Linear Closure Approximation Method for Classical Statistical Mechanics

JOHN M. RICHARDSON

Hughes Research Laboratories, Malibu, California

AND

LEO C. LEVITT*

Department of Physics, Georgetown University, Washington, D.C.

A description is given of a general closure principle involving the minimization of the mean square error. The procedure based upon this principle can be applied to the truncation of the BBGKY hierarchy at various stages and to the approximation of unwanted terms arising in the equation of motion method by linear combinations of the observables to be retained. On a general level the significance of the closure principle is described in terms of the geometry of function space, and several useful general properties of the principle are derived. A discussion is devoted to the relation between the closure error (i.e., the least mean square error) and the error in the end result (e.g., the free energy, the radial distribution function, etc.); however, the results, while providing some insight, are not sufficiently refined to provide upper bounds to errors in all problems of statistical mechanics where the method is applicable. On the level of specific application it is shown that the principle yields results identical to the random phase approximation and to the linearized version of the Kirkwood superposition approximation in two special cases. Later sections of the paper describe in greater than usual generality, the formalism connecting thermodynamic properties and other equilibrium properties with the microscopic equations of motion in which closure approximations have been introduced. Two illustrative examples of the application of the over-all method were made to the case of a classical system of electrons in a uniform background of compensating charge, one leading to the well-known results of Debye and the other to a more accurate and elaborate theory developed in quantitative detail elsewhere.

I. INTRODUCTION

THE purpose of this paper is to apply a general closure principle, discussed in detail elsewhere by one of the authors,[1] to some representative problems in classical statistical mechanics. The closure principle is described in Sec. III, and is called least mean square closure. It consists of only two input elements: one is a set of observables (or more generally, a

manifold of observables) in terms of which another function of coordinates and moments is to be linearly approximated. The second input element is an averaging operation to be employed in calculating the mean of the square of the error in the above approximation. When these input elements are fixed, the remainder of the procedure is determined—the minimization of the mean square error with respect to the choice of coefficients in the linear approximant.

In the present paper the viewpoint is mainly methodological. Our closure principle yields two well-known closure approximations as special cases,

---

i.e., the random phase approximation and the linearized Kirkwood superposition approximation; this does not mean that it is to be viewed solely as the process of deriving well-known approximations from relatively unfamiliar starting points. The consistency of the principle with widely accepted approximations lends added confidence in applying the principle to closure approximations in other situations. Therefore, the special cases presented here are intended, at least in part, as a "calibration" of the general method.

However, the least mean square error closure principle provides more than extrapolation of approximation methodology from the familiar into the unfamiliar. It gives a new and different understanding of frequently employed closure approximations (e.g., the random phase approximation and the linearized Kirkwood superposition approximation). It also promises a satisfactory approach to the problem of error estimation, which has been a serious deficiency in previous work on closure techniques.

In this paper the application of the closure principle is described on two levels: a general level characterized by a general observable manifold, and a specific level on which particular manifolds and particular physical systems are considered. In Secs. II, III, and IV the method is described and discussed for the case of a general observable manifold. In Sec. V, we consider an observable manifold spanned by symmetric sums of single-particle functions of position and an averaging operation corresponding to a free-particle canonical ensemble. It is shown that the resultant closure principle, when applied to pair functions, is identical to the random phase approximation. Section VI treats a different manifold spanned by symmetric sums of translationally invariant pair functions of position. With the same averaging operation as before, the closure principle applied to triplet functions is equivalent to the linearized Kirkwood superposition approximation.[2] Sections VII and VIII deal with the calculation of thermodynamic properties, using the classical equation of motion method for a general observable manifold and incorporating the least mean square error closure procedure. The final section applies the previous results to the specific case of a system composed of classical electrons in a uniform background of compensating charge.

II. STATES AND OBSERVABLES

We consider classical systems containing a fixed number N of particles of one type. The state of a system is defined by the set of coordinates and

moments of the N particles

$$X = (\mathbf{r}_1, \cdots, \mathbf{r}_N; \mathbf{p}_1, \cdots, \mathbf{p}_N). \qquad (2.1)$$

A basic feature in the present method is the selection of a set of observables

$$S = (\alpha_1, \cdots, \alpha_n),$$

where each observable α_j is a function of the coordinates and moments, i.e., $\alpha_j = \alpha_j(X)$, which is generally complex. It is required that the observables be linearly independent. Furthermore, it is required that the observables be invariant to the interchange of identical particles. Although the general discussion involves a finite set of observables, extension to the case of infinite sets in special applications will involve no essential difficulties.

A fundamental concept is the observable manifold (OM), which is defined as the set of all linear combinations of observables. The OM is fundamental in the sense that it is clearly invariant to linear transformations of the type

$$\alpha_i' = \sum_{j=1}^{n} C_{ij}\alpha_j, \qquad (2.2)$$

where C_{ij} is nonsingular. Thus, the OM may be regarded as representing a large number of equivalent sets of observables. In any case, the OM may be said to be "spanned" by the set S.

It is required that the OM contain unity as an element. It is further required that the OM be invariant to complex conjugation. Invariance to other operations may be included in special cases.

III. LEAST MEAN SQUARE ERROR CLOSURE

In a large number of problems in classical statistical mechanics it is desired to approximate phase functions by linear combinations of the members of a given observable set, i.e., by a member of a given observable manifold. After averaging in a suitable statistical ensemble, the approximation is then expressed in terms of mean observables. In some cases, this takes the form of an approximation of an n-particle distribution function of one order by a linear combination of those of lower order, as shown in Secs. V and VI.

We first consider the approximation method for the case of an arbitrary phase function $\gamma = \gamma(X)$, and a general observable manifold. Let γ be approximated by a linear combination of observables in the set S, as follows:

$$\gamma = \sum_{j=1}^{n} a_j\alpha_j + \epsilon, \qquad (3.1)$$

where ϵ is the error. As the criterion of the validity of approximation we use the mean square error $\langle \epsilon^*\epsilon \rangle'$;

[2] See, for example, T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956).

here the symbol $\langle \ \rangle'$ denotes the operation of averaging in the subsidiary statistical ensemble defined by a distribution function $P'(X)$ which is sufficiently simple to allow explicit analytical treatment, but is still not drastically different from the actual ensemble involved in the problem being treated. The process of averaging in the latter ensemble is denoted by $\langle \ \rangle$. A typical example of the averaging operation $\langle \ \rangle'$ is in the process of averaging in the free-particle canonical ensemble. We frequently impose the requirement that $\langle \ \rangle'$ share the same invariance properties as the OM.

The minimization of $\langle \epsilon^* \epsilon \rangle'$ with respect to the values of the a_j yields the set of equations

$$\langle \alpha_j^* \epsilon \rangle' = \langle \alpha_j^* \gamma \rangle' - \sum_{k=1}^{n} \langle \alpha_j^* \alpha_k \rangle' a_k = 0. \qquad (3.2)$$

The solution of (3.2) is

$$a_j = \sum_k Q_{jk} \langle \alpha_k^* \gamma \rangle', \qquad (3.3)$$

where Q_{jk} is the *matrix* reciprocal of $\langle \alpha_j^* \alpha_k \rangle'$, i.e.,

$$\sum_k Q_{jk} \langle \alpha_k^* \alpha_l \rangle' = \delta_{jl}. \qquad (3.4)$$

With the optimal choice of coefficients a_j, Eq. (3.1) can be written in the form

$$\gamma = \mathfrak{p}\gamma + \epsilon, \qquad (3.5)$$

where

$$\mathfrak{p}\gamma = \sum_{jk} \alpha_j Q_{jk} \langle \alpha_k^* \gamma \rangle'. \qquad (3.6)$$

Alternatively, we can write

$$\mathfrak{p}\gamma = \int dX' P'(X') K(X, X') \gamma(X'), \qquad (3.7)$$

where

$$K(X, X') = \sum_{jk} \alpha_j(X) Q_{jk} \alpha_k^*(X') \qquad (3.8)$$

and $P'(X)$ is the subsidiary distribution function involved in the operation $\langle \ \rangle'$.

One can readily deduce the following properties of the operator $\mathfrak{p}$:

    a. $\mathfrak{p}$ is a linear operator.   (3.9)

    b. $\mathfrak{p}$ is a projection operator.   (3.10)

    c. $\mathfrak{p}$ leaves elements of the OM unchanged,

$$\mathfrak{p}\alpha_j = \alpha_j. \qquad (3.11)$$

    d. $\mathfrak{p}\gamma_1$ and $(1 - \mathfrak{p})\gamma_2$ are uncorrelated in the subsidiary ensemble, i.e.,

$$\langle (\mathfrak{p}\gamma_1)((1 - \mathfrak{p})\gamma_2) \rangle' = 0. \qquad (3.12)$$

A large number of additional properties of $\mathfrak{p}$ exist; however, the above list is sufficient for the applications treated later.

It is perhaps of interest to interpret the operation $\mathfrak{p}$ in terms of the geometry of a Hilbert space in which

the elements are all possible phase functions, the averages of whose squares exist in the subsidiary ensemble. We define the inner product of the two functions $f(X)$ and $g(X)$ by the relation

$$(f, g) = \langle f^* g \rangle'. \qquad (3.13)$$

Clearly, the norm (i.e., "distance" of f from the origin) is given by

$$\|f\|^2 = \langle f^* f \rangle'. \qquad (3.14)$$

The observable manifold (OM) is then the $(n + 1)$-dimensional hyperplane passing through the "origin" (the function O) and the "points" $\alpha_j(X)$. We consider next an arbitrary "point" γ, generally not a member of the OM. It can then be shown that $\mathfrak{p}\gamma$ is that point in the OM which is closest to γ in the sense of the minimum of the "distance" (norm of the difference) from γ to the point in the OM. An arbitrary point in the OM is clearly

$$\sum_j a_j \alpha_j;$$

the norm of the difference of this function and γ is

$$\left\| \gamma - \sum a_j \alpha_j \right\| = \left[\left\langle \left| \gamma - \sum_j a_j \alpha_j \right|^2 \right\rangle' \right]^{\frac{1}{2}}, \qquad (3.15)$$

the minimization of which is clearly identical to the original mean square error minimization problem.

To complete the discussion of closure we must consider the problem of minimizing an over-all "error function" dependent upon a set of errors associated with a corresponding set of functions $\gamma_\mu (\mu = 1, \cdots, m)$ to be approximated. We write

$$\gamma_\mu = \sum_j \alpha_j a_{j\mu} + \epsilon_\mu, \qquad (3.16)$$

where the coefficients are to be chosen to minimize the over-all error function

$$C = \sum W_{\mu\nu} \langle \epsilon_\mu^* \epsilon_\nu \rangle', \qquad (3.17)$$

where $W_{\mu\nu}$ is a positive definite Hermitian matrix. It can be shown that the solution is *independent* of the matrix $W_{\mu\nu}$ and that, furthermore, the solution is expressible in terms of the projection operator $\mathfrak{p}$ in the same way as before; i.e., with the optimal values of the $a_{j\mu}$, we can write

$$\sum \alpha_j a_{j\mu} = \mathfrak{p}\gamma_\mu. \qquad (3.18)$$

IV. DISCUSSION OF ERRORS

A simple rearrangement of (3.5) yields the following expression for the error ϵ involved in the replacement of the phase function γ by the optimal point $\mathfrak{p}\gamma$ in the OM:

$$\epsilon = \mathfrak{q}\gamma, \qquad (4.1)$$

where $\mathfrak{q} = 1 - \mathfrak{p}$ is the projection operator which is complementary to $\mathfrak{p}$. The least mean square error,

employing the subsidiary averaging operation $\langle\ \rangle'$, is clearly

$$\langle|\epsilon|^2\rangle' = \langle|q\gamma|^2\rangle', \qquad (4.2)$$

a quantity which can be computed directly in most cases of interest. It is important to note that the average error (again using the subsidiary averaging operation) is simply

$$\langle\epsilon\rangle' = 0; \qquad (4.3)$$

furthermore,

$$\langle\alpha_j^*\epsilon\rangle' = 0, \qquad (4.4)$$

for any observable α_j in the OM.

Let us now turn to the consideration of the operation of averaging in the actual statistical ensemble involved. Let us denote the corresponding distribution by $P(X)$ and the averaging operation by $\langle\ \rangle$. If we had used $\langle\ \rangle$ in computing the least mean square error, i.e., $\langle\ \rangle' = \langle\ \rangle$, it would follow directly from (4.3) that

$$\langle\epsilon\rangle = 0; \qquad (4.5)$$

thus we would conclude that

$$\langle\gamma\rangle = \langle p\gamma\rangle. \qquad (4.6)$$

Therefore, in this case the truncation approximation would be *exact* after averaging. It is thus clear that in the general case where the two averaging operations are not identical, the magnitude of the error $\langle\epsilon\rangle$ is proportional to the deviation of $P(X)$ from $P'(X)$.

Since the distribution functions corresponding to $\langle\ \rangle$ and $\langle\ \rangle'$ are P and P', respectively, we can write

$$\langle\epsilon\rangle = \langle\epsilon P/P'\rangle'. \qquad (4.7)$$

Using a trivial generalization of the Schwartz inequality, we obtain the result

$$|\langle\epsilon\rangle|^2 \le \langle|\epsilon|^2\rangle'\langle(P/P')^2\rangle'. \qquad (4.8)$$

However, it is possible to derive a stronger inequality. Employing the results (3.12) and (4.1), it can be shown that

$$\langle\epsilon P/P'\rangle' = \langle\epsilon q(P/P')\rangle' \qquad (4.9)$$

and hence that

$$|\langle\epsilon\rangle|^2 \le \langle|\epsilon|^2\rangle'\langle[q(P/P')]^2\rangle'. \qquad (4.10)$$

The last inequality provides a generally closer upper bound to $\langle\epsilon\rangle$ than does (4.8), since it can be shown that

$$\langle[(qP/P')]^2\rangle' \le \langle(P/P')^2\rangle'. \qquad (4.11)$$

Thus, the expressions (4.8) and (4.10) provide upper bounds to $\langle\epsilon\rangle$ (the error averaged in the actual ensemble) in terms of $\langle\epsilon^2\rangle'$, the square error averaged in the subsidiary ensemble.

Let us assume that the actual ensemble is defined by the distribution function

$$P(X) = \exp[\beta(A - H)], \qquad (4.12)$$

and the subsidiary ensemble by

$$P'(X) = \exp[\beta(A_0 - H_0)]. \qquad (4.13)$$

Let us assume that

$$H = H_0 + gH_1,$$

where g is a parameter of smallness. Let us further assume that $H_1, H_1^2, \cdots, H_1^r$ lie in the OM. It then follows that in the expansion of P/P' in powers of gH_1, the application of q will annihilate all terms up through the rth. Thus, employing the Taylor-Maclaurin remainder formula, we obtain

$$q\left(\frac{P}{P'}\right) = \frac{1}{(r+1)!}\exp[\beta(A - A_0 - g\theta H_1)] \cdot (gH_1)^{r+1}, \qquad (4.14)$$

where θ is a function of gH_1 satisfying the inequality $0 < \theta < 1$. Assuming that γ is independent of g, it follows from (4.10) that $\langle\epsilon\rangle$ is of the order of g^{r+1} if the subsidiary average of the square of (4.14) exists.

Unfortunately, the above inequalities are not sufficiently strong for application to estimation of the error associated with all γ's and all Hamiltonians H met in problems of physical interest (for example, the Coulomb interaction problem defies application of the above inequalities and gives trivial results— i.e., that the error is bounded by infinity). However, the above results provide some insight into the nature of the least mean square error closure procedure.

V. RANDOM PHASE APPROXIMATION

In this section and Sec. VI we demonstrate that the least mean square error closure principle, with suitable choices of OM's and subsidiary averaging operations, is equivalent to two approximations commonly employed in statistical mechanics, viz., the random phase approximation and the linearized Kirkwood superposition approximation, respectively. In this section we investigate application of an OM spanned by single-particle functions of position and a subsidiary averaging operation $\langle\ \rangle'$ corresponding to a canonical ensemble of systems of free particles. It is shown that such a procedure applied to the approximation of a given pair function by an element in the above OM gives a result identical to the random phase approximation. When the result is averaged in a canonical ensemble of interacting particles, the approximation is equivalent to the neglect of second-order correlations.

Consider an observable manifold containing all symmetric sums of functions of the positions of individual members of a set of identical particles, confined to a volume Ω. Such a manifold is spanned

by the set of functions

$$\rho(\mathbf{r}) = \sum_{s=1}^{N} \delta(\mathbf{r} - \mathbf{r}_s), \qquad (5.1)$$

where $\mathbf{r}$ is a reference vector assuming all possible positions in the volume Ω, and where the $\mathbf{r}_s$ are the positions of the N particles. An alternative set is

$$\rho_k = \sum_{s=1}^{N} e^{-i\mathbf{k}\cdot\mathbf{r}_s}, \qquad (5.2)$$

where the wave vector $\mathbf{k}$ is defined on a lattice given by the usual cyclic (periodic) conditions at the boundaries of the volume Ω now assumed to be a cube. For the sake of simplicity the latter set of functions is used in the subsequent analysis.

We assume that the subsidiary averaging operation $\langle\ \rangle'$ is defined by the distribution function

$$P'(X) = \Omega^{-N}P(\mathbf{p}_1, \cdots, \mathbf{p}_N), \qquad (5.3)$$

where $P(\mathbf{p}_1, \cdots, \mathbf{p}_N)$ is the equilibrium distribution of momenta and where the positions $\mathbf{r}_1, \cdots, \mathbf{r}_N$ are uniformly distributed in the volume Ω. The least mean square error closure is given by the projection operator $\mathfrak{p}$ defined by

$$\mathfrak{p}\gamma(X) = \int dX' P'(X')K(X, X')\gamma(X'). \qquad (5.4)$$

With the present choice of OM and $\langle\ \rangle'$, the kernel $K(X, X')$ is given by

$$K(X, X') = 1 + N^{-1}\sum_k {}' \rho_k(X)\rho_k^*(X')$$
$$= 1 + N^{-1}\sum_{ss'} [\Omega\delta(\mathbf{r}_s - \mathbf{r}_{s'}) - 1], \qquad (5.5)$$

where the last summation is over all pairs of particle labels s and s' (including $s = s'$). Since the OM contains only those functions which are symmetric with respect to the permutation of particles, the projection operator $\mathfrak{p}$ not only replaces an arbitrary phase function by an optimal linear combination of functions of single-particle positions, but it also plays the role of a symmetrizer. For example, it can be shown that

$$\mathfrak{p}g(\mathbf{r}_1) = \frac{1}{N}\sum_{s=1}^{N} g(\mathbf{r}_s), \qquad (5.6)$$

where $g(\mathbf{r}_s)$ is an arbitrary function of the position $\mathbf{r}_s$ of particle s.

