,\mathbf{v}_2\cdot\mathbf{r}}{r^2}\right) - \frac{\gamma^2}{2c^2}\frac{m_1 m_2(m_1+m_2)}{r^2} + \cdots,$$

(γ = constant of gravitation); and (b) the electrodynamic derivatives in $\mathbf{F}_i$, D_i^{2p}, are computed over the gravitational motion as "unperturbed motion." The final equations of motion are series of powers in γ and in ϵ containing in the mixed terms the coupling of the two physical phenomena. The hypothesis joining gravitation and electrodynamics is therefore not so much a simple addition of gravitational and electromagnetic effects (in any case compulsory in nonrelativistic limit) as it is a complete specification of the latter through the primacy of the former to tell what is the nature of motion in the first place; the electrodynamics, in short, sits as a contingency on any prior description of the motion of uncharged particles.

III. THE LIE-KÖNIGS THEOREM

We now take from Sec. II the example, Eqs. (8), of equations of motion describing relativistic forces up to order ϵ. A further characteristic ϵ-structure which must eventually enter the forces can be seen by reckoning D as the full derivative $\dot{\mathbf{v}}_i\cdot\partial/\partial\mathbf{v}_i + d$ in $\mathbf{F}_{i0}$ instead of the truncated d used previously in securing the expansion in ϵ. [In fact all terms in $\dot{\mathbf{v}}_j$ occurring in the differentiations in $\mathbf{F}_i$ in Eq. (7) should properly be separated out and considered to be inertia-like on the same footing as $m_i\dot{\mathbf{v}}_i$.] The equations of motion are then

$$\dot{\mathbf{v}}_i = -\frac{\epsilon}{m_i}\Lambda_i\left\{\left(\dot{\mathbf{v}}_i\cdot\frac{\partial}{\partial\mathbf{v}_i} + d\right)\frac{\mathbf{v}_j/c^2}{r[1-(\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}} + \frac{\partial}{\partial\mathbf{r}_i}\frac{(1-\mathbf{v}_1\cdot\mathbf{v}_2/c^2)}{r[1-(\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}}\right\}, \tag{10}$$

and stem from the variational principles, Eqs. (6), in which enter the four-vectors (9) with $\mathbf{v}_i = \mathbf{v}_i(t)$ instead of constant $\mathbf{v}_i$. It is just the use of separate covariant action principles that eventually, upon the full computation of potentials here only approximated, must guarantee invariance of the individual world-lines. Calling Ω_i the dyadic,

$$\Omega_i \equiv \frac{\partial}{\partial\mathbf{v}_i}\frac{\mathbf{v}_j}{r[1-(\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}} = \frac{(1-(\mathbf{v}_j\times\hat{r})^2/c^2)I - (\mathbf{v}_j\mathbf{v}_j - \mathbf{v}_j\cdot\hat{r}\mathbf{v}_j\hat{r})/c^2}{r[1-(\mathbf{v}_j\times\hat{r})^2/c^2]^{\frac{1}{2}}},$$

[11] L. Infeld and J. Plebanski, *Motion and Relativity* (Pergamon Press, Inc., New York, 1960); V. Fock, *The Theory of Space, Time, and Gravitation* (Pergamon Press, Inc., New York, 1959).

Eqs. (10) are

$$\dot{\mathbf{v}}_1 + (\epsilon/m_1c^2)\Lambda_1\cdot\tilde{\Omega}_2\cdot\dot{\mathbf{v}}_2 = (\epsilon/m_1)\mathbf{F}_{10}(\mathbf{r}_k, \mathbf{v}_k),$$

$$\dot{\mathbf{v}}_2 + (\epsilon/m_2c^2)\Lambda_2\cdot\tilde{\Omega}_1\cdot\dot{\mathbf{v}}_1 = (\epsilon/m_2)\mathbf{F}_{20}(\mathbf{r}_k, \mathbf{v}_k),$$

or in decoupled form,

$$\begin{aligned}\dot{\mathbf{v}}_1 &= \left(I - \frac{\epsilon^2}{m_1m_2c^4}\Lambda_1\cdot\tilde{\Omega}_2\cdot\Lambda_2\cdot\tilde{\Omega}_1\right)^{-1} \\ &\quad\times\left(\frac{\epsilon}{m_1}\mathbf{F}_{10} - \frac{\epsilon^2}{m_1m_2c^2}\Lambda_1\cdot\tilde{\Omega}_2\cdot\mathbf{F}_{20}\right), \\ \dot{\mathbf{v}}_2 &= \left(I - \frac{\epsilon^2}{m_1m_2c^4}\Lambda_2\cdot\tilde{\Omega}_1\cdot\Lambda_1\cdot\tilde{\Omega}_2\right)^{-1} \\ &\quad\times\left(\frac{\epsilon}{m_2}\mathbf{F}_{20} - \frac{\epsilon^2}{m_1m_2c^2}\Lambda_2\cdot\tilde{\Omega}_1\cdot\mathbf{F}_{10}\right).\end{aligned} \tag{11}$$

These forces are some physically significant parts of the whole Newtonian-reduced electrodynamic forces, and illustrate something of the order of complexity of the latter, well beyond that of the $\mathbf{F}_{i0}$.

Let us consider now quite generally the problem of reducing a system of first-order differential equations

$$\dot{x}_i = X_i(x_1, \cdots, x_n) \qquad (i = 1, \cdots, n) \tag{12}$$

to Hamiltonian form (we use x_i for the time being for general variables, not to be confused with dynamical coordinates). First, we can force Eqs. (12) to result from a variational principle

$$\delta\int_{t_1}^{t_2} J(x_i, \dot{x}_i)\,dt = 0$$

only if J is linear in the $\dot{x}_i$,

$$J = \sum U_i(x)\dot{x}_i - W(x), \tag{13}$$

requiring then

$$\sum_l\left(\frac{\partial U_k}{\partial x_l} - \frac{\partial U_l}{\partial x_k}\right)\dot{x}_l + \frac{\partial W}{\partial x_k} = 0.$$

To invert for $\dot{x}_l$ and equate to X_l, we must have the skew-symmetric $U_{k,l} - U_{l,k}$ nonsingular, which rules out odd-order systems. Then for n even, assuming $U_{k,l} - U_{l,k}$ is not "accidentally" singular, U_k and W have to satisfy

$$\sum_l X_l\frac{\partial U_k}{\partial x_l} - X_l\frac{\partial U_l}{\partial x_k} = -\frac{\partial W}{\partial x_k}. \tag{14}$$

Writing

$$X_l\frac{\partial U_l}{\partial x_k} = \frac{\partial}{\partial x_k}(U_lX_l) - U_l\frac{\partial X_l}{\partial x_k},$$

the last step is, in a transparent notation,

$$\mathbf{X}\cdot\nabla_x\mathbf{U} + \tilde{X}\cdot\mathbf{U} = \nabla_x(\mathbf{X}\cdot\mathbf{U} - W),$$

$$(\tilde{X})_{ij} \equiv \partial X_i/\partial x_i, \qquad \mathbf{U} \equiv (U_1, \cdots, U_n).$$

The determination of W, $\mathbf{U}$ can be effected separately by requiring $W = \mathbf{X}\cdot\mathbf{U}$ within an additive term having vanishing gradient, when $\mathbf{U}$ is fixed by

$$\mathbf{X}\cdot\nabla_x\mathbf{U} + \tilde{X}\cdot\mathbf{U} = 0.$$

This is the usual condition[6] for $\int \mathbf{U}\cdot\delta\mathbf{x}$ to be an integral invariant (time independent) for $\dot{\mathbf{x}} = \mathbf{X}$. Solving for $\partial U_k/\partial x_1$, we obtain the Kowalewskian system of partial differential equations for $\mathbf{U}$,

$$\partial U_k/\partial x_1 = -X_1^{-1}\{(\mathbf{X}\cdot\nabla)'U_k + \tilde{X}_{kl}U_l\}$$

$$((\mathbf{X}\cdot\nabla)' = \mathbf{X}\cdot\nabla - X_1\,\partial/\partial x_1),$$

which is completely integrable and in general will permit nonsingular $U_{k,l} - U_{l,k}$.

It is therefore always possible to have an action principle lie back of $\dot{\mathbf{x}} = \mathbf{X}$. When we place $n = 2m$, the action integral may be written

$$\int U_1\,dx_1 + \cdots + U_{2m}\,dx_{2m} - W\,dt.$$

And now we can seek independent functions $P_i(x)$, $Q_i(x)$ $(i = 1, \cdots, m)$ of the original variables that admit

$$U_1\,dx_1 + \cdots + U_{2m}\,dx_{2m} = P_1\,dQ_1 + \cdots + P_m\,dQ_m.$$

The reduction is always possible (Pfaff's problem[12]), placing the original variational principle and equations of motion in Hamiltonian form,

$$\delta\int\left(\sum_1^m P_i\dot{Q}_i - H\right)dt = 0,$$

$$H(P, Q) = W(x(P, Q)); \qquad \dot{Q}_i = \partial H/\partial P_i,$$

$$\dot{P}_i = -\partial H/\partial Q_i.$$

This constitutes the Lie–Königs Theorem.

A direct proof of the existence of P, Q, without relying on the theory of Pfaffians, is readily formulated by simply introducing a change of variables $x_i = \omega_i(y_1, \cdots, y_n)$ and writing J as

$$J = \sum_k\left(\sum_i U_i\,\partial\omega_i/\partial y_k\right)\dot{y}_k - W.$$

Then putting $\sum U_i\,\partial\omega_i/\partial y_k = y_{k+1}$ or 0 for odd k or even k (so that odd or even y are Q or P above), an analysis shows $y_i(x)$ to be determined from a Kowalewskian system.

For any Newtonian-order equations of motion like Eqs. (8) or (11) broken down as

[12] A. R. Forsyth, *Theory of Differential Equations* Dover Publications, Inc., New York, 1959), Vol. 1.

$$\dot{\mathbf{r}}_1 = \mathbf{v}_1, \qquad \dot{\mathbf{r}}_2 = \mathbf{v}_2,$$
$$\dot{\mathbf{v}}_1 = \mathbf{F}_1(\mathbf{r}_k, \mathbf{v}_k), \qquad \dot{\mathbf{v}}_2 = \mathbf{F}_2(\mathbf{r}_k, \mathbf{v}_k),$$

we are assured by the Lie–Königs theorem of a Hamiltonian representation. What we are not at all assured of is such a representation with a preselection of canonical coordinates Q. Suppose indeed we try to force $Q_s = x_s$, $s = 1 \cdots 6$, using the index s to enumerate particle coordinates x_1, y_1, z_1, x_2, y_2, z_2. This means in Eq. (13)

$$J = U_s(x, v)\dot{x}_s + 0\,\dot{v}_s - W(x, v),$$

with $U_s(x, v)$ the presumed canonical mate to x_s (summation on s understood), and from Eq. (14),

$$\left(v_s \frac{\partial}{\partial x_s} + F_s \frac{\partial}{\partial v_s}\right) U_k = \partial Z/\partial x_k,$$
$$U_k = \partial Z/\partial v_k, \qquad (Z \equiv v_s U_s - W), \tag{15}$$

or

$$\left(v_s \frac{\partial}{\partial x_s} + F_s \frac{\partial}{\partial v_s}\right) \frac{\partial Z}{\partial v_k} = \frac{\partial Z}{\partial x_k}. \tag{16}$$

That is, we are asking simply for the existence of a Lagrangian $Z(x, v)$ in back of the primitive equations of motion in order to guarantee a Hamiltonian formulation of them by what comes down to be the customary route.

We can now phrase the whole question of the Hamiltonian–Lorentzian compatibility in the following terms: when the F_s are restricted to providing Lorentz-invariant equations of motion $\dot{v}_s = F_s$ describing invariant world-lines, are the partial-differential systems (15) or (16)—necessary and sufficient for an Hamiltonian scheme with position a canonical variable,—integrable?

The integrability conditions, framed as differential conditions on F_s, fall within the algebraic theory of differential equations.[13] Equations (15) for instance are a so called orthonomic system when cotes are assigned in Riquier's way to the unknowns so as to order them U_s, Z and to the independent variables so as to order them v_s; x_s and the system is written

$$\partial U_k/\partial v_1 = * \qquad (k = 1, \cdots, 6),$$
$$\partial Z/\partial v_k = * \qquad (k = 1, \cdots, 6),$$

with right-hand members, *, involving only parametric derivatives. The explicit statement of all the integrability conditions requires appreciable analytical labor, but it quickly becomes apparent that they speak toward only highly particularized F_s—it can be doubted on sight (see below) whether anything like the complex F_s of Eq. (8) (not to say something like those of Eq. (11)) could satisfy them in a consistently ϵ-ordered computation. There is, above all, the zero-interaction theorem to tell without ambiguity that the integrability conditions cannot be met: there is no Hamiltonian theory of motion with position being canonical coordinate.

A closer view of what are at least necessary conditions for integrability may be obtained as follows: By Eqs (15) we have

$$\frac{\partial U_k}{\partial v_l} = \frac{\partial^2 Z}{\partial v_l\, \partial v_k} = \frac{\partial U_l}{\partial v_k},$$

or the matrix $(U)_{kl} \equiv \partial U_k/\partial v_l$ must be symmetric. Differentiating the first of Eqs. (15) with respect to v_l gives

$$U_k^l + F_{sl}U_{ks} + \Delta U_{kl} = U_l^k \tag{17}$$

in the notation

$$\Delta \equiv v_s\, \partial/\partial x_s + F_s\, \partial/\partial v_s,$$
$$U_k^l \equiv \partial U_k/\partial x_l, \qquad F_{sl} \equiv \partial F_s/\partial v_l$$

(that is, Δ is a time derivative with the motion $\dot{v}_s = F_s$ impressed into it). Reversing k and l and subtracting and also adding the result to Eq. (17) produces

$$F_{sl}U_{ks} - F_{sk}U_{ls} = 2(U_l^k - U_k^l),$$
$$\Delta U_{kl} = -\tfrac{1}{2}(F_{sl}U_{ks} + F_{sk}U_{ls}). \tag{18}$$

Now differentiating with respect to x_l,

$$\frac{\partial F_s}{\partial x_l}\frac{\partial U_k}{\partial v_s} + \Delta \frac{\partial U_k}{\partial x_l} = \frac{\partial^2 Z}{\partial x_l\, \partial x_k},$$

and reversing k and l and subtracting,

$$\Delta(U_l^k - U_k^l) = F_s^l U_{ks} - F_s^k U_{ls}$$
$$(F_s^l \equiv \partial F_s/\partial x_l). \tag{19}$$

We therefore find, using Eqs. (18) and (19),

$$\Delta(F_{sl}U_{ks} - F_{sk}U_{ls}) = 2(F_s^l U_{ks} - F_s^k U_{ls}).$$

The left-hand member here is, using Eq. (18) for ΔU_{kl},

$$F'_{sl}U_{ks} - \tfrac{1}{2}F_{sl}(F_{ts}U_{kt} + F_{tk}U_{st})$$
$$-F'_{sk}U_{ls} + \tfrac{1}{2}F_{sk}(F_{ts}U_{lt} + F_{tl}U_{st})$$

[13] J. F. Ritt, *Differential Equations from the Algebraic Standpoint* (American Mathematical Society Colloquium Publications, New York, 1932); J. M. Thomas, *Ann. Math.* **30**, 285 (1929); **35**, 306 (1934); C. Riquier, *Les Systèmes d'Equations aux Dérivées Partielles* (Gauthier-Villars, Paris, 1910). The integrability conditions can also be formulated from Helmholtz's theorem [P. Havas, Nuovo Cimento Suppl. **5**, 363 (1957)] on the conditions for any equations $G_i(x_i, \dot{x}_i, \ddot{x}_i) = 0$ to stem from a Lagrangian; but it is essential that G_i be taken as a linear combination $\beta_{is}(\ddot{x}_s - F_s(x, \dot{x}))$ with $\beta_{is} = \beta_{is}(x, \dot{x})$, $|\beta| \neq 0$, rather than any single $\ddot{x}_s - F_s$.

(summation on t as well as s; $F'_{sl} = \Delta F_{sl}$) in which the terms

$$\tfrac{1}{2}U_{st}(F_{sk}F_{tl} - F_{sl}F_{tk})$$

collectively contribute nothing, whence

$$(F'_{sl} - 2F^{l}_{s} - \tfrac{1}{2}F_{st}F_{tl})U_{ks} = (F'_{sk} - 2F^{k}_{s} - \tfrac{1}{2}F_{st}F_{tk})U_{ls}.$$

Defining now

$$(F)_{kl} = -\tfrac{1}{2}\,\partial F_k/\partial v_l, \qquad (G)_{kl} = -2\{\Delta(F)_{kl} + \partial F_k/\partial x_l + (F^2)_{kl}\}, \tag{20}$$

we have altogether the matrix equations

$$U = \tilde{U} \qquad (|U| \neq 0), \tag{21}$$

$$\Delta U \equiv U' = UF + \tilde{F}U, \tag{22}$$

$$GU = U\tilde{G}, \tag{23}$$

viz., U must be nonsingular and symmetric and simultaneously must satisfy the differential equation (22) and the algebraic equation (23).

To see what this entails, apply Δ to Eq. (23),

$$G'U + GU' = U'\tilde{G} + U\tilde{G}',$$

or by Eq. (22)

$$(G' + G\tilde{F} - \tilde{F}G)U = U(\tilde{G}' + F\tilde{G} - \tilde{G}F).$$

By defining the operator α through

$$\alpha A = \Delta A + A\tilde{F} - \tilde{F}A,$$

the latter equation is

$$(\alpha G)U = U\widetilde{(\alpha G)}$$

and then repeating the operation Δ successively,

$$(\alpha^2 G)U = U\widetilde{(\alpha^2 G)},$$

$$(\alpha^3 G)U = U\widetilde{(\alpha^3 G)}$$

$$\vdots \qquad\qquad \vdots \qquad .$$

This displays U as having to satisfy an infinite chain of algebraic equations.

More simply, we notice α has the derivative property

$$\alpha A_1A_2 = A_1A_2' + A_1'A_2 + A_1A_2\tilde{F} - \tilde{F}A_1A_2$$
$$= A_1(A_2' + A_2\tilde{F} - \tilde{F}A_2) + (A_1' + A_1\tilde{F} - \tilde{F}A_1)A_2$$
$$= A_1(\alpha A_2) + (\alpha A_1)A_2$$

so that α applied to Eq. (23) provides

$$GU(F + \tilde{F}) = U(F + \tilde{F})\tilde{G} \tag{24}$$

after using

$$\alpha U = U' + U\tilde{F} - \tilde{F}U = U(F + \tilde{F}).$$

Thereupon, from Eq. (24), placing $GU = U\tilde{G}$ in the left member, we obtain

$$U[(F + \tilde{F})\tilde{G} - \tilde{G}(F + \tilde{F})] = 0.$$

Owing to U being nonsingular, the [ ] must vanish or

$$G(F + \tilde{F}) = (F + \tilde{F})G.$$

That is, *in order for position to be canonical in an Hamiltonian formulation of prescribed Newtonian-order equations of motion* $\dot{v}_s = F_s(x_k, v_k)$, *it is necessary that the functional matrices* G *and the symmetric-part-of-*F [as defined by Eq. (20)] *commute.*

Returning now necessarily to the Lie–Königs viewpoint in which not x, p but $Q(x, v)$, $P(x, v)$ are canonical, some difficulties in any attempt at quantization immediately come up: Q, P are non-unique beyond simply contact transformation; even if Q in nonrelativistic limit is the position variable (a requirement that could be expected to help in the question of uniqueness), one cannot expect to phrase commutation rules straightforwardly from classical Poisson-bracket relations. Rather Q_i, Q_j could probably not be expected to commute, nor P_i, P_j, nor could Q_i, P_j be expected to have any simple commutator; the commutation rules would seem themselves to have to have a dynamical content or dependence which is absent in one-particle relativistic mechanics or in many-particle nonrelativistic mechanics.

IV. CONCLUSION

Let us summarize the main points of the present discussion and indicate some of the questions it raises.

Using electrodynamics as a base, covariant equations of motion of particles travelling along invariant curves can be formally produced. These cannot be fitted into the so-called relativistic generator formalism of Dirac, which disallows world-line invariance excepting trivally straight world lines. Yet the Lie–Königs theorem guarantees an Hamiltonian statement of motion, but not one as in the Dirac formalism for which position is a canonical coordinate. The classical family of motions is wider (admitting super-light velocities) than is customarily considered, and raises questions as to their possible significance in quantum theory though not in purely classical theory.

One may conjecture that if, as seems to be the case, the relativistic generator formalism cannot cope

with electrodynamics, then it is perhaps too restrictive altogether as a general tool in relativistic many-particle theory. To discard world-line invariance is in any case a drastic step, whose consequences it nonetheless remains important to understand, but it is not any necessary step. A rather general way, apart from Dirac's formalism, for measuring any particle dynamics, covariant or otherwise, for an Hamiltonian fitting preserving position as canonical, lies in the application of the integrability conditions for a Lagrangian encompassing the dynamics.

The convergence of the formal expansions in ϵ or $1/c^2$ for Newtonian-order equations of motion in electrodynamics is entirely an open question. One can see, though, that the original infinite-order equations of motion have the character of differential-difference equations; in other, simpler contexts such equations do admit an order-reduction of the type that has been introduced.

Finally, the outstanding question of principle is: granting a covariant Hamiltonian formulation of particle dynamics that has to exclude position as canonical, what may be the rules for quantization? It is enough to broach the question, for instance asking for universal (?) commutation rules and the encompassment of all classical orbits in a correspondence limit, to see how mutually centrifugal appear the requirements of relativity and quantum theory, and how particular must be their concordance, if there be any. Moreover the obscurity of the relativistic quantum mechanics of interacting particles may be portraying a corresponding obscurity even for noninteracting particles, insofar as the mechanics in both cases should have a common basis; it is not clear, for example, that canonical coordinates $Q(x, v)$, suited to interaction, necessarily degenerate to just position coordinates x upon removal of interaction, even though x *a priori* may certainly be taken to be canonical for free particles. Like the grin of the Cheshire cat, interaction might be leaving its impress when it is not there, so to say preparing the particles to interact, and also preparing them differently for different interactions.

V. ACKNOWLEDGMENTS

I am greatly obliged to the University of Delaware Research Foundation and to the National Science Foundation for support in this work over the past two summers, and to Dr. Robert N. Hill for valuable critical comments.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Quantum Electrodynamics with Vanishing Bare Fermion Mass

S. K. Bose and S. N. Biswas
Centre for Advanced Studies in Theoretical Physics and Astrophysics, University of Delhi, Delhi, India
(Received 5 October 1964; final manuscript received 4 January 1965)

Quantum electrodynamics is studied in the approximation suggested by Johnson, Baker, and Willey. The (approximate) Dyson equation of this model is transformed into a second-order, nonlinear differential equation by the application of a method due, originally, to Green. The solutions of this equation are then studied under various simplifying assumptions.

THE question of the existence of finite, nonperturbative solutions of the equations of Quantum Electrodynamics has recently been raised by Johnson, Baker, and Willey.[1] These authors have suggested an approximation scheme which is characterized by the properties that: (i) the bare mass of the electron is zero, (ii) the matrix elements attain their "correct" values at high energies, and that (iii) the scheme is formulated in the Landau gauge. Those authors have further shown that their model admits of finite solutions and have obtained the asymptotic form of the electron propagator valid for large values of energy. This model was subsequently studied by Maris, Herscovitz, and Jacob[2] who obtained, under an additional simplifying assumption (see below), an explicit solution for the electron propagator. This form was found to be consistent with the general requirements which a prop-

[1] K. Johnson, M. Baker, and R. Willey, Phys. Rev. Letters 11, 518 (1963).

[2] Th. Maris, V. Herscovitz, and G. Jacob, Phys. Rev. Letters 12, 313 (1964).

with electrodynamics, then it is perhaps too restrictive altogether as a general tool in relativistic many-particle theory. To discard world-line invariance is in any case a drastic step, whose consequences it nonetheless remains important to understand, but it is not any necessary step. A rather general way, apart from Dirac's formalism, for measuring any particle dynamics, covariant or otherwise, for an Hamiltonian fitting preserving position as canonical, lies in the application of the integrability conditions for a Lagrangian encompassing the dynamics.

The convergence of the formal expansions in ϵ or $1/c^2$ for Newtonian-order equations of motion in electrodynamics is entirely an open question. One can see, though, that the original infinite-order equations of motion have the character of differential-difference equations; in other, simpler contexts such equations do admit an order-reduction of the type that has been introduced.

Finally, the outstanding question of principle is: granting a covariant Hamiltonian formulation of particle dynamics that has to exclude position as canonical, what may be the rules for quantization? It is enough to broach the question, for instance asking for universal (?) commutation rules and the encompassment of all classical orbits in a correspondence limit, to see how mutually centrifugal appear the requirements of relativity and quantum theory, and how particular must be their concordance, if there be any. Moreover the obscurity of the relativistic quantum mechanics of interacting particles may be portraying a corresponding obscurity even for noninteracting particles, insofar as the mechanics in both cases should have a common basis; it is not clear, for example, that canonical coordinates $Q(x, v)$, suited to interaction, necessarily degenerate to just position coordinates x upon removal of interaction, even though x *a priori* may certainly be taken to be canonical for free particles. Like the grin of the Cheshire cat, interaction might be leaving its impress when it is not there, so to say preparing the particles to interact, and also preparing them differently for different interactions.

V. ACKNOWLEDGMENTS

I am greatly obliged to the University of Delaware Research Foundation and to the National Science Foundation for support in this work over the past two summers, and to Dr. Robert N. Hill for valuable critical comments.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Quantum Electrodynamics with Vanishing Bare Fermion Mass

S. K. Bose and S. N. Biswas
Centre for Advanced Studies in Theoretical Physics and Astrophysics, University of Delhi, Delhi, India
(Received 5 October 1964; final manuscript received 4 January 1965)

Quantum electrodynamics is studied in the approximation suggested by Johnson, Baker, and Willey. The (approximate) Dyson equation of this model is transformed into a second-order, nonlinear differential equation by the application of a method due, originally, to Green. The solutions of this equation are then studied under various simplifying assumptions.

THE question of the existence of finite, nonperturbative solutions of the equations of Quantum Electrodynamics has recently been raised by Johnson, Baker, and Willey.[1] These authors have suggested an approximation scheme which is characterized by the properties that: (i) the bare mass of the electron is zero, (ii) the matrix elements attain their "correct" values at high energies, and that (iii) the scheme is formulated in the Landau gauge. Those authors have further shown that their model admits of finite solutions and have obtained the asymptotic form of the electron propagator valid for large values of energy. This model was subsequently studied by Maris, Herscovitz, and Jacob[2] who obtained, under an additional simplifying assumption (see below), an explicit solution for the electron propagator. This form was found to be consistent with the general requirements which a prop-

[1] K. Johnson, M. Baker, and R. Willey, Phys. Rev. Letters 11, 518 (1963).

[2] Th. Maris, V. Herscovitz, and G. Jacob, Phys. Rev. Letters 12, 313 (1964).

agator of a local field theory is supposed to satisfy, for instance, the propagator was more singular on the light cone than the corresponding free-field propagator in accordance with Lehman's theorem. The purpose of the present note is to point out that a method developed previously by Green[3] in connection with the Bethe–Saltpeter equation, and which consists in transforming the relevant integral equation to a differential equation by the application of the d'Alembertian operator, can be applied to the present model. We start with the Dyson equation for the electron propagator in the approximation suggested by Johnson *et al.*[1] and transform it into a second-order, nonlinear differential equation[4] in momentum space by the application of Green's method. This differential equation [Eq. (5) of text] is still too difficult to solve. However, solutions may be obtained under additional simplifying assumptions which are valid in limited regions of energy. In this paper we describe a few such cases. We obtain, in particular, the solutions obtained in Refs. 1 and 2. We also discuss the solutions near the origin and around the mass shell.

Let us consider the Dyson equation for the fermion propagator in the approximation suggested in Ref. 1,

$$\frac{1}{S(p)} = i\gamma_\mu p_\mu - i\frac{\alpha}{4\pi^3}\int \gamma_\mu S(k)\gamma_\nu \times \left[\delta_{\mu\nu} - \frac{(p-k)_\mu (p-k)_\nu}{(p-k)^2}\right]\frac{d^4k}{(p-k)^2 - i\epsilon}. \tag{1}$$

In the above, $S(p)$ is the Fermion propagator and α the fine-structure constant. It is actually more convenient to study the mass operator m defined as

$$S^{-1}(p) = i\gamma_\mu p_\mu + m(p^2). \tag{2}$$

From (1) and (2) the integral equation for $m(p^2)$ follows:

$$m(p^2) = -\frac{3\alpha i}{4\pi^3}\int \frac{m(k^2)}{m^2(k^2) + k^2 - i\epsilon}\cdot\frac{d^4k}{(p-k)^2 - i\epsilon}. \tag{3}$$

Equation (3) is our starting point. We will now transform this equation into a differential equation by the application of Green's method. This is done as follows: One first analytically continues Eq. (3) into an Euclidean region of momenta by Wick's[5] procedure of rotation of the contour in k° plane; one then applies the d'Alembartian operator $\Box^2$ on both sides of this equation. One obtains thus

$$\begin{aligned}\Box_p^2 m(p^2) &= -i\frac{3\alpha}{4\pi^3}\int \frac{m(k^2)}{m^2(k^2)+k^2}\,\Box_p^2\frac{1}{(p-k)^2}\,d^4k \\ &= -\frac{3\alpha}{\pi}\int \frac{m(k^2)}{m^2(k^2)+k^2}\,\delta^4(p-k)\,d^4k \\ &= -\frac{3\alpha}{\pi}\frac{m(p^2)}{m^2(p^2)+p^2}.\end{aligned} \tag{4}$$

Noticing that (because of Lorentz invariance) only the "S-wave part" of the left-hand side of Eq. (4) survives, we obtain the desired differential equation

$$S\frac{d^2m}{dS^2} + 2\frac{dm}{dS} + \lambda\frac{m}{m^2+S} = 0. \tag{5}$$

We have, in the above, introduced the notation $S = p^2$ and $\lambda = (3\alpha)/4\pi$. Equation (5), being of second order, needs two boundary conditions for the specification of its solutions. Both of these follow from the integral equation (3). These are (i) $m(S)$ vanishes at infinity, $S \to \infty$ and (ii) $m(S)$ is bounded at origin, $S = 0$. Differential equation (5), together with these boundary conditions, is completely equivalent to the integral equation (3). We also notice that Eq. (5) is dilationally invariant, i.e., invariant under the transformation

$$S = l^2 S, \qquad m \to lm. \tag{6}$$

Hence if $m(S)$ is a solution, so is $lm(l^2S)$. This is an immediate consequence of the hypothesis of vanishing bare mass. This further shows that Eq. (5) really has an infinity of solutions with an arbitrary scale. To remove this arbitrariness one further puts the requirement that $S(p)$ should have a pole at the physical mass m_1, i.e., the mass operator should satisfy the condition

$$m(-m_1^2) = m_1. \tag{7}$$

With these preliminary remarks let us now pass on to study the solutions of our differential equation.

As already stated earlier, Eq. (5) is still too difficult to solve (although a numerical solution might be feasible). We are, therefore, led to search for further approximations on Eq. (5) so as to simplify its structure. But first we note that the asymptotic form of the solution at high energies can be obtained in a straightforward manner. In view of the boundary condition, Eq. (5) simplifies in the asymptotic region into

$$S\frac{d^2m}{dS^2} + 2\frac{dm}{dS} + \lambda\frac{m}{S} = 0. \tag{8}$$

[3] H. S. Green, Phys. Rev. **97**, 540 (1955).

[4] Use of (nonlinear) differential equations, to study dilationally invariant electrodynamics, has been previously made by R. Haag and Th. Maris, Phys. Rev. **132**, 2325 (1963). Our method is closely analogous to that of these authors.

[5] G. C. Wick, Phys. Rev. **96**, 1124 (1954).

Equation (8) has two solutions, the one satisfying the proper boundary condition is given by

$$m(S) = S^{-\frac{1}{2}[i\mp(1-4\lambda)^{\frac{1}{2}}]}; \qquad S \to \infty, \tag{9}$$

one[6] of which [viz. the one with negative sign before the square root $(1 - 4\lambda)^{\frac{1}{2}}$] agrees with the solution obtained previously by Johnson *et al.*[1] Solution (9) is nonanalytic in the coupling constant λ around $\lambda = \frac{1}{4}$. It may be recalled in this connection that on the basis of intuitive reasoning one expects[7] quantum electrodynamics to be singular around $\lambda = 0$. We further notice that (9) puts a restriction on λ, namely that this solution exists for λ satisfying the condition $0 < \lambda < \frac{1}{4}$. The corresponding bounds for the fine-structure constant is $0 < \alpha < \frac{1}{3}\pi$.

The main difficulty in solving Eq. (5) is clearly its nonlinear character. We have seen above that because of the boundary condition, the nonlinear term disappears in the asymptotic region. In the general case one could think of an approximation to linearize equation (5). This is achieved by replacing $m^2(S)$ occuring in the denominator of the nonlinear term in (5) by m_1^2. Equation (5) is thus replaced by

$$S\frac{d^2m}{dS^2} + 2\frac{dm}{dS} + \lambda\frac{m}{m_1^2 + S} = 0. \tag{10}$$

Equation (10) is a standard hypergeometric equation and its solution satisfying the proper boundary conditions can be immediately written down. For $S \leq m_1^2$ this is given by

$$m(S) = m_1\lambda\Gamma\left(\frac{1+\nu}{2}\right)\Gamma\left(\frac{1-\nu}{2}\right){}_2F_1 \times \left(\frac{1+\nu}{2}, \frac{1-\nu}{2}; 2; -\frac{S}{m_1^2}\right). \tag{11}$$

In the above $\nu = (1 - 4\lambda)^{\frac{1}{2}}$ and Γ is the usual gamma function. Solution (11) was previously derived by Maris *et al.*[3] who have discussed its physical content. Once again this solution gives the bounds on the coupling constant $0 < \lambda < \frac{1}{4}$ which was obtained from the asymptotic solution (9). Further this solution is consistent with Lehman's theorem regarding the degree of singularity on the light cone. The ease with which solution (11) is obtained immediately shows us the power of the present method and in particular, the advantage in working with differential equations in momentum space rather than those in coordinate space as is done, for instance, in Ref. 2.

Let us now consider the behavior of Eq. (5) near the origin, $S = 0$. It is clear that in this region the term Sd^2m/dS^2 is negligible provided only that d^2m/dS^2 is bounded. On the basis of general consideration involving the analyticity of $m(S)$ in a cut plane one expects $m(S)$ to be analytic in the neighborhood of the origin and hence to have a bounded second derivative. However, this condition is by no means guaranteed by Eq. (5). Therefore, we conclude that to the extent one can consistently obtain a solution for $m(S)$ which possesses a bounded second derivative at $S = 0$, One can throw away the term Sd^2m/dS^2 while discussing the behavior of Eq. (5) near the origin. Thus Eq. (5) simplifies to

$$2dm/dS + \lambda m/(m^2 + S) = 0. \tag{12}$$

Let us now consider the transformation

$$e^x = S/m_1^2. \tag{13}$$

Equation (12) reduces to

$$dm/dx + \tfrac{1}{2}\lambda m_1^2 e^x m/(m^2 + e^x m_1^2) = 0. \tag{14}$$

A further transformation

$$e^{-x/2}m = f \tag{15}$$

brings Eq. (14) into the form

$$\frac{df}{dx} + \frac{1}{2}f + \frac{\lambda}{2}f\frac{m_1^2}{f^2 + m_1^2} = 0. \tag{16}$$

From (16) one immediately obtains

$$-\frac{x}{2} = \int \frac{df\,(f^2 + m_1^2)}{f[f^2 + m_1^2(1+\lambda)]}. \tag{17}$$

The integration in (17) is elementary. Performing this and transforming back to the original variables, we obtain

$$m^{2/\lambda}[m^2 + S(1+\lambda)] = m_1^{2[(1+\lambda)/\lambda]}. \tag{18}$$

Thus the solutions of the nonlinear differential equation (12) are those which satisfy Eq. (18). For specified values of λ, Eq. (18) can be solved. For instance for $\lambda = 0$, the solution simply reads $m(S) = m_1$ so that the propagator $S(p)$ reduces to a free propagator (see below). For $\lambda = 1$, the solution of (18) has the form

$$m(S) = \{(S^2 + m_1^4)^{\frac{1}{2}} - S\}^{\frac{1}{2}}, \tag{19}$$

so that the condition that $m(S)$ be analytic around $S = 0$ is consistently satisfied. However for $\lambda = 2$, Eq. (18) reduces to a cubic equation whose solution develops a singularity at $S = 0$. This shows that

[6] This lack of uniqueness of the solution may be due to our use of the boundary condition as boundedness conditions. We have not been able to find out a more precise form of the boundary conditions.

[7] F. J. Dyson, Phys. Rev. 85, 631 (1952).

arbitrarily large values of λ are not allowed, in the present theory. In the general case, a numerical solution of (18) might be feasible.

We will now consider the differential equation (5) in the neighborhood of the mass shell. In this region, the term dm/dS is negligible compared to the other terms. This is because dm/dS is proportional to the vertex operator $\Lambda_\mu(p, p)[= \Gamma_\mu - \gamma_\mu]$, which vanishes on the mass shell. Thus Eq. (5) reduces to

$$Sd^2m/dS^2 + \lambda m/(m^2 + S) = 0. \tag{20}$$

Using the variables x and f introduced earlier, Eq. (20) becomes

$$d^2f/dx^2 - \tfrac{1}{4}f + \lambda m_1^2 f/(f^2 + m_1^2) = 0. \tag{21}$$

We now define

$$df/dx = p, \tag{22}$$

so that Eq. (21) reduces to

$$pdp/df = -f[\lambda m_1^2/(f^2 + m_1^2) - \tfrac{1}{4}]. \tag{23}$$

Integrating (23), we obtain

$$p^2 = \tfrac{1}{4}f^2 - \lambda m_1^2 \log (1 + f^2/m_1^2). \tag{24}$$

using (22) and (24) we obtain the "solution" in the form of a quadrature

$$x = \int \frac{df}{[\frac{1}{4}f^2 - \lambda m_1^2 \log (1 + f^2/m^2)]^{\frac{1}{2}}}. \tag{25}$$

This is almost as far as we can go; to proceed further we are forced to mutilate the integral (25). But first we note that for $\lambda = 0$, (25) yields

$$\tfrac{1}{2}x = \log (f/m_1), \tag{26}$$

so that the corresponding expression for the fermion propagator is

$$\begin{aligned} S(p) &= (ip_\mu\gamma_\mu - S/m_1)^{-1} \\ &\simeq (ip_\mu\gamma_\mu + m_1)^{-1}. \end{aligned} \tag{27}$$

Solution (27) shows that in the limit of vanishing interaction, the fermion behaves like a free particle, with its physical mass, which is assumed to be non-zero. This is an interesting situation whose implication is not very clear to the present authors. One might think that in a theory, such as the present one, in which the entire mass of the electron is supposed to be of electromagnetic origin, the limit of vanishing (electromagnetic) interaction would yield a massless electron. This is however, clearly not so in the present case, unless m_1 is identically zero, in which case the entire basis of the present scheme collapses. Since it has not been yet shown rigorously that m_1 is indeed nonzero, we feel that this whole question is completely open. Let us now go back to Eq. (25). We have not succeeded in performing the integral in the general case. We, therefore, arbitrarily expand the logarithm and retain only the first two terms. We obtain thus

$$\begin{aligned} x &\simeq \int \frac{df}{[\frac{1}{4}f^2 - \lambda f^2 + \frac{1}{2}\lambda f^4/m_1^2]^{\frac{1}{2}}} \\ &= 2(1 - 4\lambda)^{-\frac{1}{2}} \\ &\quad \times \log \frac{(1 - 4\lambda + 2\lambda f^2/m_1)^{\frac{1}{2}} - (1 - 4\lambda)^{\frac{1}{2}}}{(1 - 4\lambda + 2\lambda f^2/m_1)^{\frac{1}{2}} + (1 - 4\lambda)^{\frac{1}{2}}}, \end{aligned} \tag{28}$$

whence we obtain finally

$$\frac{S}{m_1^2} = \left\{\frac{[1 - 4\lambda + 2\lambda m^2/S]^{\frac{1}{2}} - (1 - 4\lambda)^{\frac{1}{2}}}{[1 - 4\lambda + 2\lambda m^2/S]^{\frac{1}{2}} + (1 - 4\lambda)^{\frac{1}{2}}}\right\}^{(1-4\lambda)^{-\frac{1}{2}}}. \tag{29}$$

Once again, the "solution" turns out to be non-analytic in λ around $\lambda = \frac{1}{4}$ and the same bounds for λ, viz. $0 < \lambda < \frac{1}{4}$, obtains.

Finally it should be emphasised that the original integral equation (3) is meaningful only for large S, where the corresponding differential equation (5) linearizes. The solutions of (5) around $S = 0$ and near the mass shell may, therefore, be physically unreliable. In particular, Eq. (3) does not include the correct perturbation graphs near the mass shell, nevertheless, we have discussed these solutions for the sake of completeness and also because Eq. (5) mathematically simplifies in these regions.

We are thankful to Professor R. C. Majumdar for his interest in the present work.

Scattering of Surface Waves by a Submerged Circular Cylinder*

HAROLD LEVINE
Department of Mathematics, Stanford University
(Received 25 January 1965)

The effect of a rigid circular cylinder, wholly immersed within and lying parallel to the free surface of an incompressible and inviscid fluid, on straight-crested surface waves passing overhead is investigated. A mode of analysis is developed, on the hypotheses of small amplitude and time-periodic fluid motions, that encompasses all directions of incidence of the primary wave; and is used to extend results previously obtained in the case of normal incidence. It is shown, in particular, that the absence of surface-wave reflection at normal incidence gives way to a partial reflection for other primary directions, which in turn verges on completeness as the direction of the incoming wave becomes more closely aligned with that of the cylinder axis.

1. INTRODUCTION

THE remarkable conclusion, arrived at originally by Dean (1948)[1] and subsequently confirmed by Ursell (1950)[2] and Ogilvie (1963)[3] is well known—namely, that surface waves of small amplitude experience no reflection when they pass over, in a transverse heading, a submerged and fixed circular cylinder whose axis lies parallel to the mean free surface. However, the circumstances that delimit its validity of this statement are as yet incompletely specified. To assist with the latter appraisal, a generalization of the problem treated by these authors is pursued here, namely by relinquishing the assumption of normal incidence, and details of the resulting wave complex are obtained in a manner that embraces uniformly all primary directions from the broadside on to a near grazing one; this generalization is effected by means of a formulation and mode of analysis which differs from those previously employed (and partly incapable of the requisite modification, as will be appreciated in due course).

The setting for our investigation is a heavy, frictionless, incompressible fluid of unbounded depth and lateral extension, having a plane free surface in the state of equilibrium. A circular and impenetrable cylinder of radius a is assumed to be fully immersed in the fluid, with its (infinitely elongated) axis at the fixed depth $h(> a)$ below the free surface. If a coordinate origin is located on the cylinder (or z) axis and the x, y axes are oriented as shown in the plane section at right angles thereto (Fig. 1), the level of the undisturbed free surface becomes $y = -h$.

The spectrum of time-periodic, irrotational motions of the fluid includes straight-crested, small-amplitude surface waves that can propagate freely and independently of one another; and it would be anticipated that any of these surface waves undergoes scattering in the presence of an obstacle in the fluid, say the immersed cylinder, as manifest by a reduction in the amplitude of the wave after passing over the latter and the concomitant appearance of a second or reflected wave. Dean addressed himself to the aspects of this encounter when the incoming surface wave proceeds at right angles to the axis of the cylinder, the state of affairs then being entirely two-dimensional and amenable to the techniques of complex analysis since the spatial part of the velocity potential obeys the Laplace equation. His approach entailed a conformal mapping of the whole fluid section onto a finite annular domain whose concentric circular boundaries

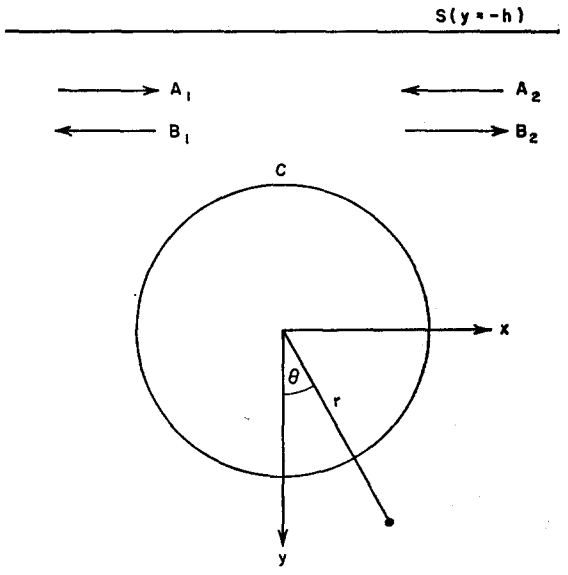


FIG. 1. Trace of submerged cylinder.

* This work was supported in part by Office of Naval Research Contract Nonr-225(74) at Stanford University.
[1] W. R. Dean, Proc. Cambridge Phil. Soc. 44, 483 (1948).
[2] F. Ursell, Proc. Cambridge Phil. Soc. 46, 141, 153 (1950).
[3] T. F. Ogilvie, J. Fluid Mech. 16, 451 (1963).

refer to the traces of the free and cylinder surfaces S, C, respectively; and then involved the characterization of harmonic functions in this simpler geometry, subject to a linearized free-surface condition on the image of S and the requirement that the image of C correspond to a streamline, as befits a vanishing normal velocity on the rigid-cylinder contour. It proved feasible to construct (infinite) sequences of functions individually compatible with the former condition, though only a finite number of them were utilized by Dean in securing partial satisfaction of the streamline condition for the purposes of a numerical calculation; nonetheless, he was able to infer the absence of any reflected wave in general, given the aforesaid conditions and the attendant linear boundary-value-problem framework, and thus to conclude that the sole effect of the cylinder at large distances therefrom is comprised in a phrase shift between the incident and transmitted wave motions.

In a pair of companion papers, Ursell dealt with the same problem by a more elaborate analysis and also succeeded in establishing the uniqueness of its solution for all values of the parameters κa, κh, where κ designates the wavenumber of free-surface motions. He dispensed with the preliminary mapping, and gave attention instead to the characterizarion of arbitrary orders of multipole (line) sources in a fluid having a free surface, the object being ultimately to determine what combination of such sources, placed at the center of the cylinder, is necessary to achieve the compensation of the normal velocity due to the primary wave at the cylinder boundary. The equations for the requisite source strengths are of an infinite linear nature, and Ursell was able to draw upon an established theory for discussing their solution in terms of absolutely convergent infinite determinants when the values of κa are not small enough to make a power series solution feasible. Ogilvie employs a similar basis in his account of the interaction between surface waves and submerged cylinders capable of executing definite movements, the overall patterns of motion continuing to retain a two-dimensional aspect.

Our approach also relies on the concept of a source function, but it is only the simple (or fundamental) variety that figures therein, as embodied in the free-surface Green's function. With the help of the latter entity, the differential version of any surface-wave boundary-value problem relating to an immersed and fixed cylinder may be recast in terms of an integral equation for the velocity potential on the cylindrical contour; this equation is, in fact, an expression of the boundary condition at the contour, with the velocity potential measuring the local strength of a (double) source distribution overlaying the whole contour and acting so as to compensate for the nonvanishing contribution of the primary wave to the normal velocity thereat. If the contour velocity potential is represented by a Fourier series (in the angular coordinate ϑ), the integral equation allows ready conversion to an infinite linear system for the expansion coefficients, and with their determination the problem is formally solved; although this system differs in form and detail from that obtained by Ursell, both share a common fitness for explicit solution when the values of κa are small.

Among the advantages stemming from the employment of a Green's function are a compactness of representation for all aspects of the fluid motion and a flexibility in regard to accommodating altered circumstances of excitation. In particular, it will be seen how directly the modification is effected should the primary surface wave approach the cylinder obliquely, without recourse to the construction of a whole new family of source functions and in contrast to the difficulties in following up Dean's mapping procedure when the behavior of the velocity potential in any plane at right angles to the cylinder axis no longer conforms to the Laplace equation. To bring out the foregoing aspects in detail, we shall discuss the problems arising out of normal and oblique incidence separately, thus also permitting, in the former case, a ready comparison with previously derived results.

2. FORMULATION OF THE SCATTERING PROBLEM AT NORMAL INCIDENCE

If the time-periodic motions at the surface of, and within, the incompressible fluid are described in terms of the velocity potential

$$\Phi(\mathbf{r}, t) = \mathrm{Re}\,\{\varphi(\mathbf{r})e^{-i\omega t}\}, \tag{1}$$

attention may henceforth be focussed on the complex-valued space factor, $\varphi(\mathbf{r})$, such that

$$\nabla^2\varphi(\mathbf{r}) = 0. \tag{2}$$

The linearized free surface boundary condition then takes the form (cf. Fig. 1)

$$\partial\varphi/\partial y + \kappa\varphi = 0, \qquad y = -h, \tag{3}$$

where

$$\kappa = \omega^2/g \tag{4}$$

and g is the gravitational constant; quiescence at the remote depths of the fluid, $y \to \infty$, implies that the velocity—or gradient of φ—vanishes.

The particular harmonic functions

$$\varphi_{\pm}(x, y) = A_{\pm}e^{\pm i\kappa x-\kappa y}, \qquad y \geq -h \tag{5}$$

that comply with the two-dimensional version of (2) and the conditions at the boundaries of the fluid, serve to generate, via (1), a pair of oppositely directed waves having the surface profiles (or elevations)

$$\xi_{\pm}(x, t) = \mathrm{Re}\,\{(i\omega/g)\varphi_{\pm}(x, -h)e^{-i\omega t}\}. \tag{6}$$

To study the interaction of these surface waves with the submerged cylinder depicted in Fig. 1, we introduce an auxiliary or Green's function, $G(x, y; x', y')$, according to the specifications

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)G(x, y; x', y') = -\delta(x - x')\delta(y - y')$$

$$(-\infty < x, x' < \infty, -h < y, y' < \infty) \tag{7}$$

and

$$(\partial/\partial y + \kappa)G = 0, \qquad y = -h, \tag{7'}$$

as well as a null limit when $y \to \infty$; it is familiar that all of the latter are vouchsafed for by the representation[4]

$$G(x, y; x', y') = -\frac{1}{2\pi}\log\frac{R}{R^*} + \frac{1}{\pi}\int_0^{\infty}\cos\zeta(x - x')e^{-\zeta(y+y'+2h)}\frac{d\zeta}{\zeta - \kappa}, \tag{8}$$

where

$$R^2 = (x - x')^2 + (y - y')^2$$

and

$$R^{*2} = (x - x')^2 + (y + y' + 2h)^2$$

evidently refer to the (squares of the) distances separating the observation point (x, y) from the source point (x', y') and its image in the free surface, $(x', -y' - 2h)$. The singularity of the integrand in (8) at $\zeta = \kappa$ calls for some remedial procedure or interpretation, and if an outgoing (surface) wave or radiation condition be invoked we have only to deform the contour slightly below this point; then it follows that

$$G(x, y; x', y') = -(1/2\pi)\log(R/R^*) + ie^{i\kappa|x-x'|-\kappa(y+y'+2h)} + \frac{1}{\pi}\int_0^{\infty}e^{-\eta|x-x'|}\{\eta\cos\eta(y + y' + 2h) - \kappa\sin\eta(y + y' + 2h)\}d\eta/(\eta^2 + \kappa^2) \tag{9}$$

[4] F. John, Commun. Pure Appl. Math. 3, 45 (1950).

and the second term (arising from the aforesaid singularity) represents the surface wave or dominant component for $|x - x'| \to \infty$.

Choosing the functions $\varphi(x, y)$ and $G(x, y; x', y')$ as the subjects for an application of Green's (symmetric) integral theorem within the entire fluid domain surrounding the rigid cylinder, at which $\partial_r\varphi = 0$, $-\pi \leq \vartheta \leq \pi$, the outcome

$$\varphi(x, y) = \varphi^{\mathrm{inc}}(x, y) + \int_C \varphi(\vartheta')\frac{\partial}{\partial r'}G(x, y; x', y')]_{r'=a}\,ds' \tag{10}$$

expresses the potential factor at any point in terms of the (assigned) incoming surface waves and the (unknown) values of the same potential on the cylinder boundary. If we ascribe a fore–aft symmetry to the incoming waves or excitation, say

$$\varphi^{\mathrm{inc}}(x, y) = \cos\kappa x\, e^{-\kappa y},$$

and designate the appertaining potential function with the index s, then

$$\varphi_s(x, y) = \cos\kappa x\, e^{-\kappa y} + \int_C \varphi_s(\vartheta')\frac{\partial}{\partial r'}G(x, y; x', y')]_{r'=a}\,ds'. \tag{11}$$

Far from the cylinder, where the surface wave component of the Green's function (9) predominates, it appears that

$$\varphi_s(x, y) \simeq e^{i\kappa x-\kappa y}\left[\tfrac{1}{2} + ie^{-2\kappa h} \times \int_C \varphi_s(\vartheta)\partial_r(e^{-i\kappa x-\kappa y})_{r=a}\,ds\right] + \tfrac{1}{2}e^{-i\kappa x-\kappa y}, \quad x \to \infty,$$

$$\varphi_s(x, y) \simeq \tfrac{1}{2}e^{i\kappa x-\kappa y} + e^{-i\kappa x-\kappa y}\left[\tfrac{1}{2} + ie^{-2\kappa h} \times \int_C \varphi_s(\vartheta)\partial_r(e^{i\kappa x-\kappa y})_{r=a}\,ds\right], \quad x \to -\infty, \tag{12}$$

whence the amplitude factors (Fig. 1) of the incoming and outgoing surface waves can be identified:

$$A_1 = \tfrac{1}{2}, \quad B_1 = \tfrac{1}{2} + ie^{-2\kappa h}\int_C \varphi_s(\vartheta)\partial_r(e^{+i\kappa x-\kappa y})_{r=a}\,ds,$$

$$A_2 = \tfrac{1}{2}, \quad B_2 = \tfrac{1}{2} + ie^{-2\kappa h}\int_C \varphi_s(\vartheta)\partial_r(e^{-i\kappa x-\kappa y})_{r=a}\,ds.$$

Since $\varphi_s(\vartheta)$ is an even function of ϑ (or x),

$$\int_C \varphi_s(\vartheta)\partial_r(e^{\pm i\kappa x-\kappa y})_{r=a}\,ds = \int_C \varphi_s(\vartheta)\partial_r(\cos\kappa x\, e^{-\kappa y})_{r=a}\,ds,$$

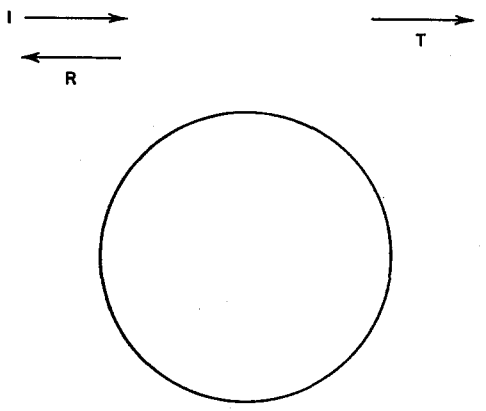


FIG. 2. Single-wave excitation.

and thus

$$A_1 = A_2 = \tfrac{1}{2}, \tag{13}$$

$$B_1 = B_2 = \tfrac{1}{2} + ie^{-2\kappa h}\int_C \varphi_s(\vartheta)\partial_r(\cos \kappa x\, e^{-\kappa y})_{r=a}\, ds.$$

To determine $\varphi_s(\vartheta)$, and thereby pave the way for a complete specification of the corresponding fluid motion, we subject the function $\varphi_s(x, y)$ in (11) to the boundary condition at the cylinder and obtain an integral equation

$$\partial_r(\cos \kappa x\, e^{-\kappa y})_{r=a} = \int_C \varphi_s(\vartheta')K(\vartheta, \vartheta')\, ds' \tag{14}$$

with the kernel

$$K(\vartheta, \vartheta') = -\partial_r\partial_{r'}G(x, y; x', y')]_{r=r'=a}. \tag{15}$$

For an incoming wave configuration of the opposite (i.e., antisymmetry) namely

$$\varphi^{\text{inc}}(x, y) = i \sin \kappa x\, e^{-\kappa y},$$

the accompanying potential function

$$\varphi_a(x, y) = i \sin \kappa x\, e^{-\kappa y} + \int_C \varphi_a(\vartheta')\frac{\partial}{\partial r'}G(x, y; x', y')]_{r'=a}\, ds' \tag{16}$$

has the asymptotic behaviors

$$\varphi_a(x, y) \simeq e^{i\kappa x-\kappa y}\left[\tfrac{1}{2}+ie^{-2\kappa h}\int_C \varphi_a(\vartheta)\partial_r(e^{-i\kappa x-\kappa y})_{r=a}\, ds\right] - \tfrac{1}{2}e^{-i\kappa x-\kappa y}, \qquad x \to \infty, \tag{17}$$

$$\varphi_a(x, y) \simeq \tfrac{1}{2}e^{i\kappa x-\kappa y} + e^{-i\kappa x-\kappa y}\left[-\tfrac{1}{2} + ie^{-2\kappa h}\int_C \varphi_a(\vartheta)\partial_r(e^{i\kappa x-\kappa y})_{r=a}\, ds\right], \qquad x \to -\infty.$$

Inasmuch as $\varphi_a(\vartheta)$ is an odd function of ϑ (or x),

$$\int_C \varphi_a(\vartheta)\partial_r(e^{-i\kappa x-\kappa y})_{r=a}\, ds = -\int_C \varphi_a(\vartheta)\partial_r(e^{i\kappa x-\kappa y})_{r=a}\, ds = -i\int_C \varphi_a(\vartheta)\partial_r(\sin \kappa x\, e^{-\kappa y})_{r=a}\, ds$$

and the incoming–outgoing wave amplitude factors comprised in (17), viz.:

$$\bar{A}_1 = -\bar{A}_2 = \tfrac{1}{2}, \tag{18}$$

$$\bar{B}_2 = -\bar{B}_1 = \tfrac{1}{2} + e^{-2\kappa h}\int_C \varphi_a(\vartheta)\partial_r(\sin \kappa x\, e^{-\kappa y})_{r=a}\, ds$$

bear the expected interrelation. Calling upon the boundary condition at the cylinder once again, this time as it concerns the wavefunction (16), we deduce an integral equation for $\varphi_a(\vartheta)$,

$$i\partial_r(\sin \kappa x\, e^{-\kappa y})_{r=a} = \int_C \varphi_a(\vartheta')K(\vartheta, \vartheta')\, ds', \tag{19}$$

with the same kernel encountered before [cf. (15)].

The outcome of superposing these symmetric and antisymmetric configuration furnishes a wave pattern as schematized in Fig. 2, wherein the reflection and transmission coefficients R, T for a lone incoming wave [with amplitude factor $1 \cdot e^{i\kappa x-\kappa y}$] take the forms

$$\begin{aligned} R &= B_1 + \bar{B}_1 \\ &= ie^{-2\kappa h}\int_C [\varphi_s(\vartheta) + \varphi_a(\vartheta)]\partial_r(e^{i\kappa x-\kappa y})_{r=a}\, ds \\ &= ie^{-2\kappa h}\int_C \varphi_s(\vartheta)\partial_r(\cos \kappa x\, e^{-\kappa y})_{r=a}\, ds \\ &\quad - e^{-2\kappa h}\int_C \varphi_a(\vartheta)\partial_r(\sin \kappa x\, e^{-\kappa y})_{r=a}\, ds \end{aligned} \tag{20}$$

and

$$\begin{aligned} T &= B_2 + \bar{B}_2 \\ &= 1 + ie^{-2\kappa h}\int_C [\varphi_s(\vartheta) + \varphi_a(\vartheta)]\partial_r(e^{-i\kappa x-\kappa y})_{r=a}\, ds \\ &= 1 + ie^{-2\kappa h}\int_C \varphi_s(\vartheta)\partial_r(\cos \kappa x\, e^{-\kappa y})_{r=a}\, ds \\ &\quad + e^{-2\kappa h}\int_C \varphi_a(\vartheta)\partial_r(\sin \kappa x\, e^{-\kappa y})_{r=a}\, ds. \end{aligned} \tag{21}$$

If we introduce the expansions

$$\varphi_s(\vartheta) = \sum_{n=0}^{\infty} a_n \cos n\vartheta, \qquad \varphi_a(\vartheta) = \sum_{n=0}^{\infty} b_n \sin n\vartheta \qquad (-\pi \le \vartheta \le \pi) \tag{22}$$

of the requisite symmetry, then

$$R = e^{-2\kappa h} \sum_{n=0}^{\infty} \{ia_n I_n - b_n J_n\}$$
$$T = 1 + e^{-2\kappa h} \sum_{n=0}^{\infty} \{ia_n I_n + b_n J_n\}, \tag{23}$$

and

where

$$I_n = \int_C \cos n\vartheta \, \partial_r(\cos \kappa x \, e^{-\kappa y})_{r=a} \, ds, \tag{24}$$

$$J_n = \int_C \sin n\vartheta \, \partial_r(\sin \kappa x \, e^{-\kappa y})_{r=a} \, ds. \tag{25}$$

After multiplying by $\cos n\vartheta$, $\sin n\vartheta$ in the integral equations (14), (19) for $\varphi_s(\vartheta)$, $\varphi_a(\vartheta)$, respectively, and integrating around the cylindrical contour, it appears that the expansion coefficients a_n , b_n of (22) are determined by the linear systems

$$I_n = \sum_{m=0}^{\infty} a_m K_{mn}, \qquad iJ_n = \sum_{m=0}^{\infty} b_m L_{mn}, \qquad n = 0, 1, \cdots, \tag{26}$$

in which

$$K_{mn} = \int_C \cos m\vartheta K(\vartheta, \vartheta') \cos n\vartheta' \, ds \, ds' = K_{nm} \tag{27}$$

and

$$L_{mn} = \int \sin m\vartheta K(\vartheta, \vartheta') \sin n\vartheta' \, ds \, ds' = L_{nm}. \tag{28}$$

3. APPROXIMATION TO THE TRANSMISSION COEFFICIENT AT NORMAL INCIDENCE

Our scheme for analyzing the broadside encounter between a lone surface wave and the submerged cylinder is now formally complete, and the next stage in carrying it forward evidently requires an evaluation of the definite integrals defined by (24), (25), (27), and (28). To this end, we rely upon the

Lemma. Let $u(x, y)$ be a solution of the reduced wave equation (without sources),

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2 + k^2)u(x, y) = 0 \tag{29}$$

on and within the circle of radius a; if the plane polar coordinates r, ϑ are referred to an origin at the center of the circle, viz.: $x = r \sin \vartheta$, $y = r \cos \vartheta$, and the symbolic operator angle D is defined by

$$\cos D = (1/ik)\partial/\partial y, \qquad \sin D = (1/ik)\partial/\partial x, \tag{30}$$

then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial r} u(r, \vartheta)\right]_{r=a} \cdot \begin{matrix}\cos \\ \sin\end{matrix} n\vartheta \, d\vartheta = i^n \left[\frac{d}{dr} J_n(kr)\right]_{r=a} \cdot \begin{matrix}\cos \\ \sin\end{matrix} nDu(0), \tag{31}$$

where $J_n(z)$ is the ordinary Bessel function and $u(0)$ designates the value of $u(x, y)$ at the center of the circle.

A corresponding statement for regular harmonic functions is the direct consequence of proceeding to the limit $k \to 0$ in (31); since

$$J_n'(z) = (nz^{n-1}/2^n n!) + O(z^{n+1}),$$

and

$$\cos nD = 2^{n-1} \cos^n D - \binom{n}{1} 2^{n-3} \cos^{n-2} D + \cdots,$$

while

$$\sin nD = \sin D\{2^{n-1} \cos^{n-1} D - \binom{n-2}{1} 2^{n-3} \cos^{n-3} D + \cdots\},$$

we find that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial u(r, \vartheta)}{\partial r}\right)_{r=a} \cos n\vartheta \, d\vartheta = \frac{a^{n-1}}{2(n-1)!} \frac{\partial^n}{\partial y^n} u(0) \tag{32}$$

and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial u(r, \vartheta)}{\partial r}\right)_{r=a} \sin n\vartheta \, d\vartheta = \frac{a^{n-1}}{2(n-1)!} \frac{\partial}{\partial x} \frac{\partial^{n-1}}{\partial y^{n-1}} u(0). \tag{33}$$

The latter formulas can be immediately applied to the evaluation of the single integrals (24), (25), wherein $\cos \kappa x \, e^{-\kappa y}$, $\sin \kappa x \, e^{-\kappa y}$ evidence a regular plane harmonic character. Thus,

$$I_n = \int_C \cos n\vartheta \, \partial_r(\cos \kappa x \, e^{-\kappa y})_{r=a} \, ds = 2\pi a \frac{a^{n-1}}{2(n-1)!} \frac{d^n}{dy^n} (e^{-\kappa y})_{y=0} = (-1)^n \frac{\pi(\kappa a)^n}{(n-1)!}, \tag{34}$$

whereas

$$J_n = \int_C \sin n\vartheta \, \partial_r(\sin \kappa x \, e^{-\kappa y})_{r=a} \, ds = 2\pi a \frac{a^{n-1}}{2(n-1)!} \frac{\partial}{\partial x} \frac{\partial^{n-1}}{\partial y^{n-1}} (e^{-\kappa y} \sin \kappa x)_{x, y=0} = (-1)^{n-1} \pi(\kappa a)^n/(n-1)! = -I_n; \tag{35}$$

it will be noted, in particular, that $I_0 = J_0 = 0$, and so the sums in (23) effectively commence at $n = 1$.

As a preliminary to the use of the generalized

mean-value relations (32), (33) in connection with the double integrals (27), (28), we write

$$G(x, y; x', y') = -(1/2\pi)\log R + g(x, y; x', y') \tag{36}$$

and thereby isolate both the singular part of the Green's function (when (x, y), (x', y') span any common domain) and the regular harmonic part [cf. (8)]

$$g(x, y; x', y') = \frac{1}{2\pi}\log R^* + \frac{1}{\pi}\int_0^\infty \cos\zeta(x - x')e^{-\zeta(y+y'+2h)}\frac{d\zeta}{\zeta - \kappa}, \tag{37}$$

if $y, y' > -h$. Since

$$\log R = \log r_> - \sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r_<}{r_>}\right)^n \cos n(\vartheta - \vartheta'),$$

where $r_>$, $r_<$ designate the greater and lesser of r, r' respectively, it now follows that

$$K_{nn} = \frac{n\pi}{2} - \pi^2\frac{a^{2n}}{[(n-1)!]^2}\frac{\partial^n}{\partial y^n}\frac{\partial^n}{\partial y'^n}g(0, y; 0, y')|_{y=y'=0}, \tag{38}$$

while

$$K_{mn} = -\pi^2\frac{a^{m+n}}{(m-1)!(n-1)!}\frac{\partial^m}{\partial y^m}\frac{\partial^n}{\partial y'^n} \times g(0, y; 0, y')|_{y=y'=0}, \qquad m \neq n. \tag{39}$$

Moreover,

$$\frac{\partial^n}{\partial y^n}\frac{\partial^n}{\partial y'^n}\log R^*|_{\substack{y,y'\to 0\\ x=x'=0}} = \frac{d^{2n}}{dy^{2n}}\log(y + 2h)]_{y=0} = \frac{d^{2n-1}}{dy^{2n-1}}\left(\frac{1}{y+2h}\right)_{y=0} = -\frac{(2n-1)!}{(2h)^{2n}},$$

and

$$\frac{\partial^n}{\partial y^n}\frac{\partial^n}{\partial y'^n}\left[\int_0^\infty \cos\zeta(x - x')e^{-\zeta(y+y'+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{\substack{y,y'\to 0\\ x=x'=0}} = \frac{d^{2n}}{dy^{2n}}\left[\int_0^\infty e^{-\zeta(y+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{y=0} = \frac{1}{2^{2n}}\frac{d^{2n}}{dh^{2n}}\int_0^\infty e^{-2\zeta h}\frac{d\zeta}{\zeta - \kappa} = \kappa^{2n}\frac{d^{2n}}{d\mu^{2n}}F(\mu)|_{\mu=2\kappa h}$$

with

$$F(\mu) = \int_0^\infty e^{-\mu\zeta}\frac{d\zeta}{\zeta - 1}; \tag{40}$$

hence

$$K_{nn} = \frac{n\pi}{2} + \frac{\pi}{2}\left(\frac{a}{2h}\right)^{2n}\frac{(2n-1)!}{[(n-1)!]^2} - \pi\left(\kappa a\right)^{2n}\frac{1}{[(n-1)!]^2}\frac{d^{2n}}{d\mu^{2n}}F(\mu)|_{\mu=2\kappa h}, \tag{41}$$

and, similarly,

$$K_{mn} = (-1)^{m+n}\frac{\pi}{2}\left(\frac{a}{2h}\right)^{m+n}\frac{(m+n-1)!}{(m-1)!(n-1)!} - \pi\left(\kappa a\right)^{m+n}\frac{1}{(m-1)!(n-1)!}\frac{d^{m+n}}{d\mu^{m+n}} \times F(\mu)|_{\mu=2\kappa h}, \qquad m \neq n, \tag{42}$$

which imply, in particular, that $K_{0n} = 0$, all n. Furthermore,

$$L_{nn} = \frac{n\pi}{2} - \pi^2\frac{a^{2n}}{[(n-1)!]^2}\frac{\partial}{\partial x}\frac{\partial}{\partial x'}\frac{\partial^{n-1}}{\partial y^{n-1}}\frac{\partial^{n-1}}{\partial y'^{n-1}} \times g(x, y; x', y')|_{\substack{x,y\\ x',y'\to 0}} \tag{43}$$

and, inasmuch as

$$\frac{\partial}{\partial x}\frac{\partial}{\partial x'}\frac{\partial^{n-1}}{\partial y^{n-1}}\frac{\partial^{n-1}}{\partial y'^{n-1}}\log R^*|_{\substack{x,y\\ x',y'\to 0}} = -\frac{\partial^2}{\partial x^2}\frac{\partial^{2n-2}}{\partial y^{2n-2}}\log[x^2 + (y+2h)^2]^{\frac{1}{2}}_{x,y\to 0} = -\frac{1}{2}\frac{\partial^{2n-2}}{\partial y^{2n-2}}\frac{\partial}{\partial x}\left\{\frac{2x}{x^2 + (y+2h)^2}\right\}_{x,y\to 0} = -\frac{d^{2n-2}}{dy^{2n-2}}\left(\frac{1}{(y+2h)^2}\right)_{y=0} = \frac{d^{2n-1}}{dy^{2n-1}}\left(\frac{1}{y+2h}\right)_{y=0},$$

while

$$\frac{\partial}{\partial x}\frac{\partial}{\partial x'}\frac{\partial^{n-1}}{\partial y^{n-1}}\frac{\partial^{n-1}}{\partial y'^{n-1}} \times \left[\int_0^\infty \cos\zeta(x - x')e^{-\zeta(y+y'+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{\substack{x,y\\ x',y'\to 0}} = -\frac{\partial^2}{\partial x^2}\frac{\partial^{2n-2}}{\partial y^{2n-2}}\left[\int_0^\infty \cos\zeta x\, e^{-\zeta(y+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{x,y\to 0} = \frac{d^{2n-2}}{dy^{2n-2}}\left[\int_0^\infty \zeta^2 e^{-\zeta(y+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{y=0} = \frac{d^{2n}}{dy^{2n}}\left[\int_0^\infty e^{-\zeta(y+2h)}\frac{d\zeta}{\zeta - \kappa}\right]_{y=0},$$

it becomes apparent that

$$K_{nn} = L_{nn};$$

this equality continues to hold, in fact, whatever the values of the indices, viz.:

$$K_{mn} = L_{mn}, \qquad \text{all } m, n. \tag{44}$$

Since $J_n = -I_n$ and $K_{mn} = L_{mn}$, the equations (26) for the Fourier expansion coefficients a_n, b_n can be displayed in the forms

$$I_n = \sum_{m=1}^{\infty} a_m K_{mn} \quad \text{and} \quad I_n = i \sum_{m=1}^{\infty} b_m K_{mn}, \qquad n = 1, 2, \cdots$$

whose comparison reveals that

$$a_m = ib_m. \tag{45}$$

Accordingly, we find

$$R = e^{-2\kappa h} \sum_{n=1}^{\infty} \{ia_n I_n - b_n J_n\} \equiv 0, \tag{46}$$

thus corroborating the conclusion reached by Dean, and furthermore

$$T = 1 + 2ie^{-2\kappa h} \sum_{n=1}^{\infty} a_n I_n, \tag{47}$$

which constitutes a simplified version of the form in (23).

To obtain specific details concerning the latter quantity, there remains the task of determining the Fourier coefficients a_n; we forego here any attempts at substantial analytic generality in this regard and content ourselves with recording a first approximation, based on the supposition that $a_1 \neq 0$, $a_2 = a_3 = \cdots = 0$. Then $a_1 = I_1/K_{11}$ and

$$T^{(1)} = 1 + 2ia_1 I_1 e^{-2\kappa h} = 1 + 2i[(I_1)^2/K_{11}]e^{-2\kappa h}, \tag{48}$$

where

$$I_1 = \int_C \cos\vartheta \, \partial_r(\cos \kappa x \, e^{-\kappa v})_{r=a} \, ds = -\pi\kappa a \tag{49}$$

and

$$K_{11} = -\int_C \cos\vartheta \, \partial_r \, \partial_{r'} G|_{r=r'=a} \cos\vartheta' \, ds \, ds' = \frac{\pi}{2} + \frac{\pi}{2}\left(\frac{a}{2h}\right)^2 - \pi(\kappa a)^2 d^2/d\mu^2 F(\mu)|_{\mu=2\kappa h}. \tag{50}$$

A complex structure attaches to the function $F(\mu)$ since the defining integral, (40), must be conducted along a path that deviates from the real ζ-axis in the neighborhood of the singularity at $\zeta = 1$; as noted earlier, the radiation condition favors bypassing the singularity from below, and it becomes a simple matter to confirm that accordingly,

$$F(\mu) = \pi i e^{-\mu} + \int_0^{\infty} \{\eta \cos \mu\eta - \sin \mu\eta\} \frac{d\eta}{\eta^2 + 1}, \tag{51}$$

where the second, and entirely real, contribution is expressible by means of the exponential integral, viz.:

$$G(\mu) = \int_0^{\infty} \{\eta \cos \mu\eta - \sin \mu\eta\} \frac{d\eta}{\eta^2 + 1} = -e^{-\mu} \mathcal{E}i^*(\mu), \tag{52}$$

with

$$\mathcal{E}i^*(x) = \mathcal{E}i(-x) + 2 \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)(2n+1)!}, \qquad 0 < x < \infty, \tag{53}$$

and

$$\mathcal{E}i(-x) = -\int_x^{\infty} \frac{e^{-t}}{t} \, dt.$$

Evidently, $d^2F/d\mu^2 = \pi i e^{-\mu} + d^2G/d\mu^2$, and therefore

$$\begin{aligned} T^{(1)} &= 1 + 2i(\pi\kappa a)^2 e^{-2\kappa h} \left\{ \frac{\pi}{2} + \pi\left(\frac{a}{h}\right)^2 \times \left[\frac{1}{8} - (\kappa h)^2 \left(\frac{d^2F}{d\mu^2}\right)_{\mu=2\kappa h}\right]\right\}^{-1} \\ &= 1 + 4i\pi(\kappa a)^2 e^{-2\kappa h} \left\{ 1 + \frac{1}{4}\left(\frac{a}{h}\right)^2 - 2(\kappa a)^2 \times \left[\pi i e^{-2\kappa h} + \left(\frac{d^2G}{d\mu^2}\right)_{\mu=2\kappa h}\right]\right\}^{-1} \\ &= \left[1 + \frac{1}{4}\left(\frac{a}{h}\right)^2 - 2(\kappa a)^2 \left(\frac{d^2G}{d\mu^2}\right)_{\mu=2\kappa h} + 2i\pi(\kappa a)^2 e^{-2\kappa h}\right]\left[1 + \frac{1}{4}\left(\frac{a}{h}\right)^2 - 2(\kappa a)^2 \left(\frac{d^2G}{d\mu^2}\right)_{\mu=2\kappa h} - 2i\pi(\kappa a)^2 e^{-2\kappa h}\right]^{-1}, \end{aligned} \tag{54}$$

with an absolute magnitude equal to unity, as befits the absence of reflection.

A second approximation to the transmission coefficient, $T^{(2)}$, follows from the assumption that a_1, $a_2 \neq 0$, $a_3 = a_4 = \cdots = 0$, whence

$$a_1 = \frac{I_1 K_{22} - I_2 K_{12}}{K_{11}K_{22} - K_{12}^2}, \qquad a_2 = \frac{I_2 K_{11} - I_1 K_{12}}{K_{11}K_{22} - K_{12}^2}$$

and

$$T^{(2)} = 1 + 2ie^{-2\kappa h}(a_1 I_1 + a_2 I_2)$$
$$= 1 + 2ie^{-2\kappa h}(K_{11}K_{22} - K_{12}^2)^{-1}$$
$$\times (I_1^2 K_{22} + I_2^2 K_{11} - 2I_1 I_2 K_{12}); \qquad (55)$$

owing to the fact that I_n has a magnitude proportional to $(\kappa a)^n$, the efficacy of this successive approximation scheme is contingent on small values of κa. If the circumstances are otherwise and a large number of coefficients a_n in the Fourier expansion of the contour potential have comparable magnitudes, we are prompted to seek alternative (analytic) representations for the latter function through modification of, or approximation in, the integral equation (14) and also to consider the utilization of variational procedures.

4. FORMULATION OF THE SCATTERING PROBLEM AT OBLIQUE INCIDENCE

Let us now inquire after the effect of the submerged cylinder upon an incoming surface wave whose direction of travel makes the (glancing) angle $\frac{1}{2}\pi - \alpha$, $0 \leq \alpha < \frac{1}{2}\pi$, relative to its axis; the limit $\alpha \to 0$ thus corresponds to the earlier case of normal incidence, and $\alpha \to \frac{1}{2}\pi$ signals the approach to a grazing encounter. If we maintain the prior orientation of the coordinate system (Fig. 1), and characterize the incoming or primary wave by the potential function

$$\varphi^{\text{inc}}(x, y, z) = e^{i\kappa(x\cos\alpha + z\sin\alpha) - \kappa y}$$
$$= e^{i\kappa z\sin\alpha}\psi^{\text{inc}}(x, y), \qquad (56)$$

it follows that

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2 - \kappa^2\sin^2\alpha)\psi^{\text{inc}} = 0. \qquad (57)$$

Inasmuch as the cylinder maintains an invariable aspect relative to any plane $z = \text{const}$, and the condition of vanishing normal velocity holds uniformly at its surface, the complete (i.e., primary + secondary) potential function $\varphi(x, y, z)$ evidences the same z dependence associated with the primary component, viz.:

$$\varphi(x, y, z) = e^{i\kappa z\sin\alpha}\psi(x, y), \qquad (58)$$

and thus, for each plane section normal to the cylinder axis, the problem centers on solution of the differential equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \kappa^2\sin^2\alpha\right)\psi(x, y) = 0. \qquad (59)$$

To describe a solution compatible with the boundary condition at both the free and cylinder surfaces in the manner hitherto employed, we require a Green's function akin to that defined by (7), though specified now in terms of the equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \nu^2\right)G(x, y; x', y')$$
$$= -\delta(x - x')\delta(y - y'). \qquad (60)$$

For the determination of this function, we may commence by subjecting (60) to a Fourier integral transformation in respect of the x-coordinate, thereby arriving at the ordinary differential equation

$$(d^2/dy^2 - k^2)\bar{G}(y, y', \zeta) = -\delta(y - y') \qquad (61)$$

[after successive integrations by parts which bring in null limit contributions when it is assumed that ν has an infinitesimal positive imaginary part], where

$$\bar{G}(y, y', \zeta)$$
$$= \int_{-\infty}^{\infty} e^{-i\zeta(x - x')}G(x - x', y, y')\, d(x - x') \qquad (62)$$

and

$$k^2 = \zeta^2 + \nu^2. \qquad (63)$$

The everywhere-continuous solution of (61) that satisfies the (free surface) condition

$$(d/dy + \kappa)\bar{G} = 0, \qquad y = 0, \qquad (64)$$

and vanishes for $y \to \infty$, can be expressed in the forms

$$\bar{G}(y, y', \zeta)$$
$$= \begin{cases} A(e^{ky} + (k + \kappa)/(k - \kappa)e^{-ky})e^{-ky'}, & y < y', \\ A(e^{ky'} + (k + \kappa)/(k - \kappa)e^{-ky'})e^{-ky}, & y > y', \end{cases}$$

where $k = +(\zeta^2 + \nu^2)^{\frac{1}{2}}$, and compliance with the discontinuity requirement for the derivative

$$d\bar{G}/dy]_{y'-0}^{y'+0} = -1$$

implies that $A = \frac{1}{2}k$. Hence

$$\bar{G}(y, y', \zeta)$$
$$= \frac{1}{2k}\left\{e^{-k(y_> - y_<)} + \frac{k + \kappa}{k - \kappa}e^{-k(y_> + y_<)}\right\}, \qquad (65)$$

if $y_>$, $y_<$ designate the greater and lesser of y, y' respectively, and by employing the inverse of the transformation (62) we arrive at an integral representation of the desired Green's function, namely

$$G(x, y; x', y') = \frac{1}{4\pi}\int_{-\infty}^{\infty} e^{i\zeta(x-x')} \times \left\{\frac{e^{-k(y_>-y_<)} - e^{-k(y+y')}}{k} + \frac{2}{k-\kappa}e^{-k(y+y')}\right\} d\zeta$$

$$= \frac{1}{2\pi}[K_0(\nu R) - K_0(\nu R^*)] + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\zeta(x-x')}\frac{e^{-k(y+y')}}{k-\kappa}\,d\zeta, \qquad (66)$$

where

$$R^2 = (x-x')^2 + (y-y')^2,$$

$$R^{*2} = (x-x')^2 + (y+y')^2,$$

and $K_0(z)$ denotes the modified Bessel function. It is readily confirmed that the above expression reverts to the form (8) in the limit $\nu \to 0$, bearing in mind that the free-surface condition has been (temporarily) enforced at $y = 0$ rather than $y = -h$.

Before completing these preliminaries relative to the Green's function, there remains the task of identifying a surface wave component therein, for which purpose the integral in (66) invites closer attention. The stipulation of a vanishing value for $\arg k = \arg(\zeta^2 + \nu^2)^{\frac{1}{2}}$ at the extremities of the path of integration fixes the assignments for $\arg k$ elsewhere as depicted in Fig. 3, if a pair of distinct vertical branch cuts be drawn away from the points $\zeta = \pm i\nu$. Since $k = \kappa$ implies that $\zeta^2 + \nu^2 = \kappa^2$ or $\zeta = \pm(\kappa^2 - \nu^2)^{\frac{1}{2}}$, we are obliged to conduct the path of integration so as to bypass these singularities of the integrand; if $\nu < \kappa$, and the latter are situated on the real axis of the ζ-plane, the indented contour shown is the one consistent with an outgoing wave behavior for the function whose integral is under consideration.

When $x - x' > 0$, the aforesaid integration contour can be deformed upwards in the ζ-plane and around the sides of the branch cut extending from

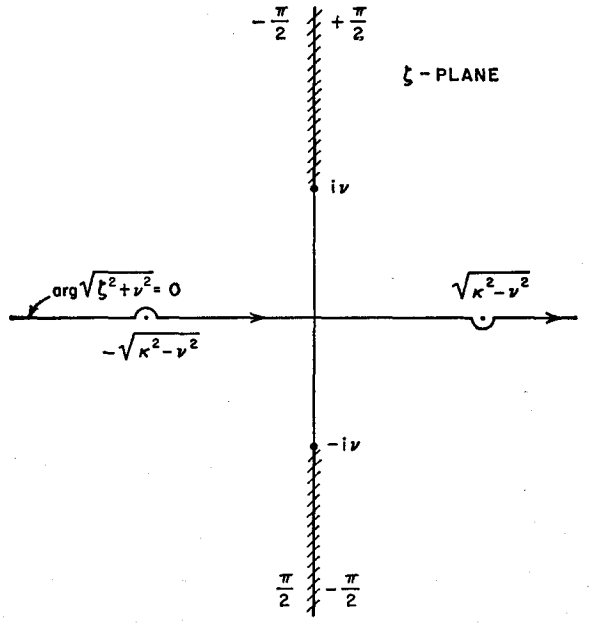


FIG. 3. Integration contour.

$\zeta = i\nu$ to $\zeta = i\infty$, with provision for the pole at $\zeta = +(\kappa^2 - \nu^2)^{\frac{1}{2}}$ which is thereby encountered. The (discrete) residue contribution arising from the latter source becomes apparent if the integrand is displayed in the form

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\zeta(x-x')-k(y+y')} \times \frac{(\zeta^2+\nu^2)^{\frac{1}{2}} + \kappa}{[\zeta - (\kappa^2-\nu^2)^{\frac{1}{2}}][\zeta + (\kappa^2-\nu^2)^{\frac{1}{2}}]}\,d\zeta$$

and thus proves to be

$$i\kappa(\kappa^2-\nu^2)^{-\frac{1}{2}} \exp[i(\kappa^2-\nu^2)^{\frac{1}{2}}(x-x')] \times \exp[-\kappa(y+y')]. \qquad (67)$$

Accordingly, the original integral is expressible by the sum of (67) and the (continuous) branch cut contribution, viz.:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\zeta(x-x')}\frac{e^{-k(y+y')}}{k-\kappa} = \frac{i\kappa}{(\kappa^2-\nu^2)^{\frac{1}{2}}}\exp[i(\kappa^2-\nu^2)^{\frac{1}{2}}(x-x')]\exp[-\kappa(y+y')]$$
$$+ \frac{1}{2\pi}\int_{\nu}^{\infty} e^{-\eta(x-x')}\frac{e^{-i(\eta^2-\nu^2)^{\frac{1}{2}}(y+y')}}{i(\eta^2-\nu^2)^{\frac{1}{2}}-\kappa}\,id\eta + \frac{1}{2\pi}\int_{\infty}^{\nu} e^{-\eta(x-x')}\frac{e^{i(\eta^2-\nu^2)^{\frac{1}{2}}(y+y')}}{-i(\eta^2-\nu^2)^{\frac{1}{2}}-\kappa}\,id\eta$$
$$= \frac{i\kappa}{(\kappa^2-\nu^2)^{\frac{1}{2}}}\exp[i(\kappa^2-\nu^2)^{\frac{1}{2}}(x-x')]\exp[-\kappa(y+y')]$$
$$+ \frac{1}{\pi}\int_{\nu}^{\infty} e^{-\eta(x-x')}\{(\eta^2-\nu^2)^{\frac{1}{2}}\cos[(\eta^2-\nu^2)^{\frac{1}{2}}(y+y')]$$
$$- \kappa\sin[(\eta^2-\nu^2)^{\frac{1}{2}}(y+y')]\}(\eta^2+\kappa^2-\nu^2)^{-1}\,d\eta, \quad x-x'>0. \qquad (68)$$

On the other hand, if $x - x' < 0$ and the transformation of our original integral is effected by analogous manipulations in the lower half of the ζ-plane, the sole modification of (68) amounts to an interchange of x and x', which is to say that the restriction appended thereto can be lifted after the replacement of $x - x'$ by $|x - x'|$.

Hence, as an alternative to the representation (66), we obtain [for $-\infty < x, x' < \infty, 0 < y, y' < \infty$]

$$G(x, y; x', y') = \frac{1}{2\pi}[K_0(\nu R) - K_0(\nu R^*)]$$
$$+ i\kappa(\kappa^2 - \nu^2)^{-\frac{1}{2}} \exp i(\kappa^2 - \nu^2)^{\frac{1}{2}} |x - x'| e^{-\kappa(y+y')}$$
$$+ \frac{1}{\pi}\int_\nu^\infty e^{-\eta|x-x'|}\{(\eta^2 - \nu^2)^{\frac{1}{2}} \cos[(\eta^2 - \nu^2)^{\frac{1}{2}}(y + y')]$$
$$- \kappa \sin[(\eta^2 - \nu^2)^{\frac{1}{2}}(y + y')]\}(\eta^2 + \kappa^2 - \nu^2)^{-1} d\eta, \quad (69)$$

wherein all of the terms have direct antecedents in the expression (8), reached after a passage to the limit $\nu \to 0$; evidently the second term of (69) predominates for $|x - x'| \to \infty$ and has the characteristic surface wave attenuation with increasing values of y. If the free surface lies at the level $y = -h$, we have only to substitute $y + y' + 2h$ for $y + y'$ in each of the preceding equations for the Green's function.