It follows directly from (3.11) that

$$\mathfrak{p}\rho_k = \rho_k, \qquad (5.7)$$

and in particular that

$$\mathfrak{p}\rho_0 = \rho_0, \qquad (5.8)$$

or

$$\mathfrak{p}1 = 1. \qquad (5.9)$$

By direct computation it can be shown that where

$\mathbf{k}$ and $\mathbf{k}'$ are both nonvanishing, the following relation holds:

$$\mathfrak{p}e^{i(\mathbf{k}\cdot\mathbf{r}_1 + \mathbf{k}'\cdot\mathbf{r}_2)} = 0, \quad \mathbf{k}, \mathbf{k}' \neq 0. \qquad (5.10)$$

These simple results form the basis for the derivation of several further results of more direct interest in classical statistical mechanics.

We consider first the phase function

$$\rho_{kk'} \equiv \rho_k\rho_{k'} - \rho_{k+k'}. \qquad (5.11)$$

Using the results of the previous paragraph, it can be readily shown that

$$\mathfrak{p}\rho_{kk'} = N(\rho_k\delta_{k'0} + \rho_{k'}\delta_{k0} - N\delta_{k0}\delta_{k'0}). \qquad (5.12)$$

The substitution of $\rho_{kk'}$ by $\mathfrak{p}\rho_{kk'}$ is recognizable, in either the averaged form or the above unaveraged form, as the random phase approximation. Transforming the above result from $(\mathbf{k}, \mathbf{k}')$ space to $(\mathbf{r}, \mathbf{r}')$ space gives

$$\mathfrak{p}\rho(\mathbf{r}, \mathbf{r}') = n(\rho(\mathbf{r}) + \rho(\mathbf{r}') - n), \qquad (5.13)$$

where $n = N/\Omega$ and where

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{ss'} {}' \delta(\mathbf{r} - \mathbf{r}_s)\delta(\mathbf{r}' - \mathbf{r}_{s'}), \qquad (5.14)$$

in which the prime on the summation denotes the omission of "self" terms for which $s = s'$.

The error incurred in the replacement of $\rho_{kk'}$ by $\mathfrak{p}\rho_{kk'}$ is found to be

$$\left.\begin{aligned}\epsilon_{kk'} &\equiv \mathfrak{q}\rho_{kk'} = \rho_{kk'}, \quad \text{if} \quad k, k' \neq 0\\ &= 0, \quad \text{otherwise}\end{aligned}\right\}. \qquad (5.15)$$

Transforming this result into $(\mathbf{r}, \mathbf{r}')$ space, we obtain

$$\epsilon(\mathbf{r}, \mathbf{r}') \equiv \mathfrak{q}\rho(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r}, \mathbf{r}') - n[\rho(\mathbf{r}) + \rho(\mathbf{r}')] + n^2. \qquad (5.16)$$

As demonstrated in Sec. III for a general OM, it is a fact that $\epsilon(\mathbf{r}, \mathbf{r}')$ not only vanishes under the averaging operation $\langle\ \rangle'$, but also possesses vanishing moments with respect to $\rho(\mathbf{r}'')$ (or, alternatively, ρ_k); that is,

$$\langle\rho(\mathbf{r}'')\epsilon(\mathbf{r}, \mathbf{r}')\rangle' = 0. \qquad (5.17)$$

However, these results no longer hold if in the above expression the averaging operation $\langle\ \rangle'$ corresponding to an equilibrium ensemble of noninteracting particles is replaced by averaging operation $\langle\ \rangle$, corresponding to an equilibrium ensemble of systems of interacting particles. Here we obtain the result

$$\langle\epsilon(\mathbf{r}, \mathbf{r}')\rangle = n^2[g(|\mathbf{r} - \mathbf{r}'|) - 1], \qquad (5.18)$$

assuming the absence of external forces. In the above expression $g(r)$ is the radial distribution function. Hence, the approximation

$$\langle\rho(\mathbf{r}, \mathbf{r}')\rangle \simeq \langle\mathfrak{p}\rho(\mathbf{r}, \mathbf{r}')\rangle \qquad (5.19)$$

is equivalent to the neglect of second-order correlations.

VI. LINEARIZED KIRKWOOD SUPERPOSITION APPROXIMATION

In this section it is shown that the use of an OM spanned by translationally invariant pair functions of position and the use of the same subsidiary averaging operation (i.e., averaging in a free-particle canonical ensemble) yields a closure approximation which when applied to a triplet function of position is, after canonical averaging, identical to the linearized Kirkwood superposition approximation.

In this case we choose an OM spanned by the functions

$$1, \sigma_k (\equiv \rho_k \rho_{-k} - \rho_0), \qquad (6.1)$$

where $\mathbf{k}$ assumes the same discrete set of vector values as before, except that now we retain only half of $\mathbf{k}$ space, i.e., the largest set not transformable into itself by reflection ($\mathbf{k} \to -\mathbf{k}$). This OM is equivalently spanned by the functions

$$1, \sum_{ss'}{}' \delta(\mathbf{r} - \mathbf{r}_s + \mathbf{r}_{s'}), \qquad (6.2)$$

where $\mathbf{r}$ assumes all positions in the volume Ω. Still employing the same subsidiary averaging operation $\langle\ \rangle'$ as before, the projection operator for the new OM is defined by the kernel

$$K(X, X') = 1 + \frac{1}{2N(N-1)}$$

$$\times \sum_{s_1 s_2}{}' \sum_{s_3 s_4}{}' [\Omega \delta(\mathbf{r}_{s_1} - \mathbf{r}_{s_2} + \mathbf{r}'_{s_3} - \mathbf{r}'_{s_4}) - 1]. \quad (6.3)$$

We consider the operation by $\mathfrak{p}$ on a triplet function of position. It is sufficient to consider $\exp(i\mathbf{k}_1 \cdot \mathbf{r}_1 + i\mathbf{k}_2 \cdot \mathbf{r}_2 + i\mathbf{k}_3 \cdot \mathbf{r}_3)$, where $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$ are all non-vanishing. We obtain the result

$$\mathfrak{p} \exp(i\mathbf{k}_1 \cdot \mathbf{r}_1 + i\mathbf{k}_2 \cdot \mathbf{r}_2 + i\mathbf{k}_3 \cdot \mathbf{r}_3) = 0. \quad (6.4)$$

With the use of this result, a straightforward calculation yields

$$\mathfrak{p}\rho_{k_1}\rho_{k_2}\rho_{k_3} = \delta^K(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$$

$$\times (\sigma_{k_1} + \sigma_{k_2} + \sigma_{k_3} + N), \quad (6.5)$$

where $\delta^K(\mathbf{u})$ is the Kronecker delta function of $\mathbf{u}$.

The canonical average of (6.5) may be shown, by transforming back to $\mathbf{r}$ space, to be identical to the linearized Kirkwood superposition approximation. However, the relation between (6.4) and the linearized Kirkwood superposition approximation may be shown more easily. Let us write the configurational distribution function for a triplet of particles in the form

$$f^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \Omega^{-3}[1 + w^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + w^{(2)}(\mathbf{r}_2, \mathbf{r}_3)$$

$$+ w^{(2)}(\mathbf{r}_3, \mathbf{r}_1) + w^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)], \quad (6.6)$$

where $w^{(2)}$ is defined by the relation $f^{(2)}(\mathbf{r}_1, \mathbf{r}_2) =$

$\Omega^{-2}[1 + w^{(2)}(\mathbf{r}_1, \mathbf{r}_2)]$. The approximation in question is the neglect of $w^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. By direct computation we obtain the result

$$\langle \exp[i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2 + \mathbf{k}_3 \cdot \mathbf{r}_3)] \rangle$$

$$= \Omega^{-3} \iiint d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 w^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$$

$$\times \exp[i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2 + \mathbf{k}_3 \cdot \mathbf{r}_3)]. \quad (6.7)$$

According to (6.4), the replacement of $\exp[i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2 + \mathbf{k}_3 \cdot \mathbf{r}_3)]$ by the same quantity operated upon by $\mathfrak{p}$ is the same as the neglect of this quantity altogether; therefore, it follows that the average of this quantity must also be neglected. According to (6.7), this is equivalent to the neglect of $w^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. Thus the approximation of a triplet function by $\mathfrak{p}$ operating on this function is equivalent to the linearized Kirkwood superposition approximation.

The above analysis shows that the linearized Kirkwood superposition principle has another basis (i.e., the least mean square error closure principle using a pair function OM and free-particle subsidiary averaging). Although the Kirkwood superposition principle is superior to its linearized version in certain qualitative features, it is wrong to conclude that the validity of the linearized version is equivalent to the validity of linearization. It is our conviction that the linearized version is valid in a broader domain than is the process of linearization.

The applications of the general closure principle illustrated in this section and in Sec. V can obviously be extended to higher order—for example, the approximation of a four-particle distribution function by a linear combination of lower-order distribution functions. The treatment of higher-order closure approximations will be treated in a later communication.

VII. A GENERAL HAMILTONIAN AND ITS RELATION TO THERMODYNAMIC PROPERTIES

The application of the linear closure techniques described above to the approximate calculation of thermodynamic properties requires the use of an appropriate formalism connecting the average values of observables with the free energy. In the description of this aspect of the problem we consider the general set of observable $S = (\alpha_1, \cdots, \alpha_n)$. The first step is the construction of a Hamiltonian which is more general than that of direct physical interest. Accordingly, we write

$$H = H_0 + \sum_j b_j \alpha_j + \tfrac{1}{2} \sum_{jk} c_{jk} \alpha_j^* \alpha_k, \quad (7.1)$$

where H_0 is a Hamiltonian of simpler structure than

H and the quantities b_j and c_{jk} are parameters which are generally complex. It is essential that the domain of variation of these parameters includes the points representing the Hamiltonian(s) of physical interest. As a matter of formal convenience, it is assumed that c_{jk} and c_{kj} are not necessarily equal. Furthermore, the parameters will not generally be constrained in such a way that H is real, although the parameters of direct physical significance must of course satisfy such a constraint.

The appropriate constraints can be expressed as follows. We first note that the assumed invariance of the OM to complex conjugation implies that α_j^* is a linear combination of the α_k, i.e.,

$$\alpha_j^* = \sum_k C_{jk}\alpha_k. \qquad (7.2)$$

The conditions for the reality of H can then be written

$$b_j^* = \sum b_k D_{kj}, \qquad (7.3a)$$

$$c_{jk}^* = c_{kj}, \qquad (7.3b)$$

where D_{jk} is the matrix inverse of C_{jk}.

In the petite canonical ensemble, the Helmholtz free energy A, corresponding to H, is defined by

$$e^{-\beta A} = \int dX e^{-\beta H} \qquad (7.4)$$

and the average value of an arbitrary function of coordinates and momenta $f(X)$ is

$$\langle f \rangle = \int dX f e^{-\beta H} \Big/ \int dX e^{-\beta H}$$

$$= \int dX f e^{\beta(A-H)}. \qquad (7.5)$$

In this treatment the customary factor $(h^{3N}N!)^{-1}$ in front of the integrals is omitted for convenience.

Clearly, the Helmholtz free energy is a function of β, b_j, and c_{jk}. It is well known that

$$\partial(\beta A)/\partial\beta = \langle H \rangle, \qquad (7.6a)$$

and it is easy to show that

$$\partial A/\partial b_j = \langle \partial H/\partial b_j \rangle = \langle \alpha_j \rangle, \qquad (7.6b)$$

$$\partial A/\partial c_{jk} = \langle \partial H/\partial c_{jk} \rangle = \tfrac{1}{2}\langle \alpha_j^* \alpha_k \rangle. \qquad (7.6c)$$

Thus, the differential of βA is given by

$$d(\beta A) = \langle H \rangle \, d\beta + \beta$$
$$\times \left(\sum_j \langle \alpha_j \rangle \, db_j + \tfrac{1}{2}\sum_{jk} \langle \alpha_j^* \alpha_k \rangle \, dc_{jk} \right). \qquad (7.7)$$

It is easy to show that

$$\partial\langle\alpha_j\rangle/\partial b_k = \partial\langle\alpha_k\rangle/\partial b_j = -\beta\langle\Delta\alpha_j\Delta\alpha_k\rangle, \qquad (7.8)$$

where $\Delta\alpha_j = \alpha_j - \langle\alpha_j\rangle$. Multiplication by C_{ij} and summation on j yields

$$\partial\langle\alpha_j^*\rangle/\partial b_k = \partial\langle\alpha_k\rangle/\partial \bar{b}_j = -\beta\langle\Delta\alpha_j^*\Delta\alpha_k\rangle, \qquad (7.9)$$

where $\bar{b}_j$ is defined by

$$\bar{b}_j = \sum_k b_k D_{kj}. \qquad (7.10)$$

If the constraints (7.3a) ensuring the reality of H are imposed, then $\bar{b}_j$ and b_j^* are equal. Using (7.6c) and (7.9), we can derive the result

$$\frac{\partial A}{\partial c_{jk}} = \frac{1}{2}\left[\langle\alpha_j^*\rangle\langle\alpha_k\rangle - \frac{1}{\beta}\frac{\partial\langle\alpha_j^*\rangle}{\partial b_k} \right]. \qquad (7.11)$$

Certain differential relations connecting the Helmholtz free energy with the average observables have now been established. We now consider the integration of these relations in order to obtain explicit expressions for the Helmholtz free energy.

Procedure 1: Let the Hamiltonian of interest be given by

$$\hat{H} = H_0 + \sum_j \bar{b}_j \alpha_j.$$

Here it has been assumed that the set of observables is sufficiently extensive that the term

$$\sum_j \bar{b}_j \alpha_j$$

includes the interaction energy. The procedure here is to integrate (7.6b) along the following path in parameter space:

$$\beta = \text{const},$$
$$c_{jk} = 0, \qquad (7.12)$$
$$b_j = \lambda\bar{b}_j, \quad 0 \le \lambda \le 1.$$

We obtain

$$\hat{A} = A_0 + \sum_j \bar{b}_j \int_0^1 d\lambda\langle\alpha_j\rangle, \qquad (7.13)$$

where

$$e^{-\beta A_0} = \int dX e^{-\beta H_0}. \qquad (7.14)$$

It is understood that the $\langle\alpha_j\rangle$ are evaluated at $\beta = \beta$, $a_j = \lambda a_j$, $c_{jk} = 0$.

Procedure 2: In this case the Hamiltonian of interest is assumed to be given by

$$\hat{H} = H_0 + \sum_j \bar{b}_j \alpha_j + \tfrac{1}{2}\sum_{jk} \hat{c}_{jk}\alpha_j^*\alpha_k.$$

Here, as a minimal requirement, the set of observables α_j need only be sufficiently extensive for the quadratic term in $\hat{H}$ to represent the interaction energy. Equation (7.7) is now integrated along the path

$$\beta = \text{const},$$
$$b_j = \lambda\bar{b}_j,$$
$$c_{jk} = \lambda\hat{c}_{jk}, \qquad (7.15)$$
$$0 \le \lambda \le 1.$$

We obtain

$$\hat{A} = A_0 + \int_0^1 d\lambda \left\langle \frac{\partial H}{\partial \lambda} \right\rangle$$

$$= A_0 + \int_0^1 d\lambda \left[\sum_j \hat{b}_j \langle \alpha_j \rangle + \tfrac{1}{2} \sum_{kj} \hat{c}_{jk} \right.$$

$$\left. \times \left(\langle \alpha_j^* \rangle \langle \alpha_k \rangle - \frac{1}{\beta} \frac{\partial \langle \alpha_j^* \rangle}{\partial b_k} \right) \right]. \quad (7.16)$$

In order to carry out the explicit calculation, one must know the quantities $\langle \alpha_j \rangle$ and $\partial \langle \alpha_j \rangle / \partial b_k^*$. The complex conjugates can be obtained through the use of (7.2) and (7.3a). It is, of course, understood that in (7.16) the quantities $\langle \alpha_j \rangle$ and $\partial \langle \alpha_j^* \rangle / \partial b_k$ are evaluated at a point on the contour (7.15) corresponding to λ.

VIII. THE "EQUATION OF MOTION" METHOD

The key feature of this method is the estimation of the average observables as functions of the parameters β, b_j, and c_{jk} by the use of equations of motion of the α_j truncated by the linear closure technique discussed in Secs. III through VI.

As a representative point in phase space moves according to Hamilton's equations [using the general Hamiltonian (7.1)], the corresponding time dependence of the observables α_j is given by

$$\dot{\alpha}_j = [\alpha_j, H] = \mathfrak{L}\alpha_j, \quad (8.1)$$

where $[u, v]$ is the Poisson bracket of u and v and $\mathfrak{L}$ is the Liouville operator. If the right-hand side of (8.1) were a linear combination of the α_j (i.e., if it lay in the OM), the exact solution would be obtainable directly. Furthermore, the canonical average defined by (7.5) of both sides of (8.1) would vanish, and one could then deduce the $\langle \alpha_j \rangle$ exactly. However, all of the $\mathfrak{L}\alpha_j, j = 1, \cdots, n$ lie in the OM only in special trivial cases. In most cases of interest, we are obliged to approximate the expressions $\mathfrak{L}\alpha_j$ in some manner. Using the least mean square error closure procedure, we obtain

$$\dot{\alpha}_j \simeq \mathrm{p}\mathfrak{L}\alpha_j = \sum_{kl} \alpha_k Q_{kl} \langle \alpha_l^* \mathfrak{L}\alpha_j \rangle', \quad (8.2)$$

where Q_{jk} is the matrix reciprocal of $\langle \alpha_j^* \alpha_k \rangle$ and $\langle \ \rangle'$ denotes, as before, the subsidiary averaging process used in the computation of the mean square error. There is an interesting correspondence between Eq. (8.2) and the work of Zwanzig[3] on the approximate eigenvalues and eigenfunctions of the Liouville operator $\mathfrak{L}$. If in his formalism we take his weighting function to be the subsidiary distribution function $P'(X)$ and restrict his approximate eigenfunction to be a point in the OM, then the formal solution of his variational problem is the same as the equation giving the normal modes of (8.2).

The averaging of (8.2) in the actual ensemble then gives the approximate result

$$\langle \mathrm{p}\mathfrak{L}\alpha_j \rangle = \sum_{kl} \langle \alpha_k \rangle Q_{kl} \langle \alpha_l^* \mathfrak{L}\alpha_j \rangle' = 0. \quad (8.3)$$

Since the equations above are homogeneous in the $\langle \alpha_k \rangle$, it would appear that the correct solution would be $\langle \alpha_k \rangle = 0$ in all but singular cases. However, since the OM is required to contain unity, it follows that at least one linear combination of the α_k is known in advance. Assuming for the sake of simplicity that only one linear combination of the α_k is known and that this one linear combination is simply a single member α_1 of the observable set (element of the OM), it follows that the equation labeled by $j = 1$ in the set (8.3) is redundant and should be discarded. The remaining set $(j = 2, \cdots, n)$ are not homogeneous in the quantities $\langle \alpha_2 \rangle, \cdots, \langle \alpha_n \rangle$, and hence they can be solved for these quantities (except perhaps for singular cases).

In many problems the $\mathfrak{L}\alpha_j$ will have vanishing projections on the OM. In this case, it is appropriate to consider the alternative approximate equations

$$\ddot{\alpha}_j \simeq \mathrm{p}\mathfrak{L}^2\alpha_j \quad (8.4)$$

and replace $\mathfrak{L}$ by $\mathfrak{L}^2$ everywhere in the equations subsequent to (8.2).

In many cases, the exact form of the canonically averaged equations

$$\langle \mathfrak{L}\alpha_j \rangle = 0 \quad (\text{or } \langle \mathfrak{L}^2\alpha_j \rangle = 0) \quad (8.5)$$

is well known (e.g., the thermal BBGKY hierarchy). In this case the least mean square error closure approximation can be applied in its canonically averaged form; that is, $\langle \mathfrak{L}\alpha_j \rangle$ (or $\langle \mathfrak{L}^2\alpha_j \rangle$) is to be replaced by $\langle \mathrm{p}\mathfrak{L}\alpha_j \rangle$ (or $\langle \mathrm{p}\mathfrak{L}^2\alpha_j \rangle$).

IX. TREATMENT OF THE CLASSICAL ELECTRON GAS

In this section, the closure approximations discussed in Secs. V and VI are applied through the use of the formalism of Secs. VII and VIII to the calculation of the thermodynamic properties of an illustrative system. We have chosen a classical system of N electrons in a uniform background of compensating charge. The Hamiltonian of such a system is

$$\hat{H} = H_0 + \tfrac{1}{2} \sum_k{}' \hat{v}_k(\rho_k \rho_{-k} - \rho_0), \quad (9.1)$$

where

$$H_0 = \frac{1}{2m} \sum_{s=1}^N P_s^2 \quad (9.2)$$

is the kinetic energy of the system of electrons. The quantities ρ_k are defined by

$$\rho_k = \sum_{s=1}^N e^{-i\mathbf{k}\cdot\mathbf{r}_s}, \quad (9.3)$$

[3] R. Zwanzig, Phys. Rev. **144**, 170 (1966).

where the r_s are the positions of the electrons labeled by $s = 1, 2, \cdots, N$. The quantity $\hat{v}_k$ is the Fourier transform of the interaction potential and, in the present case of Coulomb interaction, it is given by

$$\hat{v}_k = 4\pi e^2/\Omega k^2, \qquad (9.4)$$

where Ω is the volume of the system. The discrete spectrum of values of $\mathbf{k}$ is determined by the usual periodic (cyclic) boundary conditions. The prime on the summation in (9.1) denotes the omission of the term for $\mathbf{k} = 0$, which implies electrical neutrality for the system as a whole.

Although Procedure 2 (see Sec. VII) is conceptually more complicated, it is simpler to work out; for this reason, it is considered first. Here, the observables are chosen to be the function ρ_k for all $\mathbf{k}$. The subsidiary averaging operation $\langle \quad \rangle'$ is chosen to be the averaging in the free-particle canonical ensemble, i.e.,

$$\langle g(X) \rangle' = \int dX g(X) e^{-\beta H_0} \bigg/ \int dX e^{-\beta H_0}. \quad (9.5)$$

In the case of Procedure 2 we consider a more general Hamiltonian than (9.1):

$$H = H_0 + \sum_k \phi_k^* \rho_k + \tfrac{1}{2} \sum_k{}' v_k(\rho_k \rho_{-k} - \rho_0), \quad (9.6)$$

where the term $\sum \phi_k^* \rho_k$ corresponds to an arbitrary potential of external force and where $v_k = \xi \hat{v}_k$.

The next part of the procedure is to consider the equations of motion of the ρ_k using the general Hamiltonian H. Because the ρ_k are even in the momenta, it follows that $\mathfrak{L}\rho_k$ will be odd in the momenta; therefore, the latter quantities will have vanishing projections in the OM. For these reasons we consider the second time derivatives of the ρ_k:

$$\ddot{\rho}_k = \mathfrak{L}^2 \rho_k. \qquad (9.7)$$

Application of the least mean square error closure principle to the right-hand side of (9.7) gives

$$\ddot{\rho}_k \simeq \mathfrak{p}\mathfrak{L}^2 \rho_k = -(c^2 k^2 + \omega_p^2 \xi)\rho_k - (k^2 \rho_0/m)\phi_k, \quad (9.8)$$

where $c = 1/(\beta m)^{\frac{1}{2}}$ is the isothermal velocity of sound and

$$\omega_p = (4\pi e^2 N/m\Omega)^{\frac{1}{2}}$$

is the plasma frequency. After canonical averaging we obtain the approximate result

$$0 = -(c^2 k^2 + \omega_p^2 \xi)\langle \rho_k \rangle - (k^2 \rho_0/m)\phi_k$$

or

$$\langle \rho_k \rangle = -\beta \rho_0 \phi_k/(1 + \xi/\lambda_D^2 k^2), \qquad (9.9)$$

where $\lambda_D = (4\pi e^2 \beta N/\Omega)^{\frac{1}{2}}$. The above equations [(9.8) and (9.9)] have been linearized with respect to the ϕ_k since the final result involves only small values of the ϕ_k. With the use of (7.6c) and (7.9) of Sec. VII, we obtain

$$\partial A/\partial \xi = -\tfrac{1}{2} N e^2 \xi^{\frac{1}{2}}/\lambda_D \qquad (9.10)$$

and

$$A - A_0 = \int_0^1 d\xi \frac{\partial A}{\partial \xi} = -\frac{1}{3}\frac{Ne^2}{\lambda_D}, \qquad (9.11)$$

the well-known result of Debye.[4]

In the case of Procedure 1, we choose a different set of observables, viz., ρ_0 and $\sigma_k = \rho_k \rho_{-k} - \rho_0$ for all unvanishing $\mathbf{k}$, but employ the same subsidiary averaging operation $\langle \quad \rangle'$. The general Hamiltonian is now chosen to be

$$H = H_0 + \tfrac{1}{2} \sum{}' v_k \sigma_k, \qquad (9.12)$$

where, as before, $v_k = \xi \hat{v}_k$.

The equation of motion for the σ_k (using the second time derivative for the same reason as before) is

$$\ddot{\sigma}_k = \mathfrak{L}^2 \sigma_k. \qquad (9.13)$$

Applying canonical averaging, we obtain

$$\left.\begin{aligned}
\langle \ddot{\sigma}_k \rangle = 0 = &-(2k^2/m\beta)\langle \sigma_k \rangle \\
&- \frac{\mathbf{k}}{m\Omega} \cdot \sum_q{}' \mathbf{q} v_q (\langle \rho_q \rho_{-q+k} \rho_{-k} \rangle + \langle \rho_{-q} \rho_{q-k} \rho_k \rangle)
\end{aligned}\right\}. $$
$$(9.14)$$

The least mean square error closure procedure applied to the unaveraged product $\rho_q \rho_{-q+k} \rho_{-k}$ gives

$$\mathfrak{p}\rho_q \rho_{-q+k} \rho_{-k} = \sigma_q + \sigma_{-q+k} + \sigma_{-k} + \rho_0. \quad (9.15)$$

It was also shown in Sec. VI that this result is the Fourier transform of the linearized version of the Kirkwood superposition approximation. Substitution of this approximation for the triple product of ρ's gives

$$0 = \frac{k^2}{m}\left(\frac{1}{\beta} + \frac{N v_k}{\Omega}\right)\langle \sigma_k \rangle - \frac{k^2 N^2}{m\Omega}$$
$$+ \frac{1}{m\Omega}\mathbf{k} \cdot \sum_q \mathbf{q} v_k \langle \sigma_k \rangle, \quad q \neq \pm k. \quad (9.16)$$

Transforming the last result into $\mathbf{r}$ space we obtain the more familiar result

$$\frac{d^2 f}{dr^2} + \left(\frac{2}{r} - \frac{\gamma \Lambda_D}{r^2}\right)\frac{df}{dr} - \frac{f}{\Lambda_D^2} = 4\pi(f - 1)\delta(\mathbf{r}), \quad (9.17)$$

where $f(r) = 1 - g(r)$, the radial distribution function, is related to the $\langle \sigma_k \rangle$ by the expression

$$g(r) = \frac{1}{N(N-1)}\sum_k \langle \sigma_k \rangle e^{i\mathbf{k}\cdot\mathbf{r}}. \qquad (9.18)$$

In (9.17) $\Lambda_D = (4\pi\beta\rho e^2 \xi)^{-\frac{1}{2}}$ is the Debye length and $\gamma = e^2 \beta \Lambda_D^{-1}$ is the dimensionless plasma parameter, both quantities corresponding to an electron charge reduced by a factor $\xi^{\frac{1}{2}}$.

In another paper,[5] Eq. (9.17) is derived (directly from the BBGKY hierarchy) and solved to yield thermodynamic properties over a wide range of γ.

[4] See, for example, I. Z. Fischer, *Statistical Theory of Liquids* (University of Chicago Press, Chicago, 1964).
[5] L. C. Levitt, J. M. Richardson, and E. R. Cohen, Phys. Fluids **10**, 406 (1967).

Expansion of the T Matrix for Resonance Collisions*

DAVID A. MICHA†

Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

The eigenfunctions of the kernel of the Lippmann–Schwinger collision equation, corresponding to outgoing waves in all channels, are used to obtain an expansion of the T matrix valid for multichannel collisions, including rearrangements. From this expansion, the transition amplitude in the general case of overlapping resonances is obtained and a characterization of bound states and resonance states is given.

I. INTRODUCTION

ONE of the most important methods for the description of resonances in collision processes was developed originally by Kapur and Peierls[1] for finite range potentials; it is based on a modification of the usual boundary conditions for scattering. Later on Siegert[2] described a similar method for a single channel s-wave collision that had the advantage of avoiding the dependence of the resonance parameters on the initial relative kinetic energy of the particles and on the radius chosen to impose the boundary conditions. His approach was extended by Humblet and Rosenfeld[3] to the multichannel case, by using a Mittag–Leffler expansion for the collision matrix. More recently Herzenberg and Mandl[4] proposed an alternative expansion in terms of the eigenfunctions of a modified Schrödinger equation with the Kapur–Peierls boundary conditions, which is more explicit than the Mittag–Leffler expansion.

In the present contribution we want to preserve these advantages, using the Lippmann–Schwinger[5] equations of collision theory as a starting point. The origin of the modified Schrödinger equation is then clearly seen, the expression for the T matrix is valid for multichannel collisions, including rearrangements, and it provides a unified description of bound states and resonances. Our main purpose is to present a simple and yet general formalism to describe resonance collisions, rather than to provide mathematically rigorous proofs. Section II deals with the expansion of the T matrix, which is used in Sec. III to obtain the resonance formula and to characterize bound states and resonances. Section IV discusses briefly the connection with other approaches.

II. EXPANSION OF THE T MATRIX

Let us consider a collision process described by a Hamiltonian

$$H = H_0 + V = K + h + V, \tag{1}$$

where K is the relative kinetic energy of the incoming particles, h the Hamiltonian for their internal motion, and V the interaction potential. We can also write, in terms of the variables for the outgoing particles,

$$H = H_0' + V' = K' + h' + V', \tag{2}$$

with the corresponding meaning for K', h', and V'. Defining a channel as the set of all the internal states corresponding to given particles, we restrict the treatment to channels with not more than two particles, whose interaction potential is zero for the interparticle distances $r > r_0$, or $r' > r_0'$. If P is an operator equal to one when $r \leq r_0$ and $r' \leq r_0'$ and zero otherwise, then

$$V = PVP \quad \text{and} \quad V' = PV'P. \tag{3}$$

The transition amplitude between the states $\varphi_a(E)$ and $\varphi_{a'}(E)$, solutions of

$$(E - H_0)\varphi_a = 0 \quad \text{and} \quad (E - H_0')\varphi_{a'} = 0 \tag{4}$$

for the noninteracting incoming and outgoing particles, respectively, is given by the T matrix element $T_{a'a}$. It is the same[6] whether we use the operator

$$T = V + V'G_0'T, \tag{5}$$

with $G_0'(E^+) = (E + i\epsilon - H_0')^{-1}$, $\epsilon \to 0+$, or the similar operator

$$T' = V' + T'G_0V. \tag{6}$$

† Present address: Department of Physics and Institute for Radiation Physics and Aerodynamics, University of California, San Diego, California.

[1] P. L. Kapur and R. Peierls, Proc. Roy. Soc. (London) A166, 277 (1938).
[2] A. J. F. Siegert, Phys. Rev. 56, 750 (1939).
[3] J. Humblet, Mem. Soc. Roy. Sci. Liege XII, 9 (1952); J. Humblet and L. Rosenfeld, Nucl. Phys. 26, 529 (1961); see also R. G. Newton, J. Math. Phys. 1, 319 (1960).
[4] A. Herzenberg and F. Mandl, Phys. Letters 6, 288 (1965); A. Herzenberg, K. L. Kwok, and F. Mandl, Proc. Phys. Soc. (London) 84, 477 (1964).
[5] B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 469 (1950); B. A. Lippmann, ibid. 102, 264 (1956).
[6] M. Gell-Mann and M. L. Goldberger, Phys. Rev. 91, 398 (1953).

In what follows we use Eq. (6), which is equivalent to

$$T'(z) = V'[1 - G_0(z)V]^{-1}, \qquad (7)$$

where z indicates a complex energy. We can now generalize previous treatments by introducing the solutions of the equation

$$\eta(z)|\psi(z)\rangle = G_0(z)V|\psi(z)\rangle \qquad (8)$$

with the boundary conditions

$$\langle\psi(z)|\, P\, |\psi(z)\rangle = N(z) \quad \text{(finite)} \qquad (9)$$

and with $|\psi(z)\rangle$ giving asymptotically outgoing waves for *all* the channel states. This is an eigenvalue problem with complex boundary conditions which defines eigenfunctions $|\psi_n(z)\rangle$ and complex eigenvalues $\eta_n(z)$. Indicating with $\mathcal{K}$ the time reversal operator, we write

$$\mathcal{K}\,|\psi_n(z)\rangle = |\bar{\varphi}_n(z^*)\rangle \qquad (10)$$

and obtain from Eq. (8), for the eigenvalue $\eta_m(z)$,

$$\langle\bar{\psi}_n(z^*)|\, V\, |\psi_m(z)\rangle\eta_m(z)$$
$$= \langle\bar{\psi}_n(z^*)|\, VG_0(z)V\, |\psi_m(z)\rangle. \quad (11)$$

Applying $\mathcal{K}$ to the Hermitian conjugate of Eq. (8) written for $\eta_n(z)$, and then taking the scalar product with $V\,|\psi_m(z)\rangle$, we get, after subtracting the result from Eq. (11),

$$[\eta_m(z) - \eta_n(z)]\langle\bar{\psi}_n(z^*)|\, V\, |\psi_m(z)\rangle = 0, \qquad (12)$$

which shows that we can always choose

$$\langle\bar{\psi}_n(z^*)|\, V\, |\psi_m(z)\rangle = \delta_{nm}h_m(z) \qquad (13)$$

to replace the usual orthogonality condition. We also find immediately from

$$\eta_m(z) = (\langle\psi_m(z^*)|\, V\, |\psi_m(z)\rangle)^{-1}$$
$$\times \langle\psi_m(z^*)|\, VG_0(z)V\, |\psi_m(z)\rangle \quad (14)$$
that

$$\eta_m(z^*) = [\eta_m(z)]^*. \qquad (15)$$

Using Eq. (15) we can write the completeness relation for the eigenfunctions of Eq. (8) in the form

$$1 = \sum_m |\psi_m(z)\rangle[h_m(z)]^{-1}\langle\bar{\psi}_m(z^*)|\, V, \qquad (16)$$

valid when applied to any state having, in general, incoming and outgoing waves at infinity.[1] From this relation we obtain an expansion for $T'(z)$:

$$T'(z) = \sum_m [1 - \eta_m(z)]^{-1}V'\, |\psi_m(z)\rangle$$
$$\times [h_m(z)]^{-1}\langle\bar{\psi}_m(z^*)|\, V. \quad (17)$$

Equation (17) is the main result of this section. It provides a general expression for $T'(z)$ from which the transition amplitude could be obtained by putting

$z = E^+ = E + i\epsilon$, $\epsilon \to 0+$. Nevertheless, the importance of this equation seems to arise from its connection with the description of resonance collisions, which is studied in the following section.

III. THE RESONANCE FORMULA

Let $\varphi_a(z)$ be the state obtained from $\varphi_a(E)$ by giving complex values to the energy and restricting the corresponding wavenumbers to values with a positive real part. The matrix elements

$$T_{a'a}(z) = \langle\varphi_{a'}(z)|\, T'(z)\, |\varphi_a(z)\rangle \qquad (18)$$

will, in general, have a cut for positive real z since for those values the operator to the right in Eq. (8) is unbounded. Also, it is seen from Eq. (17) that $T_{a'a}(z)$ will have poles for $z_n = E_n - \tfrac{1}{2}i\Gamma_n$ with E_n and Γ_n real, such that

$$\eta_n(z_n) = 1, \qquad (19)$$

in which case Eq. (8) reduces to

$$(z_n - H)\, |\psi_n(z_n)\rangle = 0. \qquad (20)$$

Accordingly, E_n and Γ_n are interpreted as the energy and linewidth for the pair of particles in the compound state $|\psi_n(z_n)\rangle$. Calling W_α the αth eigenvalue of h, the complex channel wavenumbers k_α corresponding to the energy z are given by

$$k_\alpha = [(2\mu/\hbar^2)(z - W_\alpha)]^{\frac{1}{2}} = \kappa_\alpha - i\lambda_\alpha, \qquad (21)$$

where μ is the reduced mass of relative motion and we choose $\kappa_\alpha \geq 0$ to make k_α a uniform function of z. In terms of these quantities we can write

$$E_n = W_\alpha + (\hbar^2/2\mu)(\kappa_{\alpha n}^2 - \lambda_{\alpha n}^2)$$
$$\text{and} \quad \Gamma_n = (2\hbar^2/\mu)\kappa_{\alpha n}\lambda_{\alpha n}. \quad (22)$$

Besides, Eqs. (15) and (19) imply $\eta_n(z_n^*) = 1$, so that z_n^* will also be a resonance pole. Indicating with k_n the set of wavenumbers $\{k_{\alpha n}\}$ and putting $|\psi_n(z_n)\rangle = |\psi_{k_n}\rangle$, we find that the compound states $|\psi_{-k_n^*}\rangle$ are the ones corresponding to z_n^*. They contain only incoming waves at infinity and describe the compound state in (20) when time is reversed.[7] It follows from this that the restriction $\kappa_\alpha \geq 0$ does not imply any lack of generality.