With this stock of information at hand, we can resume the analysis of the scattering of an obliquely incident surface wave by the submerged cylinder. As in the case of broadside incidence, it proves convenient to fashion the sought-for asymmetric configuration from a pair of others that have an even or odd character, respectively, in any plane section normal to the cylinder axis. More precisely, an even designation attaches to the potential factor

$$\psi_s(x, y) = \cos(\kappa x \cos\alpha)e^{-\kappa y}$$
$$+ \int_C \psi_s(\vartheta')\frac{\partial}{\partial r'} G(x, y; x', y')_{r'=a}\, ds' \quad (70)$$

whose first term underlies the incoming wave combination

$$\tfrac{1}{2}[e^{i\kappa(x\cos\alpha + z\sin\alpha) - \kappa y} + e^{i\kappa(-x\cos\alpha + z\sin\alpha) - \kappa y}]$$

that is symmetric in x, and whose second term, indicative of the presence of the cylinder, involves a Green's function of the equation (60) with the special parameter $\nu = \kappa \sin\alpha$, viz.:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \kappa^2 \sin^2\alpha\right)G(x, y; x', y')$$
$$= -\delta(x - x')\delta(y - y'). \quad (71)$$

Far from the cylinder C, where the surface wave component

$$\frac{i}{\cos\alpha} e^{i\kappa\cos\alpha|x-x'|}e^{-\kappa(y+y'+2h)}$$

of the Green's function dominates [cf. (69)], we find

$$\psi_s(x, y) \simeq e^{i\kappa x\cos\alpha - \kappa y}\left[\frac{1}{2} + \frac{ie^{-2\kappa h}}{\cos\alpha}\right.$$
$$\left.\times \int_C \psi_s(\vartheta)\partial_r(e^{-i\kappa x\cos\alpha - \kappa y})_{r=a}\, ds\right]$$
$$+ \tfrac{1}{2}e^{-i\kappa x\cos\alpha - \kappa y}, \qquad x \to \infty, \quad (72)$$
$$\psi_s(x, y) \simeq \tfrac{1}{2}e^{i\kappa x\cos\alpha - \kappa y} + e^{-i\kappa x\cos\alpha - \kappa y}\left[\frac{1}{2} + \frac{ie^{-2\kappa h}}{\cos\alpha}\right.$$
$$\left.\times \int_C \psi_s(\vartheta)\partial_r(e^{i\kappa x\cos\alpha - \kappa y})_{r=a}\, ds\right], \quad x \to -\infty,$$

and thus, relying on the fact that $\psi_s(\vartheta)$ is an even function of $\vartheta(-\pi \le \vartheta \le \pi)$, the incoming and outgoing wave amplitude factors comprised in (72) belong to the symmetric scheme

$$A_1 = A_2 = \tfrac{1}{2},$$
$$B_1 = B_2 = \tfrac{1}{2} + \frac{ie^{-2\kappa h}}{\cos\alpha} \quad (73)$$
$$\times \int_C \psi_s(\vartheta)\partial_r[\cos(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\, ds.$$

For the opposite symmetry pattern, which refers to an incoming wave combination

$$\tfrac{1}{2}[e^{i\kappa(x\cos\alpha + z\sin\alpha) - \kappa y} - e^{i\kappa(-x\cos\alpha + z\sin\alpha) - \kappa y}]$$

that is antisymmetric in x, the concomitant potential factor takes the form

$$\psi_a(x, y) = i\sin(\kappa x\cos\alpha)e^{-\kappa y}$$
$$+ \int_C \psi_a(\vartheta')\frac{\partial}{\partial r'} G(x, y; x', y')|_{r'=a}\, ds', \quad (74)$$

and bespeaks the aggregate of (correspondingly related) individual wave amplitude factors

$$\bar{A}_1 = -\bar{A}_2 = \tfrac{1}{2},$$
$$\bar{B}_2 = -\bar{B}_1 = \frac{1}{2} + \frac{e^{-2\kappa h}}{\cos\alpha}$$
$$\times \int_C \psi_a(\vartheta)\partial_r[\sin(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\, ds. \quad (75)$$

On direct superposition of this pair of symmetry configurations, we deduce that for the lone incoming wave (56) the relative amplitude factors (or reflection

and transmission coefficients) of the two outgoing waves are

$$R = B_1 + \bar{B}_1 = \frac{ie^{-2\kappa h}}{\cos\alpha} \times \int_C \psi_s(\vartheta)\partial_r[\cos(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds - \frac{e^{-2\kappa h}}{\cos\alpha}\int_C \psi_a(\vartheta)\partial_r[\sin(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds \quad (76)$$

and

$$T = B_2 + \bar{B}_2 = 1 + \frac{ie^{-2\kappa h}}{\cos\alpha} \times \int_C \psi_s(\vartheta)\partial_r[\cos(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds + \frac{e^{-2\kappa h}}{\cos\alpha}\int_C \psi_a(\vartheta)\partial_r[\sin(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds. \quad (77)$$

Pursuant to the expansions

$$\psi_s(\vartheta) = \sum_{n=0}^{\infty} a_n\cos n\vartheta, \qquad \psi_a(\vartheta) = \sum_{n=0}^{\infty} b_n\sin n\vartheta, \qquad -\pi \le \vartheta \le \pi, \quad (78)$$

that incorporate the symmetries characteristic of the respective cylinder contour potentials, it follows that

$$R = \frac{e^{-2\kappa h}}{\cos\alpha}\sum_{n=0}^{\infty}\{ia_nI_n - b_nJ_n\} \quad (79)$$

and

$$T = 1 + \frac{e^{-2\kappa h}}{\cos\alpha}\sum_{n=0}^{\infty}\{ia_nI_n + b_nJ_n\},$$

where

$$I_n = \int_C \cos n\vartheta\;\partial_r[\cos(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds \quad (80)$$

and

$$J_n = \int_C \sin n\vartheta\;\partial_r[\sin(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a}\,ds. \quad (81)$$

If these same expansions are coupled with the integral equations for the contour potentials that stem from (70), (74), namely

$$\partial_r[\cos(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a} = \int_C \psi_s(\vartheta')K(\vartheta,\vartheta')\,ds',$$

$$i\partial_r[\sin(\kappa x\cos\alpha)e^{-\kappa y}]_{r=a} = \int_C \psi_a(\vartheta')K(\vartheta,\vartheta')\,ds',$$

$$K(\vartheta,\vartheta') = -\partial_r\partial_{r'}G(x,y;x',y')|_{r=r'=a},$$

the consequences include a pair of linear systems

$$I_n = \sum_{m=0}^{\infty} a_mK_{mn}, \qquad K_{mn} = \int_C \cos m\vartheta K(\vartheta,\vartheta')\times\cos n\vartheta'\,ds\,ds', \qquad m,n = 0,1,\cdots, \quad (82)$$

$$iJ_n = \sum_{m=0}^{\infty} b_mL_{mn}, \qquad L_{mn} = \int_C \sin m\vartheta K(\vartheta,\vartheta')\times\sin n\vartheta'\,ds\,ds' \quad (83)$$

that determine the coefficients a_m, b_m. There is a manifest similarity between the schemes for determining the reflection and transmission coefficients at broadside and oblique incidence which rest upon the internally complete groups of equations (23)–(28) and (79)–(83); whatever the differences in outcome must therefore be ascribed to the elements I_m, J_m, K_{mn}, and L_{mn} that figure in the respective cases.

5. APPROXIMATION TO THE REFLECTION AND TRANSMISSION COEFFICIENTS AT OBLIQUE INCIDENCE

For particulars regarding the newly defined elements, a preliminary modification of the Lemma (29)–(31) is in order; thus, on replacing k by $i\kappa\sin\alpha$ therein, it follows that solutions of the equation

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2 - \kappa^2\sin^2\alpha)u(x,y) = 0, \quad (84)$$

regular throughout the circular domain $0 \le r \le a$, $-\pi \le \vartheta \le \pi$, conform to the generalized mean-value relations

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{\partial}{\partial r}u(r,\vartheta)\right]_{r=a}\cdot\begin{matrix}\cos\\ \sin\end{matrix} n\vartheta\,d\vartheta = (-1)^n\left[\frac{d}{dr}I_n(\kappa r\sin\alpha)\right]_{r=a}\cdot\begin{matrix}\cos\\ \sin\end{matrix} nDu(0), \quad (85)$$

where

$$\cos D = -\frac{1}{\kappa\sin\alpha}\frac{\partial}{\partial y}, \qquad \sin D = -\frac{1}{\kappa\sin\alpha}\frac{\partial}{\partial x} \quad (86)$$

and $x = r\sin\vartheta$, $y = r\cos\vartheta$; here $I_n(z)$ designates the Bessel function of purely imaginary argument, viz.:

$$I_n(z) = e^{-\frac{1}{2}n\pi i}J_n(iz).$$

The integrals (80), (81) may be directly evaluated by recourse to (85), (86), and the initial representatives ($n = 0, 1, 2$) prove to be

$$I_0 = 2\pi\kappa a\sin\alpha I_1(\kappa a\sin\alpha),$$
$$I_1 = -2\pi\kappa aI_1'(\kappa a\sin\alpha), \quad (87)$$
$$I_2 = (2\pi\kappa a/\sin\alpha)(2-\sin^2\alpha)I_2'(\kappa a\sin\alpha),$$

and

$$J_0 = 0,$$
$$J_1 = 2\pi\kappa a \cos\alpha I_1'(\kappa a \sin\alpha), \tag{88}$$
$$J_2 = -\frac{4\pi\kappa a}{\sin\alpha}\cos\alpha I_2'(\kappa a \sin\alpha),$$

where the primes connote argument derivatives; it is readily confirmed that these expressions are in harmony with the previously obtained values in the limit $\alpha \to 0$, though the proportionality relation expressed in (35) does not apply when $\alpha \neq 0$.

As in the case of broadside incidence, we effect an evaluation of the double integrals characterizing K_{mn} and L_{mn} after the resolution

$$G(x, y; x', y') = (1/2\pi)K_0(\kappa R \sin\alpha) + g(x, y; x', y'),$$

which isolates the singular and regular components of the Green's function (71) for all dispositions of the points (x, y) and (x', y') on or within the immersed contour C. The former possesses a Fourier expansion

$$\frac{1}{2\pi}K_0(\kappa R \sin\alpha) = \frac{1}{2\pi}I_0(\kappa r_< \sin\alpha)K_0(\kappa r_> \sin\alpha) + \frac{1}{\pi}\sum_{n=1}^{\infty} I_n(\kappa r_< \sin\alpha)K_n(\kappa r_> \sin\alpha)\cos n(\vartheta - \vartheta'),$$

that lends itself to the determination of (harmonic) averages on the contour, while the latter

$$g(x, y; x', y') = -\frac{1}{2\pi}K_0(\kappa R^* \sin\alpha) + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\zeta(x-x')} \times \frac{\exp[-(\zeta^2 + \kappa^2\sin^2\alpha)^{\frac{1}{2}}(y + y' + 2h)]}{(\zeta^2 + \kappa^2\sin^2\alpha)^{\frac{1}{2}} - \kappa}\, d\zeta,$$

is amenable to the computation of such averages by means of (85), (86). Proceeding in this fashion, there obtains

$$K_{00} = \frac{I_0^2}{2\pi}\left\{\frac{K_1(\kappa a \sin\alpha)}{I_1(\kappa a \sin\alpha)} + K_0(2\kappa h \sin\alpha) - F(2\kappa h, \sin\alpha)\right\}, \tag{89}$$

$$K_{01} = K_{10} = \frac{I_0 I_1}{2\pi}\left\{\sin\alpha K_1(2\kappa h \sin\alpha) + \frac{d}{d\mu}F(\mu, \sin\alpha)\Big]_{\mu=2\kappa h}\right\} \tag{90}$$

and

$$K_{11} = \frac{I_1^2}{2\pi}\left\{-\frac{1}{2}\sin^2\alpha\frac{K_1'(\kappa a \sin\alpha)}{I_1'(\kappa a \sin\alpha)} + \sin^2\alpha\left[K_0(2\kappa h \sin\alpha) + \frac{K_1(2\kappa h \sin\alpha)}{2\kappa h \sin\alpha}\right] - \frac{d^2}{d\mu^2}F(\mu, \sin\alpha)|_{\mu=2\kappa h}\right\}, \tag{91}$$

whereas $L_{00} = L_{01} = L_{10} = 0$ and

$$L_{11} = \frac{I_1^2}{2\pi}\left\{-\frac{1}{2}\sin^2\alpha\frac{K_1'(\kappa a \sin\alpha)}{I_1'(\kappa a \sin\alpha)} + \frac{\sin\alpha}{2\kappa h}K_1(2\kappa h \sin\alpha) + \left(\sin^2\alpha - \frac{d^2}{d\mu^2}\right)F(\mu, \sin\alpha)|_{\mu=2\kappa h}\right\}, \tag{92}$$

etc., with

$$F(\mu, \sin\alpha) = \int_{-\infty}^{\infty}\frac{\exp[-\mu(\zeta^2 + \sin^2\alpha)^{\frac{1}{2}}]}{(\zeta^2 + \sin^2\alpha)^{\frac{1}{2}} - 1}\, d\zeta. \tag{93}$$

On emulating the integration procedure employed for the establishment of the alternative Green's function representation (69), we deduce that

$$F(\mu, \sin\alpha) = \frac{2\pi i}{\cos\alpha}e^{-\mu} + 2\int_{\sin\alpha}^{\infty}\frac{(\eta^2 - \sin^2\alpha)^{\frac{1}{2}}}{\eta^2 + \cos^2\alpha}\cos[\mu(\eta^2 - \sin^2\alpha)^{\frac{1}{2}}]\, d\eta - 2\int_{\sin\alpha}^{\infty}\frac{\sin[\mu(\eta^2 - \sin^2\alpha)^{\frac{1}{2}}]}{\eta^2 + \cos^2\alpha}\, d\eta. \tag{94}$$

Moreover, the combination of (real) infinite integrals in (94) can be recast in terms of a single integral over a finite range, viz.:

$$F(\mu, \sin\alpha) = \frac{2\pi i}{\cos\alpha}e^{-\mu} + 2K_0(\mu\sin\alpha) - \frac{2e^{-\mu}}{\cos\alpha}\log\frac{1 + \cos\alpha}{\sin\alpha} - 2e^{-\mu}\int_0^{\mu} e^x K_0(x\sin\alpha)\, dx. \tag{95}$$

Accordingly, the complex nature of the quantities K_{00}, K_{01}, K_{11} and L_{11} is displayed by the forms

$$K_{00} = -i(I_0^2/\cos\alpha)e^{-2\kappa h} + (I_0^2/2\pi)k_{00},$$
$$K_{01} = -i(I_0 I_1/\cos\alpha)e^{-2\kappa h} - (I_0 I_1/2\pi)k_{01}, \tag{96}$$
$$K_{11} = -i(I_1^2/\cos\alpha)e^{-2\kappa h} + (I_1^2/2\pi)k_{11},$$

and

$$L_{11} = -iI_1^2\cos\alpha e^{-2\kappa h} + (I_1^2/2\pi)l_{11},$$

whose real parts find explicit expression in terms of the appertaining factors

$$k_{00} = \frac{K_1(\kappa a \sin\alpha)}{I_1(\kappa a \sin\alpha)} - K_0(2\kappa h \sin\alpha) + \frac{2e^{-2\kappa h}}{\cos\alpha}\log\frac{1+\cos\alpha}{\sin\alpha} + 2e^{-2\kappa h}\int_0^{2\kappa h} e^x K_0(x\sin\alpha)\,dx,$$

$$k_{01} = 2K_0(2\kappa h\sin\alpha) + \sin\alpha K_1(2\kappa h\sin\alpha) - \frac{2e^{-2\kappa h}}{\cos\alpha}\log\frac{1+\cos\alpha}{\sin\alpha} - 2e^{-2\kappa h}\int_0^{2\kappa h} e^x K_0(x\sin\alpha)\,dx, \quad (97)$$

$$k_{11} = -\tfrac{1}{2}\sin^2\alpha\frac{K_1'(\kappa a\sin\alpha)}{I_1'(\kappa a\sin\alpha)} - 2K_0(2\kappa h\sin\alpha) - 2\sin\alpha K_1(2\kappa h\sin\alpha) - \sin^2\alpha\left\{K_0(2\kappa h\sin\alpha) + \frac{K_1(2\kappa h\sin\alpha)}{2\kappa h\sin\alpha}\right\} + \frac{2e^{-2\kappa h}}{\cos\alpha}\log\frac{1+\cos\alpha}{\sin\alpha} + 2e^{-2\kappa h}\int_0^{2\kappa h} e^x K_0(x\sin\alpha)\,dx,$$

and

$$l_{11} = -\tfrac{1}{2}\sin^2\alpha\frac{K_1'(\kappa a\sin\alpha)}{I_1'(\kappa a\sin\alpha)} - 2K_0(2\kappa h\sin\alpha) - 2\sin\alpha K_1(2\kappa h\sin\alpha) - (\sin\alpha/2\kappa h)K_1(2\kappa h\sin\alpha) + 2\cos\alpha e^{-2\kappa h}\log[(1+\cos\alpha)/\sin\alpha] + 2\cos^2\alpha e^{-2\kappa h}\int_0^{2\kappa h} e^x K_0(x\sin\alpha)\,dx.$$

For a first approximation to the reflection and transmission coefficients, we employ truncated versions of the linear systems (82), (83) based on the hypotheses that a_0, a_1 and b_1 are the sole nonvanishing coefficients in the contour potential expansions. Then

$$a_0 = \frac{I_0K_{11} - I_1K_{01}}{K_{00}K_{11} - K_{01}^2}, \quad a_1 = \frac{I_1K_{00} - I_0K_{01}}{K_{00}K_{11} - K_{01}^2}, \quad b_1 = i\frac{J_1}{L_{11}}$$

and

$$R^{(1)} = (e^{-2\kappa h}/\cos\alpha)[i(a_0I_0 + a_1I_1) - b_1J_1] = i\frac{e^{-2\kappa h}}{\cos\alpha}\left[\frac{I_0^2K_{11} + I_1^2K_{00} - 2I_0I_1K_{01}}{K_{00}K_{11} - K_{01}^2} - \cos^2\alpha\frac{I_1^2}{L_{11}}\right], \quad (98)$$

whereas

$$T^{(1)} = 1 + i\frac{e^{-2\kappa h}}{\cos\alpha}\frac{I_0^2K_{11} + I_1^2K_{00} - 2I_0I_1K_{01}}{K_{00}K_{11} - K_{01}^2} + \cos^2\alpha\frac{I_1^2}{L_{11}}. \quad (99)$$

Since

$$I_0^2K_{11} + I_1^2K_{00} - 2I_0I_1K_{01} = (I_0^2I_1^2/2\pi)(k_{00} + k_{11} + 2k_{01})$$

and

$$K_{00}K_{11} - K_{01}^2 = -i(I_0^2I_1^2/2\pi\cos\alpha)e^{-2\kappa h}(k_{00} + k_{11} + 2k_{01}) + [I_0^2I_1^2/(2\pi)^2](k_{00}k_{11} - k_{01}^2),$$

it appears that

$$R^{(1)} = -\left[1 + \frac{i}{2\pi}\cos\alpha\frac{k_{00}k_{11} - k_{01}^2}{k_{00} + k_{11} + 2k_{01}}e^{2\kappa h}\right]^{-1} + \frac{\cos\alpha}{\cos\alpha + (i/2\pi)l_{11}e^{2\kappa h}}. \quad (100)$$

if $\alpha \to 0$, $l_{11} \to k_{11}$, $k_{00} \to \infty$, and

$$(k_{00}k_{11} - k_{01}^2)/(k_{00} + k_{11} + 2k_{01}) \to k_{11},$$

whence $R^{(1)} \to 0$, as anticipated; if $\alpha \to \frac{1}{2}\pi$, $R^{(1)} \to -1$, in keeping with the fact that a straight-crested surface wave cannot travel for unlimited distances over, and parallel to the axis of a submerged circular cylinder.

The transmission coefficient $T^{(1)}$, given by (99), reduces to the prior approximation, (48), at normal incidence and, in consort with $R^{(1)}$, upholds the requirement of energy balance among the primary and secondary surface wave components.

Transformation from a Linear Momentum to an Angular Momentum Basis for Relativistic Particles of Nonzero Mass and Any Spin

H. E. Moses
MIT Lincoln Laboratory, Lexington, Massachusetts*
(Received 14 January 1965)

The infinitesimal generators of the inhomogeneous Lorentz group have been given in a basis in which the components of the linear-momentum operators are diagonal and in another basis in which the square of the angular momentum is diagonal for all unitary irreducible ray representations of the group. In a previous paper we showed how the two bases were related for representations corresponding to zero mass and any finite spin. In the present paper we show how the two bases are related for representations corresponding to nonzero mass and any spin. Thus this paper and the preceding one enable us to expand relativistic plane waves into relativistic spherical waves and vice-versa for particles of any spin and any mass.

In the previous paper we used the relation between the linear- and angular-momentum bases to integrate the infinitesimal generators in the angular-momentum basis and thereby obtain closed expressions which show how the angular momentum of particles of zero mass and any finite spin transform under changes of frame of reference. Similar use could be made of the transformation of the present paper. Results of such use will be given in a later paper.

1. INTRODUCTION AND SUMMARY

In Ref. 1 we showed how the linear- and angular-momentum bases were related for particles of zero mass and any finite spin. In the present paper it is our objective to show how the two bases are related for relativistic particles of nonzero mass and any spin. In order to minimize the preliminaries, we shall write this paper as a direct extension of Ref. 1, which we shall hereafter refer to as Part I and refer to formulas there as needed.

In the case of nonzero mass, we let the abstract vector Φ be represented by the complex function $f(\mathbf{p}, \gamma) = f(p_1, p_2, p_3, \gamma)$ in the linear-momentum basis where the variables p_i have the same range of values as in Part I and the variable γ can take on only the $2s + 1$ discrete values $\gamma = -s, -s + 1, \cdots, s - 1, s$. As before s is the spin of the particle and is thus either a nonnegative integer or a positive half-odd integer.

In the angular momentum basis, the abstract vector Φ is represented by the complex function $F(E, j, m, \alpha)$ where the variable E is a continuous real variable which has the range $\mu < E < \infty$ when the sign of the energy $\epsilon = +1$ and which has the range $-\infty < E < -\mu$ for $\epsilon = -1$, where μ is the mass of the particle. The variable α can take on the values $-s, -s + 1, \cdots, s - 1, s$; the quantum number j takes on the values (for a fixed value of α) $j = |\alpha|, |\alpha| + 1, |\alpha| + 2, \cdots$; and for a fixed value of j, the quantum number m takes on the $2j + 1$ values $m = -j, -j + 1, \cdots, j - 1, j$.

In terms of the two bases the norm of Φ is given by

$$(\Phi, \Phi) = \sum_{\gamma} \int \frac{d\mathbf{p}}{p} |f(\mathbf{p}, \gamma)|^2 = \sum_{\alpha, j, m} \int \frac{dE}{k(E)} |F(E, j, m, \alpha)|^2,$$

where $k(E) = (E^2 - \mu^2)^{\frac{1}{2}}$, and the integrations and summations are taken over the ranges of definition of the variables.

As in Part I we write

$$\Phi \underset{p}{\leftrightarrow} f(\mathbf{p}, \gamma) \underset{i}{\leftrightarrow} F(E, j, m, \alpha),$$

$$A\Phi \underset{p}{\leftrightarrow} (Af)(\mathbf{p}, \gamma) \underset{i}{\leftrightarrow} (AF)(E, j, m, \alpha),$$

where A is an operator in the Hilbert space. Thus far, the chief difference between the nonzero-mass representations and the zero-mass representations of Part I is the appearance of the additional variables γ and α. The appearance of these variables adds considerably to the difficulty of finding the transformation between the function $f(p, \gamma)$ and the function $F(E, j, m, \alpha)$.

We now show how the infinitesimal generators act in the linear-momentum basis. For the sake of definiteness we shall pick a particular representation for the spin operators which appear in this basis. We introduce three operators S_i $(i = 1, 2, 3)$ which are defined as follows:

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

[1] H. E. Moses, J. Math. Phys. 6, 928 (1965).

$$(S_3 f)(\mathbf{p}, \gamma) = \gamma f(\mathbf{p}, \gamma),$$
$$((S_2 + iS_1)f)(\mathbf{p}, \gamma) = [(s - \gamma)(s + \gamma + 1)]^{\frac{1}{2}} f(\mathbf{p}, \gamma + 1),$$
$$((S_2 - iS_1)f)(\mathbf{p}, \gamma) = [(s + \gamma)(s - \gamma + 1)]^{\frac{1}{2}} f(\mathbf{p}, \gamma - 1). \quad (1.1)$$

It is easily seen that the operators S_i are Hermitian angular momentum operators which constitute an irreducible set characterized by $S_1^2 + S_2^2 + S_3^2 = s(s + 1)I$. Indeed (1.1) gives a standard representation for angular momentum or spin operators.

The infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group in the linear momentum bases are (Refs. 2 and 3)

$$(P_i f)(\mathbf{p}, \gamma) = p_i f(\mathbf{p}, \gamma),$$
$$(Hf)(\mathbf{p}, \gamma) = \epsilon\omega(p) f(\mathbf{p}, \gamma),$$
$$(J_i f)(\mathbf{p}, \gamma) = -i(\mathbf{p} \times \nabla)_i f(\mathbf{p}, \gamma) + (S_i f)(\mathbf{p}, \gamma),$$
$$(\mathcal{J}_i f)(\mathbf{p}, \gamma) = \epsilon[i\omega(p)\nabla_i f(\mathbf{p}, \gamma) + \sum_{jk} \frac{\epsilon_{ijk} p_j}{\omega(p) + \mu} (S_k f)(\mathbf{p}, \gamma)] \quad (1.2)$$

where $\nabla_i = \partial/\partial p_i$, $p = |\mathbf{p}|$, $\omega(p) = (p^2 + \mu^2)^{\frac{1}{2}}$, and ϵ_{ijk} is the usual antisymmetric three-index symbol.

The infinitesimal generators as given in the angular momentum basis are so complicated that we do not reproduce them here but refer to equations (2.1)–(2.10) of Ref. 4. However, we shall have to make some changes in order that the results of Ref. 4 are given in our notation. First of all, the function $\varphi(E, j, m, \alpha)$ of Ref. 4 is to be replaced by $F(E, j, m, \alpha)$. The quantity p of Ref. 4 is to be replaced by $k(E) = (E^2 - \mu^2)^{\frac{1}{2}}$. Finally, we note that if A is an operator, $A\varphi(E, j, m, \alpha)$ is now to be written $(AF)(E, j, m, \alpha)$.

We note also an erratum in Ref. 4, namely the following: In the present paper we use μ as being the mass of the particle which we always take as positive; in Ref. 4 we must replace μ by $\epsilon\mu$ to be consistent with the notation of the present paper.

We are now in a position to give the transformation between the functions $f(\mathbf{p}, \gamma)$ and $F(E, j, m, \alpha)$ which is the principal point of the present paper. We define the polynomial $S(j, m, \alpha, x)$ by equation (1.4) of Part I. Then

$$f(\mathbf{p}, \gamma) = \frac{1}{\pi^{\frac{1}{2}}} \frac{1}{p} \sum_{j,m,\alpha} (-i)^{m-\gamma} (-1)^{\alpha+\gamma+m+s} (\tfrac{1}{2})^{\gamma+m+1} \times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}} \left(1 - \frac{p_3}{p}\right)^{\gamma-\alpha} \left(1 + \frac{p_3}{p}\right)^{\gamma+\alpha} \times \left(\frac{p_1 + ip_2}{p}\right)^{m-\gamma} S\left(j, m, \alpha, \frac{p_3}{p}\right) S\left(s, \gamma, \alpha, \frac{p_3}{p}\right) F(\epsilon\omega(p), j, m, \alpha), \quad (1.3)$$

$$F(E, j, m, \alpha) = \frac{k(E)}{\pi^{\frac{1}{2}}} \sum_{\gamma=-s}^{s} (-1)^{\alpha+\gamma+m+s} (\tfrac{1}{2})^{\gamma+m+1} (i)^{m-\gamma} \times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}} \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, (\sin\theta)^{m-\gamma+1} e^{-i(m-\gamma)\varphi} \times (1 - \cos\theta)^{\gamma-\alpha} (1 + \cos\theta)^{\gamma+\alpha} S(j, m, \alpha, \cos\theta) S(s, \gamma, \alpha, \cos\theta) \times f(k(E) \sin\theta \cos\varphi, k(E) \sin\theta \sin\varphi, k(E) \cos\theta, \gamma). \quad (1.4)$$

In (1.3) the summation is taken over the entire range of definition of the variables. Also we have introduced the usual polar angles θ, φ by $\mathbf{p} = p(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$.

As in Part I we can use the relations (1.3) and (1.4) to integrate the infinitesimal generators in the angular momentum basis and thereby show explicitly how the angular momentum of a particle changes under changes of frame of reference. We shall postpone this discussion for another paper. It should be mentioned, however, that in Ref. 5 which was written before (1.3) and (1.4) were derived, the infinitesimal generators in the angular momentum basis were integrated numerically to show how the angular momentum changes under changes of origin, this change being called "displacement broadening of angular

[2] L. L. Foldy, Phys. Rev. **102**, 568 (1956).
[3] Iu. M. Shirokov, Zh. Eksperim. i Teor. Fiz. **33**, 1196 (1957) [English transl.: Soviet Phys.—JETP **13**, 240 (1961)].
[4] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 294 (1964).
[5] H. E. Moses and S. C. Wang, Nuovo Cimento **36**, 788 (1965).

momentum." Also some calculations are given for small displacements of origin.

The remainder of the present paper presents the derivation of (1.3) and (1.4). In contrast to our presentation in Part I where we verified the corresponding relations for massless paricles by using properties of the Jacobi polynomials, we shall in the present paper derive the expressions from first principles. A simplified form of the procedure of the present paper was used to derive the expressions for the massless case of Part I.

2. USE OF BRA AND KET NOTATION AND THE TRANSFORMATION FUNCTION

It will be convenient to use Dirac's bra and ket notation. Accordingly, abstract vectors in Hilbert space will be denoted by kets. To describe the linear-momentum basis we introduce the set of kets $|\mathbf{p}, \gamma\rangle$ and the corresponding set of bras. To describe the angular momentum basis we introduce the set of kets $|E, j, m, \alpha\rangle$ and the corresponding bras. Then if the state which was designated in Sec. 1 by Φ is now designated by $|\Phi\rangle$, we relate the functions $f(\mathbf{p}, \gamma)$ and $F(E, j, m, \alpha)$ to $|\Phi\rangle$ by

$$f(\mathbf{p}, \gamma) = \langle p, \gamma \mid \Phi\rangle \tag{2.1}$$

and

$$F(E, j, m, \alpha) = \langle E, j, m, \alpha \mid \Phi\rangle. \tag{2.2}$$

The requirement that the inner product be preserved in the two bases leads to the completeness relations

$$\sum_{\gamma} \int \frac{d\mathbf{p}}{p} |\mathbf{p}, \gamma\rangle\langle\mathbf{p}, \gamma| = \sum_{j,m,\alpha} \int \frac{dE}{k(E)} |E, j, m, \alpha\rangle\langle E, j, m, \alpha| = I, \tag{2.3}$$

where I is the identity operator. It follows that

$$f(\mathbf{p}, \gamma) = \sum_{j,m,\alpha} \int \langle \mathbf{p}, \gamma \mid E, j, m, \alpha\rangle \times \frac{dE}{k(E)} F(E, j, m, \alpha), \tag{2.4}$$

and

$$F(E, j, m, \alpha) = \sum_{\gamma} \int \langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle \frac{d\mathbf{p}}{p} f(\mathbf{p}, \gamma). \tag{2.5}$$

It will be our objective to find the transformation function $\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle$. This objective will be accomplished by considering operators A constructed from the set of infinitesimal generators of the group and evaluating $\langle E, j, m, \alpha| A |\mathbf{p}, \gamma\rangle$ in two ways: (1) by considering this quantity to be a product of the bra $\langle E, j, m, \alpha|$ and the ket $A |\mathbf{p}, \gamma\rangle$, (2) by considering this quantity to be a product of the bra $\langle E, j, m, \alpha| A$ and the ket $|\mathbf{p}, \gamma\rangle$. On equating these two expressions for a sufficient number of operators A, we shall find that the transformation function is completely determined up to a constant factor. This constant is then determined by the completeness relation (2.3).

We now show how some of the infinitesimal generators of the group act on the kets $|\mathbf{p}, \gamma\rangle$. These results are obtained by transcribing (1.1) and (1.2) into ket notation.

$$S_3 |\mathbf{p}, \gamma\rangle = \gamma |\mathbf{p}, \gamma\rangle,$$

$$(S_2 - iS_1) |\mathbf{p}, \gamma\rangle = [(s - \gamma)(s + \gamma + 1)]^{\frac{1}{2}} |\mathbf{p}, \gamma + 1\rangle,$$

$$(S_2 + iS_1) |\mathbf{p}, \gamma\rangle = [(s + \gamma)(s - \gamma + 1)]^{\frac{1}{2}} |\mathbf{p}, \gamma - 1\rangle,$$

$$P_i |\mathbf{p}, \gamma\rangle = p_i |\mathbf{p}, \gamma\rangle,$$

$$H |\mathbf{p}, \gamma\rangle = \epsilon\omega(p) |\mathbf{p}, \gamma\rangle,$$

$$J_3 |\mathbf{p}, \gamma\rangle = i(\mathbf{p} \times \nabla)_3 |\mathbf{p}, \gamma\rangle + S_3 |\mathbf{p}, \gamma\rangle,$$

$$(J_2 - iJ_1) |\mathbf{p}, \gamma\rangle = \left[ip_3 \frac{\partial}{\partial p_1} - p_3 \frac{\partial}{\partial p_2} + (p_2 - ip_1) \frac{\partial}{\partial p_3}\right] |\mathbf{p}, \gamma\rangle + (S_2 - iS_1) |\mathbf{p}, \gamma\rangle,$$

$$(J_2 + iJ_1) |\mathbf{p}, \gamma\rangle = \left[ip_3 \frac{\partial}{\partial p_1} + p_3 \frac{\partial}{\partial p_2} - (p_2 + ip_1) \frac{\partial}{\partial p_3}\right] |\mathbf{p}, \gamma\rangle + (S_2 + iS_1) |\mathbf{p}, \gamma\rangle,$$

$$\mathcal{J}_3 |\mathbf{p}, \gamma\rangle = -i\epsilon\omega(p) \frac{\partial}{\partial p_3} |\mathbf{p}, \gamma\rangle + \frac{i\epsilon}{2[\omega(p) + \mu]} [(p_2 - ip_1)(S_2 + iS_1) |\mathbf{p}, \gamma\rangle - (p_2 + ip_1)(S_2 - iS_1) |\mathbf{p}, \gamma\rangle]. \tag{2.6}$$

We shall also need to know how some of the operators act on the bras $\langle E, j, m, \alpha|$. In particular,

$$\langle E, j, m, \alpha| H = E\langle E, j, m, \alpha|,$$

$$\langle E, j, m, \alpha| J_3 = m\langle E, j, m, \alpha|,$$

$$\langle E, j, m, \alpha| (J_2 - iJ_1) = [(j + m)(j - m + 1)]^{\frac{1}{2}}\langle E, j, m - 1, \alpha|,$$

$$\langle E, j, m, \alpha| (J_2 + iJ_1) = [(j - m)(j + m + 1)]^{\frac{1}{2}}\langle E, j, m + 1, \alpha|,$$

$$\langle E, j, j, \alpha| P_3 = k(E)\left[\frac{\alpha}{j+1}\langle E, j, j, \alpha| + \frac{1}{j+1}\left[\frac{(2j+1)(j-\alpha+1)(j+\alpha+1)}{2j+3}\right]^{\frac{1}{2}}\langle E, j, j+1, \alpha|\right],$$

$$\langle E, s, s, s| \mathcal{J}_3 = \frac{i}{s+1}\left\{sk(E)\frac{\partial}{\partial E}\langle E, s, s, s| + \frac{\epsilon\mu}{2k(E)}2s\langle E, s, s, s-1|\right\}$$

$$+\frac{i}{s+1}[2s+3]^{-\frac{1}{2}}\left\{(2s+1)^{\frac{1}{2}}\left[(s+1)\frac{E}{k(E)} + k(E)\frac{\partial}{\partial E}\right]\langle E, s+1, s, s,| + \frac{\epsilon\mu}{2k(E)}(4s)^{\frac{1}{2}}\langle E, s+1, s, s-1|\right\},$$

$$\langle E, j, m, \alpha|\, \mathbf{P}\cdot\mathbf{J} = k(E)\alpha\langle E, j, m, \alpha|. \tag{2.7}$$

Finally, we shall need to know how the operator T_3 defined by Eq. (2.13) of Ref. 4 acts on $\langle E, j, j, \alpha|$,

$$\begin{aligned}\langle E, j, j, \alpha|\, T_3 = &-\frac{1}{2(j+1)}[(j-\alpha+1)(j+\alpha)(s-\alpha+1)(s+\alpha)]^{\frac{1}{2}}\langle E, j, j, \alpha-1|\\ &-\frac{1}{2(j+1)}\left[\frac{(j-\alpha+1)(j-\alpha+2)(s-\alpha+1)(s+\alpha)}{2j+3}\right]^{\frac{1}{2}}\langle E, j+1, j, \alpha-1|\\ &-\frac{1}{2(j+1)}[(j-\alpha)(j+\alpha+1)(s-\alpha)(s+\alpha+1)]^{\frac{1}{2}}\langle E, j, j, \alpha+1|\\ &+\frac{1}{2(j+1)}\left[\frac{(j+\alpha+1)(j+\alpha+2)(s-\alpha)(s+\alpha+1)}{2j+3}\right]^{\frac{1}{2}}\langle E, j+1, j, \alpha+1|.\end{aligned} \tag{2.8}$$

Exponentials of the angular momentum operators will prove useful. From (2.6) and (2.7) we obtain, by simple expansion of the exponential as a power series in the scalar β,

$$\exp[i\beta S_3]\,|\mathbf{p}, \gamma\rangle = e^{i\beta\gamma}\,|\mathbf{p}, \gamma\rangle,$$

$$\exp[\beta(S_2 - iS_1)]\,|\mathbf{p}, \gamma\rangle = \sum_{r=0}^{s-\gamma}\frac{\beta^r}{r!}\left[\frac{(s-\gamma)!(s+\gamma+r)!}{(s+\gamma)!(s-\gamma-r)!}\right]^{\frac{1}{2}}|\mathbf{p}, \gamma+r\rangle,$$

$$\exp[\beta(S_2 + iS_1)]\,|\mathbf{p}, \gamma\rangle = \sum_{r=0}^{s+\gamma}\frac{\beta^r}{r!}\left[\frac{(s+\gamma)!(s-\gamma+r)!}{(s-\gamma)!(s+\gamma-r)!}\right]^{\frac{1}{2}}|\mathbf{p}, \gamma-r\rangle, \tag{2.9}$$

$$\langle E, j, m, \alpha|\exp[i\beta J_3] = e^{i\beta m}\langle E, j, m, \alpha|,$$

$$\langle E, j, m, \alpha|\exp[\beta(J_2 - iJ_1)] = \sum_{r=0}^{j+m}\frac{\beta^r}{r!}\left[\frac{(j+m)!(j-m+r)!}{(j-m)!(j+m-r)!}\right]^{\frac{1}{2}}\langle E, j, m-r, \alpha|,$$

$$\langle E, j, m, \alpha|\exp[\beta(J_2 + iJ_1)] = \sum_{r=0}^{j-m}\frac{\beta^r}{r!}\left[\frac{(j-m)!(j+m+r)!}{(j+m)!(j-m-r)!}\right]^{\frac{1}{2}}\langle E, j, m+r, \alpha|. \tag{2.10}$$

An operator identity which will be useful is the following: Let A and B be any two operators. Then

$$e^{-A}Be^{A} = \sum_{n=0}^{\infty}\frac{\{B, A\}^{(n)}}{n!}, \tag{2.11}$$

where $\{B, A\}^{(n)}$ is defined by induction by means of commutators

$$\{B, A\}^{(n)} = [\{B, A\}^{(n-1)}, A], \qquad \{B, A\}^{(0)} = B. \tag{2.11a}$$

3. EXPLICIT E AND γ DEPENDENCE OF THE TRANSFORMATION FUNCTIONS

In the present section we obtain the explicit E and γ dependence of the transformation functions. It will be convenient to introduce the operator T defined by

$$T = \exp[-(P_2 - iP_1)^{-1}(P - P_3)(S_2 - iS_1)] \tag{3.1}$$

and a new set of kets $|\mathbf{p}, \gamma)$ defined by

$$|\mathbf{p}, \gamma) = T\,|\mathbf{p}, \gamma\rangle. \tag{3.2}$$

In (3.1) the operator $P = (P_1^2 + P_2^2 + P_3^2)^{\frac{1}{2}}$. We first find the E and γ dependence of $\langle E, j, m, \alpha|\mathbf{p}, \gamma)$ and then that of $\langle E, j, m, \alpha|\mathbf{p}, \gamma\rangle$ from

$$\langle E, j, m, \alpha\,|\,\mathbf{p}, \gamma\rangle = \langle E, j, m, \alpha|\,T^{-1}\,|\mathbf{p}, \gamma). \tag{3.3}$$

First we note that since H and P_i commute with T it follows from Eqs. (2.6) that

$$H\,|\mathbf{p}, \gamma) = \epsilon\omega(p)\,|\mathbf{p}, \gamma), \qquad P_i\,|\mathbf{p}, \gamma) = p_i\,|\mathbf{p}, \gamma). \tag{3.4}$$

Now, from the first of Eqs. (3.4),

$$\langle E, j, m, \alpha|\,H\,|\mathbf{p}, \gamma) = \epsilon\omega(p)\langle E, j, m, \alpha\,|\,\mathbf{p}, \gamma), \tag{3.5}$$

while from the first of Eqs. (2.7),

$$\langle E, j, m, \alpha|\,H\,|\mathbf{p}, \gamma) = E\langle E, j, m, \alpha\,|\,\mathbf{p}, \gamma). \tag{3.6}$$

Thus from (3.5) and (3.6)

$$(E - \epsilon\omega(p))\langle E, j, m, \alpha \mid \mathbf{p}, \gamma) = 0. \tag{3.7}$$

The general solution of (3.7) is

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma) = \delta(E - \epsilon\omega(p))F(j, m, \alpha \mid \mathbf{p}, \gamma), \tag{3.8}$$

where F is a function of its arguments which is still to be determined. We have thus obtained the E-dependence of the function $\langle E, j, m, \alpha|\mathbf{p}, \gamma)$.

Now from (2.6)

$$\mathbf{P}\cdot\mathbf{J} = \mathbf{P}\cdot\mathbf{S}. \tag{3.9}$$

Thus on using (2.11) we have

$$T^{-1}(\mathbf{P}\cdot\mathbf{J})T = T^{-1}(\mathbf{P}\cdot\mathbf{S})T = \tfrac{1}{2} \times (P_2 - iP_1)(S_2 + iS_1) + PS_3. \tag{3.10}$$

From (3.10), the second of Eqs. (3.4), and the first three equations of (2.6) we find

$$(\mathbf{P}\cdot\mathbf{J})\ |\mathbf{p}, \gamma) = \tfrac{1}{2}(p_2 - ip_1)[(s + \gamma)(s - \gamma + 1)]^{\frac{1}{2}} \times |\mathbf{p}, \gamma - 1) + p\gamma\ |\mathbf{p}, \gamma). \tag{3.11}$$

The simplicity of (3.11) is the principal reason for the introduction of the basis $|\mathbf{p}, \gamma)$. The expression for $(\mathbf{P}\cdot\mathbf{J})\ |\mathbf{p}, \gamma\rangle$ is more complicated.

From (3.11) we have

$$\langle E, j, m, \alpha|\ (\mathbf{P}\cdot\mathbf{J})\ |\mathbf{p}, \gamma) = \tfrac{1}{2}(p_2 - ip_1) \times [(s + \gamma)(s - \gamma + 1)]^{\frac{1}{2}}\langle E, j, m, \alpha \mid \mathbf{p}, \gamma - 1) + p\gamma\langle E, j, m, \alpha \mid \mathbf{p}, \gamma). \tag{3.12}$$

But also from the last of Eqs. (2.7)

$$\langle E, j, m, \alpha|\ (\mathbf{P}\cdot\mathbf{J})\ |\mathbf{p}, \gamma) = k(E)\alpha\langle E, j, m, \alpha \mid \mathbf{p}, \gamma). \tag{3.13}$$

Let us equate the right-hand sides of (3.12) and (3.13). On substituting (3.8) for $\langle E, j, m, \alpha|\mathbf{p}, \gamma)$ we obtain the following recursion relation for $F(j, m, \alpha|\mathbf{p}, \gamma)$ in terms of γ:

$$F(j, m, \alpha \mid \mathbf{p}, \gamma - 1) = \frac{2}{\eta}\frac{(\alpha - \gamma)}{[(s + \gamma)(s - \gamma + 1)]^{\frac{1}{2}}}F(j, m, \alpha \mid \mathbf{p}, \gamma), \tag{3.14}$$

where

$$\eta = (p_2 - ip_1)/p. \tag{3.15}$$

It is clear from (3.14) that $F(j, m, \alpha|\mathbf{p}, \gamma) = 0$ if $\gamma < \alpha$. It is also clear that for $\gamma \geq \alpha$, $F(j, m, \alpha|\mathbf{p}, \gamma)$ can be found from $F(j, m, \alpha|\mathbf{p}, \alpha) = G(j, m, \alpha|\mathbf{p})$ by induction. We thus represent the general solution of the recursion relation (3.14) by

$$F(j, m, \alpha \mid \mathbf{p}, \gamma) = H_\alpha(\gamma)\left(\frac{-\eta}{2}\right)^{\gamma-\alpha} \times \frac{1}{(\gamma - \alpha)!}\left[\frac{(s + \gamma)!(s - \alpha)!}{(s + \alpha)!(s - \gamma)!}\right]^{\frac{1}{2}} \times G(j, m, \alpha \mid \mathbf{p}), \tag{3.16}$$

where $H_\alpha(\gamma)$ is defined by

$$H_\alpha(\gamma) = 0 \quad \text{if} \quad \gamma < \alpha \qquad = 1 \quad \text{if} \quad \gamma \geq \alpha. \tag{3.17}$$

Thus

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma) = \delta(E - \epsilon\omega(p))H_\alpha(\gamma)\left(\frac{-\eta}{2}\right)^{\gamma-\alpha} \times \frac{1}{(\gamma - \alpha)!}\left[\frac{(s + \gamma)!(s - \alpha)!}{(s + \alpha)!(s - \gamma)!}\right]^{\frac{1}{2}} \times G(j, m, \alpha \mid \mathbf{p}). \tag{3.18}$$

Now we wish to use (3.3) to find $\langle E, j, m, \alpha|\mathbf{p}, \gamma)$. Since $(S_2 - iS_1)$ commutes with T we have

$$(S_2 - iS_1)\ |\mathbf{p}, \gamma) = [(s - \gamma)(s + \gamma + 1)]^{\frac{1}{2}}\ |\mathbf{p}, \gamma + 1). \tag{3.19}$$

Thus the operator $(S_2 - iS_1)$ acts on $|\mathbf{p}, \gamma)$ in precisely the same way that it acts on $|\mathbf{p}, \gamma\rangle$. On expanding T^{-1} in a power series in $(S_2 - iS_1)$ and on using (3.19) and the second of Eqs. (3.4), one obtains

$$T^{-1}\ |\mathbf{p}, \gamma) = \sum_{r=0}^{s-\gamma}\left(1 - \frac{p_3}{p}\right)^r\eta^{-r} \times \left[\frac{(s - \gamma)!(s + \gamma + r)!}{(s + \gamma)!(s - \gamma - r)!}\right]^{\frac{1}{2}}\frac{1}{r!}\ |p, \gamma + r). \tag{3.20}$$

We now substitute (3.20) into (3.3), use (3.18) and absorb $[(s - \alpha)!/(s + \alpha)!]^{\frac{1}{2}}$ into the still unknown function $G(j, m, \alpha|\mathbf{p})$.

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(E - \epsilon\omega(p))\eta^{\gamma-\alpha} \times [(s - \gamma)!/(s + \gamma)!]^{\frac{1}{2}}T(s, \gamma, \alpha, p_3/p)G(j, m, \alpha \mid \mathbf{p}), \tag{3.21}$$

where

$$T(s, \gamma, \alpha, x) = \sum_{r=\gamma}^{s} H_\alpha(r)(1 - x)^{r-\gamma}\left(\frac{-1}{2}\right)^{r-\alpha} \times \frac{1}{(r - \alpha)!(r - \gamma)!}\frac{(s + r)!}{(s - r)!}. \tag{3.22}$$

We now show, however, that

$$T(s, \gamma, \alpha, x) = (-\tfrac{1}{2})^{\gamma-\alpha}[(s + \gamma)!/(s - \alpha)!]S(s, \gamma, \alpha, x). \tag{3.23}$$

Our proof will be by induction.

From (1.11) of Part I

$$S(s, s, \alpha, x) = 1. \tag{3.24}$$

From (3.22) which defines $T(s, \gamma, \alpha, x)$ we find

$$T(s, s, \alpha, x) = (-\tfrac{1}{2})^{s-\alpha}[(2s)!/(s-\alpha)!]. \tag{3.25}$$

Thus (3.23) is valid for $\gamma = s$.

Let us assume that (3.23) holds for a particular value of γ and prove that it will then hold for $\gamma - 1$.

We can easily show from (3.22) that

$$dT(s, \gamma, \alpha, x)/dx = -T(s, \gamma + 1, \alpha, x) \tag{3.26}$$

for all γ.

Then from (3.26) and (3.23) of the present paper and (2.6) of Part I

$$\frac{d}{dx} T(s, \gamma - 1, \alpha, x)$$

$$= \left(-\frac{1}{2}\right)^{\gamma-\alpha-1} \frac{(s+\gamma-1)!}{(s-\alpha)!} \frac{d}{dx} S(s, \gamma - 1, \alpha, x). \tag{3.27}$$

Then integrating with respect to x,

$$T(s, \gamma - 1, \alpha, x)$$

$$= (-\tfrac{1}{2})^{\gamma-\alpha-1}[(s+\gamma-1)!/(s-\alpha)!]$$

$$\times S(s, \gamma - 1, \alpha, x) + K(s, \gamma, \alpha), \tag{3.28}$$

where K is a constant of integration. If we can show that $K = 0$ we shall have verified (3.23) for all γ by induction.

From (3.22) we have

$$T(s, \gamma, \alpha, 1)$$

$$= H_\alpha(\gamma)(-\tfrac{1}{2})^{\gamma-\alpha}[(s+\gamma)!/(s-\gamma)!(\gamma-\alpha)!], \tag{3.29}$$

while from (1.10) Part I

$$S(s, \gamma, \alpha, 1)$$

$$= H_\alpha(\gamma)[(s-\alpha)!/(\gamma-\alpha)!(s-\gamma)!]. \tag{3.30}$$

We then set $x = 1$ in (3.28) and use (3.29) and (3.30) in the resulting expression. It is seen that $K(s, \gamma, \alpha) = 0$. Thus (3.23) has been proved.

On using (3.23) in (3.21) we obtain the principal expression of this section, after absorbing the factor $(s-\alpha)!$ into $G(j, m, \alpha|\mathbf{p})$.

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle$$

$$= \delta(E - \epsilon\omega(p))(-\tfrac{1}{2}\eta)^{\gamma-\alpha}[(s-\gamma)!(s+\gamma)!]^{\frac{1}{2}}$$

$$\times S(s, \gamma, \alpha, p_3/p)G(j, m, \alpha \mid \mathbf{p}), \tag{3.31}$$

where the function G is still to be determined.

4. USE OF THE OPERATOR J_3

We now evaluate $\langle E, j, m, \alpha \,|\exp[i\nu J_3]|\, \mathbf{p}, \gamma\rangle$ in two ways and further simplify the expression for the transformation function.

It will noted that

$$\exp[i\nu J_3]\,|\mathbf{p}, \gamma\rangle = e^{i\nu\gamma}\,|\mathbf{p}', \gamma\rangle, \tag{4.1}$$

where $\mathbf{p}'$ is given in terms of $\mathbf{p}$ by

$$\begin{aligned} p_1' &= p_1 \cos\nu + p_2 \sin\nu, \\ p_2' &= p_2 \cos\nu - p_1 \sin\nu, \\ p_3' &= p_3. \end{aligned} \tag{4.2}$$

Relation (4.1) can be verified by differentiating with respect to ν and then setting $\nu = 0$. Then one has the sixth equation of (2.6).

Also we have

$$\langle E, j, m, \alpha| \exp[i\nu J_3] = e^{i\nu m}\langle E, j, m, \alpha|. \tag{4.3}$$

From (4.1) and (4.3) we obtain

$$e^{i\nu m}\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = e^{i\nu\gamma}\langle E, j, m, \alpha \mid \mathbf{p}', \gamma\rangle, \tag{4.4}$$

which, on substituting for (3.31), leads to an equation for the yet undetermined function G:

$$G(j, m, \alpha \mid \mathbf{p}) = e^{i(\alpha-m)\nu}G(j, m, \alpha \mid \mathbf{p}'). \tag{4.5}$$

We can easily find a solution of this equation

Let us pick ν such that

$$\begin{aligned} \cos\nu &= p_1/(p_1^2 + p_2^2)^{\frac{1}{2}}, \\ \sin\nu &= p_2/(p_1^2 + p_2^2)^{\frac{1}{2}}, \end{aligned} \tag{4.6}$$

and hence

$$e^{i\nu} = i\eta(1 - p_3^2/p^2)^{-\frac{1}{2}}, \tag{4.7}$$

$$e^{i(\alpha-m)\nu} = (i\eta)^{\alpha-m}(1 - p_3^2/p^2)^{\frac{1}{2}(m-\alpha)}. \tag{4.8}$$

Also

$$G(j, m, \alpha \mid \mathbf{p}') = G(j, m, \alpha \mid (p_1^2 + p_2^2)^{\frac{1}{2}}, 0, p_3). \tag{4.9}$$

Thus $G(j, m, \alpha|\mathbf{p}')$ may be taken as a function of the variables j, m, α, p, p_3. Let us define $F(j, m, \alpha|p, p_3)$ by

$$F(j, m, \alpha, \mid p, p_3)$$

$$= (1 - p_3^2/p^2)^{\frac{1}{2}(\alpha-m)}G(j, m, \alpha \mid \mathbf{p}'). \tag{4.10}$$

Then the solution of (4.5) is

$$G(j, m, \alpha \mid \mathbf{p}) = (i\eta)^{\alpha-m}F(j, m, \alpha \mid p, p_3) \tag{4.11}$$

and expression (3.31) for the transformation function simplifies down to

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(E - \epsilon\omega(p))(-\tfrac{1}{2})^{\gamma-\alpha}\eta^{\gamma-m}(i)^{\gamma-m}[(s-\gamma)!\,(s+\gamma)!]^{\frac{1}{2}} \times S\left(s, \gamma, \alpha, \frac{p_3}{p}\right)F(j, m, \alpha \mid p, p_3), \tag{4.12}$$

where the function F is still to be determined.

5. USE OF THE OPERATOR $J_2 + iJ_1$

We shall use

$$\exp[\beta(J_2 + iJ_1)] \mid \mathbf{p}, \gamma\rangle = \exp[\beta(S_2 + iS_1)] \mid \mathbf{p}', \gamma\rangle, \tag{5.1}$$

where $\mathbf{p}'$ is given in terms of $\mathbf{p}$ by

$$\begin{aligned} p_1' &= (1 + \tfrac{1}{2}\beta^2)p_1 - \tfrac{1}{2}i\beta^2 p_2 + i\beta p_3, \\ p_2' &= -\tfrac{1}{2}i\beta^2 p_1 + (1 - \tfrac{1}{2}\beta^2)p_2 + \beta p_3, \\ p_3' &= -i\beta p_1 - \beta p_2 + p_3. \end{aligned} \tag{5.2}$$

In (5.1) the kets $|\mathbf{p}, \gamma\rangle$, which were defined originally for real values of p_i only, have been analytically continued to complex values of p_i through the use of complex rotations. Equation (5.1) can be verified by differentiating with respect to β and setting $\beta = 0$. One obtains the last of Eqs. (2.6).

One can show that on defining $(p')^2$ by

$$(p')^2 = (p_1')^2 + (p_2') + (p_3')^2, \tag{5.3}$$

that

$$(p')^2 = p^2 \tag{5.4}$$

and hence

$$p' = p \tag{5.5}$$

by taking the proper square root.

From (5.1) and the third of Eqs. (2.9)

$$\langle E, j, m, \alpha| \exp[\beta(J_2 + iJ_1)] \,|\mathbf{p}, -s\rangle = \langle E, j, m, \alpha \mid \mathbf{p}', -s\rangle. \tag{5.6}$$

But from the last of Eqs. (2.10)

$$\langle E, j, m, \alpha| \exp[\beta(J_2 + iJ_1)] \,|\mathbf{p}, -s\rangle = \sum_{r=0}^{j-m} \frac{\beta^r}{r!}\left[\frac{(j-m)!\,(j+m+r)!}{(j+m)!\,(j-m-r)!}\right]^{\frac{1}{2}} \times \langle E, j, m+r, \alpha \mid \mathbf{p}, -s\rangle. \tag{5.7}$$

We can therefore equate the right-hand sides of (5.6) and (5.7).

We now substitute (4.12) into the resulting equation using the fact that

$$S(s, -s, \alpha, x) = (-1)^{s+\alpha}2^{-2s}(1-x)^{s+\alpha}(1+x)^{s-\alpha} \tag{5.8}$$

which follows from Eq. (1.10) of Part I. We shall also set $\beta = p_3/(p_2 + ip_1)$ in this equation. We note that with this choice of β

$$p_3' = 0,$$

$$\eta' = \frac{p_2' - ip_1'}{p'} = \eta\left(1 - \frac{p_3}{p}\right)^{-1}\left(1 + \frac{p_3}{p}\right)^{-1}. \tag{5.9}$$

Indeed the first of equations (5.9) motivated our choice of β.

Let us define $G(j, m, \alpha|p)$ by

$$G(j, m, \alpha \mid p) = \left[\frac{(j-m)!}{(j+m)!}\right]^{\frac{1}{2}}F(j, m, \alpha \mid p, 0). \tag{5.10}$$

The result of these substitutions into the equation obtained from (5.6) and (5.7) is the following equation for the hitherto unknown function $F(j, m, \alpha|p, p_3)$, namely

$$\sum_{r=0}^{j-m}\frac{1}{r!}\left(\frac{p_3}{p}\right)^r(-i)^r\left[\frac{(j+m+r)!}{(j-m-r)!}\right]^{\frac{1}{2}} \times \left(1 - \frac{p_3}{p}\right)^{-(m-\alpha+r)}\left(1 + \frac{p_3}{p}\right)^{-(m+\alpha+r)} \times F(j, m+r, \alpha \mid p, p_3) = G(j, m, \alpha \mid p). \tag{5.11}$$

To solve (5.11) we shall introduce the function $H(m)$ defined by

$$H(m) = \left[\frac{(j+m)!}{(j-m)!}\right]^{\frac{1}{2}}\left(1 - \frac{p_3}{p}\right)^{-(m-\alpha)} \times \left(1 + \frac{p_3}{p}\right)^{-(m+\alpha)}\left(-\frac{ip_3}{p}\right)^m F(j, m, \alpha \mid p, p_3). \tag{5.12}$$

In defining $H(m)$ we have suppressed all variables except m since Eq. (5.11) is a set of equations in this variable only. Likewise we define

$$G(m) = (-ip_3/p)^m G(j, m, \alpha \mid p). \tag{5.13}$$

Then the equation (5.11) becomes an equation for $H(m)$ which is

$$\sum_{r=0}^{j-m}\frac{H(r+m)}{r!} = G(m). \tag{5.14}$$

In a certain sense, the solution of (5.14) represents a kernel of the problem of finding the transformation function.

The solution of (5.14) is

$$H(m) = \sum_{r=0}^{j-m}\frac{G(j-r)}{(j-m-r)!}(-1)^{j-m-r}. \tag{5.15}$$

To verify that (5.15) is the solution is somewhat tricky. But to avoid a lengthy proof we shall leave the verification to the reader.

We now have the solution of (5.11) which is

$$F(j, m, \alpha \mid p, p_3) = \left[\frac{(j-m)!}{(j+m)!}\right]^{\frac{1}{2}}\left(1-\frac{p_3}{p}\right)^{m-\alpha}\left(1+\frac{p_3}{p}\right)^{m+\alpha} \times \sum_{r=0}^{j-m}\left(i\frac{p_3}{p}\right)^{j-m-r}\frac{G(j, m-r, \alpha \mid p)}{(j-m-r)!}. \quad (5.16)$$

Equation (5.16) can be written

$$F(j, m, \alpha \mid p, p_3) = \left[\frac{(j-m)!}{(j+m)!}\right]^{\frac{1}{2}}\left(1-\frac{p_3}{p}\right)^{m-\alpha} \times \left(1+\frac{p_3}{p}\right)^{m+\alpha} R\left(j, m, \alpha, \frac{p_3}{p}, p\right), \quad (5.17)$$

where

$$R(j, m, \alpha, x, p) = \sum_{r=0}^{j-m}\frac{(ix)^r}{r!}D(j, m+r, \alpha \mid p), \quad (5.17a)$$

with

$$D(j, m, \alpha \mid p) = C(j, j+2m, \alpha \mid p). \quad (5.17b)$$

Thus from (4.12) we can write the final formula of the present section,

$$\langle E, j, m, \alpha \mid p, \gamma\rangle = \delta(E-\epsilon\omega(p))(-\tfrac{1}{2})^{\gamma-\alpha} \times i^{\alpha-m}\eta^{\gamma-m}[(s-\gamma)!\,(s+\gamma)!]^{\frac{1}{2}} \times \left[\frac{(j-m)!}{(j+m)!}\right]^{\frac{1}{2}}\left(1-\frac{p_3}{p}\right)^{m-\alpha}\left(1+\frac{p_3}{p}\right)^{m+\alpha} \times S\left(s, \gamma, \alpha, \frac{p_3}{p}\right)R\left(j, m, \alpha, \frac{p_3}{p}, p\right). \quad (5.18)$$

6. USE OF THE OPERATOR $J_2 - iJ_1$

We note first that

$$\exp[\beta(J_2 - iJ_1)]\,|\mathbf{p}, \gamma\rangle = \exp[\beta(S_2 - iS_1)]\,|\mathbf{p}', \gamma\rangle, \quad (6.1)$$

where $\mathbf{p}'$ is related to $\mathbf{p}$ by the following expression:

$$\begin{aligned} p_1' &= (1+\tfrac{1}{2}\beta^2)p_1 + \tfrac{1}{2}i\beta^2 p_2 + i\beta p_3, \\ p_2' &= \tfrac{1}{2}i\beta^2 p_1 + (1-\tfrac{1}{2}\beta^2)p_2 - \beta p_3, \\ p_3' &= -i\beta p_1 + \beta p_2 + p_3. \end{aligned} \quad (6.2)$$

Expression (6.1) can be verified by differentiating with respect to β and setting $\beta = 0$. Then one obtains the seventh of Eqs. (2.6).

On using (6.1), and the second of Eqs. (2.9),

$$\langle E, j, -j, \alpha|\exp[\beta(J_2 - iJ_1)]\,|\mathbf{p}, s\rangle = \langle E, j, -j, \alpha \mid \mathbf{p}', s\rangle. \quad (6.3)$$

But on using the second of Eq. (2.10) we have

$$\langle E, j, -j, \alpha|\exp[\beta(J_2 - iJ_1)]\,|\mathbf{p}, s\rangle = \langle E, j, -j, \alpha \mid \mathbf{p}, s\rangle. \quad (6.4)$$

Thus

$$\langle E, j, -j, \alpha \mid \mathbf{p}, s\rangle = \langle E, j, -j, \alpha \mid \mathbf{p}', s\rangle. \quad (6.5)$$

On using

$$S(s, s, \alpha, x) = 1, \quad (6.6)$$

which follows from (1.11) of Part I, we substitute (5.18) into (6.5).

We then set

$$\beta = -(p_3/p)\eta^{-1} \quad (6.7)$$

which leads to

$$p_3' = 0, \quad p' = p, \quad \eta' = \eta. \quad (6.8)$$

We obtain the following relation for $R(j, -j, \alpha, p_3/p, p)$:

$$R(j, -j, \alpha, p_3/p, p) = (1-p_3/p)^{j+\alpha} \times (1+p_3/p)^{j-\alpha}R(j, -j, \alpha, 0, p). \quad (6.9)$$

Let us define $F(j, \alpha, p)$ by

$$F(j, \alpha, p) = (-1)^{j+\alpha}R(j, -j, \alpha, 0, p). \quad (6.10)$$

Then

$$R(j, -j, \alpha, x, p) = (1+x)^{j-\alpha}(1-x)^{j+\alpha} \times (-1)^{j+\alpha}F(j, \alpha, p). \quad (6.11)$$

Equation (6.11) leads to the following theorem:

$$R(j, m, \alpha, x, p) = 2^{j-m}(j+m)!\,S(j, m, \alpha, x)F(j, \alpha, p). \quad (6.12)$$

Equation (6.12) is proved by induction by using

$$dR(j, m, \alpha, x, p)/dx = iR(j, m+1, \alpha, x, p), \quad (6.13)$$

which follows from (5.17a) and by using Eq. (1.10) of Part I.

Thus we have the following formula for the transformation function after absorbing some factors which depend only on j and α into $F(j, \alpha, p)$,

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(E-\epsilon\omega(p))(-\tfrac{1}{2})^{m+\gamma} \times \eta^{\gamma-m}[(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!]^{\frac{1}{2}} \times \left(1-\frac{p_3}{p}\right)^{m-\alpha}\left(1+\frac{p_3}{p}\right)^{m+\alpha}S\left(j, m, \alpha, \frac{p_3}{p}\right) \times S\left(s, \gamma, \alpha, \frac{p_3}{p}\right)F(j, \alpha, p). \quad (6.14)$$

7. USE OF THE OPERATOR P_3

On using the fourth equation of (2.6) and the fifth equation of (2.7) we obtain on evaluating $\langle E, j, j, \alpha|P_3|\mathbf{p}, s\rangle$ two ways:

$$k(E)\Big[\frac{\alpha}{j+1}\langle E, j, j, \alpha \mid p, s\rangle + \frac{1}{j+1}\Big[\frac{(2j+1)(j-\alpha+1)(j+\alpha+1)}{2j+3}\Big]^{\frac{1}{2}} \times \langle E, j+1, j, \alpha \mid p, s\rangle\Big] = p_3\langle E, j, j, \alpha \mid \mathbf{p}, s\rangle. \tag{7.1}$$

We now substitute (6.14) into (7.1) and use (6.6) and

$$S(j+1, j, \alpha, x) = (j+1)x - \alpha, \tag{7.2}$$

which follows from (1.11) of Part I to obtain the following recursion relation for $F(j, \alpha, p)$:

$$F(j+1, \alpha, p) = \Big[\frac{2j+3}{(2j+1)(j-\alpha+1)(j+\alpha+1)}\Big]^{\frac{1}{2}} F(j, \alpha, p). \tag{7.3}$$

This equation simplifies somewhat if we use $D(j, \alpha, p)$ defined by

$$D(j, \alpha, p) = (2j+1)^{-\frac{1}{2}}F(j, \alpha, p). \tag{7.4}$$

Then

$$D(j+1, \alpha, p) = D(j, \alpha, p)[(j-\alpha+1)(j+\alpha+1)]^{-\frac{1}{2}}. \tag{7.5}$$

Let

$$G(\alpha, p) = [(2\,|\alpha|)!]^{\frac{1}{2}}D(|\alpha|\,, \alpha, p). \tag{7.6}$$

Then one can show that the general solution of (7.4) is

$$D(j, \alpha, p) = [(j-\alpha)!(j+\alpha)!]^{-\frac{1}{2}}G(\alpha, p), \tag{7.7}$$

and hence

$$F(j, \alpha, p) = [(2j+1)/(j-\alpha)!\,(j+\alpha)!]^{\frac{1}{2}}G(\alpha, p). \tag{7.8}$$

Thus we come to the final formula of this section, namely that, for the transformation function,

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(E - c\omega(p))(-\tfrac{1}{2})^{m+\gamma}\eta^{\gamma-m} \times \Big[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(j-\alpha)!\,(j+\alpha)!}\Big]^{\frac{1}{2}} \times \Big(1-\frac{p_3}{p}\Big)^{m-\alpha}\Big(1+\frac{p_3}{p}\Big)^{m+\alpha} S\Big(j, m, \alpha, \frac{p_3}{p}\Big) \times S\Big(s, \gamma, \alpha, \frac{p_3}{p}\Big)G(\alpha, p). \tag{7.9}$$

8. USE OF THE OPERATOR T_3

The operator T_3 is defined by Eq. (2.13) of Ref. 4 as being given by

$$T_3 = S_3 - (\mathbf{P}\cdot\mathbf{J})P^{-2}P_3, \tag{8.1}$$

where $p^{-2} = (\mathbf{P}\cdot\mathbf{P})^{-1}$.

But on using the first of Eqs. (2.6), the fourth of equations (2.6) and the last of Eqs. (2.7) and the fact that

$$k(E)\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = p\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle, \tag{8.2}$$

which follows from the presence of the Dirac delta in the transformation function we have

$$\langle E, j, j, \alpha|\, T_3\,|\mathbf{p}, s\rangle = \langle E, j, j, \alpha|\, S_3 - (\mathbf{P}\cdot\mathbf{J})P^{-2}P_3\,|\mathbf{p}, s\rangle = s\langle E, j, j, \alpha \mid \mathbf{p}, s\rangle - (p_3/p)\alpha\langle E, j, j, \alpha \mid \mathbf{p}, s\rangle. \tag{8.3}$$

But in the angular momentum representation we can use Eq. (2.8) of the present paper to obtain another expression for $\langle E, j, j, \alpha\,|T_3|\,\mathbf{p}, s\rangle$. Equating the two expressions for this quantity and substituting (7.9) into the resulting equation leads to the following equation for $G(\alpha, p)$—after a considerable amount of tedious but more-or-less straightforward reduction—

$$[(s-\alpha+1)(s+\alpha)]^{\frac{1}{2}}G(\alpha-1, p) + [(s-\alpha)(s+\alpha+1)]^{\frac{1}{2}}G(\alpha+1, p) = -2sG(\alpha, p), \tag{8.4}$$

which is a recursion relation in the variable α. It is readily verified that the solution of this equation is

$$G(\alpha, p) = (-1)^{s+\alpha}[(s+\alpha)!\,(s-\alpha)!]^{-\frac{1}{2}}G(p), \tag{8.5}$$

where

$$G(p) = [(2s)!]^{\frac{1}{2}}G(-s, p). \tag{8.6}$$

Thus substituting (8.5) into (7.9) we have the following expression for the transformation function:

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(E - c\omega(p)) \times (-1)^{\alpha+\gamma+m+s}(\tfrac{1}{2})^{\gamma+m}\eta^{\gamma-m} \times \Big[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\Big]^{\frac{1}{2}} \times \Big(1-\frac{p_3}{p}\Big)^{m-\alpha}\Big(1+\frac{p_3}{p}\Big)^{m+\alpha} \times S(j, m, \alpha, p_3/p)S(s, \gamma, \alpha, p_3/p)G(p). \tag{8.7}$$

9. USE OF THE OPERATOR $\mathcal{J}_3$

From the second, third, and the last of Eqs. (2.6) we have

$$\langle E, s, s, s|\ \mathcal{J}_3\ |\mathbf{p}, s\rangle = i\epsilon\omega(p)(\partial/\partial p_3) \times \langle E, s, s, s \mid \mathbf{p}, s\rangle + i\epsilon/2[\omega(p) + \mu] \times (p_2 - ip_1)(2s)^{\frac{1}{2}}\langle E, s, s, s \mid \mathbf{p}, s-1\rangle. \quad (9.1)$$

It will be convenient to use δ-function relation

$$\delta(E - \epsilon\omega(p)) = [\omega(p)/p]\delta(k(E) - p) \quad (9.2)$$

which follows from familiar δ-function identities.

Then on absorbing the factor $\omega(p)/p$ into the still unknown function $D(p)$, our expression (8.7) for the transformation function becomes

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = \delta(k(E) - p) \times (-1)^{\alpha+\gamma+m+s}(\tfrac{1}{2})^{\gamma+m}\eta^{\gamma-m} \times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}} \times \left(1 - \frac{p_3}{p}\right)^{m-\alpha}\left(1 + \frac{p_3}{p}\right)^{m+\alpha} S\left(j, m, \alpha, \frac{p_3}{p}\right) \times S\left(s, \gamma, \alpha, \frac{p_3}{p}\right)D(p). \quad (9.3)$$

Then from (9.1), (9.3), (6.6), and (7.2) we obtain

$$\begin{aligned}\langle E, s, s, s|\ \mathcal{J}_3\ |\mathbf{p}, s\rangle &= i(\tfrac{1}{2})^{2s}(2s+1)^{\frac{1}{2}}\epsilon\left(1 + \frac{p_3}{p}\right)^{2s} \times \left\{\delta'(k-p)\frac{p_3}{p}\omega(p)D(p)\right. \\ &+ \delta(k-p)\left[-\frac{2s}{p}\left(1 - \frac{p_3}{p}\right)D(p)\omega(p)\right. \\ &\left.\left. + \frac{ps}{\omega(p)+\mu}\left(1 - \frac{p_3}{p}\right)D(p) - \omega(p)D'(p)\frac{p_3}{p}\right]. \quad (9.4)\right.\end{aligned}$$

In (9.3) $k = k(E)$ and the prime on a function means a derivative with respect to the argument.

We now want to evaluate $\langle E, s, s, s\ |\mathcal{J}_3|\ \mathbf{p}, s\rangle$ a second way using the sixth of Eqs. (2.7). It is convenient to replace the variable E by k. We note

$$\partial/\partial E = (E/k)\partial/\partial k. \quad (9.5)$$

Thus the sixth of Eqs. (2.7) leads to

$$\begin{aligned}\langle E, s, s, s|\ \mathcal{J}_3\ |\mathbf{p}, s\rangle &= [i/(s+1)]\left\{sE\frac{\partial}{\partial k}\langle E, s, s, s \mid \mathbf{p}, s\rangle\right. \\ &\left. + (\epsilon\mu/2k)2s\langle E, s, s, s-1 \mid \mathbf{p}, s\rangle\right\} \\ &+ \frac{i}{s+1}[2s+3]^{-\frac{1}{2}}\left\{[2s+1]^{\frac{1}{2}}\left[(s+1)\frac{E}{k}\right.\right. \\ &\left.\left. + E\frac{\partial}{\partial k}\right]\langle E, s+1, s, s \mid \mathbf{p}, s\rangle\right. \\ &\left. + \frac{\epsilon\mu}{2k}(4s)^{\frac{1}{2}}\langle E, s+1, s, s-1 \mid \mathbf{p}, s\rangle\right\}. \quad (9.6)\end{aligned}$$

We now use (9.3) in (9.6). We shall also use (6.6), (7.2), and the following symbolic function relation

$$E(k)\delta'(k-p) = \epsilon\omega(p)\delta'(k-p) - [\epsilon p/\omega(p)]\delta(k-p), \quad (9.7)$$

which follows from the well-known symbolic function identity

$$f(x)\delta'(x-y) = f(y)\delta(x-y) - f'(y)\delta(x-y), \quad (9.8)$$

to obtain—after a considerable amount of reduction—

$$\begin{aligned}\langle E, s, s, s|\ \mathcal{J}_3\ |\mathbf{p}, s\rangle &= i\epsilon(\tfrac{1}{2})^{2s}(2s+1)^{\frac{1}{2}}\left(1 + \frac{p_3}{p}\right)^{2s}D(p) \\ &\times \left\{\delta'(k-p)\omega(p)\frac{p_3}{p} - \delta(k-p)\frac{1}{p\omega(p)}\right. \\ &\left. \times \left[s\left(1 - \frac{p_3}{p}\right)(\omega^2 + \mu\omega) - \mu^2\frac{p_3}{p}\right]\right\}. \quad (9.9)\end{aligned}$$

When we compare (9.9) with (9.4) we obtain a differential equation for the yet unknown function $D(p)$:

$$D'(p) = -\mu^2 D(p)/p(\mu^2 + p^2). \quad (9.10)$$

It is easy to integrate this equation. The solution is

$$D(p) = C\omega(p)/p, \quad (9.11)$$

where C is constant.

Thus the transformation function is now given by

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = C\delta(k-p)[\omega(p)/p] \times (-1)^{\alpha+\gamma+m+s}(\tfrac{1}{2})^{\gamma+m}\eta^{\gamma-m} \times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}} \times \left(1 - \frac{p_3}{p}\right)^{m-\alpha}\left(1 + \frac{p_3}{p}\right)^{m+\alpha} \times S(j, m, \alpha, p_3/p)S(s, \gamma, \alpha, p_3/p). \quad (9.12)$$

From (9.2) this result may also be written

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = C\delta(E - \epsilon\omega(p))$$
$$\times (-1)^{\alpha+\gamma+m+s}(\tfrac{1}{2})^{\gamma+m}\eta^{\gamma-m}$$
$$\times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}}$$
$$\times \left(1-\frac{p_3}{p}\right)^{m-\alpha}\left(1+\frac{p_3}{p}\right)^{m+\alpha}$$
$$\times S\left(j, m, \alpha, \frac{p_3}{p}\right)S\left(s, \gamma, \alpha, \frac{p_3}{p}\right). \tag{9.13}$$

10. USE OF THE NORMALIZATION CONDITION

We now need only obtain the constant C to give the transformation function uniquely. We shall use the normalization condition

$$\langle E, j, m, \alpha \mid E', j', m', \alpha'\rangle$$
$$= \sum_{\gamma=-s}^{s}\int \langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle \frac{d\mathbf{p}}{\omega(p)}$$
$$\times \langle p, \gamma \mid E', j', m', \alpha'\rangle$$
$$= k(E)\delta(E - E')\delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}, \tag{10.1}$$

which follows from the completeness relations (2.3). In particular, let us take $j = j' = m = m' = \alpha = \alpha' = s$ in (10.1) and substitute (9.13) in the resulting equation. After a certain amount of reduction one is led to the following equation:

$$4\pi\,|C|^2\,\frac{(2s+1)!}{(4s+1)!}\sum_{\gamma=-s}^{s}\frac{(3s+\gamma)!}{(s+\gamma)!} = 1. \tag{10.2}$$

But from

$$\sum_{q=0}^{r}\frac{(r+q)!}{q!} = \frac{(2r+1)!}{(r+1)!}, \tag{10.3}$$

which can be proved in a variety of ways, we can show by a suitable change of variable of summation

$$\sum_{\gamma=-s}^{s}\frac{(3s+\gamma)!}{(s+\gamma)!} = \frac{(4s+1)!}{(2s+1)!}. \tag{10.4}$$

Thus, choosing C real and positive (which choice we may always make),

$$C = 1/2\sqrt{\pi}. \tag{10.5}$$

Finally, we write

$$\eta^{\gamma-m} = (i)^{m-\gamma}\left(\frac{p_1 - ip_2}{p}\right)^{m-\gamma}$$
$$\times \left(1-\frac{p_3}{p}\right)^{\gamma-m}\left(1+\frac{p_3}{p}\right)^{\gamma-m}. \tag{10.6}$$

Then on substituting (10.5) and (10.6) into (9.12) and (9.13) we obtain as our final expression for the transformation function

$$\langle E, j, m, \alpha \mid \mathbf{p}, \gamma\rangle = (1/\sqrt{\pi})\delta(E - \epsilon\omega(p))$$
$$\times (-1)^{\alpha+\gamma+m+s}(i)^{m-\gamma}(\tfrac{1}{2})^{\gamma+m+1}\left(\frac{p_1 - ip_2}{p}\right)^{m-\gamma}$$
$$\times \left(1-\frac{p_3}{p}\right)^{\gamma-\alpha}\left(1+\frac{p_3}{p}\right)^{\gamma+\alpha}$$
$$\times \left[\frac{(s-\gamma)!\,(s+\gamma)!\,(j-m)!\,(j+m)!\,(2j+1)}{(s-\alpha)!\,(s+\alpha)!\,(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}}$$
$$\times S\left(j, m, \alpha, \frac{p_3}{p}\right)S\left(s, \gamma, \alpha, \frac{p_3}{p}\right). \tag{10.7}$$

Equation (10.7) is then substituted into (2.4) and (2.5) to obtain (1.3) and (1.4).

Homogeneous Solutions of the Einstein–Maxwell Equations

ISTVÁN OZSVÁTH
Southwest Center for Advanced Studies, Dallas, Texas
(Received 30 November 1964)

In this paper the solutions of the Einstein–Maxwell equations are investigated under the assumption that the metric of the space–time and the electromagnetic field are invariant under the transformations of a four-parametric, simply transitive group.

The results can be summarized as follows: In the case of null electromagnetic fields there are two different possibilities; If $\Lambda = 0$, all the solutions are Robinson waves; if $\Lambda \neq 0$, there exists only one solution, first given here by (6.26). There exist no other solutions for null electromagnetic fields. In the case of nonnull electromagnetic fields two solutions are found. One metric is known having been first given by Robinson; we give a new solution of type I. The question as to whether there are solutions different from these remains open.

1. INTRODUCTION

In a previous paper[1] we developed a method to find homogeneous solutions of the Einstein field equations with incoherent matter. We now want to apply this method to find homogeneous solutions of the Einstein–Maxwell equations. In order to keep this introduction short we shall assume that the reader is familiar with the first two sections of that paper.

The Einstein–Maxwell equations are

$$R_{jk} = \Lambda g_{jk} + T_{jk}, \tag{1.1a}$$

$$T_{jk} = 2(F_{jl}F_k{}^l - \tfrac{1}{4}g_{jk}F_{lm}F^{lm}), \tag{1.1b}$$

$$F^{ik}{}_{;k} = 0, \qquad F^{*ik}{}_{;k} = 0, \tag{1.1c}$$

$$F_{jk} = -F_{kj}, \tag{1.1d}$$

where Λ is the cosmological constant, and the semicolon denotes covariant derivation with respect to the metric g_{jk}. This problem in general has a large literature which we will not review here. We only mention Witten's work.[2]

We call a four-dimensional normal hyperbolic Riemannian space R_4 (signature -2) homogeneous, if there exists a four-parametric simply transitive group G_4 of transformations, which leaves the metric g_{jk} invariant. We have seen in Ref. 1 that in the case of a homogeneous space–time we can always find four covariant vector fields

$$e^a{}_j = e^a{}_j(x^k), \tag{1.2}$$

where $j = 0, 1, 2, 3$ is the coordinate index and $a = 0, 1, 2, 3$ labels the vectors, called reciprocal vectors of G_4, such that

$$g_{jk} = e^a{}_j g_{ab} e^b{}_k \tag{1.3}$$

where

$$g_{ab} = \text{diag}\,(+1, -1, -1, -1), \tag{1.4}$$

and the equations

$$e^a{}_{j,k} - e^a{}_{k,j} = -C_{bc}{}^a e^b{}_j e^c{}_k \tag{1.5}$$

and the condition

$$\det(e^a{}_j) \neq 0 \tag{1.6}$$

are satisfied. The constants $C_{ab}{}^c = -C_{ab}{}^c$ are the structure constants of G_4 satisfying the Jacobi identities

$$C_{ab}{}^f C_{fc}{}^d + C_{bc}{}^f C_{fa}{}^d + C_{ca}{}^f C_{fb}{}^d = 0. \tag{1.7}$$

With the help of the equations

$$e^a{}_j e^j{}_b = \delta^a{}_b \quad \text{or} \quad e^a{}_j e^k{}_a = \delta^k{}_j, \tag{1.8}$$

we define the contravariant vector fields $e^j{}_a$, called invariant vectors of G_4 satisfying the equations

$$e^j{}_a e^k{}_{b,j} - e^j{}_b e^k{}_{a,j} = -C_{ab}{}^c e^j{}_c \tag{1.9}$$

and the condition

$$\det(e^j{}_a) \neq 0. \tag{1.10}$$

The transformations of G_4 carry the vector fields $e^a{}_j$ and $e^j{}_a$ into themselves.[3] We have therefore an invariant tetrad of covariant vectors $e^a{}_j$ and an invariant tetrad of contravariant vectors $e^j{}_a$ in each point of our Riemannian space R_4, and because of the conditions (1.6) and (1.10) it is possible to assign unique tetrad components to each tensor of R_4. The *invariant* tensors, such as for example R_{jk}, T_{jk}, and F_{jk}, have *constant* tetrad components, given by

$$R_{ab} = R_{jk}e^j{}_a e^k{}_b, \qquad T_{ab} = T_{jk}e^j{}_a e^k{}_b, \qquad F_{ab} = F_{jk}e^j{}_a e^k{}_b, \tag{1.11}$$

[1] I. Ozsváth, J. Math. Phys. **6**, 590 (1965).

[2] L. Witten, "A Geometric Theory of the Electromagnetic and Gravitational Fields," in *Gravitation*, edited by L. Witten (John Wiley & Sons, Inc., New York, 1962).

[3] L. P. Eisenhart, *Continuous Groups of Transformations*, (Princeton University Press, Princeton, New Jersey 1933).

and therefore

$$R_{ik} = R_{ab}e^a{}_i e^b{}_k, \qquad T_{ik} = T_{ab}e^a{}_i e^b{}_k, \qquad F_{ik} = F_{ab}e^a{}_i e^b{}_k. \tag{1.12}$$

The covariant derivative of an invariant tensor can be expressed in the tetrad. For example:

$$u_{k;l} = u_{a:b}e^a{}_i e^b{}_k \quad \text{or} \quad F_{ik;l} = F_{ab:c}e^a{}_i e^b{}_k e^c{}_l, \tag{1.13}$$

where

$$u_{a:b} = u^c A_{cab} \quad \text{or} \quad F_{ab:c} = F^d{}_b A_{dac} + F_a{}^d A_{dbc} \tag{1.14}$$

and the "Ricci rotation symbols" A_{abc} are defined by

$$A_{abc} = -\tfrac{1}{2}(C_{bca} + C_{cab} - C_{abc}). \tag{1.15}$$

The raising and lowering of the tetrad indices are carried out by the tetrad components of the metric

$$g_{ab} = g^{ab} = \text{diag}\,(+1, -1, -1, -1) \tag{1.16}$$

defined, similarly to (11), by

$$g_{ab} = g_{ik}e^i{}_a e^k{}_b, \qquad g^{ab} = g^{ik}e^a{}_i e^b{}_k.$$

Suppose we have a homogeneous solution of (1.1) with an F_{ik} which is invariant under the group in question, then by transvecting (1.1) with $e^i{}_a e^k{}_b$ we get the algebraic equations

$$R_{ab} = \Lambda g_{ab} + T_{ab}, \tag{1.17a}$$

$$T_{ab} = 2(F_{ac}F_b{}^c - \tfrac{1}{4}g_{ab}F_{cd}F^{cd}), \tag{1.17b}$$

$$F^{ab}{}_{:b} = 0, \qquad F^{*ab}{}_{:b} = 0, \tag{1.17c}$$

$$F_{ab} = -F_{ba}, \tag{1.17d}$$

where R_{ab} is given by

$$R_{ab} = A'{}_{ag}A^g{}_{bf} - C'{}_{gf}A^g{}_{ab} \tag{1.18}$$

and these equations are necessarily satisfied. Suppose now that we have a suitable set of constants F_{ab} and $C_{ab}{}^c$ satisfying the equations (1.7) and (1.17), then by integrating (1.4), which are integrable since the Jacobi identities are the integrability conditions of that system, we can find a set of vectors e^a_i, and using (1.3) and (1.12) we can construct the metric and the electromagnetic field, which then satisfy (1.1). It is therefore obvious that *the problem of finding a homogeneous solution of (1.1) with an invariant electromagnetic field is equivalent to the problem of finding a solution of the algebraic equations (1.7), (1.17) for the unknown constants F_{ab} and $C_{ab}{}^c$.* We want to deal with this problem in the next section.

2. SPINOR FORMS OF THE BASIC EQUATIONS

Just as in the case of incoherent matter, it is easier to handle our algebraic problem if we use null tetrads instead of the real orthonormal tetrads $e^i{}_a$ and $e^a{}_i$. We introduce null tetrads $\sigma^i{}_{AA'}$ and $\sigma^{AA'}{}_i$ as special linear combinations of the original orthonormal tetrads, namely

$$\sigma^i{}_{AA'} = e^i{}_a\sigma^a{}_{AA'}, \tag{2.1}$$

$$\sigma^{AA'}{}_i = \sigma^{AA'}{}_a e^a{}_i \tag{2.2}$$

where

$$\sigma^a{}_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \tag{2.3}$$

$$\sigma^{AA'}{}_a \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}. \tag{2.4}$$

In formula (2.3) [(2.4)] a is the row [column] index and the pairs $AA', \cdots = 1, 2$) are the column [row] indices, related to the indices 0, 1, 2, 3 by the following correspondence:

$$\begin{array}{cccccc} a & 0 & 1 & 2 & 3 \\ AA' & 11' & 12' & 21' & 22'. \end{array} \tag{2.5}$$

(The numerical values of A and A' are independent of each other.)[4] Using $\sigma^i{}_{AA'}$ and $\sigma^{AA'}{}_i$ as basic tetrads, the tetrad components of the tensors of our Riemannian manifold become spinors. (Strictly speaking, they are not *the* spinors, but the "dyad components" of the spinors. See Ref. 5.) For example

$$R_{AA'BB'} = R_{ik}\sigma^i{}_{AA'}\sigma^k{}_{BB'}, \qquad T_{AA'BB'} = T_{ik}\sigma^i{}_{AA'}\sigma^k{}_{BB'}, \qquad F_{AA'BB'} = F_{ik}\sigma^i{}_{AA'}\sigma^k{}_{BB'} \tag{2.6}$$

which are analogous to (1.11). It follows further that

$$R_{AA'BB'} = R_{ab}\sigma^a{}_{AA'}\sigma^b{}_{BB'}, \qquad T_{AA'BB'} = T_{ab}\sigma^a{}_{AA'}\sigma^b{}_{BB'}, \qquad F_{AA'BB'} = F_{ab}\sigma^a{}_{AA'}\sigma^b{}_{BB'}. \tag{2.7}$$

Equations (2.7) give the correspondence between the two sorts of "tetrad components." The invariant tensors of our manifold evidently have constant spinor components. For example,

$$g_{ab} \leftrightarrow \epsilon_{AB}\epsilon_{A'B'} \quad \text{and} \quad g^{ab} \leftrightarrow \epsilon^{AB}\epsilon^{A'B'}, \tag{2.8}$$

where ϵ_{AB}, $\epsilon_{A'B'}$ and ϵ^{AB}, $\epsilon^{A'B'}$ are skew symmetric and

$$\epsilon_{12} = \epsilon_{1'2'} = \epsilon^{12} = \epsilon^{1'2'} = 1. \tag{2.9}$$

[4] R. Penrose and W. Rindler, *Application of Spinors to Relativity* (Cambridge University Press, New York, to be published).

[5] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

The raising and lowering of spinor indices is carried out according to the typical rules

$$\varphi^A = \epsilon^{AB}\varphi_B, \qquad \varphi_B = \varphi^A\epsilon_{AB}, \tag{2.10}$$
$$\psi^{A'} = \epsilon^{A'B'}\psi_{B'}, \qquad \psi_{B'} = \psi^{A'}\epsilon_{A'B'}.$$

Since $F_{ab} = -F_{ba}$ and F_{ab} is real the spinor, $F_{AA'BB'} \leftrightarrow F_{ab}$ can be split in the following way:

$$F_{AA'BB'} = \tfrac{1}{2}(F_{AB}\epsilon_{A'B'} + \bar{F}_{A'B'}\epsilon_{AB}), \tag{2.11}$$

where

$$F_{AB} = F_{BA}. \tag{2.12}$$

We have used here the notation typified by

$$\overline{\phi_{AB'}} = \bar{\phi}_{A'B} \tag{2.13}$$

for the complex conjugate spinor. (In *this* equation $A = A'$, etc., numerically.) Since the quantities A_{abc} are real tetrad tensors and skew symmetric in the first two indices, the spinor $A_{AA'BB'CC'} \leftrightarrow A_{abc}$ can be written in the form

$$A_{AA'BB'CC'} = \tfrac{1}{2}(A_{ABCC'}\epsilon_{A'B'} + \bar{A}_{A'B'C'C}\epsilon_{AB}), \tag{2.14}$$

where $A_{ABCC'}$, which we will call the "Ricci rotation spinor," has the symmetry property

$$A_{ABCC'} = A_{BACC'}. \tag{2.15}$$

The tetrad tensor C_{abc} has the same skew symmetry in its first two indices as A_{abc}; therefore the spinor $C_{AA'BB'CC'} \leftrightarrow C_{abc}$ can be split similarly, $C_{AA'BB'CC'} = \frac{1}{2}(C_{ABCC'}\epsilon_{A'B'} + \bar{C}_{A'B'C'C}\epsilon_{AB})$ and the "structure constant spinor" $C_{ABCC'}$ is symmetric in its first two indices. Using (1.15) we get the expression

$$C_{ABCC'} = A_{C(AB)C'} + \bar{A}_{C'(A}\epsilon_{B)C}. \tag{2.16}$$

We used here the definition

$$A_{AA'} = A_{AB}{}^{B}{}_{A'}. \tag{2.17}$$

Using (1.17b) and (2.11) we get for the spinor $T_{AA'BB'} \leftrightarrow T_{ab}$ the expression

$$T_{AA'BB'} = -F_{AB}\bar{F}_{A'B'}. \tag{2.18}$$

Before proceeding we want to make a remark concerning the operation ":" defined by (1.13). This operation can be carried over by translating (1.14) by means of (2.7). We get, for example,

$$u_{AA':BB'} = \tfrac{1}{2}(u^P{}_{A'}A_{PABB'} + u_A{}^{P'}\bar{A}_{P'A'B'B}).$$

It will be convenient later to extend this operation for quantities like ξ_A or $\bar{\xi}_{A'}$ by saying $\xi_{A:BB'} = \frac{1}{2}\xi^P A_{PABB'}$ or $\bar{\xi}_{A':BB'} = \frac{1}{2}\bar{\xi}^{P'}\bar{A}_{P'A'B'B}$ and $(\xi_A + \eta_B)_{:BB'} = \xi_{A:BB'} + \eta_{A:BB'}$, and $(\xi_A\eta_B)_{:CC'} = \xi_{A:CC'}\eta_B + \xi_A\eta_{B:CC'}$ (we suppose here that the ξ's and η's are constant); similarly for quantities like ξ^A or $\bar{\xi}^{A'}$. Then it follows that

$$\epsilon_{AB:CC'} = 0, \qquad \epsilon_{A'B':CC'} = 0,$$
$$\epsilon^{AB}{}_{:CC'} = 0, \qquad \epsilon^{A'B'}{}_{:CC'} = 0.$$

It is easily seen that the quantities $\xi_{A:BB'}$ defined above are the "dyad components of the corresponding spinor" $\xi_{\alpha;\beta\beta'}$ (denoting *the* spinor indices by Greek letters). This concept was first introduced by Newman and Penrose (cf. Ref. 5). Calling the *scalars* ξ_A, F_{AB}, etc., *spinors* instead of calling them *dyad components of the corresponding spinor* is in fact a looseness of terminology—the same as when we call the *scalars* u_a, F_{ab}, etc., *tensors* instead of calling them *tetrad components of the corresponding tensors*. We do it for the sake of brevity and it should not cause confusion here.