Expanding the nth term of $T_{a'a}(z)$ in a Laurent's series at z_n, we get

$$T_{a'a}(z) = T_{a'a}^{(n)}(z) + T'_{a'a}(z), \qquad (23)$$

---

[7] The compound states $|\psi k_n^*\rangle$ must be ruled out because, in accordance with the last paragraph of this section, they should correspond to bound states if the $|\psi_{k_n}\rangle$ were resonance states, and reciprocally.

where

$$T_{a'a}^{(n)}(z) = (z - z_n)^{-1}(-d\eta_n/dz_n)^{-1}$$
$$\times \langle\varphi_{a'}(z_n)| V' |\psi_n(z_n)\rangle[h_n(z_n)]^{-1}$$
$$\times \langle\bar\psi_n(z_n^*)| V |\varphi_a(z_n)\rangle \quad (24)$$

is the singular contribution or the resonance part, while $T_{a'a}'(z)$ is regular around z_n and it is called the potential part. To find $d\eta_n/dz_n$ we write Eq. (8) in the equivalent form

$$\{H_0 + [\eta_n(z)]^{-1}V - z\} |\psi_n(z)\rangle = 0, \quad (25)$$

which implies

$$0 = (d/dz)\langle\bar\psi_n(z^*)| [P(z - H_0)\eta_n(z) - V] |\psi_n(z)\rangle. \quad (26)$$

Differentiation for $z = z_n$ leads immediately to

$$-d\eta_n/dz_n = [h_n(z_n)]^{-1}[\langle\bar\psi_n(z_n^*)| P |\psi_n(z_n)\rangle$$
$$- \langle\bar\psi_n(z_n^*)| [P, H_0] |d\psi_n(z)/dz_n\rangle. \quad (27)$$

This quantity is different from zero except for very particular potentials, so that z_n can be considered in general as a single root of Eq. (19). Since $T_{a'a}^{(n)}(z)$ is well-defined for real z, we can obtain the transition amplitude putting $z = E^+$ in Eq. (23), provided $\Gamma_n < |E_n - E_{n\pm1}|$. If, on the other hand, $\Gamma_n \geq |E_n - E_{n\pm1}|$, the resonance behavior at energy E will change due to contributions from overlapping resonances contained in $T_{a'a}'(z)$. Expanding each term of $T_{a'a}(z)$ contributing to the overlap in a Laurent's series at z_n, the general resonance formula is

$$T_{a'a}(E) = T_{a'a}''(E) + \sum_n{}' \left(E - E_n + \frac{i}{2}\Gamma_n\right)^{-1}$$
$$\times \frac{\langle\varphi_{a'}(z_n)| V' |\psi_n(z_n)\rangle\langle\bar\psi_n(z_n^*)| V |\varphi_a(z_n)\rangle}{\langle\bar\psi_n(z_n^*)| P |\psi_n(z_n)\rangle - \langle\bar\psi_n(z_n^*)| [P, H_0] |d\psi_n/dz_n\rangle} \quad (28)$$

with the sum extending over overlapping resonances. This equation contains as particular cases the Bethe–Peierls formula[8] for a resonance at very low energies and the Breit–Wigner formula[9] for narrow resonances. Furthermore, E_n and Γ_n are given explicitly in terms of $|\psi_n(z_n)\rangle$ by the relations

$$E_n = [N_n(z_n)]^{-1}(\langle\psi_n(z_n)| V |\psi_n(z_n)\rangle$$
$$+ \langle\psi_n(z_n)| \tfrac{1}{2}(PH_0 + H_0P) |\psi_n(z_n)\rangle) \quad (29a)$$

and

$$\Gamma_n = [N_n(z_n)]^{-1}\langle\psi_n(z_n)| i[P, H_0] |\psi_n(z_n)\rangle, \quad (29b)$$

which are obtained by taking the scalar product of Eq. (20) with $\langle\psi_n(z_n)| P$. By a similar procedure we

also obtain

$$\langle\bar\psi_n(z_n^*)| V |\varphi_a(z_n)\rangle = \langle\bar\psi_n(z_n^*)| [P, H_0] |\varphi_a(z_n)\rangle \quad (30a)$$

and

$$\langle\varphi_{a'}(z_n)| V' |\psi_n(z_n)\rangle = \langle\varphi_{a'}(z_n)| [H_0', P] |\psi_n(z_n)\rangle. \quad (30b)$$

Since the operator $[P, H_0]$ (the operator $[H_0', P]$) is different from zero only at $r = r_0$ (at $r' = r_0'$), the left-hand sides of Eqs. (29b), (30a), and (30b) can be expressed in terms of the values of the wavefunctions at the potential boundaries, which are usually called channel amplitudes.

We can now get some more results and the connection with previous developments by making explicit use of the condition of outgoing waves at infinity. Indicating with $u_\alpha(\mathbf{x})$ the αth eigenfunction of h, we define, for the incoming particles,

$$u_c(S) = Y_l^m(\omega)u_\alpha(\mathbf{x}), \quad (31)$$

where ω gives the direction of the interparticle vector $\mathbf{r}$ and $c = (\alpha lm)$. With similar definitions for the outgoing particles, the wavefunction $\langle\mathbf{r}, \mathbf{x} | \psi\rangle = \psi(\mathbf{r}, \mathbf{x})$ is given by

$$\psi_{\text{int}}(\mathbf{r}, \mathbf{x}) = P\psi(\mathbf{r}, \mathbf{x}) = \sum_c \frac{y_c(r)}{r} u_c(S),$$
$$= \sum_{c'} \frac{y_{c'}(r')}{r'} u_{c'}(S') \quad (32a)$$

and

$$\psi_{\text{ext}}(\mathbf{r}, \mathbf{x}) = (1 - P)\psi(\mathbf{r}, \mathbf{x}),$$
$$= \sum_c \frac{O_c(r)}{r} u_c(S) + \sum_{c'} \frac{O_{c'}(r')}{r'} u_{c'}(S') \quad (32b)$$

in the interior and exterior regions of space defined by P and $1 - P$, respectively. In Eq. (32b)

$$O_c(r) = ik_\alpha r h_l^{(1)}(k_\alpha r) \underset{r\to\infty}{\sim} \exp i(k_\alpha r - \tfrac{1}{2}l\pi), \quad (33)$$

where $h_l^{(1)}$ is the spherical Hankel function of first class,[10] and we have used the orthogonality property of u_c and $u_{c'}$ in the exterior region.[11] From the continuity of the logarithmic derivative of $y_c(r)$ at $r = r_0$, we get

$$\left[\frac{dy_c/dr}{y_c(r)}\right]_{r=r_0} = \left[\frac{dO_c/dr}{O_c(r)}\right]_{r=r_0} = L_l(k_\alpha r_0), \quad (34)$$

with $L_l(k_\alpha r_0) \sim ik_\alpha$ for $|k_\alpha r_0| \gg 1$. Using this quantity

[8] H. A. Bethe and R. Peierls, Proc. Roy. Soc. (London) A149, 176 (1935).
[9] G. Breit and E. Wigner, Phys. Rev. 49, 519 (1936).
[10] The changes required to deal with the Coulomb interaction in resonance scattering can be found in Ref. 3 and in J. Humblet, Nucl. Phys. 50, 1 (1964).
[11] M. L. Goldberger and K. M. Watson, Collision Theory (John Wiley & Sons, Inc., New York, 1964), Appendix C.

we can write Eq. (29b) for the linewidth in the form

$$\Gamma_n = \left(\sum_c \int_0^{r_0} dr\, |y_{cn}(r)|^2\right)^{-1} \sum_c \left(-\frac{\hbar^2}{2\mu}\right)$$
$$\times \int_0^\infty dr\, y_{cn}(r)^* i\left[\theta(r_0 - r), \frac{d^2}{dr^2}\right] y_{cn}(r),$$

$$= \left(\sum_c \int_0^{r_0} dr\, |y_{cn}(r)|^2\right)^{-1}$$
$$\times \sum_c \left(-\frac{\hbar^2}{2\mu}\right) i(L_l - L_l^*)\, |y_{cn}(r_0)|^2, \quad (35)$$

where $\theta(r_0 - r)$ is the step function and the subscript n indicates $z = z_n$. Equation (35) can be transformed by means of the relations[3]

$$L_l(k_{an}r_0) - L_l(k_{an}r_0)^* = i(k_{an} + k_{an}^*)M_l(k_{an}r_0)$$
$$+ (k_{an}^2 - k_{an}^{*2})N_l(k_{an}r_0), \quad (36a)$$

$$M_l(k_{an}r_0) = |h_0^{(1)}(k_{an}r_0)/h_l^{(1)}(k_{an}r_0)|^2, \quad (36b)$$

$$N_l(k_{an}r_0) = |r_0 k_{an}^2 h_l^{(1)}(k_{an}r_0)|^{-1}$$
$$\times \sum_{p=0}^P [2(l - 2p) - 1]\, |h_{l-2p-1}^{(1)}(k_{an}r_0)|^2, \quad (36c)$$

with $N_0(k_{an}r_0) = 0$ and $P = \frac{1}{2}l - \frac{1}{4}[3 + (-1)^l]$. The result is

$$\Gamma_n = \sum_\alpha \left(\sum_{lm} w_{cn}\right)\Gamma_{an} \Big/ \sum_c w_{cn}, \quad (37)$$

$$\Gamma_{an} = (\hbar^2/\mu)\kappa_{an} \sum_{lm} M_l\, |y_{cn}(r_0)|^2 \Big/ \sum_{lm} w_{cn}, \quad (38)$$

where

$$w_{cn} = \int_0^{r_0} dr\, |y_{cn}(r)|^2 + N_l\, |y_{cn}(r_0)|^2.$$

Here Γ_{an} is the partial linewidth for decay of the nth compound state into two particles in state α. We can compare Eq. (37) with the second Eq. (22) to obtain, provided $\kappa_{an} \neq 0$,

$$\lambda_{an} = \frac{1}{2}\sum_{lm} M_l\, |y_{cn}(r_0)|^2 \Big/ \sum_{lm} w_{cn} \quad (39)$$

so that λ_{an} is nonnegative for $\kappa_{an} \neq 0$. The same will hold for any other state $b \neq a$ since Γ_n will then be nonnegative, as seen from Eq. (22). In this case $|\psi_n(z_n)\rangle$ will describe a system decaying in all channels and it is called a *resonance* state. With respect to W_α,

the resonance may be called *proper* if $\kappa_{an} > \lambda_{an}$ (or $E_n - W_\alpha > 0$) and *virtual* if $\kappa_{an} < \lambda_{an}$ (or $E_n - W_\alpha < 0$). The solutions with λ_{an} negative will exist only for $\kappa_{an} = 0$. In such a case $\Gamma_n = 0$ and $|\psi_n(E_n)\rangle$ will represent a *bound* state with energy $E_n = W_\alpha - \hbar^2\lambda_{an}^2/(2\mu)$.

IV. DISCUSSION

Our treatment has been based on the solution of Eq. (8), that is, the eigenvalue equation for the kernel of the T matrix equation, with the boundary conditions specified by Eq. (9) and by the requirement of outgoing waves for all the channel states. As a result we have arrived naturally at the modified Schrödinger equation (25), previously used in the literature. We could have started our study with Eq. (5) rather than Eq. (6) to obtain similar results in terms of the outgoing-particle variables. The present approach should be compared with those of Meetz[12] and of Weinberg,[13] which are based essentially on a symmetrized form of Eq. (8) and impose as a boundary condition the finiteness of $\langle \psi_n| V |\psi_n\rangle$.[14] The definition of resonance energy and linewidth in both cases are different, and the one given here, which coincides with the Siegert–Humblet definition, leads to a simpler computational problem for determining E_n and Γ_n, namely, the solution of Eq. (20) with the boundary conditions of Eq. (8). We have restricted ourselves to finite range potentials. The effect of an infinite range in V and V' will depend on their dependence on r and r', but Eq. (17) can be expected to hold if these potentials go to zero at infinity faster than any exponential.

ACKNOWLEDGMENTS

[12] K. Meetz. J. Math. Phys. 3, 690 (1962).

[13] S. Weinberg, Phys. Rev. 131, 440 (1963), and the modifications introduced in M. Scadron, S. Weinberg, and J. Wright, *ibid.* 135, B202 (1964).

[14] The work in Refs. 12 and 13 has been extended in R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Company, Inc., New York, 1966), Chap. 9.

A Class of Eigenvalues of the Fine-Structure Constant and Internal Energy Obtained from a Class of Exact Solutions of the Combined Klein–Gordon–Maxwell–Einstein Field Equations*

A. DAS† AND C. V. COFFMAN

Department of Mathematics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received 2 September 1966)

This paper deals with the combined Klein–Gordon–Maxwell–Einstein field equations, which govern completely and self-consistently the spinless, charged, gravitating matter distribution. One of the theorems that have been proved here states that, from a static, purely gravitational universe, a class of electrogravitational universes containing a stationary matter field can be constructed, provided a single differential equation is satisfied. The construction of the electrogravitational universe from the Schwarzchild solution hinges on the solubility of the ordinary differential equation

$$U'' + \alpha^2 (x \operatorname{csch}^2 x)^2 U^3 = 0,$$

where the prime denotes differentiation and α^2 stands for the fine-structure constant. Next, the following nonlinear eigenvalue problem related to this differential equation has been posed. Are there some positive values of α corresponding to which solutions $U(x)$ exist such that (i) U is analytic and positive in $x \in (0, \infty]$ ($\Rightarrow$ the volume element has one sign), (ii) $U(0) = 0$ (this condition is physically unpleasant but forced by the differential equation itself), (iii) $U'(\infty) = 0$ ($\Rightarrow$ no force at the center of spherically symmetric mass and charge distributions), (iv) $U''(0) = \alpha$ ($\Rightarrow$ the total charge of the material distribution is α)? The answer is "yes" and it has been rigorously proved that there exists a unique solution of the problem. The corresponding value of α comes out to be $1.4343(hc)^{\frac{1}{2}}$, which, unfortunately, does not agree with the experiment (the discrepancy may be attributed to the neglect of the second quantization). If the restriction in $U(x)$ to be positive is withdrawn, then a countable number of solutions exist with the corresponding eigenvalues for the electronic charge, internal energy, and mass. These solutions give rise to universes which are topologically inequivalent to Euclidean space and contain a finite number of shells. It should be mentioned that the present eigenvalue problem appears as a consequence of the "Weyl–Majumdar" condition on the electrogravitational universe. There may well exist other eigenvalue problems for the fine-structure constant within the framework of the Klein–Gordon–Maxwell–Einstein field equations without the "Weyl–Majumdar" condition. ". . . es kann dann in jedem Punkte das Krümmungsmass in drei Richtungen einen beliebigen Werth haben, wenn nur die ganze Krümmung jedes messbaren Raumtheils nicht merklich von Null verschieden ist . . ."—Riemann.

I. INTRODUCTION

VARIOUS attempts have been made to arrive at a completely field-theoretic, nonsingular description of matter. Wheeler[1] and his school have offered a purely geometrical description of matter. Finkelstein[2] has presented extended models of particles with internal rotational motions. Some nonsingular models of particles from the solutions of the field equations, derivable from the quadratic action principle in general relativity, have been constructed by Lanczos.[3]

One of the authors (A. D.)[4] introduced the complex scalar field in general relativity to replace the usual and phenomenological description of matter in the right-hand sides of the electromagnetic and gravitational equations. De[5] has followed up the investigation of the combined De Klein–Gordon–Maxwell–Einstein field equations. The present paper starts with the same set of combined field equations to be investigated under more general conditions than in Das[4] and De.[5]

The first of the theorems states that the system of the coupled, nonlinear, partial differential equations, representing the combined fields, is derivable from a variational principle and the second proves the determinateness of the system of the equations. The third theorem deals with the static electrogravitational fields generated by stationary matter for which no specific spatial symmetry is assumed. It is shown that the Weyl–Majumdar[6] relationship

$$g_{44} = [1 \pm (4\pi)^{\frac{1}{2}} A_4]^2$$

[1] C. W. Misner and J. A. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1957); J. A. Wheeler, *ibid.* **2**, 604 (1957).

[2] D. Finkelstein, Phys. Rev. **100**, 924 (1955); also see L. de Broglie, D. Bohm, P. Hillion, F. Halbwachs, T. Takabayasi, and J. P. Vigier, Phys. Rev. **129**, 438 (1963).

[3] C. Lanczos, Rev. Mod. Phys. **29**, 334 (1957).

[4] A. Das, J. Math. Phys. **4**, 45 (1963). For the description of matter by the Dirac field in general relativity, see A. Das, Proc. Roy. Soc. (London) **A267**, 1 (1962).

[5] N. De, Progr. Theoret. Phys. (Kyoto) **33**, 545 (1965); Nuovo Cimento **39**, 986 (1965).

[6] H. Weyl, Ann. Phys. (Paris) **54**, 117 (1917); S. D. Majumdar, Phys. Rev. **72**, 390 (1947). Also see H. E. J. Curzon, Proc. London Math. Soc. **23**, 477 (1925); A. Papapetrou, Proc. Roy. Irish Acad. **A51**, 191 (1947).

between the metric tensor g_{44} and electrostatic potential A_4 implies $\alpha^2 = m^2$, where α, m are the charge and mass parameters, respectively. This condition physically means that inside the matter there is a balance of forces between the electrostatic repulsion and the gravitational attraction. The fourth theorem proves that the vanishing of the curvature invariant in the 3-space, conformal to the Weyl–Majumdar electrogravitational 3-space, implies that (i) the original 3-space is flat, (ii) the wavefunction is a constant, and (iii) the combined field equations boil down to a single differential equation. The fifth theorem, which is the most pertinent for the subsequent sections, proves that from a given static, purely gravitational universe one can construct the Weyl–Majumdar electrogravitational universes containing matter, if a single differential equation is satisfied.[5] Three examples are presented, of which the first one starts from the flat universe (trivially gravitational) and the relevant differential equation is $\nabla^2 v = -\lambda^2 v^3$, where ∇^2 is the Euclidean Laplace operator and λ^2 is a positive constant. The solutions of this equation have infinite oscillations; the corresponding electrogravitational universes consist of an infinite number of shells and hence seem to be physically not feasible. In the third example, the purely gravitational universe is due to the finite number of concentric and coplanar rings, and the corresponding differential equation (4.17) which has to be satisfied is quite complicated. The investigation of this equation is postponed for a future occasion.

The second example deals with the well-known spherically symmetric universe of Schwarzchild, and the related differential equation is[4]

$$U'' + \alpha^2 (x \operatorname{csch}^2 x)^2 U^3 = 0.$$

Rigorous and exhaustive investigations of this differential equation have been carried out. The fundamental theorem, which solves a nonlinear eigenvalue problem, proves that there exists a unique positive constant α and a unique solution $U(x)$ of the last differential equation in $x \in [0, \infty)$, such that (i) $U > 0$ for $0 < x < \infty$ ($\Rightarrow$ the volume element is one, signed), (ii) $U(0) = 0$,

(iii) $$\lim_{x \to \infty} U'(x) = 0$$

($\Rightarrow$ no force at the center of spherically symmetric matter distribution), (iv) $U'(0) = \alpha$ ($\Rightarrow$ the total charge is α). The proof of this theorem is preceded by 13 lemmas, of which the fifth one solves the initial value problem $U'(0) = a$, $U(0) = 0$ of the differential equation with help of the contractive mapping theorem. The seventh theorem states that there exist

solutions which have finite numbers of zeros and asymptotically approach straight lines.

The theoretical bounds for the eigenvalue have been ascertained and the actual number $\alpha = 1.4343$ is obtained by solving the boundary value problem with the help of the computer. The disagreement of this number with the experimental value $\alpha^2 = \frac{1}{137}$ may be due to neglecting the second quantization and other relevant fields. It may also be possible that the correct value will come out of a different eigenvalue problem when the Weyl–Majumdar requirement is discarded.

The next section is devoted to the discussion of the physical, geometrical, and topological properties of the universes permitted by the differential equation. The universe corresponding to the nonnegative solution $U(x)$ is topologically Euclidean. The universes corresponding to the solutions with a finite number of zeros are topologically inequivalent to the Euclidean space and are comprised of onionlike shells.

It should be mentioned that in recent years many investigators[7] have not considered the usual topology for the physical space.

The main theme of this paper is posing the nonlinear eigenvalue problem for the fine-structure constant (and observable internal energy) and showing that a countable number of eigenvalues exist. It is an interesting outcome that the logarithmically divergent electrostatic self-energy does not affect the internal energy levels. Physical properties of these particles are found to be similar to geons. There is no reason why these particles will not be created in very high-energy interactions (such as in a nova or the quasi-stellar sources), and therefore their existence is predicted.

II. DEFINITIONS AND NOTATIONS

(1) V_4 denotes a four-dimensional Riemannian manifold and physically represents the space–time universe of the events. A point $x \in V_4$ has the real coordinates x^i (where i and the other Roman indices take 1, 2, 3, 4). V_3 denotes a x^4-constant submanifold in V_4 and represents a spatial universe. A point $\mathbf{x} \in V_3$ has the real coordinates x^α (α and other Greek indices take 1, 2, 3).

(2) V_4 has index of inertia -2, i.e., the metric form

$$\Phi \stackrel{\text{def}}{=} g_{ab}(x)\, dx^a\, dx^b$$

[7] J. L. Synge, Dublin Lectures (1947); J. P. Vigier, Phys. Rev. Letters 17, 39 (1966); D. Finkelstein and C. W. Misner, Ann. Phys. 6, 230 (1959); E. C. Zeeman, "The Topology of Minkowski Space," Cambridge University Preprint (1965); D. Finkelstein, J. Math. Phys. 7, 1218 (1966); D. Atkinson and M. B. Halpern, ibid. 8, 373 (1967).

is reducible at any point to $\Phi = -(dX^1)^2 - (dX^2)^2 - (dX^3)^2 + (dX^4)^2$. Here and subsequently the summation convention is followed.

(3) The Einstein tensor which represents the energy-momentum-stress density is defined by

$$G_{ij} \stackrel{\text{def}}{=} R_{ij} - \tfrac{1}{2}g_{ij}R,$$

where R_{ij} and R stand for the Ricci tensor and the curvature scalar.

(4) The vector field $A^i(x)$ defined on V_4 represents the electromagnetic potential and the corresponding intensity field is defined by

$$F_{ij} \stackrel{\text{def}}{=} \nabla_j A_i - \nabla_i A_j = A_{i,j} - A_{j,i}.$$

Here ∇_i and $\nabla_{,i}$ denote, respectively, the covariant and the partial differentiation with respect to x^i.

(5) The complex scalar field $\psi(x)$ is defined on V_4 and $\psi^*\psi$ represents, in a way, the matter density (asterisk stands for the complex conjugation).

(6) The combined Klein–Gordon–Maxwell–Einstein field equations in V_4 are defined as

$$K \stackrel{\text{def}}{=} [D^i D_i + m^2]\psi(x) = 0,$$

$$M^i \stackrel{\text{def}}{=} \nabla_j F^{ij} + (4\pi)^{\frac{1}{2}} i\alpha(D^{*i}\psi^* \cdot \psi - \psi^* D^i \psi) = 0,$$
$$(2.1a)$$

$$(F) \quad E_{ij} \stackrel{\text{def}}{=} G_{ij} + 8\pi[D_i^*\psi^* \cdot D_j\psi + D_j^*\psi^* \cdot D_i\psi$$
$$- g_{ij}(D^{*a}\psi^* \cdot D_a\psi - m^2\psi^*\psi)$$
$$- F_{ik}F^k_{\cdot j} + \tfrac{1}{4}g_{ij}F_{ab}F^{ab}] = 0,$$

where

$$D_i \stackrel{\text{def}}{=} \nabla_i + (4\pi)^{\frac{1}{2}} i\alpha A_i,$$

α and m are the charge and the mass parameters. The units are so chosen that $\hbar = c = G = 1$, and all the physical quantities are expressed as pure numbers.

There are also the additional constraints on g_{ab} and A_i, the so-called coordinate conditions and the Lorentz gauge condition:

$$(F') \quad \begin{array}{l} C_i(g_{ab}) = 0, \\ \mathfrak{L} \stackrel{\text{def}}{=} \nabla_i A^i = 0. \end{array} \quad (2.1b)$$

Physically, these equations represent completely and self-consistently the spinless, charged, gravitating, unquantized matter field under its own electromagnetic and gravitational interactions.

III. THEOREMS ON THE COMBINED KLEIN–GORDON–MAXWELL–EINSTEIN EQUATIONS

Theorem I: Let D be a bounded simply-connected domain of V_4 which is covered by a single coordinate system. Let $\partial(D)$ be the orientable, piecewise-smooth

boundary of D. Let the fields $\psi(x)$, $A_i(x)$, $g^{ij}(x)$ be C^2 with $\det g_{ij} < 0$ in D and on $\partial(D)$. Let an invariant action integral on D be defined[8] as

$$A \stackrel{\text{def}}{=} \int_D [R + 16\pi(D^{*a}\psi^* \cdot D_a\psi - m^2\psi^*\psi)$$
$$- 4\pi F^{ab}F_{ab}]\, d_4v,$$

where d_4v is the invariant volume element in V_4. Then the Euler–Lagrange equations which are obtainable from the variational equation $\delta A = 0$, with $\delta\psi(\partial(D)) = \delta A^i(\partial(D)) = \delta g^{ab}(\partial(D)) = 0$, are the field equations (F).

Theorem II: Let D be a domain of V_4 which is covered by a single coordinate system, where the functions $\psi(x)$ is C^2, $A_i(x)$ is C^3, g_{ij} is C^n $(n \geq 3)$. The system of coupled, nonlinear, partial differential equations (F) with the constraint (F') is a determinate system in D.

Proof. The number of 16 unknown functions can be exhibited as

$$1(\text{Re } \psi) + 1(\text{Im } \psi) + 4(A_i) + 10(g_{ij}).$$

The number of equations is 21, namely,

$$1(K) + 1(K^*) + 4(M^i) + 10(E_{ij}) + 4(C_i) + 1(\mathfrak{L}).$$

However, there are five differential identities[9]:

$$1(\nabla_i M^i = 0) + 4(\nabla_j E^{ij} = 0).$$

So there are 16 independent equations. Therefore the system is determinate.

IV. THEOREMS ON STATIC ELECTRO-GRAVITATIONAL FIELD GENERATED BY STATIONARY MATTER FIELD

Definition 7: The static electric field is characterized by the following:

$$A_\alpha(\mathbf{x}) = 0, \quad \varphi \stackrel{\text{def}}{=} A_4(\mathbf{x}), \quad F_{\alpha\beta} = 0, \quad F_{4\alpha} = \varphi_{,\alpha}.$$

Definition 8: The static gravitational field is characterized by[10]

$$\Phi = g_{\alpha\beta}(\mathbf{x})\, dx^\alpha\, dx^\beta + f(\mathbf{x})(dx^4)^2.$$

---

[8] The integrand is invariant under general coordinate transformation and also under the gauge transformation

$$\psi' = \psi e^{i(4\pi)^{\frac{1}{2}}\alpha\varphi(x)}, \quad A'_a = A_a - \varphi_{,a}.$$

[9] A. Das, Ref. 4. These identities represent, respectively, the differential conservations of charge–current vector and energy-momentum-stress tensor.

[10] Such a universe admits one parameter group of motion along x^4 lines. Also, V_3 is a totally geodesic hypersurface of V_4. Cf. L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1949), p. 183.

Definition 9: The stationary matter field is characterized by

$$\psi = \chi(\mathbf{x})e^{iEx^4}, \quad \chi = \chi^*,$$

where E is a positive number representing the energy of the matter field.

Lemma I: Let $\mathbf{D}$ be a domain of V_3, a x^4-constant hypersurface of the static V_4. Let $\chi(\mathbf{x})$, $\varphi(\mathbf{x})$, $g_{\alpha\beta}(\mathbf{x})$, $f(\mathbf{x})$ be C^2 with $g \stackrel{\text{def}}{=} \det g_{\alpha\beta} < 0, f(\mathbf{x}) > 0$ in $\mathbf{D}$. Then the field equations $M^\alpha = E_{\alpha 4} = 0$ identically (physically meaning, there is no current or momentum flow). Also, $(F) \Rightarrow (\mathbf{F})$ in $\mathbf{D}$, where $(\mathbf{F})$ is the following:

$$Ke^{-iEx^4} = (-g)^{-\frac{1}{2}}[(-g)^{\frac{1}{2}}g^{\alpha\beta}\chi_{,\alpha}]_{,\beta}$$
$$+ [m^2 - f^{-1}(E + (4\pi)^{\frac{1}{2}}\alpha\varphi)^2]\chi = 0,$$

$$M^4 = (-g)^{-\frac{1}{2}}[(-g)^{\frac{1}{2}}g^{\alpha\beta}\varphi_{,\alpha}]_{,\beta} - f^{-1}g^{\alpha\beta}f_{,\alpha}\varphi_{,\beta}$$
$$+ 2(4\pi)^{\frac{1}{2}}\alpha[E + (4\pi)^{\frac{1}{2}}\alpha\varphi]\chi^2 = 0,$$

$$(\mathbf{F}) \quad \tilde{E}_{44} = \frac{1}{2}[(-g)^{\frac{1}{2}}\{(-g)^{\frac{1}{2}}g^{\alpha\beta}f_{,\alpha}\}_{,\beta} - f^{-1}g^{\alpha\beta}f_{,\alpha}f_{,\beta}]$$
$$+ 8\pi[2\{E + (4\pi)^{\frac{1}{2}}\alpha\varphi\}^2\chi^2$$
$$- m^2f\chi^2 - \frac{1}{2}g^{\alpha\beta}\varphi_{,\alpha}\varphi_{,\beta}] = 0, \quad (4.1)$$

$$\tilde{E}_{\alpha\beta} = R_{\alpha\beta} + 8\pi[2\chi_{,\alpha}\chi_{,\beta} - m^2g_{\alpha\beta}\chi^2 - f^{-1}\varphi_{,\alpha}\varphi_{,\beta}$$
$$+ \frac{1}{2}f^{-1}g_{\alpha\beta}g^{\gamma\delta}\varphi_{,\gamma}\varphi_{,\delta}] = 0,$$

where

$$\tilde{E}_{ij} \stackrel{\text{def}}{=} E_{ij} - \frac{1}{2}g_{ij}E^k_{.k}.$$

Definition 10: The Weyl–Majumdar[6] static electrogravitational universe is characterized by

$$f(\mathbf{x}) = F[\varphi(\mathbf{x})] = [1 \pm (4\pi)^{\frac{1}{2}}\varphi]^2.$$

Theorem III: Let $\mathbf{D}$ be a domain of V_3 which is a x^4-constant hypersurface of a Weyl–Majumdar universe. If the assumptions and the consequent field equations $(\mathbf{F})$ of the Lemma I hold in $\mathbf{D}$, then $\alpha^2 = m^2 = E^2$. Moreover, $(\mathbf{F}) \Rightarrow (\mathscr{F})$ in $\mathbf{D}$, where $(\mathscr{F})$ is the following:

$$Ke^{-iEx^4} = (\bar{g}^{-\frac{1}{2}}\bar{g}^{\alpha\beta}\chi_{,\alpha})_{,\beta} = 0,$$

$$(\mathscr{F})V^5\tilde{E}_{44} = \bar{g}^{-\frac{1}{2}}(\bar{g}^{\frac{1}{2}}\bar{g}^{\alpha\beta}V_{,\alpha})_{,\beta} + 8\pi^2\alpha^2\chi^2V^3 = 0, \quad (4.2)$$

$$\tilde{E}_{\alpha\beta} - \bar{g}_{\alpha\beta}\tilde{E}_{44} = \bar{R}_{\alpha\beta} + 16\pi\chi_{,\alpha}\chi_{,\beta} = 0,$$

where $V = f^{-\frac{1}{2}}$, $\bar{g}_{\alpha\beta} = -V^{-2}g_{\alpha\beta}$, $\bar{R}_{\alpha\beta}$ is the Ricci tensor constructed out of $\bar{g}_{\alpha\beta}$.

Proof: The Weyl–Majumdar condition $f(\mathbf{x}) = [1 \pm (4\pi)^{\frac{1}{2}}\varphi(\mathbf{x})]^2$ yields

$$f_{,\alpha} = \pm 2(4\pi f)^{\frac{1}{2}}\varphi_{,\alpha}. \quad (4.3)$$

Substituting (4.3) into (4.1), the following is obtained:

$$-\frac{1}{2}(4\pi)^{-\frac{1}{2}}f^{\frac{1}{2}}\chi^{-2}M^4 \pm (8\pi)^{-1}\chi^{-2}\tilde{E}_{44}$$
$$= \pm 4\pi(\alpha^2 - m^2)[\varphi(\mathbf{x})]^2 - (4\pi)^{\frac{1}{2}}(2m^2 + \alpha^2 \mp 3E\alpha)$$
$$\times \varphi(\mathbf{x}) \pm (2E^2 - m^2 \mp E\alpha) = 0. \quad (4.4)$$

Since (4.4) is satisfied for $\forall \mathbf{x} \in \mathbf{D}$, the coefficients of φ^2, φ must vanish separately, i.e.,

$$\left. \begin{array}{r} \alpha^2 - m^2 = 0, \\ 2m^2 + \alpha^2 \mp 3E\alpha = 0, \\ 2E^2 - m^2 \mp E\alpha = 0. \end{array} \right\} \Rightarrow \alpha^2 = m^2 = E^2. \quad (4.5)$$

From the physical conditions the choice $m = E > 0$ should be made. The condition $\alpha = \pm m$ physically implies an equilibrium inside matter due to the mutual actions of the electrostatic repulsion and the gravitational attraction. The second part of the theorem follows if (4.3) and (4.5) are substituted in (4.1).

Theorem IV: Let V_3 be a x^4-constant hypersurface of a Weyl–Majumdar universe V_4 with

$$\Phi = -V^2\bar{g}_{\alpha\beta}dx^\alpha dx^\beta + V^{-2}(dx^4)^2.$$

Let the conditions of the Lemma I and the field equations $(\mathbf{F})$ hold in V_3. If we define a manifold $\bar{V}_3$, which is conformal to V_3, by the metric form $\bar{\Phi} = \bar{g}_{\alpha\beta}dx^\alpha dx^\beta$, then the vanishing of the curvature invariant $\bar{R} \equiv \bar{R}^\alpha_{.\alpha}$ with $\chi \neq 0$ implies that (i) χ is a constant, (ii) $\bar{V}_3$ is flat, and (iii) $(\mathscr{F}) \Rightarrow \nabla^2V = -\lambda^2V^3$, where the constant $\lambda = (4\pi)^{\frac{1}{2}}\alpha\chi$ and ∇^2 is the Euclidean Laplace operator.

Proof: From the previous theorem it follows that $\alpha^2 = m^2 = E^2$ and $(\mathscr{F})$ holds in V_3. Then

$$\tilde{E}^\alpha_{.\alpha} - 3\tilde{E}_{44} = \bar{R} + 16\pi\bar{g}^{\alpha\beta}\chi_{,\alpha}\chi_{,\beta} = 0. \quad (4.6)$$

The vanishing of $\bar{R}$ yields

$$\bar{g}^{\alpha\beta}\chi_{,\alpha}\chi_{,\beta} = 0. \quad (4.7)$$

From the assumption V_4 has index of inertia -2, it follows that V_3 is a negative-definite and $\bar{V}_3$ is a positive-definite Riemannian manifold.

Therefore $\exists$ a coordinate system in $\mathbf{D}$ of $\bar{V}_3 \ni$

$$\bar{\Phi} = e^{P_1(\mathbf{x})}(dx^1)^2 + e^{P_2(\mathbf{x})}(dx^2)^2 + e^{P_3(\mathbf{x})}(dx^3)^2.$$

In this coordinate system (4.7) becomes

$$e^{-P_1}(\chi_{,1})^2 + e^{-P_2}(\chi_{,2})^2 + e^{-P_3}(\chi_{,3})^2 = 0. \quad (4.8)$$

The positive definiteness of (4.8) implies $\chi_{,1} = \chi_{,2} = \chi_{,3} = 0 \Rightarrow \chi$ is a constant. But χ is a scalar field. Therefore χ is a constant in any coordinate system. So, in general,

$$E_{\alpha\beta} - \bar{g}_{\alpha\beta}\tilde{E}_{44} = \bar{R}_{\alpha\beta} = 0. \quad (4.9)$$

In three-dimensional Riemannian space like $\bar{V}_3$, the necessary and sufficient condition for flatness is the equation (4.9). Therefore $\bar{V}_3$ is Euclidean, and the only surviving equation of $(\mathcal{F})$ reduces to

$$\tilde{E}_{44} = \nabla^2 V + \lambda^2 V^3 = 0, \qquad (4.10)$$

where the constant $\lambda^2 = 4\pi\alpha^2\chi^2$, and ∇^2 is the Euclidean Laplace operator.

Definition 11: Purely gravitational field equations are defined as (F_0): $R_{ij} = 0$. The universe $_0V_4$, where (F_0) is satisfied except at finite number of singularities, is called a purely gravitational universe.

Lemma 2: Let the metric form

$$\Phi_0 = -e^{-\omega(\mathbf{x})}g_{\alpha\beta}(\mathbf{x})\,dx^\alpha\,dx^\beta + e^{\omega(\mathbf{x})}(dx^4)^2$$

represent a static gravitational universe. Then the field equations $(F_0) \Rightarrow (\mathbf{F}_0)$, where $(\mathbf{F}_0)$ is the following:

$$\begin{aligned}
_0\tilde{E}_{\alpha\beta} &= \mathcal{R}_{\alpha\beta} + \tfrac{1}{2}\omega_{,\alpha}\omega_{,\beta} = 0, \\
-2_0\tilde{E}_{44} &= g^{-\frac{1}{2}}(g^{\frac{1}{2}}g^{\alpha\beta}\omega_{,\alpha})_{,\beta} = 0, \qquad (4.11) \\
_0\tilde{E}_{4\alpha} &= 0,
\end{aligned}$$

where $\mathcal{R}_{\alpha\beta}$ is the Ricci tensor constructed out of $g_{\alpha\beta}$ and $g = \det g_{\alpha\beta}$.