One sees easily that the equation $F^{ab}{}_{:b} = 0$ reduces to

$$F^{AB}{}_{:BC'} = 0. \tag{2.19}$$

The same holds for $F^{*ab}{}_{:b} = 0$. Equations (2.19) are therefore the Maxwell equations in spinor form.[6] Translating (1.14) into spinors and using the aforementioned postulates we get

$$F_{AB:CC'} = \tfrac{1}{2}(F^D{}_B A_{DACC'} + F_A{}^D A_{DBCC'}). \tag{2.20}$$

We have seen in Ref. 1 that the curvature spinors χ_{ABCD} and $\phi_{ABC'D'}$ can be expressed by

$$\chi_{ABCD} = \tfrac{1}{4}\{A_{PACP'}A^P{}_{BD}{}^{P'} + A_{PADP'}A^P{}_{BC}{}^{P'}\} + \tfrac{1}{2}A_{ABPP'}C_{CD}{}^{PP'}, \tag{2.21}$$

$$\phi_{ABC'D'} = \tfrac{1}{4}\{A_{PAQC'}A^P{}_B{}^Q{}_{D'} + A_{PAQD'}A^P{}_B{}^Q{}_{C'}\} + \tfrac{1}{2}A_{ABPP'}\bar{C}_{C'D'}{}^{P'P}, \tag{2.22}$$

and their symmetry properties

$$\chi_{ABCD} = \chi_{CDAB}, \tag{2.23a}$$

$$\lambda = \tfrac{1}{2}\chi_{AB}{}^{AB} = \tfrac{1}{2}\bar{\chi}_{A'B'}{}^{A'B'} = \bar{\lambda}, \tag{2.23b}$$

$$\phi_{ABC'D'} = \bar{\phi}_{C'D'AB}, \tag{2.23c}$$

are equivalent to the Jacobi identities. The spinor $R_{AA'BB'} \leftrightarrow R_{ab}$ is given by

$$R_{AA'BB'} = \lambda\epsilon_{AB}\epsilon_{A'B'} - \phi_{ABA'B'} \tag{2.24}$$

(see Ref. 6). From (1.17a) we get therefore the equations

$$\Lambda = \lambda = \tfrac{1}{2}\chi_{AB}{}^{AB}, \tag{2.25a}$$

[6] R. Penrose, Ann. Phys. (N. Y.) 10, 171 (1960).

$$\phi_{ABA'B'} = F_{AB}\bar{F}_{A'B'}. \tag{2.25b}$$

We now observe the following facts. If (2.25a) is satisfied then (2.23b) is also satisfied, since Λ is by definition real. If (2.25b) is satisfied then (2.23c) is also satisfied, since the right-hand side of (2.25c) is, by construction, Hermitian. The only independent part of (2.23) is (2.23a), which is equivalent to

$$C_{ABCC'}(A^{CC'} + \bar{A}^{C'C}) = 0, \tag{2.26}$$

as we have seen in Ref. 1. The spinor form of the basic equations is therefore

$$F^{AB}{}_{:BC'} = 0, \tag{2.27a}$$

$$\phi_{ABA'B'} = F_{AB}\bar{F}_{A'B'}, \tag{2.27b}$$

$$\Lambda = \tfrac{1}{2}\chi_{AB}{}^{AB}, \tag{2.27c}$$

$$C_{ABCC'}(A^{CC'} + \bar{A}^{C'C}) = 0. \tag{2.27d}$$

Our algebraic problem is to solve these equations for the "unknowns" $A_{ABCC'}$ and F_{AB}.

3. The Explicit Algebraic Systems

We will denote our unknowns, the components of $A_{ABCC'}$, by the following symbols:

$A_{ABCC'}$:

	CC'			
AB	$11'$	$12'$	$21'$	$22'$
11	κ	σ	ρ	τ
(12)	ϵ	β	α	γ
22	π	μ	λ	ν

(3.1)

(see Ref. 5). Maxwell's equations have the form

$$\begin{aligned}
(\pi - 2\alpha)F_{11} + 2\rho F_{12} - \kappa F_{22} &= 0,\\
(\mu - 2\gamma)F_{11} + 2\tau F_{12} - \sigma F_{22} &= 0,\\
-\lambda F_{11} + 2\pi F_{12} + (\rho - 2\epsilon)F_{22} &= 0,\\
-\nu F_{11} + 2\mu F_{12} + (\tau - 2\beta)F_{22} &= 0.
\end{aligned} \tag{3.2}$$

We have to distinguish two different cases: (1) the null case: $\frac{1}{2}F_{AB}F^{AB} = 0$; (2) The nonnull case: $\frac{1}{2}F_{AB}F^{AB} \neq 0$, which we will discuss separately. We realize that the σ's are fixed up to spinor transformations, i.e., we are at liberty to use transformations with constant complex components of the form

$$\mathfrak{L}^{A}{}_{B} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3.3}$$

with

$$\det(\mathfrak{L}^{A}{}_{B}) = +1$$

in order to simplify, for example, the component of F_{AB}.

(a) *The Null Case*

Using (3.3) we can always arrange that

$$F_{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{3.4}$$

and we still have the freedom to use transformations of the form

$$\mathfrak{L}^{A}{}_{B} = \begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}, \tag{3.4a}$$

(where C is an arbitrary complex number) which leave (3.4) unchanged, in order to simplify $A_{ABCC'}$ in a convenient stage of our calculations. The form (3.4) of F_{AB} means geometrically that we choose one of the vectors of our basic null tetrad in the direction of the null vector associated with the null electromagnetic field. Maxwell's equations take the simple form

$$\pi = 2\alpha,\ \mu = 2\gamma,\ \lambda = 0,\ \nu = 0 \tag{3.5}$$

and the equation $\phi_{222'2'} = 0$ gives

$$\gamma = 0. \tag{3.6}$$

Equations (2.27) which are not identically satisfied are

$$\alpha(\bar{\beta} - 3\alpha) = 0, \tag{3.7a}$$

$$3\alpha(\bar{\alpha} + \tau) + \beta(\bar{\beta} + \bar{\tau}) = 0, \tag{3.7b}$$

$$\alpha(3\rho + \bar{\epsilon}) + \beta\bar{\sigma} - \bar{\beta}\epsilon = 0, \tag{3.7c}$$

$$\tau(\tau + \beta - \bar{\alpha}) = 0, \tag{3.7d}$$

$$\begin{aligned}
\tau(\bar{\rho} - \rho + \bar{\epsilon} - \epsilon) - \rho(3\bar{\alpha} + \beta + \tau)\\
- \sigma(\bar{\beta} + \bar{\tau} - \alpha) = 0,
\end{aligned}$$

$$\tau\bar{\kappa} - \rho(\rho + \epsilon + \bar{\epsilon}) + \kappa(\alpha + \bar{\beta}) - \sigma\bar{\sigma} = 2, \tag{3.7e}$$

$$\Lambda = -\tfrac{1}{4}[\alpha\bar{\alpha} + (\beta - \tau)(\bar{\beta} - \bar{\tau}) - 2\alpha(\beta + \tau)], \tag{3.7f}$$

$$\begin{aligned}
\sigma(\alpha + \bar{\beta} - \bar{\tau}) + (2\epsilon - \rho)(\bar{\alpha} + \beta - \tau)\\
+ (2\beta - \tau)(\rho + \bar{\rho} - \epsilon - \bar{\epsilon}) = 0,
\end{aligned} \tag{3.7g}$$

$$2\alpha(\bar{\alpha} + \beta - \tau) + \tau(\alpha + \bar{\beta} - \bar{\tau}) = 0. \tag{3.7h}$$

(b) *The Nonnull Case*

Using (3.3) we can always arrange that

$$F_{AB} = \begin{pmatrix} 0 & F \\ F & 0 \end{pmatrix} \tag{3.8}$$

and we are left with the freedom

$$\mathfrak{L}^{A}{}_{B} = \begin{pmatrix} A & 0 \\ 0 & A^{-1} \end{pmatrix}, \qquad A \neq 0 \quad \text{complex}. \tag{3.9}$$

The Maxwell equations take the simple form

$$\rho = 0,\ \tau = 0,\ \pi = ,\ \mu = 0 \tag{3.10}$$

and the equations (2.27) read as follows:

$$\kappa(3\alpha + \bar{\beta}) - \sigma\bar{\sigma} = 0, \tag{3.11a}$$

$$\nu(3\beta + \bar{\alpha}) - \lambda\bar{\lambda} = 0, \tag{3.11b}$$

$$\kappa(3\gamma + \bar{\gamma}) + \sigma(3\alpha - \bar{\beta}) = 0, \quad (3.11c)$$

$$\nu(3\epsilon + \bar{\epsilon}) + \lambda(3\beta - \bar{\alpha}) = 0, \quad (3.11d)$$

$$\kappa\bar{\nu} + \sigma(3\gamma - \bar{\gamma}) = 0, \quad (3.11e)$$

$$\nu\bar{\kappa} + \lambda(3\epsilon - \bar{\epsilon}) = 0, \quad (3.11f)$$

$$\kappa\lambda + \gamma\bar{\kappa} - \beta\bar{\sigma} + \epsilon\bar{\beta} - \alpha\bar{\epsilon} + 2\alpha\epsilon = 0, \quad (3.11g)$$

$$\nu\sigma + \epsilon\bar{\nu} - \alpha\bar{\lambda} + \gamma\bar{\alpha} - \beta\bar{\gamma} + 2\beta\gamma = 0, \quad (3.11h)$$

$$\Lambda + F\bar{F} = \tfrac{1}{2}(\kappa\nu - \alpha\bar{\alpha} - \beta\bar{\beta} + 2\alpha\beta), \quad (3.11i)$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(\sigma\lambda + \gamma\bar{\epsilon} + \bar{\gamma}\epsilon + 2\gamma\epsilon), \quad (3.11j)$$

$$\kappa(\gamma + \bar{\gamma}) + \sigma(\bar{\beta} - \alpha) - 2(\bar{\alpha}\epsilon + \beta\bar{\epsilon}) = 0, \quad (3.11k)$$

$$\nu(\epsilon + \bar{\epsilon}) + \lambda(\bar{\alpha} - \beta) - 2(\bar{\beta}\gamma + \alpha\bar{\gamma}) = 0. \quad (3.11l)$$

4. SOLUTIONS OF THE ALGEBRAIC SYSTEM OF THE NULL CASE

In this section we obtain *all* the solutions of Eqs. (3.7). Looking at these equations, one can see easily that there are only two different cases possible, namely Case a: $\alpha = 0$, $\tau = 0$ and case b: $\alpha\tau \neq 0$. [The case $\alpha = 0$, $\tau \neq 0$ is not possible, since (3.7d) $\rightarrow$ $\tau = -\beta$ and (3.7h) $\rightarrow$ $\tau = \beta$ and therefore $\tau = 0$, which is a contradiction. The case $\alpha \neq 0$, $\tau = 0$ can also be excluded since (3.7h) $\rightarrow$ $\bar{\beta} = -\alpha$ and (3.7a) $\rightarrow$ $\bar{\beta} = 3\alpha$ and therefore $\alpha = 0$, which is a contradiction.]

We now consider Case (a).

$$\alpha = 0, \tau = 0. \quad (4.1)$$

From (3.7b) we get

$$\beta = 0 \quad (4.2)$$

and two other equations

$$\Lambda = 0, \rho(\rho + \epsilon + \bar{\epsilon}) + \sigma\bar{\sigma} = 2, \quad (4.3)$$

and the rest is satisfied identically. Collecting our results we have

$A_{ABCC'}$:

AB	CC' = 11′	12′	21′	22′
11	κ	σ	ρ	0
(12)	ϵ	0	0	0
22	0	0	0	0

. (4.4)

The content of (4.4) is discussed in the next section; now we examine Case (b).

From (1) and (4) it follows that

$$\beta = 3\bar{\alpha} \text{ and } \tau = -2\bar{\alpha}. \quad (4.5)$$

Substituting (4.5) into (3.7) we have the equations

$$\rho + \bar{\rho} = \epsilon - \bar{\epsilon} = 0, \quad (4.6a)$$

$$\alpha(3\rho - 2\epsilon) + 3\bar{\alpha}\bar{\sigma} = 0, \quad (4.6b)$$

$$-2\bar{\alpha}\bar{\kappa} + 4\alpha\kappa - \rho(\rho + 2\epsilon) = 2, \quad (4.6c)$$

$$\Lambda = -6\alpha\bar{\alpha}. \quad (4.6d)$$

The components of $A_{ABCC'}$ are given by

$A_{ABCC'}$:

AB	CC' = 11′	12′	21′	22′
11	κ	σ	ρ	$-2\bar{\alpha}$
(12)	ϵ	$3\bar{\alpha}$	α	0
22	2α	0	0	0

.

Applying the transformation (3.4a) to $A_{ABCC'}$

$$\tilde{A}_{ABCC'} = A_{DEFF'}\mathcal{L}^D_A\mathcal{L}^E_B\mathcal{L}^F_C\bar{\mathcal{L}}^{F'}_{C'},$$

we get for the new $\tilde{A}_{1112'}$ component the expression $\tilde{A}_{1112'} = \sigma + 2\bar{\alpha}C$. Since $\bar{\alpha} \neq 0$ we can choose C such that $\tilde{A}_{1112'} = 0$. This fact allows us to impose the condition

$$\sigma = 0 \quad (4.7)$$

without loss of generality. Then it follows from (4.6)(a) and (b) that

$$\rho = 0 \text{ and } \epsilon = 0, \quad (4.8)$$

and from (4.6)(c)

$$\alpha\kappa = 1. \quad (4.9)$$

Collecting our results we have

$A_{ABCC'}$:

AB	CC' = 11′	12′	21′	22′
11	$1/\alpha$	0	0	$-2\bar{\alpha}$
(12)	0	$3\bar{\alpha}$	α	0
22	2α	0	0	0

(4.10)

and

$$\Lambda = -6\alpha\bar{\alpha}. \quad (4.11)$$

Equations (4.4) and (4.9) contain all the solutions with null electromagnetic fields. We made the distinction between the two cases rather formal. But one observes that in the Case a, $\Lambda = 0$, and in the Case b, $\Lambda \neq 0$, both of which are invariant statements; therefore, we want to refer to these cases accordingly. We evaluate these in the section following the one immediately below.

5. SOLUTIONS OF THE ALGEBRAIC SYSTEM OF THE NONNULL CASE

In this section we obtain *some* solutions of (3.11). We make the following statement: Equations (3.11) allow solutions only in two cases: case a $\kappa = \sigma = \nu = \lambda = 0$; case b $\kappa\sigma\nu\lambda \neq 0$. We show at first that $\kappa = 0$ is equivalent to $\sigma = 0$.

That $\kappa = 0$ implies $\sigma = 0$ follows immediately from (3.11a). To prove that $\sigma = 0$ implies $\kappa = 0$ we

show that in the case of $\sigma = 0$, $\kappa \neq 0$ we don't have any solution. In this case (3.11) reduces to $\beta = -3\bar{\alpha}$ $\gamma = \nu = \lambda = \alpha\epsilon = 0$, $\Lambda = F\bar{F} = -4\alpha\bar{\alpha}$, which obviously can't be fulfilled.

We now show that $\nu = 0$ is equivalent to $\lambda = 0$.

That $\nu = 0$ implies that $\lambda = 0$ follows immediately from (3.11b). To prove that $\lambda = 0$ implies $\nu = 0$ we show that in the case of $\lambda = 0$, $\nu \neq 0$ we don't have any solution. In this case (3.11) reduces to

$$\alpha = -3\bar{\beta},\ \epsilon = \kappa = \sigma = \beta\gamma = 0,\ \Lambda = F\bar{F} = -4\beta\bar{\beta},$$

which obviously can't be fulfilled.

If we can exclude the case $\kappa = 0$, $\sigma = 0$, $\nu \neq 0$, $\lambda \neq 0$ as well as the case $\kappa \neq 0$, $\sigma \neq 0$, $\nu = 0$, $\lambda = 0$, we have proven our statement.

If $\kappa = \sigma = 0$ and $\nu\lambda \neq 0$, we get from (3.11f) $\epsilon = 0$ and from (3.11d) $\alpha = 3\bar{\beta}$ and from (3.11i, j) $\Lambda = F\bar{F} = -\beta\bar{\beta}$ which can't be fulfilled.

If $\nu = \lambda = 0$ and $\kappa\sigma \neq 0$, we get from (3.11e) $\gamma = 0$ and from (3.11c) $\beta = 3\bar{\alpha}$ and from (3.11i, j) $\Lambda = F\bar{F} = -\alpha\bar{\alpha}$ which can't be fulfilled.

Case a:

$$\kappa = \nu = \sigma = \lambda = 0. \tag{5.1}$$

Equations (3.11) read

$$(2\alpha + \bar{\beta})\epsilon - \alpha\bar{\epsilon} = 0, \tag{5.2a}$$

$$\bar{\alpha}\epsilon + \beta\bar{\epsilon} = 0, \tag{5.2b}$$

$$(2\beta + \bar{\alpha})\gamma - \beta\bar{\gamma} = 0, \tag{5.2c}$$

$$\bar{\beta}\gamma + \alpha\bar{\gamma} = 0, \tag{5.2d}$$

$$\Lambda + F\bar{F} = -\tfrac{1}{2}(\alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta), \tag{5.2e}$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(\gamma\bar{\epsilon} + \bar{\gamma}\epsilon + 2\gamma\epsilon). \tag{5.2f}$$

The determinant of the linear system for ϵ, $\bar{\epsilon}$ and γ, $\bar{\gamma}$ is

$$d = \alpha\bar{\alpha} + \beta\bar{\beta} + 2\alpha\beta.$$

If $d \neq 0$ then $\gamma = \epsilon = 0$ and therefore

$$\begin{aligned}\Lambda = F\bar{F} &= -\tfrac{1}{4}(\alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta)\\ &= -\tfrac{1}{4}(\alpha\bar{\alpha} + \beta\bar{\beta} - \alpha\beta - \bar{\alpha}\bar{\beta})\\ &= -\tfrac{1}{4}(\alpha - \bar{\beta})(\bar{\alpha} - \beta) \leq 0,\end{aligned}$$

which can't be fulfilled. We have therefore to assume that $d = \alpha\bar{\alpha} + \beta\bar{\beta} + 2\alpha\beta = 0$, but this is equivalent to $(\bar{\alpha} + \beta)(\alpha + \bar{\beta}) = 0$; therefore

$$\beta = -\bar{\alpha} \tag{5.3}$$

and the equations (5.2) read

$$\alpha(\epsilon - \bar{\epsilon}) = 0, \tag{5.4a}$$

$$\alpha(\gamma - \bar{\gamma}) = 0, \tag{5.4b}$$

$$\Lambda + F\bar{F} = -2\alpha\bar{\alpha}, \tag{5.4c}$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(\gamma + \bar{\gamma})(\epsilon + \bar{\epsilon}), \tag{5.4d}$$

$$\gamma\epsilon = \bar{\gamma}\bar{\epsilon}. \tag{5.4e}$$

From (5.4e) it follows that $\epsilon = h\bar{\gamma}$ where h is real. Substituting into (5.4c) and (5.4d) and taking into account that $F\bar{F} > 0$ we see that $h > 0$ must hold. Applying (3.9): $\tilde{A}_{ABCC'} = A_{DEFF'}\mathcal{L}^D_A\mathcal{L}^E_B\bar{\mathcal{L}}^F_C\mathcal{L}^{F'}_{C'}$ we get

$$\tilde{A}_{1211'} = AA'h\bar{\gamma},\quad \tilde{A}_{1221'} = (\bar{A}/A)\alpha,\quad \tilde{A}_{1222'} = \gamma/A\bar{A}$$

and by suitable choice of A we can arrange that $|\tilde{A}_{1211'}| = |\tilde{A}_{1222'}|$ and $\tilde{A}_{1221'}$ is real. That means that we can assume, without loss of generality, that

$$\epsilon = \bar{\gamma} \text{ and } \alpha = a \text{ real}. \tag{5.5}$$

Then our equations take the form

$$\beta = -a, \tag{5.6a}$$

$$a(\gamma - \bar{\gamma}) = 0, \tag{5.6b}$$

$$\Lambda + F\bar{F} = -2a^2, \tag{5.6c}$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(\gamma + \bar{\gamma})^2. \tag{5.6d}$$

Equation (5.6b) gives two cases: $a = 0$ or $\gamma = \bar{\gamma}$.

If

$$a = 0, \tag{5.7}$$

we have

$A_{ABCC'}$:

	CC'			
AB	$11'$	$12'$	$21'$	$22'$
11	0	0	0	0
(12)	$\bar{\gamma}$	0	0	γ
22	0	0	0	0

(5.8)

and

$$-\Lambda = F\bar{F} = \tfrac{1}{4}(\gamma + \bar{\gamma})^2. \tag{5.9}$$

If

$$\gamma = \bar{\gamma} = b, \tag{5.10}$$

we have

$A_{ABCC'}$:

	CC'			
AB	$11'$	$12'$	$21'$	$22'$
11	0	0	0	0
(12)	b	$-a$	a	b
22	0	0	0	0

(5.11)

and

$$F\bar{F} = b^2 - a^2,\ (b^2 \neq 0),\ \Lambda = -(a^2 + b^2). \tag{5.12}$$

(If $b^2 = a^2$ we get a vacuum solution with a negative Λ term.)

Case b: $\kappa\nu\sigma\lambda \neq 0$.

We did not find the general solution in this case, but it is very easy to get a solution under the assumption:

$$\nu = \bar{\kappa},\ \lambda = \bar{\sigma},\ \beta = \bar{\alpha},\ \epsilon = \bar{\gamma}. \tag{5.13}$$

In order to preserve the form of (5.13) we have to restrict (3.9) to

$$\mathfrak{L}^A{}_B = \begin{bmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{bmatrix}; \tag{5.14}$$

(5.14) allows us to impose the condition that

$$\alpha = \bar{\alpha} = a \tag{5.15}$$

without further loss of generality. Equations (3.11) read

$$4a\kappa = \sigma\bar{\sigma}, \tag{5.16a}$$

$$3\gamma + \bar{\gamma} = -2a\sigma/\kappa, \tag{5.16b}$$

$$3\gamma - \bar{\gamma} = -\kappa^2/\sigma, \tag{5.16c}$$

$$3a\gamma + (\kappa - a)(\sigma + \bar{\gamma}) = 0, \tag{5.16d}$$

$$\Lambda + F\bar{F} = \tfrac{1}{2}\kappa\bar{\kappa}, \tag{5.16e}$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(\sigma\bar{\sigma} + (\gamma + \bar{\gamma})^2), \tag{5.16f}$$

$$(\gamma + \bar{\gamma})(\kappa - 2a) = 0. \tag{5.16g}$$

Equation (5.16g) is satisfied if $\gamma + \bar{\gamma} = 0$ or $\kappa - 2a = 0$. Since $\gamma + \bar{\gamma} = 0$ gives no solution, we write

$$\kappa = 2a. \tag{5.17}$$

Substituting into (5.16) we get

$$8a^2 = \sigma\bar{\sigma}, \tag{5.18a}$$

$$3\gamma + \bar{\gamma} = -\sigma, \tag{5.18b}$$

$$3\gamma - \bar{\gamma} = -\tfrac{1}{2}\bar{\sigma}, \tag{5.18c}$$

$$\Lambda + F\bar{F} = 2a^2, \tag{5.18d}$$

$$\Lambda - F\bar{F} = -\tfrac{1}{2}(8a^2 + (\gamma + \bar{\gamma})^2). \tag{5.18e}$$

From (5.18)(b) and (c) it follows that

$$\sigma = \bar{\sigma} \text{ and } \gamma = -\tfrac{1}{4}\sigma, \tag{5.19}$$

and therefore we have

$$\sigma = 2\sqrt{2}a,\quad \gamma = -a/\sqrt{2},\quad \Lambda = -\tfrac{3}{2}a^2,\quad F\bar{F} = \tfrac{7}{2}a^2, \tag{5.20}$$

and

$A_{ABCC'}$:

AB \ CC'	11′	12′	21′	22′
11	$2a$	$2\sqrt{2}a$	0	0
(12)	$-\frac{a}{\sqrt{2}}$	a	a	$-\frac{a}{\sqrt{2}}$.
22	0	0	$2\sqrt{2}a$	$2a$

(5.21)

6. LINE ELEMENTS IN THE NULL CASE

In this section we examine the solutions of (1.1) corresponding to (4.4) and (4.10) which we call case $\Lambda = 0$, and case $\Lambda \neq 0$, respectively. We show that in the case $\Lambda = 0$ all the solutions are Robinson waves and in the case $\Lambda \neq 0$ we have a unique solution (6.26).

The totally symmetric spinor ψ_{ABCD} defined by

$$\psi_{ABCD} = \chi_{ABCD} - \tfrac{1}{6}\chi_{AB}{}^{AB}(\epsilon_{AC}\epsilon_{BD} + \epsilon_{AD}\epsilon_{BC}) \tag{6.1}$$

gives the algebraic types of the solutions (see Ref. 6). We want to calculate its components.

Case A. $\Lambda = 0$

Using (4.4) and (2.21) we see that the only nonvanishing component of ψ_{ABCD} is

$$\psi_{1111} = -\tfrac{1}{2}\sigma(3\epsilon - \bar{\epsilon} + \rho + \bar{\rho}),$$

therefore ψ_{ABCD} can be written in the form

$$\psi_{ABCD} = \psi_{(A}\psi_B\psi_C\psi_{D)}, \tag{6.2}$$

where the spinor ψ_A has the form

$$\psi_A = (\psi, 0) \tag{6.3}$$

and where ψ is a suitable complex number. We observe that F_{AB} [see (3.4)] can be written in the form

$$F_{AB} = F_{(A}F_{B)}, \tag{6.4}$$

where

$$F_A = (1, 0). \tag{6.5}$$

The meaning of (6.2)–(6.5) is that the solutions contained in (4.4) are all of type null and the null vectors $\psi_i \leftrightarrow \psi_A$ and $F_i \leftrightarrow F_A$ are colinear. Using the formula

$$\psi_A\psi_{A:BB'} = \tfrac{1}{2}\psi^P A_{PABB'} \tag{6.6}$$

and (4.4), we see that the only nonvanishing component of this spinor is given by $\psi_{1:11'} = \psi^2 A_{2111'} = -\frac{1}{2}\epsilon\psi$, which means that

$$\psi_{A:BB'} = \omega\psi_A\psi_B\bar{\psi}_{B'}, \tag{6.7}$$

where ω is a suitable complex constant. The null vector $\psi_i \leftrightarrow \psi_A$ given by

$$\psi_i = \psi_A\bar{\psi}_{A'}\sigma^{AA'}{}_i \tag{6.8}$$

has the covariant derivative

$$\psi_{i;k} = (\psi_A\bar{\psi}_{A'})_{:BB'}\sigma^{AA'}{}_i\sigma^{BB'}{}_k = (\omega + \bar{\omega})\psi_i\psi_k. \tag{6.9}$$

Therefore by suitable choice of the function $\lambda(\chi^k)$ we can arrange that the null vector $k_i = \lambda\psi_i$ satisfies the condition

$$k_{i;k} = 0. \tag{6.10}$$

This result we can express by the following:

Theorem 1. All homogeneous solutions of the Einstein–Maxwell equations with electromagnetic null fields and $\Lambda = 0$ are Robinson waves.

We have seen that the covariant constant null vector of gravitational field coincides with the null vector associated with the electromagnetic null field. This is a simple consequence of a

General Theorem Found by I. Robinson (unpublished): If the Riemannian space–time R_4 is such that (i) the Einstein–Maxwell equations in vacuum [we mean here the equations

$$R_{ik} = 2(F_{il}F_k{}^l - \tfrac{1}{4}g_{ik}F_{lm}F^{lm}),$$
$$F_{ik} = -F_{ki}, \qquad F^{ik}{}_{;k} = 0, \qquad F^{*ik}{}_{;k} = 0]$$

are satisfied, and (ii) there exists a covariant constant null vector, then it follows that: (a) the Weyl tensor of R_4 is of type null; (b) the electromagnetic field is a null field; (c) both principal null directions are *the* covariant constant one. R_4 is generally described as Robinson wave or plane-fronted wave.

We don't want to examine here the question how many different groups are contained in (4.4) and how many different solutions belong to this class.

Case B. $\Lambda \neq 0$

Using (4.10), (2.21), and (6.1) we see that the only nonvanishing component of ψ_{ABCD} is

$$\psi_{1111} = 3\bar{\alpha}/\alpha;$$

therefore, exactly as before, it follows that the solution (we will see that there is only one) is of type null and the null vectors ψ_i and F_i are colinear. The components of the spinor $\psi_{A:BB'}$ are given by

$\psi_{A:BB'}$:

A \ BB'	11′	12′	21′	22′
1	0	$-\frac{3}{2}\bar{\alpha}\psi$	$-\frac{\alpha}{2}\psi$	0
2	$-\alpha\psi$	0	0	0

and

$$\bar{\psi}_{A':BB'} = \tfrac{1}{2}\bar{\psi}^{P'}\bar{A}_{P'A'B'B} \quad \text{by}$$

$\bar{\psi}_{A':BB'}$:

A' \ BB'	11′	12′	21′	22′
1′	0	$-\frac{\bar{\alpha}}{2}\bar{\psi}$	$-\frac{3}{2}\alpha\bar{\psi}$	0
2′	$-\bar{\alpha}\bar{\psi}$	0	0	0

The tetrad components of $\psi_i \leftrightarrow \psi_A$ with respect to the orthonormal tetrad $e^a{}_i$ are given by

$$\psi_a = \sigma^{AA'}{}_a\psi_A\bar{\psi}_{A'} = \frac{1}{\sqrt{2}}|\psi|^2(1, 0, 0, 1),$$

and using

$$\psi_{a:b} = \sigma^{AA'}{}_a\sigma^{BB'}{}_b(\psi_A\bar{\psi}_{A'}){}_{:BB'}$$
$$= \sigma^{AA'}{}_a\sigma^{BB'}{}_b(\psi_{A:BB'}\bar{\psi}_{A'} + \psi_A\bar{\psi}_{A':BB'}),$$

we get

$$\psi_{a:b} = |\psi|^2\begin{bmatrix} 0 & -2A & 2a & 0 \\ -A & 0 & 0 & -A \\ a & 0 & 0 & a \\ 0 & -2A & 2a & 0 \end{bmatrix} \tag{6.11}$$

where

$$\alpha = A + i\alpha. \tag{6.12}$$

One sees immediately that

$$\psi_{a:b}\psi^b = 0, \qquad \theta = \tfrac{1}{2}\psi^a{}_{:a} = 0,$$
$$\omega^2 = \tfrac{1}{2}\psi_{[a:b]}\psi^{a:b} = 0, \tag{6.13}$$

and

$$\sigma^2 = \tfrac{1}{2}\{\psi_{(a:b)}\psi^{a:b} - \tfrac{1}{2}(\psi^a{}_{:a})^2\} = 0 \tag{6.14}$$

(see Ref. 7). The rays are therefore geodesic, having vanishing shear, expansion, and rotation, but are not covariant constant.

We want to show that we have a single solution in this case. Using (4.10) and (2.16) we can calculate the components of $C_{ABCC'}$ and then

$$C_{abc} = \sigma^{AA'}{}_a\sigma^{BB'}{}_b\sigma^{CC'}{}_c\tfrac{1}{2}\{C_{ABCC'}\epsilon_{A'B'} + \bar{C}_{A'B'C'C}\epsilon_{AB}\}$$
$$C_{ab}{}^c = g^{cd}C_{abc} \qquad \text{[see (1.16)]}, \tag{6.15}$$

which are structure constants of the group. We give the results of this calculation, writing down the commutator relations of the group:

$$(X_2, X_3) = (a/2\sqrt{2}\,|\alpha|^2)\{(8\,|\alpha|^2 + 1)X_0 - (4\,|\alpha|^2 + 1)X_3\},$$
$$(X_3, X_1) = (A/2\sqrt{2}\,|\alpha|^2)\{(8\,|\alpha|^2 + 1)X_0 - (4\,|\alpha|^2 + 1)X_3\},$$
$$(X_1, X_2) = \sqrt{2}\{aX_1 + AX_2\},$$
$$(X_1, X_0) = (A/2\sqrt{2}\,|\alpha|^2)\{(4\,|\alpha|^2 - 1)X_0 - (8\,|\alpha|^2 - 1)X_3\},$$
$$(X_2, X_0) = -(a/2\sqrt{2}\,|\alpha|^2)\{(4\,|\alpha|^2 - 1)X_0 - (8\,|\alpha|^2 - 1)X_3\},$$
$$(X_3, X_0) = 0. \tag{6.16}$$

[7] P. Jordan, J. Ehlers, and R. Sachs, Akad. Wiss. Lit. (Mainz), Abhandl. Math.-Nat. Kl. Nr. 1 1961.

Introducing new operators Y_a given by

$$Y_a = X_b A^b{}_a, \tag{6.17}$$

where $A^b{}_a$ is given by

$$A^b{}_a = \begin{pmatrix} \frac{1}{8|\alpha|^2} & 0 & 0 & \left(1+\frac{1}{8|\alpha|^2}\right) \\ 0 & \frac{a}{\sqrt{2}|\alpha|^2} & -\frac{A}{\sqrt{2}|\alpha|^2} & 0 \\ 0 & \frac{A}{\sqrt{2}|\alpha|^2} & \frac{a}{\sqrt{2}|\alpha|^2} & 0 \\ -\frac{1}{8|\alpha|^2} & 0 & 0 & \left(1-\frac{1}{8|\alpha|^2}\right) \end{pmatrix}, \tag{6.18}$$

we get the new commutator relations

$$\begin{aligned} &(Y_0, Y_1) = 0, && (Y_1, Y_3) = 0, \\ &(Y_3, Y_0) = 0, && (Y_0, Y_2) = 3Y_0, \\ &(Y_1, Y_2) = Y_1, && (Y_3, Y_2) = -Y_3. \end{aligned} \tag{6.19}$$

A solution of (1.5) corresponding to (6.19) is given by

$$e^a{}_i = \begin{pmatrix} e^{-x^2+x^3} & 0 & 0 & 0 \\ 0 & e^{-x^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{x^2} \end{pmatrix}. \tag{6.20}$$

The reciprocal vectors of (6.16) are given by

$$\bar{e}^a{}_i = A^a{}_b e^b{}_i \tag{6.21}$$

and the metric by

$$g_{ik} = \bar{e}^a{}_i \bar{g}_{ab} \bar{e}^b{}_k = e^c{}_i g_{cd} e^d{}_k, \tag{6.22}$$

where

$$\bar{g}_{ab} = \text{diag}\,(+1, -1, -1, -1) \tag{6.23}$$

and

$$g_{cd} = A^a{}_c \bar{g}_{ab} A^b{}_d. \tag{6.24}$$

Using our formulas we get by computation

$$g_{ab} = \frac{1}{2|\alpha|^2} \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{pmatrix}. \tag{6.25}$$

The metric takes the form

$$\begin{aligned} ds^2 = (1/2|\alpha|^2)\{&e^{-2x^2}[dx^0\,dx^3 - (dx^1)^2] \\ &+ e^{2x^2}(dx^3)^2 - (dx^2)^2\}, \end{aligned} \tag{6.26}$$

$$\Lambda = -6|\alpha|^2.$$

The electromagnetic field is given by

$$F_{ik} = k_{[i} w_{k]} \tag{6.27}$$

where

$$k_i = (0, 0, 0, 2e^{x^2}), \tag{6.28}$$

$$w_i = (0, (a/\sqrt{2}|\alpha|^2)e^{-x^2}, -A/\sqrt{2}|\alpha|^2, 0)$$

[see (6.12)]. (6.29)

Integrating the Killing equations

$$g_{ik,l}\xi^l + g_{lk}\xi^l{}_{,i} + g_{il}\xi^l{}_{,k} = 0,$$

we can see that the maximal group of (6.26) is a five-parametric one, given by the commutator relations

$$\begin{aligned} &(X_0, X_1) = 0, && (X_1, X_3) = 0, \\ &(X_3, X_0) = 0, && (X_0, X_2) = 3X_0, \\ &(X_1, X_2) = X_1, && (X_3, X_2) = -X_3, \\ &(X_0, X_4) = 0, && (X_1, X_4) = 0, \\ &(X_2, X_4) = -2X_4, && (X_3, X_4) = 0, \end{aligned} \tag{6.30}$$

having in our coordinate system the following infinitesimal generators:

$$\xi^i{}_a = \begin{pmatrix} 1 & 0 & 3x^0 & 0 & 2x^1 \\ 0 & 1 & x^1 & 0 & x^3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -x^3 & 1 & 0 \end{pmatrix}.$$

We can summarize our results as follows:

Theorem 2. The Einstein–Maxwell equations with null electromagnetic field and $\Lambda \neq 0$ have a single solution given by (6.26).

7. LINE ELEMENTS IN THE NONNULL CASE

Equation (5.8) contains the group with the commutator relations:

$$\begin{aligned} &(X_1, X_0) = 0, && (X_2, X_0) = 0, \\ &(X_1, X_2) = 0, && (X_0, X_3) = -\sqrt{2}aX_0, \\ &(X_1, X_3) = \sqrt{2}bX_2, && (X_2, X_3) = -\sqrt{2}bX_1, \end{aligned} \tag{7.1}$$

where

$$\gamma = a + ib \qquad (a \neq 0 \text{ since } a^2 = F\bar{F} = -\Lambda).$$

A set of reciprocal vectors is given by

$$e^a{}_i = \begin{bmatrix} e^{\sqrt{2}az^3} & 0 & 0 & 0 \\ 0 & \cos\sqrt{2}bx^3 & \sin\sqrt{2}bx^3 & 0 \\ 0 & -\sin\sqrt{2}bx^3 & \cos\sqrt{2}bx^3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.2}$$

and the line element takes the form

$$ds^2 = e^{2\sqrt{2}az^3}(dx^0)^2 - (dx^3)^2 - (dx^1)^2 - (dx^2)^2. \tag{7.3}$$

The group, corresponding to (4.21) is given by the commutator relations:

$$\begin{aligned} &(X_1, X_0) = 0, && (X_2, X_0) = 0, \\ &(X_1, X_2) = -\sqrt{2}bkX_2, && (X_0, X_3) = -\sqrt{2}bX_0, \\ &(X_1, X_3) = 0, && (X_2, X_3) = 0, \end{aligned} \tag{7.4}$$

where we introduced the new parameter k by the equation

$$a = bk. \tag{7.5}$$

Then from (5.12)

$$F\bar{F} = b^2(1-k^2), \quad \Lambda = -b^2(1+k^2) \quad (k^2 \leq 1). \tag{7.6}$$

A set of reciprocal vectors is given by

$$e^a{}_i = \frac{1}{\sqrt{2}b}\begin{bmatrix} e^{z^3} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-kz^1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.7}$$

and the line element has the form

$$\begin{aligned} ds^2 = (1/2b^2)[&e^{2z^3}(dx^0)^2 - (dx^3)^2 \\ &- (dx^1)^2 - e^{-2kz^1}(dx^2)^2]; \end{aligned} \tag{7.8}$$

(7.8) contains (7.3) at $k = 0$. This solution is known and was first given by Robinson.[8]

We now evaluate (5.21). We show that the solution belonging to (5.21) is of type I. A Riemannian space–time is of type I if there exist four different spinors ξ^A satisfying the equation

$$\psi_{ABCD}\xi^A\xi^B\xi^C\xi^D = 0,$$

where ψ_{ABCD} is defined by (5.1). That means that the polynomial

$$\psi_{1111}x^4 + 4\psi_{1112}x^3 + 6\psi_{1122}x^2 + 4\psi_{1222}x + \psi_{2222} = 0,$$

where $x = \xi^1/\xi^2$ has four different roots. One calculates easily that in our case this polynomial is given by

$$\begin{aligned} x^4 - \frac{4\sqrt{2}}{3}x^3 - x^2 - \frac{4\sqrt{2}}{3}x + 1 \\ = (x^2 - \tfrac{1}{3}(2\sqrt{2} + [35]^{\frac{1}{2}})x + 1) \\ \times (x^2 - \tfrac{1}{3}(2\sqrt{2} - [35]^{\frac{1}{2}})x + 1) = 0, \end{aligned}$$

and has obviously four different roots.

The corresponding group is given by the commutator relations

$$\begin{aligned} &(X_0, X_1) = 0, && (X_0, X_2) = 0, \\ &(X_1, X_2) = 0, && (X_0, X_3) = aX_0, \\ &(X_1, X_3) = -2a(\sqrt{2}X_0 + X_1), \\ &(X_2, X_3) = -2aX_2. \end{aligned} \tag{7.9}$$

Introducing new operators Y_a given by (6.17) where $A^a{}_b$ has the form

$$A^b{}_a = \frac{1}{a}\begin{bmatrix} 1 & -2\sqrt{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.10}$$

we get the new commutator relations

$$\begin{aligned} &(Y_0, Y_1) = 0, && (Y_0, Y_2) = 0, \\ &(Y_1, Y_2) = 0, && (Y_0, Y_3) = Y_0, \\ &(Y_1, Y_3) = 2Y_1, && (Y_2, Y_3) = -Y_2. \end{aligned} \tag{7.11}$$

A set of reciprocal vectors to (7.11) is given by

$$e^a{}_i = \begin{bmatrix} e^{-z^3} & 0 & 0 & 0 \\ 0 & e^{-2z^3} & 0 & 0 \\ 0 & 0 & e^{2z^3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.12}$$

and g_{ab} defined by (6.24) takes the form

$$g_{ab} = \frac{1}{a^2}\begin{bmatrix} 1 & -2\sqrt{2} & 0 & 0 \\ -2\sqrt{2} & 7 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{7.13}$$

and the metric is given by

$$\begin{aligned} ds^2 = a^{-2}\{&e^{-2z^3}(dx^0)^2 - 4\sqrt{2}e^{-3z^3}\,dx^0\,dx^1 \\ &+ 7e^{-4z^3}(dx^1)^2 + e^{2z^3}(dx^2)^2 - (dx^3)^2\} \end{aligned} \tag{7.14}$$

[see (5.22)], and

$$\Lambda = -\tfrac{3}{2}a^2, \qquad F\bar{F} = \tfrac{7}{2}a^2. \tag{7.15}$$

[8] I. Robinson, Bulletin de l'Academie Polonaise des Sciences, Serie des Sci. Math., Astron., et Phys. 7, No. 6, 351 (1959).

The electromagnetic field has the form

$$F_{ik} = 2a\{k_{[i}l_{k]} \cos\varphi + \tfrac{1}{2}\eta_{ikpq}k^{[p}l^{q]} \sin\varphi\} \quad (7.16)$$

where φ is given by $F = ae^{i\varphi}$ and

$$k_i = (1/a\sqrt{2})(e^{-x^3}, -2\sqrt{2}e^{-2x^3}, 0, 1), \quad (7.17)$$

$$l_i = (1/a\sqrt{2})(e^{-x^3}, -2\sqrt{2}e^{-2x^3}, 0, -1). \quad (7.18)$$

The question remains open whether one could find solutions in the case $\kappa\sigma\lambda\nu \neq 0$ different from (7.14).

ACKNOWLEDGMENTS

I am deeply indebted to I. Robinson for his interest in this work. I wish to express my gratitude to R. Penrose and W. Rindler for valuable discussions about spinors. Special thanks go to W. Rindler for his help in preparing this paper for publication. This work has been supported in part by NASA under grant No. NsG269-62, and in part by the U.S. Air Force Office of Scientific Research, under grant No. AF-AFOSR-903-65.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

All Homogeneous Solutions of Einstein's Field Equations with Incoherent Matter and Electromagnetic Radiation

ISTVÁN OZSVÁTH

Southwest Center for Advanced Studies, Dallas, Texas

(Received 30 November 1964)

In this paper I prove the theorem that the Einstein Field equations with dust and electromagnetic null field have only two homogeneous solutions. I have given a discussion of these two solutions in in another paper.

1. INTRODUCTION

IN another paper [1] I found two homogeneous solutions of the equations

$$R_{ik} = -Au_iu_k + (\Lambda + \tfrac{1}{2}A)g_{ik} + 2F_{il}F_k{}^l, \quad (1.1a)$$

$$u_iu^i = 1, \quad F_{ik} = -F_{ki}, \quad F_{ik}F^{ik} = 0, \quad (1.1b)$$

$$F^{ik}{}_{;k} = 0, \quad F^{*ik}{}_{;k} = 0 \quad (1.1c)$$

($A > 0$ is proportional to the density of the incoherent matter and Λ is the cosmological constant), given by

$$ds^2 = (dx^0)^2 - (dx^1)^2 - \tfrac{1}{4}p^{-1}\{(3p+1) + (p-1)\cos^2 x^1\}(dx^2)^2 - 2\cos x^1\, dx^2\, dx^3 - (dx^3)^2, \quad (1.2a)$$

$$p \geq 1, \quad A = 2/(3p+1), \quad \Lambda = \tfrac{1}{2}Ap, \quad (1.2b)$$

$$k_i = -[(p+1)(p-1)/p(3p+1)]^{\frac{1}{2}} \times (1, 0, \cos x^1, 1), \quad (1.2c)$$

$$w_i = (0, -\sin[(3p+1)/4p]^{\frac{1}{2}}(x^3 + x^0), [(3p+1)/4p]^{\frac{1}{2}} \sin x^1 \cos[(3p+1)/4p]^{\frac{1}{2}}(x^3 + x^0), 0), \quad (1.2d)$$

$$F_{ik} = k_{[i}w_{k]}, \quad (1.2e)$$

$$u_i = \frac{1}{2}p^{-\frac{1}{2}}((p+1), 0, -(p-1)\cos x^1, -(p-1)), \quad (1.2f)$$

and

$$ds^2 = -(dx^0)^2 - (dx^1)^2 + [(p-1)/4p]e^{2x^1}(dx^2)^2 + 2e^{x^1}\, dx^2\, dx^3 + (dx^3)^2, \quad (1.3a)$$

$$p \leq -1, \quad A = -2/(3p+1), \quad \Lambda = \tfrac{1}{2}Ap, \quad (1.3b)$$

$$k_i = [(p+1)(p-1)/p(3p+1)]^{\frac{1}{2}} \times (-1, 0, e^{x^1}, 1), \quad (1.3c)$$

$$w_i = \left(0, -\cos\left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}(x^0 - x^3), \left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}e^{x^1}\sin\left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}(x^0 - x^3), 0\right), \quad (1.3d)$$

$$F_{ik} = k_{[i}w_{k]}, \quad (1.3e)$$

$$u_i = 2(-p)^{-\frac{1}{2}}((p+1), 0, (p-1)e^{x^1}, (p-1)). \quad (1.3f)$$

I stated without proof Theorem 1.

Theorem 1. (1.2) and (1.3) are the only homogeneous solutions of (1.1) with vanishing shear and

[1] I. Ozsvath, Two Rotating Universes with Dust and Electromagnetic Field (Volume to be published in honor of V. Hlavaty).

The electromagnetic field has the form

$$F_{ik} = 2a\{k_{[i}l_{k]} \cos\varphi + \tfrac{1}{2}\eta_{ikpq}k^{[p}l^{q]} \sin\varphi\} \quad (7.16)$$

where φ is given by $F = ae^{i\varphi}$ and

$$k_i = (1/a\sqrt{2})(e^{-x^3}, -2\sqrt{2}e^{-2x^3}, 0, 1), \quad (7.17)$$

$$l_i = (1/a\sqrt{2})(e^{-x^3}, -2\sqrt{2}e^{-2x^3}, 0, -1). \quad (7.18)$$

The question remains open whether one could find solutions in the case $\kappa\sigma\lambda\nu \neq 0$ different from (7.14).

ACKNOWLEDGMENTS

I am deeply indebted to I. Robinson for his interest in this work. I wish to express my gratitude to R. Penrose and W. Rindler for valuable discussions about spinors. Special thanks go to W. Rindler for his help in preparing this paper for publication. This work has been supported in part by NASA under grant No. NsG269-62, and in part by the U.S. Air Force Office of Scientific Research, under grant No. AF-AFOSR-903-65.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

All Homogeneous Solutions of Einstein's Field Equations with Incoherent Matter and Electromagnetic Radiation

ISTVÁN OZSVÁTH

Southwest Center for Advanced Studies, Dallas, Texas

(Received 30 November 1964)

In this paper I prove the theorem that the Einstein Field equations with dust and electromagnetic null field have only two homogeneous solutions. I have given a discussion of these two solutions in in another paper.

1. INTRODUCTION

IN another paper [1] I found two homogeneous solutions of the equations

$$R_{ik} = -Au_iu_k + (\Lambda + \tfrac{1}{2}A)g_{ik} + 2F_{il}F_k{}^l, \quad (1.1a)$$

$$u_iu^i = 1, \quad F_{ik} = -F_{ki}, \quad F_{ik}F^{ik} = 0, \quad (1.1b)$$

$$F^{ik}{}_{;k} = 0, \quad F^{*ik}{}_{;k} = 0 \quad (1.1c)$$

($A > 0$ is proportional to the density of the incoherent matter and Λ is the cosmological constant), given by

$$ds^2 = (dx^0)^2 - (dx^1)^2 - \tfrac{1}{4}p^{-1}\{(3p+1) + (p-1)\cos^2 x^1\}(dx^2)^2 - 2\cos x^1\, dx^2\, dx^3 - (dx^3)^2, \quad (1.2a)$$

$$p \geq 1, \quad A = 2/(3p+1), \quad \Lambda = \tfrac{1}{2}Ap, \quad (1.2b)$$

$$k_i = -[(p+1)(p-1)/p(3p+1)]^{\frac{1}{2}} \times (1, 0, \cos x^1, 1), \quad (1.2c)$$

$$w_i = (0, -\sin[(3p+1)/4p]^{\frac{1}{2}}(x^3 + x^0), [(3p+1)/4p]^{\frac{1}{2}} \sin x^1 \cos[(3p+1)/4p]^{\frac{1}{2}}(x^3 + x^0), 0), \quad (1.2d)$$

$$F_{ik} = k_{[i}w_{k]}, \quad (1.2e)$$

$$u_i = \frac{1}{2}p^{-\frac{1}{2}}((p+1), 0, -(p-1)\cos x^1, -(p-1)), \quad (1.2f)$$

and

$$ds^2 = -(dx^0)^2 - (dx^1)^2 + [(p-1)/4p]e^{2x^1}(dx^2)^2 + 2e^{x^1}\, dx^2\, dx^3 + (dx^3)^2, \quad (1.3a)$$

$$p \leq -1, \quad A = -2/(3p+1), \quad \Lambda = \tfrac{1}{2}Ap, \quad (1.3b)$$

$$k_i = [(p+1)(p-1)/p(3p+1)]^{\frac{1}{2}} \times (-1, 0, e^{x^1}, 1), \quad (1.3c)$$

$$w_i = \left(0, -\cos\left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}(x^0 - x^3), \left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}e^{x^1}\sin\left(\frac{3p+1}{4p}\right)^{\frac{1}{2}}(x^0 - x^3), 0\right), \quad (1.3d)$$

$$F_{ik} = k_{[i}w_{k]}, \quad (1.3e)$$

$$u_i = 2(-p)^{-\frac{1}{2}}((p+1), 0, (p-1)e^{x^1}, (p-1)). \quad (1.3f)$$

I stated without proof Theorem 1.

Theorem 1. (1.2) and (1.3) are the only homogeneous solutions of (1.1) with vanishing shear and

[1] I. Ozsvath, Two Rotating Universes with Dust and Electromagnetic Field (Volume to be published in honor of V. Hlavaty).

nonvanishing density of matter. "Vanishing shear" means

$$u_{(i;k)} = 0. \tag{1.4}$$

I now want to consider all the homogeneous solutions of (1.1) and in the course of this investigation there will emerge a proof of the theorem stated above.

One calls a space–time homogeneous if there exists a four-parametric, simply transitive group G_4 of transformations which leaves the metric invariant. It is well known that in a homogeneous space–time one can find four contravariant vector fields $e^i{}_a$ and four covariant vector fields $e^a{}_i$ (a labels the vectors and j the coordinates and $a, \cdots, j, \cdots = 0, 1, 2, 3$), called the invariant and reciprocal vectors of G_4, respectively; they are connected by the equations

$$e^i{}_a e^a{}_k = \delta^i{}_k \quad \text{or} \quad e^a{}_i e^i{}_b = \delta^a{}_b, \tag{1.5}$$

such that the metric has the form

$$g_{ik} = e^a{}_i g_{ab} e^b{}_k, \tag{1.6}$$

where

$$g_{ab} = \text{diag}\,(+1, -1, -1, -1). \tag{1.7}$$

The functions $e^i{}_a$ and $e^a{}_i$ satisfy certain partial differential equations expressing that the vector fields $e^i{}_a$ and $e^a{}_i$ are transformed into themselves under G_4.

Using the quantities $\sigma^i{}_{AB'}$ and $\sigma^{AB'}{}_i$ defined by

$$\sigma^i{}_{11'} = (1/\sqrt{2})(e^i{}_0 + e^i{}_3), \quad \sigma^i{}_{12'} = (1/\sqrt{2})(e^i{}_1 + ie^i{}_2),$$
$$\sigma^i{}_{21'} = (1/\sqrt{2})(e^i{}_1 - ie^i{}_2), \quad \sigma^i{}_{22'} = (1/\sqrt{2})(e^i{}_0 - e^i{}_1)$$

and

$$\sigma_i^{11'} = (1/\sqrt{2})(e_i^0 + e_i^3), \qquad \sigma_i^{12'} = (1/\sqrt{2})(e_i^1 - ie_i^2),$$
$$\sigma_i^{21'} = (1/\sqrt{2})(e_i^1 + ie_i^2), \qquad \sigma_i^{22'} = (1/\sqrt{2})(e_i^0 - e_i^3), \tag{1.8}$$

we can introduce spinors and we find that

1. The components of the spin connection denoted by $\frac{1}{2}A_{ABCD'} = \frac{1}{2}A_{BACD'}$ are constant.

2. The components of a spinor $T_{AC'BD'} \leftrightarrow T_{ik}$ are also constant if T_{ik} is an invariant tensor field, i.e., transforms into itself under G_4.

Exactly as in Ref. 2 one sees that the problem of finding homogeneous solutions of (1.1) is equivalent to the problem of finding a suitable set of constants

$$A, \quad \Lambda, \quad A_{ABCD'}, \quad u_{AB'}, \quad F_{AB} = F_{BA}$$

defined by

$$u_{AB'} \leftrightarrow u_i, \qquad F_{AC'BD'} = \tfrac{1}{2}(F_{AB}\epsilon_{C'D'} + \bar{F}_{C'D'}\epsilon_{AB}) \leftrightarrow F_{ik} \tag{1.9}$$

satisfying the algebraic equations

$$\phi_{ABC'D'} = \tfrac{1}{2}A(u_{AC'}u_{BD'} + u_{AD'}u_{BC'}) + F_{AB}\bar{F}_{C'D'}, \tag{1.10a}$$

$$\Lambda + \tfrac{1}{4}A = \tfrac{1}{2}\chi_{AB}{}^{AB}, \tag{1.10b}$$

$$u_{AB'}u^{AB'} = 1, \qquad F_{AB}F^{AB} = 0, \tag{1.10c}$$

$$F^{AB}{}_{:BC'} = 0, \tag{1.10d}$$

$$C_{ABCD'}(A^{CD'} + \bar{A}^{D'C}) = 0, \tag{1.10e}$$

where

$$A_{AD'} = A_{AB}{}^{B}{}_{D'}, \tag{1.11}$$

$$C_{ABCD'} = A_{C(AB)D'} + \bar{A}_{D'(A}\epsilon_{B)C}, \tag{1.12}$$

$$\chi_{ABCD} = \tfrac{1}{4}(A_{PACQ'}A^{P}{}_{BD}{}^{Q'} + A_{PADQ'}A^{P}{}_{BC}{}^{Q'}) + \tfrac{1}{2}A_{ABPQ'}C_{CD}{}^{PQ'}, \tag{1.13}$$

$$\phi_{ABC'D'} = \tfrac{1}{4}(A_{PAQC'}A^{P}{}_{B}{}^{Q}{}_{D'} + A_{PAQD'}A^{P}{}_{B}{}^{Q}{}_{C'}) + \tfrac{1}{2}A_{ABPQ'}\bar{C}_{C'D'}{}^{Q'P}, \tag{1.14}$$

and

$$F_{AB:CD'} = \tfrac{1}{2}(F^{P}{}_{B}A_{PACD'} + F_{A}{}^{P}A_{PBCD'}). \tag{1.15}$$

From the fact that A is constant and from the twice contracted Bianchi identities and from (1.1) it follows that

$$u^i{}_{;i} = 0 \tag{1.16}$$

and

$$u_{i;k}u^k = 0. \tag{1.17}$$

The spinor equivalents of these equations are

$$u^{AB'}{}_{:AB'} = 0 \tag{1.18}$$

and

$$u_{AB':CD'}u^{CD'} = 0, \tag{1.19}$$

where

$$u_{AB':CD'} = \tfrac{1}{2}(u^{P}{}_{B'}A_{PACD'} + u_{A}{}^{P'}\bar{A}_{P'B'D'C}). \tag{1.20}$$

2. THE EXPLICIT SYSTEM

In writing down our equations explicitly we are at liberty to choose the time-like unit vector $e^i{}_0$ such that

$$e^i{}_0 = u^i \tag{2.1}$$

and the spacelike unit vector $e^i{}_3$ such that the null vector k^i defined by F^{ik} takes the form

$$k^i \propto e^i{}_0 + e^i{}_3. \tag{2.2}$$

[2] I. Ozsvath, Homogeneous Solutions of the Einstein–Maxwell Equation J. Math. Phys. 6, 590 (1965).

It then follows that $u_{AB'}$ and F_{AB} have the form

$$u_{AB'} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{2.3}$$

$$F_{AB} = \begin{pmatrix}0 & 0\\ 0 & F\end{pmatrix} \quad (F \neq 0 \text{ a complex constant}). \tag{2.4}$$

We still have the freedom to make rotations in the 1–2 plane, i.e., we have the spin transformations

$$\mathfrak{L}^A{}_B = \begin{bmatrix}e^{i\varphi} & 0\\ 0 & e^{-i\varphi}\end{bmatrix} \tag{2.5}$$

at our disposal to simplify the components of $A_{ABCD'}$ at a convenient stage of our calculations.

We see immediately that the equations (1.10c) are satisfied. We denote our unknowns, the components of $A_{ABCD'}$ by the following symbols:

$A_{ABCD'}$:

AB \ CD'	11′	12′	21′	22′
11	κ	σ	ρ	τ
(12)	ϵ	β	π	γ
22	π	μ	λ	ν

(2.6)

One sees immediately that Maxwell's equations take the form

$$\kappa = 0,\ \sigma = 0,\ \rho = 2\epsilon,\ \tau = 2\beta. \tag{2.7}$$

From (1.18) and (1.19) we see that

$$\epsilon + \bar{\epsilon} + \gamma + \bar{\gamma} = \mu + \bar{\mu} = 0, \tag{2.8a}$$

$$\pi + \nu = 2\bar{\beta}. \tag{2.8b}$$

From the equation $\phi_{111'1'} = \frac{1}{2}A > 0$ we get

$$-\epsilon(3\epsilon + \bar{\epsilon}) = \tfrac{1}{2}A. \tag{2.9}$$

Since A is real it follows that $(\epsilon + \bar{\epsilon})(\epsilon - \bar{\epsilon}) = 0$ and from $A > 0$ we get

$$\epsilon + \bar{\epsilon} = 0. \tag{2.10}$$

Equation (2.8a) leads to

$$\gamma + \bar{\gamma} = 0, \qquad \mu + \bar{\mu} = 0. \tag{2.11}$$

Using (2.5) we can arrange without loss of generality that

$$\beta + \bar{\beta} = 0. \tag{2.12}$$

The equations (2.8b) and (2.9) take the form

$$\pi + \nu = -2\beta, \tag{2.13}$$

$$-4\epsilon^2 = A. \tag{2.14}$$

We get from $\phi_{121'1'} = 0$ and $\phi_{111'2'} = 0$ the equations

$$3\pi - \alpha + \beta = 0 \quad \text{and} \quad \pi + \alpha - 9\beta = 0, \tag{2.15}$$

and therefore, using (2.13), we have

$$\pi = 2\beta,\ \nu = -4\beta,\ \alpha = 7\beta. \tag{2.16}$$

Now it is very easy to compute the other equations of (1.10)(1). There are as follows:

$$32\beta^2 - \lambda\bar{\lambda} - \mu^2 = A + 2F\bar{F}, \tag{2.17a}$$

$$60\beta^2 - 4\epsilon(\gamma + \mu) = A, \tag{2.17b}$$

$$10\beta^2 - \epsilon\lambda = 0, \tag{2.17c}$$

$$\beta\{4\epsilon - 3\mu - 6\gamma + 7\lambda\} = 0, \tag{2.17d}$$

$$\beta\{-4\epsilon + 5\mu + 2\gamma - 5\lambda\} = 0. \tag{2.17e}$$

One sees easily that (1.10e) are already contained in (2.17d, e). The equation (1.10b) takes the form

$$\Lambda + \tfrac{1}{4}A = 11\beta^2 - \epsilon\gamma. \tag{2.18}$$

Our explicit algebraic system is therefore (2.14), (2.17), and (2.18), and we have to find all the solutions of this system. Before doing that we calculate the components of the shear spinor defined by

$$\sigma_{ABC'D'} \leftrightarrow u_{(i;k)}. \tag{2.19}$$

We find

$\sigma_{ABC'D'}$:

AB \ $C'D'$	1′1′	(1′2′)	2′2′
11	0	2β	$-\frac{1}{2}\bar{\lambda}$
(12)	-2β	0	$-2\bar{\beta}$
22	$\frac{1}{2}\lambda$	2β	0

(2.20)

If $\beta = 0$, then, since $\epsilon \neq 0$, it follows from (2.17c) that $\lambda = 0$, and that the shear vanishes. Therefore we can distnguish two different cases,

Case 1: $\beta \neq 0$; i.e., shear is present;

Case 2: $\beta = 0$; i.e., the shear vanishes.

We now discuss these cases separately.

Case 1

$$\beta \neq 0. \tag{2.21}$$

From (2.17c, d) we get

$$\lambda = \tfrac{2}{3}(3\gamma - \epsilon),\ \mu = \tfrac{2}{3}(2\gamma + \epsilon), \tag{2.22}$$

and substituting into our system we arrive at

$$A = -4\epsilon^2, \tag{2.23a}$$

$$5F\bar{F} = 2[8\beta^2 + (\gamma - \epsilon)^2 + 4\epsilon^2], \tag{2.23b}$$

$$25\beta^2 = \epsilon(3\gamma - \epsilon), \tag{2.23c}$$

$$\Lambda = 11\beta^2 + \epsilon(\epsilon - \gamma). \tag{2.23d}$$

One sees that the right-hand side of (2.23b) ≤ 0, and the left-hand side ≥ 0, therefore a solution

could exist only if the magnetic field vanishes, but then from $8\beta^2 + (\gamma - \epsilon)^2 + 4\epsilon^2 = 0$ it follows that $\beta = \epsilon = \gamma = 0$ which is a contradiction. We can therefore state theorem 2.

Theorem 2. The equations (1.1) have no homogeneous solution with nonvanishing shear.

Case 2

$$\beta = 0. \tag{2.24}$$

It then follows from (2.17)(3) that

$$\lambda = 0, \tag{2.25}$$

and we are left with the equations

$$A = -4\epsilon^2,\ A + 2F\bar{F} = -\mu^2,$$
$$A = -4\epsilon(\gamma + \mu),\ \Lambda = \epsilon(\epsilon - \gamma). \tag{2.26}$$

It follows immediately that $\epsilon = \gamma + \mu$ or

$$\gamma = \epsilon - \mu. \tag{2.27}$$

Collecting our results, we can write

$A_{ABCD'}$:

AB	CD' 11′	12′	21′	22′
11	0	0	2ϵ	0
(12)	ϵ	0	0	$\epsilon-\mu$
22	0	μ	0	0

(2.28)

$$A = -4\epsilon^2,\ F\bar{F} = \tfrac{1}{2}(2\epsilon - \mu)(2\epsilon + \mu),\ \Lambda = \epsilon\mu,$$

where ϵ and μ are arbitrary imaginary numbers subject to the conditions

$$\epsilon \neq 0,\ (2\epsilon - \mu)(2\epsilon + \mu) \geq 0. \tag{2.29}$$

Introducing new parameters by the equations

$$\epsilon = i\sqrt{2}a,\ \mu = -2\sqrt{2}iap, \tag{2.30}$$

we get

$$A = 8a^2,\ F\bar{F} = 4a^2(p + 1)(p - 1) \text{ and } \Lambda = \tfrac{1}{2}Ap \tag{2.31}$$

with the condition

$$a \neq 0,\ |p| \geq 1. \tag{2.32}$$

We have seen in Ref. 2 that the Ricci rotation coefficients A_{abc} of the tetrad $e^i{}_a$ can be calculated by

$$A_{abc} = \sigma^{AA'}{}_a\sigma^{BB'}{}_b\sigma^{CC'}{}_c\tfrac{1}{2}(A_{ABCC'}\epsilon_{A'B'} + \bar{A}_{A'B'C'C}\epsilon_{AB}), \tag{2.33}$$

where

$$\sigma^{AA'}{}_a = \sigma^{AA'}{}_i e^i{}_a, \tag{2.34}$$

which can be written in matrix form

$$\sigma^{AA'}{}_a = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}. \tag{2.35}$$

The result of this calculation is given by

A_{abc}:

ab	c 1	2	3	4
23	0	$-a(p+1)$	0	0
31	0	0	$-a(p+1)$	0
12	$2a(p+1)$	0	0	$-2ap$
10	0	0	$-a(p-1)$	0
20	0	$a(p-1)$	0	0
30	0	0	0	0

(2.36)

Applying the formulas

$$C_{abc} = -(A_{cab} - A_{cba}),\quad C_{ab}{}^c = g^{cf}C_{abf} \tag{2.37}$$

[see (1.7)], we get the group with the commutator relations

$$(X_2, X_3) = -a(3p + 1)X_1,$$
$$(X_3, X_1) = -a(3p + 1)X_2,$$
$$(X_1, X_2) = -2a\{(p - 1)X_0 + (p + 1)X_3\}, \tag{2.38}$$
$$(X_1, X_0) = -a(3p + 1)X_2,$$
$$(X_2, X_0) = a(3p + 1)X_1, \qquad (X_3, X_0) = 0,$$

and the metric is given by (1.6) if $e^a{}_i$ denotes a set of reciprocal vectors of (2.38). We have seen in Ref. 1 that we are thus led to (1.2) and (1.3) only, and, since (2.28) is the only solution of (1.10) with nonvanishing density of matter, we have established Theorem 1. Theorems 1 and 2 lead to Theorem 3.

Theorem 3. (1.2) and (1.3) are the only homogeneous solutions of (1.1).

ACKNOWLEDGMENT

I am deeply indebted to R. Penrose and I. Robinson for their interest in this work. Special thanks go to W. Rindler for valuable discussions and for his help in preparing this paper for publication. This research was partially supported by NASA under grant number NsG 269-62 and by the Air Force Office of Scientific Research under grant number AF-AFOSR-709-64.

Particlelike Solutions to Nonlinear Scalar Wave Theories

GERALD ROSEN
Southwest Research Institute, San Antonio, Texas
(Received 11 January 1965)

According to a recent theorem proved by Derrick, no absolutely stable time-independent particlelike solution of finite energy is obtainable from a large class of Lorentz-covariant scalar wave theories. We study a solvable nonlinear scalar wave theory and derive a rigorous *metastable* particlelike solution of finite energy, a quasistatic solution having a rate of dissolution which is free to be arbitrarily small relative to the associated particle rest mass. Derrick's theorem notwithstanding, the specific example presented here suggests that particlelike quasistatic solutions to a nonlinear scalar wave theory may still be of some relevancy to meson field physics, where no absolutely stable but instead metastable elementary particles are present.

I. INTRODUCTION

IN a recent paper, Derrick[1] proves a theorem that precludes the existence of static (time-independent) *stable* solutions of finite energy for a wide class of nonlinear wave equations, namely, for scalar (or pseudoscalar) field theories derived from a Lagrangian density of the generic form

$$\mathcal{L} = (\dot{\theta})^2 - (\nabla\theta)^2 - f(\theta), \tag{1.1}$$

where the admissible wave field $\theta = \theta(\mathbf{x}, t)$ is a piecewise C^2 scalar function with respect to $\mathbf{x}$ and t and $f(\theta)$ is a certain piecewise C^2 function of θ. [The continuity classes prescribed here for the functions $\theta(\mathbf{x}, t)$ and $f(\theta)$ are sufficient for Derrick's original proof and for the dynamical stability generalization[2] shown in Appendix A.] According to Derrick's result, the nonlinear wave equation which follows from a Lagrangain density (1.1),

$$\ddot{\theta} - \nabla^2\theta + \tfrac{1}{2}f'(\theta) = 0, \tag{1.2}$$

has no time-independent localized solution $\theta = \theta_0(\mathbf{x})$ that is stable[2] and with a finite static field energy (a finite "particle rest mass"),

$$E_0 = E_0[\theta_0] \equiv \int [(\nabla\theta_0)^2 + f(\theta_0)]\, d^3\mathbf{x}. \tag{1.3}$$

Thus, within the realm of purely classical field theory, the theorem of Derrick asserts that no stable spinless particlelike solution is obtainable from a large class of Lorentz-covariant scalar wave equations.

It is important to note that Derrick's result does not preclude the existence of *metastable* particlelike solutions of finite energy, solutions having rates of dissolution which are very small relative to their quantum-theoretic characteristic frequencies $E_0/\hbar$ based on the rest mass energy (1.3). This indeed is the actual situation in nature, for there are a certain number of well-established metastable spinless elementary particles (π and K mesons), but no stable spinless elementary particle is known to exist. In the light of Derrick's theorem, the question arises as to whether metastable (quasi-static) particlelike solutions of finite energy can be derived from a nonlinear scalar field theory based on a Lagrangian density (1.1). An affirmative answer to this question is given in the present paper. By concentrating attention on a specific nonlinear scalar wave theory, we are able to present an example of a static particlelike solution which can indeed be metastable if an appropriately large value is assigned to a certain constant of integration. In other words, we obtain a rigorous quasistatic particlelike solution having a rate of dissolution which is free within the classical field theory to be arbitrarily small relative to the associated particle rest mass energy E_0.

II. A SOLVABLE NONLINEAR SCALAR WAVE THEORY

We consider the theory based on the Lagrangian density (1.1) with

$$f(\theta) \equiv -g\theta^6 \qquad (g \equiv \text{positive physical constant}), \tag{2.1}$$

and so the associated scalar wave equation (1.2) takes the form

$$-\ddot{\theta} + \nabla^2\theta + 3g\theta^5 = 0. \tag{2.2}$$

[1] G. H. Derrick, J. Math. Phys. **5**, 1252 (1964). Also see U. Enz, Phys. Rev. **131**, 1392 (1963), and papers cited therein.

[2] That the second variation of the energy functional (1.3) about θ_0 should be nonnegative is the stability criterion evoked by Derrick. In Appendix A we show that Derrick's necessary condition for a stable θ_0, the requirement $\delta^2 E_0 \geq 0$ about θ_0, is in fact necessary and sufficient for a θ_0 that is *dynamically stable* in the more general sense of Liapunov.

The singularity-free static and spherically symmetric solution to Eq. (2.2) is given by[3]

$$\theta = \theta_0 \equiv Z(Z^4 g + r^2)^{-\frac{1}{4}}, \tag{2.3}$$

in which $r \equiv |\mathbf{x}|$ and the "size parameter" Z is a free (positive or negative) real constant of integration.[4] It is an elementary matter to verify that (2.3) satisfies Eq. (2.2):

$$\nabla^2\theta_0 = \frac{1}{r}\frac{d^2}{dr^2}(r\theta_0) = \frac{Z}{r}\frac{d^2}{dr^2}\left(\frac{Z^4 g}{r^2} + 1\right)^{-\frac{1}{4}}$$

$$= \frac{Z^5 g}{r}\frac{d}{dr}(Z^4 g + r^2)^{-\frac{3}{4}} = -3Z^5 g(Z^4 g + r^2)^{-5/2}$$

$$\equiv -3g\theta_0^5. \tag{2.4}$$

The static field energy (1.3) associated with the solution (2.3) is also computed easily:

$$E_0 = 4\pi\int_0^\infty \left[\left(\frac{d\theta_0}{dr}\right)^2 - g\theta_0^6\right] r^2\, dr$$

$$= 4\pi Z^2 \int_0^\infty \frac{(r^4 - Z^4 g r^2)}{(Z^4 g + r^2)^3}\, dr$$

$$= \frac{4\pi}{g^{\frac{1}{2}}}\int_0^\infty \frac{(s^4 - s^2)\, ds}{(1 + s^2)^3} = \frac{\pi^2}{2g^{\frac{1}{2}}}. \tag{2.5}$$

That the static field energy $E_0 = \pi^2/2g^{\frac{1}{2}}$ is entirely independent of the size parameter Z in the solution (2.3) was to be expected, because the general expression for the static field energy (1.3) with (2.1) is a scale-invariant quantity,[4] and therefore the rest mass of any of the particlelike solutions is prefixed in the theory. Also of some interest is the nonanalytic character of E_0 about $g = 0$, showing that the nonlinear term in (2.2) is not generally amenable to a rigorous perturbation-theory treatment even if g is taken arbitrarily small.

Let us now consider the dynamical stability of the solution (2.3). With the perturbed general solution about θ_0 given by

$$\theta = \theta_0 + \sum_{k,l,m} \frac{\xi_{kl}}{r}\,\mathrm{Re}\,[c_{klm}e^{ikt}Y_l^m], \tag{2.6}$$

in which the ξ's are real functions of r, the c's are complex constants (small in magnitude but otherwise arbitrary), and the Y's are the well-known complex spherical harmonics; the linearization of (2.2) with (2.6) produces an eigenvalue equation for the ξ's

$$\frac{d^2\xi_{kl}}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2} + 15g\theta_0^4\right]\xi_{kl} = 0 \tag{2.7}$$

which must be supplemented here with the appropriate boundary conditions for a singularity-free localized perturbation,

$$\xi_{kl}(0) = 0, \qquad \lim_{r\to\infty}\left[\frac{\xi_{kl}(r)}{r}\right] = 0. \tag{2.8}$$

Equations (2.7) and (2.8) constitute a Sturm–Liouville–Schrödinger eigenvalue problem in which k^2 plays the role of an "energy" eigenvalue and the quantity

$$-15g\theta_0^4 = -15Z^4 g(Z^4 g + r^2)^{-2} \tag{2.9}$$

acts like an attractive "potential." In conformity with Derrick's theorem, there is a ground state with a negative energy eigenvalue associated with the effective potential (2.9). That is, there exists an $l = 0$ eigenfunction ξ_{k0} with k^2 a minimum and negative in value, thus with k purely imaginary, and so the associated perturbation term in (2.6) generally grows exponentially with time in a dynamically unstable fashion. By performing some straightforward analysis, the "ground state" eigenvalue $\min k^2 \equiv -\lambda_0^2$ (< 0) is determined approximately in Appendix B, from which we obtain the approximate rate of exponential dissolution of the solution (2.3),

$$\lambda_0 \cong (1.9)/Z^2 g^{\frac{1}{2}}. \tag{2.10}$$

We note that λ_0^{-1} is of the order of the characteristic time for propagation of infinitesimal disturbances through the particlelike solution (2.3) ("particle radius" of the order $Z^2 g^{\frac{1}{2}}$), and so the result (2.10) is consonant with naïve physical intuition. It follows from (2.5) and (2.10) that the dissolution rate to rest energy ratio

$$\lambda_0/E_0 \cong (0.39)/Z^2 \tag{2.11}$$

can be made arbitrarily small by letting the absolute value of the size parameter $|Z|$ take on a sufficiently large value.

In summary then, the static particlelike solution (2.3) has a finite rest energy and is metastable provided that $|Z|$ is large, corresponding to a solution which is relatively small in maximum field magnitude but relatively large in spatial extension. Such a solution, one which is not highly localized or concentrated about a point in space but rather

[3] The static and spherically symmetric specialization of (2.2) produces a so-called Emden equation, familiar in astrophysics, e.g., S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (University of Chicago Press, Chicago, Illinois, 1939). However, Eq. (2.2) itself is quite distinct from the dynamical equations ordinarily encountered in astrophysics.

[4] It should be noted that the "size parameter" Z is a constant of homology, stemming from the scale invariance of Eq. (2.2), $\theta(\mathbf{x}, t) \to \mu\theta(\mu^2\mathbf{x}, \mu^2 t)$ $(\mu \neq 0)$. In general, this scale invariance gives rise to equivalence classes for the solutions to Eq. (2.2) with Z parameterizing the members of a particular equivalence class of solutions in Eq. (2.3).

global in character, is indeed more in harmony with the qualitative notion of a "classical particle" that is obtained by applying a correspondence principle argument to the quantum field theoretic description of a one-particle state. If the term "particlelike" is understood to embrace time-independent solutions of finite energy that have a rather global (instead of a highly localized) character, the specific model theory considered here suggests that such "particle-like" solutions to nonlinear scalar wave theories may still be of some relevancy in meson field physics.

Note added in proof: A paper by R. H. Hobart [Proc. Phys. Soc. (London) **82**, 201 (1963)] has come to the author's attention. Although not cited by Derrick,[1] this paper by Hobart establishes the instability of nonsingular time-independent *spherically symmetric* solutions to equations having the generic form (1.2). Derrick's theorem follows as a natural extension of Hobart's result for spherically symmetric solutions.

APPENDIX A: EQUIVALENCE OF STABILITY CRITERIA FOR GENERAL STATIC SOLUTIONS OF FINITE ENERGY

Here we show that Derrick's necessary condition for a stable θ_0, the static energy requirement

$$\delta^2 E \equiv \tfrac{1}{2}[(d^2/d\epsilon^2)E_0[\theta_0 + \epsilon\omega]]_{\epsilon=0} = \int [(\nabla\omega)^2 + \tfrac{1}{2}f''(\theta_0)\omega^2]\, d^3\mathbf{x} \geq 0 \qquad \text{(A1)}$$

with both $\theta_0 = \theta_0(\mathbf{x})$ and $\omega = \omega(\mathbf{x})$ independent of time and piecewise C^2 functions with respect to $\mathbf{x}$, is in fact a necessary and sufficient condition for a θ_0 that is *dynamically stable* in the sense of Liapunov.[5] To derive the dynamical stability criterion, we make the perturbed field depend on time by putting

$$\theta(\mathbf{x}, t) = \theta_0(\mathbf{x}) + \omega(\mathbf{x}) \cos kt, \qquad \text{(A2)}$$

where the constant k may be either purely real or purely imaginary and $|\omega(\mathbf{x})| \ll |\theta_0(\mathbf{x})|$ for all values of $\mathbf{x}$. By substituting (A2) into (1.2) and retaining only the terms linear in ω, we obtain an eigenvalue equation for k^2 and ω,

$$(\nabla^2 - \tfrac{1}{2}f''(\theta_0) + k^2)\omega = 0, \qquad \text{(A3)}$$

which can be recast in the form of a variational principle,

$$\delta k^2 = 0\,, \qquad k^2 \equiv \int [(\nabla\omega)^2 + \tfrac{1}{2}f''(\theta_0)\omega^2]\, d^3\mathbf{x} \left[\int \omega^2\, d^3\mathbf{x}\right]^{-1}. \qquad \text{(A4)}$$

[5] See, for example: W. Hahn, *Theory and Application of Liapunov's Direct Method* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963), pp. 5–10.

Now if $f''(\theta_0)$ is piecewise continuous and if k^2 is negative for a certain admissible (piecewise C^2 function) ω, k^2 will be stationary about some negative value of k^2; we shall then have a purely imaginary eigenvalue for k and thus a dynamically unstable static solution $\theta_0(\mathbf{x})$, according to (A2). Hence, with the hypothesis that $f(\theta)$ is piecewise C^2, a comparison of (A1) and (A4) shows that Derrick's static energy requirement $\delta^2 E_0 \geq 0$ for all admissible ω is actually necessary and sufficient for a dynamically stable solution θ_0.

APPENDIX B: GROUND STATE EIGENVALUE FOR EQS. (2.7) AND (2.8)

By substituting (2.9) into Eq. (2.7) and setting

$$\rho \equiv r/Z^2 g^{\frac{1}{2}}, \quad \gamma \equiv Z^2 g^{\frac{1}{2}}\lambda_0 \equiv Z^2 g^{\frac{1}{2}}(-\min k^2)^{\frac{1}{2}}, \qquad \text{(B1)}$$

we obtain the dimensionless eigenvalue equation for the "ground state"

$$\frac{d^2\xi_\gamma}{d\rho^2} + \left[\frac{15}{(1+\rho^2)^2} - \gamma^2\right]\xi_\gamma = 0. \qquad \text{(B2)}$$

Here, the "ground state" eigenfunction $\xi_\gamma \equiv \xi_{0k}$ with $l = 0$ and $k^2 \equiv -\lambda_0^2$ a (negative valued) minimum is associated with the rate-controlling unbounded perturbation term in (2.6). That Eq. (B2) has a bound "ground state" solution with a real $\gamma > 0$ such that

$$\xi_\gamma(0) = 0, \qquad \lim_{\rho\to\infty}\left[\frac{e^{\gamma\rho}\xi_\gamma(\rho)}{\rho}\right] = 0 \qquad \text{(B3)}$$

is confirmed most readily by considering the eigenfunctions associated with a simple (mathematically tractable) potential (e.g., a square-well attractive potential) $\tilde{V}(\rho)$ such that $0 \geq \tilde{V}(\rho) \geq -15(1+\rho^2)^{-2}$ for all real positive values of ρ. Alternatively, by considering the $\gamma = 0$ eigenfunction associated with Eqs. (B2) and (B3), an eigenfunction given explicitly in closed form by the algebraic expression[6]

$$\xi_0 = (\rho - \rho^3)(1+\rho^2)^{-\frac{3}{2}}, \qquad \text{(B4)}$$

which exhibits a node at $\rho = 1$, we infer[7] the existence of one (unique) lower energy state, necessarily with $\gamma > 0$ and *without* a node occurring for some interior value of ρ. Since the "ground state" eigenfunction ξ_γ and its eigenvalue γ cannot be obtained by exact mathematical analysis, we work out two mutually corroborating approximate solutions of the eigenvalue problem in the following paragraphs. The first approximate solution is based on a novel

[6] E. Kamke, *Differentialgleichungen Lösungsmethoden und Lösungen* (Akademische Verlags., Leipzig, 1956), p. 494.

[7] For example: R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. I, p. 458.

heuristic method, while the second approximate solution involves a more direct and foolproof Rayleigh–Ritz procedure.

First, by putting

$$\xi_\gamma \equiv e^{-\gamma\rho}(1+\rho^2)^{-\frac{1}{4}}\zeta_\gamma \tag{B5}$$

into Eq. (B2), we find an equation for the new dependent variable ζ_γ,

$$(1+\rho^2)\frac{d^2\zeta_\gamma}{d\rho^2} - [2\gamma(1+\rho^2)+6\rho]\frac{d\zeta_\gamma}{d\rho} + (12+6\gamma\rho)\zeta_\gamma = 0, \tag{B6}$$

while the boundary conditions (B3) take the form

$$\zeta_\gamma(0) = 0, \qquad \lim_{\rho\to\infty}\left[\frac{\zeta_\gamma(\rho)}{\rho^4}\right] = 0. \tag{B7}$$

Note that the algebraic denominator in Eq. (B5), suggested by the exact $\gamma=0$ solution (B4), eliminates the second-order character of the $\rho = \pm i$ poles manifest in (B2) with only simple zeros evident in the coefficients of the transformed Eq. (B6). The general Fuchsian theory guarantees that the relevant solution of (B6) is analytic about $\rho = 0$ and is thus expressible as a convergent power series for $|\rho| < 1$,

$$\zeta_\gamma = \sum_{n=1}^{\infty} a_n\rho^n, \tag{B8}$$

in which the a's are given by a recurrence relation derived from (B6),

$$a_{n+2} = \frac{2\gamma}{(n+2)}a_{n+1} - \frac{(n-4)(n-3)}{(n+2)(n+1)}a_n + \frac{2\gamma(n-4)}{(n+2)(n+1)}a_{n-1} \quad (n\geq 2), \tag{B9}$$

with

$$a_1 \equiv 1, \qquad a_2 = \gamma, \qquad a_3 = \tfrac{2}{3}\gamma^2 - 1. \tag{B10}$$

From (B9) and (B10) it follows that

$$a_4 = \tfrac{1}{3}\gamma^3 - \gamma, \quad a_5 = \tfrac{2}{15}\gamma^4 - \tfrac{1}{2}\gamma^2, \quad a_6 = \tfrac{2}{45}\gamma^5 - \tfrac{1}{6}\gamma^3, \quad a_7 = \tfrac{4}{315}\gamma^6 - \tfrac{4}{105}\gamma^4 - \tfrac{1}{42}\gamma^2. \tag{B11}$$

Now observe that the eigenvalue condition

$$\gamma = \tfrac{1}{2}(15)^{\frac{1}{2}} \cong 1.93 \tag{B12}$$

makes $a_5 = a_6 = 0$, and so (B8) reduces to the form

$$\zeta_{(15)^{1/2}} = \rho + \tfrac{1}{2}(15)^{\frac{1}{2}}\rho^2 + \tfrac{3}{2}\rho^3 + \tfrac{1}{8}(15)^{\frac{1}{2}}\rho^4 + \tfrac{15}{336}\rho^7 + O(\rho^8), \tag{B13}$$

with the terms up to order ρ^6 being put in closest possible accord with the functional form for $\rho \sim 1$ suggested by the second boundary condition in (B7). We regard the first four terms in (B13) as an asymptotic expansion for $\zeta_{(15)^{1/2}}$ with ρ of the order or not much greater than unity, an approximate form for ζ_γ in closest possible agreement with the asymptotic behavior required by the second boundary condition in (B7). Thus, we can tentatively regard (B12) and (B13) as an approximate solution to the "ground state" eigenvalue problem.

To corroborate the preceding analysis, let us set up a variational principle for the solution to Eq. (B2) and then apply a Rayleigh–Ritz approximation procedure. Here it is convenient to introduce the new independent and dependent variables

$$\phi \equiv \tan^{-1}\rho \qquad (0 \leq \phi \leq \tfrac{1}{2}\pi), \qquad \omega_\gamma(\phi) \equiv (\cos\phi)\xi_\gamma(\rho). \tag{B14}$$

In terms of these new quantities, (B2) is transformed to the equation

$$\frac{d^2\omega_\gamma}{d\phi^2} + \left[16 - \frac{\gamma^2}{(\cos\phi)^4}\right]\omega_\gamma = 0, \tag{B15}$$

which leads to the variational principle

$$\delta\gamma^2 = 0 \quad \gamma^2 \equiv \int_0^{\frac{1}{2}\pi}\left[16\omega_\gamma^2 - \left(\frac{d\omega_\gamma}{d\phi}\right)^2\right]d\phi \tag{B16}$$

with ω_γ subject to the normalization condition

$$\int_0^{\frac{1}{2}\pi}\frac{(\omega_\gamma)^2}{(\cos\phi)^4}\,d\phi = 1 \tag{B17}$$

and the boundary conditions

$$\omega_\gamma(0) = 0, \qquad \omega_\gamma(\tfrac{1}{2}\pi) = 0. \tag{B18}$$

We seek an approximate solution of the form

$$\omega_\gamma = 2\pi^{-\frac{1}{2}}(\cos\phi)^2(\alpha\sin 2\phi + \beta\sin 4\phi), \tag{B19}$$

where α and β are variational parameters, constrained by (B17) to satisfy

$$\alpha^2 + \beta^2 = 1. \tag{B20}$$

By putting (B19) into the definition part of (B16) we have

$$\gamma^2 = 3\alpha^2 + 3\alpha\beta - \tfrac{1}{2}\beta^2, \tag{B21}$$

and thus obtain the maximizing conditions for γ^2

$$(6-2\gamma^2)\alpha + 3\beta = 0, \qquad 3\alpha - (1+2\gamma^2)\beta = 0, \tag{B22}$$

which produce

$$\gamma = \tfrac{1}{2}[5 + (85)^{\frac{1}{2}}]^{\frac{1}{2}} \cong 1.88, \tag{B23}$$

as well as the mixing ratio $\alpha/\beta = \tfrac{1}{6}[7 + (85)^{\frac{1}{2}}] \cong 2.70$.

A comparison of (B12) and (B23) shows that the two approximate solutions of the eigenvalue problem are mutually consistent and give $\gamma \cong 1.90$ to better than 2%. Inverting the definition of λ_0 in (B1), we finally obtain the estimate stated in (2.10).

On Projective Representations of Finite Groups

P. RUDRA
Saha Institute of Nuclear Physics, Calcutta, India
(Received 9 November 1964)

The algebra of projective representation, belonging to a factor system, has been presented. Exact expressions for the projection operators, Kronecker (inner) direct-product representation, have been obtained. Formulas for obtaining the characters of all the inequivalent irreducible projective representations and the Clebsch–Gordan coefficients have been derived.

1. INTRODUCTION

VECTOR representation theory of finite groups have long been formulated[1,2] from the view point of Physics. Applications of projective (or ray) representation[2] in physical problems have so far been few and far between.[3,4] Moreover, the expressions for different algebraic relations of the representations have not been presented in a systematic manner. Many of the expressions for the vector representations will be suitably modified when we deal with the projective representations. Here we have derived expressions for the transformation rule for the basis functions, projection operators, orthogonality relations, reduction of a reducible representation into direct sum of irreducible components, Kronecker (inner) direct product, and Clebsch–Gordan coefficients of the projective representation, when the factor systems are known. The scheme of presentation of the material follows that of Wigner[1] and Hamermesh[2] in the case of vector representations.

The application of the projective representation theory in the space groups of crystals will be dealt with subsequently.[5]

2. BASIC CONCEPTS AND RELATIONS[2]

If corresponding to each $P \in G$ (where G is the group) we have a matrix $\Gamma(P)$ such that

$$\Gamma(P)\Gamma(Q) = \omega_{P,Q}\Gamma(PQ)$$

with

$$|\omega_{P,Q}| = 1 \quad \text{for all} \quad P, Q \in G, \tag{1}$$

we call the set of matrices $\Gamma(P)$'s a "projective (or ray) representation" of the group belonging to the particular factor system $\omega_{P,Q}$. If the correspondence is one-to-one, we call the representation "faithful."

The factors $\omega_{P,Q}$'s satisfy the relation

$$\omega_{P,Q}\omega_{PQ,R} = \omega_{P,QR}\omega_{Q,R}. \tag{2}$$

Conversely, if there are g^2 nonvanishing constants satisfying Eq. (2), this set form a projective representation. (g is the order of the group G).

The concepts of equivalence and reducibility is the same as in vector representations. Associated with a particular factor system, there will be a finite number of inequivalent irreducible representation. Here we are concerned with the properties of the inequivalent irreducible representations, if the factor system is given.

The matrix $\Gamma(E)$, associated with the identity element is a unit matrix, having the dimension of the representation. Hence

$$\Gamma(P)\Gamma(E) = \omega_{P,E}\Gamma(P)$$

and

$$\Gamma(E)\Gamma(P) = \omega_{E,P}\Gamma(P)$$

will give rise to the relation, since $\Gamma(E)$ commutes with all the matrices and its product with any matrix will give the same matrix,

$$\omega_{P,E} = \omega_{E,P} = 1 \tag{3}$$

for all $P \in G$. Also, since

$$\Gamma(P)\Gamma(P^{-1}) = \omega_{P,P^{-1}}\Gamma(E)$$

and

$$\Gamma(P^{-1})\Gamma(P) = \omega_{P^{-1},P}\Gamma(E),$$

we have

$$\Gamma(P)^{-1} = [\omega_{P,P^{-1}}]^{-1}\Gamma(P^{-1}) = [\omega_{P^{-1},P}]^{-1}\Gamma(P^{-1}) \tag{4}$$

and hence

$$\omega_{P,P^{-1}} = \omega_{P^{-1},P}. \tag{4a}$$

The character is no more a class function; but

[1] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, Inc., New York, 1959).
[2] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., 1962).
[3] W. Döring, Z. Naturforsch. **14**, 343 (1959).
[4] E. Brown, Phys. Rev. **133**, A1038 (1964).
[5] P. Rudra, J. Math. Phys. **6**, 1278 (1965) (following paper).

the absolute value of the character of elements belonging to the same class is equal, since

$$\text{Tr } \Gamma(P^{-1}QP) = (\omega_{P^{-1},P}/\omega_{P^{-1},Q}\omega_{P^{-1}Q,P}) \text{ Tr } \Gamma(Q). \quad (5)$$

3. FUNCTIONS FORMING THE REPRESENTATION: PROJECTION OPERATOR

We first define a set of linear operators O_R, forming a projective representation of the group. Wigner defined a similar sort of linear operators for vector representations. These new operators will be in one-to-one correspondence with the projective representation matrices. Corresponding to each element $R \in G$, we define O_R such that its action on a function $\psi(x)$ is

$$O_R\psi(x) = [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{1/2g}\psi(R^{-1}x). \quad (6)$$

The factors $\omega_{P,Q}$'s are one of the g roots of unity[2] and are expressed as exp $i\xi(P, Q)$, where $0 \leq \xi(P, Q) < 2\pi$. The expression $[\prod_{P\in G} \omega_{R,P}\omega_{P,R}]$ then takes the form exp $[i \sum_{P\in G} \{\xi(R, P) + \xi(P, R)\}]$ and the multiplicative term on the right-hand side is taken as

$$\exp [(i/2g) \sum_{P\in G} \{\xi(R, P) + \xi(P, R)\}].$$

Thus the apparent ambiguity of the $2g$ roots is avoided. This particular term will occur again in future and everywhere this meaning has to be assigned.