Theorem V: Given a domain $\mathbf{D}_0$ of a x^4-constant hypersurface $_0V_3$ of static, purely gravitational universe $_0V_4 \ni$ the metric fields $\omega(\mathbf{x})$, $g_{\alpha\beta}(\mathbf{x})$ are C^2 with $g > 0$ in $\mathbf{D}_0$, then a domain $\mathbf{D}$ of a hypersurface V_3 of a Weyl–Majumdar universe V_4 containing the stationary matter field can be constructed with $\chi(\mathbf{x}) = \tfrac{1}{2}(8\pi)^{-\frac{1}{2}}\omega(\mathbf{x})$, $\bar{g}_{\alpha\beta} = g_{\alpha\beta}$, provided $\exists$ a $V(\mathbf{x}) \neq 0$ which is C^2 and satisfies

$$\bar{g}^{-\frac{1}{2}}(\bar{g}^{\frac{1}{2}}\bar{g}^{\alpha\beta}V_{,\alpha})_{,\beta} + 8\pi\alpha^2\chi^2 V^3 = 0 \qquad (4.12)$$

in $\mathbf{D}_0$.

Comparing (4.2) and (4.11), the proof of the theorem follows immediately.

Corollary I: Given a domain $\mathbf{D}_0$ of $_0V_3$ of a flat $_0V_4$ [where $(\mathbf{F}_0)$ is satisfied trivially] with the metric form

$$\Phi_0 = -e^{-k}(dx^\alpha\,dx^\alpha) + e^k(dx^4)^2$$

(k being a constant), a domain $\mathbf{D}$ of V_3 of a Weyl–Majumdar V_4 can be constructed with

$$\chi = \tfrac{1}{4}(2\pi)^{-\frac{1}{2}}k, \quad \Phi = -V^2(dx^\alpha\,dx^\alpha) + V^{-2}(dx^4)^2,$$

provided $\exists$ a $V(\mathbf{x}) \neq 0$ in $\mathbf{D}$ that satisfies

$$\nabla^2 V + \lambda^2 V^3 = 0, \qquad (4.13)$$

where the constant $\lambda^2 = \tfrac{1}{4}\alpha^2 k^2$. [Note that (4.13) is same as (4.10).]

Corollary II: Given a domain $\mathbf{D}_0$ of a spherically symmetric $_0V_3$ in $_0V_4$ with the Schwarzchild's metric form[11]

$$\begin{aligned}
\Phi_0 = &-(1 + k/x^1)^4[(dx^1)^2 + (x^1\,dx^2)^2 \\
&+ (x^1\sin x^2\,dx^3)^2] + [(x^1 - k)/(x^1 + k)]^2(dx^4)^2,
\end{aligned}$$

k being a constant, a domain $\mathbf{D}$ of V_3 of a Weyl–Majumdar V_4 can be constructed with

$$\chi(x^1) = (8\pi)^{-\frac{1}{2}}\ln[(x^1 - k)/(x^1 + k)],$$

$$\begin{aligned}
\Phi = &-V^2[1 - (k/x^1)^2]^2[(dx^1)^2 + (x^1\,dx^2)^2 \\
&+ (x^1\sin x^2\,dx^3)^2] + V^{-2}(dx^4)^2,
\end{aligned}$$

provided $\exists$ a $V(\mathbf{x}) \neq 0$ which is C^2 in $\mathbf{D}$ and satisfies

$$\begin{aligned}
&\frac{d^2V}{(dx^1)^2} + \frac{2}{x^1}\left[1 + \frac{k^2}{\{(x^1)^2 - k^2\}}\right]\frac{dV}{dx^1} \\
&+ \alpha^2\left[\left\{1 - \left(\frac{k}{x^1}\right)^2\right\}\ln\frac{\{(x^1 - k)\}}{\{(x^1 + k)\}}\right]^2 V^3 = 0. \quad (4.14)
\end{aligned}$$

With a coordinate transformation

$$x = -\ln[(x^1 - k)/(x^1 + k)], \quad U(x) = 2kV(x^1),$$

the last equation and the metric form go over to[4]

$$\begin{aligned}
\Phi = &-U^2[(\operatorname{csch}^2 x\,dx)^2 + (\operatorname{csch} x\,dx^2)^2 \\
&+ (\operatorname{csch} x\sin x^2\,dx^3)^2] + U^{-2}(2k\,dx^4)^2,
\end{aligned}$$

$$U'' + \alpha^2(x\operatorname{csch}^2 x)^2 U^3 = 0, \qquad (4.15)$$

where the prime denotes differentiation with respect to x. Changing the names of the coordinates to $\theta = x^2$, $\varphi = x^3$, $t = 2kx^4$, the metric assumes the following form:

$$\begin{aligned}
\Phi = &-U^2[\operatorname{csch}^4 x(dx)^2 + \operatorname{csch}^2 x(d\theta^2 + \sin^2\theta\,d\varphi^2)] \\
&+ U^{-2}\,dt^2. \quad (4.16)
\end{aligned}$$

Corollary III: Let a domain $\mathbf{D}_0$ of an axially symmetric, static, purely gravitational universe due to N concentric rings at $x^2 = 0$ "plane" with "radii" $0 < \rho_1 < \rho_2 < \cdots < \rho_N$ and masses M_1, $M_2\cdots$, M_N, be given by the metric form

$$\begin{aligned}
\Phi_0 = &-e^{-2\lambda(x^1,x^2)}[e^{2\nu(x^1,x^2)}\{(dx^1)^2 + (dx^2)^2\} \\
&+ (x^1\,dx^3)^2] + e^{2\lambda}(dx^4)^2,
\end{aligned}$$

where

$$\lambda(x^1, x^2) = -2\sum_{n=1}^{N}\left(\frac{M_n}{p_n}\right)K\left[1 - \left(\frac{q_n}{p_n}\right)^2\right],$$

$$p_n = (x^1 + \rho_n)^2 + (x^2)^2, \quad q_n = (x^1 - \rho_n)^2 + (x^2)^2,$$

---

[11] Isotropic coordinates have been used. The case of the cylindrically-symmetric static gravitational field due to a finite rod can be transformed into the Schwarzchild's form [cf. N. Rosen, Rev. Mod. Phys. **21**, 503 (1948)]. Thus Corollary II covers the case of finite rod too.

K is the elliptic integral of the first kind with the modulus $1 - (q_n/p_n)^2$.

$v(x^1, x^2)$ is given by the line integral

$$v(x^1, x^2) = \int_C [x^1\{(\lambda_{,1})^2 - (\lambda_{,2})^2\} \, dx^1 + 2\lambda_{,1}\lambda_{,2} \, dx^2],$$

where the path C is a piecewise-smooth curve and it does not pass through any of the ringlike singularities of $\lambda(x^1, x^2)$.

Then a domain $\mathbf{D}$ of V_3 of a Weyl–Majumdar V_4 can be constructed with

$$\chi(x^1, x^2) = (8\pi)^{-\frac{1}{2}} \sum_{n=1}^{N} \left(\frac{M_n}{p_n}\right) K\left[1 - \left(\frac{q_n}{p_n}\right)^2\right],$$

$$\Phi = -V^2[e^{2v}\{(dx^1)^2 + (dx^2)^2\} + (x^1 dx^3)^2] + V^{-2}(dx^4)^2,$$

provided ∃ a $V(x^1, x^2) \neq 0$ which is C^2 in $\mathbf{D}$ and satisfies

$$V_{,11} + (x^1)^{-1}V_{,1} + V_{,22}$$

$$= -\alpha^2 e^{2v}\left[\sum_{n=1}^{N}\left(\frac{M_n}{p_n}\right) K\left[1 - \left(\frac{q_n}{p_n}\right)^2\right]\right]^2 V^3. \quad (4.17)$$

Remarks: In the one-dimensional case the general solution of (4.13) is in terms of the elliptic function $V(x^1) = -a \, dn[2a(\frac{1}{2}\lambda x^1 + b), \frac{1}{2}]$. In the spherically symmetric case this equation boils down to one of the Emden equations. In general, a solution of (4.13) will have infinite numbers of zeros implying infinite numbers of singularities in the metric form. This feature makes this class of solutions physically unacceptable. The differential equation (4.17) is complicated, and its study is postponed for a future occasion. The subsequent section is solely devoted to the exhaustive investigation of the differential equation (4.15).

V. EXISTENCE AND UNIQUENESS THEOREMS FOR THE DIFFERENTIAL EQUATION
$$U'' + \alpha^2(x \operatorname{csch}^2 x)^2 U^3 = 0$$

This section is devoted to the proof of the following fundamental result concerning the differential equation which appears in the title above.

Theorem VI: There exists a uniquely determined positive constant α and a uniquely determined function $U = U(x)$ in $C^2[0, \infty)$ such that

(i) $U(x)$ is positive for $0 < x < \infty$,

(ii) U satisfies for $0 < x < \infty$ the differential equation

$$U'' + \alpha^2(x \operatorname{csch}^2 x)^2 U^3 = 0, \quad (5.1)$$

(iii) U satisfies the boundary conditions

$$U(0) = 0, \quad U'(0) = \alpha, \quad (5.2)$$

and

$$\lim_{x \to \infty} U'(x) = 0. \quad (5.3)$$

Conditions (5.1) and (5.3) imply that $U(x)$ has a finite limit as $x \to \infty$:

$$\lim_{x \to \infty} U(x) < \infty. \quad (5.4)$$

If $U(x)$ is any solution of the equation (5.1), then $y(x) = \alpha U(x)$ is a solution of

$$y'' + (x \operatorname{csch}^2 x)^2 y^3 = 0. \quad (5.5)$$

Thus, for the sake of simplicity, in most of the analysis to follow we direct our attention to Eq. (5.5). For convenience we put $p(x) = (x \operatorname{csch}^2 x)^2$.

Lemma 3: Let $y(x)$ be a solution of (5.5) defined on an interval I in $(0, \infty)$. Then the function

$$\Phi(x) = (y'(x))^2 + \tfrac{1}{2}p(x)(y(x))^4$$

is a decreasing function of x on I.

Proof: If we differentiate $\Phi(x)$ and use the fact that y is a solution of (5.5), we obtain

$$\Phi'(x) = \tfrac{1}{2}p'(x)(y(x))^4$$

$$= x \operatorname{csch}^4 x \, (1 - 2x \coth x)(y(x))^4.$$

Since $2x \coth x > 1$ on $(0, \infty)$, it follows that $p'(x)$, and hence also $\Phi'(x)$, is negative on $(0, \infty)$.

Lemma 4: Let $x_0 > 0$ and let a and b be any two real numbers. Then on the interval $[x_0, \infty)$ there exists a unique solution $y(x)$ of the initial value problem

$$y'(x_0) = a, \quad y(x_0) = b \quad (5.6)$$

for (5.5).

Proof: Local existence and local uniqueness follow from the fact that the term $p(x)y^3$ satisfies a local Lipschitz condition in y. If $y(x)$ is a solution of the initial value problem on an interval $[x_0, x_1)$, $0 < x_0 < x_1 < \infty$, then, by Lemma 3, both y and y' remain bounded on $[x_0, x_1)$; therefore y and y' can be extended continuously to $[x_0, x_1]$. It follows that $y(x)$ can be uniquely extended to the right indefinitely as a solution of (5.5).

We also need the following result.

Lemma 5: For any real number a there exists one and only one function $y(x)$ which is of class C^2 in $[0, \infty)$ and is such that

$$y'(0) = a, \quad y(0) = 0, \quad (5.7)$$

and which furthermore satisfies (5.5) in $(0, \infty)$.

Proof: It follows from Lemma 3 that $y(x) \equiv 0$ in $[0, \infty)$ is the unique C^2 solution of (5.5) and (5.7) when $a = 0$. Because of the fact that $-y(x)$ is a solution of (5.5) whenever $y(x)$ is, it suffices to prove the lemma for $a > 0$. We suppose from now on that a is a fixed positive number. Let $x_0 > 0$ be so chosen that

$$24a^2 \int_0^{x_0} x^3 p(x)\, dx < 1. \quad (5.8)$$

[Notice that $x^3 p(x) = x^5 \operatorname{csch}^4 x$ tends to zero as x tends to zero.]

Let $C = C[0, x_0]$ denote the Banach space of continuous functions $u(x)$ on $[0, x_0]$ with the usual norm

$$\|u\| = \max_{0 \le x \le x_0} |u(x)|$$

for $u \in C$. Let $B = \{u \in C : \|u\| < 2a\}$, and define an operator T on B by

$$[Tu](x) = a - \int_0^x \left(1 - \frac{s}{x}\right) s^3 p(s) u^3(s)\, ds. \quad (5.9)$$

For $u, v \in B$, from (5.9) we obtain

$$\|Tu\| \le a + 8a^3 \int_0^{x_0} s^3 p(s)\, ds \quad (5.10)$$

and

$$\|Tu - Tv\| \le \int_0^{x_0} s^3 p(s)\, |u(s) - v(s)|$$
$$\cdot |u^2(s) + u(s)v(s) + v^2(s)|\, ds$$
$$\le 12a^2 \|u - v\| \int_0^{x_0} s^3 p(s)\, ds.$$

It follows from this inequality and (5.8) and (5.10) that T is a contractive mapping of B into B. Therefore, by the contractive mapping theorem of Banach,[12] there is a unique function $w \in B$ satisfying, for $0 \le x \le x_0$,

$$w(x) = a - \int_0^x \left(1 - \frac{s}{x}\right) s^3 p(s) w^3(s)\, ds. \quad (5.11)$$

The function $y(x) = xw(x)$ is a solution of (5.5) and (5.7) on $[0, x_0]$. By Lemma 4 it can be extended to $[0, \infty)$. If, on the other hand, y is any solution of (5.5) and (5.7) on $[0, \infty)$, then for x_0 sufficiently small the restriction to $[0, x_0]$ of $w(x) = x^{-1}y(x)$ belongs to B and satisfies (5.11). Local uniqueness of a solution of (5.5) and (5.7) thus follows from the uniqueness in B of the fixed point of T. Uniqueness in the large follows from Lemma 4.

We denote by $y(x, a)$ the solution of (5.5) which satisfies the condition (5.7). For every real a, $y(x, a)$

[12] A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis* (Graylock Press, Rochester, New York, 1957), p. 43.

is defined for $0 \le x < \infty$; it is clear that $y(x, -a) = -y(x, a)$.

Lemma 6: If a is any real number, then, for $0 \le x < \infty$,

$$|y(x, a)| \le |a|\, x. \quad (5.12)$$

Proof: When a solution $y(x)$ of (5.5) satisfies the conditions (5.7), the associated function $\Phi(x)$ defined in Lemma 3 is continuous at $x = 0$. It follows from Lemma 3 that for $y = y(x, a)$

$$(y'(x))^2 \le \Phi(x) \le \Phi(0) = a^2.$$

Therefore $|y'(x, a)| \le |a|$ for all x on $(0, \infty)$. Now (5.12) follows, using (5.7), from an integration of this last inequality.

Lemma 7: Let $y = y(x) = y(x, a)$ for some nonzero real number a. Then $y(x)$ has, at most, finitely many zeros on $[0, \infty)$ and $y(x)$ and $y'(x)$ satisfy the asymptotic formulas

$$y(x) = \beta x + \gamma + o(1) \quad (5.13)$$

and

$$y'(x) = \beta + o(1), \quad (5.14)$$

as $x \to \infty$, where

$$\beta = a - \int_0^\infty p(x) y^3(x)\, dx, \quad \gamma = \int_0^\infty x p(x) y^3(x)\, dx. \quad (5.15)$$

Proof: Lemma 6 shows that the improper integrals in (5.15) are convergent. Our discussion of (5.11) shows that

$$y(x) = x\left(a - \int_0^x \left(1 - \frac{s}{x}\right) p(s) y^3(s)\, ds\right).$$

Hence

$$y(x) = x\left(a - \int_0^\infty p(s) y^3(s)\, ds\right) + \int_0^\infty s p(s) y^3(s)\, ds$$
$$+ \int_x^\infty (x - s) p(s) y^3(s)\, ds,$$

and (5.13) and (5.14) follow. The $o(1)$ terms in these two asymptotic formulas can thus be estimated more explicitly. Indeed from the equation immediately above and Lemma 6 there follows

$$|y(x) - (\beta x + \gamma)| \le |a|^3 \int_x^\infty (s - x) s^3 p(s)\, ds \quad (5.16)$$

and

$$|y'(x) - \beta| \le |a|^3 \int_x^\infty s^3 p(s)\, ds. \quad (5.17)$$

Notice also that, if β and γ are both zero, then

$$y(x) = \int_x^\infty (x - s) p(s) y^3(s)\, ds.$$

Thus for $0 < x_0 < \infty$ we have

$$\max_{x_0 \leq x < \infty} |y(x)| \leq \max_{x_0 \leq x < \infty} |y(x)| \int_{x_0}^{\infty} |a|^2 s^3 p(s) \, ds,$$

and we must therefore have $y(x) \equiv 0$. A similar argument shows that, in fact, the constants in (5.13) uniquely determine a solution of (5.5).

From Lemmas 4 and 5 it follows that the zeros of $y(x)$ are isolated in $[0, \infty)$. Since $a \neq 0$, β and γ are not both zero, and thus because of (5.13) all of the zeros of $y(x)$ lie in some bounded subinterval of $[0, \infty)$. Consequently, $y(x)$ has only finitely many zeros. This completes the proof of Lemma 7.

The sequence of results just proved is required primarily for the proof of the uniqueness assertion in Theorem VI. At this point we give a proof of the existence assertion of Theorem VI. The work to follow depends strongly on the results of Moore and Nehari.[13]

We let $\mathfrak{L}$ denote the class of those functions $y(x)$ which are locally absolutely continuous and not identically zero on $[0, \infty)$, whose first derivatives belong to $L^2[0, \infty)$, and which vanish at $x = 0$. From Holder's inequality we have, for $y \in \mathfrak{L}$,

$$|y(x)| \leq x^{\frac{1}{2}} \left(\int_0^{\infty} [y'(s)]^2 \, ds \right)^{\frac{1}{2}}. \tag{5.18}$$

Therefore we can define a functional $J(y)$ for y in $\mathfrak{L}$ by

$$J(y) = \left(\int_0^{\infty} (y')^2 \, dx \right)^2 \left(\int_0^{\infty} p y^4 \, dx \right)^{-1}. \tag{5.19}$$

From (5.18) there follows, for $y \in \mathfrak{L}$,

$$J(y) \geq \left(\int_0^{\infty} x^2 p \, dx \right)^{-1}. \tag{5.20}$$

For $n = 1, 2, \cdots$, let $p_n(x)$ be defined by $p_n(x) = p[1/(n+1)]$, $0 \leq x \leq 1/(n+1)$, $p_n(x) = p(x)$, $x \geq 1/(n+1)$. For each positive integer n let $\mathfrak{L}_n$ denote the class of functions in $\mathfrak{L}$ which do not vanish identically on $[0, n]$. Then on each $\mathfrak{L}_n$ we define

$$J_n(y) = \left(\int_0^n (y')^2 \, dx \right)^2 \left(\int_0^n p_n y^4 \, dx \right)^{-1}. \tag{5.21}$$

The results of Moore and Nehari[13] imply that for each n there is a function $y_n(x)$ of class C^2 on $[0, n]$, positive on $(0, n)$, and satisfying the following conditions.

---

[13] R. A. Moore and Z. Nehari, Trans. Am. Math. Soc. 93, 30 (1959).

(1) y_n minimizes J_n with respect to $\mathfrak{L}_n$:

$$J_n(y_n) = \inf_{y \in \mathfrak{L}_n} J_n(y) = M_n.$$

(y_n as yet is only defined on $[0, n]$; hence it does not belong to $\mathfrak{L}$.)

(2) y_n satisfies the differential equation

$$y_n'' + p_n y_n^3 = 0 \tag{5.22}$$

and the boundary conditions

$$y_n(0) = y_n'(n) = 0. \tag{5.23}$$

We note that if $y \in \mathfrak{L}$ satisfies (5.5), then

$$\int_0^{\infty} (y')^2 \, dx = \int_0^{\infty} p y^4 \, dx,$$

so that

$$J(y) = \int_0^{\infty} (y')^2 \, dx \tag{5.24}$$

for y, a solution of (5.5). Similarly, for the functions y_n we have, because of (5.22),

$$J_n(y_n) = \int_0^n (y_n')^2 \, dx. \tag{5.25}$$

We extend the definition of $y_n(x)$ to $[0, \infty)$ by putting $y_n(x) = y_n(n)$ for $x \geq n$. Thus extended, each y_n belongs to $\mathfrak{L}_n$, and y_n minimizes J_n in $\mathfrak{L}_n$. It is clear that

$$M_n = J_n(y_n) \geq J_{n+1}(y_n) \geq J_{n+1}(y_{n+1}) = M_{n+1}. \tag{5.26}$$

For a given $y \in \mathfrak{L}$ we obviously have

$$\lim_{n \to \infty} J_n(y) = J(y); \tag{5.27}$$

therefore,

$$\inf_{y \in \mathfrak{L}} J(y) = \lim_{n \to \infty} M_n = M_{\infty}. \tag{5.28}$$

By (5.25) and (5.26)

$$\int_0^{\infty} (y_n')^2 \, dx = M_n \leq M_1, \quad n = 1, 2, \cdots, \tag{5.29}$$

so that, by (5.18),

$$y_n(x) \leq (M_1 x)^{\frac{1}{2}}, \quad n = 1, 2, \cdots. \tag{5.30}$$

Since $y_n'(x) = 0$ for $x \geq n$, one has

$$y_n'(x) = -\int_x^n p_n(s) y_n^3(s) \, ds, \quad 0 \leq x \leq n. \tag{5.31}$$

From which it follows, using (5.30), that the sequence $\{y_n(x)\}$ is uniformly bounded on $(0, \infty)$; in fact,

$$|y_n'(x)| \leq M_1^{\frac{3}{2}} \int_x^{\infty} p(s) s^{\frac{3}{2}} \, ds. \tag{5.32}$$

Equicontinuity of both $\{y_n\}$ and $\{y_n'\}$ follows from (5.30), (5.31), and (5.22). Thus, by Ascoli's theorem,[14]

---

[14] A. N. Kolmogorov and S. V. Fomin, Ref. 12, p. 54.

there exists a subsequence $\{y_{n_k}(x)\}$ converging uniformly on compact intervals to a nonnegative solution $y_\infty(x)$ of (5.5); the subsequence of derivatives $\{y'_{n_k}\}$ converges to y'_∞ uniformly on compact intervals. The term on the right-hand side of (5.32) is square integrable as a function of x, so it follows from Lebesgue's dominated convergence theorem that

$$\int_0^\infty (y'_\infty(x))^2 \, dx = \lim_{k \to \infty} \int_0^\infty (y'_n(x))^2 \, dx = M_\infty.$$

By (5.20), $M_\infty > 0$. Therefore y_∞ is not the trivial solution of (5.5). Thus we have the following result.

Lemma 8: There exists a function y_∞ in the class $\mathfrak{L}$ such that

$$J(y_\infty) = \min_{y \in \mathfrak{L}} J(y).$$

The function y_∞ is a solution of (5.5) on $(0, \infty)$, is positive in $(0, \infty)$, and

$$y_\infty(0) = \lim_{x \to \infty} y'_\infty(x) = 0. \qquad (5.33)$$

The proof of the uniqueness assertion in Theorem VI depends on a study of the initial value problem (5.7) for (5.5)—in particular, on a study of the dependence of the solution $y(x, a)$ on the parameter a. This study is based on the use of the *variational equation* for (5.5):

$$\Delta'' + 3p(x)y^2(x)\Delta = 0. \qquad (5.34)$$

Solutions of the variational equation have the same asymptotic behavior as solutions of (5.5). More specifically we have the following.

Lemma 9: Let a be any real number and let $y(x) = y(x, a)$ in (5.34). Then (5.34) has two linearly independent solutions $\Delta_1(x)$ and $\Delta_2(x)$ satisfying, respectively, as $x \to \infty$,

$$\Delta_1(x) = x + o(1), \quad \Delta'_1(x) = 1 + o(1), \quad (5.35)$$

$$\Delta_2(x) = 1 + o(1), \quad \Delta'_2(x) = o(1). \qquad (5.36)$$

$\Delta_2(x)$ is uniquely determined by (5.36).

Proof: Because of (5.12) we clearly have

$$\int_0^\infty xp(x)y^2(x) \, dx < \infty.$$

The existence of Δ_1 and Δ_2 then follows from a theorem in Hartman.[15] The uniqueness of Δ_2 follows from the fact that any solution of (5.34) must be a linear combination of Δ_1 and Δ_2.

Assume that $q(x)$ is continuous for $0 \le x < \infty$. The differential equation

$$v'' + q(x)v = 0 \qquad (5.37)$$

is said to be *disconjugate* on the interval J in $[0, \infty)$ if no nontrivial solution of (5.37) has more than one zero in J. We have occasion to use the following disconjugacy criterion, which is a consequence of the Sturm comparison theorem.

Lemma 10: Let J be a finite or infinite open interval in $[0, \infty)$. Then (5.37) is disconjugate on J if there exists a function w which is positive and of class C^2 on J and satisfies

$$w'' + q(x)w \le 0 \quad \text{on} \quad J.$$

If the above condition holds and if either J is unbounded or if $w'' + q(x)w \not\equiv 0$ on J, then (5.37) is disconjugate on the closure of J.

Proof: See Hartman.[16]

Lemma 11: Let a and $y(x)$ in (5.34) be as in Lemma 9. Let $\Delta_2(x)$ be the unique solution of (5.34) satisfying (5.36) as $x \to \infty$. If Δ_2 changes sign in $(0, \infty)$ and if x_0 is the largest positive zero of Δ_2, then $[x_0, \infty)$ is a maximal interval of disconjugacy for (5.34).

Proof: Assume that $x_0 > 0$ can be defined as above. Then $\Delta_2(x)$ is positive on (x_0, ∞); and by the Sturm comparison theorem[17] no solution of (5.34) can have more than one zero in $[x_0, \infty)$ and (5.34) is disconjugate there. Let Δ_3 be a solution of (5.34) with a zero in (x_0, ∞). Then Δ_3 is linearly independent of Δ_2 and, therefore, by Lemma 9, must have the asymptotic behavior

$$\Delta_3(x) \sim \mu x, \quad \mu \ne 0, \quad \text{as } x \to \infty.$$

We can choose Δ_3 so that $\mu > 0$, Δ_3 will then be positive for large x; and since it can have just one zero in $[x_0, \infty)$, it will be negative at x_0. Let $0 \le x_1 < x_0$. Then for any positive number ϵ, $\Delta = \Delta_2 - \epsilon \Delta_3$ will be positive at x_0; and since Δ_2 changes sign at x_0, if ϵ is sufficiently small, Δ will have a zero in (x_1, x_0). However, because of (5.36), Δ has the asymptotic behavior $\Delta \sim -\epsilon \mu x$ as $x \to \infty$; consequently Δ has at least two zeros in $[x_1, \infty)$. Thus (5.34) fails to be disconjugate in $[x_1, \infty)$ for any $x_1 < x_0$.

We now adopt the following notation. For a real number a, let $\Delta(x, a)$ denote the solution of (5.34), with $y(x) = y(x, a)$, which satisfies the initial conditions

$$\Delta'(0) = 1, \quad \Delta(0) = 0. \qquad (5.38)$$

[15] P. Hartman, *Ordinary Differential Equations* (John Wiley & Sons, Inc., New York, 1964), p. 380.

[16] P. Hartman, Ref. 15, p. 362.
[17] P. Hartman, Ref. 15, p. 335.

Let $\beta = \beta(a)$ and $\gamma = \gamma(a)$ be defined by (5.15), with $y(x) = y(x, a)$. Because of Lemma 9, $\Delta(x, a)$ has the following asymptotic behavior as $x \to \infty$:

$$\Delta(x, a) = \mu x + \nu + o(1), \quad \Delta'(x, a) = \mu + o(1). \tag{5.39}$$

Lemma 12: For any real number a, $y(x, a)$ and $y'(x, a)$ are differentiable with respect to a and

$$(\partial/\partial a)y(x, a) = \Delta(x, a), \quad (\partial/\partial a)y'(x, a) = \Delta'(x, a). \tag{5.40}$$

Furthermore, $\beta(a)$ and $\gamma(a)$ are differentiable with respect to a and

$$(d/da)\beta = \mu, \quad (d/da)\gamma = \nu, \tag{5.41}$$

where $\mu = \mu(a)$ and $\nu = \nu(a)$ are the coefficients in (5.39).

Proof: The first assertion of this lemma does not follow from the standard theorem on differentiability with respect to initial conditions because of the singularity of the coefficient $p(x)$ at $x = 0$. The proof of the standard theorem[18] can, however, easily be adapted to this case. We omit these details here. One obtains in fact

$$y(x, a') - y(x, a) = (a' - a)[\Delta(x, a) + o(a' - a)] \tag{5.42}$$

and

$$y'(x, a') - y'(x, a) = (a' - a)[\Delta'(x, a) + o(a' - a)], \tag{5.43}$$

uniformly on compact intervals.

The coefficients μ and ν satisfy

$$\mu(a) = 1 - 3\int_0^\infty p(x)y^2(x, a)\Delta(x, a)\, dx,$$

$$\nu(a) = 3\int_0^\infty xp(x)y^2(x, a)\Delta(x, a)\, dx.$$

So if we use (5.42) and (5.43) in (5.15), (5.41) follows.

Lemma 13: Let a be such that $y(x, a)$ is positive on $(0, \infty)$ and $\beta(a) = 0$. Then $\Delta(x, a)$ has precisely one zero in $(0, \infty)$, and $\mu(a) < 0$.

Proof: Let a be as in the statement of the lemma and let $y = y(x, a)$, $\Delta = \Delta(x, a)$. We show first that Δ has at least one zero on $(0, \infty)$. From (5.5) and (5.34) we have

$$y''\Delta - \Delta''y = 2py^3\Delta.$$

---

[18] P. Hartman, Ref. 15, pp. 95–96.

Integration of this inequality, using (5.7) and (5.38), yields

$$y'(x)\Delta(x) - \Delta'(x)y(x) = 2\int_0^x p(s)y^3(s)\Delta(s)\, ds.$$

Upon letting $x \to \infty$, we obtain, since $\beta(a) = 0$,

$$-\mu(a)\gamma(a) = 2\int_0^\infty p(s)y^3(s)\Delta(s)\, ds. \tag{5.44}$$

The assumption that $\Delta(x) > 0$ on $(0, \infty)$ implies that

$$\mu(a) = \lim_{x \to \infty} \Delta'(x, a) \geq 0$$

and also that the integral on the right in (5.44) will be positive. Since $\gamma(a) > 0$, this yields a contradiction; so $\Delta(x)$ must change sign at least once on $(0, \infty)$.

Now let x_1 be the uniquely determined point where $y'(x) = y(x)$. In order to show that Δ has at most one zero in $(0, \infty)$ we show that (5.34) [with $y(x) = y(x, a)$] is disconjugate in $[x_1, \infty)$ and also in $[0, x_1]$. Take $w = y - y'$; then $w > 0$ on (x_1, ∞) and

$$w'' + 3py^2w = (2p + p')y^3.$$

Since

$$(2p + p') = x^2 \operatorname{csch}^4 x[2 + (2/x) - 4\coth x] < 0$$

on $(0, \infty)$, it follows from Lemma 10 that (5.34) is disconjugate on (x_1, ∞). Next take $w_1(x) = -\varphi(x)w(x)$ where $\varphi(x)$ is a positive C^2 function on $(0, \infty)$. Then w_1 satisfies

$$w_1'' + 3py^2w_1 = -p\varphi[2 + (\ln p)' + 2(\ln \varphi)']y^3 - (2\varphi' - \varphi'')y' - \varphi''y. \tag{5.45}$$

Since $w < 0$ on $(0, x_1)$, in order to prove that (5.34) is disconjugate on $[0, x_1]$ it suffices to find a positive C^2 function which satisfies, on $(0, \infty)$,

$$2 + (\ln p)' + 2(\ln \varphi)' > 0, \tag{5.46}$$

$$2\varphi' - \varphi'' > 0, \tag{5.47}$$

$$\varphi'' > 0. \tag{5.48}$$

Take $\varphi(x) = xe^x$. It is easily verified that (5.47) and (5.48) hold and that $2 + (\ln p)' + 2(\ln \varphi)' = 4[1 + (1/x) - \coth x] > 0$ on $(0, \infty)$, so that (5.46) holds. Thus from (5.45) we have

$$w_1'' + 3py^2w_1 < 0 \quad \text{on} \quad (0, x_1).$$

So, by Lemma 10, (5.34) is disconjugate on $[0, x_1]$. It follows that Δ has precisely one positive zero x_2 and $x_2 > x_1$. Since $x_2 > x_1$, $[x_2, \infty)$ cannot be a maximal interval of disconjugacy for (5.34). Hence it follows from Lemma 11 that Δ is linearly independent of the solution $\Delta_2(x)$ of (5.34), which has the asymptotic

behavior (5.36). Thus we must have $\mu(a) < 0$. This completes the proof of Lemma 13.

Let A denote the set of all positive real numbers a for which $y(x, a)$ has at least one zero in $(0, \infty)$ and let B denote the set of all positive real numbers a for which $y(x, a)$ is positive in $(0, \infty)$ and $\beta(a) > 0$. From the continuous dependence of $y(x, a)$ and $\beta(a)$ on a it follows that A and B are open. We show that A and B have a *single common limit point* a_0, and $A = (a_0, \infty)$ and $B = (0, a_0)$. The uniqueness assertion of Theorem VI follows in an elementary way once this has been shown. First let a be such that $y(x, a)$ is positive on $(0, \infty)$ and $\beta(a) = 0$. By Lemma 8 there is at least one such a, and by Lemmas 12 and 13 $(d/da)\beta(a) < 0$. It follows that for a' in a neighborhood of a, $\beta(a') > 0$ for $a' < a$ and $\beta(a') < 0$ for $a' > a$. Clearly, if $a' > 0$ and $\beta(a') < 0$, we have $a' \in A$. Thus it follows that $a' \in A$ for $a' > a$ and a' sufficiently near to a. On the other hand, using the fact that $\beta(a') > 0$ when $a' < a$ and a' is sufficiently close to a, it follows from (5.42) and (5.16) that $y(x, a')$ is positive on $(0, \infty)$, and thus that $a' \in B$ when $a' < a$ and a' is sufficiently close to a. In other words, we have shown that a is the right end point of an interval in B and the left end point of an interval in A. Now suppose that a is any positive number which does not belong either to A or to B. Then $y(x, a)$ must be positive on $(0, \infty)$ and $\beta(a)$ must be zero; so, by the argument just given, a must be the left end point of an interval in B and the right end point of an interval in A. Since A and B are open, it follows that there can be just one such point a_0, and that $A = (a_0, \infty)$, $B = (0, a_0)$.

In view of the definitions of A and B, the characterization of these sets which we have just obtained implies that there is precisely one solution $y(x)$ of (5.5) which vanishes and has a finite derivative at $x = 0$, is positive on $(0, \infty)$, and for which

$$\lim_{x \to \infty} y'(x) = 0.$$

The unique solution $U(x)$ of the eigenvalue problem (5.1)–(5.3) is obtained by taking $U(x) = \alpha^{-1} y(x)$, where

$$\alpha = (y'(0))^{\frac{1}{2}} = a_0^{\frac{1}{2}}.$$

Since the solution $y(x) = y(x, a_0)$ of (5.5) is actually the solution whose existence was proved in Lemma 8, we have the following theoretical upper bound for α:

$$\alpha = (y'(0))^{\frac{1}{2}} \leq \left(M_\infty^{\frac{3}{2}} \int_0^\infty x^{\frac{3}{2}} p(x) \, dx \right)^{\frac{1}{2}};$$

cf. the derivation of (5.32). An arbitrary test function in $\mathfrak{L}$ can be used in J to obtain an upper estimate for
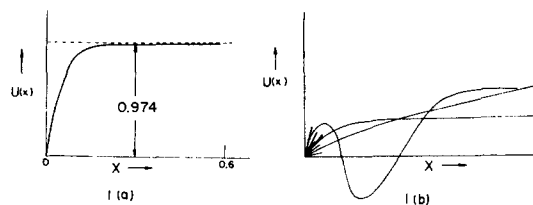

FIG. 1. The solutions $U(x)$. (a) The solution $U(x)$ corresponding to the fundamental theorem. [The method used to compute $U(x)$ is described in the Appendix.] (b) The qualitative plot of the solutions $U(x)$ corresponding to the initial value problem $U(0) = 0$, $U'(0) = a$. The bold lines correspond to the critical initial slopes of the solutions which cross x axis a finite number of times and then tend asymptotically to a flat line. The amplitudes of the oscillations increase, whereas the crossing slopes decrease, with x.

M_∞. A lower bound for α can be obtained as follows. Let a be any positive number; then

$$y'(x, a) = a - \int_0^x p(s) y^3(s) \, ds.$$

So from (5.12) it follows that

$$y'(x, a) \geq a \left(1 - a^2 \int_0^x p(s) s^3 \, ds \right)$$

on an interval $(0, x_1)$, provided $y(x)$ remains positive on that interval. Consequently, if

$$a < \left(\int_0^\infty p(s) s^3 \, ds \right)^{-\frac{1}{2}}, \tag{5.49}$$

then y' and y must remain positive on $(0, \infty)$. Therefore $a \in B$ when (5.49) holds. This gives

$$\alpha = a_0^{\frac{1}{2}} \geq \left(\int_0^\infty p(s) s^3 \, ds \right)^{-\frac{1}{4}}.$$

The value of α obtained from the computer was 1.4343. For a graph of the function $U(x)$ see Fig. 1.

We conclude this section with a theorem concerning the existence of additional solutions of the problem (5.1)–(5.3).