Successive operations of two operators O_R and O_S is given by

$$O_SO_R\psi(x) = O_S\phi(x) = [\prod_{P\in G} \omega_{S,P}\omega_{P,S}]^{1/2g}\phi(S^{-1}x)$$

$$= \left[\prod_{P\in G} \frac{\omega_{S,P}\omega_{P,S}\omega_{R,P}\omega_{P,R}}{\omega_{SR,P}\omega_{P,SR}}\right]^{1/2g} O_{SR}\psi(x)$$

$$= \omega_{S,R}O_{SR}\psi(x)$$

Thus,

$$O_SO_R = \omega_{S,R}O_{SR}. \quad (7)$$

The basis functions forming the μth irreducible representation will thus satisfy

$$O_R \,|\mu, i\rangle = \Gamma_\mu(R)_{ji} \,|\mu, j\rangle \quad (8)$$

where a repeated latin index will everywhere mean a summation over that dummy index. The $\Gamma_\mu(R)_{ji}$'s are the matrix elements of the μth irreducible representation, belonging to the particular factor system. To prove this we have

$$O_SO_R \,|\mu, i\rangle = \Gamma_\mu(R)_{ji}O_S \,|\mu, j\rangle$$

$$= [\Gamma_\mu(S)\Gamma_\mu(R)]_{ki} \,|\mu, k\rangle.$$

From (7), again

$$O_SO_R \,|\mu, i\rangle = \omega_{S,R}O_{SR} \,|\mu, i\rangle$$

$$= \omega_{S,R}\Gamma_\mu(SR)_{ki} \,|\mu, k\rangle.$$

Thus

$$\Gamma_\mu(S)\Gamma_\mu(R) = \omega_{S,R}\Gamma_\mu(SR),$$

and so the $\Gamma_\mu(R)$'s, defined in Eq. (8) form an irreducible representation of G, belonging to the factor system $\omega_{S,R}$.

The Wigner projection operators will take a modified form to tally with projective representations; we shall define it as

$$P_\mu^{(i)} = \sum_{R\in G} \frac{1}{\omega_{R^{-1},R}} \Gamma_\mu(R^{-1})_{ii}O_R. \quad (9)$$

Thus $P_\mu^{(i)}$ acting on an arbitrary function ψ will give the ith basis of the μth irreducible representation of the group, for arbitrary but fixed value of j. Then we can write

$$|\mu, i\rangle = \sum_{R\in G} \frac{1}{\omega_{R^{-1},R}} \Gamma_\mu(R^{-1})_{ji}O_R\psi. \quad (10)$$

Proof follows easily:

$$O_S \,|\mu, i\rangle$$

$$= \sum_{R\in G} \frac{1}{\omega_{R^{-1},R}} \Gamma_\mu(R^{-1})_{ji}O_SO_R\psi$$

$$= \sum_{R\in G} \frac{\omega_{S,R}}{\omega_{R^{-1},R}} \Gamma_\mu(R^{-1}S^{-1}S)_{ji}O_{SR}\psi$$

$$= \sum_{R\in G} \frac{\omega_{S,R}}{\omega_{R^{-1},R}\omega_{R^{-1}S^{-1},S}} \Gamma_\mu(S)_{ki}\Gamma_\mu(R^{-1}S^{-1})_{jk}O_{SR}\psi$$

$$= \Gamma_\mu(S)_{ki} \sum_{SR\in G} \frac{1}{\omega_{R^{-1}S^{-1},SR}} \Gamma_\mu(R^{-1}S^{-1})_{jk}O_{SR}\psi$$

$$= \Gamma_\mu(S)_{ki} \,|\mu, k\rangle.$$

This form of the projection operator will be useful in obtaining the matrices forming the irreducible representations. For that purpose, instead of the exact $\Gamma_\mu(R^{-1})_{ji}$'s, the corresponding characters have to be used when a combination of the basis functions would be obtained.

4. UNITARY REPRESENTATIONS

If the matrices forming the representation of the group are unitary, i.e., if $\Gamma(P)^{-1} = \Gamma(P)^\dagger$ for every $P \in G$, we say that the representation is unitary.

As in the case of vector representations, the projective representations for a finite group can always be chosen to be unitary.

Defining for arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\{\mathbf{x} \mid \mathbf{y}\} = \frac{1}{g}\sum_{R\in G}\langle\Gamma(R)\mathbf{x} \mid \Gamma(R)\mathbf{y}\rangle,$$

we have

$$\begin{aligned}\{\Gamma(S)\mathbf{x} \mid \Gamma(S)\mathbf{y}\} &= \frac{1}{g}\sum_{R\in G}\langle\Gamma(R)\Gamma(S)\mathbf{x} \mid \Gamma(R)\Gamma(S)\mathbf{y}\rangle \\ &= \frac{1}{g}\sum_{R\in G}\omega^*_{R,S}\omega_{R,S}\langle\Gamma(RS)\mathbf{x} \mid \Gamma(RS)\mathbf{y}\rangle \\ &= \frac{1}{g}\sum_{RS\in G}\langle\Gamma(RS)\mathbf{x} \mid \Gamma(RS)\mathbf{y}\rangle \\ &= \{\mathbf{x} \mid \mathbf{y}\}.\end{aligned}$$

The rest of the proof is similar to that for vector representations.[2]

5. SCHUR'S LEMMAS: ORTHOGONALITY AND OTHER RELATIONS

The Lemmas of Schur are exactly valid for projective representations, since the proof depends on the operation of the operators O_R's on a function, whose nature is the same as in the case of vector representations. Thus we have:

Lemma I. If Γ and Γ' be two irreducible projective representations of a group G, belonging to the same factor system, and which have different dimensions, then if there be a matrix A satisfying $\Gamma(R)A = A\Gamma'(R)$ for all $R \in G$, then $A = 0$.

Lemma Ia. If Γ and Γ' be irreducible representations of the same dimension of the group G, belonging to the same factor system and if there be a matrix A, satisfying $\Gamma(R)A = A\Gamma'(R)$ for all $R \in G$, then, either $\Gamma(R) \approx \Gamma'(R)$ or $A = 0$.

Lemma II. If Γ be an irreducible representation of a group G, and if there be a matrix A satisfying $\Gamma(R)A = A\Gamma(R)$ for all $R \in G$ then, $A =$ (constant) $\times I$, where I is a unit matrix having the same dimension as Γ.

To obtain the orthogonality relations, we construct a matrix,

$$A = \sum_{S\in G}\frac{1}{\omega_{S,S^{-1}}}\Gamma(S^{-1})X\Gamma(S),$$

where X is any arbitrary matrix. It readily follows from the basic relations of the factors $\omega_{P,Q}$,

$$\Gamma(R)A = A\Gamma(R) \quad \text{for all} \quad R \in G.$$

Thus from Schur's Lemma II, $A = \lambda I$.

We choose X, so that the only nonvanishing element is

$$X_{lm} = \delta_{ml} \quad \text{and} \quad \lambda = \lambda_{ml}.$$

Thus,

$$\sum_{S\in G}\frac{1}{\omega_{S,S^{-1}}}\Gamma(S^{-1})_{jl}\Gamma(S)_{mi} = \lambda_{ml}\delta_{ml}.$$

From this it easily follows that

$$\sum_{S\in G}\frac{1}{\omega_{S,S^{-1}}}\Gamma(S)_{im}\Gamma(S^{-1})_{lj} = \frac{g}{n}\delta_{ij}\delta_{ml}, \tag{11}$$

which takes for unitary matrices the usual form

$$\sum_{S\in G}\Gamma(S)_{im}\Gamma(S)^*_{jl} = \frac{g}{n}\delta_{ij}\delta_{ml}, \tag{11a}$$

where n is the dimension of the representation. We can generalize these formulas for two irreducible representations, belonging to the same factor system and obtain

$$\sum_{S\in G}\frac{1}{\omega_{S,S^{-1}}}\Gamma_\mu(S)_{im}\Gamma_\nu(S^{-1})_{lj} = \frac{g}{n_\mu}\delta_{\mu\nu}\delta_{ij}\delta_{ml} \tag{12}$$

and for unitary matrices,

$$\sum_{S\in G}\Gamma_\mu(S)_{im}\Gamma_\nu(S)^*_{jl} = \frac{g}{n_\mu}\delta_{\mu\nu}\delta_{ij}\delta_{ml}. \tag{12a}$$

Here n_μ is the dimension of the μth irreducible representation. The relation between the characters will be

$$\sum_{S\in G}\frac{1}{\omega_{S,S^{-1}}}\chi_\mu(S)\chi_\nu(S^{-1}) = g\delta_{\mu\nu}, \tag{13}$$

which has the well-known form for unitary matrices,

$$\sum_{S\in G}\chi_\mu(S)\chi_\nu(S)^* = g\delta_{\mu\nu}. \tag{13a}$$

Any reducible projective representation of the finite group G, belonging to a particular factor system, will again be expressible, according to Maschke's theorem, as a direct sum of irreducible representations, belonging to the same factor system;

$$\Gamma(R) = \sum_\mu a_\mu\Gamma_\mu(R), \tag{14}$$

so that the orthogonality relations will give, for the number of repetitions of the μth irreducible projective representation,

$$a_\mu = \frac{1}{g}\sum_{R\in G}\frac{1}{\omega_{R,R^{-1}}}\chi(R)\chi_\mu(R^{-1}), \tag{15}$$

which reduces, for unitary matrices, to

$$a_\mu = \frac{1}{g}\sum_{R\in G}\chi(R)\chi_\mu(R)^*. \tag{15a}$$

The regular representation, belonging to the factor system $\omega_{P,Q}$ is built up in a way similar to that for the vector representation, except for a factor

$[\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{\frac{1}{2g}}$ for the matrix $\Gamma_{reg}(R)$. Thus if $RS_i = S_{i_i}$ $(i = 1, \cdots, g)$,

$$\Gamma_{reg}(R)_{il} = [\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{1/2g}\delta_{li_i}. \tag{16}$$

Proceeding with this form of the regular representation, we obtain

$$\sum_{\mu} n_\mu^2 = g. \tag{17}$$

6. A METHOD FOR OBTAINING THE CHARACTER OF THE ELEMENTS OF A GROUP

A formula suitable for obtaining the characters of all the elements of the group, for all the irreducible projective representation, belonging to a particular factor system $\omega_{P,Q}$ would be here obtained. From a similar formula, applicable to vector representations, Bethe[6] obtained the character table for the irreducible representations of the point group.

We define for this purpose

$$\Gamma_i^m = \frac{g_i}{g}\sum_{R\in G}\frac{1}{\omega_{R^{-1},R}}\Gamma(R^{-1})\Gamma(A_i^m)\Gamma(R). \tag{18}$$

A_m^i is some mth element in the ith class of the group, having g_i elements. It can easily be shown that Γ_i^m commutes with the representation matrices of all the elements of the group. That is,

$$\Gamma(S)\Gamma_i^m = \Gamma_i^m\Gamma(S) \quad \text{for all} \quad S \in G.$$

Hence, it follows from Schur's Lemma II that $\Gamma_i^m = \lambda_i^m I$. Also, the product $\Gamma_i^m\Gamma_j^n$ will consist of integral sums of different Γ_i^p's,

$$\Gamma_i^m\Gamma_j^n = C_{ijl}^{mnp}\Gamma_l^p \tag{19}$$

with

$$C_{ijl}^{mnp} = C_{jil}^{nmp}.$$

Since

$$\text{Tr}\ \Gamma_i^m = g_i\ \text{Tr}\ \Gamma(A_i^m) \equiv g_i\chi_i^m \quad \text{[no summation]},$$

we obtain

$$\lambda_i^m = g_i\chi_i^m/\chi_1,$$

where χ_1 is the character of the identity element, which is equal to the dimensionality of the irreducible representation. From Eq. (19), the following working formula would be obtained:

$$g_ig_j\chi_i^m\chi_j^n = \chi_1 C_{ijl}^{mnp}g_l\chi_l^p. \tag{20}$$

In the left-hand-side, there is no summation over i, j. From Eq. (20) and the orthogonality relation, the character table for all the inequivalent irreducible projective representations of the group, belonging to the particular factor system can be obtained. It is to be noted that Γ_i^m is not, in general, independent of m, as it is for vector representation. If any particular Γ_k^q vanishes, the character of the corresponding group element $\chi_k^q = 0$ for all the inequivalent irreducible projective representation, belonging to the factor system.

7. KRONECKER (INNER) DIRECT PRODUCT: CLEBSCH-GORDAN COEFFICIENTS

The Kronecker (inner) direct-product representation is defined in the case of projective representation by

$$\Gamma_{\mu\otimes\nu}(R)_{ik,jl} \equiv (\Gamma_\mu(R)\otimes\Gamma_\nu(R))_{ik,jl}$$
$$= [\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{-1/2g}\Gamma_\mu(R)_{ij}\Gamma_\nu(R)_{kl}. \tag{21}$$

That $\Gamma_{\mu\otimes\nu}(R)$ is a (in general) reducible representation of the group, belonging to the particular factor system, follows from the exact calculation

$$\Gamma_{\mu\otimes\nu}(RS)_{ik,jl} = (\omega_{R,S})^{-1}[\Gamma_{\mu\otimes\nu}(R)\Gamma_{\mu\otimes\nu}(S)]_{ik,jl}$$

or

$$\Gamma_{\mu\otimes\nu}(R)\Gamma_{\mu\otimes\nu}(S) = \omega_{R,S}\Gamma_{\mu\otimes\nu}(RS).$$

The character for this product representation is then given by

$$\chi_{\mu\otimes\nu}(R) = [\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{-1/2g}\chi_\mu(R)\chi_\nu(R). \tag{22}$$

As in the case of vector representation, here also, the product of the jth basis of the μth irreducible projective representation and the lth basis of the νth representation will also form the (j, l)th basis for the product representation $\Gamma_{\mu\otimes\nu}$.

Proof: Let

$$|\mu, j\rangle \equiv \psi_\mu^j(x) \quad \text{and} \quad |\nu, l\rangle \equiv \phi_\nu^l(x)$$

and

$$|\mu, j; \nu, l\rangle \equiv \psi_\mu^j(x)\phi_\nu^l(x).$$

Then using Eqs. (6), (8), and (21), we have

$$O_R\ |\mu, j; \nu, l\rangle = O_R[\psi_\mu^j(x)\phi_\nu^l(x)]$$
$$= [\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{1/2g}\psi_\mu^j(R^{-1}x)\phi_\nu^l(R^{-1}x)$$
$$= [\prod_{P\in G}\omega_{R,P}\omega_{P,R}]^{-1/2g}[O_R\psi_\mu^j(x)][O_R\phi_\nu^l(x)]$$
$$= \Gamma_{\mu\otimes\nu}(R)_{ik,jl}\ |\mu, i; \nu, k\rangle.$$

Since $\Gamma_{\mu\otimes\nu}$ is, in general, a reducible representation, this can be expanded as a direct sum of the irreducible components, in a Clebsch–Gordan series:

$$\Gamma_{\mu\otimes\nu}(R) = \sum_\sigma(\mu\nu\sigma)\Gamma_\sigma(R). \tag{23}$$

[6] Hans A. Bethe, Ann. Physik 3, 133 (1929). English Translation: "Splitting of Terms in Crystals," Consultants Bureau, New York.

Using Eqs. (13), (22), and (23), we obtain

$$(\mu\nu\sigma) = \frac{1}{g}\sum_{R\in G}\frac{1}{\omega_{R,R^{-1}}}\chi_{\mu\otimes\nu}(R)\chi_\sigma(R^{-1})$$

$$= \frac{1}{g}\sum_{R\in G}\frac{1}{\omega_{R,R^{-1}}}\Big[\prod_{P\in G}\omega_{R,P}\omega_{P,R}\Big]^{-1/2g}\chi_\mu(R)\chi_\nu(R)\chi_\sigma(R^{-1}). \tag{24}$$

For unitary matrices, this equation takes the form

$$(\mu\nu\sigma) = \frac{1}{g}\sum_{R\in G}\Big[\prod_{P\in G}\omega_{R,P}\omega_{P,R}\Big]^{-1/2g}\chi_\mu(R)\chi_\nu(R)\chi_\sigma(R)^*. \tag{24a}$$

Since $|\mu, j; \nu, l\rangle = \psi_\mu^j\phi_\nu^l$ forms the basis of the (in general) reducible representation $\Gamma_{\mu\otimes\nu}$, expressible as a direct sum of the different irreducible representations, the sth basis function $|\sigma, \tau_\sigma, s\rangle$, belonging to the τ_σth repetition of the σth irreducible representation in the Clebsch–Gordan series expansion of $\Gamma_{\mu\otimes\nu}$, can be expressed as the linear combination of different $|\mu, j; \nu, l\rangle$'s with suitable coefficients;

$$|\sigma, \tau_\sigma, s\rangle = \langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle\,|\mu, j; \nu, l\rangle,$$

$$\text{where}\quad j = 1, \cdots, n_\mu \quad\text{and}\quad l = 1, \cdots, n_\nu. \tag{25}$$

$\langle\mu, j; \nu, l|\sigma, \tau_\sigma, s\rangle$'s are called the Clebsch–Gordan coefficients. The reverse transformations are also valid.

$$|\mu, j; \nu, l\rangle = \sum_{\sigma,\tau_\sigma}\langle\sigma, \tau_\sigma, s \mid \mu, j; \nu, l\rangle\,|\sigma, \tau_\sigma, s\rangle. \tag{26}$$

The orthogonality conditions of the Clebsch–Gordan coefficients remain as usual.[2]

$$\langle\sigma', \tau_{\sigma'}, s' \mid \mu, j; \nu, l\rangle\langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle = \delta_{\sigma\sigma'}\delta_{\tau_\sigma\tau_{\sigma'}}\delta_{ss'},$$

$$\sum_{\sigma,\tau_\sigma}\langle\mu, j'; \nu, l' \mid \sigma, \tau_\sigma, s\rangle\langle\sigma, \tau_\sigma, s \mid \mu, j; \nu, l\rangle = \delta_{jj'}\delta_{ll'}. \tag{27}$$

For unitary representations, these reduce to

$$\langle\sigma, \tau_\sigma, s \mid \mu, j; \nu, l\rangle = \langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle^*,$$

$$\langle\mu, j; \nu, l \mid \sigma', \tau_{\sigma'}, s'\rangle^*\langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle = \delta_{\sigma\sigma'}\delta_{\tau_\sigma\tau_{\sigma'}}\delta_{ss'}, \tag{27a}$$

$$\sum_{\sigma,\tau_\sigma}\langle\mu, j'; \nu, l' \mid \sigma, \tau_\sigma, s\rangle^*\langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle = \delta_{jj'}\delta_{ll'}.$$

The Wigner formula, connecting the Clebsch–Gordan coefficient and the matrix elements of the irreducible representations would have to be modified in the same way as the Kronecker (inner) direct product. Algebraic calculations give, finally,

$$\langle\mu, j; \nu, l \mid \sigma, \tau_\sigma, s\rangle\Gamma_{\mu\otimes\nu}(R)_{ik,jl} = \Gamma_{\sigma,\tau_\sigma}(R)_{s's}\langle\mu, i; \nu, k \mid \sigma, \tau_\sigma, s'\rangle. \tag{28}$$

Using the suitable orthogonality relations, we arrive at the form

$$\sum_{R\in G}\frac{1}{\omega_{R,R^{-1}}}\Gamma_{\mu\otimes\nu}(R)_{m_1n_1,m_2n_2}\Gamma_{\sigma,\tau_\sigma}(R^{-1})_{s_2s_1} = \frac{g}{n_\sigma}\sum_{\tau_\sigma}\langle\mu, m_1; \nu, n_1 \mid \sigma, \tau_\sigma, s_1\rangle \times\langle\sigma, \tau_\sigma, s_2 \mid \mu, m_2; \nu, n_2\rangle \tag{29}$$

and for unitary representations,

$$\sum_{R\in G}\Gamma_{\mu\otimes\nu}(R)_{m_1n_1,m_2n_2}\Gamma_{\sigma,\tau_\sigma}(R)^*_{s_1s_2} = \frac{g}{n_\sigma}\sum_{\tau_\sigma}\langle\mu, m_1; \nu, n_1 \mid \sigma, \tau_\sigma, s_1\rangle \times\langle\mu, m_2; \nu, n_2 \mid \sigma, \tau_\sigma, s_2\rangle^*. \tag{29a}$$

For simply reducible groups, where the Kronecker (inner) direct product of two irreducible representations of a group contains an irreducible representation only once, we have the simplified relation

$$\sum_{R\in G}\frac{1}{\omega_{R,R^{-1}}}\Gamma_{\mu\otimes\nu}(R)_{m_1n_1,m_2n_2}\Gamma_\sigma(R^{-1})_{s_2s_1} = \frac{g}{n_\sigma}\langle\mu, m_1; \nu, n_1 \mid \sigma, s_1\rangle\langle\sigma, s_2 \mid \mu, m_2; \nu, n_2\rangle \tag{30}$$

and for unitary representation,

$$\sum_{R\in G}\Gamma_{\mu\otimes\nu}(R)_{m_1n_1,m_2n_2}\Gamma_\sigma(R)^*_{s_1s_2} = (g/n_\sigma)\langle\mu, m_1; \nu, n_1 \mid \sigma, s_1\rangle\langle\mu, m_2; \nu, n_2 \mid \sigma, s_2\rangle^*. \tag{30a}$$

In actual calculations, we have to replace everywhere

$$\Gamma_{\mu\otimes\nu}(R)_{m_1n_1,m_2n_2}$$

by

$$\Big[\prod_{P\in G}\omega_{R,P}\omega_{P,R}\Big]^{-1/2g}\Gamma_\mu(R)_{m_1m_2}\Gamma_\nu(R)_{n_1n_2}.$$

8. CONCLUSION

Here we have presented the algebra of the projective representations of finite groups. In the next paper[5] we shall apply it in obtaining the irreducible representations of space groups for special points of the Brillouin zone.

ACKNOWLEDGMENT

I shall like to thank Professor D. K. Ray for his interest in the work.

On Irreducible Representations of Space Groups

P. RUDRA
Saha Institute of Nuclear Physics, Calcutta, India
(Received 9 November 1964)

A logical extension has been made of Seitz–Koster's method of space-group representations to the points on the surface of the Brillouin zone, using projective representations. This method has been applied in the case of the space group D_3^4.

1. INTRODUCTION

EVER since Bouckaert, Smoluchowski, and Wigner's treatment[1] of solutions of the Schrödinger equation in crystals, attention has been drawn to space groups. Seitz[2] and Koster[3] have developed the theory of representations for space groups from the irreducible representations of the invariant subgroup of pure translations. The theory has been completed for symmorphic space groups and for all the points except for those on the surface of the Brillouin zone in case of nonsymmorphic space groups. Zak[4] has given a method to obtain the characters of nonsymmorphic space groups, based on the general theory[5] of obtaining the representation of a total group, from the representations subduced by the irreducible representation of its subgroups of index two or three. Döring and Zehler[6] have obtained the character table for diamond in a similar method. Here we shall extend Seitz[2] and Koster's[3] method to the points on the surface of the Brillouin zone using the method of projective representation.[7,8] Döring[9] has previously used the idea of projective representation, to obtain all possible projective representations for crystallographic point groups.

We first give a summary of Seitz and Koster's method and then give the extension of the method using projective representations. Finally we give this method for the space group D_3^4.

2. GENERAL THEORY OF SPACE-GROUP REPRESENTATIONS[2,3,10]

We define the space-group operations in a slightly modified form than Seitz and Koster. The operation of the space-group elements $\{\mathbf{c}|\gamma\}$ on any vector $\mathbf{x}$ is such that

$$\{\mathbf{c} \mid \gamma\}\mathbf{x} = \mathbf{c} + \gamma\mathbf{x}, \tag{1}$$

where $\mathbf{c}$ is a translation and γ is a point-group operation. The combination laws and the inverse operation are given by

$$\{\mathbf{c}_2 \mid \gamma_2\}\{\mathbf{c}_1 \mid \gamma_1\} = \{\mathbf{c}_2 + \gamma_2\mathbf{c}_1 \mid \gamma_2\gamma_1\}$$

and

$$\{\mathbf{c} \mid \gamma\}^{-1} = \{-\gamma^{-1}\mathbf{c} \mid \gamma^{-1}\}. \tag{2}$$

The pure translational elements $\{\mathbf{R}_n|E\}$ form an invariant subgroup of the space group except when there is a magnetic field present, with which we are not concerned here.[11] For each point of the first Brillouin zone, we have an inequivalent irreducible representation, which is one dimensional:

$$\Gamma_{\mathbf{k}}(\mathbf{R}_n) = \exp i\mathbf{k}\cdot\mathbf{R}_n. \tag{3}$$

For the total space groups these will not, of course, form the irreducible representations. But the matrices corresponding to the elements of this invariant subgroup can be brought to the diagonal form

$$\Gamma_{\mathbf{k}}\{\mathbf{R}_n \mid E\} = \begin{bmatrix} [\exp i\mathbf{k}_1\cdot\mathbf{R}_n]I & 0 & \cdots & 0 \\ 0 & [\exp i\mathbf{k}_2\cdot\mathbf{R}_n]I & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & [\exp i\mathbf{k}_p\cdot\mathbf{R}_n]I \end{bmatrix}, \tag{4}$$

[1] L. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).
[2] F. Seitz, Ann. Math. 37, 17 (1936).
[3] G. F. Koster, "Notes on Group Theory," Solid State and Molecular Theory Group, Massachusetts Institute of Technology, Technical Report No. 8 (1956).
[4] J. Zak, J. Math. Phys. 1, 165 (1960).
[5] H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963).
[6] W. Döring and V. Zehler, Ann. Physik 13, 214 (1953).
[7] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).
[8] P. Rudra, J. Math. Phys. 6, 1273 (1965) (preceding paper).
[9] W. Döring, Z. Naturforshung 14, 343 (1959)
[10] A good review is given by A. V. Sokolov and V. P. Shirokovskji, Usp. Fiz. Nauk 71, 485 (1960) [English transl.: Soviet Phys.— Usp. 3, 551 (1961)].
[11] J. Zak, Phys. Rev. 134, A 1602, A 1607 (1964).

where the $\mathbf{k}_i$'s satisfy the relation $\mathbf{k}_i = \alpha_i\mathbf{k}$, where α_i's are point-group portions of some specified space-group elements and $\alpha_1 = E$, so that $\mathbf{k}_1 = \mathbf{k}$. The unit matrices with each exponential factor has suitable dimension.

Among all the space-group elements of G, there are some elements $\{\mathbf{b}|\beta\}$ which has the property for the particular $\mathbf{k}$ vector

$$\beta\mathbf{k} = \mathbf{k} + \mathbf{K}_q, \tag{5}$$

so that

$$\exp i\beta\mathbf{k}\cdot\mathbf{R}_n = \exp i\mathbf{k}\cdot\mathbf{R}_n, \tag{5a}$$

where $\mathbf{K}_q$ is any reciprocal lattice vector. All such $\{\mathbf{b}|\beta\}$'s form a subgroup, K, of the total space group G. Expanding in cosets, we get

$$G = K + \{\mathbf{a}_2 \mid \alpha_2\}K + \cdots + \{\mathbf{a}_m \mid \alpha_m\}K. \tag{6}$$

The α_i's introduced before are these α's, so that $p = m$. An n-dimensional representation thus breaks up in the orbits of $\alpha_i\mathbf{k}$'s each having d dimensions, so that, $n = md$. These $\alpha_i\mathbf{k}$'s ($i = 1, \cdots, m$) are said to be the stars of the particular $\mathbf{k}$ vector. d is the dimension of the irreducible representations of the point-group portions of K, denoted by $G_0(\mathbf{k})$.

The total matrix for $\Gamma\{\mathbf{b}|\beta\}$ can be so transformed that the only nonvanishing block in its first column, lies in the first row, and this is denoted by $\Gamma_{\mathbf{k}}\{\mathbf{b}|\beta\}_{11}$ for different $\mathbf{k}$ vectors. This $\Gamma_{\mathbf{k}}\{\mathbf{b}|\beta\}_{11}$'s are sufficient to give the matrices of all the elements, $\Gamma_{\mathbf{k}}\{\mathbf{c}|\gamma\}$ of G.

The group element $\{\mathbf{a}_j|\alpha_j\}$ has a nonvanishing block in the first column, only in the jth row, and this can be put equal to $\Gamma_{\mathbf{k}}\{\mathbf{0}|E\}_{11}$.

For any other space-group element $\{\mathbf{c}|\gamma\}$, the nonvanishing block in the lth column is the mth row and this is

$$\Gamma_{\mathbf{k},\mu}\{\mathbf{c} \mid \gamma\}_{ml} = \Gamma_{\mathbf{k},\mu}\{\mathbf{b} \mid \beta\}_{11}, \tag{7}$$

where

$$\{\mathbf{c} \mid \gamma\}\{\mathbf{a}_l \mid \alpha_l\} = \{\mathbf{a}_m \mid \alpha_m\}\{\mathbf{b} \mid \beta\}. \tag{7a}$$

μ denotes different irreducible representations of $G_0(\mathbf{k})$.

For symmorphic space groups

$$\Gamma_{\mathbf{k},\mu}\{\mathbf{b} \mid \beta\}_{11} = [\exp i\mathbf{k}\cdot\mathbf{b}]\Gamma_\mu(\beta), \tag{8}$$

where $\Gamma_\mu(\beta)$'s are all the nonequivalent irreducible vector representations of $G_0(\mathbf{k})$. For nonsymmorphic space groups also, the form of $\Gamma_{\mathbf{k},\mu}\{\mathbf{b}|\beta\}_{11}$, given in Eq. (8), with all the vector representations of $G_0(\mathbf{k})$, form the irreducible representations of G, for all the points inside the Brillouin zone, but not so, for points on the surface of the Brillouin zone.

3. APPLICATION OF PROJECTIVE REPRESENTATIONS

For treating the irreducible representations of G, corresponding to $\mathbf{k}$ vectors on the surface of the Brillouin zone, we shall show that if we set $\Gamma_\mu(\beta)$ as a projective representation of $G_0(\mathbf{k})$ belonging to the factor system, given below, then $\Gamma_{\mathbf{k},\mu}\{\mathbf{b}|\beta\}_{11}$ expressed in the form given in Eq. (8), form a vector representation of K group.

In order that $\Gamma_{\mathbf{k},\mu}\{\mathbf{b}|\beta\}_{11}$ form an irreducible vector representation of K, the following condition must be satisfied:

$$\Gamma_{\mathbf{k},\mu}\{\mathbf{b}_i \mid \beta_i\}_{11}\Gamma_{\mathbf{k},\mu}\{\mathbf{b}_j \mid \beta_j\}_{11} = \Gamma_{\mathbf{k},\mu}\{\mathbf{b}_i + \beta_i\mathbf{b}_j \mid \beta_i\beta_j\}_{11}. \tag{9}$$

If we write $\Gamma_{\mathbf{k},\mu}\{\mathbf{b}|\beta\}_{11} = [\exp i\mathbf{k}\cdot\mathbf{b}]\ \Gamma_\mu(\beta)$, then $\Gamma_\mu(\beta)$'s must satisfy

$$\Gamma_\mu(\beta_i)\Gamma_\mu(\beta_j) = \omega_{\beta_i,\beta_j}\Gamma_\mu(\beta_i\beta_j), \tag{10}$$

where

$$\omega_{\beta_i,\beta_j} = \exp i(\beta_i^{-1}\mathbf{k} - \mathbf{k})\cdot\mathbf{b}_j. \tag{11}$$

Now since

$$\omega_{\beta_i,\beta_j}\omega_{\beta_i\beta_j,\beta_l} = \omega_{\beta_i,\beta_j\beta_l}\omega_{\beta_j,\beta_l}, \tag{12}$$

the $\Gamma_\mu(\beta)$'s form an irreducible projective representation[7,8] of $G_0(\mathbf{k})$ belonging to the factor system, given in (11). Actually this is the general case for every point of the Brillouin zone. Since, however, $\beta_i^{-1}\mathbf{k} - \mathbf{k} = \mathbf{0}$ for all points inside the Brillouin zone, for both symmorphic and nonsymmorphic G, $\omega_{\beta_i,\ \beta_j}$'s become 1 for all β_i, $\beta_j \in G_0(\mathbf{k})$ and $\Gamma_\mu(\beta)$'s become the vector representations. For symmorphic groups, since all $\mathbf{b}$'s are *lattice translations* $\mathbf{R}_n$, even the surface points of the Brillouin zone are characterized by vector representation. For only those points of the Brillouin zone surface, whose K group contains a nonprimitive translation, the projective representations are to be used, the factor system, being given by Eq. (11). The factor systems being known, the characters are obtained easily.[8] To obtain the matrices of the irreducible projective representations, we operate on any reducible basis the corresponding projection operator,[8] replacing the matrix element by the character. This will give a combination of the basis functions, from which the basis functions separately, and hence the matrix elements themselves, can be obtained.

4. IRREDUCIBLE REPRESENTATIONS OF THE SPACE GROUP D_3^4

We now apply the method to obtain the inequivalent irreducible representations of the nonsym-

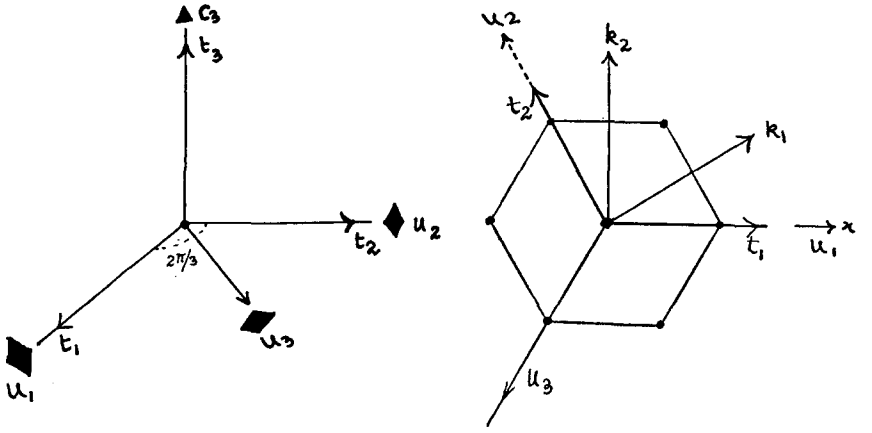


FIG. 1. Lattice of D_3^4 with the position of the rotation axes, and the projection on the (t_1, t_2) plane.

morphic space group D_3^4. We give only those representations corresponding to points on the Brillouin zone surface, for which the considerations of projective representations have to be brought in.

The description of the space group is[12] $D_3^4(C\,3_1\,2\,1)$. The lattice type is Γ_h. The basic translation vectors are shown in Fig. 1. They are

$$\mathbf{t}_1 = a_0\begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \mathbf{t}_2 = a_0\begin{bmatrix}-\frac{1}{2}\\ \sqrt{3}/2\\0\end{bmatrix} \qquad \mathbf{t}_3 = c_0\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

The group operations are

$$\{\mathbf{R}_n \mid E\},\quad \{\mathbf{R}_n + \tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\},\quad \{\mathbf{R}_n + \tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\},$$
$$\{\mathbf{R}_n + \tfrac{1}{3}\mathbf{t}_3 \mid U_1\},\quad \{\mathbf{R}_n + \tfrac{2}{3}\mathbf{t}_3 \mid U_2\},\quad \{\mathbf{R}_n \mid U_3\}.$$

The unit vectors in the reciprocal lattice space are

$$\mathbf{k}_1 = \frac{4\pi}{\sqrt{3}a_0}\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right),\quad \mathbf{k}_2 = \frac{4\pi}{\sqrt{3}a_0}(0, 1, 0),$$
$$\mathbf{k}_3 = 2\pi c_0^{-1}(0, 0, 1).$$

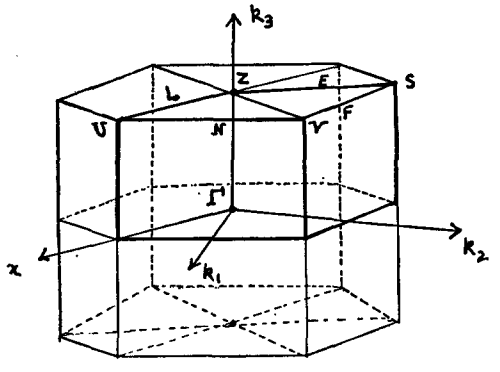


FIG. 2. Brillouin zone for D_3^4.

[12] G. Ya. Lyubarskii, *The Application of Group Theory in Physics* (Pergamon Press, Inc., New York, 1960) (English translation).

The Brillouin zone and the particular $\mathbf{k}$ points of interest are shown in Fig. 2. Of these the points Z, U, V, and S have the same type of irreducible representation. The lines L and F and the lines E and N have similar types.

Point Z. Here, $K = G$. Thus

$$Z_\mu\{\mathbf{b} \mid \beta\} = \Gamma_{\mathbf{k},\mu}\{\mathbf{b} \mid \beta\}_{11} = [\exp i\mathbf{k}\cdot\mathbf{b}]\Gamma_\mu(\beta)$$

when the $\Gamma_\mu(\beta)$'s belong to the factor system tabulated in Table I. The character table for this projective representation is given in Table II.

Line L. Here

$$K \ni \{\mathbf{R}_n \mid E\},\quad \{\mathbf{R}_n + \tfrac{1}{3}\mathbf{t}_3 \mid U_1\}$$

and so,

$$G = \{\mathbf{0} \mid E\}K + \{\tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\}K + \{\tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\}K.$$

TABLE I. Factor system ω_{β_i,β_j} for the $G_0(\mathbf{k})$ at the Z point.

	β_j					
β_i	E	C_3^1	C_3^2	U_1	U_2	U_3
E	1	1	1	1	1	1
C_3^1	1	1	1	1	1	1
C_3^2	1	1	1	1	1	1
U_1	1	$e^{2\pi i/3}$	$e^{4\pi i/3}$	$e^{4\pi i/3}$	$e^{2\pi i/3}$	1
U_2	1	$e^{2\pi i/3}$	$e^{4\pi i/3}$	$e^{4\pi i/3}$	$e^{2\pi i/3}$	1
U_3	1	$e^{2\pi i/3}$	$e^{4\pi i/3}$	$e^{4\pi i/3}$	$e^{2\pi i/3}$	1

TABLE II. Character table for the projective representations of $G_0(\mathbf{k})$ at the Z point.

	β					
$\chi_\mu(\beta)$	E	C_3^1	C_3^2	U_1	U_2	U_3
χ_1	1	$e^{4\pi i/3}$	$e^{2\pi i/3}$	$e^{2\pi i/3}$	$e^{4\pi i/3}$	1
χ_2	1	$e^{4\pi i/3}$	$e^{2\pi i/3}$	$-e^{2\pi i/3}$	$-e^{4\pi i/3}$	-1
χ_3	2	$-e^{4\pi i/3}$	$-e^{2\pi i/3}$	0	0	0

TABLE III. Factor system $\omega\beta_i,\ \beta_j$ for the $G_0(\mathbf{k})$ on the line L.

	β_j	
β_i	E	U_1
E	1	1
U_1	1	$e^{4\pi i/3}$

TABLE V. Factor system $\omega\beta_i,\ \beta_j$ for the $G_0(\mathbf{k})$ on the line E.

	β_j	
β_i	E	U_2
E	1	1
U_2	1	$e^{2\pi i/3}$

TABLE IV. Character table for the projective representations of $G_0(\mathbf{k})$ on the line L.

	β	
$\chi_\mu(\beta)$	E	U_1
χ_1	1	$e^{2\pi i/3}$
χ_2	1	$-e^{2\pi i/3}$

TABLE VI. Character table for the projective representations of $G_0(\mathbf{k})$ on the line E.

	β	
$\chi_\mu(\beta)$	E	U_2
χ_1	1	$e^{\pi i/3}$
χ_2	1	$-e^{\pi i/3}$

The factor system for $G_0(\mathbf{k})$ is given in Table III and the corresponding characters (which coincide with the representations in this one-dimensional case) is given in Table IV. The matrices of all the group elements can be obtained from the following ones:

$$L_\mu\{\mathbf{R}_n \mid E\} = \begin{bmatrix} [\exp i\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) & 0 & 0 \\ 0 & [\exp iC_3^1\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) & 0 \\ 0 & 0 & [\exp iC_3^2\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) \end{bmatrix},$$

$$L_\mu\{\tfrac{1}{3}\mathbf{t}_3 \mid U_1\} = \begin{bmatrix} [\exp i\mathbf{k}\cdot\tfrac{1}{3}\mathbf{t}_3]\Gamma_\mu(U_1) & 0 & 0 \\ 0 & 0 & [\exp(-iC_3^1\mathbf{k}\cdot\tfrac{2}{3}\mathbf{t}_3)]\Gamma_\mu(U_1) \\ 0 & [\exp(-iC_3^2\mathbf{k}\cdot\tfrac{2}{3}\mathbf{t}_3)]\Gamma_\mu(U_1) & 0 \end{bmatrix},$$

$$L_\mu\{\tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\} = \begin{bmatrix} 0 & 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \\ 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \end{bmatrix},$$

$$L_\mu\{\tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\} = \begin{bmatrix} 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \\ 0 & 0 & \Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \end{bmatrix}.$$

Line E. Here

$$K \ni \{\mathbf{R}_n \mid E\}, \qquad \{\mathbf{R}_n + \tfrac{2}{3}\mathbf{t}_3 \mid U_2\}.$$

Thus

$$G = \{\mathbf{0} \mid E\}K + \{\tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\}K + \{\tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\}K.$$

The factor system for $G_0(\mathbf{k})$ and the corresponding characters (which again coincide with the representations) are given in Tables V and VI. The matrices are

$$E_\mu\{\mathbf{R}_n \mid E\} = \begin{bmatrix} [\exp i\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) & 0 & 0 \\ 0 & [\exp iC_3^1\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) & 0 \\ 0 & 0 & [\exp iC_3^2\mathbf{k}\cdot\mathbf{R}_n]\Gamma_\mu(E) \end{bmatrix},$$

$$E_\mu\{\tfrac{2}{3}\mathbf{t}_3 \mid U_2\} = \begin{bmatrix} \Gamma_\mu(U_2) & 0 & 0 \\ 0 & 0 & [\exp(-i\mathbf{k}\cdot\tfrac{1}{3}\mathbf{t}_3)]\Gamma_\mu(U_2) \\ 0 & [\exp(-i\mathbf{k}\cdot\tfrac{1}{3}\mathbf{t}_3)]\Gamma_\mu(U_2) & 0 \end{bmatrix},$$

$$E_\mu\{\tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\} = \begin{bmatrix} 0 & 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \\ 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \end{bmatrix},$$

$$E_\mu\{\tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\} = \begin{bmatrix} 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \\ 0 & 0 & \Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \end{bmatrix}.$$

ACKNOWLEDGMENT

I wish to thank Professor D. K. Ray for his constant interest in the work and many helpful discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Separation of the Interaction Potential into Two Parts in Treating Many-Body Systems. I. General Theory and Applications to Simple Fluids with Short-Range and Long-Range Forces*

J. L. LEBOWITZ, G. STELL, AND S. BAER†
Belfer Graduate School of Science, Yeshiva University, New York, New York
(21 December 1964)

Systematic methods are developed for investigating the correlation functions and thermodynamic properties of a classical system of particles interacting via a pair potential $v(r) = q(r) + w(r)$. The method is then applied to the case in which $w(r)$ is a "Kac potential" $w(r, \gamma) = \gamma^\nu\varphi(\gamma r)$ (ν the dimensionality of the space) whose range γ^{-1} is very long compared to the range of $q(r)$. Our work is related closely to the work of Kac, Uhlenbeck, and Hemmer. The main new feature of our method is the separation of the correlations, e. g., the two-particle Ursell function $\mathfrak{F}(r)$, into a short-range part $\mathfrak{F}^s(r, \gamma)$ and a long-range part $\mathfrak{F}^L(y, \gamma)$, $y \equiv \gamma r$; r the distance between the particles. The two parts of $\mathfrak{F}$ are defined in terms of their representation by graphs with density (or fugacity) vertices and K- and Φ-bonds, $K(r) = e^{-\beta q} - 1$, $\Phi = -\beta w$. A resummation of these graphs then yields a simple graphical representation for the long-range part of the correlation functions in terms of graphs with Φ-bonds and "hypervertices" made up of the short-range part of the correlations. This representation is then used in this paper to make separate expansions of $\mathfrak{F}^s(r, \gamma)$ and $\mathfrak{F}^L(y, \gamma)$ and through them of the thermodynamic parameters in powers of γ. Explicit calculations of the Helmboltz free energy is carried out to a higher order in γ than done previously by Hemmer and it is shown how to carry out the calculation, in principle, to any order. The general method is further applied (in separate articles) to lattice gases, plasmas, and to the special problem of critical phenomena.

I. INTRODUCTION

THIS paper is the first in a series dealing with classical equilibrium systems. The system discussed in this paper, where the general formalism is developed, is a fluid of point particles interacting via pair potentials $v(r)$. [Later papers will deal with lattice gases[1] (Ising spin systems), and plasmas.[2]]

The problem, as usual, is to obtain the thermodynamic properties and low-order correlation functions of the fluid from the properties of the interatomic potential $v(r)$ believed to consist of a very

* Supported by the U. S. Air Force Office of Scientific Research, the National Science Foundation, and the U. S. Atomic Energy Commission.

† On leave of absence from the Hebrew University, Jerusalem.

[1] S. Baer, J. L. Lebowitz, G. Stell, and W. Theumann (to be published in J. Math. Phys.).

[2] J. L. Lebowitz and G. Stell, Bull. Am. Phys. Soc. 9, 105 (1964); (article to be published in J. Math. Phys.).

$$E_\mu\{\tfrac{2}{3}\mathbf{t}_3 \mid U_2\} = \begin{bmatrix} \Gamma_\mu(U_2) & 0 & 0 \\ 0 & 0 & [\exp(-i\mathbf{k}\cdot\tfrac{1}{3}\mathbf{t}_3)]\Gamma_\mu(U_2) \\ 0 & [\exp(-i\mathbf{k}\cdot\tfrac{1}{3}\mathbf{t}_3)]\Gamma_\mu(U_2) & 0 \end{bmatrix},$$

$$E_\mu\{\tfrac{2}{3}\mathbf{t}_3 \mid C_3^1\} = \begin{bmatrix} 0 & 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \\ 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \end{bmatrix},$$

$$E_\mu\{\tfrac{1}{3}\mathbf{t}_3 \mid C_3^2\} = \begin{bmatrix} 0 & [\exp i\mathbf{k}\cdot\mathbf{t}_3]\Gamma_\mu(E) & 0 \\ 0 & 0 & \Gamma_\mu(E) \\ \Gamma_\mu(E) & 0 & 0 \end{bmatrix}.$$

ACKNOWLEDGMENT

I wish to thank Professor D. K. Ray for his constant interest in the work and many helpful discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Separation of the Interaction Potential into Two Parts in Treating Many-Body Systems. I. General Theory and Applications to Simple Fluids with Short-Range and Long-Range Forces*

J. L. LEBOWITZ, G. STELL, AND S. BAER†
Belfer Graduate School of Science, Yeshiva University, New York, New York
(21 December 1964)

Systematic methods are developed for investigating the correlation functions and thermodynamic properties of a classical system of particles interacting via a pair potential $v(r) = q(r) + w(r)$. The method is then applied to the case in which $w(r)$ is a "Kac potential" $w(r, \gamma) = \gamma^\nu\varphi(\gamma r)$ (ν the dimensionality of the space) whose range γ^{-1} is very long compared to the range of $q(r)$. Our work is related closely to the work of Kac, Uhlenbeck, and Hemmer. The main new feature of our method is the separation of the correlations, e. g., the two-particle Ursell function $\mathfrak{F}(r)$, into a short-range part $\mathfrak{F}^s(r, \gamma)$ and a long-range part $\mathfrak{F}^L(y, \gamma)$, $y \equiv \gamma r$; r the distance between the particles. The two parts of $\mathfrak{F}$ are defined in terms of their representation by graphs with density (or fugacity) vertices and K- and Φ-bonds, $K(r) = e^{-\beta q} - 1$, $\Phi = -\beta w$. A resummation of these graphs then yields a simple graphical representation for the long-range part of the correlation functions in terms of graphs with Φ-bonds and "hypervertices" made up of the short-range part of the correlations. This representation is then used in this paper to make separate expansions of $\mathfrak{F}^s(r, \gamma)$ and $\mathfrak{F}^L(y, \gamma)$ and through them of the thermodynamic parameters in powers of γ. Explicit calculations of the Helmboltz free energy is carried out to a higher order in γ than done previously by Hemmer and it is shown how to carry out the calculation, in principle, to any order. The general method is further applied (in separate articles) to lattice gases, plasmas, and to the special problem of critical phenomena.

I. INTRODUCTION

THIS paper is the first in a series dealing with classical equilibrium systems. The system discussed in this paper, where the general formalism is developed, is a fluid of point particles interacting via pair potentials $v(r)$. [Later papers will deal with lattice gases[1] (Ising spin systems), and plasmas.[2]]

The problem, as usual, is to obtain the thermodynamic properties and low-order correlation functions of the fluid from the properties of the interatomic potential $v(r)$ believed to consist of a very

* Supported by the U. S. Air Force Office of Scientific Research, the National Science Foundation, and the U. S. Atomic Energy Commission.

† On leave of absence from the Hebrew University, Jerusalem.

[1] S. Baer, J. L. Lebowitz, G. Stell, and W. Theumann (to be published in J. Math. Phys.).

[2] J. L. Lebowitz and G. Stell, Bull. Am. Phys. Soc. 9, 105 (1964); (article to be published in J. Math. Phys.).

strongly repulsive short-range part and a weaker attractive part of longer range. Of particular interest are the properties of the system at liquid densities and the nature of the ever present phase transition. For these purposes the usual virial expansion is not of much theoretical use. As has been shown recently it is possible for the virial expansion to diverge at a density which is either lower or higher than the density at which a phase transition occurs.[3] Also the method often used in the theory of classical fluids, that of approximate integral equations for the radial distribution function, appears to be quite useful for describing the properties of systems interacting via short-range repulsive forces (e.g., hard spheres) but has been less useful for representing the effect of the attractive part of the intermolecular potential,[4] the part responsible for the existence of the liquid.

This suggests developing a theory, or method, which would use as a reference system, or zero-order term, a system of particles interacting via the short-range repulsive part of the potential only: in contrast to the virial expansion which has the zero density gas as a reference system. One such method that has already been used, successfully, for fluids and spin systems employes a high temperature expansion.[5] The method we shall consider is somewhat different and consists essentially of an expansion in the ratio of the ranges of the short-range and long-range part of the interatomic potential. The idea of separating the intermolecular potential into a short-range repulsive and long-range attractive part goes back to van der Waals[6] who used it to derive the famous equation of state bearing his name.

A precise mathematical formulation of van der Waals' idea has been given by Kac[7] and utilized extensively by Kac, Uhlenbeck, and Hemmer.[7] They considered a one-dimensional system with a pair potential $v(r) = q(r) + w(r, \gamma)$,

$$q(r) = \begin{cases} \infty, & r < \delta, \\ 0, & r > \delta, \end{cases}$$

$$w(r, \gamma) = \gamma\varphi(\gamma r) = (\alpha/2)\gamma e^{-\gamma r}, \quad \alpha < 0. \tag{1.1}$$

Kac[7] first showed that as long as γ is finite, there is no phase transition (this is true in general in one dimension for potentials which fall off reasonably fast with distance). However, in the limit $\gamma \to 0$, KUH[7] found the pressure p as a function of density ρ and temperature β^{-1} to be given by the van der Waals equation of state combined with Maxwell's equal area construction,

$$\lim_{\gamma \to 0} p(\rho, \gamma) = \rho\beta^{-1}/(1 - \rho\delta) + \tfrac{1}{2}\alpha\rho^2$$
$$\textit{plus Maxwell's rule.} \tag{1.2}$$

Unfortunately, their actual method of solution depends very much on the exact form of the potential and on it being one dimensional. The main new feature of their potential is the strict separation of $v(r)$ into a short-range part and a *truly long-range* part in the limit $\gamma \to 0$. It should be emphasized that the limit $\gamma \to 0$ is taken *after* the size of the system has been made infinite.

The work of Kac, Uhlenbeck, and Hemmer has been extended recently by Lebowitz and Penrose[8] to higher dimensions and to more general interparticle potentials of the form $v(r) = q(r) + \gamma^{\nu}\varphi(\gamma r)$, where ν is the dimensionality of the space. In this work, which is also related to recent work by van Kampen,[9] it is proven rigorously for a wide class of q's and φ's that in the "van der Waals limit", $\gamma \to 0$, the equation of state assumes the form

$$\lim_{\gamma \to 0} p(\rho, \gamma) = p^0(\rho) + \tfrac{1}{2}\alpha\rho^2$$
$$\textit{plus Maxwell's rule,}$$

where $p^0(\rho)$ is the pressure in the reference system, $w(r) = 0$, and $\alpha \equiv \int \varphi(\mathbf{y})d\mathbf{y}$.

In actual physical systems the potential does not, of course, have infinite range in the above sense. The separation of the potential should still be useful though whenever many particles can fit in within the range of one particle's attractive potential. This was, in fact, the central idea behind the earlier work of Brout for lattice systems which was further

[3] J. L. Lebowitz and O. Penrose, J. Math. Phys. **5**, 841 (1964); O. Penrose, *ibid.* **4**, 1312 (1963).

[4] A major obstacle has been the difficulty in obtaining the solutions of these equations in a tractable enough form to be sure just what they predict about liquids, especially in transition and critical regions. Furthermore even where these equations have been solved numerically it is difficult to assess their worth because of uncertainty of the exact form of intermolecular potentials of real systems and lack of molecular dynamical or Monte Carlo results comparable in accuracy to corresponding hard-sphere results.

[5] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954); E. B. Smith and B. Alder, *ibid.* **30**, 1190 (1959); H. L. Frisch, E. Praestgaard, and J. L. Lebowitz, Bull. Am. Phys. Soc. **9**, (1964); C. Domb, Advan. Phys. Soc. **9**, No. 34, 35 (1960); M. E. Fisher, J. Math. Phys. **4**, 278 (1963).

[6] J. D. van der Waals, Dissertation, Leiden (1873); L. S. Ornstein, dissertation, Leiden (1908).

[7] M. Kac, Phys. Fluids, **2**, 8 (1959). M. Kac. G. Uhlenbeck, and P. C. Hemmer, (abbreivated KUH), J. Math. Phys. **4**, 216, (1963); UHK, *ibid.* p. 229; HKU, *ibid.* **5**, 60 (1964). "KUH" will refer in the text to all four of these articles.

[8] J. L. Lebowitz and O. Penrose, A Rigorous Treatment of the van der Waals–Maxwell Theory of the Liquid-Vapour Transition (to be published in J. Math. Phys.).

[9] N. van Kampen, Phys. Rev. **135**, 362 (1964).

developed and generalized by Horwitz and Callen, Englert, and Coopersmith and Brout.[10]

Following the work of KUH, Lebowitz and Percus[11] investigated the asymptotic form of the radial distribution function in a fluid with particular reference to systems whose interparticle potential has a weak long-range part. Their results, though not rigorous, applied to more general systems (e.g., plasmas) and agreed with the results of KUH.

It is the purpose of our work to exploit further the idea of separating the potential into a short-range and a long-range part. To this end we develop general systematic methods for expanding the correlation functions (and through them the thermodynamic functions) about their values in a reference system whose particles interact only via the short-range part of the potential. Our method utilizes the language of graphs[12] and is related closely to the work of Hemmer.[13] When applied to the lattice systems it also turns out to be related closely to the work of Horwitz and Callen, Englert,[10] and Stillinger,[14] although we have used the usual Mayer cluster expansions as a starting point, rather than introduce separate formalisms as those authors did. We believe our method clarifies the relations between the expansions they developed and the cluster series commonly used to treat continuum fluids. Our work also makes contact with several other recent treatments of systems with long-range forces[14] and we shall discuss these in the appropriate place.

In Sec. II, we develop a general graphical formalism for a system whose interparticle potential is separated, essentially arbitrarily, into a sum of two terms, $v(r) = q(r) + w(r)$. This analysis is entirely formal and its usefulness only becomes apparent in Sec. III where w is specified to have the form $\gamma^{\nu}\varphi(\gamma r)$. The ordering of graphs introduced in Sec. II is then given meaning in terms of an ordering in the parameter γ. An expansion of the two body Ursell function, the direct correlation function, and of the thermodynamic functions in powers of γ is carried out in Secs. IV, V, and VI. In Sec. V, is also introduced a new auxillary function W which is more convenient for some purposes than the direct correlation function. In Sec. VII we discuss briefly some of the uses of our general formalism for lattice gases as well as the limitation of the γ-expansion to regions of density and temperature in which the system is in a single phase. In the Appendices we prove a lemma that we use in obtaining our graphical formalism, and carry its development further.

II. GENERAL FORMALISM

We consider a system of particles interacting via a pair potential $v(r)$ having the form

$$v(r) = q(r) + w(r). \tag{2.1}$$

The functions $v(r)$ and $q(r)$ are assumed to satisfy the conditions necessary for the existence of a stable thermodynamic system,[15] but are otherwise arbitrary. The particles may also be subject to an external one-body potential $u(\mathbf{r})$. The system is represented by a grand canonical ensemble with a temperature $T = (k\beta)^{-1}$ and fugacity z.

The l-particle distribution function $n_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ is defined as the probability density for finding l *distinct* particles at positions $\mathbf{r}_1, \cdots, \mathbf{r}_l$. We define similarly[16] $\hat{n}_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ as the probability density for finding l particles, not necessarily distinct, at positions $\mathbf{r}_1, \cdots, \mathbf{r}_l$. Thus,

$$\hat{n}_1(\mathbf{r}) = n_1(\mathbf{r}),$$

$$\hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) = n_2(\mathbf{r}_1, \mathbf{r}_2) + n_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2), \cdots. \tag{2.2}$$

The l-particle Ursell functions $F_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ are defined[16] in terms of the n_i, $i = 1, \cdots, l$ in such a way that they vanish whenever their arguments decompose into two or more independent sets

$$F_1(\mathbf{r}_1) = n_1(\mathbf{r}_1),$$

$$F_2(\mathbf{r}_1, \mathbf{r}_2) = n_2(\mathbf{r}_1, \mathbf{r}_2) - n_1(\mathbf{r}_1)n_1(\mathbf{r}_2). \tag{2.3}$$

We define $\hat{F}_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ to be the same functions of $\hat{n}_i$ as F_l is of the n_i, i.e.,

[10] J. Mayer, J. Chem. Phys. **18**, 1426 (1950); R. Brout, Phys. Rev. **115**, 824 (1959); **118**, 1009 (1960); G. Horwitz and H. B. Callen, *ibid.* **124**, 1757 (1961); F. Englert, *ibid.* **129**, 567, (1963); M. Coopersmith and R. Brout, *ibid.* **130**, 2539 (1963).

[11] J. L. Lebowitz and J. K. Percus, J. Math. Phys. **4**, 248 (1963).

[12] An ordering of graphs suitable for long-range potentials was used in a somewhat *ad hoc* fashion in the work of Ref. 10. Its use as a tool in a strict γ-expansion appears to have been first used in Ref. 11. (cf. footnote 11 and 14 in that reference).

[13] P. C. Hemmer, J. Math. Phys. **5**, 75 (1964).

[14] F. H. Stillinger, Phys. Rev. **135**, A1646 (1964); A. J. F. Siegert, *Statistical Physics* **3**, *Brandeis Summer Institute 1962* (W. A. Benjamin, Inc., New York, 1963); A. J. F. Siegert, "On the Ising Model with Long-Range Interaction," Northwestern University preprint, 1962; B. Muhlschlegel and H. Zittartz, Z. Physik **175**, 553 (1963); G. A. Baker, Jr., Phys. Rev. **126**, 2072 (1962); E. Helfand, J. Math. Phys. **5**, 127 (1964); J. Percus and G. Yevick, Phys. Rev. **136**, B290 (1964); C. Bloch and J. Langer, J. Math. Phys. **6**, 554 (1965).

[15] D. Ruelle, Helv. Phys. Acta. **36**, (1963). M. Fisher, "The Free Energy of a Macroscopic System," Arch. Ratl. Mech. and Analysis **17**, 377 (1964). These conditions on the potential also guarantee the convergence of the fugacity and virial expansions in a finite domain, cf. Ref. 3. This gives some meaning to our graphical manipulations.

[16] J. L. Lebowitz and J. K. Percus, J. Math. Phys. **4**, 1495 (1963). The definitions on the various correlation functions used here as well as their representation as variational derivatives of the grand partition function is given in Sec. II of this reference.

$$\hat{F}_1(\mathbf{r}_1) = \hat{n}_1(\mathbf{r}_1),$$
$$\hat{F}_2(\mathbf{r}_1, \mathbf{r}_2) = \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) - \hat{n}_1(\mathbf{r}_1)\hat{n}_1(\mathbf{r}_2) \tag{2.4}$$
$$= F_2(\mathbf{r}_1, \mathbf{r}_2) + n_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2).$$

The representation of the n's and F's by means of graphs with Mayer f-bonds and either fugacity vertices $z_1(\mathbf{r}) = z \exp[-\beta u(\mathbf{r})]$ or density vertices $n_1(\mathbf{r})$ is well known.[17,18] For the purposes of our analysis we shall break up each f-bond into "short-range" K-bonds and "long-range" Φ-bonds,[13]

$$f(r_{ij}) = e^{-\beta v(r_{ij})} - 1$$
$$= K(r_{ij})\left[\sum_{m=0}^{\infty} \frac{1}{m!} \Phi^m(r_{ij})\right]$$
$$+ \sum_{m=1}^{\infty} \frac{1}{m!} \Phi^m(r_{ij}), \tag{2.5}$$

where

$$K(r) = e^{-\beta q(r)} - 1, \quad \Phi(r) = -\beta w(r).$$

Pictorially the K-bonds will be represented by dotted lines and the Φ-bonds by solid lines. Our graphs will then consist of points or vertices representing the functions $z_1(\mathbf{r})$ or $n_1(\mathbf{r})$ and K- and Φ-bonds. Between any pair of vertices there can be zero or one K-bond and any number (including zero) of Φ-bonds.[13] We shall call these graphs *composite* graphs. In conformity with the usual graphical notation, vertices colored black represent unlabeled *field points* over which integrations are performed while white vertices represent labeled points (*root points*). Each graph is associated with its corresponding integral (over the field points) divided by $\sigma \prod (t_{ij}!)$, where σ is the symmetry of the graph, t_{ij} is the number of Φ-bonds between the points i and j. The product of t_{ij}'s is taken over all pairs of vertices. As an illustration, we have, for example

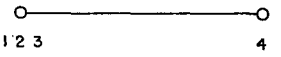
$$= \frac{1}{2!} \int w_1(\mathbf{r}_1) K(\mathbf{r}_1 - \mathbf{x})$$
$$\times \Phi(\mathbf{r}_1 - \mathbf{x}) w_1(\mathbf{x}) [\Phi(\mathbf{x} - \mathbf{r}_2)]^2 w_1(\mathbf{r}_2)\, d\mathbf{x}, \tag{2.6}$$

where $w_1(\mathbf{r})$ stands for either $z_1(\mathbf{r})$ or $n_1(\mathbf{r})$.

In terms of the usual graph language we then have[13,17]

$F_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ = the sum of all composite irreducible (or connected) graphs with $n_1(\mathbf{x})$-vertices [or $z_1(\mathbf{x})$-vertices] having l white vertices labeled by 1, 2, $\cdots$, l, respectively. (2.7)

By irreducible we mean both connected and free of *articulation vertices*, i.e., vertices whose removal would separate the graph into two or more parts, one of which is free of white vertices. The delta-functions that arise in the relationship between the F_l and the $\hat{F}_l$ can be represented by letting any subset of k white vertices coalesce to form a single white vertex labeled with k numbers. Such a vertex represents the function w_1 times the product of delta functions in the differences of the k arguments, e.g.,



$$= w_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2)$$
$$\times \delta(\mathbf{r}_1 - \mathbf{r}_3)\Phi(r_{14})w_1(\mathbf{r}_4). \tag{2.8}$$

We thus have, when $l \geq 2$,

$\hat{F}_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ = the sum of all composite irreducible (or connected) graphs with $n_1(\mathbf{x})$-vertices (or $z_1(\mathbf{x})$-vertices) having $m \leq l$ white vertices, each labeled by a subset of the set of numbers $\{1, \cdots, l\}$. The subsets are disjoint and exhaust the set $\{1, \cdots, l\}$. When $l = 1$ this characterization is still true for the connected graphs with $z_1(\mathbf{x})$-vertices. (2.9)

The characterization given in (2.7) lends itself immediately to an expansion of the F's in powers of the density or powers of the fugacity. This is however not what we are interested in. We desire (for reasons indicated in the introduction and elucidated later) a representation of the F's, and $\hat{F}$'s, in terms of graphs which contain only Φ-bonds. In order to accomplish this we shall consider graphs consisting of Φ-bonds and *hypervertices*. A hypervertex, which represents a function $w_k(\mathbf{r}_1, \cdots, \mathbf{r}_k)$, can be pictured as a *large* circle, along the circumference of which are attached k vertices (or points). We shall call k the *order* of the hypervertex. The small vertices can be either black or white and correspond, respectively, to *field points* over which integrations are performed and to labeled points (*root points*), e.g., 

. Each field point has *one* and only one Φ-bond coming out of it going to another vertex. A graph is associated with its corresponding integral (over the field points) divided by $\sigma \prod (t_{ij}!)$ defined previously, (treating each hypervertex as a

[17] G. Stell in *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz, (W. A. Benjamin Company, Inc., 1964).

[18] G. E. Uhlenbeck and G. W. Ford in *Studies in Statistical Mechanics* edited by J. de Boer and G. E. Uhlenbeck, (North-Holland Publishing Company, Inc., Amsterdam, 1962).

point for the purpose of counting). As an illustration

$$= \frac{1}{2!3!}\int w_3(\mathbf{r}_1, \mathbf{x}_2, \mathbf{x}_3)\Phi(\mathbf{x}_2, \mathbf{x}_4)\Phi(\mathbf{x}_3, \mathbf{x}_5) \times w_5(\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8)\Phi(\mathbf{x}_6, \mathbf{x}_9)\Phi(\mathbf{x}_7, \mathbf{x}_{10}) \times \Phi(\mathbf{x}_8, \mathbf{x}_{11})w_4(\mathbf{x}_9, \mathbf{x}_{10}, \mathbf{x}_{11}, \mathbf{r}_2)\, d\mathbf{x}. \quad (2.10)$$

The usual graph theory, with its point vertices, may be recovered from our formalism by setting $w_k(\mathbf{r}_1, \cdots, \mathbf{r}_k) = w_1(\mathbf{r}_1)\delta(\mathbf{r}_1-\mathbf{r}_2)\delta(\mathbf{r}_1-\mathbf{r}_3)\cdots\delta(\mathbf{r}_1-\mathbf{r}_k)$.

We shall now divide each of the F_l's and $\hat{F}_l$'s into two parts F_l^s and F_l^L. This division is defined in terms of their graphical representation *given in* (2.7), i.e., $F_l^s(\mathbf{r}_1, \cdots, \mathbf{r}_l)$, ($F_l$ short range), *is the subset of all composite graphs in* $F_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ *in which there is a path, consisting of K-bonds alone, connecting the labeled points* $\mathbf{r}_1, \cdots, \mathbf{r}_l$. Then

$$F_l(\mathbf{r}_1, \cdots, \mathbf{r}_l) = F_l^s(\mathbf{r}_1, \cdots, \mathbf{r}_l) + F_l^L(\mathbf{r}_1, \cdots, \mathbf{r}_l). \quad (2.11)$$

The second term in (2.11) will be called the long-range part of F_l. A similar definition applies to $\hat{F}_l$,

$$\hat{F}_l = \hat{F}_l^s + \hat{F}_l^L \quad (2.12)$$

with labeled points replaced by *distinct* labeled points in the above characterization, i.e., the graph ${}_{1}\overset{\circ}{,}{}_{2} = \delta(\mathbf{r}_1 - \mathbf{r}_2)$ belongs to $\hat{F}_2^s$ clearly

$$F_1^s(\mathbf{r}) = \hat{F}_1^s(\mathbf{r}) = n_1(\mathbf{r}). \quad (2.13)$$

The definition of F_l^s and F_l^L is clearly independent of whether composite graphs with $n_1(\mathbf{x})$ or $z_1(\mathbf{x})$ vertices are used.

The F_l^s may themselves be divided into a subset F_l' containing those graphs *with* $z_1(\mathbf{x})$*-vertices* in which there is a path *consisting of K-bonds alone,* connecting *all* the vertices and a remainder F_l''. Thus,

$$F_l^s(\mathbf{r}_1, \cdots, \mathbf{r}_l) = F_l'(\mathbf{r}_1, \cdots, \mathbf{r}_l) + F_l''(\mathbf{r}_1, \cdots, \mathbf{r}_l) \quad (2.14)$$

and similarly for the $\hat{F}_l^s$. Finally F_l' will contain a subset of graphs $F_l'^0$ in which there are no Φ-bonds. This $F_l'^0$, is just the value of F_l considered as a functional of $z_1(x)$ and φ when $\varphi = 0$.

A little thought (see Appendix A for details) now shows that a resummation of the graphs in (2.9) yields the following prescription for the $\hat{F}_l$,

> $\hat{F}_l(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ = the sum of all irreducible (connected) graphs with Φ-bonds and $\hat{F}_k^s(\mathbf{x}_1, \cdots, \mathbf{x}_k)$-hypervertices ($\hat{F}_k'(\mathbf{x}_1, \cdots, \mathbf{x}_k)$-hypervertices), having l white vertices labeled by 1, 2, $\cdots$ l, respectively. (2.15)

In considering the connectedness or irreducibility of a graph each hypervertex is to be thought of as a single point. The simplest graph in (2.15) consists of a single hyper-vertex. In the case where $q(r)$ *vanishes,* $K(r) = 0$, *then* both the F_l^s and F_l' vanish for $l > 1$, and

$$\hat{F}_l^s(\mathbf{x}_1, \cdots, \mathbf{x}_l) = n_1(\mathbf{x}_1)\delta(\mathbf{x}_1 - \mathbf{x}_2)\cdots\delta(\mathbf{x}_1 - \mathbf{x}_l),$$
$$\hat{F}_l'(\mathbf{x}_1, \cdots, \mathbf{x}_l) = z_1(\mathbf{x}_1)\delta(\mathbf{x}_1 - \mathbf{x}_2)\cdots\delta(\mathbf{x}_1 - \mathbf{x}_l), \quad (2.16)$$

and (2.15) reduces to the usual expansion in Φ-bonds and density (fugacity) vertices for systems with interparticle potential $w(r)$. Equation (2.15) may thus be considered a generalization of these expansions to the case where there is an extra term $q(r)$ in the interparticle potential, i.e., the reference system is no longer one in which the potential is zero but one in which the potential is q.

The relation (2.15) expressing $\hat{F}_l^L$ in terms of graphs with Φ-bonds and $\hat{F}_k^s$ hypervertices may be supplemented by an explicit formal expression for $\hat{F}_k^s$. To accomplish this we introduce the notion of the *very long-range* part of F_l, F_l^{sL}; where

> $F_l^{sL}(\mathbf{r}_1, \cdots, \mathbf{r}_l)$ = the subset of F_l^L consisting of all these graphs in which there is no path consisting of K-bonds alone connecting *any* pair of labeled points $\mathbf{r}_1, \cdots, \mathbf{r}_l$. (2.17)

We then have

$$\hat{F}_l^s(\mathbf{r}_1, \cdots, \mathbf{r}_l) = \mathfrak{N}\left\{\exp\left[\sum_{k=2}(1/k!)\int d\mathbf{x}_1 \cdots d\mathbf{x}_k \times F_k^{sL}(\mathbf{x}_1, \cdots \mathbf{x}_k)\delta^k \Big/ \prod_{i=1}^{k}\delta n_1(\mathbf{x}_i)\right]\right\} \times \hat{F}_l^0(\mathbf{r}_1, \cdots \mathbf{r}_l; [n_1(\mathbf{x})]), \quad (2.18)$$

where $\delta/\delta n_1(\mathbf{y})$ means, as usual, the variational derivative with respect to $n_1(\mathbf{y})$ and $\hat{F}_l^0$ is the value of F_l considered as a functional of the density $n_1(\mathbf{x})$ and Φ when $\Phi = 0$. The script $\mathfrak{N}$ indicates a normal order in which all variational derivatives go to the right before evaluation. Equation (2.18) can be most easily obtained by noting the relationship between functional differentiation and a graphical operation[17], since graphically it is a simple identity.

In the case of the ordinary cluster series it was found profitable by a number of workers[17] to con-

sider a resummation of the $n_1(\mathbf{x})$-vertex $f(\mathbf{x}_1, \mathbf{x}_2)$-bond graphical expansions which yields $n_1(\mathbf{x})$-vertex and $[F_2(\mathbf{x}_1, \mathbf{x}_2)/n_1(\mathbf{x}_1)n_1(\mathbf{x}_2)]$-bond expansions. This resummation yields expansions in terms of graphs that are characterized by the absence of articulation pairs of points as well as articulation points. (An articulation pair is a pair of vertices whose removal will disconnect the graph into two or more pieces such that one of the pieces contains at least one unlabeled vertex but no labeled vertices.) It is natural to ask if there is any analogous resummation that we can perform on our graphs with $\hat{F}^*$-hypervertices and Φ-bonds and if such a resummation has any use. The answer to the first question is yes; we shall indicate in a later article why the answer to the second question may also be yes. The new graphs we shall consider will have L-bonds where

$$L(\mathbf{r}_1, \mathbf{r}_2) = \Phi(r_{12}) + \int d\mathbf{r}_3\, d\mathbf{r}_4\, \Phi(r_{13})\hat{F}_2(\mathbf{r}_3, \mathbf{r}_4)\Phi(r_{42}). \tag{2.19}$$

If we use L-bonds instead of Φ-bonds, then instead of (2.15), we find, when $l \geq 2$

$$\hat{F}_l(\mathbf{r}_1, \cdots, \mathbf{r}_l) = \text{the sum of all irreducible graphs with } L\text{-bonds and } \hat{F}_k^*(\mathbf{x}_1, \cdots, \mathbf{x}_k)\text{-hypervertices such that } l \text{ white vertices are labeled by } 1, 2, \cdots, l, \text{ respectively, and the graphs are free of articulation pairs of vertices. (Note that this last requirement is quite different from the restriction that the graphs be free of articulation pairs of } \textit{hyper}\text{-vertices.)} \tag{2.20}$$

Equation (2.20) is obtained from (2.15) on the basis of exactly the same reasoning that yields the $n_1(\mathbf{x})$-vertex, (F_2/n_1n_1)-bond expansion for F_l, $l \geq 3$, from its n_1-vertex, f-bond expansion.[17]

III. SHORT- AND LONG-RANGE FORCES

In the last section the division of the potential into two parts q and w was entirely arbitrary. We now specialize to the case where q and w are truly short range and long range. For this purpose we follow Kac[7] and introduce a parameter γ into w such that γ^{-1} measures the range of w,

$$w(r, \gamma) = \gamma^3\varphi(\gamma r). \tag{3.1}$$

[More generally $\gamma^\nu\varphi(\gamma r)$, where ν is the dimensionality of the space.] The functions q and φ are assumed to have the property that a system of particles interacting via a pair potential $q(r)$ or $v(r)$ has a stable thermodynamic limit[15] for any $\gamma > 0$. [This requires, in particular,[15] that there exists some δ, such that $|q(r)| \leq D/e^{3+\epsilon}$ for $r > \delta$.] We shall require in addition that

$$\text{(a)} \quad \int w(r, \gamma)\, d\mathbf{r} = \int \varphi(y)\, d\mathbf{y} = \alpha < \infty, \tag{3.2}$$

$$\text{(b)} \quad |q(r)| < Ae^{-\lambda r} \quad \text{for} \quad r > \delta;\ A, \lambda \geq 0, \tag{3.3}$$

(c) $\varphi(y)$ is piecewise analytic and bounded for all $y \geq 0$. We have in mind here a $\varphi(y)$ of the type

$$\varphi(y) = (\alpha/8\pi)e^{-y} \quad \text{or} \quad \varphi(y) = \bar{\alpha}(\pi)^{-3/2}e^{-y^2}. \tag{3.4}$$

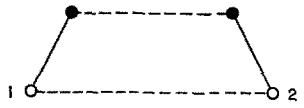
Condition (a) is essential for the existence of the van der Waals limit[8] $\gamma \to 0$. Conditions (b) and (c) are only necessary for the existence (formally at least) of a power-series expansion in γ, [cf. Eqs. (3.6) and (3.7)], which we now discuss. The reason for introducing the parameter γ is to be able to consider the case in which $\gamma \ll \delta^{-1}$ i.e. $\varphi(\gamma r)$ is very long range. Our interest now is to obtain an expansion in ascending orders of γ for thermodynamic quantities and distribution functions of the system, and in particular for the two-particle Ursell function $F_2(\mathbf{r}_1, \mathbf{r}_2)$ or $\hat{F}_2(\mathbf{r}_1, \mathbf{r}_2)$.

In order to carry out this expansion we shall first go to the thermodynamic limit, i.e., let the size of the system become infinite. Assuming for simplicity the absence of any external potential, $u(\mathbf{r}) = 0$, the one-particle density is constant, $n_1(\mathbf{r}) = \rho$, and $F_2(\mathbf{r}_1, \mathbf{r}_2) = F_2(r_{12})$ where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. Now in order for an expansion of $F_2(r, \gamma)$, in powers of γ, to be useful it must have the property that truncation of the expansion after a finite number of terms yield reasonable approximations for those properties of the system which are of physical interest, e.g., x-ray scattering cross sections, equation of state, etc. For this to be the case it is essential that each term in the expansion approach zero as $r \to \infty$, as does $F_2(r)$ itself. On the other hand, consider the value of $F_2(r, \gamma)$ at very low densities $\rho^2(\exp.\{-\beta[q(r) + \gamma^3\varphi(\gamma r)]\} - 1)$. On expansion in γ this would give terms proportional to r^n. What is clearly necessary is to treat the short- and long-range parts of $F_2(r)$, F_2^s and F_2^L, on a different footing. We thus write

$$F_2(r, \gamma) = F_2^s(r, \gamma) + F_2^L(y, \gamma), \qquad y \equiv \gamma r \tag{3.5}$$

and expand them separately as functions of r and y, in powers of γ.

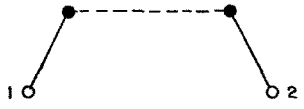
In order to illustrate the difference in treatment of F_2^s and F_2^L we consider two graphs of (2.7) with ρ-vertices, the first belonging to $F_2^s(r_{12}, \gamma)$ and the second to $F_2^L(y_{12}, \gamma)$. Thus, after some change of variables,



$$= \rho^4\beta^2 K(r_{12})\gamma^3 \iint \varphi(\mathbf{x}_1 + \gamma\mathbf{r}_{12})\varphi(\mathbf{x}_1 + \gamma\mathbf{x}_2)K(\mathbf{x}_2)\, d\mathbf{x}_1\, d\mathbf{x}_2$$

$$= \gamma^3\beta^2\rho^4 K(r_{12})\left\{\iint K(\mathbf{x}_2)[\varphi(\mathbf{x}_1)]^2\, d\mathbf{x}_1\, d\mathbf{x}_2 + \gamma \iint K(\mathbf{x}_2)\varphi(\mathbf{x}_1)[\mathbf{r}_{12} + \mathbf{x}_2]\cdot\nabla\varphi(\mathbf{x}_1)\, d\mathbf{x}_1\, d\mathbf{x}_2 + O(\gamma^2)\right\}, \tag{3.6}$$

while



$$= \gamma^3\beta^2\rho^4 \iint \varphi(\mathbf{y}_{12} + \mathbf{x}_1)K(\mathbf{x}_2)\varphi(\mathbf{x}_1 + \gamma\mathbf{x}_2)\, d\mathbf{x}_1\, d\mathbf{x}_2$$

$$= \gamma^3\beta^2\rho^4\left\{\int d\mathbf{x}_2\, K(\mathbf{x}_2) \int \varphi(\mathbf{y}_{12} + \mathbf{x}_1)[\varphi(\mathbf{x}_1) + \gamma\mathbf{x}_2\cdot\nabla\varphi(\mathbf{x}_1)]\, d\mathbf{x}_1 + O(\gamma^2)\right\}. \tag{3.7}$$

In either case the graph is of $O(\gamma^3)$, but in one case the coefficient of γ^3 is a function of r_{12}, and in the other case it is a function of $y_{12} = \gamma r_{12}$. To obtain the general order of a graph appearing in (2.7) or (2.9) we note first that the integrand of a graph containing t Φ-bonds in $O(\gamma^{3t})$. On the other hand, each "free integration", i.e., one not tied down by a K-bond,[13] brings in a factor γ^{-3}. The number of free integrations m is obtained by erasing all the Φ-bonds and counting the number of remaining disjoint components in the K-graph not containing any root point. The graph is then $O[(\gamma^3)^{t-m}]$.

This suggests ordering the composite graphs, with *density vertices* appearing in $\hat{F}_l$, according to the value of $t - m$. For $\hat{F}_l^s$ $m = 0, 1, \cdots$, and $t \geq m$, while for $\hat{F}_l^L$ $m = 0, 1, \cdots$, and $t \geq m + 1$. Thus

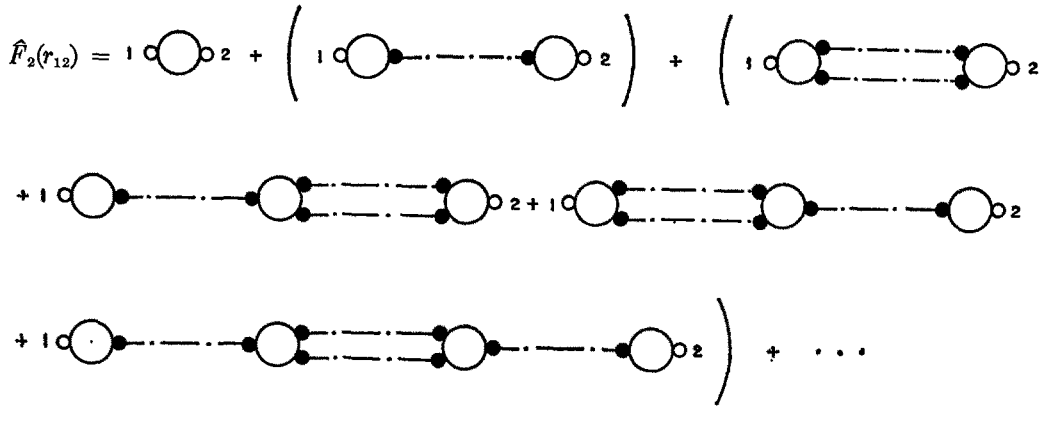
$$\begin{aligned} \hat{F}_l^s &= \hat{F}_{l[0]}^s + \hat{F}_{l[1]}^s + \cdots, \\ \hat{F}_l^L &= \hat{F}_{l[1]}^L + \hat{F}_{l[2]}^L + \cdots, \end{aligned} \tag{3.8}$$

where for each k, $\hat{F}_{l[k]}$ is the sum of all graphs for which $t - m = k$ in the expansion of F_l, (short or long range). An analysis similar to that given in (3.6) and (3.7) shows that $\hat{F}_{l,[m]}$ can be expanded in a power series in γ, starting with γ^{3m},

$$\hat{F}_{l[m]}(\gamma) = \sum_{j=3m}^{\infty} \gamma^j \hat{F}_{l[m]j}. \tag{3.9}$$

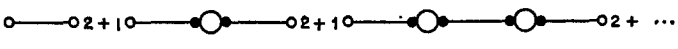
In particular $\hat{F}_{l[0]}^s$ is the value of $\hat{F}_l$ considered as a functional of ρ and φ when $\varphi = 0$. We shall call this ordering of the graphs in (2.7) γ-ordering.

A related, and sometimes useful[1], ordering which also has the property that each successive term starts with a higher power of γ may be applied to the graphs appearing in (2.15). These are ordered according to the difference between the number of Φ-bonds and the number of *hypervertices containing no labeled vertices in the graph*. Thus,



$$= \Gamma_0(r_{12}, \gamma) + \Gamma_1(y_{12}, \gamma; \{\hat{F}_2^s\}) + \Gamma_2(y_{12}, \gamma; \{\hat{F}_2^s, \hat{F}_3^s\}) + \cdots, \tag{3.10}$$

where the dot–dash line, a $\mathcal{C}$-bond, introduced here for convenience, is the sum of all chains,

$\rho^2\mathfrak{C}(y_{12}, \gamma) =$ 

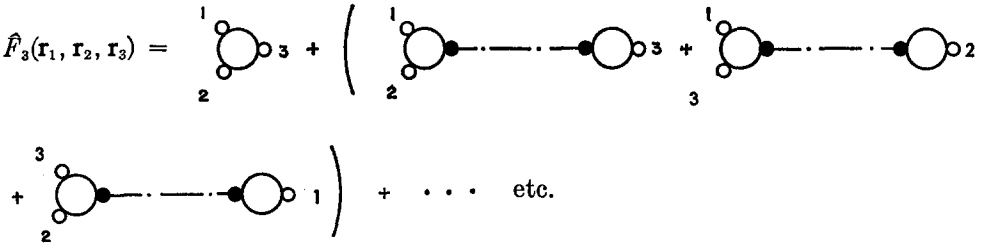
$$= \rho^2\Big[\Phi(y_{12}) + \iint \Phi(\gamma r_{13})\hat{F}_2^s(r_{34}, \gamma)\Phi(\gamma r_{41})\, d\mathbf{r}_3\, d\mathbf{r}_4 + \cdots\Big], \tag{3.11}$$

and is evaluated explicitly in the next section [cf. Eq. (4.3)]. In (3.10),

$$\Gamma_n = O(\gamma^{3n}) = \sum_{i=0}^{\infty} \gamma^{3n+i}\Gamma_{n,i}, \tag{3.12}$$

$$\Gamma_0 = \hat{F}_2^s, \tag{3.13}$$

and we indicated explicitly the functional dependence of Γ_n on $\hat{F}_l^s$, for $l \leq n + 1$. We shall call this ordering of the graphs Γ-ordering. In this ordering we similarly have

$\hat{F}_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) =$ 

etc. (3.14)

We note here that the graphical expansion of $\hat{F}_l$ in terms of graphs with $\mathfrak{C}$-bonds is characterized in the same way as its expansion in terms of graphs with Φ-bonds [Eq. (2.15)] with the added restriction that no graphs contain a hypervertex with two black vertices and no white vertices.

IV. EXPANSION OF F_2 IN POWERS OF γ

In this section we shall obtain explicit expressions for the first few terms in the expansion of $\hat{F}_2$ in powers of γ. These will be similar to those obtained by Hemmer[13] although Hemmer did not make an explicit separation between the short-range and long-range part of the correlations which appears to us essential for clarity. Since the series expansion of $\hat{F}_2^s(r_{12}, \gamma)$ and $\hat{F}_2^L(\gamma r_{12}, \gamma)$ is unique it does not matter when retaining terms up to a certain power in γ, say γ^5, whether one uses the first two terms in the γ- or Γ-ordering, as these will only differ in terms of higher order in γ. We shall use both orderings interchangeably.

For clarity of notation we shall now replace F_2 by $\mathfrak{F}$ (and make similar replacements for all its superscripted and subscripted versions). The series expansion of $\hat{\mathfrak{F}}^s$ will then have the form, [cf. (3.8), (3.9)]

$$\hat{\mathfrak{F}}^s(r_{12}, \gamma) \equiv \Gamma_0 = \hat{\mathfrak{F}}_0(r_{12}) + \gamma^3\mathfrak{F}_3^s(r_{12}) + \gamma^4\mathfrak{F}_4^s(r_{12}) + \cdots, \tag{4.1}$$

where $\hat{\mathfrak{F}}_0(r_{12})$ is the value of $\hat{F}_2$ when $\varphi = 0$, which differs from $\mathfrak{F}_0(r_{12})$ by the terms $\rho\delta(\mathbf{r}_{12})$. There are no $O(\gamma)$ or $O(\gamma^2)$ terms in (4.1) for the type of φ considered here, i.e., satisfying (3.3) and (3.4), [cf. Eq. (6.6)]. The series expansion of $\hat{\mathfrak{F}}^L(y_{12}, \gamma) = \mathfrak{F}^L(y_{12}, \gamma)$ coincides with that of Γ_1 up to $O(\gamma^6)$, [cf. (3.12)],

$$\mathfrak{F}^L(y_{12}, \gamma) = \sum_{i=3}^{\infty} \gamma^i\mathfrak{F}_i^L(y_{12}) = \Gamma_1(y_{12}, \gamma) + O(\gamma^6). \tag{4.2}$$

Utilizing (3.10) and (3.11) we readily find

$$\Gamma_1(y, \gamma) = (2\pi)^{-3}\int e^{i\mathbf{k}\mathbf{y}}\gamma^3[\tilde{\Gamma}_0(\gamma k, \gamma)]^2\mathfrak{C}(k, \gamma)\, d\mathbf{k}$$

$$= (2\pi)^{-3}\int e^{i\mathbf{k}\mathbf{y}}\frac{\gamma^3[\tilde{\Gamma}_0(\gamma k, \gamma)]^2\tilde{\Phi}(k)}{1 - \tilde{\Gamma}_0(\gamma k, \gamma)\tilde{\Phi}(k)}\, d\mathbf{k}$$

$$= \left(\frac{\gamma}{2\pi}\right)^3(\chi^0)^2\int e^{i\mathbf{k}\mathbf{y}}\left[\frac{\tilde{\Phi}(k)}{1 - \chi^0\tilde{\Phi}(k)}\right]d\mathbf{k} + O(\gamma^4), \tag{4.3}$$

where

$$\tilde{\Gamma}_0(k, \gamma) = \int e^{-i\mathbf{k}\cdot\mathbf{r}}\Gamma_0(r, \gamma)\, d\mathbf{r}, \tag{4.4}$$

$$\tilde{\Phi}(k) = \gamma^{-3}\int e^{-i\mathbf{k}\cdot\mathbf{y}}\Phi(y)\, d\mathbf{y} = -\beta\int e^{-i\mathbf{k}\cdot\mathbf{y}}\varphi(y)\, d\mathbf{y}, \tag{4.5}$$

and

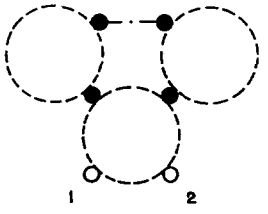
$$\chi^0 = \rho/\left(\beta\frac{dp^0}{d\rho}\right) = \int \hat{\mathfrak{F}}_0(r)\, d\mathbf{r} = \tilde{\Gamma}_{0,0}(0); \tag{4.6}$$

$p^0(\rho, \beta)$ being the pressure and $\beta\chi^0/\rho^2$ the isothermal compressibility in the *absence* of φ. [p^0 differs from $p_0(\rho, \beta)$ which is the van der Waals pressure obtained when γ is set equal to zero *after* the thermodynamic limit has been taken, cf. Eq. (6.2)].

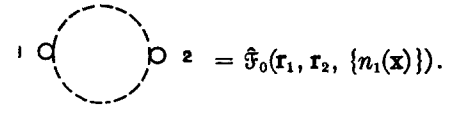
Considering now the next order term in $\hat{\mathfrak{F}}^s$ we note that up to $O(\gamma^6)$ it coincides with $\hat{\mathfrak{F}}^s_{[1]}$ which is given according to (2.18) by

$$\hat{F}^s_{[1]}(\mathbf{r}_{12}, \gamma) = \frac{1}{2}\int \mathfrak{F}^L_{[3]}(\gamma\mathbf{x}_{12}, \gamma) \times \frac{\delta^2\hat{\mathfrak{F}}_0(\mathbf{r}_1, \mathbf{r}_2; \{n_1(\mathbf{x})\}}{\delta n_1(\mathbf{x}_1)\delta n_1(\mathbf{x}_2)}\bigg|_{n_1=\rho} d\mathbf{x}_1\, d\mathbf{x}_2 \qquad (4.7)$$

since for the two-particle Ursell function $\mathfrak{F}^L = \mathfrak{F}^{vL}$. Equation (4.7) may be represented graphically in the form

$\hat{\mathfrak{F}}^s_{[1]} =$  (4.8)

with



$= \hat{\mathfrak{F}}_0(\mathbf{r}_1, \mathbf{r}_2, \{n_1(\mathbf{x})\})$.

Expanding (4.7) in powers of γ we find

$$\hat{\mathfrak{F}}^s_{[1]} = \tfrac{1}{2}\gamma^3\mathfrak{F}^L_3(\gamma r_{12})\partial^2\hat{\mathfrak{F}}_0(r_{12})/\partial\rho^2 + O(\gamma^4) = \tfrac{1}{2}\gamma^3\mathfrak{F}^L_3(0)\partial^2\mathfrak{F}_0(r_{12})/\partial\rho^2 + O(\gamma^4), \qquad (4.9)$$

where use was made in the last equality of the relation [cf. Eq. (2.4)] $\partial^2\hat{\mathfrak{F}}_0/\partial\rho^2 = \partial^2\mathfrak{F}_0/\partial\rho^2$. The reason for writing out the intermediate equality in (4.9) is that it permits the combining of $\mathfrak{F}^s_3(r_{12})$ with $\mathfrak{F}^L_3(\gamma r_{12})$ to give

$$\mathfrak{F}(r) = \mathfrak{F}_0(r) + \tfrac{1}{2}\gamma^3(\partial^2/\partial\rho^2)[n_2^0(r)]\mathfrak{F}^L_3(\gamma r) + O(\gamma^4), \qquad (4.10)$$

where $n_2^0(r)$ is the value of n_2 in the reference system. By combining (4.3) with (4.9) and carrying out explicitly the functional differentiation there we can obtain $\mathfrak{F}$ up to $O(\gamma^6)$ in terms of the properties of the reference system, i.e., one for which $\varphi = 0$. Continuing in this manner it is easy to express $\mathfrak{F}^L$ through $O(\gamma^{11})$ in terms of functions of the reference system. To go beyond this we need to utilize the higher-order terms in (2.18) and the analysis soon gets very complicated.

V. EXPANSION OF THE DIRECT CORRELATION FUNCTION AND A RELATED AUXILIARY FUNCTION

It is convenient for many purposes to introduce the direct correlation function $C(r_{12})$ of Ornstein and Zernike[19], defined by the relation

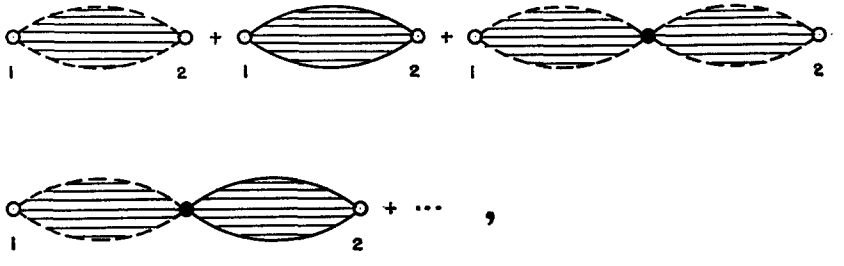
$$\mathfrak{F}(r_{12}) = \rho^2 C(r_{12}) + \rho\int C(r_{13})\mathfrak{F}(r_{32})\, d\mathbf{r}_3. \qquad (5.1)$$

The graphical representation of $\rho^2 C(r_{12})$, in terms of composite graphs with ρ vertices, is similar to that[17] of $\mathfrak{F}(r_{12})$ given in (2.7), with the added restriction that no vertex be a *cutting* or nodal vertex whose removal separates the graph into two or more parts with each of the two white vertices in different parts.

We now divide $C(r_{12})$ like $\mathfrak{F}$, into a short-range and long-range part, (according to whether or not there is a path consisting of short range K-bonds alone connecting the points $\mathbf{r}_1$ and $\mathbf{r}_2$), and write the long-range part as a function of $y = \gamma r$,

$$C(r_{12}) = C^s(r_{12}, \gamma) + C^L(y_{12}, \gamma). \qquad (5.2)$$

The relationship between C^s, C^L, and $\mathfrak{F}^s$ is most readily obtained from the graphical interpretation of Eq. (5.1) as representing $\mathfrak{F}$ by the sum of all distinct repeated convolutions involving C^s and C^L

$\mathfrak{F}(r_{12}) =$  (5.3)

where  represents C^s, 

[19] L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amsterdam, 17, 793 (1914).

represents C^L, and vertices represent ρ. It is now clear that since $\mathfrak{F}^s(r_{12})$ consists of those graphs in $\mathfrak{F}$ in which there is a short-range path from $\mathbf{r}_1$ an $\mathbf{r}_2$ it must be represented by all the convolutions of C^s with itself.

$\mathfrak{F}^s(r_{12}) =$  , (5.4)

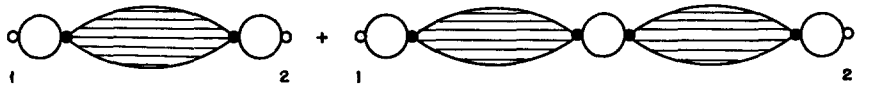
or

$$\mathfrak{F}^s(r_{12}) = \rho^2 C^s(r_{12}) + \rho \int C^s(r_{13})\mathfrak{F}^s(r_{32})\, d\mathbf{r}_3. \tag{5.5}$$

In a similar manner $\mathfrak{F}^L$ will consist of the sum of all graphs containing C^L bonds separated by either a single ρ vertex or by all possible convolutions of C^s with itself, i.e., the C^L bonds will be separated by the hypervertex Γ_0,

$\Gamma_0(r_{12}) = \hat{\mathfrak{F}}^s(r_{12}) = \mathfrak{F}^s(r_{12}) + \rho\delta(\mathbf{r}_1 - \mathbf{r}_2) =$  . (5.6)

Thus,

$\mathfrak{F}^L(r_{12}) =$ 

$$= \int \hat{\mathfrak{F}}^s(r_{13})C^L(y_{34})\hat{\mathfrak{F}}^s(r_{42})\, d\mathbf{r}_3\, d\mathbf{r}_4 + \int \hat{\mathfrak{F}}^s(r_{13})C^L(y_{34})\hat{\mathfrak{F}}^L(y_{42})\, d\mathbf{r}_3\, d\mathbf{r}_4. \tag{5.7}$$

Equations (5.5) and (5.7) may be thought of as the generalization of (5.1) from a system whose particles interact *only with the potential* w to a system with interactions $q + w$. Of course when $q = 0$, C^s and $\mathfrak{F}^s$ vanish, $\Gamma_0(r_{12})$ becomes equal to $\rho\delta(\mathbf{r}_1 - \mathbf{r}_2)$, and (5,7) becomes the usual equation for C.

An inspection of the graphs in (5.5) and (5.1) enables us now to obtain simple relations between terms in the γ-ordering of $\mathfrak{F}$ [cf. Eq. (3.9)], and a similar ordering of C, i.e., $C_{[m]}$ is the sum of all those composite graphs appearing in C in which the number of Φ-bonds less the number of free integration is m. Thus,

$$\mathfrak{F}^s_{[m]}(r_{12}) = \rho^2 C^s_{[m]}(r_{12}) + \rho \sum_{k=0}^{m} \int C^s_{[k]}(r_{13})\mathfrak{F}^s_{[m-k]}(r_{32})\, d\mathbf{r}_3 \tag{5.8}$$

and

$$\mathfrak{F}^L_{[m]}(y_{12}) = \sum_{i+j+k=m} \int \hat{\mathfrak{F}}^s_{[i]}(r_{13})C^L_{[j]}(y_{34})\hat{\mathfrak{F}}^s_{[k]}(r_{42})\, d\mathbf{r}_3\, d\mathbf{r}_4 + \gamma^{-3} \sum_{i+j+k=m+1} \int \hat{\mathfrak{F}}^s_{[i]}(r_{13})C^L_{[j]}(y_{34})\hat{\mathfrak{F}}^L_{[k]}(y_{42})\, d\mathbf{r}_3\, d\mathbf{y}_4, \tag{5.9}$$

where it should be remembered that in carrying out the convolution of C^L and $\mathfrak{F}^L$ there is introduced an extra free integration. Introducing the Fourier transforms $\tilde{C}$ and $\tilde{\mathfrak{F}}$ we find from (5.5) and (5.8),

$$1 - \rho\tilde{C}^s(k, \gamma) = [1 + \rho^{-1}\tilde{\mathfrak{F}}^s(k, \gamma)]^{-1} = [1 + \rho^{-1}\tilde{\mathfrak{F}}_0(k)]^{-1} \times \left[1 + [\rho + \tilde{\mathfrak{F}}_0(k)]^{-1} \sum_{n=1}^{\infty} \tilde{\mathfrak{F}}^s_{[n]}(k, \gamma)\right]^{-1}, \tag{5.10}$$

which gives, upon equating terms of the same order

$$1 - \rho\tilde{C}_0(k) = [1 + \rho^{-1}\tilde{\mathfrak{F}}_0(k)]^{-1}, \tag{5.11}$$

$$\tilde{C}^s_{[1]}(k, \gamma) = [\rho + \tilde{\mathfrak{F}}_0(k)]^{-2}\tilde{\mathfrak{F}}^s_{[1]}(k, \gamma), \cdots \text{etc.}, \tag{5.12}$$

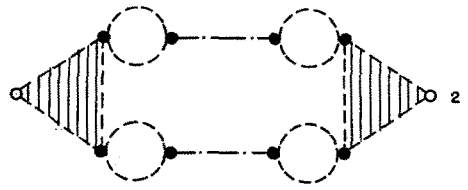
with $\mathfrak{F}^s_{[1]}$ given by (4.8).

In analyzing the long-range parts of C and $\mathfrak{F}$ we must remember, cf. Eq. (4.5), that in any function of $y = \gamma r$, Fourier transformation brings down a factor γ^3, i.e.,

$$\tilde{C}^L(k, \gamma) = \gamma^{-3} \int e^{-i\mathbf{k}\mathbf{y}} C^L(y\gamma)\, d\mathbf{y} \tag{5.13}$$

and similarly for all other functions of y. We then find

$$\tilde{\mathfrak{F}}^L(k, \gamma) = \tilde{C}^L(k, \gamma)[\rho + \tilde{\mathfrak{F}}^s(\gamma k, \gamma)]^2\{1 - [\rho + \tilde{\mathfrak{F}}^s(\gamma k, \gamma)]\tilde{C}^L(k, \gamma)\}^{-1} = (\rho + \tilde{\mathfrak{F}}_0)^2\left[1 + (\rho + \tilde{\mathfrak{F}}_0)^{-1} \sum_{n=1}^{\infty} \tilde{\mathfrak{F}}^s_{[n]}\right]^2 \sum_{n=1}^{\infty} \tilde{C}^L_{[n]} \times [1 - (\rho + \tilde{\mathfrak{F}}_0)\tilde{C}^L_{[1]}]^{-1}\{1 - [1 - (\rho + \tilde{\mathfrak{F}}_0)\tilde{C}^L_{[1]}]^{-1} \sum_{k+m\geq 1} \tilde{\hat{\mathfrak{F}}}^s_{[k]}\tilde{C}^L_{[m+1]}\}^{-1}, \tag{5.14}$$

$\rho^2 C^L_{[2]}(y_{12}, \gamma) =$ 

$$= \frac{\rho^2}{2} \int \cdots \int d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{x}_3\, d\mathbf{x}_4\, \hat{K}^0_3(\mathbf{r}_1, \mathbf{x}_1, \mathbf{x}_2) \mathfrak{F}^L_{[1]}(\gamma|\, \mathbf{x}_1 - \mathbf{x}_3\,|) \mathfrak{F}^L_{[1]}\gamma(|\, \mathbf{x}_2 - \mathbf{x}_4\,|)\, \hat{K}^0_3(\mathbf{x}_3, \mathbf{x}_4, \mathbf{r}_2), \quad (5.19)$$

where the argument of the short-range functions is γk and

$$\mathfrak{F}^L(y, \gamma) = \frac{\gamma^3}{(2\pi)^3} \int e^{i\mathbf{k}\cdot\mathbf{y}} \tilde{\mathfrak{F}}^L(k, \gamma)\, d\mathbf{k}. \quad (5.15)$$

Expanding (5.14) yields

$$\tilde{\mathfrak{F}}^L_{[1]} = (\rho + \tilde{\mathfrak{F}}_0)^2 \tilde{C}^L_{[1]} [1 - (\rho + \tilde{\mathfrak{F}}_0)\tilde{C}^L_{[1]}]^{-1}, \quad (5.16)$$

$$\tilde{\mathfrak{F}}^L_{[2]} = [(\rho + \tilde{\mathfrak{F}}_0)^2 \tilde{C}^L_{[2]} + \gamma^3 \tilde{\mathfrak{F}}^s_{[1]}] \times [1 - (\rho + \tilde{\mathfrak{F}}_0)\tilde{C}^L_{[1]}]^{-2} - \gamma^3 \tilde{\mathfrak{F}}^s_{[1]}, \text{ etc.} \quad (5.17)$$

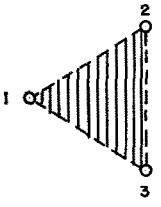
The only graph contributing to $C^L_{[1]}$ is one consisting of a single Φ-bond,

$$C^L_{[1]}(y, \gamma) = \Phi(y, \gamma) = -\beta\gamma^3 \varphi(y), \quad (5.18)$$

so that to this order C^L is equal to $-\beta$ times the long-range potential.[11] Combining (5.18) with (5.16) yields, of course, the same result (except for a factor ρ^2) as the first integrand in (4.3) when Γ_0 there is replaced by $\tilde{\mathfrak{F}}_0$.

The evaluation of $C^L_{[2]}$ is also straightforward. Graphically,

where



$$= \rho^3 \hat{K}^0_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \rho^3 [K^0_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \rho^{-2}\delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(\mathbf{r}_1 - \mathbf{r}_3)]$$

$$= \rho^3 \delta/\delta n_1(\mathbf{r}_3)[C_0(\mathbf{r}_1, \mathbf{r}_2; \{n_1(\mathbf{x})\}) - n_1^{-1}(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2)]_{n_1(\mathbf{x})=\rho}$$

$$= -\rho^3 \delta^2 \ln z_1^0(\mathbf{r}_1, \{n_1(\mathbf{x})\})/\delta n_1(\mathbf{r}_2)\delta n_1(\mathbf{r}_3)|_{n_1(\mathbf{x})=\rho}. \quad (5.20)$$

$\rho^3 K^0_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ is the sum of all diagrams, with ρ vertices and K-bonds, having three distinct labeled vertices and no vertices, black or white, whose removal would separate the graph into two or more parts. K_3 is the "natural" generalization of the direct correlation function C, the superscript zero indicating as usual that it is to be evaluated in the reference system $\varphi = 0$. The second equality follows from the definition of the direct correlation function as a variational derivative,[11]

$$C(\mathbf{r}_1, \mathbf{r}_2; \{n_1(\mathbf{x})\}) = \delta \ln [n_1(\mathbf{r}_1)/z_1(\mathbf{r}_1, \{n_1(\mathbf{x})\})]/\delta n_1(\mathbf{r}_2), \quad (5.21)$$

with $z_1(\mathbf{r}_1) = ze^{-\beta u(\mathbf{r}_1)}$ (u the external potential) considered as a functional of the nonuniform density $n_1(\mathbf{x})$,while z_1^0 is again this functional in the absence of φ. Substituting (5.20) into (5.19) and carrying through an analysis similar to that done in Eq. (4.8) we find that the lowest-order term in $C^L{}_{[2]}$, $C^L{}_{[2],6}$ is given by

$$C^L_{[2],6}(y_{12}) = \tfrac{1}{2}[\mathfrak{F}^L_3(y_{12})]^2 \times \left[\iint \delta^2 \ln z_1^0(\mathbf{r}_1)/\delta n_1(\mathbf{x}_1)\delta n_1(\mathbf{x}_2)|_{n_1=\rho}\, d\mathbf{x}_1,\, d\mathbf{x}_2\right]^2$$

$$= \tfrac{1}{2}[\mathfrak{F}^L_3(y_{12})]^2 [\partial^2 \beta\mu^0/\partial\rho^2]^2, \quad (5.22)$$

where $\mu^0(\rho, \beta)$ is the value of the chemical potential in the reference system, $\rho\partial\mu^0/\partial\rho = \partial p^0/\partial\rho$. By substituting (5.22) into (5.17) and utilizing previously obtained results we may obtain an explicit expression for $\mathfrak{F}^L_{[2],6}$.

Note added in proof: In terms of the Γ-ordering given at the end of Sec. III we have simply[1]

$$\tilde{\mathfrak{F}}^L(\mathbf{k}, \gamma) = \tilde{\Gamma}_1(\mathbf{k}, \gamma) + \tilde{\Gamma}_2(\mathbf{k}, \gamma) + O(\gamma^4)$$

and

$$\tilde{\Gamma}_2(\mathbf{k}, \gamma) = \frac{1}{2}\left(\frac{\partial^2 \beta\mu^0}{\partial\rho^2}\right)^{-2} [1 - x^0 \tilde{\Phi}(\mathbf{k})]\tilde{T}(k) + O(\gamma^4)$$

where $\tilde{T}(k)$ is the Fourier transform of $\mathfrak{C}^2(y)$.

Equation (5.22) is of some interest for its own

sake since in terms of our ordering it makes more precise the popular notion that $C(r) - \Phi(r)$ should be short ranged compared to $\mathfrak{F}(r)$. To second order in the γ-ordering, $C(r) - \Phi(r)$ has "one-half" the range of $\mathfrak{F}(r)$ when the coefficient $(\partial^2\beta\mu^0/\partial\rho^2)$ is not zero. However, the coefficient will be zero for some important special cases—for example, for a field-free Ising model (or lattice gas at $\rho = \frac{1}{2}$)—and in such cases $C - \Phi$ vanishes to second order in the γ ordering.

The function $C(r)$ is introduced in considering fluids as a convenient auxiliary function. The fact that $C(r)$ itself cannot be characterized as a sum of graphs with $\hat{F}_n^s$-hypervertices and Φ-bonds robs it of some of its convenience for our purposes. For this reason we introduce another auxiliary function by subtracting $\Phi(r)$ from $C(r)$ and repeatedly convoluting the result with itself. The sum of $C(r) - \Phi(r)$ and all such convolutions we shall denote by $W(r)$.

$$W(r_{12}) = C(r_{12}) - \Phi(r_{12}) + \int d\mathbf{r}_3[C(r_{13}) - \Phi(r_{13})]\rho[C(r_{32}) - \Phi(r_{32})] + \cdots, \quad (5.23)$$

or in terms of Fourier transforms with respect to r (rather than γr)

$$\tilde{W}(k) = [\tilde{C}(k) - \tilde{\Phi}(k)]/\{1 - \rho[\tilde{C}(k) - \tilde{\Phi}(k)]\}. \quad (5.24)$$

Alternatively, we could define $\tilde{W}(k)$ directly in terms of $\tilde{\mathfrak{F}}(k)$:

$$\tilde{\mathfrak{F}}(k) = \rho^2\left[\frac{\tilde{W}(k) + \tilde{\Phi}(k)[1 + \rho\tilde{W}(k)]}{1 - \rho\tilde{\Phi}(k)[1 + \rho\tilde{W}(k)]}\right] = -\rho + \frac{\rho[1 + \rho\tilde{W}(k)]}{1 - \rho\tilde{\Phi}(k)[1 + \rho\tilde{W}(k)]}. \quad (5.25)$$

It is easy to verify that separating W in the usual way into short- and long-range parts we have $W = W^s + W^L$ where $W^s(r, \gamma)\rho^2 = \mathfrak{F}^s(r, \gamma)$

$\rho^2 W^L(y_{12}, \gamma)$ = the sum of all irreducible graphs with Φ-bonds and $\hat{F}_k^s$-hypervertices such that each graph contains two white vertices labeled by 1 and 2, respectively, and there are no *cutting bonds*, i.e., bonds whose removal separates the graph into two parts, each of which contains a white vertex. (5.26)

In terms of L-bonds rather than Φ-bonds, we have the same representation, except that we must add the restriction that the graphs contain no articulation pairs of vertices. The resulting series is a natural generalization of an expansion that Stillinger[14] has introduced to treat lattice systems and reduces to his expansion for such systems.

VI. EXPANSION OF THERMODYNAMIC FUNCTIONS

In order to obtain the thermodynamic functions as an expansion in powers of γ it is possible either to start directly with a graphical representation[13] of the Helmholtz free energy (or some related function, cf. App. C) or to use one of the several methods which connect thermodynamic properties with integrals over the two-body distribution functions which we have already investigated. We shall begin here by mentioning two commonly used relations of the latter type: the virial theorem and the fluctuation theorem. Both of these methods permit, utilizing our previous results on $\mathfrak{F}$, calculation of the equations of state through $O(\gamma^5)$, a result already obtained by Hemmer.[13] We shall then develop a new method which will utilize our previous results to yield the Helmholtz free energy (and the equation of state) through $O(\gamma^8)$, (or more generally through $O(\gamma^{3\nu-1})$ where ν is the dimensionality of the space.)

The virial theorem may be written in the following form:

$$p = \rho kT - \frac{2}{3}\pi\int_0^\infty q'(r)[\rho^2 + \mathfrak{F}(r)]r^3\,dr + 2\pi\rho^2\int_0^\infty \varphi(y)y^2\,dy - \frac{2}{3}\pi\int_0^\infty \varphi'(y)\mathfrak{F}^L(y)y^3\,dy - \gamma^4\frac{2}{3}\pi\int_0^\infty \varphi'(\gamma r)\mathfrak{F}^s(r)r^3\,dr, \quad (6.1)$$

where q' and φ' are the derivatives of $q(r)$ and $\varphi(y)$ with respect to r and y and we have put the term $\rho^2 + \mathfrak{F}(r) = n_2(r)$, together in the term containing $q'(r)$ to enable us to use the proper limiting procedure in case $q(r)$ has discontinuities (e.g., hard core, square well). Equation (6.1) yields to zero order in γ, the "van der Waals" equation of state,[8]

$$p = \left\{\rho kT - \frac{2}{3}\pi\int q'(r)[\rho^2 + \mathfrak{F}_0(r)]r^3\,dr\right\} + \frac{1}{2}\rho^2\int\varphi(y)\,dy + O(\gamma^3) = p^0 + \tfrac{1}{2}\rho^2\alpha + O(\gamma^3) = p_0 + O(\gamma^3). \quad (6.2)$$

To obtain the $O(\gamma^n)$ corrections to p_0 from (6.1) we need to know both $\mathfrak{F}_i^s$ and $\mathfrak{F}_i^L$ to the same order. Our previous calculations of $\mathfrak{F}_{[1]}^s$ and $\mathfrak{F}_{[1]}^L$ thus enable us, at least in principle, to obtain p through $O(\gamma^5)$ from the properties of the reference system; to go to

higher order we must use $\mathfrak{F}_6^s$ which can be obtained in principle from (2.18).

We next consider the fluctuation theorem [cf. Eq. (4.6)]. This is best expressed in terms of the direct correlation function and may be written in the form[11]

$$\beta\frac{dp}{d\rho} = 1 - \rho\int C^s(r,\gamma)\,d\mathbf{r} - \gamma^{-3}\rho\int C^L(y,\gamma)\,d\mathbf{y}$$
$$= 1 - \rho\tilde{C}^s(0,\gamma) - \rho\tilde{C}^L(0,\gamma). \tag{6.3}$$

While (6.3) is simpler in structure than (6.1) it requires knowledge of C^L to order $n+3$ to obtain p to $O(\gamma^n)$. To zeroth order we, of course, have again as in (6.2)

$$\beta\frac{dp_0}{d\rho} = 1 - \rho\int C_0(r)\,d\mathbf{r} + \beta\rho\int\varphi(y)\,d\mathbf{y} = \beta\frac{dp^0}{d\rho} + \beta\rho\alpha. \tag{6.4}$$

The next-order term in p may now be obtained from (5.12), (4.9), and (5.22). Again we cannot go higher than $O(\gamma^5)$ with the available information on $\mathfrak{F}$ and C.

In order to make further progress we now utilize the functional relation between the Helmholtz free energy and the two particle distribution function n_2. We have, for a uniform system,

$$n_2(r_{12}) = 2\rho\delta A(\rho,\{v\})/\delta v(r_{12}), \tag{6.5}$$

where A is the nonideal-gas part of Helmholtz free energy per particle [the ideal-gas part making no contribution to (6.5) anyway] considered as a functional of the interparticle potential $v(r)$ and the density. Equation (6.5) is a special case of the more general relation holding also for a nonuniform system,

$$\int n_2(\mathbf{r}_{12},\mathbf{r}_2)\,d\mathbf{r}_2 = 2\delta\mathfrak{A}(\{n_1(\mathbf{x})\},\{v\})/\delta v(r_{12}) \tag{6.6}$$

where $\mathfrak{A}$ is the Helmholtz free energy (non-ideal part) of the whole system. Equation (6.6) may be derived either from the graphical representation[17] of $\mathfrak{A}$ or more directly from the definition of $-\beta\mathfrak{A}$, for a system with a fixed number of particles, as the logarithm of the canonical partition function.

We may now rewrite (6.5) in the form

$$\mathfrak{F}(r_{12}) = -2\beta\rho\frac{\delta[A-\frac{1}{2}\rho\alpha]}{\delta\Phi(r_{12})}, \tag{6.7}$$

since $v = q - \beta^{-1}\Phi$. "Inversion" of (6.7) then yields

$$A = A^0 + \tfrac{1}{2}\rho\alpha - \frac{1}{2\beta\rho}\int d\mathbf{r}\int\mathfrak{F}(r,\rho,\{\Phi\})\delta\Phi, \tag{6.8}$$

where $A^0(\beta,\rho)$ is the free energy per particle (nonideal part) in the reference system $\varphi = 0$, and we have indicated explicitly the functional dependence of $\mathfrak{F}$ on Φ. In order to carry out the functional integration indicated in (6.8) we consider a process in which Φ is "turned on" from zero to its final value[4] via a parameter λ. We write

$$\Phi(r,\lambda) = \lambda\Phi(r) = -\lambda\beta\gamma^3\varphi(\gamma r). \tag{6.9}$$

For $\lambda = 0$, the system is in its reference state and for $\lambda = 1$ the system is in its final state. Equation (6.8) may now be written in the form

$$A = A^0 + \tfrac{1}{2}\rho\alpha - \frac{1}{2\beta\rho}\int d\mathbf{r}\int_0^1\mathfrak{F}(r,\rho,\lambda)\Phi(r)\,d\lambda$$
$$= A^0 + \int_0^1 d\lambda\left\{\frac{1}{2}\rho^{-1}\int n_2(r,\lambda)\gamma^3\varphi(\gamma r)\,d\mathbf{r}\right\}. \tag{6.10}$$

We note that in the case where $q(r)$ is a simple hard-core potential, $q(r) = \infty$, for $r < \delta$, and vanishes for $r > \delta$, (including the case $\delta = 0$, or $q = 0$), the term in the bracket in the second equality in (6.10) is just the average potential energy per particle E while the parameter λ enters only in the combination $\lambda\beta$. Equation (6.10) becomes *in this case*

$$A = A^0 + \beta^{-1}\int_0^\beta E(\rho,\beta)\,d\beta, \tag{6.11}$$

which is usual thermodynamic relations between A and E.

Returning now to the general case we may write (6.10) in the form

$$A = A_0 + \tfrac{1}{2}\rho^{-1}\gamma^3\int d\mathbf{r}\,\varphi(\gamma r)\int_0^1\mathfrak{F}^s(r,\gamma,\lambda)\,d\lambda + \tfrac{1}{2}\rho^{-1}\int d\mathbf{y}\,\varphi(\mathbf{y})\int_0^1\mathfrak{F}^L(y,\gamma,\lambda)\,d\lambda, \tag{6.12}$$

where

$$A_0 = A^0 + \tfrac{1}{2}\rho\alpha \tag{6.13}$$

is the van der Waals form of the Helmholtz free energy. We see now that our knowledge of $\mathfrak{F}^s$ to $O(\gamma^5)$ and $\mathfrak{F}^L$ to $O(\gamma^8)$ permits us to compute A to to $O(\gamma^8)$. The calculation to $O(\gamma^6)$ is straightforward and gives[20]

$$\beta A = \beta A_0 - \tfrac{1}{2}\beta\gamma^3\varphi(0) + \rho^{-1}\frac{1}{2}\left(\frac{\gamma}{2\pi}\right)^3\int d\mathbf{k}\ln[1-\tilde{\mathfrak{F}}_0(\gamma k)\tilde{\Phi}(k)]$$
$$- \tfrac{1}{4}\rho^{-1}(\gamma/2\pi)^6\iint d\mathbf{k}\,d\mathbf{k}'\left[\frac{1}{3}\frac{1}{\chi^0}\left(\frac{\partial\chi^0}{\partial\rho}\right)^2\frac{\chi^0\tilde{\Phi}(k')\chi^0\tilde{\Phi}(|\mathbf{k}-\mathbf{k}'|)\chi^0\tilde{\Phi}(k)}{[1-\chi^0\tilde{\Phi}(k)][1-\chi^0\tilde{\Phi}(k')][1-\chi^0\tilde{\Phi}(|\mathbf{k}-\mathbf{k}'|)]}\right.$$
$$\left.+\frac{1}{2}\frac{\partial^2\chi^0}{\partial\rho^2}\frac{\chi^0\tilde{\Phi}(k)\chi^0\tilde{\Phi}(k')}{[1-\chi^0\tilde{\Phi}(k)][1-\chi^0\tilde{\Phi}(k')]}\right] + O(\gamma^7). \tag{6.14}$$

[20] Hemmer's (Ref. 13) expression for A through $O(\gamma^5)$ does not contain the term $-\frac{1}{2}\beta\gamma^3\varphi(0)$. This term makes no contribution to the pressure, cf. (6.16).

Equation (6.14) may be evaluated explicitly for the one-dimensional system investigated by Kac, Uhlenbeck, and Hemmer, cf. Eq. (1.1), to yield

$$\beta A = \ln \rho/(1-\delta\rho) + \tfrac{1}{2}\rho\alpha\beta - \tfrac{1}{4}\gamma\alpha\beta - \tfrac{1}{2}\gamma\rho^{-1}\{1-[1+\alpha\beta\rho(1-\delta\rho)^2]^{\frac{1}{2}}\} + \frac{1}{4}\gamma^2 \frac{\alpha^2\beta^2\rho(1-\delta\rho)^4}{36(1+\alpha\beta\rho(1-\delta\rho)^2)^2}[9\delta(2-3\delta\rho) + \alpha\beta(1-\delta\rho)^2(1+12\delta\rho-18\delta^2\rho^2)] + O(\gamma^3). \quad (6.15)$$

The pressure, internal energy, specific heat and other thermodynamic functions may be obtained directly from (6.14). The specific heat per particle is thus given (in units of Boltzmann's constant) by

$$C = C_0 - \left(\frac{\gamma}{2\pi}\right)^{\nu}\frac{1}{2\rho}\int \frac{\beta^{-2}\frac{\partial^2\chi^0}{\partial(\beta^{-1})^2}\tilde{\Phi}(\mathbf{k})}{1-\chi^0\tilde{\Phi}(\mathbf{k})}\,d\mathbf{k} + \left(\frac{\gamma}{2\pi}\right)^{\nu}\frac{1}{2\rho}\int \frac{\left[\frac{\partial}{\partial\beta}(\beta\chi^0)\tilde{\Phi}(\mathbf{k})\right]^2}{(1-\chi^0\tilde{\Phi}(\mathbf{k}))^2}\,d\mathbf{k} + O(\gamma^{\nu+1}), \quad (6.16)$$

where C_0 is the specific heat of the van der Waal's system which coincides with that of the reference system, C^0, in the *uniform* state. It should be noted here though that as we approach the van der Waal's coexistence curve[8], which coincides with $[1-\chi^0\tilde{\Phi}(0)] = 0$, the coefficient of (γ^ν) in C will become infinite, with the nature of the singularity depending on the number of dimensions ν. Similar singularities will occur in the expansion of all other thermodynamic functions and will get worse for higher terms in the expansion. This expansion therefore breaks down completely near the critical point[7,11,13]: the zero-order term however remaining exact[8], (c.f. Sec. VII).

VII. DISCUSSION

In Sec. II, and parts of Sec. V, of this paper we have set up a general formalism, utilizing graphs, for systems whose interparticle potential consists of two parts $v(r) = q(r) + w(r)$. This formalism uses a system with interparticle potential $q(r)$ as a reference system. We then obtained the first few terms in a γ-expansion of the two-particle correlations, and of the free energy, for the case where w is a Kac type long-range potential of the form $\gamma^\nu\varphi(\gamma r)$, ($\nu$ the dimensionality of the space). It thus serves as a natural extension of Ref. 8 where only the limit $\gamma \to 0$ is considered. In obtaining these few terms in the γ-expansion, the general formalism was, however, utilized only to a limited extent and is almost unnecessary (cf. Hemmer, Ref. 13). The main practical value of the general formalism lies (for us) in lending itself readily to approximations that are not simply the result of considering successively higher powers of γ. Thus in the case of a lattice gas discussed in the next paper[1] we use the Γ-ordering of the graphs and evaluate the hypervertices by requiring that the various Ursell functions obtained successively in this ordering, satisfy certain conditions. The results, even in the lowest order, are an improvement over the mean-field and spherical-model approximations. Also investigation of the behavior of fluids and lattice gases with long (but not infinite) range forces near the critical point appears to be facilitated by approximations for C and W suggested by our Γ-ordering. To make the connection between our expansions and the lattice-system expansions of Brout, Horwitz and Callen, Englert, and Stillinger, we can let $q(\mathbf{r})$ express that impossibility of multiple particle-occupancy of sites and $w(r)$ describe the rest of the interparticle interaction. (For convenience, we use lattice-gas language here rather than spin-system language.) Thus $q(\mathbf{r}_{ij}) = \infty$ when $r_{ij} = 0$, i.e., when i and j refer to the same site, and $q(\mathbf{r}_{ij}) = 0$ otherwise. The functions $w(\mathbf{r}_{ij})$ can be set equal to zero for $\mathbf{r}_{ij} = 0$. For this lattice system, some of the expressions we have derived reduce to expressions already obtained by the authors cited. For example, our $\hat{F}'_k$ and $\hat{F}^*_k$ reduce to Englert's semi-invariants M^0_k and renormalized semi-invariants M_k, respectively.[10] When $\rho = \frac{1}{2}$ the system corresponds to the field-free spin system considered by Stillinger[14] and our $W(\mathbf{r}_{ij})$ reduces to his $W(\mathbf{r}_{ij})$ (for $\mathbf{r}_{ij} \neq 0$). We defer further examination of these correspondences until our next paper.

We now discuss briefly the possible range of validity of our expressions in γ. We distinguish here the case $\alpha < 0$ and $\alpha > 0$. The van der Waals free energy per particle A obtained formally by letting $\gamma \to 0$, (after taking the thermodynamic limit), considered as a function of the volume per particle ρ^{-1} has the form[8]

$$A_0(\rho^{-1}) = A^0(\rho^{-1}) + \tfrac{1}{2}\alpha/\rho^{-1}. \quad (7.1)$$

For $\alpha = 0$, A_0 coincides with A^0 while for $\alpha > 0$, $A^0(\rho^{-1})$ will be a monotonically decreasing concave function[8] of ρ^{-1}. This follows from our assumption that the reference system has a thermodynamic limit so that $A^0(\rho^{-1})$ is a monotonically decreasing nonconvex function[15] of ρ^{-1}. Hence the system characterized by A_0, the "van der Waals system", will show no first-order phase transition since $A_0(\rho^{-1})$ has no linear region. However if the reference system exists in two phases in the density range $\rho_1 < \rho < \rho_2$,

then $\chi^0(\rho)$ is infinite in this range and our whole expansion scheme breaks down [cf. Eqs. (4.3)–(4.6)]. The expansion in γ can thus be meaningful even for small γ only outside this range of densities.

In the second, physically much more relevant, case where $\alpha < 0$, $A_0(\rho^{-1})$ *could* be convex for some range of ρ^{-1} (and some range of the temperature T), and there would be a loop in the p vs. ρ^{-1} curve. This nonphysical behavior is caused[8,9], by the assumption that $n_1(r)$ is constant [cf. going from (6.6) to (6.7)] while in reality $n_1(r)$ would correspond to the coexistence of two phases since that results in a smaller free energy. In this case the expansion in γ with a uniform density can be meaningful only for those densities for which the van der Waals system have a uniform density. This still leaves open the question of the behavior of the system near the critical point when approached from the gas side, the coefficients of the expansion in γ becoming *infinite* there. KUH[7] have shown, for their model, that the γ expansion breaks down in the vicinity of the critical point (cf. also Refs. 11 and 13). It appears to us that this breakdown is specifically associated with the non-analytic (in γ) behavior of the *short-range* part of the distribution functions, and that approximations in the critical region can only be adequately made when the effects of the long- and short-range parts of the distribution functions are considered separately. We shall investigate these questions carefully in future installments of this work.

4. ACKNOWLEDGMENTS

We would like to thank O. Penrose, J. Percus, and E. Lieb for valuable discussions, and we are especially indebted to F. Stillinger for his critical reading of a preliminary version of this article.

APPENDIX A. PROOF OF EQ. (2.15)

Equation (2.15) can be obtained by a resummation from Eq. (2.9) in much the same way that the ordinary $n_1(\mathbf{x})$-vertex, f-bond expansion of F_l can be obtained by resumming the $z_1(\mathbf{x})$-vertex, f-bond expansion.[17] In both cases the only problem is to verify that all counting comes out correctly—the topological aspect of the recharacterization is trivial. We prove here a lemma sufficient to establish (2.15); it is a straightforward generalization of a lemma found in Ref. 17 which was used there to establish the usual virial expansion of F_l, following a method devised by Salpeter and others.[17]

As noted in Sec. 2, we can define $\hat{F}_k^s(\mathbf{x}_1, \cdots, \mathbf{x}_k)$ [or $\hat{F}'_k(\mathbf{x}_1, \cdots, \mathbf{x}_k)$] as a sum of graphs simply by adding to the rhs of (2.9) the stipulation that between any pair of white vertices (or vertices) there can be found a path consisting of K-bonds if $z_1(r)$ vertices are used. We denote a typical graph in this new sum by $\Gamma_\alpha^l(x_1, \cdots, x_l)$ and write

$$\hat{F}_l^s(\mathbf{x}_1, \cdots, \mathbf{x}_l) = \sum_{\alpha \geq 1} \Gamma_\alpha^l(\mathbf{x}_1, \cdots, \mathbf{x}_l)$$

$$[\text{or} \quad \hat{F}'_l(\mathbf{x}_1, \cdots, \mathbf{x}_l) = \sum_{\alpha}^{l} \Gamma_\alpha^l(\mathbf{x}_1, \cdots, \mathbf{x}_l)]. \quad \text{(A1)}$$

Let G be any graph described on the rhs of (2.15) where *connected* and $\hat{F}_k^s$ (or *irreducible* and $\hat{F}'_k$) are considered, and let T be the set of all distinct graphs $T_1, T_2, \cdots$, each of which is obtained by replacing each hypervertex $\hat{F}_l^s$ (or $\hat{F}'_L$) in G by some Γ_α^l. The replacement is to be made in the obvious fashion, i.e., the set of bonds that connects $\hat{F}_l^s$ (or $\hat{F}'_l$) to the rest of G is reattached to the white vertices of Γ_α^l and all of the vertices among these which are replacing black vertices of $\hat{F}_l^s$ (or $\hat{F}'_l$) are then stripped of their labels and blackened. We note that in forming a particular T_1 from G, the same Γ_α^l can be used in replacing several, or all, of the hypervertices $\hat{F}_L^s$ (or $\hat{F}'_l$). We can now prove our Lemma:

$$G = \text{the sum of graphs in } T.$$

Proof. The graph G stands for $1/\sigma_G$ times the integral, where σ_G is the symmetry number of G. We consider the sum of integrals that results when the sum given by (A.1) is substituted for $\hat{F}_l^s$ (or $\hat{F}'_l$) whenever the latter appears in this integral. The result is a product of sums of integrals, which can be rewritten as a single sum if all of the indicated multiplication is done. If the graph G has n black vertices, then a typical term t in this new sum will be some integral $\mathcal{I}$ times $1/\sigma_G\sigma_{\alpha_1}\sigma_{\alpha_2} \cdots \sigma_{\alpha_n}$ where σ_{α_i} is the symmetry of $\Gamma_{\alpha_i}^l$ and the integrand of $\mathcal{I}$ is some number—call it N—times a product of n_1's (or z_1's), Φ's, K's, and factors of the form $\prod (t_{ij}!)$. (See the paragraph above Eq. (2.6) for the definition of t_{ij}). The number N will not necessarily be one, since $\mathcal{I}$ need not correspond to a graph with a symmetry number equal to $\sigma_G\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}$. However, there will be other terms in the sum that are indistinguishable from this one in the sense that they are all products of the same factor $1/\sigma_G\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}$ times integrals that differ from one another (and $\mathcal{I}$) only by a different labeling of the dummy variables of integration. Suppose that there are a total of s such terms, including t. They can be summed up and written as a single term, $s/\sigma_G\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}$ times $\mathcal{I}$. The statement of our lemma amounts to the claim that $\sigma_G\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}/s$ is the symmetry number σ_I of the graph corresponding to I. To verify this

claim, we characterize σ_I in terms of a second labeling. We do this by noting that if we consider the graph obtained from G by labeling the hypervertices 1, 2, $\cdots$ n, then s is the number of distinguishable ways of labeling the hypervertices 1, $\cdots$, n with the labels α_1, $\cdots$, α_n (The labels α_i and α_j may be the same for $i \neq j$ i.e. Γ_{α_i} may be the same graph as Γ_{α_j}. If all the Γ_{α_i}'s are different, $s = n!$ If all the Γ_{α_i}'s are the same, $s = 1$.) Now σ_I can be expressed as $\tilde{\sigma}\sigma_{\alpha_1} \cdots \sigma_{\alpha_i}$ where $\tilde{\sigma}$ is the number of permutations among the $\alpha_1, \alpha_2, \cdots, \alpha_n$ (or equivalently, among the labels 1, 2, $\cdots$, n) that leave the doubly labeled graph invariant. But we have also the equation

$$s\tilde{\sigma} = \sigma_G. \tag{A2}$$

Equation (A2) is an expression of the fact that the group of permutations among the labels 1, 2, $\cdots$, n that leave the doubly labeled graph invariant is a subgroup of the larger group of permutations that leave the singly labeled groups invariant, and the ratio of the orders of the two groups is just the number of distinguished ways that the double labeling can be done. This follows immediately from the appropriate application of Lagrange's theorem relating the order of a group, the order of a subgroup, and the number of cosets induced by the subgroup.

Using Eq. (A2) we can write $\sigma_I = \tilde{\sigma}\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}$ as $\sigma_G\sigma_{\alpha_1} \cdots \sigma_{\alpha_n}/s$ and we are through. Equation (2.15) follows immediately from the lemma.

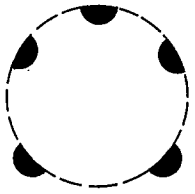
APPENDIX B. EXPANSION OF $\hat{F}_k$ IN TERMS OF REFERENCE SYSTEM HYPERVERTICES

In this Appendix we show more explicitly how Eq. (2.18) can be used to obtain the full expansion for the $\hat{F}_k$ in terms of Φ and functionals of $n_1(x)$ which depend entirely upon the properties of the "unperturbed" system for which $\Phi \equiv 0$.

Combining (2.15) and (2.18), leads to a prescription for $\hat{F}_k$ that can be represented graphically in terms of Φ-bonds and *new* hypervertices which represent the functions

$$\delta^l \hat{F}_k^0/\delta_{n_1}(\mathbf{x}_{k+1}) \cdots \delta_{n_1}(\mathbf{x}_{k+l}) \equiv H_{k,l}(\mathbf{x}_1, \cdots, \mathbf{x}_{k+l}). \tag{B1}$$

Here the $H_{k,l}$-hypervertices must be thought of as having finite extent as we shall consider graphs in which these hypervertices may be connected to each other, as well as to Φ-bonds. In any such connected graph an articulation hypervertex will contain two kinds of vertices—"out" vertices that connect it to the rooted part of a graph and "in" vertices from which are hung pieces that would separate from all the roots if the hypervertex were deleted. (Any labeled vertex is automatically an "out" vertex, and a hypervertex that is not an articulation hypervertex has only "out" vertices.) The $H_{k,l}$ defined above will be represented pictorically by a hypervertex which has k "out" vertices and l "in" vertices.

Thus  $= H_{2,3}$ and we have (B2)

$\hat{F}_m(\mathbf{r}_1, \cdots, \mathbf{r}_m)$ = the sum of all distinct connected graphs, rooted at m points, consisting of Φ-bonds and $H_{k,l}$-hyper-vertices such that there are no articulation vertices (even though there are articulation hypervertices). A vertex belonging to a hypervertex can either have a single bond incident upon it (in which case it must be an "out" vertex) or be shared by another hypervertex (in which case it must be the "out" vertex of one hypervertex and the "in" vertex of the other). Two hypervertices can share at most a single vertex and, as before, a single Φ-bond cannot have both its ends attached to a single hypervertex. (B3)

These expressions involving $H_{k,l}$ are similar to the expressions of Brout[10], and Coopersmith and Brout[10], in the sense that like the vertex functions of those authors, the $H_{k,l}$ are functions of ρ and β and depend only on the unperturbed system for which $\Phi \equiv 0$, (In this respect the expansions of Brout, and Coopersmith and Brout differ from those of Horwitz and Callen and Englert.) However, our expansions are valid for an arbitrary system whereas Coopersmith and Brout were unable to generalize their lattice-system expansions to a continuum fluid without invoking a set of superposition approximations.

APPENDIX C. HYPERVERTEX EXPANSION OF THE THERMODYNAMIC POTENTIALS

Here we supplement the discussion in Sec. VI by considering directly the graphical characterization of the logarithm of Ξ, the grand partition function.

Starting with the standard characterization[17,18] of $\ln \Xi(\{z_1(\mathbf{x})\})$ as the sum of all unrooted composite connected graphs with K- and Φ-bonds and at least two $z_1(\mathbf{x})$ vertices we first isolate a subset $(\ln \Xi)' \equiv F_0'$

which consists of all those graphs in $\ln \Xi$ for which there is a short-range path between any pair of vertices. $F_0'(\{z_1(x)\})$ serves as a generating functional for the F_l' [defined in (2.14)] in precisely the same manner as $\ln \Xi(\{z_1(x)\})$ does for the F_l[16,17],

$$F_l'(\mathbf{r}_1, \cdots, \mathbf{r}_m) = \prod_{i=1}^{l} z_1(x_i) \frac{\delta^l F_0'}{\delta z_1(x_1) \cdots \delta z_1(x_l)}. \quad \text{(C1)}$$

We can now carry through a functional Taylor expansion of $F_0'(\{z_1(x)\})$ about its value at $z_1 = 0$ to yield

$$F_0' = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l!} \int F_l'(\mathbf{x}_1, \cdots, \mathbf{x}_l)\, dx_1 \cdots dx_l. \quad \text{(C2)}$$

Equation (C2) remains valid when either the primes are removed from the F_l or when the F_l' are replaced by $F_l'^0$, the values of the Ursell functions in the reference system. In these cases $F_0 = \ln \Xi(\{z_1(x)\})$ and $F_0'^0 = \ln \Xi^0(\{z_1(x)\})$ where $\Xi^0(z_1(x))$ is the value of $\Xi(\{z_1(\mathbf{x})\})$ when $\Phi \equiv 0$. Having isolated F_0' we now have

$\ln \Xi(\{z_1(\mathbf{x})\}) = F_0'$ + the sum of all unrooted connected graphs composed of Φ-bonds and at least two $\hat{F}_k'$-hypervertices. (C3)

To find the $\ln \Xi$ in terms of Φ and functions that depend only on the properties of the unperturbed reference system in which $\Phi \equiv 0$, we use an identity relating the F_k' and the $F_k'^0$:

$$F_k'(\mathbf{x}_1, \cdots, \mathbf{x}_k) = \exp\left\{\frac{1}{2}\int d\mathbf{y}_1\, d\mathbf{y}_2 \left[z_1(\mathbf{y}_1) z_1(\mathbf{y}_2) \times \Phi(\mathbf{y}_1, \mathbf{y}_2) \frac{\delta^2}{\delta z_1(\mathbf{y}_1)\, \delta z_1(\mathbf{y}_2)} \right] \right\} F_k'^0(\mathbf{x}_1 \cdots \mathbf{x}_k). \quad \text{(C4)}$$

In order to obtain $\ln \Xi$ from (C3) and (C4) we must be given the fugacity z of the system, since the $F_k'^0$ are functions of z. It is somewhat more convenient to be able to express $\ln \Xi$ in terms of ρ directly. We can do this by working through the $\hat{F}_k^s$ and the functions $G_k(\mathbf{x}_1 \cdots \mathbf{x}_k)$ defined by

$$G_l(\mathbf{y}_1, \cdots, \mathbf{y}_l) = \frac{1}{l!} \int \prod_i \Phi(\mathbf{x}_i - \mathbf{y}_i) \times \hat{F}_l(\mathbf{x}_1, \cdots \mathbf{x}_l)\, d\mathbf{x}_1, \cdots, d\mathbf{x}_l. \quad \text{(C5)}$$

The $\hat{F}_l^s$, G_l and $\hat{F}_l'$ are related by an expression very similar to (2.18). It is

$$\hat{F}_l^s(\mathbf{y}_1, \cdots, \mathbf{y}_l) = \mathfrak{N}\left\{ \exp\left[\sum_{k\geq 1} \frac{1}{k!} \int d\mathbf{x}_1 \cdots d\mathbf{x}_k \times G_k(\mathbf{x}_1, \cdots, \mathbf{x}_k) \frac{\delta}{\Pi \delta \ln z_1(\mathbf{x}_i)} \right] \right\} \hat{F}_l'(\mathbf{y}_1, \cdots, \mathbf{y}_l) \quad \text{(C6)}$$

where $\mathfrak{N}$ has the meaning of a normal product as in (2.18). For $l > 0$, (C6) is a graphical identity and for $l = 0$ it defines $\hat{F}_0^s$, which can also be obtained by letting $l = 0$ and $\hat{F}_k^0 = \{\ln \Xi\}^0$ in (2.18).

We can now use the same combinational analysis that has been used by several authors (see Sec. 7 in Ref. 17) to show that $\ln \Xi$ and the excess Helmholtz free energy, $\mathcal{A}$ are related by the equation

$$\ln \Xi = \int d\mathbf{x}\, n_1(\mathbf{x}) - \beta \mathcal{A} + \int dx\, n_1(\mathbf{x}) \frac{\delta\, \beta \mathcal{A}}{\delta n_1(\mathbf{x})}. \quad \text{(C7)}$$

Here $-\beta \mathcal{A}$ = the sum of all irreducible simple unrooted graphs with n_1-vertices and at least one f-bond. This analysis yields in our case the similar-looking equation

$$\ln \Xi = \hat{F}_0^s + S - \sum_{n\geq 1} \int d\mathbf{x}_1, \cdots, d\mathbf{x}_n G_n \hat{F}_n^s, \quad \text{(C8)}$$

where

S = the sum of all irreducible unrooted graphs with $\hat{F}_n^s$ hypervertices and at least one Φ-bond. (C9)

We note that

$$\delta S / \delta \hat{F}_n^s(\mathbf{x}_1, \cdots, \mathbf{x}_n) = G_n(\mathbf{x}_1, \cdots, \mathbf{x}_n) \quad \Phi \text{ and } \hat{F}_l^s \text{ for } l \neq n \text{ fixed} \quad \text{(C10)}$$

so that (C8) can be rewritten in strict analogy with (C7) as

$$\ln \Xi = \hat{F}_0^s + S - \sum_{n\geq 1} \int d\mathbf{x}_1 \cdots d\mathbf{x}_n\, \hat{F}_n^s \frac{\delta S}{\delta \hat{F}_n^s}.$$

Furthermore, when $q(r) = 0$ so that $v(r) = w(r)$, we have

$$S = -\beta a; \qquad \hat{F}_0^s = \int d\mathbf{x}\, n_1(\mathbf{x}),$$

$$\hat{F}_l^s(\mathbf{x}_1, \cdots, \mathbf{x}_l) = n_1(\mathbf{x}_1) \prod_{1\leq i\leq l} \delta(\mathbf{x}_1 - \mathbf{x}_i) \text{ for } l \geq 2$$

and (C8) reduces to (C7). [When $v(r) = 0$, $S \equiv 0$ and $\hat{F}_0^s = \ln \Xi$.]

Since the $\hat{F}_k^s$ and hence the G_k can be expressed in terms of Φ and the functions $H_{k,l} = \delta^l \hat{F}_k^0 / \Pi\, \delta n_1(\mathbf{x}_i)$ by using (2.18) as indicated in Appendix B we have succeeded in obtaining $\ln \Xi$ in terms of Φ and functionals of $n_1(x)$ and β that are defined completely in terms of the reference system.

Many of the equations discussed in this Appendix reduce to lattice-system equations derived by Englert, and Bloch and Langer, where $\mathbf{r}$ is restricted to a set of discrete lattice sets. However some of our equations, notably (2.18), are new even for the special case of lattice systems.

Fokker Action Principle for Particles with Charge, Spin, and Magnetic Moment*

A. SCHILD AND J. A. SCHLOSSER
The University of Texas, Austin, Texas
and
The University of Chicago, Chicago, Illinois
(Received 12 October 1964)

A Fokker action principle is obtained for a system of particles with charge, spin, and magnetic moment, interacting through time-symmetric electromagnetic fields. Conservation laws are derived for the energy, and the linear and angular momentum of the system. A limiting process is performed, which involves a renormalization of mass, in order that the magnitude of the spin of each particle remain constant.

1. INTRODUCTION

IN a previous paper,[1] one of us outlined a program for the study of the relativistic motion and the Bohr quantization of a system of two classical particles in electromagnetic interaction. The principal device in carrying out such a program consists in starting with a Fokker action principle. This results in time-symmetric (half-retarded plus half-advanced) interactions which permit periodic motions to which Bohr quantization can be applied. A Fokker action principle also gives, in a natural manner, finite expressions for the energy, linear momentum, and angular momentum of a system, which automatically include contributions of the electromagnetic field.

In this paper we develop a model of a classical particle which has charge, spin, and magnetic moment.[2] We obtain a Fokker action principle[3] for a system of such particles, and derive conservation laws and expressions for the 4-momentum and angular momentum of the system.

During the motion the spins of the particles change not only in direction but also in magnitude. In order to have a classical model of elementary particles with constant magnitude S of spin, a limiting procedure is adopted where the moment of inertia of each particle tends to zero while S remains finite. This requires a renormalization of the mass. Renormalized equations of motion and renormalized expressions for the conserved quantities are obtained.

A study is in progress of some special motions of two particles in circular orbits. The results of this study will be published in a later paper.

2. KINEMATICS OF A SPINNING PARTICLE

The metric of Minkowski space–time is

$$ds^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$$
$$= -(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2. \qquad (2.1)$$

Greek suffixes range and sum over 1, 2, 3, 4; the the space–time coordinates $x^\mu = (\mathbf{r}, t)$ are real; the Minkowski metric $\eta_{\mu\nu}$ is used to raise and lower tensor suffixes; the scalar product notation $A \cdot B = A_\mu B^\mu = \eta_{\mu\nu} A^\mu B^\nu$ will be used.

The motion of a spinning particle is characterized by

$$x^\mu(u),\ a_{(m)}{}^\mu(u), \qquad (2.2)$$

where

$$a_{(m)} \cdot a_{(n)} = \eta_{mn}, \qquad a_{(m)} \cdot \dot{x} = 0. \qquad (2.3)$$

The translational motion of the particle is determined by its timelike world line $x^\mu(u)$, u being some parameter. The dot denotes differentiation with respect to the parameter u, so that $\dot{x}^\mu = dx^\mu/du$ is the 4-velocity with respect to u.

The rotational motion of the particle is determined by the orthonormal triad of vectors $a_{(m)}{}^\mu$ orthogonal to the world line. Latin suffixes serve to label the different vectors of the triad, and range and sum over 1, 2, 3; $\eta_{mn} = -\delta_{mn}$ serves to raise and lower triad labels, e.g., $a_{(m)}{}^\mu = \eta_{mn} a^{(n)\mu} = -a^{(m)\mu}$. Since

* This research has been sponsored by the Aerospace Research Laboratory, Office of Aerospace Research, U. S. Air Force.

[1] A. Schild, Phys. Rev. **131**, 2762 (1963).

[2] The history of such models is long. The classical papers are J. Frenkel, Z. Phys. **37**, 243 (1926); L. H. Thomas, Nature **117**, 514 (1926); Phil. Mag. **3**, 1 (1927); H. A. Krammers, Physica **1**, 825 (1934). Some recent papers, which include references to earlier work, are: H. C. Corben, Nuovo Cimento **20**, 529 (1961); **28**, 202 (1963); Phys. Rev. **121**, 1833 (1961); **131**, 2219 (1963); **134**, B832 (1964); Proc. Natl. Acad. Sci. U. S. **48**, 387, 1559, 1746 (1962); J. B. Hughes, Nuovo Cimento Suppl. **20**, 89, 148 (1961); P. Nyborg, *ibid.* **23**, 47 (1962); **26**, 821 (1962); R. Schiller, Phys. Rev. **125**, 1100, 1109, 1116 (1962).

[3] K. Schwarzschild, Nachr. Akad. Wiss. Gottingen Math. Physik. Kl. IIa **1903**, 128, 132, 245; H. Tetrode, Z. Physik **10**, 317 (1922); A. D. Fokker, *ibid.* **58**, 386 (1929); Physica **9**, 33 (1929); **12**, 145 (1932); J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945); **21**, 425 (1949); J. W. Dettman and A. Schild, Phys. Rev. **95**, 1057 (1954); A. Schild, *ibid.* **92**, 1009 (1953); A. Schild, in *Proceedings of the International School of Physics "Enrico Fermi" Course* 20: *Evidence for Gravitational Theories* (Academic Press Inc., New York, 1963), pp. 69–115.

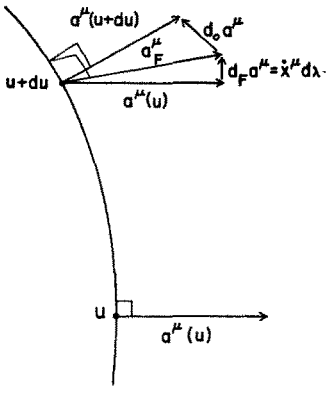


FIG. 1. The Firmi differential.

$a_{(m)}{}^\mu$ and $\dot{x}^\mu$ span the four dimensions of space–time, we have the completeness relation

$$\delta_\mu^\nu = a_{(m)\mu}a^{(m)\nu} + \dot{x}_\mu\dot{x}^\nu/\dot{x}\cdot\dot{x}. \tag{2.4}$$

The triad $a_{(m)}{}^\mu$ plays the role of the body-fixed triad in the Newtonian theory of a rotating rigid body. In Newtonian theory a body is nonrotating, its angular velocity is zero, if the vectors $a_{(m)}{}^\mu$ remain constant. In relativity theory, with the metric (2.1) of Minkowski space, we cannot use the constancy of the components $a_{(m)}{}^\mu$, i.e., the parallel transport of the triad $a_{(m)}{}^\mu$ along the world line $x^\mu(u)$ of the particle. This is because a vector which is originally orthogonal to a curve will not, in general, remain orthogonal to it under parallel transport. We therefore replace parallel transport by Fermi transport. Fermi transport of a vector $a^\mu(u)$, orthogonal to a curve at the point u, to a neighboring point $u + du$ is achieved by adding to the parallel-transported vector a component $d_Fa^\mu = \dot{x}^\mu d\lambda$ tangential to the curve, so that the vector $a_F^\mu(u + du) = a^\mu(u) + d_Fa^\mu$ is orthogonal to the tangent vector $\dot{x}^\mu + \ddot{x}^\mu du$. This gives $d_Fa^\mu = -\dot{x}^\mu\ddot{x}\cdot a du/\dot{x}\cdot\dot{x}$. Consider now a vector field $a^\mu(u)$ defined along the curve $x^\mu(u)$ and which is everywhere orthogonal to the curve, so that

$$\dot{x}\cdot a = 0, \tag{2.5}$$

and therefore $\ddot{x}\cdot a = -\dot{x}\cdot\dot{a}$. The Fermi differential $d_0a^\mu = da^\mu - d_Fa^\mu$ measures the deviation of the vector $a^\mu(u + du)$ from the Fermi-transported vector $a_F^\mu(u + du)$, as shown in Fig. 1. The Fermi derivative ${}^\circ a^\mu = d_0a^\mu/du$ is given by

$${}^\circ a^\mu = \dot{a}^\mu - \dot{x}^\mu\dot{x}\cdot\dot{a}/\dot{x}\cdot\dot{x}, \tag{2.6}$$

and it clearly satisfies

$${}^\circ a\cdot\dot{x} = 0. \tag{2.7}$$

The angular velocity of the spinning particle with respect to the parameter u is measured by the Fermi derivative ${}^\circ a_{(m)}{}^\mu$ of the triad, and is defined by

$$\tilde{\omega}^{\mu\nu} = {}^\circ a_{(m)}{}^\mu a^{(m)\nu} = -\tilde{\omega}^{\nu\mu}, \tag{2.8}$$

or, equivalently, by

$${}^\circ a_{(m)}{}^\mu = \tilde{\omega}^{\mu\nu}a_{(m)\nu}. \tag{2.9}$$

Clearly,

$$\dot{x}_\mu\tilde{\omega}^{\mu\nu} = 0 = \tilde{\omega}^{\mu\nu}\dot{x}_\nu. \tag{2.10}$$

In the instantaneous rest frame of the particle, in which $\dot{x}^\mu = (\mathbf{0}, \dot{x}^4)$, we have $a_{(m)}{}^\mu = (\mathbf{a}_{(m)}, 0)$, ${}^\circ a_{(m)}{}^\mu = (\dot{\mathbf{a}}_{(m)}, 0)$, $\tilde{\omega}^{14} = \tilde{\omega}^{24} = \tilde{\omega}^{34} = 0$. If we define $\boldsymbol{\omega} = (\tilde{\omega}^{23}, \tilde{\omega}^{31}, \tilde{\omega}^{12})$, Eq. (2.9) becomes $\dot{\mathbf{a}}_{(m)} = \tilde{\boldsymbol{\omega}} \times \mathbf{a}_{(m)}$, the usual equation for an angular velocity.

The physical angular velocity of the particle, or angular velocity for short, is the angular velocity $\omega^{\mu\nu}$ with respect to the proper time s. It is given by

$$\omega^{\mu\nu} = \tilde{\omega}^{\mu\nu}/(\dot{x}\cdot\dot{x})^{\frac{1}{2}}. \tag{2.11}$$

Our model for a particle with charge, spin, and magnetic moment will be as follows: The particle is characterized by four constants, the (rest) mass m, the electric charge e, the moment of inertia I, and the gyromagnetic ratio G. The spin $\sigma^{\mu\nu}$ is given by

$$\sigma^{\mu\nu} = I\omega^{\mu\nu}, \tag{2.12}$$

the magnetic moment $\mu^{\mu\nu}$ by

$$\mu^{\mu\nu} = G\sigma^{\mu\nu}. \tag{2.13}$$

This corresponds to a classical rigid body with equal principal moments of inertia, and with a magnetic moment proportional to the spin. For a classical Dirac electron, the gyromagnetic ratio is $G = e/m$.[4]

3. VARIATIONAL PRINCIPLE FOR A PARTICLE IN AN EXTERNAL ELECTROMAGNETIC FIELD

We shall now study the motion of a spinning particle in an external electromagnetic field with 4-potential $A_\mu = (-\mathbf{A}, \Phi)$ and field strengths

$$F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}, \tag{3.1}$$

where the comma denotes partial differentiation with respect to coordinates, e.g., $A_{\mu,\nu} = \partial A_\mu/\partial x^\nu$. The electric and magnetic field components are $\mathbf{E} = (F_{14}, F_{24}, F_{34})$ and $\mathbf{H} = (F_{23}, F_{31}, F_{12})$.

For the action of the particle we choose

$$J_*^{**} = \int_{u^*}^{u^{**}} \tilde{L}\, du, \tag{3.2}$$

$$\tilde{L} = m(\dot{x}\cdot\dot{x})^{\frac{1}{2}} + eA\cdot\dot{x} - \tfrac{1}{4}I\tilde{\omega}_{\mu\nu}\tilde{\omega}^{\mu\nu}/(\dot{x}\cdot\dot{x})^{\frac{1}{2}} - \tfrac{1}{2}IG\tilde{\omega}_{\mu\nu}F^{\mu\nu}. \tag{3.3}$$

[4] The referee kindly points out that this last statement can be challenged. In many respects, a spinning charged particle with $G = 0$ behaves like a classical Dirac electron; cf. H. C. Corben, Phys. Rev. 121, 1833 (1961).

This expression for the action is independent of the choice of the parameter u along the world line of the particle. If the arc length is taken as parameter, it becomes

$$J_*^{**} = \int_{s^*}^{s^{**}} L\,ds = \int_{s^*}^{s^{**}} [m\,ds + eA\cdot dx - \tfrac{1}{4}I\omega_{\mu\nu}\omega^{\mu\nu}\,ds - \tfrac{1}{2}IG\omega_{\mu\nu}F^{\mu\nu}\,ds]. \quad (3.4)$$

In the element of action $\tilde{L}\,du$, the first term is $[m - \frac{1}{2}mv^2 + O(v^4)]dt$ and represents the translational kinetic energy. In the instantaneous rest frame, the third term becomes $-\frac{1}{2}I\omega^2 dt$ and represents the rotational kinetic energy, the second and fourth terms become $e\Phi dt$ and $-\boldsymbol{\mu}\cdot\mathbf{H}\,dt$ and represent the potential energy of the charge e and of the magnetic moment $\boldsymbol{\mu} = IG\boldsymbol{\omega}$ in the external electromagnetic field.

The translational motion of the particle is varied by a variation $\delta x^\mu(u)$ of its world line. The rotational motion is varied by variations $\delta a_{(m)}{}^\mu(u)$ of the body triad. The new triad $a_{(m)}{}^\mu + \delta a_{(m)}{}^\mu$ must be orthonormal and orthogonal to the new world line, whose tangent vector is $\dot{x}^\mu + \delta\dot{x}^\mu$. The most general variation of the triad, which satisfies these conditions, is given by

$$\delta a_{(m)}{}^\mu = -\dot{x}^\mu a_{(m)}\cdot\delta\dot{x}/\dot{x}\cdot\dot{x} + \delta b_{(mn)}a^{(n)\mu}, \quad (3.5)$$

$$\delta b_{(mn)} = -\delta b_{(nm)}. \quad (3.6)$$

The independent variations are the four components of $\delta x^\mu(u)$ and the three components of the skew symmetric expression $\delta b_{(mn)}(u)$.

The equations of motion of the particle are obtained from the principle of stationary action. It states that, for the physical motion,

$$\delta J_*^{**} = 0, \quad (3.7)$$

for variations δx^μ and $\delta b_{(mn)}$ which are arbitrary, except that they must vanish at the two endpoints.

It is convenient to retain at first the general parameter u. Since the action J, Eqs. (3.2) and (3.3), is independent of the parametrization, we can always arrange it so that the physical and the varied world lines have the same initial parameter value u^* and the same final value u^{**}. As soon as the variation has been performed, we simplify the results by identifying u with the arc length s along the physical world line. Thus, after the variation has been performed, we shall have

$$\dot{a} = da/ds, \qquad \dot{x}\cdot\dot{x} = 1, \qquad \tilde{\omega}^{\mu\nu} = \omega^{\mu\nu}. \quad (3.8)$$

We then have

$$\delta\tilde{\omega}^{\mu\nu} = -\delta\dot{b}^{\mu\nu} + 2\delta b^{\alpha[\nu}(\omega^{\mu]}{}_\alpha - \dot{x}^{\mu]}\ddot{x}_\alpha) - 2\ddot{x}^{[\mu}\delta\dot{x}^{\nu]} + 2\ddot{x}^{[\mu}\dot{x}^{\nu]}\dot{x}\cdot\delta\dot{x} + 2\dot{x}^{[\mu}\omega^{\nu]}{}_\alpha\delta\dot{x}^\alpha, \quad (3.9)$$

where square bracketted suffixes are to be skew symmetrized, e.g., $A^{[\mu\nu]} = \frac{1}{2}(A^{\mu\nu} - A^{\nu\mu})$, and where

$$\delta\dot{b}^{\mu\nu} = d\delta d^{\mu\nu}/ds, \qquad \delta b^{\mu\nu} = a^{(m)\mu}a^{(n)\nu}\delta b_{(mn)}, \quad (3.10)$$

so that

$$\delta b^{\mu\nu} = -\delta b^{\nu\mu}, \qquad \delta b^{\mu\nu}\dot{x}_\nu = 0. \quad (3.11)$$

For a tensor defined along the world line $x^\mu(s)$, we shall use the symbol $\perp$ to indicate that all tensor suffixes have been acted on by the projection operator $a_{(m)\mu}a^{(m)\nu} = \delta_\mu^\nu - \dot{x}_\mu\dot{x}^\nu$, $\dot{x}^\mu = dx^\mu/ds$, so that all tangential components are reduced to zero and all components orthogonal to the world line are retained unchanged. For a vector a^μ,

$$(a^\mu)_\perp = a^\mu - \dot{x}^\mu\dot{x}\cdot a, \quad (3.12)$$

and for a skew symmetric tensor $a^{\mu\nu} = -a^{\nu\mu}$,

$$(a^{\mu\nu})_\perp = a^{\mu\nu} - \dot{x}^\mu\dot{x}_\alpha a^{\alpha\nu} - a^{\mu\alpha}\dot{x}_\alpha\dot{x}^\nu. \quad (3.13)$$

We denote the orthogonal projection of $F_{\mu\nu}$ by $H_{\mu\nu}$:

$$H_{\mu\nu} = (F_{\mu\nu})_\perp. \quad (3.14)$$

The Fermi derivative of a vector field $a^\mu = (a^\mu)_\perp$, defined along the world line and orthogonal to it, is given by

$$^\circ a^\mu = (\dot{a}^\mu)_\perp, \quad (3.15)$$

and, similarly, the Fermi derivative of a skew-symmetric tensor $a^{\mu\nu} = -a^{\nu\mu} = (a^{\mu\nu})_\perp$ is defined by

$$^\circ a^{\mu\nu} = (\dot{a}^{\mu\nu})_\perp. \quad (3.16)$$

Equation (3.9) can now be rewritten in the useful form

$$\delta\tilde{\omega}^{\mu\nu} = -(\delta\dot{b}^{\mu\nu})^0 + 2\delta b^{\alpha[\nu}\omega^{\mu]}{}_\alpha - 2\ddot{x}^{[\mu}\delta\dot{x}^{\nu]} + 2\ddot{x}^{[\mu}\dot{x}^{\nu]}\dot{x}\cdot\dot{x}\,\dot{x}\cdot\delta\dot{x} + 2\dot{x}^{[\mu}\omega^{\nu]}{}_\alpha\delta\dot{x}^\alpha. \quad (3.17)$$

The variation of the action is

$$\delta J_*^{**} = \int_{s^*}^{s^{**}} ds\Big[-m\ddot{x}_\mu + e\dot{x}^\nu F_{\nu\mu} - \tfrac{1}{2}G\sigma^{\alpha\beta}F_{\alpha\beta,\mu} + \frac{d}{ds}\Big(\sigma_{\mu\nu}\ddot{x}^\nu - \frac{1}{4I}\sigma_{\alpha\beta}\sigma^{\alpha\beta}\dot{x}_\mu + IGH_{\mu\nu}\ddot{x}^\nu + GF_{\alpha\beta}\dot{x}^\alpha\sigma^\beta{}_\mu\Big)\Big]\delta x^\mu - \frac{1}{2}\int_{s^*}^{s^{**}} ds\,[{}^\circ\sigma_{\mu\nu} + IG{}^\circ H_{\mu\nu} + 2GH_{\alpha[\mu}\sigma_{\nu]}{}^\alpha]\delta b^{\mu\nu} + \Big\{\Big[m\dot{x}_\mu + eA_\mu - \sigma_{\mu\nu}\ddot{x}^\nu + \frac{1}{4I}\sigma_{\alpha\beta}\sigma^{\alpha\beta}\dot{x}_\mu - IGH_{\mu\nu}\ddot{x}^\nu - GF_{\alpha\beta}\dot{x}^\alpha\sigma^\beta{}_\mu\Big]\delta x^\mu\Big\}_{s^*}^{s^{**}} + \{\tfrac{1}{2}[\sigma_{\mu\nu} + IGH_{\mu\nu}]\delta b^{\mu\nu}\}_{s^*}^{s^{**}}. \quad (3.18)$$

Applying the principle of stationary action, the integrated terms drop out, since the variations δx^μ and $\delta b^{\mu\nu}$ vanish at the end points s^* and s^{**}. The translational equations of motion are obtained by equating to zero the coefficient of δx^μ in the first integral:

$$m\ddot{x}_\mu = e\dot{x}^\nu F_{\nu\mu} - \tfrac{1}{2}G\sigma^{\alpha\beta}F_{\alpha\beta,\mu} + \frac{d}{ds}\left(\sigma_{\mu\nu}\ddot{x}^\nu - \frac{1}{4I}\sigma_{\alpha\beta}\sigma^{\alpha\beta}\dot{x}_\mu + IGH_{\mu\nu}\ddot{x}^\nu + GF_{\alpha\beta}\dot{x}^\alpha\sigma^\beta{}_\mu\right). \tag{3.19}$$

The rotational equations of motion are obtained by equating to zero in the second integral, the skew-symmetrized and orthogonally projected components of the coefficient of $\delta b^{\mu\nu}$. These are the coefficients of the three independent variations $\delta b^{\mu\nu} = \delta b^{[\mu\nu]} = (\delta b^{\mu\nu})_\perp$. The rotational equations are

$$^\circ\sigma_{\mu\nu} = 2GH_{\alpha[\mu}\sigma^\alpha{}_{\nu]} - IG{}^\circ H_{\mu\nu}. \tag{3.20}$$

To these equations must be added

$$\sigma_{\mu\nu}\dot{x}^\nu = 0. \tag{3.21}$$

These equations of motion imply $\ddot{x}\cdot\dot{x} = 0$, so that they are consistent with the requirement $\dot{x}\cdot\dot{x} = 1$. They also imply the following first integral of the motion:

$$\tfrac{1}{2}\sigma_{\mu\nu}\sigma^{\mu\nu} + IG\sigma^{\mu\nu}H_{\mu\nu} + \tfrac{1}{2}I^2G^2H^{\mu\nu}H_{\mu\nu} = S^2 = \text{const.} \tag{3.22}$$

Thus our particle will not, in general, have a spin of constant magnitude $(\frac{1}{2}\sigma_{\mu\nu}\sigma^{\mu\nu})^{\frac{1}{2}}$ as it moves in an external electromagnetic field.

We wish to have a classical model of an elementary particle whose spin has constant magnitude. This is achieved by taking the limit of the above equations when $I \to 0$, while $\sigma_{\mu\nu}$, S^2 and

$$m' = m + S^2/2I \tag{3.23}$$

remain finite. We regard m' as the renormalized, physical rest mass. Rewriting m' as m, the renormalized equations of motion are

$$m\ddot{x}_\mu = e\dot{x}^\nu F_{\nu\mu} - \tfrac{1}{2}G\sigma^{\alpha\beta}F_{\alpha\beta,\mu} + (d/ds)(\sigma_{\mu\nu}\ddot{x}^\nu + \tfrac{1}{2}G\sigma^{\alpha\beta}H_{\alpha\beta}\dot{x}_\mu + GF_{\alpha\beta}\dot{x}^\alpha\sigma^\beta{}_\mu), \tag{3.24}$$

$$^\circ\sigma_{\mu\nu} = 2GH_{\alpha[\mu}\sigma^\alpha{}_{\nu]}, \qquad \sigma_{\mu\nu}\dot{x}^\nu = 0. \tag{3.25}$$

These equations have the first integrals $\dot{x}\cdot\dot{x} = 1$ and

$$\tfrac{1}{2}\sigma_{\mu\nu}\sigma^{\mu\nu} = S^2 = \text{const.} \tag{3.26}$$

The renormalized equations of motion agree with those given by Corben and Stehle.[5]

[5] H. Corben and P. Stehle, *Classical Mechanics* (John Wiley & Sons, Inc., New York, 1960), 2nd ed., p. 308.

4. THE FOKKER ACTION PRINCIPLE

We consider an isolated system of N spinning particles, with world lines $x_A^\mu(u_A)$ and angular velocities $\tilde{\omega}_A^{\mu\nu}(u_A)$ with respect to u_A. Capital subscripts label the different particles and range from 1 to N. We do *not* adopt the summation convention for particle labels; all summations over particles will be shown explicitly.

For the action of particle A we take

$$J_A = \int_{-\infty}^{\infty} du_A \left[m_A(\dot{x}_A\cdot\dot{x}_A)^{\frac{1}{2}} - \tfrac{1}{4}I_A\frac{\tilde{\omega}_{A\mu\nu}\tilde{\omega}_A{}^{\mu\nu}}{(\dot{x}_A\cdot\dot{x}_A)^{\frac{1}{2}}} + e_AA_A\cdot\dot{x}_A - \tfrac{1}{2}I_AG_A\tilde{\omega}_A{}^{\mu\nu}F_{A\mu\nu}\right]. \tag{4.1}$$

The following action principle is then equivalent to that of the previous section: for the physical motion of particle A,

$$\delta J_A = 0 \tag{4.2}$$

for arbitrary variations $\delta x_A{}^\mu(u_A)$, $\delta b_{A(mn)}(u_A)$ which vanish outside an arbitrary but finite interval $(\bar{u}_A,\bar{\bar{u}}_A)$ of its world line.

The electromagnetic field $A_{A\mu}$ acting on particle A is that due to the other particles of the system. In order to obtain a Fokker action principle, we must choose the time-symmetric fields of the other particles, i.e., the half-retarded plus half-advanced fields. The time-symmetric potentials due to the charges and magnetic moments of the other particles are given by

$$A_{A\mu} = \sum_{\substack{B\\(B\neq A)}}\left[e_B\int_{-\infty}^{\infty}du_B\,\delta(AB)\dot{x}_{B\mu} + I_BG_B\int_{-\infty}^{\infty}du_B\,\tilde{\omega}_{B\mu}{}^\nu\,\partial_{A\nu}\delta(AB)\right], \tag{4.3}$$

where

$$\delta(AB) = \delta[(x_A - x_B)\cdot(x_A - x_B)] = \delta(BA), \tag{4.4}$$

where $\delta[x]$ is the Dirac delta function, $\dot{x}_{B\mu} = dx_{B\mu}/du_B$, and $\partial_{A\nu} = \partial/\partial x_A{}^\nu$.

If we substitute this into the last two terms of Eq. (4.1), the interaction terms, we obtain

$$\int_{-\infty}^{\infty}du_A\,[e_AA_A\cdot\dot{x}_A - I_AG_A\tilde{\omega}_A{}^{\mu\nu}\partial_{A\nu}A_{A\mu}]$$

$$= \sum_{\substack{B\\(B\neq A)}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}du_A\,du_B\,[e_Ae_B\dot{x}_A\cdot\dot{x}_B\delta(AB) + e_AI_BG_B\dot{x}_{A\mu}\tilde{\omega}_B{}^{\mu\nu}\partial_{A\nu}\delta(AB) - e_BI_AG_A\dot{x}_{B\mu}\tilde{\omega}_A{}^{\mu\nu}\partial_{A\nu}\delta(AB) - I_AG_AI_BG_B\tilde{\omega}_A{}^{\mu\tau}\tilde{\omega}_{B\mu}{}^\sigma\partial_{A\tau}\partial_{A\sigma}\delta(AB)]. \tag{4.5}$$

Because $\delta(AB)$ is a function of the combination $x_A^\mu - x_B^\mu$,

$$\partial_{A\nu}\delta(AB) = -\partial_{B\nu}\delta(AB), \qquad \partial_{A\tau}\partial_{A\sigma}\delta(AB) = \partial_{B\tau}\partial_{B\sigma}\delta(AB). \tag{4.6}$$

It follows that each term of the sum on the right-hand side of Eq. (4.5) is symmetric in the particle labels A and B. It is this symmetry which makes possible the introduction of a single action principle for the system of N interacting particles.

Consider the Fokker action, defined by

$$J = \sum_A \int_{-\infty}^{\infty} du_A\, \Theta_A + \tfrac{1}{2}\sum_{AB}{}' \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} du_A\, du_B\, \Lambda_{AB}, \tag{4.7}$$

$$\Theta_A = m_A(\dot{x}_A\cdot\dot{x}_A)^{\frac{1}{2}} - \tfrac{1}{4}I_A\frac{\tilde{\omega}_{A\mu\nu}\tilde{\omega}_A^{\mu\nu}}{(\dot{x}_A\cdot\dot{x}_A)^{\frac{1}{2}}}, \tag{4.8}$$

$$\Lambda_{AB} = \Lambda_{BA} = e_Ae_B\dot{x}_A\cdot\dot{x}_B\delta(AB) + e_AI_BG_B\dot{x}_{A\mu}\tilde{\omega}_B^{\mu\nu}\partial_{A\nu}\delta(AB) - e_BI_AG_A\dot{x}_{B\mu}\tilde{\omega}_A^{\mu\nu}\partial_{A\nu}\delta(AB) - I_AG_AI_BG_B\tilde{\omega}_A^{\mu\tau}\tilde{\omega}_{B\mu}{}^{\sigma}\partial_{A\tau}\partial_{A\sigma}\delta(AB), \tag{4.9}$$

where the prime on $\sum'_{AB}$ denotes that the double summation is to be carried out over pairs of distinct particle labels only ($A \neq B$).

The equations of motion of the system of particles, interacting through their time symmetric electromagnetic fields, is given by the Fokker action principle

$$\delta J = 0 \tag{4.10}$$

for arbitrary variations $\delta x_A^\mu(u_A)$, $\delta b_{A(mn)}(u_A)$ which vanish outside arbitrary but finite intervals $(\bar{u}_A, \bar{\bar{u}}_A)$ of the world lines.

The part of the Fokker action J which contains the variables describing the motion of particle A, is precisely the expression J_A of Eq. (4.1). Thus when these variables are varied, the Fokker action principle gives the correct motion for particle A. Since the Fokker action is symmetric in the particles of the system, it gives the correct motion of all the particles.

The integrals in the Fokker action J do not all converge, so that J is not a well defined concept. However δJ is always well defined for any choice of finite intervals $(\bar{u}_A, \bar{\bar{u}}_A)$ outside of which the motion is not varied, so that the Fokker action principle (4.10) has meaning.

In order to derive conservation laws from Lorentz invariance we study the finite Fokker action

$$J_*^{**} \equiv \sum_A \int_{u_{A^*}}^{u_{A^{**}}} du_A\, \Theta_A + \tfrac{1}{2}\sum_{AB}{}' \int_{u_{A^*}}^{u_{A^{**}}}\int_{u_{B^*}}^{u_{B^{**}}} du_A\, du_B\, \Lambda_{AB}. \tag{4.11}$$

We now vary the finite action and, as in Sec. 3, introduce the arc lengths as parameters immediately after performing the variation, so that from here on $u_A = s_A$ and dots denote d/ds_A. We obtain

$$\delta J_*^{**} \equiv -\sum_A \int_{s_{A^*}}^{s_{A^{**}}} ds_A\,[\dot{X}_{A\mu}\delta x_A^\mu + \tfrac{1}{2}{}^\circ\sigma_{A\mu\nu}\delta b_A^{\mu\nu}] + \sum_{AB}{}'\int_{s_{B^*}}^{s_{B^{**}}} ds_B\left[Y_{AB\mu}\delta x_A^\mu - \left(\frac{\partial\Lambda_{AB}}{\partial\omega_A^{\mu\nu}}\right)_{\perp A}\delta b_A^{\mu\nu}\right]_{s_{A^*}}^{s_{A^{**}}} + \sum_A[X_{A\mu}\delta x_A^\mu + \tfrac{1}{2}\sigma_{A\mu\nu}\delta b_A^{\mu\nu}]_{s_{A^*}}^{s_{A^{**}}} + \sum_{AB}{}'\int_{s_{A^*}}^{s_{A^{**}}}\int_{s_{B^*}}^{s_{B^{**}}} ds_A\, ds_B\left[\left(\frac{\partial\Lambda_{AB}}{\partial(x_A^\mu - x_B^\mu)} - \frac{dY_{AB\mu}}{ds_A}\right)\delta x_A^\mu + \left\{\left(\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A^{\mu\nu}}\right)_{\perp A}\right)^{0A} + Z_{AB\mu\nu}\right\}\delta b_A^{\mu\nu}\right]. \tag{4.12}$$

Here $\partial\Lambda_{AB}/\partial\omega_A^{\mu\nu}$ means first differentiate Λ_{AB}, Eq. (4.9), with respect to $\tilde{\omega}_A^{\mu\nu}$, then replace $\tilde{\omega}_B^{\mu\nu}$ by $\omega_B^{\mu\nu}$, and finally skew-symmetrize with respect to μ and ν; $\perp A$ means that the orthogonal projection is taken with respect to the world line $x_A^\mu(s_A)$ of particle A; ${}^\circ A$ denotes the Fermi derivative along $x_A^\mu(s_A)$. Also,

$$X_{A\mu} \equiv (m_A + (1/4I_A)\sigma_{A\alpha\beta}\sigma_A^{\alpha\beta})\dot{x}_{A\mu} - \sigma_{A\mu\nu}\ddot{x}_A^\nu,$$

$$Y_{AB\mu} \equiv \frac{\partial\Lambda_{AB}}{\partial\dot{x}_A^\mu} + 2\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A^{\mu\nu}}\right)_{\perp A}\ddot{x}_A^\nu + 2\frac{\partial\Lambda_{AB}}{\partial\omega_A^{\alpha\beta}}\dot{x}_A^\alpha\omega_A^\beta{}_\mu,$$

$$Z_{AB\mu\nu} \equiv 2\left[\frac{\partial\Lambda_{AB}}{\partial\omega_A^{\alpha[\nu}}\omega_A^\alpha{}_{\mu]}\right]_{\perp A}. \tag{4.13}$$

We have used the identity sign to indicate that Eqs. (4.11) to (4.13) hold, whether the equations of motion are satisfied or not.

The Fokker action principle,

$$\lim_{s_{A^*}\to-\infty,\, s_{A^{**}}\to\infty} \delta J_*^{**} = 0, \tag{4.14}$$

for arbitrary variations of the motion which vanish outside arbitrary but finite intervals of the world lines, yields the equations of motion of the system,

$$-\dot{X}_{A\mu} + \sum_{B\,(B\neq A)}\int_{-\infty}^{\infty} ds_B\left(\frac{\partial\Lambda_{AB}}{\partial(x_A^\mu - x_B^\mu)} - \frac{dY_{AB\mu}}{ds_A}\right) = 0, \tag{4.15}$$

$$-\tfrac{1}{2}{}^{\circ}\sigma_{A\mu\nu} + \sum_{\substack{B \\ (B\neq A)}} \int_{-\infty}^{\infty} ds_B \times \left[\left(\left(\frac{\partial \Lambda_{AB}}{\partial \omega_A{}^{\mu\nu}}\right)_{\perp A}\right)^{\circ A} + Z_{AB\mu\nu}\right] = 0. \quad (4.16)$$

By Eqs. (4.5), (4.9), with the choice $u_B = s_B$,

$$\sum_{\substack{B \\ (B\neq A)}} \int_{-\infty}^{\infty} ds_B\, \Lambda_{AB} = e_A A_A \cdot \dot{x}_A - \tfrac{1}{2} I_A G_A \omega_A{}^{\mu\nu} F_{A\mu\nu}. \quad (4.17)$$

Substituting this into Eqs. (4.13), (4.15), and (4.16), we obtain the equations of motion in a form which corresponds to those of Sec. 3:

$$\frac{d}{ds_A}\left[\left(m_A + \frac{1}{4I_A}\sigma_{A\alpha\beta}\sigma_A{}^{\alpha\beta}\right)\dot{x}_{A\mu}\right] = e_A \dot{x}_A{}^{\alpha} F_{A\alpha\mu} - \tfrac{1}{2} G_A \sigma_A{}^{\alpha\beta} F_{A\alpha\beta,\mu} + \frac{d}{ds_A}\left[\sigma_{A\mu\nu}\ddot{x}_A{}^{\nu} + G_A F_{A\alpha\beta}\dot{x}_A{}^{\alpha}\sigma_A{}^{\beta}{}_{\mu} + I_A G_A H_{A\mu\nu}\ddot{x}_A{}^{\nu}\right], \quad (4.18)$$

$${}^{\circ}\sigma_{A\mu\nu} = 2G_A H_{A\alpha[\mu}\sigma_A{}^{\alpha}{}_{\nu]} - I_A G_A {}^{\circ}H_{A\mu\nu}. \quad (4.19)$$

As before, the following integrals of the motion are obtained:

$$\tfrac{1}{2}\sigma_{A\mu\nu}\sigma_A{}^{\mu\nu} + I_A G_A \sigma_A{}^{\mu\nu} H_{A\mu\nu} + \tfrac{1}{2} I_A{}^2 G_A{}^2 H_{A\mu\nu} H^{A\mu\nu} = S_A{}^2, \quad (4.20)$$

where S_A are constants.

5. CONSERVATION OF LINEAR MOMENTUM AND ENERGY

From the Lorentz invariance of the finite Fokker action (4.1) it follows that

$$\delta J_{*}^{**} \equiv 0 \quad (5.1)$$

is satisfied identically for variations of the motion of the system which are induced by infinitesimal Lorentz transformations. This gives rise to ten identities. In this section and the next it will be shown that, by virtue of the equations of motion, these identities can be put into the form of conservation laws.

Consider the variations induced by an infinitesimal space–time translation,

$$\delta x_A{}^{\mu} = \epsilon^{\mu}, \qquad \delta b_A{}^{\mu\nu} = 0, \quad (5.2)$$

where ϵ^{μ} is a set of four infinitesimal constants. By Eqs. (4.12) and (5.1), this gives the identities

$$-\sum_A \int_{s_A*}^{s_A**} ds_A\, \dot{X}_{A\mu} + \sum_A [X_{A\mu}]_{s_A*}^{s_A**} + \sum_{AB}{}' \int_{s_B*}^{s_B**} ds_B\, [Y_{AB\mu}]_{s_A*}^{s_A**} + \sum_{AB}{}' \int_{s_A*}^{s_A**}\int_{s_B*}^{s_B**} ds_A\, ds_B \times \left[\frac{\partial \Lambda_{AB}}{\partial(x_A{}^{\mu} - x_B{}^{\mu})} - \frac{dY_{AB\mu}}{ds_A}\right] \equiv 0. \quad (5.3)$$

Substituting for $\dot{X}_{A\mu}$ from the translational equations of motion (4.15), this becomes

$$\sum_A [X_{A\mu}]_{s_A*}^{s_A**} + \sum_{AB}{}' \int_{-\infty}^{\infty} ds_B\, [Y_{AB\mu}]_{s_A*}^{s_A**} - \sum_{AB}{}' \int_{s_A*}^{s_A**}\left(\int_{-\infty}^{s_B*} + \int_{s_B**}^{\infty}\right) ds_A\, ds_B \frac{\partial \Lambda_{AB}}{\partial(x_A{}^{\mu} - x_B{}^{\mu})} = 0. \quad (5.4)$$

In the last term, the integrand $\partial\Lambda_{AB}/\partial(x_A{}^{\mu} - x_B{}^{\mu})$ is skew-symmetric in A and B. Because of the double summation it follows that the integral operator can be skew-symmetrized to give

$$\frac{1}{2}\int_{s_A*}^{s_A**}\left(\int_{-\infty}^{s_B*} + \int_{s_B**}^{\infty}\right) - \frac{1}{2}\left(\int_{-\infty}^{s_A*} + \int_{s_A**}^{\infty}\right)\int_{s_B*}^{s_B**} \equiv \frac{1}{2}\left(\int_{s_A*}^{\infty}\int_{-\infty}^{s_B*} - \int_{-\infty}^{s_A*}\int_{s_B*}^{\infty}\right) - \frac{1}{2}\left(\int_{s_A**}^{\infty}\int_{-\infty}^{s_B**} - \int_{-\infty}^{s_A**}\int_{s_B**}^{\infty}\right). \quad (5.5)$$

Equation (5.4) can now be written

$$\left[\sum_A X_{A\mu} + \sum_{AB}{}' \int_{-\infty}^{\infty} ds_B\, Y_{AB\mu} + \frac{1}{2}\sum_{AB}{}' \left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B} - \int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right) ds_A\, ds_B \times \frac{\partial \Lambda_{AB}}{\partial(x_A{}^{\mu} - x_B{}^{\mu})}\right]_{*}^{**} = 0. \quad (5.6)$$

Since the s_A* and s_A** are two independent sets of points, we obtain the laws of conservation of linear momentum and energy,

$$P_{\mu} = \sum_A X_{A\mu} + \sum_{AB}{}' \int_{-\infty}^{\infty} ds_B\, Y_{AB\mu} + \frac{1}{2}\sum_{AB}{}' \left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B} - \int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right) ds_A\, ds_B \times \frac{\partial \Lambda_{AB}}{\partial(x_A{}^{\mu} - x_B{}^{\mu})} = \text{const}, \quad (5.7)$$

i.e., P_{μ} is independent of the choice of the points $s_1, s_2, \cdots, s_N$, one on each of the world lines of the system.

Using Eqs. (4.13) and (4.17), we find

$$P_{\mu} = \sum_A \left[\left(m_A + \frac{1}{4I_A}\sigma_{A\alpha\beta}\sigma_A{}^{\alpha\beta}\right)\dot{x}_{A\mu} - \sigma_{A\mu\nu}\ddot{x}_A{}^{\nu} + e_A A_{A\mu} - G_A F_{A\alpha\beta}\dot{x}_A{}^{\alpha}\sigma_A{}^{\beta}{}_{\mu} - I_A G_A H_{A\mu\nu}\ddot{x}_A{}^{\nu}\right]$$

$$+\frac{1}{2}\sum_{AB}{}'\left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B}-\int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right)ds_A\,ds_B \times\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\mu}-x_B{}^{\mu})}. \tag{5.8}$$

6. CONSERVATION OF ANGULAR MOMENTUM

Consider the variations induced by an infinitesimal space–time rotation:

$$\delta x_A{}^{\mu}=\epsilon^{\mu}{}_{\nu}x_A{}^{\nu}, \tag{6.1}$$

where $\epsilon_{\mu\nu}=\eta_{\mu\xi}\epsilon^{\xi}{}_{\nu}$ is a set of six skew-symmetric constants:

$$\epsilon_{\mu\nu}=-\epsilon_{\nu\mu}. \tag{6.2}$$

Under this rotation, $\delta a_{A(m)}{}^{\mu}=\epsilon^{\mu}{}_{\nu}a_{A(m)}{}^{\nu}$ and thus, by Eqs. (3.5), (3.6), (3.10), we obtain

$$\delta b_A{}^{\mu\nu}=-(\epsilon^{\mu\nu})_{\perp A}. \tag{6.3}$$

We also have

$$\delta\omega_A{}^{\mu\nu}=2\epsilon^{\alpha[\mu}\omega_A{}^{\nu]}{}_{\alpha}, \tag{6.4}$$

directly from the tensor character of $\omega_A{}^{\mu\nu}$, or else from Eq. (3.9).

First we establish a simple identity which will be required later. From the Lorentz invariance of Λ_{AB} it follows that

$$\delta\Lambda_{AB}\equiv\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\mu}-x_B{}^{\mu})}\epsilon^{\mu\nu}(x_{A\nu}-x_{B\nu})+\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\mu}}\epsilon^{\mu\nu}\dot{x}_{A\nu} +\frac{\partial\Lambda_{AB}}{\partial\dot{x}_B{}^{\mu}}\epsilon^{\mu\nu}\dot{x}_{B\nu}+2\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\nu\alpha}}\epsilon^{\mu\nu}\omega_A{}^{\alpha}{}_{\mu} +2\frac{\partial\Lambda_{AB}}{\partial\omega_B{}^{\nu\alpha}}\epsilon^{\mu\nu}\omega_B{}^{\alpha}{}_{\mu}\equiv 0. \tag{6.5}$$

Equating to zero the part of the coefficient of $\epsilon^{\mu\nu}$ which is skew symmetric in μ and ν, and rearranging terms, we find that

$$x_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial x_A{}^{\nu]}}+\dot{x}_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\nu]}}+2\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\alpha[\nu}}\omega_A{}^{\alpha}{}_{\mu]} \tag{6.6}$$

is skew symmetric in the particle labels A and B.

Substituting from Eqs. (6.1), (6.3) into Eq. (4.12), we have the identities

$$-\sum_A\int_{s_A^*}^{s_A^{**}}ds_A\,[2x_{A[\mu}\dot{X}_{A\nu]}+{}^{\circ}\sigma_{A\mu\nu}] +\sum_A[2x_{A[\mu}X_{A\nu]}+\sigma_{A\mu\nu}]_{s_A^*}^{s_A^{**}} +\sum_{AB}{}'\int_{s_B^*}^{s_B^{**}}ds_B\,2\left[x_{A[\mu}Y_{AB\nu]}-\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\mu\nu}}\right)_{\perp A}\right]_{s_A^*}^{s_A^{**}} +\sum_{AB}{}'\int_{s_A^*}^{s_A^{**}}\int_{s_B^*}^{s_B^{**}}ds_A\,ds_B\,2$$

$$\times\left[x_{A[\mu}\left(\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\nu]}-x_B{}^{\nu]})}-\frac{dY_{AB\nu]}}{ds_A}\right) +\left(\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\mu\nu}}\right)_{\perp A}\right)^{\circ}{}_A+Z_{AB\mu\nu}\right]\equiv 0. \tag{6.7}$$

Using the equations of motion (4.15), (4.16), and integrating by parts, this becomes

$$\sum_A[2x_{A[\mu}X_{A\nu]}+\sigma_{A\mu\nu}]_{s_A^*}^{s_A^{**}} +\sum_{AB}{}'\int_{-\infty}^{\infty}ds_B\,2\left[x_{A[\mu}Y_{AB\nu]}-\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\mu\nu}}\right)_{\perp A}\right]_{s_A^*}^{s_A^{**}} -\sum_{AB}{}'\int_{s_A^*}^{s_A^{**}}\left(\int_{-\infty}^{s_B^*}+\int_{s_B^{**}}^{\infty}\right)ds_A\,ds_B\,2$$

$$\times\left[x_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\nu]}-x_B{}^{\nu]})}+\dot{x}_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\nu]}} +2\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\alpha[\nu}}\omega_A{}^{\alpha}{}_{\mu]}\right]=0. \tag{6.8}$$

As was shown above, in Eq. (6.6), the last integrand is skew symmetric in A and B. We can now proceed exactly as we did from Eq. (5.4) to Eq. (5.7). We obtain the law of conservation of angular momentum:

$$L_{(0)\mu\nu}=\sum_A[2x_{A[\mu}X_{A\nu]}+\sigma_{A\mu\nu}] +\sum_{AB}{}'\int_{-\infty}^{\infty}ds_B\,2\left[x_{A[\mu}Y_{AB\nu]}-\left(\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\mu\nu}}\right)_{\perp A}\right] +\sum_{AB}{}'\left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B}-\int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right)ds_A\,ds_B$$

$$\times\left[x_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\nu]}-x_B{}^{\nu]})}+\dot{x}_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\nu]}} +2\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\alpha[\nu}}\omega_A{}^{\alpha}{}_{\mu]}\right]=\text{const}, \tag{6.9}$$

i.e., $L_{(0)\mu\nu}$ is independent of the choice of the points $s_1, s_2, \cdots, s_N$, one on each of the world lines of the system.

Using Eqs. (4.13) and (4.17), we find

$$L_{(0)\mu\nu}=\sum_A\left[2\left(m_A+\frac{1}{4I_A}\sigma_{A\alpha\beta}\sigma_A{}^{\alpha\beta}\right)x_{A[\mu}\dot{x}_{A\nu]} -2x_{A[\mu}\sigma_{A\nu]\alpha}\ddot{x}_A{}^{\alpha}+\sigma_{A\mu\nu} +2e_Ax_{A[\mu}A_{A\nu]}-2G_AF_{A\alpha\beta}\dot{x}_A{}^{\alpha}x_{A[\mu}\sigma_A{}^{\beta}{}_{\nu]} -2I_AG_Ax_{A[\mu}H_{A\nu]\alpha}\ddot{x}_A{}^{\alpha}+I_AG_AH_{A\mu\nu}\right] +\sum_{AB}{}'\left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B}-\int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right)ds_A\,ds_B$$

$$\times\left[x_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\nu]}-x_B{}^{\nu])}}+\dot{x}_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\nu]}}+2\frac{\partial\Lambda_{AB}}{\partial\omega_A{}^{\alpha[\nu}}\omega_A{}^{\alpha}{}_{\mu]}\right]. \tag{6.10}$$

The conserved quantity $L_{(0)\mu\nu}$ is the angular momentum of the system about the origin of space–time. The angular momentum $L_{(a)\mu\nu}$ about another event a^μ is obtained by replacing $x_A{}^\mu$ by $x_A{}^\mu - a^\mu$. It is then immediately seen from Eqs. (6.9) and (5.7) that

$$L_{(a)\mu\nu}=L_{(0)\mu\nu}-2a_{[\mu}P_{\nu]}. \tag{6.11}$$

The relativistic center of mass of the system can now be defined in the usual way, assuming P_μ to be timelike. It consists of those events c^μ for which

$$L_{(c)\mu\nu}P^\nu=0. \tag{6.12}$$

The solution is

$$c^\mu=L_{(0)}{}^{\mu\nu}P_\nu/P^\alpha P_\alpha+\lambda P^\mu, \tag{6.13}$$

where λ is arbitrary. These events c^μ clearly lie on a straight line parallel to P^μ, the world line of the center of mass.

7. THE RENORMALIZED EQUATIONS.

In order to have a theory for a system of particles, each with a spin of constant magnitude S_A, we perform the limiting operation and mass renormalization discussed in Sec. 3. We let $I_A\to 0$, while $\sigma_A{}^{\mu\nu}$, $S_A{}^2$ and

$$m'_A=m_A+S_A{}^2/2I_A \tag{7.1}$$

remain finite. This procedure not only gives finite equations of motion, but also finite expressions for the conserved quantities. Following the mass renormalization, we drop the prime and rewrite m'_A as m_A. We then have the following:

The equations of motions (4.18), (4.19) become

$$m_A\ddot{x}_{A\mu}=e_A\dot{x}_A{}^\alpha F_{A\alpha\mu}-\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}F_{A\alpha\beta,\mu}+(d/ds_A)[\sigma_{A\mu\nu}\ddot{x}_A{}^\nu+\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}H_{A\alpha\beta}\dot{x}_{A\mu}+G_AF_{A\alpha\beta}\dot{x}_A{}^\alpha\sigma_A{}^\beta{}_\mu], \tag{7.2}$$

$$^\circ\sigma_{A\mu\nu}=2G_AH_{A\alpha[\mu}\sigma_A{}^\alpha{}_{\nu]},\qquad \sigma_{A\mu\nu}\dot{x}_A{}^\nu=0. \tag{7.3}$$

They have the first integrals

$$\tfrac{1}{2}\sigma_{A\alpha\beta}\sigma_A{}^{\alpha\beta}=S_A{}^2. \tag{7.4}$$

The linear momentum (5.8) and the angular momentum (6.10) of the system become

$$P_\mu=\sum_A[(m_A-\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}H_{A\alpha\beta})\dot{x}_{A\mu}-\sigma_{A\mu\nu}\ddot{x}_A{}^\nu+e_AA_{A\mu}-G_AF_{A\alpha\beta}\dot{x}_A{}^\alpha\sigma_A{}^\beta{}_\mu]+\frac{1}{2}\sum_{AB}{}'\left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B}-\int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right)ds_A\,ds_B\times\frac{\partial\Lambda_{AB}}{\partial(x_A{}^\mu-x_B{}^\mu)}, \tag{7.5}$$

$$L_{(0)\mu\nu}=\sum_A[2(m_A-\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}H_{A\alpha\beta})x_{A[\mu}\dot{x}_{A\nu]}-2x_{A[\mu}\sigma_{A\nu]\alpha}\ddot{x}_A{}^\alpha+\sigma_{A\mu\nu}+2e_Ax_{A[\mu}A_{A\nu]}-2G_AF_{A\alpha\beta}\dot{x}_A{}^\alpha x_{A[\mu}\sigma_A{}^\beta{}_{\nu]}]+\sum_{AB}{}'\left(\int_{s_A}^{\infty}\int_{-\infty}^{s_B}-\int_{-\infty}^{s_A}\int_{s_B}^{\infty}\right)ds_A\,ds_B\times\left[x_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial(x_A{}^{\nu]}-x_B{}^{\nu])}}+\dot{x}_{A[\mu}\frac{\partial\Lambda_{AB}}{\partial\dot{x}_A{}^{\nu]}}+2\frac{\partial\Lambda_{AB}}{\partial\sigma_A{}^{\alpha[\nu}}\sigma_A{}^\alpha{}_{\mu]}\right], \tag{7.6}$$

where

$$\Lambda_{AB}=e_Ae_B\dot{x}_A\cdot\dot{x}_B\,\delta(AB)+e_AG_B\dot{x}_{A\mu}\sigma_B{}^{\mu\nu}\,\partial_{A\nu}\delta(AB)-e_BG_A\dot{x}_{B\mu}\sigma_A{}^{\mu\nu}\,\partial_{A\nu}\delta(AB)-G_AG_B\sigma_A{}^{\mu\zeta}\sigma_{B\mu}{}^\rho\,\partial_{A\zeta}\,\partial_{A\rho}\delta(AB). \tag{7.7}$$

Equation (6.11) remains valid for the renormalized conserved quantities, and Eq. (6.13) gives the relativistic center of mass.

We shall now check directly that the renormalized momenta are conserved as a consequence of the renormalized equations of motion. We calculate the partial derivative of P_μ, Eq. (7.5), with respect to s_A, i.e., with respect to the motion of the point chosen on the world line of particle A along that world line, leaving unchanged the other $N-1$ points on the world lines of the other particles,

$$\frac{\partial P_\mu}{\partial s_A}=\frac{d}{ds_A}[(m_A-\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}H_{A\alpha\beta})\dot{x}_{A\mu}-\sigma_{A\mu\nu}\ddot{x}_A{}^\nu-G_AF_{A\alpha\beta}\dot{x}_A{}^\alpha\sigma_A{}^\beta{}_\mu]+e_AA_{A\mu,\nu}\dot{x}_A{}^\nu-\sum_{\substack{B\\(B\neq A)}}\int_{-\infty}^{\infty}ds_B\frac{\partial\Lambda_{AB}}{\partial(x_A{}^\mu-x^{-\mu})}. \tag{7.8}$$

By Eq. (4.17), the last two terms become

$$e_AA_{A\mu,\nu}\dot{x}_A{}^\nu-e_AA_{A\nu,\mu}\dot{x}_A{}^\nu+\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}F_{A\alpha\beta,\mu}=-e_A\dot{x}_A{}^\nu F_{A\nu\mu}+\tfrac{1}{2}G_A\sigma_A{}^{\alpha\beta}F_{A\alpha\beta,\mu}.$$

Substituting this into Eq. (7.8) we find that

$$\partial P_\mu/\partial s_A=0$$

by virtue of the equations of motion (7.2). Similarly, we show that $\partial L_{(0)\mu\nu}/\partial s_A=0$, so that the conservation laws remain valid after renormalization.

Born Reciprocity Principle and Unitary Symmetry

ERNEST E. H. SHIN
National Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts*
(Received 29 June 1964)

In view of the past attempts by Born *et al.* to explain elementary particles using the Born reciprocity principle as a postulate and of the recent success of unitary symmetry schemes, it is sought to establish a contact between the reciprocity principle and unitary symmetry in connection with the problem of elementary particles.

NEARLY a decade and a half ago, Born and his collaborators[1] published a series of papers in which they attempted to explain elementary particles using the Born reciprocity principle (BRP) as a postulate and, more recently, we reported[2] a theory of quantization of elementary-particle masses using the same postulate applied to an "internal" space–time such as the one proposed by Yukawa.[3] In view of the recent apparently successful applications of the various unitary symmetry schemes among which the most noteworthy are the $SU(3)$ scheme of Gell-Mann and Ne'eman,[4] the $U(3)$ scheme of Okubo[5] and the more recent $SU(6)$ scheme by Pais and others,[6] it seems quite natural to ask whether the reciprocity postulate makes a contact with such a unitary symmetry scheme. In this paper, we show that, used as the sole postulate, the Born reciprocity principle gives unitary group $U(n)$ for the symmetry of elementary particles when the postulate is used in n-dimensional (isotropic) vector space C_n. This clearly differs from the previous attempts[1,2] which applied the postulate to four-dimensional Minkowski space instead of an isotropic n-space C_n which is adopted here.

Let the n-dimensional vector space C_n be responsible for all the "internal" physical observables such as mass, electrical charge, isospin, hypercharge and possibly the particle spins. For C_n,[7] let us define as the basic coordinates of C_n, n canonically conjugate pairs of dynamical variables $(x_\mu) = (x_1, x_2, \cdots, x_n)$ and $(p_\mu) = (p_1, p_2, \cdots, p_n)$ which satisfy the canonical commutation relations

$$[x_\mu, p^\nu] = +i\delta_\mu^\nu \qquad (\mu, \nu = 1, 2, 3, \cdots, n). \tag{1}$$

[Here, (x_μ) and (p_ν) are defined to be dimensionless.] In C_n, the dynamical observables may be represented as operators given in terms of the variables (x_μ) and (p_ν). As a postulate, we now subject the dynamics of C_n to BRP, which states that *the fundamental equations of physics are symmetric in the canonically conjugate pairs of variables.* Considering that the dynamics of C_n occupy the most fundamental realm of nature, the postulate asserts that BRP is most strongly realized in C_n in such a way that *all* "internal" dynamical observables maintain reciprocal symmetry between (x_μ) and (p_μ). This is to be contrasted with the more macroscopic realm of nature in which BRP is known to be obeyed by a few fundamental equations (e.g., commutation relations and Hamilton's principle) but not in all aspects of dynamics.

Mathematically, an operator is a "reciprocal operator" if it is invariant under the "reciprocity transformations" (RT) of Born.[1] According to Rayski,[8] the (2×2) matrices

$$\begin{pmatrix}1&0\\0&1\end{pmatrix};\quad \begin{pmatrix}0&-1\\1&0\end{pmatrix};\quad \begin{pmatrix}0&1\\-1&0\end{pmatrix};\quad \begin{pmatrix}-1&0\\0&-1\end{pmatrix} \tag{2}$$

form the common divisors for all such unimodular, unitary transformations which we denote as $R(2)$. In all, there are four such transformations in $R(2)$:

$$R_1 = I;\quad R_2 = \mathfrak{R};\quad R_3 = \mathfrak{C};\quad R_4 = \mathfrak{R} \times \mathfrak{C}, \tag{3}$$

where

$$I_{\alpha\beta} = \delta_{\alpha\beta}(\alpha, \beta = 1, 2),$$
$$\mathfrak{R}_{\alpha\beta} = -\delta_{\alpha\beta}, \tag{4}$$
$$\mathfrak{C} = \frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\\1&1\end{pmatrix}.$$

* Supported by the U. S. Air Force Office of Scientific Research.

[1] M. Born, Nature **163**, 207 (1949); M. Born and H. S. Green, Proc. Roy. Soc. Edinburgh **A92**, 470 (1949), and Nature **164**, 281 (1949); H. S. Green, Nature **163**, 208 (1949); M. Born, Rev. Mod. Phys. **21**, 463 (1949).

[2] E. E. H. Shin, Phys. Rev. Letters **10**, 196 (1963).

[3] H. Yukawa, Phys. Rev. **76**, 300, 1731 (1949); **77**, 219, 849 (1950).

[4] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).

[5] S. Okubo, Progr. Theoret. Phys. (Kyoto) **27**, 949 (1962).

[6] A. Pais, Phys. Rev. Letters **13**, 175 (1964); F. Gürsey, A. Pais, and L. Radicotti, Phys. Rev. Letters **13**, 299 (1964); F. Gürsey and L. Radicotti, Phys. Rev. Letters **13**, 173 (1964); T. K. Kuo and T. Yao, Phys. Rev. Letters **13**, 415 (1964); M. A. B. Bég and V. Singh, Phys. Rev. Letters **13**, 418 (1964).

[7] The question of whether this represents a "space–time" or a "field space" is irrelevant for the present purpose.

[8] J. Rayski, Nuovo Cimento **2**, 255 (1955).

Here, $\alpha, \beta = 1, 2$ refer, respectively, to (x_μ) and (p_μ) represented as two components of the same column (row) matrix such that

$$X_\mu = \begin{pmatrix} x_\mu \\ p_\mu \end{pmatrix},$$
$$(X_\mu)_1 = x_\mu; \quad (X_\mu)_2 = p_\mu, \tag{5}$$
$$(X'_\mu)_\alpha = \sum_{\beta=1}^{2} R_{\alpha\beta}(X_\mu)_\beta.$$

Note that these transformations preserve the vector character of the operators (x_μ) and (p_μ), and thus do not mix the indices, $\mu, \nu = 1, 2, 3, \cdots, n$. In terms of the RT of Eq. (3), we may now restate the postulate as follows: *In C_n, all dynamical operators in (x_μ, p_μ) which represent the "internal" physical system are invariant under the canonical reciprocity transformations of $R(2)$.*

This postulate then provides us with a mathemathical basis for the transition from a purely geometrical description of C_n to a physical description. Clearly, the most natural procedure for accomplishing this transition is to search for all possible operators [in terms of (x_μ) and (p_μ)] which are invariant under the RT of $R(2)$ since all "internal" physical observables may be constructed out of such operators. One need not go very far to discover that the number of such operators of the fundamental type is highly limited. The fundamental RT-invariant operators (bilinear in x and p) are given by:

$$\text{I.} \quad \delta_{\mu\nu} = i^{-1}[x_\mu, p_\nu], \tag{6}$$

$$\text{II.} \quad S_{\mu\nu} = S_{\nu\mu} = (x_\mu x_\nu + p_\mu p_\nu), \tag{7}$$

$$\text{III.} \quad T_{\mu\nu} = -T_{\nu\mu} = (x_\mu p_\nu - p_\mu x_\nu) \quad (\mu \neq \nu). \tag{8}$$

Aside from the very special operator $\delta_{\mu\nu}(= 1, \mu = \nu; = 0, \mu \neq \nu)$, there are in all n^2 number of operators which include $\frac{1}{2}n(n+1)$ number of symmetrical operators $S_{\mu\nu}$, and $\frac{1}{2}n(n-1)$ number of antisymmetrical operators $T_{\mu\nu}$. Further, the $\frac{1}{2}n(n+1)$ number of symmetrical operators contains n number of diagonal operators $(S_{11}, S_{22}, \cdots, S_{nn})$ and $\frac{1}{2}n(n-1)$ number of off-diagonal operators $S_{\mu\nu}(\mu \neq \nu)$. Note that the n diagonal operators of the form, $\frac{1}{2}S_{\mu\mu} = \frac{1}{2}(x_\mu^2 + p_\mu^2)$ are simply the n partitions of the Hamiltonian of an n-dimensional "harmonic oscillator." *All other RT-invariant operators can be constructed from these n^2 operators of the fundamental set given above.*

By virtue of the canonical commutation relation [Eq. (1)], the n^2 symmetrical and antisymmetrical operators satisfy the following commutation relations:

$$i^{-1}[S_{\mu\nu}, S_{\lambda\mu}] = \{\delta_{\mu\kappa}T_{\nu\lambda} + \delta_{\nu\kappa}T_{\mu\lambda} + \delta_{\nu\lambda}T_{\mu\kappa} + \delta_{\mu\lambda}T_{\nu\kappa}\},$$
$$[T_{\mu\nu}, T_{\lambda\kappa}] = -[S_{\mu\nu}, S_{\lambda\kappa}] \quad (\mu \neq \nu, \lambda \neq \kappa), \tag{9}$$
$$i^{-1}[T_{\mu\nu}, S_{\lambda\kappa}] = \{\delta_{\mu\lambda}S_{\nu\kappa} + \delta_{\mu\kappa}S_{\nu\lambda} - \delta_{\lambda\nu}S_{\mu\kappa} - \delta_{\nu\kappa}S_{\mu\lambda}\}.$$

These n^2 independent operators give the generators of unitary group $U(n)$, and give $(n^2 - 1)$ traceless operators which are the infinitesimal generators of unitary unimodular group $SU(n)$. Let us now use the Dirac ladder operators for the harmonic oscillator,

$$\eta^\mu = (1/\sqrt{2})(x_\mu + ip_\mu),$$
$$\bar{\eta}_\mu = (1/\sqrt{2})(x_\mu - ip_\mu) = (\eta^\mu)^\dagger, \tag{10}$$
$$[\eta^\mu, \bar{\eta}_\nu] = \delta^\mu_\nu \qquad (\mu, \nu = 1, 2, 3, \cdots, n),$$

and define n^2 mixed tensors,

$$A^\mu_\nu = (\eta^\mu \bar{\eta}_\nu), \tag{11}$$

which are given by linear combinations of the n^2 operators, $(S_{\mu\nu})$ and $(T_{\mu\nu})$, in the form

$$A^\mu_\nu = \tfrac{1}{2}\{S_{\mu\nu} - iT_{\mu\nu}\}. \tag{12}$$

These n^2 mixed tensors now satisfy the Lie equation,

$$[A^\mu_\nu, A^\lambda_\kappa] = \{\delta^\mu_\kappa A^\lambda_\nu - \delta^\lambda_\nu A^\mu_\kappa\}. \tag{13}$$

Since they are given by linear combinations of RT-invariant operators, all n^2 of (A^μ_ν) are also RT-invariant and therefore form a fundamental set of RT-invariant operators along with the very special tensors δ^μ_ν. All other RT-invariant operators can be constructed from (A^μ_ν): e.g.,

$$T^\mu_\nu = a\delta^\mu_\nu + bA^\mu_\nu + c(A \cdot A)^\mu_\nu \tag{14}$$

is RT-invariant.[9] Of the n^2 independent operators, only $(n^2 - 1)$ of them are traceless. These traceless operators are fundamental traceless RT-invariant operators and are also the infinitesimal generators of the unitary, unimodular group $SU(n)$.

Inasmuch as (x_μ) are the basic "coordinate" operators and (p_μ) the corresponding "momentum" operators of C_n, the n^2 operators (A^μ_ν) operate on (i) (x_μ), (p_μ); (ii) all operators constructed from (x_μ), (p_μ), or both; and (iii) functionals of (x_μ) or (p_μ). Let g and g^{-1} be the unitary transformation generated by one of the n^2 Hermitian operators, G. Operating on (η_μ) and $(\bar{\eta}_\mu)$, we have[10]

[9] The well-known Gell-Mann[4]–Okubo[5] mass formula is obtained by using T_3^2 as the symmetry-breaking term in $U(3)$ or $SU(3)$ scheme. Therefore, we may consider the Gell-Mann–Okubo mass formula in its operator form, to be RT-invariant within the framework of the present theory.

[10] G. A. Baker, Jr., Phys. Rev. 103, 1119 (1956).

$$\begin{aligned}
&\eta'_\mu = g_{\mu\nu}\eta_\nu, \\
&\bar{\eta}'_\mu = g^{-1}_{\mu\nu}\bar{\eta}_\nu, \\
&\eta'_\alpha\eta'_\beta \cdots \bar{\eta}'_i\bar{\eta}'_j \cdots \\
&\quad = g_{\alpha\mu}g_{\beta\nu} \cdots g^{-1}_{ii}g^{-1}_{jm}\eta_\mu\eta_\nu \cdots \bar{\eta}_l\bar{\eta}_m \cdots , \\
&\quad (\mu, \nu, \alpha, \beta, i, j, \cdots = 1, 2, 3, \cdots , n).
\end{aligned} \tag{15}$$

Associated with g and g^{-1}, we have the unitary operators α_g and $\alpha_g^{-1} = (\alpha_g)\dagger$ such that

$$\begin{aligned}
&\eta' = (g\eta) = \alpha_g\eta\alpha_g^{-1}, \\
&\delta\eta_\mu = (\eta'_\mu - \eta_\mu) \\
&\quad = \alpha_g\eta_\mu\alpha_g^{-1} - \eta_\mu \\
&\quad = i\delta c[G, \eta_\mu],
\end{aligned} \tag{16}$$

$$\begin{aligned}
&\alpha_g = (1 + i\delta cG), \\
&\alpha_g^{-1} = (1 - i\delta cG),
\end{aligned}$$

where δc is an arbitrary real infinitesimal parameter.

In short, it may be said that the BRP applied to C_n gives the symmetry of n-dimensional harmonic oscillator. Further, it is known[10] that the symmetry of n-dimensional isotropic harmonic oscillator is given by $U(n)$ or equivalently by $SU(n)$.

Finally, the author thanks Benjamin Lax, Arthur J. Freeman, Norman Horing, and Joshua Zak for their cooperation, profitable discussions and for their patience, and particularly to Professor Joshua Zak whose comments have been essential in making this paper possible. He is especially indebted to Professor Max Born for his continued encouragement. He is infinitely indebted to his wife, Shin-Ai.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Geometric Theory of Neutrinos

R. PENNEY
Ford Motor Company, Scientific Laboratory, Dearborn, Michigan
(Received 11 September 1964: final manuscript received 22 February 1965)

It is shown that the conditions $R_\mu{}^\mu = 0$, $R_{00} \geq 0$, $R_\alpha{}^\beta R_\beta{}^\lambda = \frac{1}{4}R_{\rho\tau}R^{\rho\tau}\delta_\alpha{}^\lambda$, $R_{\alpha\beta|\sigma} = (R_{\rho\tau}R^{\rho\tau})_{|\sigma}R_{\alpha\beta}$, in the limit that the scalar $R_{\alpha\beta}R^{\alpha\beta}$ vanishes, reproduce the physics of neutrinos. That is, these conditions ensure the existence of a two-component spinor which obeys the Weyl equation and represents a field of completely polarized neutrinos.

I. INTRODUCTION

THERE have been various attempts in the past to construct a neutrino theory of light,[1] yet none is generally accepted at the present date. Similarly, there have been attempts to geometrize neutrinos.[2] Wheeler[3] has drawn attention to the outstanding problem of geometrodynamics as being the need for showing that spin-$\frac{1}{2}$ is "contained" in that theory. The success of the Rainich "already-unified" field theory[4] leads one to suspect that the three fields of physics, i.e., gravitation, electromagnetism, and neutrinos, should all be unified. The problem of particles may pose great difficulties, but one should expect from the philosophy of geometrodynamics that such a simple field as the massless neutrino should be capable of geometrization. We intend to show that this can be done in a fairly general way.

Before proceeding, we should remark that our notation will be the same as that of Wheeler[3] in his book. Further, we wish to point out that the essential steps which enable us to geometrize the neutrino field were first studied by Whittaker[5] in 1936, and his paper will replace a detailed set of references. All of our manipulations can be understood by referring to Whittaker or Wheeler, and a detailed bibliography is given by Wheeler.

II. THE RAINICH THEORY

In 1925, Rainich[4] showed that there were a set of necessary and sufficient conditions that a Rie-

[1] For example, P. Jordan and R. de L. Kronig, Z. Phys. **100**, 569 (1936); M. H. L. Pryce, Proc. Roy. Soc. (London) **A165**, 247 (1938).
[2] For example, A. Inomata, Bull. Am. Phys. Soc. **9**, 86 (1964).
[3] J. A. Wheeler, *Geometrodynamics* (Academic Press, Inc., New York, 1962), p. 1.
[4] G. Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[5] E. T. Whittaker, Proc. Roy. Soc. (London) **A158**, 38 (1937).

$$\begin{aligned}
&\eta'_\mu = g_{\mu\nu}\eta_\nu, \\
&\bar{\eta}'_\mu = g^{-1}_{\mu\nu}\bar{\eta}_\nu, \\
&\eta'_\alpha\eta'_\beta \cdots \bar{\eta}'_i\bar{\eta}'_j \cdots \\
&\quad = g_{\alpha\mu}g_{\beta\nu} \cdots g^{-1}_{ii}g^{-1}_{jm}\eta_\mu\eta_\nu \cdots \bar{\eta}_l\bar{\eta}_m \cdots , \\
&\quad (\mu, \nu, \alpha, \beta, i, j, \cdots = 1, 2, 3, \cdots , n).
\end{aligned} \tag{15}$$

Associated with g and g^{-1}, we have the unitary operators α_g and $\alpha_g^{-1} = (\alpha_g)\dagger$ such that

$$\begin{aligned}
&\eta' = (g\eta) = \alpha_g\eta\alpha_g^{-1}, \\
&\delta\eta_\mu = (\eta'_\mu - \eta_\mu) \\
&\quad = \alpha_g\eta_\mu\alpha_g^{-1} - \eta_\mu \\
&\quad = i\delta c[G, \eta_\mu],
\end{aligned} \tag{16}$$

$$\begin{aligned}
&\alpha_g = (1 + i\delta cG), \\
&\alpha_g^{-1} = (1 - i\delta cG),
\end{aligned}$$

where δc is an arbitrary real infinitesimal parameter.

In short, it may be said that the BRP applied to C_n gives the symmetry of n-dimensional harmonic oscillator. Further, it is known[10] that the symmetry of n-dimensional isotropic harmonic oscillator is given by $U(n)$ or equivalently by $SU(n)$.

Finally, the author thanks Benjamin Lax, Arthur J. Freeman, Norman Horing, and Joshua Zak for their cooperation, profitable discussions and for their patience, and particularly to Professor Joshua Zak whose comments have been essential in making this paper possible. He is especially indebted to Professor Max Born for his continued encouragement. He is infinitely indebted to his wife, Shin-Ai.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Geometric Theory of Neutrinos

R. PENNEY
Ford Motor Company, Scientific Laboratory, Dearborn, Michigan
(Received 11 September 1964: final manuscript received 22 February 1965)

It is shown that the conditions $R_\mu{}^\mu = 0$, $R_{00} \geq 0$, $R_\alpha{}^\beta R_\beta{}^\lambda = \frac{1}{4}R_{\rho\tau}R^{\rho\tau}\delta_\alpha{}^\lambda$, $R_{\alpha\beta|\sigma} = (R_{\rho\tau}R^{\rho\tau})_{|\sigma}R_{\alpha\beta}$, in the limit that the scalar $R_{\alpha\beta}R^{\alpha\beta}$ vanishes, reproduce the physics of neutrinos. That is, these conditions ensure the existence of a two-component spinor which obeys the Weyl equation and represents a field of completely polarized neutrinos.

I. INTRODUCTION

THERE have been various attempts in the past to construct a neutrino theory of light,[1] yet none is generally accepted at the present date. Similarly, there have been attempts to geometrize neutrinos.[2] Wheeler[3] has drawn attention to the outstanding problem of geometrodynamics as being the need for showing that spin-$\frac{1}{2}$ is "contained" in that theory. The success of the Rainich "already-unified" field theory[4] leads one to suspect that the three fields of physics, i.e., gravitation, electromagnetism, and neutrinos, should all be unified. The problem of particles may pose great difficulties, but one should expect from the philosophy of geometrodynamics that such a simple field as the massless neutrino should be capable of geometrization. We intend to show that this can be done in a fairly general way.

Before proceeding, we should remark that our notation will be the same as that of Wheeler[3] in his book. Further, we wish to point out that the essential steps which enable us to geometrize the neutrino field were first studied by Whittaker[5] in 1936, and his paper will replace a detailed set of references. All of our manipulations can be understood by referring to Whittaker or Wheeler, and a detailed bibliography is given by Wheeler.

II. THE RAINICH THEORY

In 1925, Rainich[4] showed that there were a set of necessary and sufficient conditions that a Rie-

[1] For example, P. Jordan and R. de L. Kronig, Z. Phys. **100**, 569 (1936); M. H. L. Pryce, Proc. Roy. Soc. (London) **A165**, 247 (1938).
[2] For example, A. Inomata, Bull. Am. Phys. Soc. **9**, 86 (1964).
[3] J. A. Wheeler, *Geometrodynamics* (Academic Press, Inc., New York, 1962), p. 1.
[4] G. Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[5] E. T. Whittaker, Proc. Roy. Soc. (London) **A158**, 38 (1937).

mannian geometry should reproduce the physics of the Maxwell field without sources. These conditions are, explicitly in terms of the Ricci tensor $R_{\mu\nu}$,

$$R_\mu^\mu = 0, \tag{1}$$

$$R_\alpha^\beta R_\beta^\gamma = \tfrac{1}{4} R_{\rho\tau} R^{\rho\tau} g_\alpha^\gamma, \tag{2}$$

$$R_{00} \geq 0, \tag{3}$$

$$\alpha_{\beta|\gamma} - \alpha_{\gamma|\beta} = 0, \tag{4}$$

where the bar denotes covariant differentiation, and the geometric vector α_β is defined by

$$(R_{\rho\tau} R^{\rho\tau})\alpha_\beta \equiv (-g)^{\frac{1}{2}} \epsilon_{\beta\alpha\mu\nu} R^{\alpha\gamma|\mu} R_\gamma^\nu. \tag{5}$$

In the special case of a null Maxwell field, i.e., one for which

$$F_{\mu\nu} F^{\mu\nu} = 0, \qquad F_{\mu\nu} {}^*F^{\mu\nu} = 0, \tag{6}$$

where ${}^*F_{\mu\nu}$ is the dual tensor defined by

$${}^*F_{\mu\nu} \equiv \tfrac{1}{2}(-g)^{\frac{1}{2}} \epsilon_{\mu\nu\alpha\beta} F^{\alpha\beta}, \tag{7}$$

the Rainich conditions fail—essentially because α_β is not defined.

It is just this special case which permits a geometric theory of neutrinos, as we shall see presently.

III. NULL ELECTROMAGNETIC FIELD

For the moment, we consider flat space–time in which we have a null Maxwell field:

$$F_{\mu\nu} F^{\mu\nu} = 0, \qquad F_{\mu\nu} {}^*F^{\mu\nu} = 0, \tag{8}$$

$$F^{\mu\nu}{}_{|\nu} = 0, \qquad {}^*F^{\mu\nu}{}_{|\nu} = 0. \tag{9}$$

Following Whittaker,[5] we can easily show that such a field behaves as an elementary spinor. To see this, note that $F_{\mu\nu}$ has but four independent real components. If we form

$$S_{\mu\nu} \equiv \tfrac{1}{2}(F_{\mu\nu} + i {}^*F_{\mu\nu}), \tag{10}$$

$$A_{\mu\nu} \equiv \tfrac{1}{2}(F_{\mu\nu} - i {}^*F_{\mu\nu}), \tag{11}$$

we see that

$$F_{\mu\nu} = S_{\mu\nu} + A_{\mu\nu}, \tag{12}$$

$$S_{\mu\nu}^+ = A_{\mu\nu}, \tag{13}$$

where "+" means the complex conjugate. The tensor $S_{\mu\nu}$ has *two* independent complex components. Using Whittaker's results, we have the correspondence

$$S_{01} = \varphi_1^2 - \varphi_2^2, \qquad S_{23} = -i(\varphi_1^2 - \varphi_2^2), \tag{14}$$

$$S_{02} = -i(\varphi_1^2 + \varphi_2^2), \qquad S_{31} = -(\varphi_1^2 + \varphi_2^2), \tag{15}$$

$$S_{03} = -2\varphi_1\varphi_2, \qquad S_{12} = 2i\varphi_1\varphi_2 \tag{16}$$

between $S_{\mu\nu}$ and a spinor $(\varphi_1\varphi_2)$. That is, if the spinor transforms as

$$\varphi_1 \to \varphi_1^1 = \alpha\varphi_1 + \beta\varphi_2, \tag{17}$$

$$\varphi_2 \to \varphi_2^1 = \gamma\varphi_1 + \delta\varphi_2, \tag{18}$$

$$\alpha\delta - \beta\partial = 1 \tag{19}$$

under Lorentz transformation, then $S_{\mu\nu}$ transforms as a tensor, and conversely.

The decomposition of $F_{\mu\nu}$ into $S_{\mu\nu}$ and $A_{\mu\mu}$ is of course covariant for proper transformations, so we have a unique covariant prescription for $(\varphi_1\varphi_2)$.

We will in the sequel use two other results due to Whittaker, which we merely state here. If we consider

$$S_{\mu\nu} S^{+\rho\nu} = A^{\rho\nu} S_{\mu\nu}, \tag{20}$$

we find that

$$A^{\rho\nu} S_{\mu\nu} = -2D^\rho D_\mu \tag{21}$$

where D^μ is a vector with components

$$D^0 = |\varphi_1|^2 + |\varphi_2|^2, \qquad D^2 = i(\varphi_1\varphi_2^* - \varphi_1^*\varphi_2), \tag{22}$$

$$D' = \varphi_1\varphi_2^* + \varphi_1^*\varphi_2, \qquad D^3 = |\varphi_1|^2 - |\varphi_2|^2 \tag{23}$$

and is a null vector:

$$D_\mu D^\mu = 0. \tag{24}$$

Secondly, let us define the vector

$$H_\mu \equiv \varphi_2 \partial\varphi_1/\partial x^\mu - \varphi_1 \partial\varphi_2/\partial x^\mu, \tag{25}$$

and we then have that

$$S_{\mu\nu} D^{\mu|\rho} = 2D_\nu H^\rho, \tag{26}$$

as is verifiable by direct calculation.

IV. SUFFICIENT CONDITIONS FOR A NEUTRINO FIELD

Let us now use the relations of the previous sections to geometrize the neutrino field. Suppose for this purpose that we have a null Rainich–Riemann geometry. That is, we have

$$R = 0, \tag{27}$$

$$R_\alpha^\beta R_\beta^\gamma = 0, \tag{28}$$

$$R_{00} \geq 0. \tag{29}$$

Then as Wheeler has shown,[6] the Ricci tensor may be expressed as

$$R_{\mu\nu} = 2f_{\mu\alpha} f_\nu^\alpha = 2k_\mu k_\nu, \tag{30}$$

where k_μ is a null vector and $f_{\mu\nu}$ is a null tensor. The

[6] J. A. Wheeler, Ref. 3, p. 247.

tensor $f_{\mu\nu}$ may be expressed in terms of k_μ and a unit vector v_μ by

$$f_{\mu\nu} = k_\mu v_\nu - k_\nu v_\mu, \tag{31}$$

with

$$v^\mu k_\mu = 0, \qquad v^\mu v_\mu = 1. \tag{32}$$

The tensor $f_{\mu\nu}$ and the vector v_μ are not uniquely determined by Maxwell's equations. For example

$$v_\mu \rightarrow v'_\mu = v_\mu + \lambda k_\mu \tag{33}$$

leaves $f_{\mu\nu}$ unchanged in its properties. Also Maxwell's equations do not generally fix $f_{\mu\nu}$.

Nonetheless, one can express $R_{\mu\nu}$ as above, even if not uniquely so, and that is what we need here.

Now let us *suppose* that a particular $f_{\mu\nu}$ has been picked to obey Maxwell's equations, and a particular v_ν is chosen. Then $f_{\mu\nu}$ can be decomposed as in the previous section and is null, so we have a spinor corresponding to $f_{\mu\nu}$. Since $f_{\mu\nu}$ obeys Maxwell's equations, the spinor obeys certain differential equations.

With $f_{\mu\nu}$ expressed in terms of $(\varphi_1\varphi_2)$, we see that k_μ is simply proportional to D_μ, the vector previously introduced. Now suppose that k_μ does not vanish. Then consider the conditions

$$f_{\mu\nu}k^{\mu|\rho} = 0, \qquad {}^*f_{\mu\nu}k^{\mu|\rho} = 0. \tag{34}$$

If these conditions should be true, then $(\varphi_1\varphi_2)$ obeys the Weyl equation, as we now show. To see this, we note that the conditions imply

$$S_{\mu\nu}D^{\mu|\rho} = 0 \tag{35}$$

and therefore

$$H^\rho = 0. \tag{36}$$

Knowing this latter fact, examine the vector

$$Q_\mu \equiv \tfrac{1}{4}f_{\mu\nu}{}^{|\nu} + \tfrac{1}{4}i{}^*f_{\mu\nu}{}^{|\nu} + H_\mu \tag{37}$$

which vanishes. Explicitly, we have, e.g.,

$$Q_0 = -\varphi_1\left(\frac{\partial\varphi_1}{\partial x^1} + i\frac{\partial\varphi_1}{\partial x^2} - \frac{\partial\varphi_2}{\partial x^3} + \frac{\partial\varphi_2}{\partial x^0}\right) + \varphi_2\left(\frac{\partial\varphi_2}{\partial x^1} - i\frac{\partial\varphi_2}{\partial x^2} + \frac{\partial\varphi_2}{\partial x^3} + \frac{\partial\varphi_1}{\partial x^0}\right), \tag{38}$$

where the quantities in parentheses occur in each component. The vanishing of Q_μ therefore gives

$$\frac{\partial\varphi_1}{\partial x^1} + i\frac{\partial\varphi_1}{\partial x^2} - \frac{\partial\varphi_2}{\partial x^3} + \frac{\partial\varphi_2}{\partial x^0} = 0, \tag{39}$$

$$\frac{\partial\varphi_2}{\partial x^1} - i\frac{\partial\varphi_2}{\partial x^2} + \frac{\partial\varphi_1}{\partial x^3} + \frac{\partial\varphi_1}{\partial x^0} = 0, \tag{40}$$

unless *both* of the following are true:

$$\varphi_1^2 - \varphi_2^2 = 0, \tag{41}$$

$$2\varphi_1\varphi_2 = 0, \tag{42}$$

which is not possible unless $f_{\mu\nu}$ itself vanishes. Thus, Maxwell's equations plus the vanishing of H_ρ ensure that $(\varphi_1\varphi_2)$ obeys the Weyl equation. We will later interpret the vanishing of H_ρ, but here we only remark that it is a sufficient condition for a neutrino field.

V. GEOMATIC CONDITIONS FOR NEUTRINOS

We wish now to consider the purely geometric conditions

$$R^\mu_\mu = 0, \tag{1}$$

$$R^\beta_\alpha R^\lambda_\beta = \tfrac{1}{4}R_{\rho\tau}R^{\rho\tau}\delta^\lambda_\alpha, \tag{2}$$

$$R_{00} \geq 0, \tag{3}$$

$$R_{\alpha\beta|\sigma} = (R_{\rho\tau}R^{\rho\tau})_{|\sigma}R_{\alpha\beta}, \tag{43}$$

and show that such a geometry, or rather a subfield of such a geometry, reproduces neutrino physics.

To see this, note that the algebraic conditions are the usual ones for a Rainich geometry. The differential conditions ensure that the gradient of the complexion vanishes identically, *independently of the value of* $R_{\alpha\beta}R^{\alpha\beta}$,

$$\alpha_\beta = 0. \tag{44}$$

Thus the complexion of the field, defined by

$$\alpha = \int_0^x \alpha_\beta\, dx^\beta, \tag{45}$$

is constant everywhere. Even when we go to the limit

$$I \equiv R_{\alpha\beta}R^{\alpha\beta} \rightarrow 0 \tag{46}$$

we retain this constant complexion.

Our conditions determine a Maxwell field, just as in the usual Rainich theory, albeit this "Maxwell field" is of a peculiar kind obeying a set of nonlinear differential equations. In the limit of vanishing I, we have

$$R_{\mu\nu} = 2f_{\mu\alpha}f^\alpha_\nu = 2k_\mu k_\nu, \tag{30}$$

where $f_{\mu\nu}$ obeys Maxwell's equations.

We then consider the tensor $Q_{\alpha\rho}$ found as follows:

$$Q_{\alpha\rho} \equiv f_{\mu\alpha}k^\mu{}_{|\rho}. \tag{47}$$

Using the relations

$$f_{\mu\nu} = k_\mu v_\nu - k_\nu v_\mu, \tag{31}$$

$$v_\mu v^\mu = 1, \quad v_\mu k^\mu = 0, \quad k_\mu k^\mu = 0, \tag{48}$$

$$f_{\mu\nu}k^\nu = 0, \quad f_{\mu\nu}v^\nu = k_\mu, \tag{49}$$

we readily obtain

$$Q_{\alpha\rho} = (f_{\mu\alpha}k^\mu)_{|\rho} - f_{\mu\alpha|\rho}k^\mu = -k^\mu f_{\mu\alpha|\rho}, \tag{50}$$

$$Q_{\alpha\rho} = -f^{\mu\lambda}v_\lambda f_{\mu\alpha|\rho}, \tag{51}$$

$$Q_{\alpha\rho} = \tfrac{1}{2}R^\lambda_{\alpha|\rho}v_\lambda - v_\lambda f_{\mu\alpha}f^{\mu\lambda}{}_{|\rho} = 0. \tag{52}$$

The first term vanishes by assumption, and the second term by symmetry as we see by writing $f_{\mu\nu}$ in terms of k_μ, v_ν again. Similarly, if we consider

$$P_{\alpha\rho} \equiv {}^*f_{\mu\alpha}k^\mu{}_{|\rho}, \tag{53}$$

we see that it, too, vanishes. Thus the condition

$$R_{\mu\beta|\rho} = 0, \tag{54}$$

which obtains in the limit of vanishing I, gives

$$f_{\mu\nu}k^\mu{}_{|\rho} = 0, \quad {}^*f_{\mu\nu}k^\mu{}_{|\rho} = 0, \tag{55}$$

and so ensures that H_ρ should vanish. In turn, the vanishing of H_ρ leads to the Weyl equation for $(\varphi_1\varphi_2)$, of course.

Our purely geometric conditions allow us to determine a geometry in which there is a "Maxwell field". Having obtained this geometry, with the complexion fixed, we may proceed to the limit of a *null* field in which neutrino physics is reproduced.

VI. DUALITY ROTATIONS

As is well known, the Maxwell energy tensor, which is essentially

$$f_{\mu\alpha}f^\alpha_\nu - \tfrac{1}{4}f_{\rho\sigma}f^{\rho\sigma}g_{\mu\nu} \tag{56}$$

is invariant under the transformation

$$f_{\mu\nu} \to f'_{\mu\nu} = f_{\mu\nu}\cos\alpha + {}^*f_{\mu\nu}\sin\alpha, \tag{57}$$

which is a "duality rotation."[6]

Similarly, if we consider the Weyl equation, or, equivalently, the Lee–Yang neutrino equations[7]

$$\gamma_\mu\partial_\mu\psi = 0, \tag{58}$$

$$\psi = \gamma_5\psi, \tag{59}$$

we have invariance for "chirality rotations,"

$$\psi \to \psi' = e^{i\theta\gamma_5}\psi. \tag{60}$$

Our present theory shows that the two forms of "rotation" are equivalent. That is, a chirality rotation of ψ (which could just as easily be written in terms of φ_1, φ_2) is a duality rotation of $f_{\mu\nu}$. To see this, note

$$R_{\mu\nu} = 2f_{\mu\alpha}f^\alpha_\nu = 2k_\mu k_\nu, \tag{61}$$

where k_μ may be written as

$$k_\mu = i\bar{\psi}\gamma_\mu\psi, \tag{62}$$

with

$$\psi = \gamma_5\psi. \tag{63}$$

Obviously, k_μ is invariant under chirality rotations. $f_{\mu\nu}$, on the other hand, can be expressed as a linear combination of

$$i\bar{\psi}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\psi^c, \tag{64}$$

$$i\bar{\psi}^c(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\psi, \tag{65}$$

where ψ^c is the charge-conjugate spinor. When ψ is subjected to a chirality rotation, $f_{\mu\nu}$ undergoes a duality rotation, as is easily seen.

Thus, Wheeler's speculation on the connection between these two forms of rotation turns out to have significance.[3]

VII. THE MEANING OF $H_\rho = 0$

We have shown how to specify a Riemannian geometry to obtain a neutrino field, but this field is of a particular kind. Our neutrinos obey

$$\varphi_1\partial_\mu\varphi_2 - \varphi_2\partial_\mu\varphi_1 = 0 \tag{66}$$

in a suitable Lorentz frame. To interpret this condition, note that it is satisfied only by

$$\varphi_2 = c_0\varphi_1, \tag{67}$$

where c_0 is coordinate independent.

The general solution of Weyl's equation we may write as[7]

$$\varphi = \frac{1}{V^{\frac{1}{2}}}\sum_k a_k(t)U(\bar{k})e^{i\mathbf{k}\cdot\mathbf{r}} + b^*_k(t)V(+\bar{k})e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{68}$$

where

$$[-E + hc\sigma_i k_i]U(\bar{k}) = 0, \tag{69}$$

$$[+E + hc\sigma_i k_i]V(-\bar{k}) = 0 \tag{70}$$

are particular solutions.

Taking any particular representation of the Pauli matrices, we see that we must have two of the components k_i vanishing and one of the sets of coefficients a_k, b^*_k vanishing.

Thus we conclude that our condition implies the neutrino field is of a particular polarization type, traveling in a particular direction. For example, a particular φ satisfying our condition is

$$\varphi = \frac{1}{V^{\frac{1}{2}}}\sum_k a_k(t)U(k)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{71}$$

[7] P. Roman, *Theory of Elementary Particles* (North-Holland Publishing Company, Amsterdam, 1961), 2nd ed., p. 377.

with

$$E = hck_3, \qquad \bar{\mathbf{k}} = (00k_3). \tag{72}$$

Using this, we may readily calculate the corresponding Maxwell tensor, and see that it is the field of circularly polarized light, as might well be expected.

VIII. EXISTENCE OF THE LIMITING FIELD

We have shown that, in the limit of vanishing $R_{\alpha\beta}R^{\alpha\beta}$, we obtain our desired aim, yet the question remains whether such geometries may exist. That is, when $R_{\alpha\beta}R^{\alpha\beta}$ vanishes, we have

$$R = 0, \tag{27}$$

$$R_\alpha^\beta R_\beta^\lambda = 0, \tag{28}$$

$$R_{00} \geq 0, \tag{29}$$

$$R_{\alpha\beta|\sigma} = 0, \tag{54}$$

and one may well question whether our conditions lead to the identical vanishing of $R_{\mu\mu}$ itself.[8]

Consider, for example, the identity

$$R_{\alpha\beta|\sigma\tau} - R_{\alpha\beta|\tau\sigma} \equiv R_{\alpha\alpha\sigma\tau}R_\beta^\alpha + R_{\alpha\beta\sigma\tau}R_\alpha^\alpha, \tag{73}$$

which, for our geometry, implies sixty algebraic conditions on $R_{\mu\nu}$. It might well be expected that $R_{\mu\nu}$ must therefore vanish, but such is not the case.

To illustrate that our limiting conditions do not lead to the vanishing of $R_{\mu\nu}$, we construct an example of a Riemannian geometry in which our conditions are true, but $R_{\mu\nu}$ is nontrivial.

We suppose a Riemannian geometry which is conformally flat,[9] such that the line element is

$$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu = e^\psi \eta_{\mu\nu}\,dx^\mu\,dx^\nu, \tag{74}$$

where the $\eta_{\mu\nu}$ is the metric of flat space–time. The Christoffel symbols in our space are

$$\Gamma_{ik}^i = \gamma_{ik}^i + A_{ik}^i, \tag{75}$$

where the γ_{jk}^i are the flat-space symbols, and

$$2A_{jk}^i \equiv \delta_j^i\psi_{,k} + \delta_k^i\psi_{,j} - g^{ia}g_{jk}\psi_{|a}. \tag{76}$$

We will denote covariant derivatives (with respect to $g_{\mu\nu}$) by a slash between indices, while ordinary derivatives are implied by a comma.

The Ricci tensor for our space is easily calculated to be

$$R_{jk} = \psi_{|jk} + \tfrac{1}{2}\psi_{|j}\psi_{|k} - \tfrac{1}{2}g_{jk}[\psi_{|a}\psi^{|a} - \Box\psi], \tag{77}$$

where

$$\Box\psi \equiv g^{ab}\psi_{|ab}. \tag{78}$$

ψ is an arbitrary scalar, of course, and we may restrict it by covariant equations to attain our aim. We choose the conditions

$$\psi_{|jk} = 0, \tag{79}$$

$$\psi_{|a}\psi^{|a} = 0. \tag{80}$$

With our chosen conditions we have

$$R_{jk} = \tfrac{1}{2}\psi_j\psi_k, \tag{81}$$

$$R_{jk|l} = 0, \tag{82}$$

which fulfills our aim. We must then show that a suitable ψ may be chosen.

If we use Cartesian coordinates, our demands upon ψ are expressed by

$$\psi_{|jk} = \psi_{,jk} - \psi_{,j}\psi_{,k} + \tfrac{1}{2}g_{jk}\psi_{|b}\psi^{|b} = 0, \tag{83}$$

$$\psi_{|b}\psi^{|b} = 0, \tag{80}$$

It is easy to see that we may take

$$e^{-\psi} = a_ix^i = a_1x^1 + a_2x^2 + a_3x^3 + a_0x^0, \tag{84}$$

$$a_ia^i = a_1^2 + a_2^2 + a_3^2 - a_0^2 = 0, \tag{85}$$

and thereby have indeed a nontrivial geometry in which $R_{\mu\nu}$ is covariant constant as desired.

Perhaps we should emphasize that the example we have treated, which is a *particular* solution of our null field equations, was obtained by restricting the null vector k_μ to obey

$$k_{\mu|\nu} - k_{\nu|\mu} = 0, \tag{86}$$

which was not contained in our original conditions. It seems reasonable to assume that there may exist more general solutions of our null-field conditions, without the restriction mentioned.

At the least, we have shown by explicit counterexample that $R_{\mu\nu}$ does *not* vanish in the limit of vanishing $R_{\alpha\beta}R^{\alpha\beta}$. Our counterexample may be worthy of further study in its own right.

IX. THE LIMITING PROCESS

We have, in the preceding discussion, used the concept of a limiting geometry. That is, we have considered the possibility of constructing a Riemannian geometry corresponding to a nonnull electromagnetic field, and then passing to the limit of a null electromagnetic field.

We have not, however, detailed how one might operationally perform the limiting process in calculation. At the present juncture, we shall not

[8] The author is indebted to an anonymous referee for suggestions which led to the ensuing discussion.

[9] J. L. Synge, *Relativity: The General Theory* (North-Holland Publishing Company, Amsterdam, 1960), p. 317.

attempt to formulate the limiting process explicitly. We are only certain that such a limiting process is possible. That is, we know that Riemannian geometries exist wherein our conditions are fulfilled, whether $R_{\alpha\beta}R^{\alpha\beta}$ vanishes or not. It is possible, but highly improbable, that our limiting field cannot be reached in a continuous manner from the nonnull field.

We must admit our ignorance of the details of our limiting process, and hope that further work may shed some light on this question. On physical grounds, it would seem improbable that there should not be a continuous transition from the nonnull case to the null case.

X. INTERPRETATION OF THE RESULTS

In the preceding, we have referred to our limiting field as being appropriate for neutrino physics. Alternatively, one might interpret the formalism as a spinorial description of null electromagnetic fields.[8]

Since our calculations have all been *classical* in scope, the interpretation is not really unique. For example, there is nothing in our formalism to indicate that Fermi statistics is appropriate for our limiting field field. Thus, our use of the name "neutrino" might appear inadvisable.

However, until we fully understand the mechanics of quantizing gravitational fields, we are forced to content our selves with calculations which are classical. The "classical neutrino" is that field which obeys Weyl's equation, whence our chosen appellation.

The resolution of our two apparently conflicting interpretations is less difficult when one realizes that our theory is entirely classical. In the classical sense, we have obtained a theory of null electromagnetic fields, or of a neutrino field. The two descriptions are equivalent until quantization is performed. We are in possession only of field properties, not particle properties.

In a sense, our theory is a classical "neutrino theory of light." Certainly, we have shown how a field which is a representation of the Lorentz group corresponding to spin $\frac{1}{2}$ is contained in Riemannian geometry, and that was our aim.

Eventually, it is to be hoped, we shall understand how some feature of curved space–time, perhaps topology, may decide whether Fermi statistics for a field of neutrinos or Bose statistics for a field of coupled neutrinos is obtained.

XI. CONCLUSIONS

We have shown that a particular kind of Riemannian geometry reproduces the physics of a neutrino field.

To repeat the prescription, we specify our geometry by the conditions

$$R^{\mu}_{\mu} = 0, \tag{1}$$

$$R^{\beta}_{\alpha}R^{\lambda}_{\beta} = \tfrac{1}{4}R_{\rho\tau}R^{\rho\tau}\delta^{\lambda}_{\alpha}, \tag{2}$$

$$R_{00} \geq 0, \tag{3}$$

$$R_{\alpha\beta|\sigma} = (R_{\rho\tau}R^{\rho\tau})_{|\sigma}R_{\alpha\beta}. \tag{43}$$

The Ricci tensor for such a geometry is expressible in terms of an antisymmetric tensor $f_{\mu\nu}$ obeying Maxwell's equations. $f_{\mu\nu}$ is uniquely determined up to a constant duality rotation since the complexion gradient vanishes identically.

We then may go to the limit

$$R_{\rho\tau}R^{\rho\tau} \to 0, \tag{73}$$

which maintains the constant complexion, and retains the validity of Maxwell's equations for the tensor $f_{\mu\nu}$, which is now *null.*

$f_{\mu\nu}$ may be expressed (covariantly in a locally Lorentz frame) in terms of a spinor which thereby obeys the Weyl equation. This spinor field is a completely polarized field of neutrinos, either "particles" *or* "antiparticles."

By a constant chirality rotation, which is still allowable, the complexion of this neutrino field may be altered, and *in this sense* we may say we are describing a general neutrino field.

We have accomplished the dual aim of our program, namely to understand the null Rainich field and to geometrize the neutrino field. Perhaps further work may indicate the connection between massive particles and the Riemann geometry. At any rate, that must be the purpose of such calculations as the present one.

ACKNOWLEDGMENT

The author wishes to extend his appreciation to Professor G. Y. Rainich for the benefit of several stimulating discussions of the general philosophy of geometrization of physics.

General Operator Potential for the Two-Nucleon System

J. THARRATS, O. CERCEAU, AND O. ROJO*
Escuela de Fisica y Matematicas, Universidad Central, Caracas, Venezuela
(Received 23 November 1964)

The interaction between two nucleons is assumed to be described by a general Hermitian operator having certain symmetry and invariance properties.

It is shown that the usual phenomenological potentials employed in dealing with the two-nucleon system, i.e., nonlocal, velocity-dependent and hard-core, are related by means of unitary transformations of the interaction operator. In addition it is shown that another such transformation of the total Hamiltonian operator leads to the effective mass formalism. Some numerical results on phase shifts are compared with most accepted values.

I. INTRODUCTION

THE two-body problem in nuclear physics is still far from being solved. In the past, several phenomenological potentials, and in some cases an effective mass, have been proposed that reproduce some of the known features of the system.[1-4] In particular, phase shifts given by these potentials have been compared with the experimental ones fitted by Breit *et al.*[5] Also, some attempts to find a relationship among the phenomenological potentials have been made.[6-10]

Our object is to show that all of the potentials (as well as the effective mass formalism) are related by means of unitary transformations of the Schrödinger equation.

We shall start with the Schrödinger equation in Dirac's notation,

$$(T + \mathcal{U})|\rangle = E|\rangle, \tag{1}$$

in which $\mathcal{U}$ is a nonlocal Dirac operator (that is, in r-basis, $\langle \mathbf{r}|\ \mathcal{U}\ |\mathbf{r}'\rangle$ is nondiagonal). It will be shown that this may be transformed into any of the following equivalent forms:

$$(T + V_{\text{local}})\ |\text{shifted}\rangle = (T + V_{\text{hard-core}})|\rangle = E|\rangle, \tag{2}$$

$$(T + V_{\text{vel-depend}})|\rangle = E|\rangle, \tag{3}$$

$$[T(M^*) + V_{\text{local}}]|\rangle = E|\rangle, \tag{4}$$

where M^* is the effective mass.

In this way it will be demonstrated that velocity-dependent, nonlocal, and hard-core potentials, as well as the effective mass formalism, are related by an unitary transformation. Thus, all of the above equations have the same physical content, and which form is chosen is essentially a matter of convenience.

II. OPERATOR FORM OF $\mathcal{U}$

In this section we shall introduce $\mathcal{U}$ as a general nonlocal operator. Then it will be shown by means of a unitary transformation that a phase shift is produced in the asymptotic wavefunction, which is equivalent to the one produced by a hard-core potential.

We regard the potential as a Hermitian operator $\mathcal{U}$, with symmetry properties which will be determined in Appendix I.[11]

Let us assume that $\mathcal{U}$ is an operator that is diagonal in the basis μ. Suppose that A is a Hermitian operator so that e^{-iA} is unitary. Assume that A is such that

$$e^{-iA}\mathcal{U}e^{iA} = V(\mathbf{r}) \tag{5}$$

is a transformation that yields an operator diagonal in r-basis. Here we need to restrict $\mathcal{U}$ to have only a continuous spectrum. In general, the Schrödinger equation (1) reads (after solving for $\mathcal{U}$ in (5) and substituting)

$$(T + e^{iA}V(\mathbf{r})e^{-iA})|\rangle = E|\rangle, \tag{6}$$

and from this

$$(e^{-iA}Te^{iA} + V(\mathbf{r}))|\cdot\rangle = E|\cdot\rangle \tag{7}$$

In (7), we have $|\cdot\rangle = e^{-iA}|\rangle$; thus the state vector is modified by the shift transformation

$$|\rangle \rightarrow e^{-iA}|\rangle.$$

* Present address Instituto Politecnico Nacional, Mexico.

[1] R. Jastrow, Phys. Rev. **81**, 165 (1951).

[2] J. Gammel and R. Thaler, Phys. Rev. **107**, 291 (1957).

[3] M. Razavy, O. Rojo, and J. S. Levinger, Proc. 1st Intern. 1 Conf. Nucl. Structure, Kingston, 1960, p. 176; O. Rojo and L. Simmons, Phys. Rev. **125**, 273 (1962).

[4] W. Frahn and R. H. Lemmer, Nuovo Cimento **5**, 1564 (1957).

[5] G. Breit, H. Hull, K. Lassila, and K. D. Pyatt, Phys. Rev. **120**, 2227 (1960).

[6] M. Moshinsky, Phys. Rev. **106**, 117 (1957).

[7] J. Gammel and R. Thaler, *Progress in Elementary Particles and Cosmic Rays Physics* (North-Holland Publishing Company, Amsterdam, 1960), Vol. V.

[8] J. S. Bell, Lecture Notes, 1961 Bergen International School of Physics, Edited by Fronsdal (Benjamin and Company, New York, 1962).

[9] A. M. Green, Phys. Letters **1**, 136 (1962).

[10] M. J. Moravcsik, *The Two-Nucleon Interaction* (Clarendon Press, Oxford, England, 1963).

[11] S. Okubo and R. Marshak, Ann. Phys. (N. Y.) **4**. 166 (1958).

Assume $A = P$, where P is an operator diagonal in p-representation. This way we make sure that $\mathcal{U}$ is neither diagonal in $\mathbf{p}$- nor in $\mathbf{q}$-basis (excepting trivial cases). [Note: the linear case, $A = a_0 + a_1\mathbf{p}$, makes $\mathcal{U}$ local. Thus

$$\mathcal{U} = e^{i(a_0+a_1\mathbf{p})}V(\mathbf{r})e^{-i(a_0+a_1\mathbf{p})} = e^{ia_1\mathbf{p}}V(\mathbf{r})e^{-ia_1\mathbf{p}}.$$

Then in $\mathbf{p}$-basis we have

$$\langle\mathbf{p}'|\,\mathcal{U}\,|\mathbf{p}''\rangle = \mathcal{U}(\mathbf{p}' - \mathbf{p}'')e^{ia_1(\mathbf{p}'-\mathbf{p}'')} = f(\mathbf{p}' - \mathbf{p}'')$$

in $\mathbf{r}$-basis the last expression is written as

$$\langle\mathbf{r}'|\,\mathcal{U}\,|\mathbf{r}''\rangle = \delta(\mathbf{r}' - \mathbf{r}'')V(\mathbf{r}'' - a_1).$$

It is very easy to prove that this is the only velocity-dependent local potential. It introduces a repulsive core explicitly.]

With this approximation, we assure that the physical features of the nuclear potential are achieved. Thus, making $P = a_0 + a_1\mathbf{p} + a_2\mathbf{p}^2$, and using the Hermiticity property for P, but not necessarily the time-reversal property, we have

$$P^\dagger = a_0^* + a_1^*\mathbf{p}^\dagger + a_2^*(\mathbf{p}^2)^\dagger = a_0^* + a_1^*\mathbf{p} + a_2^*\mathbf{p}^2 = P. \quad (8)$$

From this, it follows that the a's must be real numbers.

For s- waves, in the one-dimensional case, we may write in p- basis:

$$\langle p\mid\cdot\rangle = \langle p|\,e^{-iP}|\rangle = e^{-iP}\Phi(p)$$

$$\fallingdotseq \mathfrak{F}[e^{-iP}] * \Psi(x) = \int \Psi(x - x')\mathfrak{F}[e^{-iP}]_{x'}\,dx', \quad (9)$$

where

$$\mathfrak{F}[e^{-iP}]_{x'} = (2\pi)^{-\frac{1}{2}}\int e^{ipx'}e^{-i(a_0+a_1p+a_2p^2)}\,dp = \text{const.}\, e^{i(\pi/4-a_0)}e^{-i[(a_1+x')^2/4a_2]}.$$

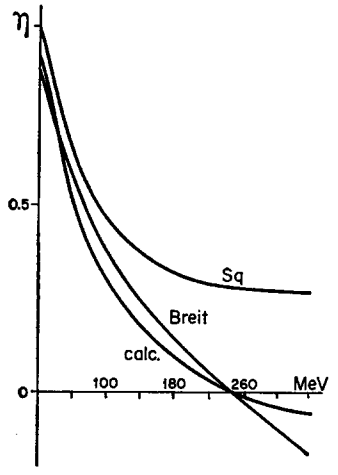


FIG. 1. Comparison between the phase shifts produced by (I) a static square potential well, with $V_0 = 20.1$ MeV., $b = 2.2$ F; (II) its transformation with the shift $e^{-ip}|>$, in first approximation, using $a_1 = 0.16$ F; (III) Breit's results.

Taking for $\Psi(x)$ the asymptotic form corresponding to the static potential, $\sin(kx + \delta)$, one arrives after integration at[12]

$$\langle x|\,e^{iP}|\rangle = ce^{i(\frac{1}{2}\pi-a_0)}\int_{-\infty}^{\infty}\sin[k(x - x') + \delta] \times\left[\cos\frac{(a_1 - x')^2}{4a_2} + i\sin\frac{(a_1 - x')^2}{4a_2}\right]dx'$$

$$= C\sin(kx + \delta - ka_1). \quad (10)$$

Observe that $\langle x|e^{iP}|\rangle$ does not depend on a_2; this is due to the fact that even terms in p do not contribute to the phase shifts. So taking the asymptotic form of $\Phi(p)$, i.e. $\frac{1}{2}[\delta(p - k) - \delta(p + k)]$ the Fourier transform of $e^{-i(a_0+a_1p+a_2p^2)}\Phi_{as}(p)$ gives the repulsive core ka_1, as contribution to the phase shift.

We see that the net effect has been the introduction of a pseudo-phase-shift in the asymptotic wavefunction, equal to ka_1. This in turn is equivalent to that introduced by a repulsive core with radius a_1. A rough approximation to Breit, s values can be obtained for the phase shifts, using a square-well for the static part of the potential, satisfying effective range theory ($V_0 = 30$ MeV., $b = 1.8$ F., $a_1 = 0.22$ F.; these values give an effective range $r_0 = 2.07$ F., scattering length $a = -27.3$ F., and shape-dependent parameter $P = +0.005$) (Fig. 1).

An alternative treatment of this problem in three dimensions by S-matrix methods leads to the same result. (We owe this suggestion to Professor M. Moshinsky).

III. EFFECTIVE POTENTIAL AND PHASE SHIFTS FOR SPECIAL VALUES OF $\mathcal{U}$

In this section we will see how velocity-dependent potentials arise by means of the unitary transformation, and solve the Schrödinger equation for a particular case.

Taking terms up to p^2 to avoid derivatives higher than second order in Eq. (1), the expression

$$\mathcal{U} = e^{-iP}Ve^{iP} = \sum_n \frac{i^n}{n!}[P[P[\cdots^{(n}\ [P, V]]$$

becomes

$$\mathcal{U} = V + i[P, V] + \tfrac{1}{2}i^2[P[P, V]] = V + a_1^2\mathbf{p}V\mathbf{p} - (a\mathbf{p}^2V + V\mathbf{p}^2a^*) \quad (11)$$

(where $a = \frac{1}{2}a_1^2 - ia_2$), after taking into account time-reversal invariance.

[12] W. Grobner and N. Hofreiter, *Integraltafeln* (Springer-Verlag, Berlin, 1961).

The real part of Eq. (11) is the velocity-dependent potential

$$\mathcal{V}_{\text{real}} = V + a_1^2 \mathbf{p} V \mathbf{p} - \tfrac{1}{2}a_1^2(\mathbf{p}^2 V + V\mathbf{p}^2). \quad (12)$$

Equation (12) resembles greatly Baker's formula[13]

$$\left\{\frac{1}{4M}[\mu(r)p^2 + 2p\mu(r)p + p^2\mu(r)] + V(r)\right\}\Psi(r) = E\Psi(r).$$

Nevertheless, his $\mu(r)$ is given now by the static potential itself, without need of a new function. Note that here, as in Baker's work, the correct formalism results from the fusion of both types of potentials, $\mathbf{p}^2\omega + \omega\mathbf{p}^2$ and $\mathbf{p}J\mathbf{p}$, used in previous works. [Note that in spite of the appearance of the "i" in Eq. (11), stays to be Hermitian, because in the potential $W + i\Omega$, Ω is anti-Hermitian. For this reason the change (16) gives again a real effective potential.]

This potential $\mathcal{V}_{\text{real}}$ is local, since it can be obtained from $e^{-ia_1\mathbf{p}}Ve^{ia_1\mathbf{p}}$, and so we consider it as the zeroth-order approximation.

The Schrödinger equation with this potential $\mathcal{V}_{\text{real}}$ is written, in r-representation, as

$$u_l'' + \left[k^2 + \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left\{\frac{2J'}{r} + J''\right\} - \frac{l(l+1)}{r^2}\right]u_l = 0, \quad (13)$$

where the effective potential is

$$W_{\text{eff}} = -\left[\frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left\{\frac{2J'}{r} + J''\right\}\right] \quad (14)$$

with $V = -V_0 J(r)$, and $\mu = MV_0/\hbar^2 a_1 > 0$.

For the case $J(r) = \exp(-r/r_0)$, Eq. (14) gives

$$W_{\text{eff}} = (V_0 M/\hbar^2)[-(1 + \tfrac{1}{2}\alpha^2)e^{-x} + \alpha^2 e^{-x}/x] \quad (15)$$

where $x = r/r_0$ and $\alpha = a_1/r_0$.

We see that using an exponential for the static part of the potential, we are led to a function that is the sum of an attractive potential plus a repulsive Yukawa (Fig. 2). Note the similarity of this graph with earlier potentials having repulsive cores.[1]

Similar effects may be obtained with an attractive Gaussian function for the static part.

Taking the whole potential, one arrives at

$$u_l'' + \left[k^2 + \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left\{\frac{2J'}{r} + J''\right\} - \frac{l(l+1)}{r^2}\right]u_l + ia_2\frac{MV_0}{\hbar^2}(2J'u_l' + J''u_l) = 0, \quad (16)$$

[13] G. Baker, Phys. Rev. 128, 1485 (1962).

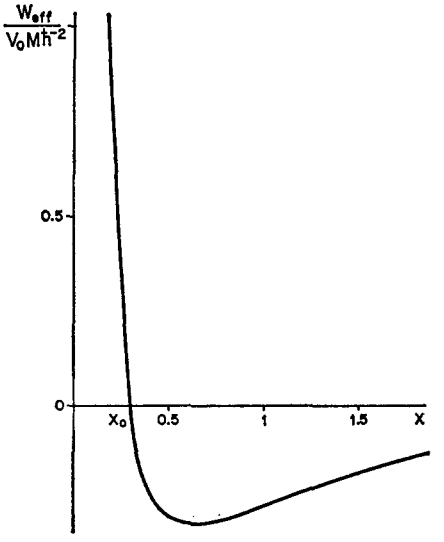


FIG. 2. Effective potential resulting from the use of an exponential for the static part, satisfying effective range theory. We use $a_1 = 0.4$ F, $r_0 = 0.62$ F, and $x = r/r_0$.

which at first glance looks like an optical potential, but which by simply changing variables,

$$u_l = v_l e^{-ia_2(MV_0/\hbar^2)J}$$

becomes

$$v_l'' + [k^2 - W_{\text{eff}} - l(l+1)/r^2]v_l = 0 \quad (17)$$

where

$$-W_{\text{eff}} = \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left\{\frac{2J'}{r} + J''\right\} + \frac{a_2^2 M^2 V_0^2}{\hbar^4}J'^2. \quad (18)$$

This shows that the complex terms are spurious.

Observe that in Eq. (15) the core value is not modified by the introduction of the new term, and only the shape of W_{eff} changes for large r.

For the special case of being J a square shape of range b, the s-phase-shifts η_s of the wavefunction are given by

$$\eta_s = \pi - kb + \tan^{-1} kb/(Ak'b \cot k'b + B) \quad (19)$$

[from Eq. (13)]. Here A and B are two constants depending on μ and $u(r)$, which is discontinuous at $r = b$. So this solution remains ambiguous, because we get the product of a distribution with a discontinuous function.

IV. HIGHER-ORDER APPROXIMATIONS TO $\mathcal{V}$

The equation $\mathcal{V} = e^{iA}Ve^{-iA}$ has been solved supposing $A \neq A(\mathbf{r})$, to exclude the possibility that $\mathcal{V} = V(\mathbf{r})$. The zeroth-order $\mathcal{V}_0$ and first-order approximation $\mathcal{V}_I$ are obtained by writing $A =$

$A(\mathbf{p}) = P$. In the first case P is a linear function of $\mathbf{p}$.

If A is not equal to $P(\mathbf{p})$, it is still possible to diagonalize A in $\mathbf{p}$-basis, as follows:

$$A = e^{iB}Pe^{-iB}. \tag{20}$$

Now, B can not be $B(\mathbf{p})$, since this leads to $\mathcal{V}_{II} = \mathcal{V}_I$. Let us put $B = B(\mathbf{r}) = V_1$, then

$$\begin{aligned}\mathcal{V}_{II} &= e^{ie^{iV_1}Pe^{-iV_1}}Ve^{-ie^{iV_1}Pe^{-iV_1}} \\ &= e^{iV_1}e^{iP}Ve^{-iP}e^{-iV_1} = e^{iV_1}\mathcal{V}_I e^{-iV_1}\end{aligned} \tag{21}$$

and if $V_1 = \text{const.}$, we have $\mathcal{V}_{II} = \mathcal{V}_I$. Using the formalism of commutators, one may write

$$\begin{aligned}\mathcal{V}_{II} &= \sum_m \frac{i^m}{m!}[V_1[V_1[\cdots[V_1, \mathcal{V}_I]] \\ &= \sum_{m,n}\frac{i^{m+n}}{m!n!}[V_1[V_1[\cdots[P[P[\cdots[P, V]]].\end{aligned} \tag{22}$$

For higher-order approximations, we follow the same pattern, assuming that $B = e^{iC}V_1(\mathbf{r})e^{-iC}$, and so on.

For example, with $C = P_1$,

$$\mathcal{V}_{III} = e^{-iP}e^{-iV_1}e^{-iP_1}\mathcal{V}_{II}e^{iP_1}e^{iV_1}e^{iP}. \tag{23}$$

If P_1 is constant, we have $\mathcal{V}_{III} = \mathcal{V}_{II}$. In this way we expect that the operators

$$\begin{aligned}P_i &= a_0^i + a_1^i\mathbf{p} + a_2^i\mathbf{p}^2 + \cdots; \\ V_i &= b_0^i + b_1^i\mathbf{r} + b_2^i\mathbf{r}^2 + \cdots\end{aligned}$$

will converge to a constant for i sufficiently large. If $P_i = \text{const.}$, the ultimate form of the potential is $\mathcal{V} = \mathcal{V}_{(i)-1}$. If $V_i = \text{const.}$, it is $\mathcal{V} = \mathcal{V}_{(i)}$

With such expressions for $\mathcal{V}_{II}, \mathcal{V}_{III}, \cdots$, it is easy to find new forms for the Schrödinger equation.

The handling of terms like $e^{iV_1}\mathbf{p}V\mathbf{p}e^{-iV_1}$, may be avoided by using a device that we employ later, when we treat the effective mass.

From $(T + \mathcal{V}_{II})|\rangle = E|\rangle$ we have

$$(e^{-iV_1}Te^{iV_1} + \mathcal{V}_I)\,|\cdot\rangle = E\,|\cdot\rangle,\ (|\cdot\rangle = e^{-iV_1}|\rangle).$$

As V_1 is local and real, $\langle\mathbf{r}|\cdot\rangle$ and $\langle\mathbf{r}|\rangle$ differ only by a phase; the Schrödinger equation then reads

$$\begin{aligned}&u_l'' + \left[k^2 + \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left\{\frac{2J'}{r} + J''\right\} - V_1^2\right. \\ &\left. - \frac{l(l+1)}{r^2}\right]u_l + i\left[\left(a_2\frac{MV_0}{\hbar^2}J'' + V_1''\right)u_l\right. \\ &\left. + \left(\frac{2a_2MV_0}{\hbar^2}J' + 2V_1'\right)u_l'\right] = 0.\end{aligned} \tag{24}$$

By taking

$$\begin{aligned}u_l &= v_l \exp\left[-i\int\left(\frac{a_2MV_0}{\hbar^2}J' + V_1'\right)dr\right] \\ &= v_l \exp\left[-i\left(\frac{a_2MV_0}{\hbar^2}J + V_1\right)\right],\end{aligned}$$

we have

$$\begin{aligned}-W_{\text{eff}} &= \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left(\frac{2J'}{r} + J''\right) \\ &\quad + \left(\frac{a_2MV_0}{\hbar^2}J' + V_1\right)^2.\end{aligned} \tag{25}$$

Finally, our potential permits us to obtain higher-order approximations, simply by adding new terms to the static potential, as we summarize here:

$$\begin{aligned}-W_{\text{eff}} &= \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left(\frac{2J'}{r} + J''\right) && \text{(zeroth order)}, \\ -W_{\text{eff}} &= \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left(\frac{2J'}{r} + J''\right) + \left(\frac{a_2MV_0}{\hbar^2}J'\right)^2 && \text{(first order)}, \\ -W_{\text{eff}} &= \frac{MV_0}{\hbar^2}J + \frac{\mu}{2}\left(\frac{2J'}{r} + J''\right) + \left(\frac{a_2MV_0}{\hbar^2}J' + V_1\right)^2 && \text{(second order)}.\end{aligned} \tag{26}$$

V. EFFECTIVE MASS

We now show that an effective mass appears when we use the second approximation to $\mathcal{V}$.

If we set $A = P(\mathbf{p})$ in Eq. (7), as T is diagonal in $\mathbf{p}$-basis, it commutes both with $\mathbf{p}$ and $A(\mathbf{p})$, and $e^{-iA}Te^{iA} = T$; hence there is not effective mass in the first approximation.

For the second approximation, we obtain

$$(e^{-iP}e^{-iV_1}Te^{iV_1}e^{iP} + V)\,|\cdot\cdot\rangle = E\,|\cdot\cdot\rangle$$

where $|\cdot\cdot\rangle$ means $e^{-iP}e^{-iV_1}|\rangle$; that is to say, the wavefunction transformed by the second approximation.

Now V is local, and

$$e^{-iP}e^{-iV_1}Te^{iV_1}e^{iP} \equiv e^{-iP}e^{-iV_1}e^{iP}Te^{-iP}e^{iV_1}e^{iP}$$

takes the place of the kinetic energy operator. As we shall see, apart from this phase shift, the second approximation gives rise to an effective mass. Let us call

$$e^{-iP}e^{-iV_1}e^{iP} = S \tag{27}$$

and define $M^{\frac{1}{2}} = m^{\frac{1}{2}}S$; then the kinetic energy term of the Schrödinger equation takes the form

$$1/M^{\frac{1}{2}}\mathbf{p}^2(1/M^{\frac{1}{2}})^{\dagger}.$$

The operator S is not Hermitian, but we will use for the definition of effective mass M^* the Hermitian part of this operator in the following way:

$$M^{*\frac{1}{2}} = m^{\frac{1}{2}}[\tfrac{1}{2}(S + S^{\dagger})]. \tag{28}$$

Then

$$\begin{aligned} M^{*\frac{1}{2}} &= m^{\frac{1}{2}}\tfrac{1}{2}[e^{-ie^{-iP}V_1e^{iP}} + e^{ie^{-iP}V_1e^{iP}}] \\ &= m^{\frac{1}{2}}\cos(e^{-iP}V_1e^{iP}) \\ &= m^{\frac{1}{2}}e^{-iP}\cos V_1e^{iP}. \end{aligned} \tag{29}$$

In this way we make sure that the eigenvalues of M^* are real. Since S is unitary, it comes out naturally that the eigenvalues of $\frac{1}{2}(S + S^{\dagger})$ are included in the interval $(-1, +1)$, and M^* will have any eigenvalue between zero and m. These are the values that would be obtained in the measurement of the mass.

From the last formula follows an expression analogous to Frahn and Lemmer's result:

$$\begin{aligned} M^* &= m\cos^2(e^{-iP}V_1e^{iP}) \\ &= m/[1 + \tan^2(e^{-iP}V_1e^{iP})], \end{aligned} \tag{30}$$

and $M^* \leq m$.

This shows that the effective mass is a nonlocal operator. In order to arrive at numerical results, observe that $\cos^2 V_1(x)$ is a function that, in good approximation, may be replaced by $\cos^2(-|x|)$ (only the one-dimensional case is to be considered). As may be shown, this function can have only one peak in the interval $(-\infty, \infty)$, in order to keep M^* less than m.

In p-space, the nondiagonal form of M^* is

$$\begin{aligned} M^*(p', p'') &= me^{-i[a_1(p'-p'')+a_3(p'^3-p''^3)+\cdots]}(2\pi)^{-\frac{1}{2}} \\ &\quad \times \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \cos^2 x e^{-i(p'-p'')x}\,dx \\ &= m\frac{(2\pi)^{\frac{1}{2}}\Gamma(3)e^{-i[a_1(p'-p'')+a_3(p'^3-p''^3)+\cdots]}}{2^3\Gamma[2 + \frac{1}{2}(p' - p'')]\Gamma[2 - \frac{1}{2}(p' - p'')]}. \end{aligned} \tag{31}$$

The diagonal part of this mass is $M^*_d = (\frac{1}{8}\pi)^{\frac{1}{2}}m = 0.621m.$, in agreement with previously used values.

VI. FINAL REMARKS

In summary, the Schrödinger equation is only a relationship between operators, in the broadest sense of quantum mechanics. The Schrödinger equation in its most general form is

$$e^{iB}T(\mathbf{p})e^{-iB}|\rangle + e^{iA}V(\mathbf{r})e^{-iA}|\rangle = E|\rangle.$$

This equation can be reduced to (a) effective potential formalism:

$$(T + e^{-iC}Ve^{iC})\,|\cdot\rangle = E\,|\cdot\rangle, \quad |\cdot\rangle = e^{-iB}|\rangle; \tag{32}$$

(b) effective mass formalism:

$$(e^{iC}Te^{-iC} + V)\,|\cdot\rangle = E\,|\cdot\rangle, \quad |\cdot\rangle = e^{-iA}|\rangle; \tag{33}$$

where $C = B - A + \frac{1}{2}[BA] + \frac{1}{12}[B[BA]] + \cdots$.

Equation (33) generalizes the Schrödinger equation with a mass $M^{\frac{1}{2}} = m^{\frac{1}{2}}e^{-iC}$, giving

$$(1/M^{\frac{1}{2}})\mathbf{p}^2[1/(M^{\frac{1}{2}})^{\dagger}] + V = E,$$

and we define effective mass M^* as the Hermitian part of M, i.e., $M^* = m\cos^2 C$.

Then we have always $M^* \leq m$.

ACKNOWLEDGMENTS

The authors are indebted to Dr. A. Galindo, of the CERN, for several helpful suggestions, and to Prof. M. Moshinsky for comments and discussions. They also wish to thank Dr. R. D. Mattuck for his reading and criticism of the manuscript.

APPENDIX

We want now to analyze the symmetry properties that characterize the potential.

It is most convenient to discuss the symmetry properties of in r-basis, where we have

$$\mathcal{U}(\mathbf{r}, \mathbf{r}') \equiv \langle\mathbf{r}|\,\mathcal{U}\,|\mathbf{r}'\rangle,$$

or in terms of the Fourier transform of this,

$$\mathcal{U}(\mathbf{r}, \mathbf{p}) = \mathfrak{F}[\mathcal{U}(\mathbf{r}, \mathbf{r}')]_{\mathbf{r}'}.$$

(It is clear that the two forms will have the same symmetry properties.)

Time-reversal invariance $\mathcal{U}(\mathbf{r}, \mathbf{p}) = \mathcal{U}^*(r, -\mathbf{p})$ requires $\mathcal{U}(\mathbf{r}, \mathbf{r}')$ to be real, since

$$\begin{aligned} \mathcal{U}(\mathbf{r}, \mathbf{p}) &= \int \langle\mathbf{r}|\,\mathcal{U}\,|\mathbf{r}'\rangle\,d\mathbf{r}'\,\langle\mathbf{r}'\,|\,\mathbf{p}\rangle \\ &= (2\pi)^{-\frac{1}{2}}\int \mathcal{U}(\mathbf{r}, \mathbf{r}')e^{i\mathbf{r}'\mathbf{p}}\,d\mathbf{r}' \end{aligned}$$

and

$$\begin{aligned} \mathcal{U}(\mathbf{r}, -\mathbf{p}) &= \int \langle\mathbf{r}|\,\mathcal{U}\,|\mathbf{r}'\rangle\,d\mathbf{r}'\,\langle\mathbf{r}'\,|\,-\mathbf{p}\rangle \\ &= (2\pi)^{-\frac{1}{2}}\int \mathcal{U}(\mathbf{r}, \mathbf{r}')e^{-i\mathbf{r}'\mathbf{p}}\,d\mathbf{r}'. \end{aligned} \tag{A1}$$

As a matter of fact, we demand the following conditions to be satisfied by the operator $\mathcal{U}$:

(a) Hermitian: (A2)

$$\mathcal{U}(\mathbf{r}, \mathbf{r}') = \mathcal{U}^*(\mathbf{r}', \mathbf{r});$$

(b) invariant under translations:

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2; \quad \text{(A3)}$$

(c) invariant under Galilean transforms:

$$\mathbf{p} = \tfrac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2); \quad \text{(A4)}$$

(d) the particles being indiscernible, must remain invariant under exchange:

$$\mathcal{U}(\mathbf{r}, \mathbf{r}') = \mathcal{U}(-\mathbf{r}, -\mathbf{r}') \quad \text{(A5)}$$

(we have assumed charge invariance to hold).

(e) invariant under time reversal.

The following relations, sometimes given as independent conditions, are consequence of these:

(1) $$\mathcal{U}(\mathbf{r}, \mathbf{p}) = \mathcal{U}^*(\mathbf{r}, -\mathbf{p});$$

This follows from (a), because

$$\mathcal{U}(\mathbf{r}, -\mathbf{p}) = \int \mathcal{U}(\mathbf{r}, \mathbf{r}')e^{-i\mathbf{r}'\mathbf{p}}\, d\mathbf{r}' = \mathfrak{F}[\mathcal{U}^*(\mathbf{r}, \mathbf{p})]_{\mathbf{p}};$$

(2) $$\mathcal{U}(\mathbf{r}, -\mathbf{r}') = \mathcal{U}^*(\mathbf{r}, \mathbf{r}')$$

[this follows from (b), by the same argument].

The reality of $\mathcal{U}(\mathbf{r}, \mathbf{r}')$, together with the given symmetry conditions, make plain the existence of the following symmetry axes for the $\mathcal{U}$ operator: the $\mathbf{r}$ and $\mathbf{r}'$ axes, and their bisectrices in the hyperspace of the moduli of $\mathbf{r}$ and $\mathbf{r}'$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Improved Method for Quantum Mechanical Three-Body Problems. I. Attractive Potentials

LEONARD EYGES
Air Force Cambridge Research Laboratories, Office of Aerospace Research, L. G. Hanscom Field, Bedford, Massachusetts
(Received 7 January 1965; final manuscript received 22 February 1965)

The quantum-mechanical ground-state problem for three identical particles bound by attractive interparticle potentials is discussed. For this problem it has previously been shown that it is advantageous to write the wavefunction in a special functional form, for which an integral equation which is equivalent to the Schrödinger equation was derived. In this paper a new method for solving this equation is presented. The method involves an expansion in a set of "effective" two-body functions; these are the eigenfunctions of a two-body problem with a potential of the same shape as the interparticle potential in the three-body problem, but of enhanced strength. The integral equation then becomes a set of coupled equations for certain functions $f_l(\kappa)$, the "coefficients" in this expansion. By choosing the effective two-body strength properly, one can optimize the convergence of this expansion so that a good approximation is got by retaining only the lowest ($l = 0$) term. The resultant single integral equation for $f_0(\kappa)$ can then be solved by an approximate method. As a test and check, the one-dimensional three-body problem with δ-function interactions is treated and the results are compared with the known exact one. It is found that the first approximation in the method, which is almost trivial to apply, yields an eigenvalue and eigenfunction accurate to a few percent. The method has also been applied to the three-dimensional problem with exponential interparticle potentials, and a comparison of the results with the best numerical calculations is given.

I. INTRODUCTION

IN a previous paper[1] we have treated the quantum-mechanical problem of finding the ground-state energy and eigenfunction for three identical particles bound by attractive interparticle potentials. In that paper the wavefunction was written in a special functional form for which we derived an integral equation which was equivalent to the Schrödinger equation. This equation [Eq. (14) of II] was then solved approximately and, if it must be said, inelegantly. The numerical results were, however, in good agreement with the results of other (variational) calculations.

In the present paper we return to that integral equation, for which we have now found a much neater method of solution. Moreover, it is one that

[1] L. Eyges, Phys. Rev. **121**, 1744 (1961); hereafter referred to as II.

As a matter of fact, we demand the following conditions to be satisfied by the operator $\mathcal{U}$:

(a) Hermitian: (A2)

$$\mathcal{U}(\mathbf{r}, \mathbf{r}') = \mathcal{U}^*(\mathbf{r}', \mathbf{r});$$

(b) invariant under translations:

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2; \quad \text{(A3)}$$

(c) invariant under Galilean transforms:

$$\mathbf{p} = \tfrac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2); \quad \text{(A4)}$$

(d) the particles being indiscernible, must remain invariant under exchange:

$$\mathcal{U}(\mathbf{r}, \mathbf{r}') = \mathcal{U}(-\mathbf{r}, -\mathbf{r}') \quad \text{(A5)}$$

(we have assumed charge invariance to hold).

(e) invariant under time reversal.

The following relations, sometimes given as independent conditions, are consequence of these:

(1) $$\mathcal{U}(\mathbf{r}, \mathbf{p}) = \mathcal{U}^*(\mathbf{r}, -\mathbf{p});$$

This follows from (a), because

$$\mathcal{U}(\mathbf{r}, -\mathbf{p}) = \int \mathcal{U}(\mathbf{r}, \mathbf{r}')e^{-i\mathbf{r}'\mathbf{p}}\, d\mathbf{r}' = \mathfrak{F}[\mathcal{U}^*(\mathbf{r}, \mathbf{p})]_{\mathbf{p}};$$

(2) $$\mathcal{U}(\mathbf{r}, -\mathbf{r}') = \mathcal{U}^*(\mathbf{r}, \mathbf{r}')$$

[this follows from (b), by the same argument].

The reality of $\mathcal{U}(\mathbf{r}, \mathbf{r}')$, together with the given symmetry conditions, make plain the existence of the following symmetry axes for the $\mathcal{U}$ operator: the $\mathbf{r}$ and $\mathbf{r}'$ axes, and their bisectrices in the hyperspace of the moduli of $\mathbf{r}$ and $\mathbf{r}'$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 8 AUGUST 1965

Improved Method for Quantum Mechanical Three-Body Problems. I. Attractive Potentials

LEONARD EYGES
Air Force Cambridge Research Laboratories, Office of Aerospace Research, L. G. Hanscom Field, Bedford, Massachusetts
(Received 7 January 1965; final manuscript received 22 February 1965)

The quantum-mechanical ground-state problem for three identical particles bound by attractive interparticle potentials is discussed. For this problem it has previously been shown that it is advantageous to write the wavefunction in a special functional form, for which an integral equation which is equivalent to the Schrödinger equation was derived. In this paper a new method for solving this equation is presented. The method involves an expansion in a set of "effective" two-body functions; these are the eigenfunctions of a two-body problem with a potential of the same shape as the interparticle potential in the three-body problem, but of enhanced strength. The integral equation then becomes a set of coupled equations for certain functions $f_l(\kappa)$, the "coefficients" in this expansion. By choosing the effective two-body strength properly, one can optimize the convergence of this expansion so that a good approximation is got by retaining only the lowest ($l = 0$) term. The resultant single integral equation for $f_0(\kappa)$ can then be solved by an approximate method. As a test and check, the one-dimensional three-body problem with δ-function interactions is treated and the results are compared with the known exact one. It is found that the first approximation in the method, which is almost trivial to apply, yields an eigenvalue and eigenfunction accurate to a few percent. The method has also been applied to the three-dimensional problem with exponential interparticle potentials, and a comparison of the results with the best numerical calculations is given.

I. INTRODUCTION

IN a previous paper[1] we have treated the quantum-mechanical problem of finding the ground-state energy and eigenfunction for three identical particles bound by attractive interparticle potentials. In that paper the wavefunction was written in a special functional form for which we derived an integral equation which was equivalent to the Schrödinger equation. This equation [Eq. (14) of II] was then solved approximately and, if it must be said, inelegantly. The numerical results were, however, in good agreement with the results of other (variational) calculations.

In the present paper we return to that integral equation, for which we have now found a much neater method of solution. Moreover, it is one that

[1] L. Eyges, Phys. Rev. **121**, 1744 (1961); hereafter referred to as II.

seems to be generalizable to the four-body and even N-body problem. We present this method here, and also compare the results it yields with exact results that are now available for the one-dimensional problem with δ-function potentials.

For convenience we begin by summarizing some of the results of II. An essential point in it was the simultaneous use of three different sets of coordinates in place of $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, the coordinate vectors of the particles. These sets were

$$(\mathbf{R}, \mathbf{r}_{12}, \boldsymbol{\varrho}_3), (\mathbf{R}, \mathbf{r}_{13}, \boldsymbol{\varrho}_2), (\mathbf{R}, \mathbf{r}_{23}, \boldsymbol{\varrho}_1),$$

where

$$\mathbf{R} = \tfrac{1}{3}(\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3)$$

and, for example,

$$\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2,$$
$$\boldsymbol{\varrho}_3 = \mathbf{r}_3 - \tfrac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2),$$

with the other variables being defined analogously. In terms of these variables we wrote the wavefunction Ψ of the system (with center-of-mass coordinate $\mathbf{R}$ split off) in the form

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \psi(\mathbf{r}_{12}, \boldsymbol{\varrho}_3) + \psi(\mathbf{r}_{13}, \boldsymbol{\varrho}_2) + \psi(\mathbf{r}_{23}, \boldsymbol{\varrho}_1). \quad (1)$$

We then defined the Fourier transform $\varphi(\mathbf{k}, \boldsymbol{\kappa})$ of $\psi(\mathbf{r}, \boldsymbol{\varrho})$ by

$$\psi(\mathbf{r}, \boldsymbol{\varrho}) = \frac{1}{(2\pi)^3} \iint \varphi(\mathbf{k}, \boldsymbol{\kappa}) e^{i(\mathbf{k}\cdot\mathbf{r} + \boldsymbol{\kappa}\cdot\boldsymbol{\varrho})}\, d\mathbf{k}\, d\boldsymbol{\kappa}$$

and were led to the basic integral equation for φ

$$\varphi(\mathbf{k}, \boldsymbol{\kappa}) = -\frac{1}{(2\pi)^3(k^2 + \frac{3}{4}\kappa^2 + k^2)} \times \iint v(r)\{\varphi(\mathbf{k}', \boldsymbol{\kappa}) e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k})} + \varphi(\tfrac{1}{2}\mathbf{k}', \mathbf{k}' - 2\boldsymbol{\kappa})(e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k}-\frac{1}{2}\boldsymbol{\kappa})} + e^{i\mathbf{r}\cdot(-\mathbf{k}'-\mathbf{k}+\frac{1}{2}\boldsymbol{\kappa})}\}\, d\mathbf{k}'\, d\mathbf{r}. \quad (2)$$

Here $E = -|E|$ is the (negative) ground state energy of the system, $V(r)$ is the interparticle potential, and

$$K^2 = m\,|E|/\hbar^2, \qquad v(r) = mV(r)/\hbar^2.$$

Our basic result, Eq. (2), is of course for the three-body problem in three dimensions. At this point, however, it is convenient to backtrack somewhat and consider the analogous problem in one dimension. This is because there has appeared[2] since the writing of II, a solution of a special, but nonetheless very useful, one-dimensional many-body problem: namely, that in which the interparticle potentials are attractive δ-functions. By writing the analog of (2) for this one-dimensional case and applying our method of solution to it we can compare the approximate answer with the exact; it goes without saying that this is a most useful check.

For the one-dimensional case everything proceeds very much as for three dimensions except for some factors of 2π and the fact that the variables are not vectors but scalars. In summary: Instead of the position coordinates x_1, x_2, x_3 we introduce a center-of-mass coordinate X, and x_{ij} and y_k, the one-dimensional analogs of $\mathbf{r}_{ij}$ and $\boldsymbol{\varrho}_k$

$$x_{ij} = x_i - x_j,$$
$$y_k = x_k - \tfrac{1}{2}(x_i + x_j);$$

we break up the wavefunction $\Psi(x_1, x_2, x_3)$ in a way similar to (1), and define $\varphi(k, \kappa)$ by

$$\psi(x, y) = \frac{1}{2\pi} \iint \varphi(k, \kappa) e^{i(kx + \kappa y)}\, dk\, d\kappa;$$

we are then led to the one-dimensional analog of (2), viz.

$$\varphi(k, \kappa) = -\frac{s}{2\pi(k^2 + \frac{3}{4}\kappa^2 + K^2)} \times \iint u(x)\{\varphi(k', \kappa) e^{ix(k'-k)} + \varphi(\tfrac{1}{2}k', k' - 2\kappa)(e^{ix(k'-k-3\kappa/2)} + e^{ix(-k'-k+3\kappa/2)})\}\, dk'\, dx. \quad (3)$$

Here we have found it convenient to separate the "strength" s of the potential from its "shape" (ux) by writing

$$v(x) = su(x). \quad (4)$$

We shall first treat Eq.(3), with the aim of clarifying the essential ideas of our viewpoint without the extraneous complication that three dimensions brings. The formal generalization to three dimensions will then be simple.

II. SOLUTION OF THE INTEGRAL EQUATION IN TERMS OF TWO-BODY FUNCTIONS

In this paper we shall try to solve Eq. (3) in terms of two-body wavefunctions, i.e., in terms of eigenfunctions of a two-body problem. Let us then consider briefly the problem of the bound state of two particles, coordinates x_1 and x_2, attracted by an interparticle potential, i.e., one which is a function

[2] J. B. McGuire, J. Math. Phys. 5, 622 (1964).

only of $x_1 - x_2$. We shall, for reasons which will be clear later, take this potential to be one with the same *shape* as the interparticle potential in the actual three-body problem we wish to solve, but of different strength. That is, referring to Eq. (4), we shall take the "effective" two-body potential $v_{eff}(x)$, whose eigenfunctions we use, to be $v_{eff}(x) = su(x)$ where s is a strength parameter[3] at our disposal.

As usual, we shall be interested only in relative motion in this potential, i.e., in the behavior of the wavefunction on the relative variable $x = x_2 - x_1$. We call $\tilde{\psi}(x)$ a general wavefunction[3] in this relative coordinate, and denote its Fourier transform by $\tilde{\varphi}(k)$. More specifically we know that there will be a complete set of such two-body wavefunctions, with corresponding eigenvalues. In general, there will be a certain number of negative-energy eigenvalues, which we call $E_l \equiv -|E_l|$, numbering them from the lowest (most negative) as the zeroth one. We call the corresponding momentum-space wavefunctions $\tilde{\varphi}_l(k)$. Then $\tilde{\varphi}_l(k)$ satisfies the equation (Schrödinger equation in momentum space)

$$\tilde{\varphi}_l(k) = -\frac{s}{2\pi(k^2 + K_l^2)} \times \iint u(x)e^{ix(k'-k)}\tilde{\varphi}_l(k')\,dk'\,dx, \qquad (5)$$

where $K_l^2 = m|E_l|/\hbar^2$.

In addition to the $\tilde{\varphi}_l(k)$, the complete set of two-body eigenfunctions will include a continuum of eigenfunctions with positive energy eigenvalues. These eigenfunctions satisfy an equation much like (5). Now our procedure in principle involves the expansion of the function $\varphi(k, \kappa)$ that occurs in (3) in terms of this complete set of two-body functions; but in practice, to avoid dealing with the troublesome continuum functions, we try to get an adequate representation in terms of only the bound-state functions. Thus we try an expansion of $\varphi(k, \kappa)$ in terms of only the $\tilde{\varphi}_l(k)$, with coefficients f_l which are functions of κ, recognizing that such an expansion cannot be exact, but that it may well be accurate enough:

$$\varphi(k, \kappa) = \sum_l f_l(\kappa)\tilde{\varphi}_l(k). \qquad (6)$$

If then we put (6) into (2) we get initially

$$-2\pi(k^2 + \tfrac{3}{4}\kappa^2 + K^2) \sum_l f_l(\kappa)\tilde{\varphi}_l(k) = s \iint u(x)\left\{\left[\sum f_l(\kappa)\tilde{\varphi}_l(k')\right]e^{ix(k'-k)} + \left[\sum f_l(k' - 2\kappa)\tilde{\varphi}_l\left(\frac{k'}{2}\right)\right](e^{ix(k'-k-\frac{1}{2}\kappa)} + e^{ix(-k'-k+\frac{1}{2}\kappa)})\right\} dx\,dk'.$$

Now we multiply this equation by $\tilde{\varphi}_n^*(k)$ and integrate with respect to k. On the right-hand side we then use (5) to get, typically,

$$\iint \tilde{\varphi}_n^*(k)u(x)e^{ix(-k+k'-\frac{1}{2}\kappa)}\,dk\,dx = -(2\pi/s)((k' - \tfrac{3}{2}\kappa)^2 + K_n^2)\tilde{\varphi}_n^*(k' - \tfrac{3}{2}\kappa).$$

With this last equation, and the fact of the orthognality of the $\tilde{\varphi}_l(k)$, we get after a little algebra

$$f_n(\kappa)\left(K^2 + \frac{3}{4}\kappa^2 - \frac{s}{\tilde{s}}K_n^2\right) + \left(1 - \frac{s}{\tilde{s}}\right)\sum_l f_l(\kappa)\int k^2\varphi_n^*(k)\varphi_l(k)\,dk = \int [(k' - \tfrac{3}{2}\kappa)^2 + K_n^2]\left(\sum_l f_l(k' - 2\kappa)\tilde{\varphi}_l\left(\frac{k'}{2}\right)\right) \times [\tilde{\varphi}_n^*(k' - \tfrac{3}{2}\kappa') + \tilde{\varphi}_n^*(-k' + \tfrac{3}{2}\kappa)]\,dk'. \qquad (7)$$

This is the basic set of equations for the unknown functions $f_l(\kappa)$. Since the set is homogeneous it presumably has solutions only for certain eigenvalues and eigenfunctions, and it is these we seek. We shall of course have to truncate Eqs. (7) to solve them, and will have to hope that one can get a good representation of $\varphi(k, \kappa)$ with only a few of the functions $\tilde{\varphi}_l$. We shall be helped in this convergence problem by the fact that we have the strength s of the two-body function at our disposal, and are at liberty to choose it as advantageously as possible for improving convergence. In fact, we shall see that with a suitable choice of s, one can get a satisfactory representation of the wavefunction by keeping only the lowest term, with $l = 0$.

To begin then, we truncate Eqs. (7) as severely as possible and assume that only $\tilde{\varphi}_0$ is large. We also use the fact that $\tilde{\varphi}_0(k) = \tilde{\varphi}_0(-k)$ and make a change of variable in the right-hand side to get the equation that will occupy us the most.

$$f_0(\kappa) = \frac{2}{\Gamma + \frac{3}{4}\kappa^2}\frac{s}{\tilde{s}}\int \tilde{\varphi}_0^*\left(\frac{k}{2} + \kappa\right)\tilde{\varphi}_0\left(k + \frac{\kappa}{2}\right) \times \left[\left(k + \frac{\kappa}{2}\right)^2 + K_0^2\right]f_0(k)\,dk, \qquad (8)$$

where

$$\Gamma = K^2 - \frac{s}{\tilde{s}}K_0^2 + \left(1 - \frac{s}{\tilde{s}}\right)\int k^2\,|\varphi_0(k)|^2\,dk. \qquad (9)$$

[3] We shall generally use the tilde ($\sim$) to designate quantities that refer to the two-body problem.

Now we discuss the most advantageous way of choosing $\bar{s}$. It appears that a good choice is that value that would yield for the *two-body* problem the (as yet unknown) three-body energy K. Stated in other words, the two-body ground state binding energy $\bar{K}_0$ is a function of $\bar{s}$, or by the same token, $\bar{s}$ is a function of $\bar{K}_0$, $\bar{s} = \bar{s}(\bar{K}_0)$. If now we imagine a two-body problem for which $\bar{K}_0 = K$, this defines a special value of $\bar{s}$, which we call $\bar{s}(K)$, and it is this potential strength that we use in defining the two-body functions. The reason that this choice of $\bar{s}$ seems to work well appears to be this: We have found for three dimensions in II that if we set $\varkappa = 0$ in $\varphi(\mathbf{k}, \varkappa)$ and expand the resulting $\varphi(\mathbf{k}, 0)$ in partial waves, the S-wave component satisfies an equation which resembles a two-body equation but for an "effective" potential with a strength appropriate to the three-body binding energy. Thus the prescription given above is advantageous for the expansion of $\varphi(\mathbf{k}, 0)$ and our hope is that it retains some advantages for the expansion of $\varphi(\mathbf{k}, \varkappa)$.

III. APPLICATION TO δ-FUNCTION POTENTIALS IN ONE DIMENSION; COMPARISON WITH EXACT SOLUTION

As a simple application and test we shall now apply the results of the preceding section to the one-dimensional three-body problem with attractive δ-function interparticle potentials. As we have mentioned, the exact solution to this problem (and to the N-body one as well) is known,[2] so that we shall be able to compare our approximate three-body results with the exact one. These exact solutions are discussed in the Appendix; here we shall simply quote the results we need.

First, for the two-body problem with interparticle potential $v(x) = -s\delta(x)$, the lowest eigenvalue $\bar{K}_0$ and the corresponding eigenfunction are

$$\bar{K}_0 = \tfrac{1}{2}s, \tag{10}$$

$$\tilde{\varphi}_0(k) = (2\bar{K}_0^3/\pi)^{\frac{1}{2}}(k^2 + \bar{K}_0^2)^{-1}. \tag{11}$$

For the three-body problem the ground-state eigenvalue K, and the function $\varphi(k, \kappa)$ defined by (3) turn out to be

$$K = s, \tag{12}$$

$$\varphi(k, \kappa) = [(K^2 + \kappa^2)(K^2 + \tfrac{3}{4}\kappa^2 + k^2)]^{-1}. \tag{13}$$

It is with these last results that we shall compare our approximate solution.

We begin by specializing Eq. (8) to the present case. In accordance with the remarks above we take $\tilde{\varphi}_0(k)$ to be the wavefunction for that two-body problem for which $\bar{K}_0 = K$. According to (11) then

$$\tilde{\varphi}_0(k) = (2K^3/\pi)^{\frac{1}{2}}(k^2 + K^2)^{-1}. \tag{14}$$

The "effective strength" parameter $\bar{s}$ is then from (10) just $\bar{s} = 2K$, and the quantity Γ turns out to be

$$\Gamma = 2K^2(1 - s/2K) + \tfrac{3}{4}\kappa^2.$$

With these results then, the integral equation (8) takes the form

$$f_0(\kappa) = \frac{s}{\pi(1 + 3\kappa^2/8K^2 - s/2K)} \times \int_{-\infty}^{\infty} \frac{f_0(k)\,dk}{(\frac{1}{2}k + \kappa)^2 + K^2}. \tag{15}$$

We shall not try here to develop very precise methods for solving (15) but rather shall try only to solve it accurately enough to show that it is correct, and to make it plausible that one could do better if one wished. Our aim will then be merely to get a good enough answer to compare with the exact solution, with the thought of checking the general procedure up to this point. An approximate method which is suitable to our purpose in that it is both simple and seems to be reasonably accurate, goes as follows.[4] We observe that if we knew the exact wavefunction we could get a transcendental equation for the energy by evaluating (15) at some special value; a convenient value is $\kappa = 0$. Thus without approximation we have

$$f_0(0) = \frac{4s}{\pi(1 - s/2K)} \int_{-\infty}^{\infty} \frac{f_0(k)\,dk}{k^2 + K^2}.$$

If then we can find a good approximation for $f_0(\kappa)$ we can use it in (15) to find the approximate eigenvalue. We get this approximation by iteration. Given some zeroth-order approximation $f_0^{(0)}(\kappa)$ we define a first approximation $f_0^{(1)}(\kappa)$ by using this under the integral sign

$$f_0^{(1)}(\kappa) = \frac{4s}{\pi(1 + 3\kappa^2/8K^2 - s/2K)} \times \int \frac{f_0^{(0)}(k)\,dk}{(k + 2\kappa)^2 + K^2}. \tag{16}$$

Of course higher-order approximations can be defined similarly. What then is a good zeroth order approximation? One possibility emerges when we recognize that it is fairly clear from Eq. (15) itself that $f_0(\kappa)$ is a function which drops off rather quickly for large κ, probably at least as $1/\kappa^4$. This suggests that a good zeroth-order iterate is simply a δ-function of κ, and it is with this that we begin.

[4] Essentially this method has been used previously for the two-body, three-dimensional case: E. E. Salpeter, Phys. Rev. 84, 1226 (1951); E. E. Salpeter and J. S. Goldstein, *ibid.* 90, 983 (1953).

We put then $f_0^{(0)}(\kappa) = \delta(\kappa)$ on the right-hand side of Eq. (15) and get, some irrelevant factors aside,

$$f_0^{(1)}(\kappa) = [(1 - s/2K + 3\kappa^2/8K^2)(K^2 + \kappa^2)]^{-1}. \quad (17)$$

If now we put this into (15) we get after some algebraic reduction the transcendental equation for K as a function of s,

$$1 = \frac{4\xi}{(1+\xi)(5-4\xi)}\left(\frac{13+4\xi}{3} - \frac{3\sqrt{3}}{(2-\xi)^{\frac{1}{2}}}\right),$$

where $\xi = s/K$. Numerically the solution of this is $\xi = 1.034$. The correct value is of course $\xi = 1.0$. The agreement seems good enough to imply that Eq. (15) is basically correct, especially since it is likely that further iterations would improve this result.

It is also interesting to compare the exact and approximate expressions for $\varphi(k, \kappa)$. The exact one is given by Eq. (A12). Putting into it the eigenvalue $K = s$, and defining

$$k' = k/s, \qquad \kappa' = \kappa/s,$$

it becomes, on normalizing it to unity at the origin,

$$\varphi_{\text{exact}} = [(1 + \kappa'^2)(1 + \tfrac{3}{4}\kappa'^2 + k'^2)]^{-1}. \quad (18)$$

The approximate expression φ_{app} is gotten by combining (14) and (17), with the approximate value for K: $K \cong 0.996s$. Again normalizing to unity at the origin we get

$$\varphi_{\text{approx}} = [(1 + 1.069\kappa'^2)(1 + 0.801\kappa'^2) \times (1 + 1.069\kappa'^2)]^{-1}. \quad (19)$$

Comparing (18) and (19) we see that, e.g., for $\kappa' = 0$ the approximate and exact functions differ by a maximum of 6.9% for k' very large. For $k' = 0$ and κ' very large the exact function goes as $1.33/\kappa^{-4}$ and the approximate one as $1.169/\kappa^{-4}$, i.e., they differ asymptotically by about 14%. For $\kappa' = k' = \frac{1}{2}$ the exact wavefunction has the value 0.556 and the approximate one, 0.52, and the corresponding numbers for $k' = \kappa' = 1$ are 0.182 and 0.130. We think that these results can be considered to represent satisfactory agreement, for the present purpose.

It is also interesting to compare the exact and approximate coordinate-space wavefunctions. The exact wavefunction can be written in any of the three pairs of coordinates we have been using. For example, in x_{12}, y_3 coordinates it is

$$\Psi_{\text{exact}} = \exp[-\tfrac{1}{2}s(|x_{12}| + |\tfrac{1}{2}x_{12} + y_3| + |\tfrac{1}{2}x_{12} - y_3|)]. \quad (20)$$

The approximate wavefunction is of the form

$$\Psi_{\text{approx}} = \psi(x_{12}, y_3) + \psi(x_{13}, y_2) + \psi(x_{23}, y_1), \quad (21)$$

where $\psi(x, y)$ is defined as the Fourier transform of $\varphi(k, \kappa)$. Using the approximate expression above for $\varphi(k, \kappa)$, we get

$$\psi_{\text{approx}}(x, y) = e^{-\frac{1}{2}s|x|}(e^{-s|y|} - e^{-\mu|y|}/\mu),$$

where

$$\mu^2 = (1 - \tfrac{1}{2}\xi)8s^2/\xi^2, \qquad \xi = 1.034.$$

It is rather ironic that we cannot do the final integral involved in calculating $\psi(x, y)$ exactly [see Eq. (A11) of the Appendix], so that we cannot compare $\psi_{\text{app}}(x, y)$ and $\psi(x, y)$ directly. What we can do however, is form Ψ_{approx} from the right-hand side of (21), and compare with Ψ_{exact}. If we do this, and for comparison with (20) put everything into x_{12}, y_3 variables we finally get

$$\begin{aligned}\Psi_{\text{approx}} = {} & e^{-\frac{1}{2}s|x_{12}|}(e^{-s|y_3|} - \mu^{-1}e^{-\mu|y_3|}) \\ & + e^{-\frac{1}{2}s|y_3+\frac{1}{2}x_{12}|}(e^{-s|\frac{3}{4}x_{12}-\frac{1}{2}y_3|} - \mu^{-1}e^{-\mu|\frac{3}{4}x_{12}-\frac{1}{2}y_3|}) \\ & + e^{-\frac{1}{2}s|y_3-\frac{1}{2}x_{12}|}(e^{-s|\frac{3}{4}x_{12}+\frac{1}{2}y_3|} - \mu^{-1}e^{-\mu|\frac{3}{4}x_{12}+\frac{1}{2}y_3|}).\end{aligned}$$

The agreement between the exact and approximate wavefunctions above turns out to be, as one might expect, about as good in general as the agreement between the exact and approximate momentum space functions $\varphi(k, \kappa)$.

IV. THE THREE-BODY PROBLEM IN THREE DIMENSIONS

Now we turn to the three-body problem in three dimensions. As in one dimension, our aim will not be to calculate very precise solutions, since at the present stage we cannot compete with the extensive computations that exist,[5,6] but rather to get solutions good enough to give confidence in our general procedure, and to make sure that nothing anomalous happens in three dimensions. The three-dimensional case is of course intrinsically more complicated than the one-dimensional one, with scalar variables becoming vectors. There is however an additional independent complication: there are no three-dimensional potentials with two-body eigenfunctions even remotely as simple as the δ-function eigenfunction in one dimension. To keep things tractable then we shall have to use approximate two-body functions, and it is hard to estimate the error they introduce.

[5] Y. C. Tang, R. C. Herndon, and E. W. Schmid, Phys. Rev. **134**, B743 (1964).
[6] M. H. Kalos, Phys. Rev. **128**, 1791 (1962).

With these remarks we turn to the three-body problem itself, that is, to the solution of Eq. (2). As we have indicated, our procedure will be essentially the same as for the one-dimensional case. Thus we write $v(r) = su(r)$ and define an effective potential $v_{eff}(r)$ by $v_{eff}(r) = \tilde{s}u(r)$, in analogy to the one-dimensional case. Also, we choose the effective $\tilde{s}$ by the same prescription as for the one-dimensional case, and call this special value of $\tilde{s}$, $\tilde{s}(K)$. There will be a complete set of momentum-space eigenfunctions for $v_{eff}(r)$; we shall call them $\tilde{\varphi}_l(k)$, where the index l is now a formal one, standing for the various quantum numbers appropriate to three dimensions. The $\tilde{\varphi}_l$ corresponding to bound states satisfy

$$\tilde{\varphi}_l(\mathbf{k}) = -\frac{\tilde{s}}{(2\pi)^3(k^2 + \tilde{K}_l^2)} \times \iint \tilde{\varphi}_l(\mathbf{k}')u(r)e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k})}\,d\mathbf{k}'\,d\mathbf{r}. \quad (22)$$

We expand $\varphi(\mathbf{k}, \boldsymbol{\kappa})$ in terms of these

$$\varphi(\mathbf{k}, \boldsymbol{\kappa}) = \sum_l \tilde{\varphi}_l(\mathbf{k})f_l(\boldsymbol{\kappa})$$

and are led to a set of equations like (7). We truncate them, retaining only the lowest term and get the basic equation

$$f_0(\boldsymbol{\kappa}) = \frac{2}{\Gamma + \frac{3}{4}\kappa^2}\frac{s}{\tilde{s}}\int \tilde{\varphi}_0^*\left(\frac{\mathbf{k}}{2} + \boldsymbol{\kappa}\right)\tilde{\varphi}_0\left(\mathbf{k} + \frac{\boldsymbol{\kappa}}{2}\right) \times \left[\left(\mathbf{k} + \frac{\boldsymbol{\kappa}}{2}\right)^2 + \tilde{K}_0^2\right]f_0(\mathbf{k})\,d\mathbf{k}, \quad (23)$$

where

$$\Gamma = \left(1 - \frac{s}{\tilde{s}(K)}\right)\left[K^2 + \int k^2\,|\tilde{\varphi}_0(\mathbf{k})|^2\,d\mathbf{k}\right].$$

The method of approximate solution is much like that in one dimension. Namely, we first set $\kappa = 0$ in Eq. (23) to get

$$f_0(0) = \frac{2}{\Gamma(K)}\frac{s}{\tilde{s}(K)} \times \int \tilde{\varphi}_0^*\left(\frac{\mathbf{k}}{2}\right)\tilde{\varphi}_0(\mathbf{k})(k^2 + K^2)f_0(\mathbf{k})\,d\mathbf{k}. \quad (24)$$

We then use some approximate $f_0(\mathbf{k})$ in this integral equation, to convert it into a transcendental one which relates the energy eigenvalue to potential strength.

The basic question then is: What is a good approximation for f_0? To get some notion of this we must have some idea of the general behavior of $\tilde{\varphi}_0(\mathbf{k})$. To be specific, we choose to work with an exponential interparticle potential since this is one for which there exist numerical calculations[5,6] to compare with our three-body results, and also because there are approximate two-body wavefunctions for it that are relatively tractable. We digress for a moment then to discuss these two-body functions.

The two-body potential is

$$v(r) = -se^{-r/d}$$

For this potential we shall use the approximate ground-state wavefunction $\tilde{\psi}_0(r)$ given by Morse and Feshbach,[7]

$$\tilde{\psi}_0(r) = \left[\frac{\alpha\beta(\alpha + \beta)}{2\pi(\beta - \alpha)^2}\right]^{\frac{1}{2}}\frac{1}{r}(e^{-\alpha r} - e^{-\beta r}),$$

where α and β are variational parameters chosen in a way we shall describe below. The corresponding radial wavefunction in momentum space is

$$\tilde{\varphi}_0(k) = \frac{1}{\pi}\left[\frac{\alpha\beta(\alpha + \beta)}{(\alpha - \beta)^2}\right]^{\frac{1}{2}}\left(\frac{1}{\alpha^2 + k^2} - \frac{1}{\beta^2 + k^2}\right). \quad (25)$$

According to a variational procedure, a fairly accurate relation between potential strength and binding energy obtains when α and β have the values

$$\alpha = \tilde{K}_0, \qquad \beta = d^{-1}[\tfrac{1}{2}(\tilde{K}_0 d + 1)]^{\frac{1}{2}},$$

which, as Morse and Feshbach show, leads to the relation

$$sd^2 = \tfrac{1}{4}\{(2\tilde{K}_0 d + 1)[1 + [2(1 + \tilde{K}_0 d)]^{\frac{1}{2}}]^2\}. \quad (26)$$

This result is, by comparison with the correct one, accurate to within a percent or so from $\tilde{K}_0 d = 0$ to $\tilde{K}_0 d = 1$. We shall use it only in this range, and with confidence only for $\tilde{K}_0 d$ less than unity, since the assumed $\tilde{\varphi}_0(k)$ behaves anomalously[8] near $\tilde{K}_0 d = 1$.

Now we return to Eq. (23). If we imagine (25) put into it we see that the function $\tilde{\varphi}_0^*(\frac{1}{2}\mathbf{k}+\boldsymbol{\kappa})\tilde{\varphi}_0(\mathbf{k}+\frac{1}{2}\boldsymbol{\kappa})$, which occurs in the integral on the right-hand side, is one which (for given $\mathbf{k}$) drops off at large $\boldsymbol{\kappa}$ as $1/\kappa^8$. Since the contributions to the integral come from a limited value of $\mathbf{k}$ one would then expect that $f_0(\boldsymbol{\kappa})$ itself drops off about as $1/\kappa^8$. In short, $f_0(\boldsymbol{\kappa})$ is a rather sharply peaked function of $\boldsymbol{\kappa}$ which, as in the one-dimensional case, makes it plausible to use a δ-function as a zeroth-order iterate. If then we put a δ-function into the right-hand side of Eq. (23) we get for $f_0^{(1)}(\boldsymbol{\kappa})$, the first iterate,

$$f_0^{(1)}(\boldsymbol{\kappa}) = [\Gamma(K) + \tfrac{3}{4}\kappa^2)^{-1}\varphi_0(\kappa)\varphi_0(\tfrac{1}{2}\kappa)(K^2 + \tfrac{1}{4}\kappa^2). \quad (27)$$

[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, New York, 1953), Vol. 2, p. 1698.

[8] It vanishes identically for $K_0 d = 1$, and becomes progressively inaccurate as $K_0 d$ increases from unity.

In this we have dropped irrelevant normalization constants, and also dropped the complex conjugation sign on f_0 since it is real. If now we put (27) into (24) we find after some algebra and simplification that it reduces to

$$1 = \frac{8}{\pi}\frac{\mu(1+\gamma)^3}{\gamma^3}\int_0^\infty \frac{y^2\,dy}{(1+y^2)(1+\frac{1}{4}y^2)(1+y^2/\gamma^2)(1+y^2/4\gamma^2)^2((1-\mu)(1+\gamma)+\frac{3}{4}y^2)}. \tag{28}$$

In this equation,

$$\mu = s/\bar{s}(K), \qquad \gamma = [\tfrac{1}{2}(1+Kd)]^{\frac{1}{2}}/Kd.$$

By definition $\bar{s}(K)$ is that two-body effective potential strength which would yield eigenvalue K for the two-body problem. By (26) then it is

$$\bar{s}(K)d^2 = \tfrac{1}{4}\{(2Kd+1)[1+[2(1+Kd)]^{\frac{1}{2}}]^2\}.$$

We have solved Eq. (28) numerically getting as a final result a relation between potential strength s and eigenvalue K, or as it turns out a relation between Kd and sd^2. The results are given in Table I along with what seem to be the quite accurate numerical results due to Kalos.

We add a few remarks on these results. First, they are rather less accurate than other numerical results in the literature[4,5] that have been achieved with one or another methods, but always with the help of digital computers. On the other hand, essentially all the work that is involved in our calculation is set out above, and the only computation necessary is in the evaluation of the integral in (28). We think that even with their imprecision, the results are good enough to lend confidence that the basic ideas of the method are correct.

It is perhaps worth spelling out the sources of inaccuracy in these results. There are three such. First is the approximation involved in the truncation of the coupled set of equations for $f_l(\kappa)$ in which we keep only the lowest-order one, Eq. (23). Then there is the further approximation involved in solving this by the iterative technique given above. Finally there is the approximation involved in Eq. (23) itself, in that we have for simplicity used in it the approximate two-body variational function given by Eq. (25). We have not yet tried to sort out the different orders of magnitudes of error implied by these separate approximations.

TABLE I. "Energy" eigenvalue Kd versus potential "strength" sd^2, for the exponential potential $v(r) = -se^{-r/d}$.

Kd	Sd^2	Sd^2 "Exact"
0.19	1.72	1.47
0.42	2.28	1.91
1.00	4.16	3.28

V. DISCUSSION

The method we have presented, although worked out in detail for the special cases above, is in fact applicable to any interparticle potential of finite range. All one needs to know is the general form of the two-body eigenfunctions $\tilde{\varphi}_l(k)$, and $\bar{s}(K_0)$, the functional dependence of the two-body energy eigenvalue on potential strength. Equation (8), or its three-dimensional counterpart which yields the energy eigenvalue, is then completely specified. Although in one dimension we have calculated only for δ-function potentials, there is no reason to expect these potentials to give uniquely good results with our method; we would expect then that one would get comparable results (eigenvalue and eigenfunction accurate to a few percent), for any one-dimensional potential with a finite range. Similarly, for three dimensions, one would think that the exponential potential we have used is a fairly typical one, so that one would expect results of comparable accuracy with any of the other short-range ones that one might use.

We have said nothing about the four and N-body problems, but it is clear that the general viewpoint presented here is also applicable to them. For the four-body problem one has, of course, to introduce further Jacobi coordinates, but having done that the decomposition of the wavefunction, and the expansion in terms of two-body functions with an appropriate "effective strength" goes much as for the three-body problem. Similar remarks apply to the N-body problem. In particular, preliminary results on the one-dimensional four-body problem with δ-function potentials are encouraging, by comparison with the exact solution.

ACKNOWLEDGMENT

The numerical solution of Eq. (15) was carried out at Parke Mathematical Laboratories, of whose staff I would particularly like to thank Dr. Charles Sherman for discussions and help.

APPENDIX

In this Appendix we derive some of the results quoted in the paper for the one-dimensional three-

body problem with δ-function potentials. The Schrödinger equation for it is

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}\right)\Psi+2s[\delta(x_1-x_2)+\delta(x_1-x_3)+\delta(x_2-x_3)]=K^2\Psi \quad \text{(A1)}$$

with $K^2=m|E|/\hbar^2$. As we have mentioned, McGuire[2] has shown by raytracing arguments that the solution of this equation is

$$\Psi=e^{-\frac{1}{2}s(|x_1-x_2|+|x_1-x_3|+|x_2-x_3|)}. \quad \text{(A2)}$$

We mention briefly how, apart from ray tracing, this solution can be verified. First one can trivially check that it satisfies Eq. (A1) for points away from the δ-function singularities, and has the eigenvalue $K=s$. For away from the singularities, say for the typical region $x_1<x_2<x_3$, the equation becomes

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}\right)\Psi=K^2\Psi$$

and Ψ becomes

$$\Psi=e^{-s(x_3-x_1)},$$

which obviously satisfies Eq. (A1). To complete this check we must verify that Eq. (A1) is satisfied when it is integrated through its δ-function singularities. Since both the wavefunction and potential are symmetric in x_1, x_2, x_3, it is clear that we need do this integration through only one of the singular potentials, say $\delta(x_1-x_2)$, since the other two will give identical results.

Consider then the singularity at $x_1-x_2=0$. For discussing this we introduce the Jacobi coordinates x, y, and X, (dropping for the moment their usual subscripts)

$$x=x_2-x_1,$$
$$y=x_3-\tfrac{1}{2}(x_1+x_2),$$
$$X=\tfrac{1}{3}(x_1+x_2+x_3).$$

In terms of these, Eq. (A1) becomes (on dropping a term in $\partial^2/\partial X^2$ which refers to the center-of-mass motion and is not of interest)

$$\left\{\frac{\partial^2}{\partial x^2}+\frac{3}{4}\frac{\partial^2}{\partial y^2}+s\left[\delta(x)+\delta\left(\frac{x}{2}+y\right)+\delta\left(\frac{x}{2}-y\right)\right]\right\}\Psi=K^2\Psi. \quad \text{(A3)}$$

In the x–y plane of these coordinates, the singularities are along the lines $x=0$, $\frac{1}{2}x=y$, $\frac{1}{2}x=-y$. These lines are plotted in Fig. 1.

Now we integrate Eq. (A3) in the x, y plane across

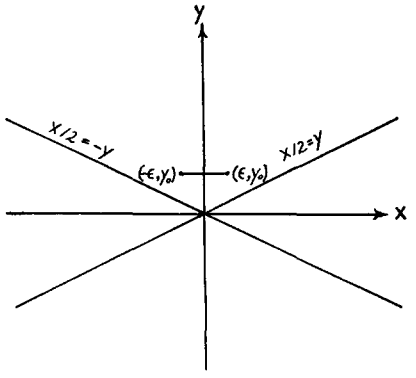


FIG. 1. The potential is singular along the three lines $x=0$, $\frac{1}{2}x=y$, $x/2=-y$.

the line $x=0$, from $(-\epsilon, y_0)$ to (ϵ, y_0), for some arbitrary y_0. We note that whatever y_0 we choose, we can always take ϵ small enough so that the condition

$$|\epsilon|<\tfrac{1}{2}y_0 \quad \text{(A4)}$$

obtains. After this integration, Eq. (A3) becomes

$$\begin{aligned}\Psi'(-\epsilon, y_0)-\Psi'(\epsilon, y_0)&+\frac{3}{4}\int_{-\epsilon}^{\epsilon}\left(\frac{\partial^2}{\partial y^2}\Psi(x,y)\right)_{y=y_0}dx\\&+s\int_{-\epsilon}^{\epsilon}\delta(x)\Psi(x,y_0)\,dx\\&+s\int_{-\epsilon}^{\epsilon}\delta\left(\frac{x}{2}+y_0\right)\Psi(x,y_0)\,dx\\&+s\int_{-\epsilon}^{\epsilon}\delta\left(\frac{x}{2}-y_0\right)\Psi(x,y_0)\,dx\\&=K^2\int_{-\epsilon}^{\epsilon}\Psi(x,y_0)\,dx.\end{aligned}$$

Most of the terms in this equation vanish in the limit $\epsilon\to 0$. The first integral on the left-hand side, and that on the right-hand side vanish because $\Psi(x, y)$ is continuous along the integration path; the terms with $\delta(\frac{1}{2}x+y_0)$ and $\delta(\frac{1}{2}x-y_0)$ in the integrand vanish because of the condition (A4). We are thus left with

$$\lim_{\epsilon\to 0}[\Psi'(-\epsilon, y_0)-\Psi'(\epsilon, y_0)]=-s\Psi(0, y_0)$$

and it is easy to see that our assumed wavefunction does satisfy this condition. This completes the verification that the Ψ of Eq. (A2) is a solution of Eq. (A1).

Although McGuire gives an expression for the N-body energy, he does not explicitly write down the N-body wavefunction. It is natural enough,

however, to guess at the generalization of (A2) to the N-body case as

$$\psi_N = e^{-\frac{1}{2}s(|x_N - x_{N-1}| + |x_N - x_{N-2}| + \cdots + |x_2 - x_1|)} . \quad \text{(A5)}$$

On assuming the ordering $x_1 < x_2 < x_3 \cdots < x_N$, we get, on differentiating the N particle Schrödinger equation,

$$K^2 = N(N^2 - 1)s^2/24. \quad \text{(A6)}$$

Moreover, it is easy to verify that the function (A5) satisfies the conditions got by integrating the Schrödinger equation across the δ-function singularities so that, in fact, it is a solution, and Eq. (A6) is its "energy" eigenvalue.

Functions as interesting as the wavefunctions Ψ themselves are the "orbitals" ψ discussed in the text. These are defined for the three-body problem, for example, as follows. The wavefunction Ψ satisfies the general integral equation, which we write in our slightly abstract notation

$$\Psi(\mathbf{P}) = \int \Psi(\mathbf{P}')G_K(\mathbf{P} - \mathbf{P}')v_t(\mathbf{P}')\, d\mathbf{P}'. \quad \text{(A7)}$$

For the present case, the total potential $v_t(\mathbf{P})$ is just

$$v_t(\mathbf{P}) = -s[\delta(x_1 - x_2) + \delta(x_1 - x_3) + \delta(x_2 - x_3)]. \quad \text{(A8)}$$

If we imagine (A8) put into (A7), we are led to three integrals, i.e., to three terms in the wavefunction. The ψ_{12} orbital is then defined to be that term which involves the integral over $\delta(x_1 - x_2)$, it being understood that both Ψ and G_K are to be expressed in x_{12}, y_3 coordinates for this purpose,

$$\psi_{12}(x_{12}, y_3) = -s \iint \Psi(x'_{12}, y'_3)\delta(x'_{12}) \times G_K(x_{12} - x'_{12}, y_3 - y'_3)\, dx'_{12}\, dy'_3. \quad \text{(A9)}$$

If then we put into (A7) the expression (A2) for $\Psi(x_{12}, y_3)$, use the integral representation for the Green's function,

$$G_K(x_{12} - x'_{12}, y_3 - y'_3) = \frac{1}{(2\pi)^2} \iint \frac{e^{i[k(x_{12} - x_{12}') + \kappa(y_3 - y_3')]}}{k^2 + \frac{3}{4}\kappa^2 + K^2}\, dk\, d\kappa,$$

and do the integration over dx'_{12}, we are led to

$$\psi_{12}(x_{12}, y_3) = \iiint \frac{e^{i[k'x_{12} + \kappa'(y_3 - y_3')] - K|y_3'|}}{k'^2 + \frac{3}{4}\kappa'^2 + K^2}\, dk'\, d\kappa'\, dy'_3. \quad \text{(A10)}$$

As we have pointed out, the subscripts in Eq. (A9) are really superfluous, since ψ_{12} is the same function of its variables as, say ψ_{13} of its variables. Thus it is convenient to drop all the subscripts in this equation and to let ψ and x, y stand, respectively, for any of the three orbitals and for the corresponding variables. If we do this in (A10) and do the integration over k' and y'_3, we get

$$\psi(x, y) \propto \int_{-\infty}^{\infty} \frac{\exp[i\kappa y - (K^2 + \frac{3}{4}\kappa^2)^{\frac{1}{2}}|x|]}{[k^2 + \kappa^2](K^2 + \frac{3}{4}\kappa^2)^{\frac{1}{2}}}\, d\kappa. \quad \text{(A11)}$$

This last integral seems to be difficult to do, but happily it is not very important to be able to do it since as usual we find it more convenient to work with $\varphi(k, \kappa)$, the Fourier transform of $\psi(x, y)$.

$$\varphi(k, \kappa) = \iint e^{i(kx + \kappa y)}\psi(x, y)\, dx\, dy.$$

If then we go back a step and Fourier transform ψ_{12} in its form (A9) we find that all the integrals can can be done and we get

$$\varphi(k, \kappa) = [(K^2 + \tfrac{3}{4}\kappa^2 + k^2)(K^2 + \kappa^2)]^{-1}. \quad \text{(A12)}$$

It is interesting to compare this result with the Fourier transform $\Phi(k, \kappa)$ of the total wavefunction Ψ. We express Ψ in x_{12}, y_3 variables

$$\Psi = \exp[-\tfrac{1}{2}K\{|x_{12}| + |\tfrac{1}{2}x_{12} + y_3| + |\tfrac{1}{2}x_{12} - y_3|\}] \quad \text{(A13)}$$

and define

$$\Phi(k, \kappa) = \iint \Psi(x_{12}, y_3)e^{i(kx_{12} + \kappa y_3)}\, dx_{12}\, dy_3. \quad \text{(A14)}$$

If we put the expression (A13) into this, the integrals can be done; but, in fact, it is easier to write

$$\Psi = \psi_{12}(x_{12}, y_3) + \psi_{13}(x_{13}, y_2) + \psi_{23}(x_{23}, y_1)$$

and then use the integral representation (A10) for ψ_{12}, etc. In either event one gets

$$\Phi(k, \kappa) = \frac{1}{K^2 + \frac{3}{4}\kappa^2 + k^2}\left\{\frac{1}{K^2 + \kappa^2} + \frac{1}{K^2 + (k - \frac{1}{2}\kappa)^2} + \frac{1}{K^2 + (k + \frac{1}{2}\kappa)^2}\right\}. \quad \text{(A15)}$$

The first term in this last expression is, of course, nothing more nor less than the Fourier transform of the ψ_{12} orbital. A comparison then of (A15) with (A12) points up the advantage in simplicity of dealing with φ, since φ defines the wavefunction quite as well as does Φ.

Concerning Space–Time, Symmetry Groups, and Charge Conservation*

E. C. G. SUDARSHAN
Physics Department, Syracuse University, Syracuse, New York
(Received 26 October 1964)

Under the restriction to a symmetry group whose generators can be linearly expressed in terms of the generators of the Lorentz group and of an internal symmetry group, we show that charge conservation implies that the symmetry group is a direct product of the Lorentz group and the internal symmetry group. Only local structure of the groups and their unitary representations are considered.

THE problem of combining relativistic invariance and internal symmetries has recently been discussed by several authors.[1] Since the multiplet masses are only approximately the same, the only way of reconciling exact invariance under both the internal symmetry group and the relativity group is to require that the elements of the two groups do not commute in general. Let us assume[2] that the full invariance group of the physical (strongly interacting) system is a Lie group (which has as subgroups the internal symmetry group, assumed to be simple, and the inhomogeneous Lorentz group) which has a Lie algebra A whose elements can be expressed as a linear combination of the elements of the Lie algebras of the internal and Lorentz groups. It can then be shown[3] that, if a complete set of commuting generators of the semisimple internal symmetry algebra S commute with the generators of the inhomogeneous Lorentz algebra L, then the algebra A is a direct sum of the algebras S and L.

In this paper we wish to extend this result to the case when only a single generator of S commutes with all elements of L. This framework is of particular interest since electric charge is conserved in all known interactions and it is a generator, or associated with a generator of the internal symmetry group which we expect to be relativistically invariant. We show that as long as we are interested in unitary representations only a direct sum algebra results; this is disappointing since there is then no possible explanation of the mass splittings within a multiplet compatible with exact invariance under the group.

Theorem. Let Q be any generator of S. If Q commutes with every element of L (and if the set of elements of S and L are closed under commutation), then every element L_A of L can be decomposed in the form with $L_A = L_A^0 + L_A^1$, L_A^1 being a linear combination of elements of S and L_A^0 commuting with every element of S. Further, both L_A^0 and L_A^1 satisfy the same commutation relations as the elements L_A of the Lorentz algebra.

Proof. Since the Lie algebra S is simple, there exists a Cartan–Weyl basis H_l, E_α such that

$$[H_l, \quad E_\alpha] = r_l(\alpha)E_\alpha, \tag{2a}$$

$$[H_l, \quad H_m] = 0, \tag{2b}$$

$$[E_\alpha, \quad E_\beta] = N_{\alpha\beta}E_{\alpha+\beta}, \qquad r(\alpha) + r(\beta) \neq 0, \tag{2c}$$

$$[E_\alpha, \quad E_{-\alpha}] = \sum_l r_l(\alpha)H_l, \tag{2d}$$

No generality is lost by taking H_1 to be a multiple of Q so that

$$[H_1, \quad L_A] = 0. \tag{3}$$

By hypothesis we can write

$$[E_\alpha, \quad L_A] = \sum_\beta a(\alpha A\beta)E_\beta + \sum_l a(\alpha Al)H_l + \sum_B a(\alpha AB)L_B.$$

Hence, if $r_1(\alpha) \neq 0$ we deduce

$$[E_\alpha, \quad L_A] = \sum_\beta a(\alpha A\beta)\delta\{r_1(\alpha) - r_1(\beta)\}E_\beta.$$

For those cases where $r_1(\alpha) = 0$, we can use (2c) to deduce the general expression

$$[E_\alpha, \quad L_A] = \sum_\beta a(\alpha A\beta)\delta\{r_1(\alpha) - r_1(\beta)\}E_\beta + \sum_m a(\alpha Am)\delta\{r_1(\alpha)\}H_m. \tag{4}$$

Similarly from (2b) and (3) we deduce

$$[H_l, \quad L_A] = \sum_\beta b(lA\beta)\delta\{r_1(\beta)\}E_\beta + \sum_m b(lAm)H_m. \tag{5}$$

* Supported in part by the United States Atomic Energy Commission.

[1] F. Lurçat and L. Michel, Nuovo Cimento **21**, 574 (1961); *Proceedings of the Coral Gables Conference on Symmetry Principles at High Energies*, edited by B. Kursunoglu and A. Perlmutter (W. J. Freeman Company, San Francisco, 1964); A. O. Barut, Nuovo Cimento **32**, 234 (1964); B. Kursunoglu, Phys. Rev. **135**, B761 (1964).

[2] Following W. D. McGlinn, Phys. Rev. Letters **12**, 469 (1964).

[3] F. Coester, M. Hamermesh, and W. D. McGlinn; Phys. Rev. **135**, B451 (1964); M. E. Mayer, H. J. Schnitzer, E. C. G. Sudarshan, R. Acharya, and M. Y. Han, Phys. Rev. **136**, B888 (1964). A Beskow and V. Ottoson, Nuovo Cimento **34**, 248 (1964).

Evaluating the double commutator

$$[H_l, [E_\alpha, \quad L_A]] = \sum_\beta a(\alpha A\beta) r_l(\beta)\delta\{r_1(\alpha) - r_1(\beta)\}E_\beta$$

using the Jacobi identity, we get the expression

$$\begin{aligned}[[H_l, \quad E_\alpha], L_A] - [[H_l, \quad L_A], E_\alpha] &= \sum_\beta a(\alpha A\beta) r_l(\alpha)\delta\{r_1(\alpha) - r_1(\beta)\}E_\beta \\ &+ \sum_m a(\alpha Am) r_l(\alpha)\delta\{r_1(\alpha)\}H_m \\ &- \sum_\beta b(lA_\beta)\delta\{r_1(\beta)\}[E_\beta, \quad E_\alpha] \\ &- \sum_m b(lAm) r_m(\alpha) E_\alpha.\end{aligned}$$

Comparing the terms in E_α in these two expressions, we get

$$\sum_m b(lAm) r_m(\alpha) = 0. \tag{6}$$

Since (6) is true for all α, and since the $r_m(\alpha)$ span an m-dimensional vector space as α varies over the allowed range, it follows that

$$b(lAm) = 0. \tag{7}$$

Two possible cases arise now.

(a) If H_1 is such that $r_1(\beta) \neq 0$ for any β it follows that[3]

$$[H_l, \quad L_A] = 0$$

and hence every element of S commutes with every element of L. Hence the statement of the theorem is trivially satisfied with

$$L_A^0 = L_A \cdot L_A^1 = 0.$$

(b) If there are some E_β for which $r_1(\beta) = 0$ we have only proved that

$$[H_l, \quad L_A] = \sum_\beta b(lA\beta)\delta\{r_1(\beta)\}E_\beta. \tag{8}$$

But making use of (2b) we can then show

$$\sum_\beta \{b(lA\beta) r_m(\beta) - b(mA\beta) r_l(\beta)\}\delta\{r_1(\beta)\}E_\beta = 0.$$

This relation entails the existence of numbers $q(A\beta)$, not necessarily nonzero, such that

$$b(lA\beta) = q(A\beta) r_l(\beta). \tag{9}$$

From (4), we can deduce

$$\begin{aligned}[[E_\alpha, \quad E_\beta], L_A] &= \sum_\gamma a(\beta A\gamma)\delta\{r_1(\beta) - r_1(\gamma)\}[E_\alpha, E_\gamma] \\ &- \sum_\gamma a(\alpha A\gamma)\delta\{r_1(\alpha) - r_1(\gamma)\}[E_\beta, \quad E_\gamma]\end{aligned}$$

which implies

$$a(\alpha+\beta\ A\ \alpha+\beta) = a(\alpha A\alpha) + a(\beta A\beta).$$

This requires the existence of numbers $p(Al)$, not necessarily nonzero, such that

$$a(\alpha A\alpha) = \sum_l p(Al) r_l(\alpha). \tag{10}$$

Now consider the quantity

$$L_A^0 = L_A + \sum_l p(Al) H_l + \sum_\beta q(A\beta) E_\beta. \tag{11}$$

It is then immediately verified that

$$[L_A^0, \quad H_l] = 0. \tag{12}$$

Consequently, we make use of (4) and (11) to write

$$[L_A^0, \quad E_\alpha] = \sum_\beta e(\alpha A_\beta) E_\beta + \sum_m e(\alpha Am) H_m;$$

we can deduce

$$\sum_\beta e(\alpha A_\beta)\{r_m(\alpha) - r_m(\beta)\}E_\beta + \sum_l r_m(\alpha) e(\alpha Al) H_l = 0,$$

so that

$$e(\alpha A\beta) = 0; e(\alpha Al) = 0.$$

Hence

$$[E_\alpha, \quad L_A^0] = 0. \tag{13}$$

If we now write the commutation relations of the Lorentz algebra in the form

$$[L_A, \quad L_B] = \sum_C \gamma_{AB}^C L_C, \tag{14}$$

we can rewrite it in the form

$$\begin{aligned}[L_A^0, \quad L_B^0] - \sum_C \gamma_{AB}^C L_C^0 &= [L_A - L_A^0, L_B - L_B^0] - \sum_C \gamma_{AB}^C (L_C - L_C^0).\end{aligned}$$

Since the expression on the left-hand side commutes with H_l, E_α while the right-hand side is linear in them, both sides must identically vanish; this gives the basic result

$$[L_A^0, \quad L_B^0] = \sum_C \gamma_{AB}^C L_C^0, \tag{15}$$

$$[L_A^1, \quad L_B^1] = \sum_C \gamma_{AB}^C L_C^1, \tag{16}$$

with

$$L_A^1 = L_A - L_A^0. \tag{17}$$

This concludes the demonstrations of the theorem.[4]

[4] After this work was completed, the author had the opportunity to learn that essentially the same results have been deduced by V. Ottoson, A. Kihlberg, and J. Nilsson, "Internal and Space–Time Symmetries," Phys. Rev. 137, B658 (1965). See also L. Michel, Phys. Rev. 137, B405 (1965).

We can deduce an important corollary from this theorem:

Corollary. In every unitary representation of the full symmetry algebra, L_A^1 must identically vanish; and hence we get a unitary representation of a direct sum of L and S only.

This follows since every unitary representation of the algebra S is the direct sum of irreducible finite dimensional unitary representations. Consequently, the quantities L_A^1 satisfying (16) have a unitary representation which is the direct sum of finite dimensional unitary representations. But the only such representations are trivial.[5]

We now make several remarks:

(1) The restriction to a simple group S can be easily relaxed to any semisimple group, the only requirement being that the generator Q must have nonvanishing "parts" in each of the simple algebras which occur in the direct sum decomposition of the semisimple algebra.

(2) The theorem is equally applicable if the Lorentz algebra L is replaced by any other Lie algebra, say the algebra of the Galilei group. In this case the corollary is no longer applicable since the Galilei group has nontrivial finite-dimensional unitary (nonfaithful) representations.

(3) We could interchange the roles of the internal symmetry algebra and the Lorentz algebra: if we require that any one element of the homogeneous Lorentz algebra, say M_{12}, commute with all elements of the symmetry algebra, then every element of the internal symmetry algebra could be expressed in the form

$$\begin{aligned} H_l &= H_l^0 + H_l^1, \\ E_\alpha &= E_\alpha^0 + E_\alpha^1, \end{aligned} \tag{18}$$

with H_l^1, E_α^1 being linear combinations of the elements of the Lorentz algebra, such that H_l^0, E_α^0 commute with every element of the Lorentz algebra.[6] We can then show these quantities satisfy the relations

$$\begin{aligned} [H_l^i,\ E_\alpha^k] &= \delta^{ik} r_l(\alpha) E_\alpha^i, \\ [H_l^i,\ H_m^k] &= 0, \\ [E_\alpha^i,\ E_{-\alpha}^k] &= \delta^{ik} \sum_l r_l(\alpha) H_l^i, \\ [E_\alpha^i,\ E_\beta^k] &= \delta^{ik} N_{\alpha\beta} E_{\alpha+\beta}^i, \end{aligned}$$

which generalize the commutation relations (2).

(4) In the demonstration above, H_l^1 and E_α^1 are linear combinations of those elements of L which commute with all elements of S. Hence if we require that the rotation subalgebra generated by M_{23}, M_{31}, M_{12} all commute with S, H_l^1 and E_α^1 must be linear combinations of the elements of L which are invariant under rotations. To see this we note that the decomposition (18) is unique since if

$$\begin{aligned} H_l &= H_l^0 + H_l^1 = H_l'^0 + H_l'^1, \\ E_\alpha &= E_\alpha^0 + E_\alpha^1 = E'^0_\alpha + E'^1_\alpha, \end{aligned}$$

then

$$H_l^0 - H_l'^0 = H_l'^1 - H_l^1$$

and

$$E_\alpha^0 - E'^0_\alpha = E'^1_\alpha - E_\alpha^1$$

must commute with every element of L but they are at the same time elements of L. Hence they belong to the center of L, which is trivial. Hence H_l^1 and E_α^1 are unique and hence must be invariant under rotations. But the only element of L invariant under rotations is the Hamiltonian (time translation generator); consequently the commutator of any two elements of H_l^1, E_α^1 vanish which implies, by virtue of (19), that they themselves vanish. Hence if the elements of S commute with space rotations, we get only a trivial direct sum structure.

The present work in conjunction with that of other authors imply the extreme difficulty of constructing a purely Lie algebra model of an exact symmetry involving mass splittings. Any such scheme would require for its success a Lie algebra whose elements cannot be expressed as linear sums of elements of the internal symmetry algebra and the Lorentz algebra only.

Note added in proof: A definitive theorem in this connection has been proved in L. S. O'Raifeartaigh, Phys. Rev. Letters **14**, 575 (1965).

ACKNOWLEDGMENTS

The essential part of this work was done during a visit to the Argonne National Laboratory. It is a pleasure to thank Professor M. Peshkin for his hospitality at the Argonne National Laboratory and Professor N. Svartholm for the opportunity to visit the Chalmers University of Technology. Very useful discussions with Professor M. Hamermesh, Dr. W. D. McGlinn, Dr. Jan Nilsson, Dr. M. Y. Han, and Professor M. E. Mayer are gratefully acknowledged.

[5] It is interesting to point out that Ottoson, Kihlberg, and Nilsson (Ref. 4) have considered nonunitary representations of the Lorentz group, relating the nonunitary nature to the instability of several members of each multiplet.

[6] This result has been deduced by Y. Tomozawa, "Internal Symmetry and the Poincaré Group," J. Math. Phys. **6**, 656 (1965).

Diffraction by a Smooth Transparent Object. II. Diffraction by a Cylindrical Cavity*

YUNG MING CHEN
Division of Mathematical Sciences, Purdue University, Lafayette, Indiana
(Received 26 October 1964)

The exact solution corresponding to the field produced by a time-harmonic line source in the presence of a circular cavity of radius $a(N < 1)$ is found and evaluated asymptotically for large ka. Upon comparing it with the geometric rays and the diffracted rays constructed in Chen, the coefficients, especially the diffraction coefficients, are thus determined.

1. INTRODUCTION

In the previous investigation of Chen,[1] the so-called geometrical theory of diffraction of Keller[2,3] is extended and applied to the problems of diffraction by a smooth transparent object of any shape. However, the diffraction coefficients for the case $N < 1$ ($N = k_2/k_1$, where k_2 is the propagation constant of the scatterer and k_1 is that of the homogeneous surrounding medium) have not yet been determined. The purpose of the present paper is to determine those diffraction coefficients. To do this, the exact solution corresponding to the field produced by a time-harmonic line source in the presence of a circular cavity of radius a ($N < 1$) is found and evaluated asymptotically for large k_2a. Upon comparing it with the geometric rays and the diffracted rays constructed by using the geometrical theory of diffraction in Chen,[1] the coefficients are thus determined.

2. GENERAL ANALYSIS

The equations satisfied by the field $u(r, \Omega)$ of a line source at $(r_0, 0)$ in the presence of the circular cavity (Fig. 2 of I) are given by (2.1)–(2.5). The unique solution of these equations is given by (2.7)–(2.10). The integrals of (2.10) can be evaluated by computing the residues at the simple poles of P_ν (2.9). (Figure 1 shows schematically the positions of the poles of P_ν.) However, the result obtained will not give a clear physical interpretation. In order to exhibit the interesting physical phenomena, we have to use several legitimate tricks.

* This research is supported by the Purdue Research Foundation, under XL Grant.

[1] Y. M. Chen, J. Math. Phys. 5, 820 (1964); hereafter referred to as I; Eq. (n.m) will indicate the mth equation of Sec. n in I.

[2] J. B. Keller, "A Geometrical Theory of Diffraction" [Symposium on Microwave Optics, Eaton Electronics Research Laboratory, McGill University, Montreal, Canada (1953)].

[3] J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).

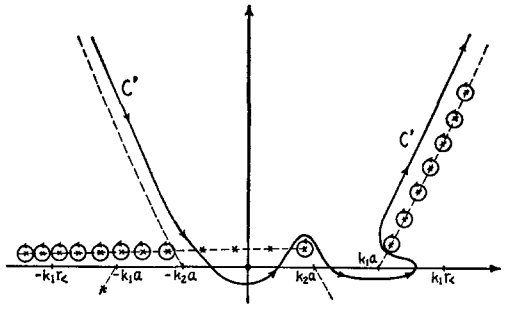


FIG. 1. The integration path C' and the positions of the simple poles of P_ν are shown schematically.

3. DIFFRACTED FIELD

We deform the integration path of (2.10) into C' (Fig. 1). Hence

$$u(r, \Omega) = \sum_{n=-\infty}^{\infty} \left\{ \frac{i}{4} \int_{C'} e^{i\nu(\Omega + 2\pi n)} (\mathfrak{I}_{10} - P_\nu \mathfrak{I}_{11})\, d\nu + \sum_s R^0_{ns} + \sum_s R^1_{ns} + \sum_s R^2_{ns} \right\}, \tag{1}$$

where R^0_{ns}, R^1_{ns}, and R^2_{ns} are the residues obtained from the simple poles with $\mathrm{Re}\,\nu < -k_2a$, $\mathrm{Re}\,\nu \sim k_2a$, and $\mathrm{Re}\,\nu > k_1a$, respectively; $\mathfrak{I}_{10}$ and $\mathfrak{I}_{11}$ are given by (2.7) and (2.8), respectively.

By using the proper asymptotic forms for those Bessel functions and Hankel functions in R^2_{ns}, we obtain

$$R^2_{ns} \simeq u^{D2}_{n\pm}(r, \Omega) = \mathfrak{N} D^2_{11s} [(r^2 - a^2)(r_0^2 - a^2)]^{-\frac{1}{4}} \times \exp\left\{ ik_1 [(r^2 - a^2)^{\frac{1}{2}} + (r_0^2 - a^2)^{\frac{1}{2}}] + i\nu_s \left(2\pi n \pm \Omega - \cos^{-1}\frac{a}{r} - \cos^{-1}\frac{a}{r_0} \right) \right\} + O\left[\frac{N}{k_1 a (1 - N^2)} \right], \tag{2}$$

where

$$D^2_{11s} = e^{-\frac{1}{3}i\pi}\left(\frac{2\pi}{k_1}\right)^{\frac{1}{3}} \times \left\{\frac{\alpha(1-N^4)^{\frac{1}{3}}e^{-\frac{1}{3}i\pi}A'(q_s e^{\frac{1}{3}i\pi})}{\beta[(1-N^2)^{\frac{1}{3}} - N^2(1+N^2)^{\frac{1}{3}}]A'(q_s)} - \left(\frac{k_1 a}{6}\right)^{\frac{1}{3}}\left[\frac{A(q_s e^{\frac{1}{3}i\pi})}{A'(q_s)}\right]\right\}, \quad (3)$$

$$\nu_s = k_1 a + q_s(\tfrac{1}{6}k_1 a)^{\frac{1}{3}}e^{\frac{1}{3}i\pi} + \cdots \quad (4)$$

q_s is a number determined by the following equation:

$$[1-(\nu_s/k_2 a)^2]^{\frac{1}{2}}\tan[(k_2^2a^2-\nu_s^2)^{\frac{1}{2}} - \nu_s\cos^{-1}\nu_s/k_2 a - \tfrac{1}{4}\pi] = \alpha e^{\frac{1}{3}i\pi}(6/k_1 a)^{\frac{1}{3}}A'(q_s)/A(q_s), \quad (5)$$

with $A(q)$ being the Airy function, and

$$\mathfrak{N} = (i/8\pi k_1)^{\frac{1}{2}}. \quad (6)$$

The $u^{D2}_{n+}(r, \Omega)$ and $u^{D2}_{n-}(r, \Omega)$ respectively represent the portions of incident wave which hit the upper and lower surface of the cavity tangentially. They decay exponentially while they travel along the interface on the side of medium 1 clockwise an angular distance $[2\pi n + \Omega - \cos^{-1}(a/r) - \cos^{-1}(a/r_0)]$ and counterclockwise an angular distance $[2\pi n - \Omega - \cos^{-1}(a/r) - \cos^{-1}(a/r_0)]$, respectively. Finally, they leave the surface tangentially towards the observing point (Fig. 4 of I). Upon comparing (2) with (1.38), we find that D_{11s} is the diffraction coefficient from medium 1 to medium 1 for the mode s.

Similarly, we have

$$R^1_{ns} \simeq u^{D1}_{n\pm}(r, \Omega) = \mathfrak{N}D_{12s}D_{21s}[(r^2-N^2a^2) \times (r_0^2-N^2a^2)]^{-\frac{1}{4}}\exp\left\{ik_1[(r^2-N^2a^2)^{\frac{1}{2}} + (r_0^2-N^2a^2)^{\frac{1}{2}} - 2a(1-N^2)^{\frac{1}{2}}] + i\tilde{\nu}_s\left(2\pi n \pm \Omega - \cos^{-1}\frac{Na}{r_0} - \cos^{-1}\frac{Na}{r_0} + 2\cos^{-1}N\right)\right\} + O\left[\frac{N}{k_1 a(1-N^2)}\right], \quad (7)$$

where

$$D_{12s}D_{21s} = \left(\frac{2\pi}{k_1}\right)^{\frac{1}{3}}e^{-\frac{1}{3}i\pi}\frac{2\beta N}{\alpha}(1-N^2)^{-\frac{1}{2}}, \quad (8)$$

$$\tilde{\nu}_s = k_2 a + \tilde{q}_s(\tfrac{1}{6}k_2 a)^{\frac{1}{3}}e^{-\frac{1}{3}i\pi} + \cdots, \quad (9)$$

and $\tilde{q}_s$ is determined by the following equation:

$$\beta N\left(\frac{6}{k_2 a}\right)^{\frac{1}{3}}e^{\frac{1}{3}i\pi}\left[\frac{e^{\frac{1}{3}i\pi}A'(\tilde{q}_s e^{-\frac{1}{3}i\pi}) + A'(\tilde{q}_s)}{e^{-\frac{1}{3}i\pi}A(\tilde{q}_s e^{-\frac{1}{3}i\pi}) + A(\tilde{q}_s)}\right] = i\alpha\left[1-\left(\frac{\tilde{\nu}_s}{k_1 a}\right)^2\right]^{\frac{1}{2}}. \quad (10)$$

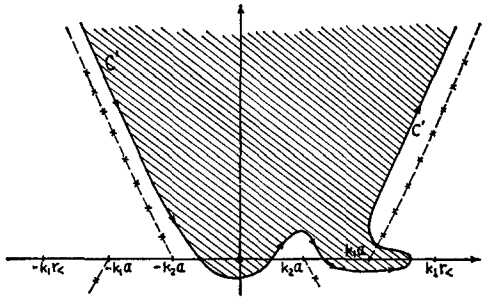


FIG. 2. The integration path C' and the positions of the simple poles of each term of the expanded P_ν are shown schematically.

The $u^{D1}_{n+}(r, \Omega)$ and $u^{D1}_{n-}(r, \Omega)$, respectively, represent the portions of incident wave which hit the upper and lower surface of the cavity at the critical angle of incidence. They decay exponentially while they travel along the interface on the side of medium 2 clockwise an angular distance $[2\pi n + \Omega - \cos^{-1}(Na/r) - \cos^{-1}(Na/r_0) + 2\cos^{-1}N]$ and counterclockwise an angular distance $[2\pi n - \Omega - \cos^{-1}(Na/r) - \cos^{-1}(Na/r_0) + 2\cos^{-1}N]$, respectively. Finally, they leave the surface at the critical angle of incidence towards the observing point. Upon comparing (7) with (1.34), we find that D_{12s} and D_{21s} are, respectively, the diffraction coefficients from medium 1 to medium 2 and from medium 2 to medium 1 for the mode s.

Because of $\mathfrak{I}_{11} \sim O[\nu^{-1}(\nu^{-2})^{\nu}]$ for $\mathrm{Re}\,\nu < -k_2 a$, all of the R^0_{ns} are of exponentially small in comparison with R^1_{ns} and R^2_{ns}. Hence R^0_{ns} can be neglected.

4. GEOMETRIC OPTICS FIELD

The geometric optics field comes from the asymptotic evaluation of the integrals appearing in (1) by saddle-point method. To show this, we must expand P_ν into a geometric series (2.11). It is important to notice that each term of the integrand of the new representation has no singularity on C' and in the shaded region of Fig. 2. After using the correct asymtpotic forms for the cylinder functions, we find that the expressions of the saddle-point equations are exactly the same as those for the case $N > 1$. Finally, we also find that the expressions of the geometric fields are exactly the same as those for the case $N > 1$.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Professor Joseph B. Keller of NYU for his encouragement.

Scalar Invariants of a Rotational System in a Lie Algebra. II. Casimir Operators

M. C. PEASE
Stanford Research Institute, Menlo Park, California
(3 December 1964)

We consider matrix equations of the form $d\mathbf{W}/dz = [\mathbf{S}, \mathbf{W}]$ where $\mathbf{S}(z)$ is a matrix-valued function of z, embedded in a given semisimple Lie algbra, L. If the set $\mathbf{X}_i$ is a basis for L, we expand $\mathbf{W}$ as $w^i\mathbf{X}_i$. We obtain general conditions on an n-index form $p_{ijk}\ldots$ such that $p = (u^iv^jw^k\cdots p_{ijk}\ldots)$ is a constant if u^i, v^j, w^k, etc., are solutions of the differential equation. It is shown that these forms are those that determine the vector invariants of the corresponding group.

We also ask what transformations will map any solution of the differential equation onto another, regardless of the detailed behavior of $\mathbf{S}(z)$. It is shown that the transformations that do this involve these same $p_{ijk}\ldots$.

With this identification, we are able to obtain procedures that can be used to calculate at least the two-index forms, p_{ij}. The existence of two-index forms, other than the Killing metric g_{ij} is shown to depend on the reducibility of the adjoint representation.

Finally, we show that these same forms generate Casimir operators—i.e., operators that commute with all elements of the algebra, and hence with the group associated with it. The Casimir operator so obtained is given explicitly by $(p_{ijk}\ldots\mathbf{X}^i\mathbf{X}^j\mathbf{X}^k\cdots)$, where the set, $\mathbf{X}^i$, is the basis of the dual algebra.

I. INTRODUCTION

In a previous paper[1] we considered matrix equations of the form

$$d\mathbf{W}/dz = [\mathbf{S}, \mathbf{W}] = \mathbf{SW} - \mathbf{WS}, \tag{1}$$

where $\mathbf{S}$ and $\mathbf{W}$ are differentiable matrix-valued functions of z that describe curves in a semisimple Lie algebra,[2-5] L, over a field of characteristic 0. We pointed out there the importance of this equation. In quantum mechanics[6] an equation of this form with t instead of z, describes the time dependence of the density matrix if $\mathbf{S}$ is $(1/ih)$ times the Hamiltonian. We also obtain Eq. (1) in coupled mode theory if we transform to the power density matrix.

In our previous paper, we defined what we called the e-forms, which are generated from the structure constants, c^k_{ij}, of the algebra by the cyclicly symmetric relations

$$e_{ijk\cdots m} = c^r_{iv}c^s_{jr}c^t_{ks}\cdots c^v_{mu}. \tag{2}$$

(We will use the summation convention throughout.)

We also defined the g-forms by symmetrization of the e-forms:

$$g_{ij} = e_{ij} = c^u_{iv}c^v_{ju},$$

$$g_{ijk} = e_{ijk} + e_{ikj},$$

$$g_{ijkh} = e_{ijkh} + e_{ijhk} + e_{ikhj} + e_{ikjh} + e_{ihjk} + e_{ihkj}. \tag{3}$$

We showed, there, that these forms generate conservation laws, or scalar invariants, of the system described by Eq. (1). If $\mathbf{U}$, $\mathbf{V}$, $\cdots$ are solutions of Eq. (1), which we express on the basis, $\mathbf{X}_i$, of the algebra as

$$\mathbf{U} = u^i\mathbf{X}_i,$$

$$\mathbf{V} = v^i\mathbf{X}_i, \tag{4}$$

$$\text{etc.},$$

then

$$p^{(n)} = (u^iv^i\cdots)e_{ij}\ldots \tag{5}$$

or

$$p^{(n)} = (u^iv^i\cdots)g_{ij}\ldots \tag{6}$$

are forms on the components that are independent of z. In particular,

$$p^{(2)} = u^iv^je_{ij} = u^iv^jg_{ij} \tag{7}$$

is the Killing form of $\mathbf{U}$ and $\mathbf{V}$, $\langle\mathbf{U}, \mathbf{V}\rangle$.

In this paper, we further develop the theory of these forms. We consider a general linear homogenous form of the nth degree, and obtain the condition that it shall be conserved by the system of

[1] M. C. Pease, J. Math. Phys. 6, 111(1965).
[2] Morton Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).
[3] P. M. Cohn, *Lie Groups*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 46 (Cambridge University Press, Cambridge, England, 1961).
[4] E. B. Dynkin, "The Structures of Semi-Simple Algebras," Usp. Math. Nauk (N. S.) 2, No. 4 (20), (1947), Am. Math. Soc. Transl. No. 17 (1950).
[5] Nathan Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962).
[6] A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1961), Vol. I, p. 331 ff.

Eq. (1), and so obtain a generalization of the previous results.

The equations that determine the operators that generate these forms are those that determine the vector invariants[7] of the group. Hence we are in fact dealing with the vector invariants.

We then show that these forms can be obtained by a quite different approach. If we ask what transformations will carry one solution of Eq. (1) into another, regardless of the detailed behavior of $\mathbf{S}(z)$, we find that we obtain the same operators as are involved in the conservation laws. To obtain the higher-order operators, we replace $\mathbf{W}$, in Eq. (1), by a matrix valued function that is the direct product of solutions of Eq. (1). We obtain a differential equation that is the equivalent of Eq. (1) for these higher-order solutions. We can then ask what transformations of these solutions will lead to a solution, independent of $\mathbf{S}(z)$. We shall find that the transformations that satisfy this condition involve, again, the same operators as generate conservation laws.

The relation between the conservation laws and the transformations gives us means for determing the operators that are involved. The transformations can be interpreted in two ways. They can be regarded as transformations of the coefficients in the expansion of $\mathbf{W}$, for example, on the given basis. Or they can be regarded as transformations of the basis of $\mathbf{W}$, but not of $\mathbf{S}(z)$. Both points of view will be exploited to develop the operators and to find some of their properties.

Finally, we shall show that these operators have still a third significance. The conditions on them are sufficient to generate Casimir operators[8] of the group of which the Lie algebra is the set of infinitesimal transformations. Thus they serve to characterize the representations of the group.

II. THE ALGEBRA AND ITS DUAL

Before tackling the problems outlined above, we need to establish the symbolism we shall use, and to summarize the underlying concepts.

As in our previous paper, we are considering a Lie algebra of $n \times n$ matrices over a field of characteristic 0, although the results can be applied equally to the corresponding abstract algebra. We do, however, assume the algebra to be finite dimensional.

We assume that a basis $\mathbf{X}_i$ has been chosen, so that the structure constants c_{ij}^k defined by

$$[\mathbf{X}_i, \mathbf{X}_j] = c_{ij}^k \mathbf{X}_k \tag{8}$$

are known. These constants are skew-symmetric in the lower indices and obey the Jacobi identity

$$c_{ij}^k = -c_{ji}^k, \tag{9}$$

$$c_{ij}^s c_{sk}^t + c_{jk}^s c_{si}^t + c_{ki}^s c_{sj}^t = 0. \tag{10}$$

An arbitrary element $\mathbf{A}$ of the algebra can, then, be expanded in terms of the basis

$$\mathbf{A} = a^i \mathbf{X}_i. \tag{11}$$

Between two such elements, $\mathbf{A}$ and $\mathbf{B}$, we can write the "Killing form," or the "scalar product of Cartan," as

$$\langle \mathbf{A}, \mathbf{B} \rangle = a^i b^j c_{iu}^v c_{jv}^u = a^i b^j g_{ij}. \tag{12}$$

In our previous paper,[1] we generalized the operators involved here by defining the e-forms given here in Eq. (2), so that

$$e_{ij} = g_{ij},$$

$$e_{ijk} = c_{iw}^u c_{ju}^v c_{kv}^w, \quad \text{etc.}$$

The key property of the e-forms was given in Lemma 1, in the previous paper, which states that, with the e-forms so defined,

$$e_{i_1 i} e_{i_2 \ldots} = e_{i_1 i \ldots} - e_{i i_2 \ldots} . \tag{13}$$

We also defined what we called the g-forms by symmetrization of the e-forms.

We are here assuming that the algebra L is semisimple. By Cartan's criterion, g_{ij} is then nonsingular, so that there exists a unique operator g^{ij} such that

$$g^{ij} g_{jk} = \delta_k^i. \tag{14}$$

The operator g^{ij} is also symmetric.

We now define the dual basis $\mathbf{X}^i$ by

$$\mathbf{X}^i = g^{ij} \mathbf{X}_j. \tag{15}$$

It follows immediately that

$$\mathbf{X}_i = g_{ij} \mathbf{X}^j. \tag{16}$$

We can show without difficulty that the set $\mathbf{X}^i$ are also the basis for a semisimple Lie algebra. Its structure constants are given by

$$[\mathbf{X}^i, \mathbf{X}^j] = c_k^{ij} \mathbf{X}^k. \tag{17}$$

If we substitute Eq. (15) in Eq. (17) and compare Eq. (8), we find that

$$c_k^{ij} = g^{iu} g^{jv} g_{kw} c_{uv}^w.$$

This result, however, can be simplified:

---

[7] H. Weyl, *Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946.)

[8] G. Racah "Sulla Caratterizzazione delli rappesentazioni irriducibili dei gruppi semicimplici di Lie" Rend. Acad. Lincei Sci. Fiz. Mat. Nat. 8, 108 (1950).

Lemma 2. The structure constants of the dual algebra are given by

$$c_k^{ij} = g^{it}c_{lk}^{i} = g^{it}c_{kl}^{j}. \tag{18}$$

This is proven by Lemma 1 [Eq. (13)], and the cyclic symmetry of the e-forms:

$$\begin{aligned} c_k^{ij} &= g^{iu}g^{jv}g_{kw}c_{uv}^{w} \\ &= g^{iu}g^{jv}(e_{vuk} - e_{uvk}) \\ &= g^{iu}g^{jv}(e_{kvu} - e_{vku}) \\ &= g^{iu}g^{jv}c_{vk}^{w}g_{wu} \\ &= g^{iv}c_{vk}^{w}\delta_w^{i} = g^{iv}c_{vk}^{i}. \end{aligned}$$

The alternate form of Eq. (18) is obtained similarly. Other relations between the normal and dual forms can be worked out similarly. Those given, however, are the ones we shall need.

III. CONSERVATION LAWS

We shall begin by considering the conservation laws of the second degree. If, as before, **U** and **V** are solutions of Eq. (1), which we express as in Eq. (4), we ask what is the condition on p_{ij} such that the scalar

$$p^{(2)} = u^i v^j p_{ij} \tag{19}$$

shall be independent of z.

Since $\mathbf{S}(z)$ is in the algebra, we can write it also as $\mathbf{S} = s^i\mathbf{X}_i$ and Eq. (1) becomes

$$dw^i/dz = s^m w^n c_{mn}^{i}. \tag{20}$$

Differentiating Eq. (19) and using Eq. (20), we find that we must have

$$dp^{(2)}/dz = s^m u^n c_{mn}^{i} v^j p_{ij} + u^i s^m v^n c_{mn}^{j} p_{ij} = 0.$$

Relabeling the dummy indices, this becomes

$$s^m u^i v^j (c_{mi}^{t} p_{tj} + c_{mj}^{t} p_{it}) = 0.$$

We wish this equation to vanish for any s^m and any initial values of u^i and v^j. Proving that the coefficients of $\mathbf{S}(z)$ obey the Lipschitz condition, we must have

$$c_{mi}^{t} p_{tj} + c_{mj}^{t} p_{it} = 0 \tag{21}$$

for all m, i, and j.

We note immediately that Eq. (22) holds for $p_{ij} = g_{ij}$, as may be proven by Lemma 1. On the other hand, Eq. (22) may also hold for other p_{ij}. For example, using the realization of the Lorentz algebra of the previous paper, Eq. (22) holds also for

$$p_{14} = p_{25} = p_{36} = p_{14} = p_{25} = p_{36} = 1, \tag{22}$$

which appeared there as one component of g_{ijkl}.

We can, now, manipulate Eq. (21) with the skew-symmetry of the c_{ij}^{k}, to obtain a sequence of relations:

$$\begin{aligned} c_{mi}^{t} p_{tj} &= c_{jm}^{t} p_{it} = -c_{mj}^{t} p_{it} \\ &= -c_{ji}^{t} p_{tm} = c_{ij}^{t} p_{tm}. \end{aligned}$$

Continuing in this way, we find that

$$\begin{aligned} c_{ij}^{t} p_{tk} &= c_{jk}^{t} p_{ti} = c_{ki}^{t} p_{tj} \\ &= c_{ij}^{t} p_{kt} = c_{jk}^{t} p_{it} = c_{ki}^{t} p_{jt}. \end{aligned} \tag{23}$$

It follows, then, that p_{ij} is either symmetric, or all its inner products with the structure contents vanish.

We now consider the more general case. We define

$$p^{(n)} = (u^i v^j w^k \cdots) p_{ijk\ldots} \tag{24}$$

and ask when this is a conservation law of Eq. (1). Differentiating $p^{(n)}$ by Eq. (20) and setting the derivative equal to zero, we obtain

$$\begin{aligned} dp^{(n)}/dz = {} & (s^m u^n c_{mn}^{i} u^j w^k \cdots) p_{ijk\ldots} \\ & + (u^i s^m v^n c_{mn}^{j} w^k \cdots) p_{ijk\ldots} \\ & + (u^i v^j s^m w^n c_{mn}^{k} \cdots) p_{ijk\ldots} \\ & + \cdots \\ = {} & (s^m u^i v^j w^k \cdots)(c_{mi}^{t} p_{tjk\ldots} + c_{mj}^{t} p_{itk\ldots} \\ & + c_{mk}^{t} p_{ijt\ldots} + \cdots) = 0. \end{aligned} \tag{25}$$

Again, if Eq. (25) is to hold for all possible s^m, u^i, v^j, etc., we must have

$$c_{mi}^{t} p_{tjk\ldots} + c_{mj}^{t} p_{itk\ldots} + c_{mk}^{t} p_{ijt\ldots} + \cdots = 0. \tag{26}$$

Again, using Lemma 1, we find that Eq. (26) is satisfied by $e_{ijk\ldots}$ and by $g_{ijk\ldots}$. It is, however, a much more general relation.

We note, finally, that the proof given establishes Eq. (26), or Eq. (21), as both a necessary and sufficient condition for the conservation law.

Equation (21) or Eq. (26) may be recognized as that determining the vector invariants of the group. That is, $p^{(2)}$ or $p^{(n)}$, from Eq. (19) or Eq. (24) are scalar-valued functions of the vectors u^i, v^j, etc. We consider the group of which L is the algebra and ask what is the condition on p_{ij} or $p_{ij}\ldots$ such that $p^{(2)}$ or $p^{(n)}$ is invariant if each vector is transformed by the group. This condition is expressed by Eq. (21) or (26).

We should note one qualification here. One normally defines the vector invariants by

$$p^{(n)} = (u^i u^j u^k \cdots) p_{ijk\ldots}$$

and obtains the forms of Eq. (19) or (26) by com-

plete polarization of this form. Doing so necessarily restricts the attention to those $p_{ijk}\ldots$ which are fully symmetric in the indices. The development given here broadens the possibilities to include all forms, whether symmetric or not.

If $p_{ijk}\ldots$ is antisymmetric in any pair of indices, then the scalar function $(u^iu^j\cdots p_{ij}\ldots)$ is identically zero, and so its conservation is trivial. However, the scalar function $(u^iv^jw^k\cdots p_{ijk}\ldots)$, where u^i, v^j, w^k, etc., are distinct solutions, is not trivial. The statement that it is conserved gives us an invariant relation between distinct solutions. If, for example, the various initial conditions on the solutions are such that this form vanishes, then it remains zero, and we have, in effect, an invariant orthogonality condition among the solutions.

IV. TRANSFORMATIONS AMONG SOLUTIONS

We now consider another problem which may seem unrelated but which will turn out to involve the same operators, $p_{ij}\ldots$. Again, we shall consider first the simplest situation—which will lead to p_{ij}—and then generalize to the higher-order situations.

We ask how we can transform on solution of Eq. (1) into another with the same $\mathbf{S}(z)$, independently of the detailed form of $\mathbf{S}(z)$.

We define p_{uv} so that

$$u^i = g^{iu}p_{uv}w^v \tag{27}$$

and require that u^i be a solution of Eq. (20). We require, then, that

$$du^i/dz = g^{iu}p_{uv}\,dw^v/dz = s^mu^nc^i_{mn}. \tag{28}$$

Using Eqs. (20) and (27), Eq. (28) becomes

$$g^{iu}p_{uv}s^mw^nc^v_{mn} = s^mc^i_{mn}g^{nu}p_{uv}w^v. \tag{29}$$

By Lemma 2, Eq. (18), the right-hand side of Eq. (29) can be written

$$s^mc^u_{nm}g^{in}p_{uv}w^v = s^mg^{iu}c^v_{um}p_{vn}w^n \tag{30}$$

by relabeling the dummy indices. This is to be true for all s^m and initial w^v. Also g^{iu} is nonsingular. Hence, equating Eq. (30) to the left side of Eq. (29), we find that the necessary and sufficient condition is that

$$c^v_{mn}p_{uv} = c^v_{um}p_{vn},$$

which is the same as Eq. (21). The operator p_{ij} that is involved in Eq. (22) also generates the conservation law of Eq. (19).

To extend this analysis to the higher-order operators, we consider the function

$$\mathbf{M} = m^{ijk\cdots}\mathbf{X}_{ijk}\ldots = (u^iv^jw^k\cdots)(\mathbf{X}_i\mathbf{X}_j\mathbf{X}_k\cdots)$$

so that $\mathbf{M}$ is the direct product of solutions of Eq. (1).

We find, then, that

$$\begin{aligned} d\mathbf{M}/dz &= (dm^{ijk\cdots}/dz)\mathbf{X}_i\mathbf{X}_j\mathbf{X}_k\cdots \\ &= (s^mu^nc^i_{mn}v^jw^k\cdots)(\mathbf{X}_i\mathbf{X}_j\mathbf{X}_k\cdots) \\ &\quad + (u^is^mv^nc^j_{mn}w^k\cdots)(\mathbf{X}_i\mathbf{X}_j\mathbf{X}_k\cdots) \\ &\quad + (u^iv^js^mw^nc^k_{mn}\cdots)(\mathbf{X}_i\mathbf{X}_j\mathbf{X}_k\cdots) \end{aligned}$$

or that

$$\begin{aligned} dm^{ijk\cdots}/dz &= s^mc^i_{mn}m^{njk\cdots} \\ &\quad + s^mc^j_{mn}m^{ink\cdots} \\ &\quad + s^mc^k_{mn}m^{ijn\cdots} + \cdots. \end{aligned} \tag{31}$$

We consider now Eq. (31) where $m^{ijk\cdots}$ may or may not be the direct product of solutions of Eq. (1). That is, we have used direct products to obtain the form of Eq. (31), but now we use Eq. (31) without restriction.

We consider the transformation

$$n^{abc\cdots} = (g^{am}g^{bn}g^{cp}\cdots)p_{mnp\cdots uvw\cdots}m^{uvw\cdots}, \tag{32}$$

and ask what condition on $p_{mnp\cdots uvw\cdots}$ will assure us that $n^{abc\cdots}$ will satisfy an equation of the form of Eq. (31).

We should note that we do not require $n^{abc\cdots}$ to be of the same degree as $m^{uvw\cdots}$.

We see that

$$\begin{aligned} dn^{ab\cdots}/dz &= (g^{am}g^{bn}\cdots)p_{mn\cdots uv\cdots}\,dm^{uv\cdots}/dz \\ &= (g^{am}g^{bn}\cdots)p_{mn\cdots uv\cdots}\{s^tc^u_{tr}m^{rv\cdots} \\ &\qquad + s^tc^v_{tr}m^{ur\cdots}\}. \end{aligned} \tag{33}$$

We want this to equal

$$\begin{aligned} &s^tc^a_{tr}n^{rb\cdots} + s^tc^b_{tr}n^{ar\cdots} + \cdots \\ &\quad = s^tc^a_{tr}(g^{rm}g^{bn}\cdots)p_{mn\cdots uv\cdots}m^{uv\cdots} \\ &\qquad + s^tc^b_{tr}(g^{am}g^{rn}\cdots)p_{mn\cdots uv\cdots}m^{uv\cdots} + \cdots \\ &\quad = s^tc^m_{rt}(g^{ar}g^{bn}\cdots)p_{mn\cdots uv\cdots}m^{uv\cdots} \\ &\qquad + s^tc^n_{rt}(g^{am}g^{br}\cdots)p_{mn\cdots uv\cdots}m^{uv\cdots} \end{aligned} \tag{34}$$

where, in the last step, we have used Lemma 2.

Equating Eq. (34) to Eq. (33), relabeling the dummy indices, and letting s^t be any curve and $m^{uv\cdots}$ have any initial value, we find that we must have

$$\begin{aligned} &c^s_{tm}p_{sn\cdots uv\cdots} + c^s_{tn}p_{ms\cdots uv\cdots} \\ &\quad + \cdots + c^s_{tu}p_{mn\cdots sv\cdots} + c^s_{tv}p_{mn\cdots us\cdots} + \cdots = 0 \end{aligned}$$

This is the same as Eq. (26).

We see, then, that the p-operators satisfying Eqs. (21) or (26) not only give us conservation laws of the system described by Eq. (1), but, when an appropriate number of their indices are raised as in Eqs. (27) or (32), they also give us transformations among solutions of either Eq. (1) or Eq. (31).

V. THE PROPERTIES OF p_{ij}

The recognition that $P_j^i = g^{it}p_{tj}$ is an operator that will transform among solutions of Eq. (1) gives us a useful device for studying the properties of p_{ij}.

We consider Eq. (20). We can express w^i as a column matrix whose dimensionality equals that of the algebra. The operator $S_n^i = s^m c_{mn}^i$ is, then, a matrix valued function of z. Likewise, in Eq. (27), we can write $P_s^i = g^{iu}p_{us}$ as a square matrix.

If, now, we follow through the steps of the preceding section, we find that for $u^i = P_j^i w^j$ to be a solution, we required that P_j^i shall commute with any possible S_j^i,

$$P_j^i S_k^j = S_j^i P_k^j. \tag{35}$$

Now Eq. (35) must be true for any S_j^i in the algebra. In particular it must be true when $\mathbf{S}$ is one of the basis matrices. If $\mathbf{S}$ is $\mathbf{X}_{(i)}$, where i now has some particular value, then $S_v^u = c_{(i)v}^u$—i.e., is the matrix whose uvth term is $c_{(i)v}^u$. This is a representation of the Ad operator,

$$Ad_{X_{(i)}} = [\mathbf{X}_{(i)}, -].$$

We recall Schur's lemma that an operator that commutes with all elements of an irreducible representation of a group must be a constant times the identity. The lemma applies equally to an irreducible representation of a Lie algebra, since such an algebra is the tangent space of the group—i.e., its infinitesimal transformations.

If, then P_j^i is to be anything but a multiple of the identity, the representation involved in S_j^i must be reducible. The extent of the reducibility involved determines the number of independent forms that p_{ij} may take. Also, it gives us a means of determining the p_{ij}.

If the field over which the algebra L is defined is the field of complex numbers, and if L is simple (i.e., if L is one of A_n, B_n, C_n, D_n, or the five exceptional algebras), then the adjoint representation is necessarily irreducible, and g_{ij} is the only two-index form. For suppose the representation were reducible; there would then be a linear vector space that would be invariant for all the operators Ad_{X_i}. Because the field is the complex numbers, this space would necessarily be expressible in terms of the basis elements of the algebra, so that it would be a subspace of the algebra. Let L_1 be such a subspace. We would have, then, that $Ad_X U = V$, where X is any element of L, and U any element of L_1, V in L_1. Since $Ad_X U = -Ad_U X = V$, L_1 would be an ideal of L. Since L is assumed simple, it does not have a proper ideal. Hence, the representation is irreducible, and g_{ij} is the only two-index form admissible.

If the field is not the complex field—if, for example, it is the real field—this argument fails since L_1 may not be a subspace of L. That this is not a trivial difficulty with the logic can be shown by example. Consider, for example, the realization of the Lorentz algebra given in our previous paper. (We do not repeat the details given there, but refer the reader to that paper.) We have already noted in Eq. (22), that there does exist a second two-index form. To see what happens to the P_j^i, we write down the adjoint representation:

$$Ad_{X_1} = \begin{bmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_1 \end{bmatrix}, \quad Ad_{X_2} = \begin{bmatrix} \mathbf{A}_2 & 0 \\ 0 & \mathbf{A}_2 \end{bmatrix},$$

$$Ad_{X_3} = \begin{bmatrix} \mathbf{A}_3 & 0 \\ 0 & \mathbf{A}_3 \end{bmatrix}, \quad Ad_{X_4} = \begin{bmatrix} 0 & -\mathbf{A}_1 \\ \mathbf{A}_1 & 0 \end{bmatrix},$$

$$Ad_{X_5} = \begin{bmatrix} 0 & -\mathbf{A}_2 \\ \mathbf{A}_2 & 0 \end{bmatrix}, \quad Ad_{X_6} = \begin{bmatrix} 0 & -\mathbf{A}_3 \\ \mathbf{A}_3 & 0 \end{bmatrix},$$

where

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We can find that the spaces

$$L_1 = \begin{bmatrix} a \\ b \\ c \\ ia \\ ib \\ ic \end{bmatrix}, \quad L_2 = \begin{bmatrix} a \\ b \\ c \\ -ia \\ -ib \\ -ic \end{bmatrix}, \tag{36}$$

where a, b, c are arbitrary complex numbers and are

invariant for all Ad_{X_i}. They are not, however, subspaces of the algebra.

We find, in particular, that the representation is indeed reducible, as may be seen by changing the basis with the similarity transformation using

$$T_j^i = \begin{bmatrix} \mathbf{I} & i\mathbf{I} \\ i\mathbf{I} & \mathbf{I} \end{bmatrix}. \tag{37}$$

The irreducible representations involved are $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, $\pm i\mathbf{A}_1$, $\pm i\mathbf{A}_2$, $\pm i\mathbf{A}_3$.

On the basis of the original $\mathbf{X}_i$, we find that any matrix of the form

$$P_j^i = \begin{bmatrix} \alpha\mathbf{I} & \beta\mathbf{I} \\ -\beta\mathbf{I} & \alpha\mathbf{I} \end{bmatrix}$$

commutes with all the Ad_{X_i}. Hence there are two linearly independent forms possible, the identity and

$$P_j^i = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix}. \tag{38}$$

The identity leads to $p_{ij} = g_{ij}$, as always. The matrix of Eq. (38) leads to the p_{ij} of Eq. (22).

We may note the emergence, in Eq. (38), of a group of which the conservation laws are one realization—in this case the cyclic group of order 4. We have, then, the somewhat curious situation of a group that is characteristic of the algebra.

As an alternative method of calculation, we consider Eq. (1). If we transform to $\mathbf{U}$, according to Eq. (27), we can write

$$\mathbf{U} = g^{iu}p_{ui}w^j\mathbf{X}_i = w^j\mathbf{Y}_j, \tag{39}$$

where

$$\mathbf{Y}_j = g^{iu}p_{ui}\mathbf{X}_i = P_j^i\mathbf{X}_i \tag{40}$$

so that we are now transforming the basis rather than the coefficients.

We require, then, that

$$d\mathbf{U}/dz = [\mathbf{S}, \mathbf{U}] = s^m w^n[\mathbf{X}_m, \mathbf{Y}_n] = (dw^i/dz)\mathbf{Y}_i. \tag{41}$$

On the other hand, we also know that

$$d\mathbf{W}/dz = [\mathbf{S}, \mathbf{W}] = s^m w^n[\mathbf{X}_m, \mathbf{X}_n] = (dw^i/dz)\mathbf{X}_i. \tag{42}$$

The coefficients s^m may take any values in the field. Also, providing that the coefficients of $\mathbf{S}$ obey the Lipschitz condition, the coefficients w^n may take any values in the field. For Eqs. (41) and (42) to be simultaneously satisfied for any values of s^m and w^n, it is necessary and sufficient that the expansion of $[\mathbf{X}_m, \mathbf{X}_n]$ in terms of $\mathbf{X}_i$ be identical, for any m and n, with the expansion of $[\mathbf{X}_m, \mathbf{Y}_n]$ in terms of $\mathbf{Y}_i$.

Since

$$[\mathbf{X}_m, \mathbf{X}_n] = c_{mn}^i\mathbf{X}_i,$$

we require that

$$[\mathbf{X}_m, \mathbf{Y}_n] = c_{mn}^i\mathbf{Y}_i. \tag{43}$$

In particular, if $\mathbf{X}_m$ is a regular element, then it generates a Cartan subalgebra as the maximum commuting subalgebra that contains it. It is evident that the transformation must leave invariant any Cartan subalgebra.

As an example, we can consider again the realization of the Lorentz algebra given in the previous paper. It is readily seen that Eq. (43) is satisfied if we set

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{X}_4, & \mathbf{Y}_4 &= -\mathbf{X}_1, \\ \mathbf{Y}_2 &= \mathbf{X}_5, & \mathbf{Y}_5 &= -\mathbf{X}_2, \\ \mathbf{Y}_3 &= \mathbf{X}_6, & \mathbf{Y}_6 &= -\mathbf{X}_3. \end{aligned} \tag{44}$$

This again leads to the p_{ij} of Eq. (22). We can also easily establish that the only transformations which leave the various Cartan subalgebras invariant are linear combinations of the identity and Eq. (44). Hence there are no other quadratic conservation laws.

Presumably, the higher-order operators can be investigated by similar techniques. We have not explored this possibility as yet.

VI. CASMIR OPERATORS

We shall now show that these same operators $p_{ij}...$ generate Casimir operators—i.e., operators that commute with the group of which the Lie algebra is the set of infinitesimal transformations. As discussed in the previous paper, we may have started with a matrix differential equation for the matricant:

$$d\mathbf{M}(z, z_0)/dz = \mathbf{S}\mathbf{M}(z, z_0), \qquad \mathbf{M}(z_0, z_0) = \mathbf{I}. \tag{45}$$

If $\mathbf{S}(z)$ is a curve in the Lie algebra, L, then $\mathbf{M}(z, z_0)$ is a set of curves in the corresponding Lie group. We discussed in that paper, also, how an equation of this form could be converted to one of the form of Eq. (1), where $\mathbf{W}(z)$ is a curve in the Lie algebra.

We are, then, concerned with operators that commute with the group in which $\mathbf{M}$ is embedded. Assuming that the representation that is involved is irreducible, Schur's lemma requires that such an operator be a constant times the identity. These operators have been used as a characterization of the representation.[8]

We will begin, again, by considering the two-index operator, p_{ij}, which satisfies Eq. (21), and later

generalize to higher orders. We consider the operator

$$\mathbf{C}^{(2)} = p_{ij}\mathbf{X}^i\mathbf{X}^j = p_{ij}g^{iu}g^{jv}\mathbf{X}_u\mathbf{X}_v. \tag{46}$$

To show that $C^{(2)}$ commutes with the group, it is sufficient to show that it commutes with any basis for the Lie algebra of infinitesimal transformations of the group. We therefore consider

$$\begin{aligned}[\mathbf{X}_i, \mathbf{C}^{(2)}] &= p_{jk}g^{iu}g^{kv}[\mathbf{X}_i, \mathbf{X}_u\mathbf{X}_v] \\ &= p_{jk}g^{iu}g^{kv}\{[\mathbf{X}_i, \mathbf{X}_u]\mathbf{X}_v + \mathbf{X}_u[\mathbf{X}_i, \mathbf{X}_v]\} \\ &= p_{jk}g^{iu}g^{kv}\{c_{iu}^s\mathbf{X}_s\mathbf{X}_v + c_{iv}^s\mathbf{X}_u\mathbf{X}_s\}.\end{aligned}$$

By Lemma 2, this becomes

$$[\mathbf{X}_i, \mathbf{C}^{(2)}] = p_{jk}g^{su}g^{kv}c_{ui}^j\mathbf{X}_s\mathbf{X}_v + p_{jk}g^{iu}g^{sv}c_{vi}^k\mathbf{X}_u\mathbf{X}_s.$$

Relabeling the dummy indices in the second term by $u \to s$, $s \to v$, $v \to k$, $k \to j$, $j \to u$, we find that

$$[\mathbf{X}_i, \mathbf{C}^{(2)}] = -g^{su}g^{kv}(c_{iu}^j p_{jk} + c_{ik}^j p_{uj})\mathbf{X}_s\mathbf{X}_v.$$

The terms in parentheses vanish by Eq. (21). Hence Eq. (21) is a sufficient condition for $\mathbf{C}^{(2)}$ to be a Casimir operator.

In the general case, we let

$$\mathbf{C}^{(n)} = (g^{iu}g^{jv}g^{kw}\cdots)p_{ijk\cdots}\mathbf{X}_u\mathbf{X}_v\mathbf{X}_w\cdots. \tag{47}$$

Then, since $Ad_{X_i} = [\mathbf{X}_i, —]$ is a derivation,

$$\begin{aligned}[\mathbf{X}_i, \mathbf{C}^{(n)}] &= (g^{iu}g^{kv}g^{hw}\cdots)p_{jkh\cdots}\{[\mathbf{X}_i, \mathbf{X}_u]\mathbf{X}_v\mathbf{X}_w\cdots \\ &\quad + \mathbf{X}_u[\mathbf{X}_i\mathbf{X}_v]\mathbf{X}_w\cdots + \mathbf{X}_u\mathbf{X}_v[\mathbf{X}_i, \mathbf{X}_w]\cdots + \cdots\} \\ &= (g^{iu}g^{kv}g^{hw}\cdots)p_{jkh\cdots}\{c_{iu}^s\mathbf{X}_s\mathbf{X}_v\mathbf{X}_w\cdots \\ &\quad + c_{iv}^s\mathbf{X}_u\mathbf{X}_s\mathbf{X}_w\cdots + c_{iw}^s\mathbf{X}_u\mathbf{X}_v\mathbf{X}_s\cdots + \cdots\} \\ &= \{(g^{it}g^{kv}g^{hw}\cdots)c_{it}^u + (g^{iu}g^{kt}g^{hw}\cdots)c_{it}^v \\ &\quad + (g^{iu}g^{kv}g^{ht}\cdots)c_{it}^w + \cdots\}p_{jkh\cdots}\mathbf{X}_u\mathbf{X}_v\mathbf{X}_w\cdots.\end{aligned}$$

Using Lemma 2, this becomes

$$\begin{aligned}&\{(g^{tu}g^{kv}g^{hw}\cdots)c_{ti}^j + (g^{iu}g^{tv}g^{hw}\cdots)c_{ti}^k \\ &\quad + (g^{iu}g^{kv}g^{tw}\cdots)c_{ti}^h + \cdots\}p_{jkh\cdots}\mathbf{X}_u\mathbf{X}_v\mathbf{X}_w\cdots \\ &= (g^{iu}g^{kv}g^{hw}\cdots)\{c_{ti}^t p_{tkh\cdots} + c_{ki}^t p_{jth\cdots} \\ &\quad + c_{hi}^t p_{jkt\cdots}\}\mathbf{X}_u\mathbf{X}_v\mathbf{X}_w\cdots \\ &= -(g^{iu}g^{kv}g^{hw}\cdots)\{c_{ti}^t p_{tkh\cdots} + c_{ik}^t p_{jth\cdots} \\ &\quad + c_{ih}^t p_{jkt\cdots}\}\mathbf{X}_u\mathbf{X}_v\mathbf{X}_w\cdots\end{aligned}$$

when we relabel the indices appropriately.

The terms in the brackets vanish by Eq. (26). Hence Eq. (26) is sufficient to assure that $\mathbf{C}^{(n)}$, defined by Eq. (47), is a Casimir operator for the group.

VII. CONCLUSIONS

We have shown, here, that the operators p_{ij} or $p_{ijk\cdots}$, defined for a given Lie algebra or group by Eq. (21) or (26), arise in several different and important contexts involving the behavior of systems embedded in the algebra or group. These application are so very different that it may seem surprising to find the same operators involved in all of them. However, the operators are those that generate the vector invariants of the underlying Lie group, and so are very intimately connected with the group structure itself.

ACKNOWLEDGMENT

We wish to acknowledge the helpful criticism of this paper by Dr. C. M. Ablow and Dr. C. Kockinos. We also wish to acknowledge the support of the U. S. Air Force for part of the work reported here.

Studies in Perturbation Theory. IX. Connection Between Various Approaches in the Recent Development—Evaluation of Upper Bounds to Energy Eigenvalues in Schrödinger's Perturbation Theory*

PER-OLOV LÖWDIN
Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida
and
Quantum Chemistry Group, Uppsala University, Uppsala, Sweden
(Received 14 December 1964)

The treatment of Schrödinger's perturbation theory based on the use of a series of inhomogeneous differential equations of iterative character is briefly surveyed. As an illustration, the method is used to derive the general expression for the expectation value of the Hamiltonian to any order which provides an upper bound for the ground-state energy. It is indicated how the well-known theory for inhomogeneous equations may be utilized also in this special case.

The solution of the Schrödinger equation by means of the partitioning technique and the concept of reduced resolvents is then treated. It is shown that the expressions obtained are most conveniently interpreted in terms of inhomogeneous differential equations. A study of the connection with the first approach reveals that the two methods are essentially equivalent, but also that the use of reduced resolvents and inverse operators may give an alternative insight in the mathematical structure of perturbation theory, particularly with respect to the "bracketing theorem" and the use of power series expansions with a remainder. In conclusion, it is emphasized that the combined use of the two methods provides a simpler and more powerful tool than any one of them taken separately.

1. INTRODUCTION

DURING the last decade, there has been a considerable amount of attention devoted to the quantum mechanical perturbation theory and its practical applications. In the classical days, Schrödinger[1] and Hylleraas[2] studied perturbation theory by considering a set of inhomogeneous equations, and recently this approach has been taken up again by Dalgarno and others.[3] This has led to an intense development of the entire field, and for a survey of the literature we would like to refer to the recent reviews by Hirschfelder *et al.*[4] and by Hall.[5] In this connection, one has studied the problem of evaluating not only the energy in the stationary states but also other properties by using improved perturbation technique.

In connection with the nuclear many-body problem there has been a considerable extension of perturbation theory based on the concepts of the wave and reaction operators, and particularly Brueckner[6] and his collaborators have contributed to this development. Between 1948 and 1951, several authors[7] discovered that the so-called "partitioning technique" could be used also in the study of perturbation theory, and this approach has been further developed in a series of papers from the Uppsala and Florida groups.[8] Particularly the concepts of wave and reaction operators have been investigated in this way.[8,9]

The method using inhomogeneous differential equations is first briefly discussed, and it is then shown how it may be used in connection with the variation principle to derive the expectation values of the Hamiltonian H to an arbitrary order. This

* This work was supported in part by the National Aeronautics and Space Administration under Research Grant No. G-512 with the University of Florida, and in part by King Gustaf VI Adolf's 70-Year Fund for Swedish Culture, Knut and Alice Wallenberg's Foundation, and by the Aerospace Research Laboratories, OAR, through the European Office of Aerospace Research (OAR), U. S. Air Force under Contract AF 61(052)-701 with Uppsala University.

[1] E. Schrödinger, Ann. Physik **80**, 437 (1926).

[2] E. A. Hylleraas, Z. Physik **48**, 469 (1928); **54**, 347 (1930); **65**, 209 (1930).

[3] A. Dalgarno and J. T. Lewis, Proc. Roy. Soc. (London) **A233**, 70 (1955); A. Dalgarno, Proc. Phys. Soc. (London) **A69**, 784 (1956); A. Dalgarno and A. L. Stewart, Proc. Roy. Soc. (London) **A238**, 276 (1956).

[4] J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, in *Advances in Quantum Chemistry*, edited by P. O. Löwdin (Academic Press, Inc., New York, 1964), Vol. 1.

[5] G. G. Hall, in *Advances in Quantum Chemistry*, edited by P.-O. Löwdin (Academic Press, Inc., New York, 1964).

[6] K. A. Brueckner, C. A. Levinson, and H. M. Mahmoud, Phys. Rev. **95**, 217 (1954); K. A. Brueckner, *ibid.* **96**, 508 (1954); **97**, 1353 (1955); **100**, 36 (1955); K. A. Brueckner and C. A. Levinson, *ibid.* **97**, 1344 (1955); H. A. Bethe, *ibid.* **103**, 1353 (1956); J. Goldstone, Proc. Roy. Soc. (London) **A239**, 267 (1957); H. A. Bethe and J. Goldstone, *ibid.* **A238**, 551 (1957); L. S. Rodberg, Ann. Phys. **2**, 199 (1957); to mention only a selection of the rich literature on this subject.

[7] H. A. Kramers, in *Studies and Essays* [presented to R. Courant on his 60th birthday (Interscience Publishers, Inc., New York 1948)], p. 205; S. Sueoka, J. Phys. Soc. Japan **4**, 361 (1949); M. H. L. Pryce, Proc. Phys. Soc. (London) **A63**, 25 (1950); M. Lax, Phys. Rev. **79**, 200A (1950); P.-O. Löwdin, J. Chem. Phys. **19**, 1396 (1951); W. Kohn, J. Chem. Phys. **17**, 670 (1949); S. F. Boys, Proc. Roy. Soc. (London) **A201**, 125 (1950). See further J. O. Hirschfelder and P.-O. Löwdin, Technical Note No. 3 from Uppsala Quantum Chemistry Group 1957; J. Mol. Phys. **2**, 229 (1959); G. Speisman, Phys. Rev. **107**, 1180 (1957); W. H. Young and N. H. March, *ibid.* **109**, 1854 (1958).

[8] P.-O. Löwdin, J. Math. Phys. **3**, 969, 1171 (1962); J. Mol. Spectry **10**, 12 (1963); **13**, 326 (1964).

[9] P.-O. Löwdin, Rev. Mod. Phys. **35**, 702 (1963).

leads to a series of expressions which provide upper bounds for the ground-state energy. After a brief survey of the partitioning technique which reveals the mathematical structure of perturbation theory in an explicit way, it is shown that the quantities involved are most conveniently interpreted in terms of solutions to inhomogeneous equations. It is established that the two approaches are completely equivalent, but also that there are certain formal mathematical differences which may be utilized as an advantage, so that the combined use of the two methods should provide a better and more powerful tool than anyone of the methods used separately.

2. PERTURBATION THEORY BASED ON THE USE OF INHOMOGENEOUS EQUATIONS

Schrödinger's Perturbation Theory

The fundamental problem in perturbation theory is the solution of the Schrödinger equation $\mathcal{H}\Psi = E\Psi$ for the stationary states of a system in the case when $\mathcal{H} = \mathcal{H}_0 + V$. It has often been customary to introduce the perturbation in the form λV instead of V and to systematize the eigenfunction Ψ and the energy E after powers of λ, but we will here instead systematize the solution to the eigenvalue problem after powers of V:

$$\Psi = \sum_{k=0}^{\infty} \varphi_k, \qquad E = E_0 + \sum_{k=1}^{\infty} \epsilon_k. \tag{1}$$

Here φ_0 is a solution to the unperturbed problem $\mathcal{H}_0\varphi_0 = E_0\varphi_0$. For the sake of simplicity we will choose the normalization

$$\langle \varphi_0 \mid \Psi \rangle = 1, \tag{2}$$

which simply implies that the higher-order corrections $\varphi_1, \varphi_2, \varphi_3, \cdots$ are orthogonal to φ_0. For the discrete eigenvalues, the final eigenfunction can then simply be renormalized to unity whenever needed. The eigenvalue problem has the form

$$(\mathcal{H}_0 + V)\Psi = E\Psi. \tag{3}$$

Following Schrödinger, one can now substitute the expansion (1) into (3) and, reorganizing the terms after powers of V, one obtains the following set of inhomogeneous equations:

$$\begin{aligned}
&(E_0 - \mathcal{H}_0)\varphi_0 = 0,\\
&(E_0 - \mathcal{H}_0)\varphi_1 = (V - \epsilon_1)\varphi_0,\\
&(E_0 - \mathcal{H}_0)\varphi_2 = (V - \epsilon_1)\varphi_1 - \epsilon_2\varphi_0,\\
&(E_0 - \mathcal{H}_0)\varphi_3 = (V - \epsilon_1)\varphi_2 - \epsilon_2\varphi_1 - \epsilon_3\varphi_0,\\
&\quad\vdots \qquad\qquad\qquad \vdots\\
&(E_0 - \mathcal{H}_0)\varphi_{n+1} = V\varphi_n - \sum_{k=0}^{n} \epsilon_{n+1-k}\varphi_k.
\end{aligned} \tag{4}$$

One can also write this system in the condensed matrix form

$$\begin{bmatrix} E^0 - \mathcal{H}_0 & & & & \\ & E_0 - \mathcal{H}_0 & & \mathbf{0} & \\ & & E_0 - \mathcal{H}_0 & & \\ & \mathbf{0} & & E_0 - \mathcal{H}_0 & \\ & & & & . \end{bmatrix} \begin{bmatrix} \varphi_0 \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ . \end{bmatrix}$$

$$= \begin{bmatrix} 0 & & & & \\ V - \epsilon_1 & 0 & & \mathbf{0} & \\ -\epsilon_2 & V - \epsilon_1 & 0 & & \\ -\epsilon_3 & -\epsilon_2 & V - \epsilon_1 & 0 & \\ . & . & . & . & . \end{bmatrix} \begin{bmatrix} \varphi_0 \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ . \end{bmatrix}. \tag{5}$$

The eigenvalue problem (3) is here equivalent with the solution of a set of inhomogeneous equations. One of the first to utilize this approach was Hylleraas in his classical study of this method, and he has later, several times returned to this problem.

The energy terms ϵ_k are easily expressed as matrix elements of the perturbation V. Multiplying equation (3) by φ_0^* to the left and integrating over the entire configuration space one obtains

$$\begin{aligned}
E &= E\langle \varphi_0 \mid \Psi \rangle = \langle \varphi_0 \mid \mathcal{H}_0 + V \mid \Psi \rangle\\
&= E_0 + \langle \varphi_0 \mid V \mid \Psi \rangle\\
&= E_0 + \sum_{k=0}^{\infty} \langle \varphi_0 \mid V \mid \varphi_k \rangle,
\end{aligned} \tag{6}$$

which immediately leads to the relation

$$\epsilon_{k+1} = \langle \varphi_0 \mid V \mid \varphi_k \rangle. \tag{7}$$

We note that we have here used the normalization (2). By means of the equation system formed it is now possible to transfer this expression so that, if all the functions $\varphi_0, \varphi_1, \cdots, \varphi_n$ up through order n have been determined, one can obtain all energy terms through order $(2n + 1)$. This can be performed by means of certain basic transformation formulas which will now be derived.

According to (4) we have the general formula

$$(E_0 - \mathcal{H}_0)\varphi_n = V\varphi_{n-1} - \sum_{l=0}^{n-1} \epsilon_{n-l}\varphi_l. \tag{8}$$

In the following, it is convenient to use the abbreviations

$$\begin{aligned}
&\langle \varphi_m \mid V \mid \varphi_n \rangle = \langle m \mid V \mid n \rangle = V_{mn};\\
&\langle \varphi_m \mid \varphi_n \rangle = \langle m \mid n \rangle.
\end{aligned} \tag{9}$$

Multiplying relation (8) to the left by φ^*_{m+1} and integrating, one obtains

$$\begin{aligned}\langle m+1|\, E_0 - \mathcal{H}_0 \,|n\rangle &= \langle m+1|\, V \,|n-1\rangle \\ &\quad - \sum_{l=0}^{n-1} \epsilon_{n-l}\langle m+1 \mid l\rangle, \\ \langle n|\, E_0 - \mathcal{H}_0 \,|m+1\rangle &= \langle n|\, V \,|m\rangle \\ &\quad - \sum_{k=0}^{m} \epsilon_{m+1-k}\langle n \mid k\rangle,\end{aligned} \tag{10}$$

where the second relation is obtained from the first by suitable changes of indices. However, since $(E_0 - \mathcal{H}_0)$ is an Hermitian operator this gives directly the equality

$$\begin{aligned}\langle m+1|\, V \,|n-1\rangle - \sum_{l=0}^{n-1} \epsilon_{n-l}\langle m+1 \mid l\rangle \\ = \langle m|\, V \,|n\rangle - \sum_{k=0}^{m} \epsilon_{m+1-k}\langle k \mid n\rangle,\end{aligned} \tag{11}$$

which leads to the basic transformation formula

$$\begin{aligned}\langle m|\, V \,|n\rangle &= \langle m+1|\, V \,|n-1\rangle \\ &\quad + \sum_{k=0}^{m} \epsilon_{m+1-k}\langle k \mid n\rangle - \sum_{l=0}^{n-1} \epsilon_{n-l}\langle m+1 \mid l\rangle.\end{aligned} \tag{12}$$

By means of this formula it is now possible to change the indices in the matrix elements of the perturbation V and relate them to the quantities ϵ_n and the matrix elements $\langle k|l\rangle$. Putting successively $m = 0, 1, 2, 3, \cdots$ we obtain a set of relations:

$$\begin{aligned}\langle 0|\, V \,|n\rangle &= \langle 1|\, V \,|n-1\rangle - \sum_{l=0}^{n-1} \epsilon_{n-l}\langle 1 \mid l\rangle; \\ \langle 1|\, V \,|n\rangle &= \langle 2|\, V \,|n-2\rangle \\ &\quad - \sum_{l=0}^{n-2} \epsilon_{n-l-1}\langle 2 \mid l\rangle + \epsilon_1\langle 1 \mid n-1\rangle; \\ \langle 2|\, V \,|n\rangle &= \langle 3|\, V \,|n-3\rangle \\ &\quad - \sum_{l=0}^{n-3} \epsilon_{n-l-2}\langle 3 \mid l\rangle + \sum_{k=1}^{2} \epsilon_{3-k}\langle k \mid n-2\rangle.\end{aligned} \tag{13}$$

Successive use of these formulas leads to the relation

$$\begin{aligned}\langle 0|\, V \,|n\rangle &= \langle p|\, V \,|n-p\rangle \\ &\quad - \sum_{k=1}^{p} \sum_{l=1}^{n-p} \epsilon_{n+1-k-l}\langle k \mid l\rangle.\end{aligned} \tag{14}$$

By means of (7), we hence obtain

$$\begin{aligned}\epsilon_{2n} &= \langle 0|\, V \,|2n-1\rangle \\ &= \langle n|\, V \,|n-1\rangle - \sum_{k=1}^{n} \sum_{l=1}^{n-1} \epsilon_{2n-k-l}\langle k \mid l\rangle, \\ \epsilon_{2n+1} &= \langle 0|\, V \,|2n\rangle \\ &= \langle n|\, V \,|n\rangle - \sum_{k=1}^{n} \sum_{l=1}^{n} \epsilon_{2n+1-k-l}\langle k \mid l\rangle,\end{aligned} \tag{15}$$

which relation shows that, if the contributions to the wavefunction have been calculated up through order n, it is possible to calculate the contributions to the energy up through order $(2n+1)$. If $n = 0, 1, 2, 3, \cdots$, one obtains, particularly,

$$\begin{aligned}\epsilon_1 &= \langle 0|\, V \,|0\rangle, \\ \epsilon_2 &= \langle 0|\, V \,|1\rangle = \langle 1|\, V \,|0\rangle, \\ \epsilon_3 &= \langle 0|\, V \,|2\rangle = \langle 1|\, V - \epsilon_1 \,|1\rangle, \\ \epsilon_4 &= \langle 0|\, V \,|3\rangle = \langle 1|\, V - \epsilon_1 \,|2\rangle - \epsilon_2\langle 1 \mid 1\rangle, \\ \epsilon_5 &= \langle 0|\, V \,|4\rangle = \langle 2|\, V - \epsilon_1 \,|2\rangle - \epsilon_3\langle 1 \mid 1\rangle \\ &\quad \cdots \\ &\quad - \epsilon_2\{\langle 1 \mid 2\rangle + \langle 2 \mid 1\rangle\}.\end{aligned} \tag{16}$$

These lower-order formulas are well-known, and general formulas of the same type as (15) have also previously been given.[10]

In concluding this section we note that Eq. (14) provides a very useful transformation formula for the matrix element of V. Putting $n = p + q$, one gets particularly the relation

$$\langle p|\, V \,|q\rangle = \epsilon_{p+q+1} + \sum_{k=1}^{p} \sum_{l=1}^{p} \epsilon_{p+q+1-k-l}\langle k \mid l\rangle, \tag{17}$$

which is of basic importance in the theory. In the following, we will use this relation to derive the expectation value of the Hamiltonian for the approximate wavefunctions of various order in Schrödinger's perturbation theory.

Variation Principle and Upper Bounds to the Energy in Schrödinger's Perturbation Theory

In the solution of the Schrödinger equation $\mathcal{H}\Psi = E\Psi$, the Rayleigh-Ritz variation principle $\delta\langle\mathcal{H}\rangle = 0$ is of fundamental importance. Let $\Phi = \Psi + \delta\Psi$ be an approximate eigenfunction and Ψ the exact eigenfunction satisfying the relation

$$(\mathcal{H} - E)\Psi = 0. \tag{18}$$

By means of this relation, the turnover rule, and the Hermitian property of the operator $(\mathcal{H} - E)$, one gets immediately

$$\frac{\langle\Phi|\, \mathcal{H} - E \,|\Phi\rangle}{\langle\Phi \mid \Phi\rangle} = \frac{\langle\delta\Psi|\, \mathcal{H} - E \,|\delta\Psi\rangle}{\langle\Phi \mid \Phi\rangle}, \tag{19}$$

or the relation

$$\begin{aligned}\langle\mathcal{H}\rangle_{\text{av}} &= \langle\Phi|\, \mathcal{H} \,|\Phi\rangle/\langle\Phi \mid \Phi\rangle \\ &= E + \langle\delta\Psi|\, \mathcal{H} - E \,|\delta\Psi\rangle/\langle\Phi \mid \Phi\rangle,\end{aligned} \tag{20}$$

showing that, if the trial function Φ is affected by an error $\delta\Psi$, the expectation value $\langle\mathcal{H}\rangle$ is affected by an error which is quadratic in the same quantity.

[10] F. Dupont-Bourdelet, J. Tillien, and J. Guy, J. Phys. Radium **21**, 776 (1960).

Applied to perturbation theory, this means that if $\delta\Psi$ is of the order $(n+1)$, then the expectation value $\langle\mathcal{H}\rangle$ is affected by an error of the order $(2n+2)$. One can also say that, if the trial function Φ is correct through order n, the expectation value $\langle\mathcal{H}\rangle$ will be correct through order $(2n+1)$.

We will now evaluate the expectation values of the energy associated with the perturbed wavefunction of the various orders explicitly. The trial function of order n is defined by the formula

$$\Phi_n = \sum_{k=0}^{n} \varphi_k, \tag{21}$$

and the associated expectation value is defined by the relation

$$\mathcal{E}_n^* = \langle\Phi_n|\ \mathcal{H}\ |\Phi_n\rangle/\langle\Phi_n\mid\Phi_n\rangle. \tag{22}$$

In the ground state, this quantity provides also an upper bound for the true eigenvalue E. This follows from (20) and the fact that the operator $(\mathcal{H}-E)$ is positive-definite.

Substitution of the trial function $\Phi_0=\varphi_0$, $\Phi_1=\varphi_0+\varphi_1$, $\Phi_2=\varphi_0+\varphi_1+\varphi_2$, $\cdots$ into Eq. (22) and the use of the relations (16) and (17) leads directly to formulas of the type

$$\begin{aligned}
\mathcal{E}_0^* &= E_0+\epsilon_1,\\
\mathcal{E}_1^* &= E_0+\epsilon_1+\epsilon_2+\epsilon_3-(\epsilon_2+\epsilon_3)\langle 1\mid 1\rangle/(1+\langle 1\mid 1\rangle),\\
\mathcal{E}_2^* &= E_0+\epsilon_1+\epsilon_2+\epsilon_3+\epsilon_4+\epsilon_5\\
&\quad -\frac{\epsilon_2\langle 2\mid 2\rangle+\epsilon_3\{\langle 1\mid 2\rangle+\langle 2\mid 1\rangle+\langle 2\mid 2\rangle\}+(\epsilon_4+\epsilon_5)\{\langle 1\mid 1\rangle+\langle 1\mid 2\rangle+\langle 2\mid 1\rangle+\langle 2\mid 2\rangle\}}{1+\langle 1\mid 1\rangle+\langle 1\mid 2\rangle+\langle 2\mid 1\rangle+\langle 2\mid 2\rangle},\\
&\quad \cdots,
\end{aligned} \tag{23}$$

showing results in full agreement with the variation principle, since they are correct through the orders 1, 3, and 5, respectively. In order to derive the general formula of this type, it is convenient to start from Eq. (8) in the modified form

$$(\mathcal{H}_0+V-E_0)\varphi_p = V\varphi_p - V\varphi_{p-1} + \sum_{l=0}^{p-1}\epsilon_{p-l}\varphi_l. \tag{24}$$

Putting $p=0,1,2,\cdots,n$, successively, one obtains

$$\begin{aligned}
(\mathcal{H}_0+V-E_0)\varphi_0 &= V\varphi_0,\\
(\mathcal{H}_0+V-E_0)\varphi_1 &= V\varphi_1 - V\varphi_0 + \epsilon_1\varphi_0,\\
(\mathcal{H}_0+V-E_0)\varphi_2 &= V\varphi_2 - V\varphi_1 + \epsilon_2\varphi_0 + \epsilon_1\varphi_1,\\
&\ \ \vdots \qquad\qquad \vdots\\
(\mathcal{H}_0+V-E_0)\varphi_n &= V\varphi_n - V\varphi_{n-1} + \sum_{l=0}^{n-1}\epsilon_{n-l}\varphi_l.
\end{aligned} \tag{25}$$

By summation of these relations, we obtain

$$(\mathcal{H}_0+V-E_0)\Phi_n = V\varphi_n + \sum_{p=1}^{n}\sum_{l=0}^{p-1}\epsilon_{p-l}\varphi_l. \tag{26}$$

Multiplying this relation to the left by Φ_m^* and using Eq. (17), one gets

$$\begin{aligned}
&\langle\Phi_n|\ \mathcal{H}_0+V-E_0\ |\Phi_n\rangle\\
&\quad=\sum_{p=0}^{n}\langle p|\ V\ |n\rangle + \sum_{k=0}^{n}\sum_{p=1}^{n}\sum_{l=0}^{n-1}\epsilon_{p-l}\langle k\mid l\rangle\\
&\quad=\sum_{p=0}^{n}\epsilon_{n+p+1} + \sum_{p=0}^{n}\sum_{k=1}^{p}\sum_{l=1}^{n}\epsilon_{n+p+1-k-l}\langle k\mid l\rangle\\
&\quad\quad+\sum_{p=1}^{n}\sum_{k=0}^{n}\sum_{l=0}^{p-1}\epsilon_{p-l}\langle k\mid l\rangle.
\end{aligned} \tag{27}$$

In order to proceed, it is convenient to write out the triple sums explicitly to see how they are constructed, and it is then easy to carry out the transformation to the following form:

$$\begin{aligned}
&\langle\Phi_n|\ \mathcal{H}_0+V-E_0-\sum_{r=1}^{2n+1}\epsilon_r\ |\Phi_n\rangle\\
&\quad=\sum_{p=0}^{n}\sum_{k=1}^{p}\sum_{l=1}^{n}\epsilon_{n+p+1-k-l}\langle k\mid l\rangle\\
&\quad\quad-\sum_{p=1}^{n}\sum_{k=1}^{n}\sum_{l=n+1-p}^{n}\epsilon_p\langle k\mid l\rangle\\
&\quad\quad-\sum_{p=0}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\epsilon_{n+p+1}\langle k\mid l\rangle
\end{aligned} \tag{28}$$

$$=-\sum_{r=2}^{2n+1}\epsilon_r\sum_{p=2n+2-r}^{2n}\ \sum_{k+l=p}\langle k\mid l\rangle. \tag{29}$$

It is instructive to consider the double sums over the matrix element $\langle k|l\rangle$ occurring in Eq. (28). These elements form a quadratic matrix of order $n\times n$, and the elements having the same sum are arranged according to Fig. 1.

Formula (28) implies that each quantity ϵ_r is modified by the double sums over the elements $\langle k|l\rangle$ taken over the diagonal $p=2n+2-r$ and over the elements situated to the right of this diagonal. This means particularly that the lower quantities ϵ_r have the following coefficients:

$$
\begin{aligned}
\epsilon_1&: \ 0, \\
\epsilon_2&: \ \langle n \mid n\rangle, \\
\epsilon_3&: \ \langle n \mid n\rangle + \langle n-1 \mid n\rangle + \langle n \mid n-1\rangle, \\
\epsilon_4&: \ \langle n \mid n\rangle + \langle n-1 \mid n\rangle \\
&\quad + \langle n \mid n-1\rangle + \langle n-2 \mid n\rangle \\
&\quad + \langle n-1 \mid n-1\rangle + \langle n \mid n-2\rangle, \qquad (30) \\
&\quad \cdots\cdots\cdots\cdots\cdots
\end{aligned}
$$

Similarly, the coefficients for ϵ_{2n-1} is the sum over all elements in the matrix except for the element $\langle 1|1\rangle$. Finally the coefficients for $\epsilon_{2n} + \epsilon_{2n+1}$ is the double sum over all elements $\langle k|l\rangle$ associated with the full $n \times n$ matrix. Introducing the notation

$$\delta_r = \sum_{p=2n+2-r}^{2n} \sum_{k+l=p} \langle k \mid l\rangle, \qquad (31)$$

we note that the quantity $\epsilon_r\delta_r$ contains terms of orders from $(2n + 2)$ up through order $(2n + r)$. Combining Eqs. (22) and (28) one has finally

$$\mathcal{E}_n^* = E_0 + \sum_{r=1}^{2n+1} \epsilon_r - \frac{\sum_{r=2}^{2n+1} \epsilon_r \delta_r}{1 + \sum_{k,l=1}^{n} \langle k \mid l\rangle}, \qquad (32)$$

in nice agreement with the variation principle, since the energy is correct through the order $(2n + 1)$. This is the formula for the successive expectation values of the Hamiltonian in Schrödinger's perturbation theory, and we note that, for the energy of the ground state of the perturbed system, it gives a series of convenient upper bounds.

Hylleraas' Variation Principle

It was pointed out by Hylleraas[11] that the individual energy terms ϵ_k fulfilled variation principles which are useful for practical calculations, and he particularly focuses his attention on the second-order energy

$$\epsilon_2 = \langle 0|\ V\ |1\rangle = \langle 1|\ V\ |0\rangle. \qquad (33)$$

The formula has the drawback that it contains the exact first-order terms φ_1 which may be found by solving the inhomogeneous equation $(E_0 - \mathcal{H}_0)\varphi_1 = (V - \epsilon_1)\varphi_0$. Hylleraas showed, however, that if $\tilde{\varphi}_1$ is an approximate solution to this equation, orthogonal to φ_0, then one has the theorem

$$\langle\tilde{\varphi}_1|\ \mathcal{H}_0 - E_0\ |\tilde{\varphi}_1\rangle + \langle\tilde{\varphi}_1|\ V - \epsilon_1\ |\varphi_0\rangle + \langle\varphi_0|\ V - \epsilon_1\ |\tilde{\varphi}_1\rangle \geq \epsilon_2. \qquad (34)$$

[11] E. A. Hylleraas, Z. Physik 65, 209 (1930).

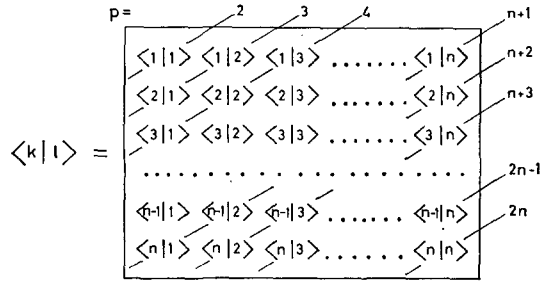


FIG. 1. Illustration of rearrangement of matrix elements according to Eq. (29).

The left-hand member gives hence an upper bound to the quantity ϵ_2, and we note that the equality sign holds only when $\tilde{\varphi}_1 = \varphi_1$. A lower bound for the quantity ϵ_2 has recently also been given by Prager and Hirschfelder.[12] The more detailed treatment of the question of upper and lower bounds for the quantities ϵ_k may be found in the broad survey by Hirschfelder *et al.*[4]

Solution of the Inhomogeneous Equations

The approach to perturbation theory outlined here is based on the solution of a set of inhomogeneous equations of the type (4), i.e.,

$$(E_0 - \mathcal{H}_0)\phi = f. \qquad (35)$$

Before trying to solve this equation, it is worthwhile to notice a general property. Multiplying this relation to the left by φ_0^* and integrating, one obtains the auxiliary condition $\langle\varphi_0|f\rangle = 0$, which is hence a necessary condition for solubility. It is sometimes convenient to write the right-hand member in the form $f = f' - \varphi_0\langle\varphi_0|f'\rangle$ with an arbitrary function f', since this condition is now automatically fulfilled. In general it is also requested that the solution ϕ should be orthogonal towards the unperturbed eigenfunction φ_0. If ϕ' is an arbitrary solution to the inhomogeneous equation (35), one can easily construct a solution satisfying the orthogonality requirement by means of the relation

$$\phi = \phi' - \varphi_0\langle\varphi_0 \mid \phi'\rangle. \qquad (36)$$

For a many-particle Hamiltonian, the solution of the inhomogeneous equation (35) represents a formidable problem, but, there are special cases for which explicit solutions may be found. The study of those cases may also give valuable hints for the development of a more general theory.

[12] S. Prager and J. O. Hirschfelder, J. Chem. Phys. 39, 3289 (1963).

One-Dimensional Case

Let us first consider a one-dimensional system with the variable x, the potential energy $U(x)$, and the unperturbed Hamiltonian

$$\mathcal{H}_0 = T + U = (-h^2/8\pi^2 m)d^2/dx^2 + U(x).$$

Instead of the inhomogeneous equation (35), we will now study the slightly more general equation

$$(\mathcal{E} - \mathcal{H}_0)\phi = f, \tag{37}$$

where $\mathcal{E}$ is a real parameter. By introducing the notations $Q(x) = 8\pi^2 m(\mathcal{E} - U)/h^2$ and $g = 8\pi^2 mf/h^2$, relation (37) takes the form

$$[d^2/dx^2 + Q(x)]\phi = g, \tag{38}$$

where the solution ϕ is subject to the condition that it should be quadratically integrable and orthogonal to φ_0, so that $\langle\phi|\varphi_0\rangle = 0$. In this connection, it is feasible to consider also the homogeneous equation

$$[d^2/dx^2 + Q(x)]\varphi = 0, \tag{39}$$

which has two linearly independent solutions, say ω_1 and ω_2. The theory of the equations (38) and (39) is extensively treated in the classical mathematical literature,[13] and one has only to apply the well-known results.

Of fundamental importance is the Wronskian W, defined by the relations

$$W(x) \equiv \omega_1(x)\omega_2'(x) - \omega_2(x)\omega_1'(x) \tag{40}$$

$$= \{\omega_1(x)\}^2(d/dx)\{\omega_2(x)/\omega_1(x)\}. \tag{41}$$

Combining (39) and (40), one obtains $W'(x) \equiv 0$, showing that the Wronskian is a constant: $W = W_0$. The relation $W_0 \neq 0$ is the necessary and sufficient condition for the linear independence of two solutions.

A nontrivial solution $\varphi(x)$ can never have zero-points of higher order, since the condition $\varphi(a) = \varphi'(a) = 0$ necessarily leads to the trivial solution $\varphi(x) \equiv 0$. All zero-points are hence single and, by means of the form (41) of the Wronskian and Rolle's theorem, one can easily show that between two zero-points to ω_2 there must be at least one zero-point to ω_1. Using the form (41), one can further easily express one of the solutions in terms of the other,

$$\omega_2(x) = W_0\omega_1(x)\int_b^x \{\omega_1(t)\}^{-2}\,dt. \tag{42}$$

Certain precautions are necessary in defining the integral, since the integration path cannot pass a zero-point of $\omega_1(x)$. However, since a zero-point of $\omega_1(x)$ corresponds to a pole of the second order for the integrand and one has the general theorem

$$\oint Z^{-2}\,dZ = 0, \tag{43}$$

the integral in (42) will be meaningful and the result the same for any complex contour which avoids the zero-points of $\omega_1(x)$. In the following, the integral will hence be defined in this way.

Let us now assume that the potential $U(x)$ is essentially different from zero only in a finite interval. Introducing the notation $\kappa^2 = -8\pi^2 m\mathcal{E}/h^2$, one obtains for $x = \pm\infty$ the asymptotic solution

$$\varphi \sim c_1 e^{-\kappa x} + c_2 e^{+\kappa x}. \tag{44}$$

It is now possible to choose the functions ω_1 and ω_2 so that they have the asymptotic behavior $e^{-\kappa x}$ and $e^{+\kappa x}$, respectively, by requesting $\omega_1(+\infty) = 0$ and $\omega_2(-\infty) = 0$.

Let us further assume that the function $g(x)$ in the right-hand member of (38) is essentially different from zero only in a finite interval. This implies that the general solution ϕ has the same asymptotic properties as φ expressed in (44).

Let us now study the particular solution ϕ^* which is characterized by the boundary condition $\phi^*(-\infty) = 0$. In general this solution will blow up for $x = +\infty$, but one can in principle evaluate the quotient

$$\lim_{x\to+\infty} \frac{\phi^*(x)}{\omega_2(x)} = a_2. \tag{45}$$

This means that the function $\phi(x) = \phi^*(x) - a_2\omega_2(x)$ is a particular solution to the inhomogeneous equation (39) which satisfies the boundary condition $\phi(\pm\infty) = 0$, necessary for the quadratic integrability (see Fig. 2). A characteristic feature of the theory is hence that it uses one of the "irregular" solutions of the homogeneous equation (39) to construct the solution ϕ desired in perturbation theory.

The solution of the inhomogeneous equation (38) can now easily be expressed in terms of ω_1 and ω_2 by using the standard method of "variation of parameters."[13] Let us put $\phi = \alpha\omega_1 + \beta\omega_2$ with the auxiliary condition $\alpha'\omega_1 + \beta'\omega_2 = 0$, where α and

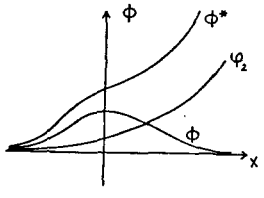


FIG. 2. Behavior of solutions to Eqs. (38) and (39).

[13] See e. g. A. R. Forsyth and W. Jacobsthal, *Lehrbuch der Differentialgleichungen* (Frederick Vieweg und Sohn, Braunschweig, Germany, 1912).

β are two functions of x to be determined. Substitution into (38) gives $\alpha'\omega_1' + \beta'\omega_2' = g$, which leads to the equation system

$$\begin{cases} \alpha'\omega_1' + \beta'\omega_2' = g, \\ \alpha'\omega_1 + \beta'\omega_2 = 0, \end{cases} \tag{46}$$

with the solution $\alpha' = -W_0^{-1}g\omega_2$, $\beta' = W_0^{-1}g\omega_1$. Hence, one has

$$\begin{aligned} \alpha(x) &= -W_0^{-1}\int^x g(\xi)\omega_2(\xi)\,d\xi, \\ \beta(x) &= W_0^{-1}\int^x g(\xi)\omega_1(\xi)\,d\xi, \end{aligned} \tag{47}$$

and the well-known solution

$$\phi(x) = \phi(a) - W_0^{-1}\int_a^x [\omega_1(x)\omega_2(\xi) - \omega_2(x)\omega_1(\xi)]g(\xi)\,d\xi. \tag{48}$$

Finally one can combine formulas (42) and (48) and, by using Dirichlet's theorem, one can derive the expression

$$\phi(x) = \phi(a) + \omega_1(x)\int_a^x \{\omega_1(t)\}^{-2}\,dt\int_a^t \omega_1(\xi)g(\xi)\,d\xi. \tag{49}$$

Choosing $a = -\infty$ and putting $\phi(-\infty) = 0$, one obtains

$$\phi(x) = \omega_1(x)\int_{-\infty}^x \{\omega_1(t)\}^{-2}\,dt\int_{-\infty}^t \omega_1(\xi)g(\xi)\,d\xi. \tag{50}$$

One can hence solve the inhomogeneous equation (38) by repeated quadrature, if one knows the specific solution $\omega_1(x)$ associated with the boundary condition $\omega_1(+\infty) = 0$ to the homogeneous equation (39). Our considerations are valid for any value of the parameter $\mathcal{E}$ and, for $\mathcal{E} = E_0$, one may replace $\omega_1(x)$ by the unperturbed eigenfunction $\varphi_0(x)$.

The spirit of the recent trend[3-5] in perturbation theory is that, in the one-dimensional case, one can solve the system of inhomogeneous equations (4) by repeated curvature simply by successively putting $f = (V - \epsilon_1)\varphi_0$, $f = (V - \epsilon_1)\varphi_1 - \epsilon_2\varphi_0$, $\cdots$ etc. into (50). For the ground state, the unperturbed function φ_0 is nodeless and all the integrals can be evaluated directly,[14] whereas, for the excited states, one has to use complex contours avoiding the poles or subtract the effect of the poles.[15] The results obtained, so far, have been encouraging, and there is little question that the general mathematical methods for treating inhomogeneous equations render a powerful tool in the practical applications of perturbation theory.

Three-Dimensional Case

In many three-dimensional problems, one can separate the variables and obtain one-dimensional equations of the type treated above. However, even in the more general case, there is a rich mathematical literature about the treatment of inhomogeneous equations which can be consulted. The results are by no means as simple as in the one-dimensional case, but it seems as if the study of the system (4) would still provide one of the most powerful approaches to perturbation theory. Here we will only make some brief comments about a particularly simple case.

Let $\mathcal{H} = T + V$ be the Hamiltonian for a single particle, and let us choose $\mathcal{H}_0$ as the Hamiltonian for a free particle, so that $\mathcal{H}_0 = T = (-h^2/8\pi^2 m)\nabla^2$. Introducing the notation $k^2 = 8\pi^2 m\mathcal{E}/h^2$ for $\mathcal{E} > 0$, one obtains from (4) inhomogeneous equations of the type

$$(\nabla^2 + k^2)\phi = g, \tag{51}$$

which have the well-known solutions

$$\phi(P) = -\frac{1}{4\pi}\int \frac{e^{\pm ikr_{PQ}}}{r_{PQ}}\,g(Q)\,dv_Q. \tag{52}$$

These are just the integrals which appear in Born's[16] treatment of scattering theory, only that the boundary conditions are different, and we note the remarkable success the use of the inhomogeneous equations has had in this case. For $\mathcal{E} < 0$, one should instead introduce the notation $\kappa^2 = -8\pi^2 m\mathcal{E}/h^2$, and the solution takes then the form

$$\phi(P) = -\frac{1}{4\pi}\int \frac{e^{\pm\kappa r_{PQ}}}{r_{PQ}}\,g(Q)\,dv_Q, \tag{53}$$

i.e., the solution is expressed in terms of so-called "Yukawa potentials." In both cases, the proper choice of φ_0 and the treatment of the boundary conditions is of essential importance. It appears that the discrete levels and the scattering states to a large extent can be treated analogously, and the closer connection ought to be studied in greater detail. It seems also worthwhile investigating whether there exist other three-dimensional Hamiltonians $\mathcal{H}_0$ for which the inhomogeneous equations

[14] W. H. Young and N. H. March, Phys. Rev. 109, 1854 (1958).

[15] W. Byers Brown and J. O. Hirschfelder, Proc. Nat. Acad. Sci. U. S. 50, 399 (1963).

[16] M. Born, Z. Physik 37, 863; 38, 803 (1926).

$(\mathcal{E} - \mathcal{H}_0)\phi = f$ have explicit solutions of a similar degree of simplicity as (53).

3. PARTITIONING TECHNIQUE

Partitioning of Secular Equation

It is feasible to start a study of the partitioning technique by considering the eigenvalue problem $\mathcal{H}\Psi = E\Psi$ in matrix form. Let us expand the eigenfunction Ψ in a complete orthonormal basis $\mathbf{f} = (f_1, f_2, f_3, \cdots)$ so that $\Psi = \sum_k f_k c_k$, where the coefficients c_k form a column vector $\mathbf{c}$. Introducing the matrix $\mathbf{H} = \mathbf{f}^\dagger \mathcal{H} \mathbf{f}$ having the elements $\mathcal{H}_{kl} = \langle f_k | \mathcal{H} | f_l \rangle$ and multiplying the eigenvalue relation $(\mathcal{H} - E)\mathbf{fc} = 0$ to the left by $\mathbf{f}^\dagger$, one obtains $\mathbf{f}^\dagger(\mathcal{H} - E)\mathbf{fc} = 0$, i.e.,

$$(\mathbf{H} - E\cdot\mathbf{1})\mathbf{c} = 0, \tag{54}$$

or $\sum_l (\mathcal{H}_{kl} - E\delta_{kl})c_l = 0$. This is the conventional matrix form of the eigenvalue problem, and the eigenvalues are determined by the secular equation $\det\{\mathcal{H}_{kl} - E\delta_{kl}\} = 0$.

In the partitioning technique, one starts from (54) but avoids the explicit use of the associated secular equation. Equation (54) is of the general form of a system of linear equations in the unknown coefficients c_k:

$$\mathbf{Mc} = 0, \tag{55}$$

where $\mathbf{M} = \mathbf{H} - E\boldsymbol{\Delta}$. Let us now partition the basis $\mathbf{f}$ into two parts a and b, for instance:

$$\underbrace{f_1, f_2, f_3,}_{a} \quad \underbrace{f_4, f_5, f_6,}_{b} \ \ldots, \tag{56}$$

with the corresponding partitioning of the matrix $\mathbf{M}$ and the column vector $\mathbf{c}$:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ba} & \mathbf{M}_{bb} \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} \mathbf{c}_a \\ \mathbf{c}_b \end{bmatrix}. \tag{57}$$

The equation system (55) can now be written in the form

$$\begin{aligned} \mathbf{M}_{aa}\mathbf{c}_a + \mathbf{M}_{ab}\mathbf{c}_b &= 0, \\ \mathbf{M}_{ba}\mathbf{c}_a + \mathbf{M}_{bb}\mathbf{c}_b &= 0. \end{aligned} \tag{58}$$

Solving $\mathbf{c}_b$ from the second equation, one obtains

$$\mathbf{c}_b = -\mathbf{M}_{bb}^{-1}\mathbf{M}_{ba}\mathbf{c}_a, \tag{59}$$

provided that the matrix $\mathbf{M}_{bb}^{-1}$ exists. Substitution into the first equation (58) leads to an expression of the form

$$\mathbf{M}'_{aa}\mathbf{c}_a = 0, \tag{60}$$

where

$$\mathbf{M}'_{aa} = \mathbf{M}_{aa} - \mathbf{M}_{ab}\mathbf{M}_{bb}^{-1}\mathbf{M}_{ba}. \tag{61}$$

Formulas (58)–(61) renders a simple recipe for "decoupling" systems of linear equations which is useful also in other connections.

The choice of the two parts (a) and (b) is quite arbitrary, and if one selects the subspace (a) to consist of a single element, say f_1, equation (60) takes the form $M'_{11}c_1 = 0$, which implies $M'_{11} = 0$ if $c_1 \neq 0$. This is equivalent with the relation

$$E = \mathcal{H}_{11} + \mathbf{H}_{1b}(E\cdot\mathbf{1}_{bb} - \mathbf{H}_{bb})^{-1}\mathbf{H}_{b1}. \tag{62}$$

The right-hand member defines a function

$$f(\mathcal{E}) \equiv \mathcal{H}_{11} + \mathbf{H}_{1b}(\mathcal{E}\cdot\mathbf{1}_{bb} - \mathbf{H}_{bb})^{-1}\mathbf{H}_{b1} \tag{63}$$

in a real or complex variable $\mathcal{E}$ with the derivative

$$f'(\mathcal{E}) \equiv -\mathbf{H}_{1b}(\mathcal{E}\cdot\mathbf{1}_{bb} - \mathbf{H}_{bb})^{-2}\mathbf{H}_{b1} < 0, \tag{64}$$

and the eigenvalues E are then represented by the zero-points of the function $y = \mathcal{E} - f(\mathcal{E})$. A closer investigation[8] shows that one obtains both the non-degenerate and the degenerate eigenvalues in this way. Since $y' = 1 - f'(\mathcal{E}) > 1$, the equation $y = 0$ has only single roots, and this implies that the partitioning technique renders a transformation from the original to the "reduced" characteristic equation for the eigenvalues. By using (64), it is further easy to show that each pair of values $\mathcal{E}$ and $\mathcal{E}_1 = f(\mathcal{E})$ bracket at least one true eigenvalue E. This "bracketing theorem" gives us a new possibility to approach the problem of upper and lower bounds to E, which will be developed in a forthcoming paper.

Partitioning by Projection Operators

In order to study the connection with the use of inhomogeneous differential equations in greater detail, we will now describe the partitioning technique expressed in terms of operators. Let $\mathcal{E}$ be a real or complex variable, and let φ be an arbitrary reference function such that $\langle\varphi|\varphi\rangle = 1$. Let further $O = |\varphi\rangle\langle\varphi|$ be the projection operator associated with φ and $P = 1 - O$ the projection operator for the orthogonal complement. They satisfy the fundamental relations

$$\begin{aligned} O^2 = O, \quad O^\dagger = O, \quad \mathrm{Tr}\,(O) = 1, \\ P^2 = P, \quad P^\dagger = P, \quad OP = PO = 0. \end{aligned} \tag{65}$$

By means of O and P, one can now easily carry out the partitioning into the two subspaces a and b. The key quantity in the matrix treatment described

above is apparently the inverse matrix $(\mathcal{E}\cdot\mathbf{1}_{bb}-\mathbf{H}_{bb})^{-1}$, which in operator theory takes the symbolic form

$$T = P/(\mathcal{E} - \mathfrak{H}), \tag{66}$$

and is called a "reduced resolvent." More strictly, T is defined by the relation $T=P[\alpha\cdot O+P(\mathcal{E}-\mathfrak{H})P]^{-1}P$, for arbitrary $\alpha \neq 0$. Since $\partial T/\partial\alpha \equiv 0$, the operator T is independent of the value of α, which is an auxiliary quantity introduced only so that the operator may properly exist. The operator T fulfills the basic relations

$$TO = OT = 0, \tag{67}$$

$$P(\mathcal{E} - \mathfrak{H})T = P. \tag{68}$$

The theory is based on the use of expressions of the type

$$\chi = Tf, \tag{69}$$

where f is an arbitrary function. From (69) and (67) follows directly $O\chi = 0$, i.e., $\langle\varphi|\chi\rangle = 0$, which implies that χ will be orthogonal to the reference function φ. From (68) follows further

$$P(\mathcal{E} - \mathfrak{H})\chi = Pf, \tag{70}$$

or

$$(\mathcal{E} - \mathfrak{H})\chi = f - \varphi\langle\varphi \mid f + \mathfrak{H}\chi\rangle, \tag{71}$$

which is indeed an inhomogeneous equation. The solution of (71) represents hence the most natural way for evaluating the function χ defined by (69), even if there may be even other possibilities.

In analogy to (59), we will now consider the function ϕ defined by the relation

$$\phi = T\mathfrak{H}\varphi, \tag{72}$$

which is a function of the type (69) for $f = \mathfrak{H}\varphi$. Substitution into (71) gives immediately the inhomogeneous equation

$$(\mathcal{E} - \mathfrak{H})\phi = \{\mathfrak{H} - \langle\varphi|\ \mathfrak{H}\ |\varphi + \phi\rangle\}\varphi. \tag{73}$$

For the discussion of this equation, it is convenient to introduce the quantities

$$\Psi_{\mathcal{E}} = \varphi + \phi = (1 + T\mathfrak{H})\varphi, \tag{74}$$

$$\mathcal{E}_1 = \langle\varphi|\ \mathfrak{H}\ |\varphi + \phi\rangle = \langle\varphi|\ \mathfrak{H} + \mathfrak{H}T\mathfrak{H}\ |\varphi\rangle. \tag{75}$$

Equation (73) takes then the form $(\mathcal{E} - \mathfrak{H})\phi = (\mathfrak{H} - \mathcal{E}_1)\varphi$, i.e., one obtains the following inhomogeneous equation:

$$(\mathfrak{H} - \mathcal{E})\Psi_{\mathcal{E}} = (\mathcal{E}_1 - \mathcal{E})\varphi. \tag{76}$$

For the special case when $\mathcal{E}_1 = \mathcal{E}$, the function $\Psi_{\mathcal{E}}$ is hence a solution to the Schrödinger equation and $\mathcal{E} = E$ is the associated eigenvalue.

This result implies that one should study the function

$$\mathcal{E}_1 = f(\mathcal{E}) \equiv \langle\varphi|\ \mathfrak{H} + \mathfrak{H}T\mathfrak{H}\ |\varphi\rangle$$

or

$$f(\mathcal{E}) \equiv \langle\varphi|\ \mathfrak{H} + \mathfrak{H}\frac{P}{\mathcal{E} - \mathfrak{H}}\mathfrak{H}\ |\varphi\rangle, \tag{77}$$

which is completely analogous to (63). According to (62), it has the derivative

$$f'(\mathcal{E}) \equiv -\langle\varphi|\ \mathfrak{H}\frac{P}{(\mathcal{E} - \mathfrak{H})^2}\mathfrak{H}\ |\varphi\rangle = -\langle\phi \mid \phi\rangle, \tag{78}$$

provided that the last integral exists, and, since this quantity is negative, one has again the "bracketing theorem." In conclusion, we note that, if the equation $y = \mathcal{E} - f(\mathcal{E})$ is solved by the Newton–Raphson procedure and one starts from the point $\mathcal{E} = \mathcal{E}_0$, the next approximation is given by the formula $\mathcal{E}_0^* = \mathcal{E}_0 - y_0/y_0'$, or

$$\mathcal{E}_0^* = \mathcal{E}_0 - (\mathcal{E}_0 - \mathcal{E}_1)/(1 + \langle\phi \mid \phi\rangle_0). \tag{79}$$

This value is identical with the expectation value of $\mathfrak{H}$ for the function $\Psi_{\mathcal{E}}$ defined by (74) for $\mathcal{E} = \mathcal{E}_0$. Hence relation (79) is equivalent with the variational expression (20).

In conclusion we note that, since ϕ is orthogonal to φ, the function $\Psi_{\mathcal{E}}$ is subject to the so-called intermediate normalization

$$\langle\varphi \mid \Psi_{\mathcal{E}}\rangle = 1, \tag{80}$$

which is valid for the entire spectrum. For the discrete energy levels, the eigenfunctions can later be renormalized to unity.

Partitioning in Perturbation Theory

In order to study the connection with perturbation theory, we will now put $\mathfrak{H} = \mathfrak{H}_0 + V$ and $\varphi = \varphi_0$, where $\mathfrak{H}_0\varphi_0 = E_0\varphi_0$. Since $T\varphi_0 = 0$, the relations (74) and (75) will be simplified to the form

$$\Psi_{\mathcal{E}} = (1 + TV)\varphi_0, \tag{81}$$

$$\mathcal{E}_1 = E_0 + \langle\varphi_0|\ V + VTV\ |\varphi_0\rangle. \tag{82}$$

It is convenient to introduce the wave operator W and the reaction operator t through the definitions

$$W = 1 + TV, \tag{83}$$

$$t = V + VTV = VW, \tag{84}$$

which gives

$$\Psi_{\mathcal{E}} = W\varphi_0, \qquad \mathcal{E}_1 = E_0 + \langle\varphi_0|\ t\ |\varphi_0\rangle. \tag{85}$$

The operator $T = P/(\mathcal{E} - \mathfrak{H})$ is essentially an inverse operator and, in this connection, we note the identity

$$(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1},$$

which by repeated use gives rise to the power series expansion

$$(A - B)^{-1} = A^{-1} \sum_{k=0}^{\infty} (BA^{-1})^k. \tag{87}$$

The denominator in T can now be divided into terms A and B in different ways, and we will particularly consider the two forms

$$\mathcal{E} - \mathfrak{H} = (\mathcal{E} - \mathfrak{H}_0) - V \tag{88}$$

$$= (E_0 - \mathfrak{H}_0) - (V - \mathcal{E} + E_0). \tag{89}$$

Expanding T in terms of powers of V, one is led to the Brillouin-type perturbation theory, whereas expansion in terms of powers of $V' = V - (\mathcal{E} - E_0)$ leads to the Schrödinger-type perturbation theory. We note that we still have the variable $\mathcal{E}$ at our disposal, and that $\mathcal{E}$ and $\mathcal{E}_1$ bracket an eigenvalue. Introducing the quantities

$$T_0 = P/(\mathcal{E} - \mathfrak{H}_0), \qquad R_0 = P/(E_0 - \mathfrak{H}_0), \tag{90}$$

and using (87), (88), and (89), one obtains the infinite expansions

$$T = T_0 \sum_{k=0}^{\infty} (VT_0)^k \tag{91}$$

$$= R_0 \sum_{k=0}^{\infty} (V'R_0)^k. \tag{92}$$

The question of the truncation of these expansions to a finite number of terms with an estimate of the remainder will be treated in a forthcoming paper.[17] Here we will instead concentrate our interest on the structure of the separate terms and their evaluation by means of inhomogeneous equations.

Brillouin-Type Perturbation Theory

Substituting expansion (91) into (81) and (82), one obtains the fundamental formulas in the Brillouin-type perturbation theory

$$\Psi_{\mathcal{E}} = \varphi_0 + T_0V\varphi_0 + T_0VT_0V\varphi_0 + \cdots = \sum_{k=0}^{\infty} (T_0V)^k\varphi_0, \tag{93}$$

$$\mathcal{E}_1 = E_0 + \langle\varphi_0|\ V + VT_0V + VT_0VT_0V + \cdots\ |\varphi_0\rangle = E_0 + \sum_{k=0}^{\infty} \langle\varphi_0|\ V(T_0V)^k\ |\varphi_0\rangle. \tag{94}$$

[17] P.-O. Löwdin, Technical Note 120, Uppsala Quantum Chemistry Group, June 1964 (unpublished).

Introducing the notations

$$\chi_k = (T_0V)^k\varphi_0, \qquad \epsilon_{k+1} = \langle\varphi_0|\ V\ |\chi_k\rangle, \tag{95}$$

one hence has

$$\Psi_{\mathcal{E}} = \sum_{k=0}^{\infty} \chi_k, \qquad \mathcal{E}_1 = E_0 + \sum_{k=1}^{\infty} \epsilon_k. \tag{96}$$

The functions χ_k are apparently related by the recursion formula $\chi_{k+1} = T_0V\chi_k$ and, substituting $\mathfrak{H} = \mathfrak{H}_0$ and $f = V\chi_k$ into (71), one gets the inhomogeneous equation

$$(\mathcal{E} - \mathfrak{H}_0)\chi_{k+1} = V\chi_k - \epsilon_{k+1}\varphi_0. \tag{97}$$

It is thus possible to evaluate the quantities χ_k and $\mathcal{E}_k$ occurring in Brillouin's perturbation theory by successively solving a series of equations of the type (97) for $k = 0, 1, 2, \cdots$, starting from $\chi_0 = \varphi_0$. This gives the connection desired.

Schrödinger-Type Perturbation Theory

Let us now substitute expansion (92) into (81) and (82) which gives the relations

$$\Psi_{\mathcal{E}} = \varphi_0 + R_0V\varphi_0 + R_0V'R_0V\varphi_0 + \cdots = \sum_{k=0}^{\infty} (R_0V')^k\varphi_0\ , \tag{98}$$

$$\mathcal{E}_1 = E_0 + \langle\varphi_0|\ V + VR_0V + VR_0V'R_0V + \cdots\ |\varphi_0\rangle = E_0 + \sum_{k=0}^{\infty} \langle\varphi_0|\ V(R_0V')^k\ |\varphi_0\rangle, \tag{99}$$

where we have also used the fact that $R_0\varphi_0 = 0$. Let us now consider the special value $\mathcal{E} = E = E_0 + \langle\varphi_0|\ t\ |\varphi_0\rangle$ and rearrange the terms in (98) and (99) after powers of V according to (1), so that

$$\Psi_{\mathcal{E}} = \sum_{k=0}^{\infty} \varphi_k; \qquad \mathcal{E}_1 = E_0 + \sum_{k=1}^{\infty} \epsilon_k. \tag{100}$$

One obtains directly

$$\begin{aligned}
\varphi_1 &= R_0V\varphi_0, \\
\varphi_2 &= R_0(V - \epsilon_1)R_0V\varphi_0, \\
\varphi_3 &= R_0(V - \epsilon_1)R_0(V - \epsilon_1)R_0V\varphi_0 - \epsilon_2R_0^2V\varphi_0 \\
&\ \ \vdots \qquad\qquad\qquad \vdots
\end{aligned} \tag{101}$$

$$\begin{aligned}
\epsilon_1 &= \langle\varphi_0|\ V\ |\varphi_0\rangle, \qquad \epsilon_2 = \langle\varphi_0|\ VR_0V\ |\varphi_0\rangle, \\
\epsilon_3 &= \langle\varphi_0|\ VR_0(V - \epsilon_1)R_0V\ |\varphi_0\rangle, \\
\epsilon_4 &= \langle\varphi_0|\ VR_0(V - \epsilon_1)R_0(V - \epsilon_1)R_0V\ |\varphi_0\rangle \\
&\qquad - \epsilon_2\langle\varphi_0|\ VR_0^2V\ |\varphi_0\rangle \\
&\ \ \vdots \qquad\qquad\qquad \vdots
\end{aligned} \tag{102}$$

The lower-order terms are found by inspection, whereas the general terms are most easily written up by using the classification of the partitioning of the integers. From (101) and (102) follows immediately the recursion relations

$$\begin{aligned}
\varphi_1 &= R_0 V \varphi_0, \\
\varphi_2 &= R_0(V - \epsilon_1)\varphi_1, \\
\varphi_3 &= R_0(V - \epsilon_1)\varphi_2 - \epsilon_2 R_0 \varphi_1, \\
&\vdots \qquad\qquad \vdots \\
\varphi_{n+1} &= R_0\Big\{(V - \epsilon_1)\varphi_n - \sum_{k=2}^{n} \epsilon_k \varphi_{n+1-k}\Big\}.
\end{aligned} \tag{103}$$

Again, the lower-order relations are immediately written by inspection, and induction leads then to the general formula which is easily proven. Use of (71) for $\mathfrak{K} = \mathfrak{K}_0$, $\mathcal{E} = E_0$, and $\varphi = \varphi_0$ shows that the definition $\chi = R_0 f$ is equivalent with the inhomogeneous equation $(E_0 - \mathfrak{K}_0)\chi = f - \varphi_0\langle\varphi_0|f\rangle$ with the auxiliary condition $\langle\varphi_0|\chi\rangle = 0$. This implies that the system (103) is completely equivalent with the system of inhomogeneous equations given in (4):

$$(E_0 - \mathfrak{K}_0)\varphi_{n+1} = (V - \epsilon_1)\varphi_n - \sum_{k=2}^{n+1} \epsilon_k \varphi_{n+1-k}, \tag{104}$$

where the term for $k = n + 1$ in the last sum comes from the projection operator P in the right-hand member of (70). It is evident that any formal expression in R_0 can be interpreted in terms of solutions of inhomogeneous differential equations in the operator $(E_0 - \mathfrak{K}_0)$, and vice versa. However, even if the two approaches are completely equivalent, many derivations may look at least formally different in the two schemes. By means of (101) and (102) and the "turn-over rule," one obtains, e.g., for the energies in (102) the following connection formulas:

$$\begin{aligned}
\epsilon_2 &= \langle\varphi_0|\ V\ |\varphi_1\rangle = \langle\varphi_1|\ V\ |\varphi_0\rangle, \\
\epsilon_3 &= \langle\varphi_1|\ V - \epsilon_1\ |\varphi_1\rangle, \\
\epsilon_4 &= \langle\varphi_1|\ V - \epsilon_1\ |\varphi_2\rangle - \epsilon_2\langle\varphi_1\ |\ \varphi_1\rangle \\
&= \langle\varphi_2|\ V - \epsilon_1\ |\varphi_1\rangle - \epsilon_2\langle\varphi_1\ |\ \varphi_1\rangle.
\end{aligned} \tag{105}$$

By induction, one finds easily the general relations which are identical with formulas (15) previously derived. The explicit relations (102) look very neat and condensed, but the final results are, of course, the same. Actually, the two derivations are much more closely parallel than one would think at first sight.

In conclusion, it should be observed that the connection with the conventional form of Schrödinger's perturbation theory is easily established by using the spectral resolution of the reduced resolvent R_0:

$$R_0 = \frac{P}{E_0 - \mathfrak{K}_0} = \sum_{k\neq 0} \frac{|\varphi_k^0\rangle\langle\varphi_k^0|}{E_0 - E_k^0}, \tag{106}$$

where the functions φ_k^0 are the eigenfunctions to $\mathfrak{K}_0$ associated with the eigenvalues. Substitution of (106) into (101) and (102) leads to the well-known expressions

$$\begin{aligned}
\varphi_1 &= \sum_{k\neq 0} \varphi_k^0 \frac{\langle\varphi_k^0|\ V\ |\varphi_0\rangle}{E_0 - E_k^0}, \\
\varphi_2 &= \sum_{k,l\neq 0} \varphi_k^0 \frac{\langle\varphi_k^0|\ V - \epsilon_1\ |\varphi_l^0\rangle\langle\varphi_l^0|\ V\ |\varphi_0\rangle}{(E_0 - E_k^0)(E_0 - E_l^0)} \\
&\vdots \qquad\qquad \vdots,
\end{aligned} \tag{107}$$

and

$$\begin{aligned}
\epsilon_1 &= \langle\varphi_0|\ V\ |\varphi_0\rangle, \\
\epsilon_2 &= \sum_{k\neq 0} \frac{\langle\varphi_0|\ V\ |\varphi_k^0\rangle\langle\varphi_k^0|\ V\ |\varphi_0\rangle}{E_0 - E_k^0}, \\
\epsilon_3 &= \sum_{k,l\neq 0} \frac{\langle\varphi_0|\ V\ |\varphi_k^0\rangle\langle\varphi_k^0|\ V - \epsilon_1\ |\varphi_l^0\rangle\langle\varphi_l^0|\ V\ |\varphi_k^0\rangle}{(E_0 - E_k^0)(E_0 - E_l^0)} \\
&\vdots \qquad\qquad \vdots,
\end{aligned} \tag{108}$$

If E_0 is the ground-state energy of the unperturbed system, one can further from (106) conclude that R_0 is a negative-definite operator and derive a series of inequalities of the form

$$\begin{aligned}
R_0 &> \frac{P}{E_0 - E_1^0} \\
R_0 &> \frac{|\varphi_1^0\rangle\langle\varphi_1^0|}{E_0 - E_1^0} + \frac{P - |\varphi_1^0\rangle\langle\varphi_1^0|}{E_0 - E_2^0}, \\
&\vdots \qquad\qquad \vdots \\
R_0 &> \sum_{k=1}^{p} \frac{|\varphi_k^0\rangle\langle\varphi_k^0|}{E_0 - E_k^0} + \frac{P - \sum_{k=1}^{p} |\varphi_k^0\rangle\langle\varphi_k^0|}{E_0 - E_{p+1}^0}.
\end{aligned} \tag{109}$$

Using these lower bounds to R_0, one can now derive a series of lower bounds to e.g., ϵ_2 starting with

$$\epsilon_2 > \frac{\langle\varphi_0|\ V^2\ |\varphi_0\rangle - \langle\varphi_0|\ V\ |\varphi_0\rangle^2}{E_0 - E_1^0}. \tag{110}$$

Combining these relations with (34), one has hence both upper and lower bounds for ϵ_2. For a more detailed discussion of these problems for the higher quantities ϵ_k, we will refer to the review of Hirschfelder *et al.*[4]

4. CALCULATION OF OTHER QUANTITIES THAN THE ENERGY

In quantum theory, the energy plays a dominating role since it is one of the physical quantities which are used to characterize the "stationary states." It is well known that, according to the variation principle, it is comparatively easy to get a good estimate of the energy, since a wavefunction of order n leads to an expectation value (20) of order $(2n + 1)$. A similar principle applies also to the "constants of motion", but for all other physical quantities associated with a stationary state one has no such tool to rely on. This implies that, if an arbitrary physical quantity is represented by the self-adjoint operator F, then the expectation value $\langle F\rangle_{av}$ is correct to the same order n as the wavefunction, and, in order to obtain a result of the same accuracy as the energy, it is then necessary to evaluate the corrections to $\langle F\rangle$ of the orders $n + 1$, $n + 2$, $\cdots$, $2n$, and $(2n + 1)$. In particular, if one has a zeroth-order wavefunction available, it is important to calculate the first-order correction to $\langle F\rangle$. In this connection, an "interchange theorem" established by Dalgarno[18] and his coworkers has proven to be of essential value.[4,5]

By using the general relation $\Psi = W\varphi_0$, where $W = 1 + TV$ is the wave operator (85), it is easy to write up the exact expression for the expectation value of F for the stationary state characterized by the wavefunction Ψ:

$$\bar{F} = \langle F\rangle_{av} = \frac{\langle\Psi|\ F\ |\Psi\rangle}{\langle\Psi\ |\ \Psi\rangle} = \frac{\langle\varphi_0|\ W^\dagger FW\ |\varphi_0\rangle}{\langle\varphi_0|\ W^\dagger W\ |\varphi_0\rangle}$$

$$= \frac{\langle\varphi_0|\ F + VTF + FTV + VTFTV\ |\varphi_0\rangle}{\langle\varphi_0|\ 1 + VT^2V\ |\varphi_0\rangle}. \quad (111)$$

Multiplication with the denominator $\langle\varphi_0|\ 1 + VT^2V\ |\varphi_0\rangle$ gives further the implicit formula

$$\bar{F} = \langle\varphi_0|\ F + VTF + FTV + VT(F - \bar{F})TV\ |\varphi_0\rangle, \quad (112)$$

which is convenient as a basis for the discussion. For T, we will now use the expansion (92) for the particular value $\mathcal{E} = E = E_0 + \sum_l \epsilon_l$, i.e.,

$$T = R_0 \sum_{k=0}^{\infty} \left\{\left(V - \sum_{l=1}^{\infty} \epsilon_l\right)R_0\right\}^k$$

$$= R_0 + R_0(V - \epsilon_1)R_0 + R_0(V - \epsilon_1)R_0(V - \epsilon_1)R_0 - \epsilon_2 R_0^2 + \cdots. \quad (113)$$

[18] W. Byers Brown and W. J. Meath, Proc. Nat. Acad. Sci. U. S. 52, 65 (1964).

Substituting (113) into (112) and rearranging the terms after powers of V, one obtains

$$\begin{aligned}
\bar{F}_0 &= \langle\varphi_0|\ F\ |\varphi_0\rangle,\\
\bar{F}_1 &= \langle\varphi_0|\ VR_0F + FR_0V\ |\varphi_0\rangle,\\
\bar{F}_2 &= \langle\varphi_0|\ VR_0(V - \epsilon_1)R_0F + FR_0(V - \epsilon_1)R_0V\\
&\qquad + VR_0(F - \bar{F}_0)R_0V\ |\varphi_0\rangle \qquad (114)\\
&\ \ \vdots \qquad\qquad \vdots\\
&\ \ \vdots \qquad\qquad \vdots\ .
\end{aligned}$$

Using the technique developed in connection with studying the general terms (101) and (102) in Schrödinger's perturbation theory, one can give the explicit expression for $\bar{F}_k$ of any order. Here we will now focus the attention on the first-order correction

$$\bar{F}_1 = \langle\varphi_0|\ VR_0F + FR_0V\ |\varphi_0\rangle. \quad (115)$$

The fact that the right-hand member is symmetric in V and F is the background for Dalgarno's "interchange theorem." Introducing the function $\eta_0 = R_0F\varphi_0$, i.e., the solution to the inhomogeneous equation

$$(E_0 - \mathcal{H}_0)\eta_0 = (F - \bar{F}_0)\varphi_0, \quad (116)$$

subject to the condition $\langle\varphi_0|\eta_0\rangle = 0$, one obtains directly

$$\bar{F}_1 = \langle\varphi_0|\ V\ |\eta_0\rangle + \langle\eta_0|\ V\ |\varphi_0\rangle. \quad (117)$$

These formulas are of particular importance for all one-electron operators F:

$$F = \sum_{i=1}^{N} F(i), \quad (118)$$

provided that the unperturbed Hamiltonian has also been chosen of the same form $\mathcal{H}_0 = \sum_i \mathcal{H}_0(i)$. In such a case, the problem $(E_0 - \mathcal{H}_0)\varphi_0 = 0$ is *separable* and has solutions in the form of Hartree-products and Slater determinants, and it is then easy to show that the same applies to the inhomogeneous equation (116). For one-electron operators F, the interchange theorem may hence render considerable simplifications.

In this connection, we note that the combined use of the reduced resolvent R_0 and the associated inhomogeneous equation gives a simple and forceful tool for treating the problem of $\bar{F}_1$. The explicit form of (115) suggests immediately the introduction of the function $\eta_0 = R_0F\varphi_0$, whereas the inhomogeneous equation (116) forms a better basis for a study of the separability property of the differential equations. The quantity $\bar{F}_1$ has now been successfully evaluated for a number of cases and, for a survey

of the current literature, we will refer to the papers by Hirschfelder *et al.*[4] and by Hall.[5]

Another important problem in this connection is to find out whether the sum $(\bar{F}_0 + \bar{F}_1)$ may again be written as an expectation value of F. This happens for instance, if the unperturbed Hamiltonian $\mathcal{H}_0$ and hence φ_0 can be chosen in such a way that

$$\bar{F}_1 = 0, \tag{119}$$

and, at the same time, the perturbation V can be kept small. In atomic theory, one has previously often used power series expansions in Z^{-1}, but one has now also tried to use expansions in $(Z - \delta)^{-1}$, where δ is a shielding constant, in order to get (119) fulfilled. So far, this problem is not yet fully solved.

The explicit form of $\bar{F}_2$ indicates that it seems difficult or impossible to generalize the interchange approach with practical success to the higher-order corrections of $\bar{F}$.

5. DISCUSSION

The purpose of this note is to show the close connection between the various approaches in the recent studies of perturbation theory based on the use of inhomogeneous equations and of reduced resolvents. The strength of the first approach is illustrated by a derivation of the explicit expression for the expectation value of $\mathcal{H}$ in the Schrödinger perturbation theory to any order, which is a quantity of importance in determining upper bounds to the energy.

The resolvent operator technique is based on the use of inverse operators, but all expressions of the type $\chi = R_0 f$ may be interpreted so that χ is the solution to the inhomogeneous equation

$$(E_0 - \mathcal{H}_0)\chi = f - \varphi_0\langle\varphi_0 \mid f\rangle, \tag{120}$$

subject to the auxiliary condition $\langle\varphi_0|\chi\rangle = 0$, which gives immediately the connection with the first approach. However, one may also utilize the special properties of inverse operators, for instance, formal rules of the type

$$(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1}, \tag{121}$$

$$(\partial/\partial\varepsilon)A^{-1} = -A^{-1}(\partial A/\partial\varepsilon)A^{-1}, \tag{122}$$

which are simple and powerful tools for transforming the expressions involved. Formula (121) is of importance in deriving perturbation expansions with a definite remainder, whereas formula (122) is basic for the derivation of (78) and the "bracketing theorem," which provides upper and lower bounds for the eigenvalues. These problems will be further discussed in a forthcoming paper.[17]

In part of the current literature, there has been a certain trend to focus the attention solely to the inhomogeneous differential equations themselves[4,18] without any explicit use of inverse operators whatsoever. From our discussion it follows, on the other hand, that the two approaches are completely equivalent, since they may be transformed into each other. There is a certain difference connected with the fact that the mathematical forms look rather different, but this should be utilized as an advantage. It is well known that mathematical expressions may be "suggestive" in various degrees to different scientists, but, even if this property is partly of a personal nature depending on the background of the scientists involved, it is highly important for the further development of the research. Instead of developing a theory solely along one line of approach, it may hence be worthwhile to consider the alternative mathematical form also of the other approach in order to get new ideas. Hence it seems likely that a *combination* of the two methods discussed may provide the simplest and most powerful tool available for treating some of the current problems in perturbation theory.

Erratum: On the Construction of a Unitary Matrix with Elements of Given Moduli

MATTS ROOS
Nordisk Institut for Teoretisk Atomfysik, Copenhagen, Denmark
[J. Math. Phys. 5, 1609 (1964)]
(Received 8 February 1965)

The theorem on p. 1611 is formulated incorrectly and should therefore be omitted. Simultaneously, the discussions concerning the inequality signs in Eqs. (12) and (14), following these equations, should be omitted. This erratum has kindly been pointed out to the author by Dr. H. Araki, Kyoto University, Japan, in a private communication.

Errata: Application of Operational Methods to The Analysis of Uniform Plasmas

SAUL SILVEN
Lockheed Missiles and Space Company, Palo Alto, California
[J. Math. Phys. 5, 557 (1964)]
(Received 4 January 1965)

IN Eq. (12) on p. 558 and Eqs. (33), (34), and (12) on p. 560, the symbols $\mathbf{i}_1$ and $\mathbf{i}_2$ should be interchanged. In Eq. (19) on p. 559 the factor $(\nu \pm a)$ should be $(-\nu \pm a)$. In Eq. (28) on p. 559, the quantity $[\frac{1}{2}(\omega_c^2 + 4\omega_p^2) \pm \omega_c]$ should be $\frac{1}{2}[(\omega_c^2 + 4\omega_p^2) \pm \omega_c]$. In the last sentence of the first paragraph on p. 560, the symbol i should be $\mathbf{i}$.