Theorem VII: For each integer $n \geq 0$ there exists a constant $\alpha = \alpha_n$ and a C^2 function $U = U_n(x)$ which has exactly n zeros in $(0, \infty)$ and satisfies (5.1), (5.2) and (5.3), with $\alpha = \alpha_n$.

The proof of this theorem is not given. For a related result the reader is referred to Theorem VII of the paper of Moore and Nehari[13] quoted above. A proof of our Theorem VII above can be based on their results in much the same way as was the proof of the existence assertion in Theorem VI.
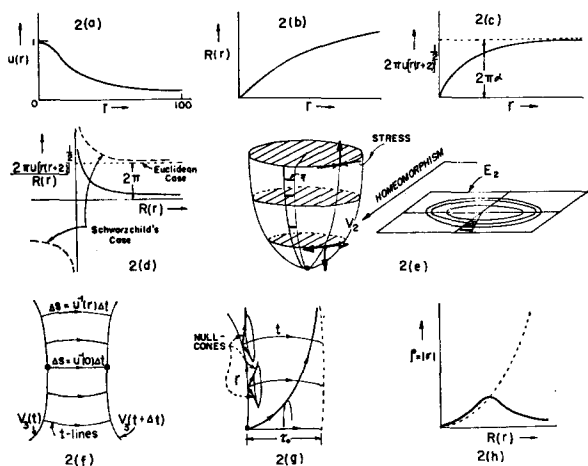
FIG. 2. The properties of the universe. (a) The exact solution of the fundamental theorem. (b) The qualitative plot of the radial length $R(r)$ versus r. (c) The qualitative plot of the circumference of a circle versus r. (d) The qualitative plot of the ratio of the circumference divided by the radial length versus $R(r)$. (e) The two-dimensional symbolic representations of V_3. To regain the suppressed dimension, the circles should be replaced by the spheres. The arrows indicate the radial and transverse stresses. The black spot at the lowest tip indicates that the elementary flatness is violated there. (f) The symbolic representation of V_4. The proper time ΔS along a t line increases as r increases. (g) The qualitative plot of the radial null lines in V_4. (h) The qualitative plot of the radial mass density $\rho = \int_0 M^4_{.4}(-g')^{\frac{1}{2}} d\theta \, d\varphi$ and the radial charge density $= \int j_4 n^4 (-g')^{\frac{1}{2}} d\theta \, d\varphi$ against r. The broken line denotes the radial density r^2 corresponding to the constant unit density.

VI. PROPERTIES OF THE UNIVERSES OBTAINABLE FROM THE SOLUTIONS OF $U'' + \alpha^2(x \, \text{csch}^2 \, x)^2 U^3 = 0$

The properties of the permissible universes discussed are threefold, viz., (i) geometrical (in the Riemannian sense), (ii) topological, and (iii) physical. For this purpose the coordinate transformation $r = \coth x - 1$, $u(r) = U(x)$, is made so that in the new system the metric (4.16) assumes a simpler form:

$$\Phi = -u^2[dr^2 + r(r+2) \times (d\theta^2 + \sin^2\theta \, d\varphi^2)] + u^{-2} dt^2. \quad (6.1)$$

In the first place the universe corresponding to the solution in the fundamental theorem (VI) is discussed. For the geometrical properties of this universe a t-constant hypersurface V_3 is dealt with first, followed by a discussion of the universe V_4 (Fig. 2).

(IA) The geometrical properties of V_3 are the following.

(a) $(-g')^{\frac{1}{2}} = u^3 r(r+2) \sin\theta > 0$,
        for $0 < r < \infty$, $0 < \theta < \pi$.

(b) The total volume is

$$\lim_{r \to \infty} 4\pi \int_0^r u^3 r'(r'+2) \, dr' = \infty$$

(logarithmically divergent).

(c) The radical distance $R(r) = \int_0^r u \, dr$ is a monotonically increasing function of r. Moreover,

$$\lim_{r \to \infty} R(r) = \infty$$

(logarithmically divergent).

[(b) and (c) show that V_3 is open.]

(d) The length of the circumference of a circle at the radial distance $R(r)$ is $2\pi u[r(r+2)]^{\frac{1}{2}}$, and this starts from zero, growing monotonically up to the finite value $2\pi\alpha$.

(e) The ratio of the circumference divided by the radial length is

$$2\pi u[r(r+2)]^{\frac{1}{2}} \Big/ \int_0^r u \, dr.$$

It starts with infinite slope, decaying monotonically to zero.

(f) The area of a sphere at the radial distance $R(r)$ is $4\pi u^2 r(r+2)$, and it starts from zero, growing monotonically as R increases up to the finite value $4\pi\alpha^2$. ["There is not so much elbow-room in distant parts as Euclid supposed."]

(g) The solid angle that a spherical surface subtends at the origin is

$$4\pi u^2 r(r+2) \Big/ \left(\int_0^r u \, dr\right)^2,$$

and it begins with the infinite slope, decaying monotonically to zero.

(h) The geodesic deviation between two adjacent radial geodesics (r lines are geodesics in V_3 and V_4) is $\eta = u[r(r+2)]^{\frac{1}{2}}\Delta\omega$ (the metric of the unit sphere being $d\omega^2 = d\theta^2 + \sin^2\theta \, d\varphi^2$). η grows monotonically from zero up to $\alpha\Delta\omega$.

(i) The Ricci curvature tensor of V_3 is

$$(R^{(3)x}_\beta) = \begin{bmatrix} 2u^{-2}[u^{-1}u'' - u^{-2}u'^2 + u^{-1}u'(r+1)r^{-1}(r+2)^{-1}] - r^{-2}(r+2)^{-2} & 0 & 0 \\ 0 & u^{-2}[u^{-1}u'' + 3u^{-1}u'(r+1)r^{-1}(r+2)^{-1}] & 0 \\ 0 & 0 & u^{-2}[u^{-1}u'' + 3u^{-1}u'(r+1)r^{-1}(r+2)^{-1}] \end{bmatrix},$$

$$\lim_{r \to \infty} (R^{(3)x}_\beta) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\alpha^{-2} & 0 \\ 0 & 0 & -\alpha^{-2} \end{bmatrix}.$$

The metric of V_3 has a singularity at $r \to \infty$, but the Ricci curvature tensor is regular there, showing that the singularity of the metric is only a coordinate singularity.

(IB) The properties of V_4 are the following.

(a) $(-g)^{\frac{1}{2}} = u^2 r(r + 2) \sin \theta \geq 0$.

(b) The total space–time volume of V_4 is

$$\lim_{\substack{r \to \infty \\ t \to \infty}} 4\pi \int_0^r \int_0^t u^2 r'(r' + 2)\, dr'\, dt' = \infty$$

(quadratically divergent).

(c) The t lines are not geodesics in V_4, except the one at the spatial origin. The distance along a t line between two t-constant hypersurfaces grows monotonically to infinity as the radial coordinate r of the t-line increases.

(d) The radial null geodesics are characterized by the equation

$$t - t_0 = \int_{r_0}^r u^2\, dr'.$$

The slope of the null curve $dr/dt = u^{-2}$ increases monotonically to infinity as the r or t coordinate increases.

(e) The surviving components of the Riemann curvature tensor[19] in V_4 are given by

$$R_{2323} = -u^2 r(r + 2) \sin^2 \theta [1 - r(r + 2) \\ \times \{u^{-1}u' + r(r + 1)/r(r + 2)\}^2],$$

$$R_{1212} = u^2 r(r + 2)[u^{-1}u'' - u^{-2}u'^2 + u^{-1}u' \\ \times (r + 1)/r(r + 2) - r^{-2}(r + 2)^{-2}],$$

$$R_{1313} = \sin^2 \theta R_{1212},$$

$$R_{1414} = u^{-2}(u^{-1}u'' - 3u^{-2}u'^2),$$

$$R_{2424} = u^{-2} r(r + 2)[u^{-2}u'^2 + u^{-1}u'(r + 1)/r(r + 2)],$$

$$R_{3434} = \sin^2 \theta R_{2424}.$$

II. The topological properties of the universe are the following: V_3 is topologically Euclidean, and in this sense it is simpler than the spatial hypersurface of Schwarzchild's universe, which has a "handle";[20] the V_3 under consideration can be embedded into a four-dimensional flat space.

III. The physical properties of the universe are the following.

(a) Let an idealized observer ("eyeless, headless mannikin!") at rest with the radial coordinate r compare his proper time with another observer at the origin.[21] He would discover that, between the same "time coordinate" lapses, his proper time runs much faster than his compatriot at the origin. The spectral lines emanating from distant sources at rest will, consequently, appear displaced towards the violet to the observer at the origin.

(b) A ray of light traveling radially from the origin along a null geodesic takes only a finite proper time

$$\tau_0 = u^{-1}(0) \int_0^\infty u^2\, dr$$

of the observer at the origin to complete the journey through the infinite distance. This is because the velocity of light (so far as the observer at the origin can assess) increases monotonically to infinity as the light recedes from the origin. [It is needless to mention that the local velocity of light according to the local measurements is always ± 1.]

(c) The condition

$$\lim_{x \to \infty} U'(x) = 0$$

(cf. Theorem VI) implies that

$$\lim_{r \to 0} \frac{du}{dr} = 0$$

in view of Eq. (5.17). For the physical meaning of this condition, consider the radial equation of motion for a neutral test particle governed by a timelike geodesic:

$$d^2 r/ds^2 = 2u^{-3}(du/dr), \quad ds^2 = -u^2\, dr^2 + u^{-2}\, dt^2.$$

From this equation it is evident that the radial acceleration and force at the origin of the spherically symmetric universe is zero.

(d) The "uniplanar" ($\theta = \frac{1}{2}\pi$) timelike geodesics, which represent the motions of the neutral test particles in V_4, are given by the inversion of the formula

$$\varphi - \varphi_0 = -\int_{r_0}^r [r(r + 2)\{u^2 h^{-2}(E^2 u^2 - 1) \\ - r^{-1}(r + 2)^{-1}\}]^{-\frac{1}{2}}\, dr,$$

$$= \int_{x_0}^x [U^2 h^{-2}(E^2 U^2 - 1) - \sinh^2 x]^{-\frac{1}{2}}\, dx,$$

where h and E represent constant areal velocity and energy, respectively.

(e) Let the combined material and electro-static stress-energy-momentum tensor be expressed as $-(8\pi)^{-1}G^a_{.b} = M^a_{.b} + \mathcal{E}^a_{.b}$. Then the surviving

[19] J. L. Synge, *Relativity: The General Theory* (North-Holland Publishing Company, Amsterdam, 1960), p. 271.

[20] M. D. Kruskal, Phys. Rev. **119**, 1743 (1960).

[21] If the Schwarzchild's radius k in (4.15) is chosen to be $k = \frac{1}{2}u(0) = 0.487$, then the "time coordinate" t will be the proper time for the observer at rest at the origin.
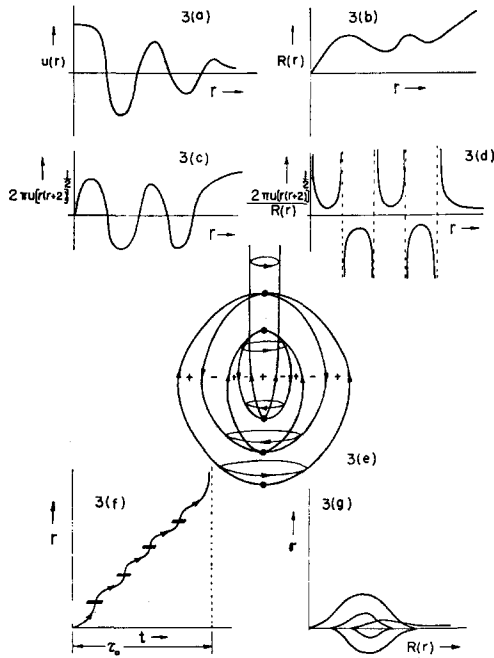
FIG. 3. Properties of the universe corresponding to $u(r)$ having five zeros. (a) The qualitative plot of $u(r)$ with five zeros (and initially flat slope) versus r. (b) The qualitative plot of the radial length $R(r)$ versus r. The curve oscillates four times. (c) The qualitative plot of the circumference of a circle versus r. The curve is oscillatory, and the negative parts should be interpreted as oppositely oriented. (d) The ratio of the circumference divided by the radial length versus r. The curve has singular behavior at five points. (e) The two-dimensional symbolic representation of V_3. The suppressed dimension can be regained by replacing the circles by the spheres. There are five shells, and a common point between two adjacent shells is denoted by a black spot where the ratio of the circumference divided by the radius becomes infinite. (f) The qualitative plot of the radial null line. The shaded slits correspond to the boundaries between two adjacent shells. So far as the observer at origin can judge, light penetrates these boundaries with infinite speed. (g) The qualitative plot of the radial charge density

$$\sigma = \int j_4 n^4 (-g')^{\frac{1}{2}} \, d\theta \, d\varphi$$

against the radial distance $R(r)$. It shows that the sign of the charge density alternates from shell to shell.

components of this tensor are given by

$$\mathcal{E}^1_{.1} = -\mathcal{E}^2_{.2} = -\mathcal{E}^3_{.3} = \mathcal{E}^4_{.4} = (8\pi)^{-1} u^{-4} u'^2,$$

$$M^1_{.1} = -M^2_{.2} = -M^3_{.3} = -(8\pi)^{-1} u^{-2} r^{-2} (r + 2)^{-2},$$

$$M^4_{.4} = (16\pi)^{-1} \alpha^2 [\ln (1 + 2/r)]^2$$
$$+ (8\pi)^{-1} u^{-2} r^{-2} (r + 2)^{-2},$$

where the prime denotes differentiation with respect to r. The electrostatic stress-energy–momentum tensor $\mathcal{E}^\alpha_{.b}$ has the usual algebraic relations for the spherically symmetric case. The material stress tensor $M^\alpha_{.\beta}$ is shear free, but not a simple pressure. There exists a sphere on which the electrostatic stress is exactly canceled by the material stress.

$M^4_{.4}$ stands for the material density. It can be expressed as $M^4_{.4} = {}_0M^4_{.4} + {}_sM^4_{.4}$, where

$${}_0M^4_{.4} = (16\pi)^{-1} [m \ln (1 + 2/r)]^2.$$

$M^4_{.4} = (4\pi)^{-1} m^2 x^2$ is the purely material density, and ${}_sM^4_{.4} = (8\pi)^{-1} u^{-2} r^{-2} (r + 2)^{-2}$ denotes the material stress-energy density. The total purely material mass $\int {}_0M^4_{.4}(-g')^{\frac{1}{2}} \, d_3v$ exactly equals m. But both the total material stress energy and the electrostatic energy are logarithmically divergent.

(f) The usual condition of the square integrability of Schrödinger's wavefunction ψ cannot be taken as such into the general relativity due to the general covariance of the theory. A condition like the finiteness of the total mass does not overcome the difficulty either—because there is no satisfactory definition of the total mass in the general relativity. The only satisfactory condition of the square integrability emerges from the definition of the total charge and by equating the charge[22] to α, i.e.,

$$\int_{V_3} j_4 n^4 \, d_3 u$$

$$= i\alpha \int_{V_3} (D^{*4}\psi^* \cdot \psi - \psi^* D^4\psi) n^4 \, d_3v,$$

$$= 2\alpha^2 \int_{V_3} \psi^* \psi g^{44} \, d_3v,$$

$$= -\int_0^\infty r(r + 2)[u''(r) + 2(r + 1)r^{-1}(r + 2)^{-1}u'] \, dr,$$

$$= -\int_0^\infty U''(x) \, dx = U'(0) = \alpha,$$

where j_4 is the charge density, n^i is the unit vector on the hypersurface V_3, and condition $U'(0) = \alpha$ has been taken into consideration in posing the nonlinear eigenvalue problem (cf. Theorem VI).

The various properties of the universe V_4, constructible out of the solution $u(r)$ which has finite numbers of zeros and asymptotically goes to zero, have been graphically summarized in Fig. 3. The main difficulty with such a model is that the metric and also the Ricci curvature tensor are singular at a zero of $u(r)$. But light can penetrate smoothly through the various shells, though the "coordinate velocity" of the penetration is infinite. A V_3 in such a universe comprises the finite number shells like an onion, with alternately positive and negative charge distributions. Unlike the V_3 corresponding to the fundamental solution, the space under consideration is not orientable—there is no well-defined interior to the sphere. The illustration of Wheeler and Misner[1] can well be applied in such a space. Suppose a hypothetical criminal is locked up inside a spherical jail in the first shell of V_3. If the radial coordinate of the sphere is increased, the sphere will eventually contract to a

[22] The common usage of α for the fine-structure constant has been changed to α^2.

point [the black spots of the Fig. 3(e)], and it will begin to gain size again, inside out, now in the second shell. If the jail bird survives this choking, he will find himself free and exterior to the prison!

Such a space $V_3 = V_3^n$, which is determined by one of the solutions U_n, $n > 0$ (see Theorem VII), is topologically equivalent to the space obtained by "joining" n 3-spheres one to another in a chain, like a sausage, and then joining the last of these to a three-dimensional hyperplane. The space which we obtain in this way clearly is not topologically a Euclidean space. In fact, at the points where two spheres are joined together, the space is not even locally Euclidean. Therefore, the question of, for example, the orientability of V_3^n becomes meaningless. The most significant topological invariant associated with V_3^n is the number n itself, i.e., the number of spheres in the above representation of V_3^n. The number n can be characterized as the third Betti number of V_3^n.

VII. CONCLUDING REMARKS

The field-theoretic models of matter as treated here are complete and self-consistent and emerge from the combination of well-accepted theories. Rigorous techniques of attacking nonlinear eigenvalue problems have been employed for the first time to determine electrogravitational structure of elementary particles. Existence of enumerable numbers of eigenvalues of the fine-structure constant, bare mass, and observable energy have been proved. It is also shown that each of these energy levels is associated with a shell structure of matter in the range of 10^{-33} cm.

(It is possible that the different coupling constants of nature are different eigenvalues of the fine-structure constant.) The total electrostatic self-energy is logarithmically divergent, but that does not affect the observable energy of the matter field. (The particle appears to be bare in spite of the heavy clothing!)

The physical properties (mass $\sim 3 \times 10^{-5}$ g, circumference $\sim 2\pi \times 2.2 \times 10^{-33}$ cm) of these particles are remarkably similar to those of the geons. There is no reason why these particles should not be created in very high-energy interactions, like in the quasistellar sources or in a nova or supernova, and hence their existence is predicted.

However, there are some drawbacks of the present theory as developed here. First, the second quantization, which has been left out, should have been taken into account to incorporate the uncertainty principle. Secondly, the attempt to interpret the solutions (other than the fundamental one) presents the following difficulty. An assumption implicit in the method of construction of universes is that the manifold V_3

is topologically equivalent to Euclidean space. A space topologically inequivalent to an open subset of Euclidean space cannot be covered without singularity by a single nonsingular coordinate system. For such a space the traditional tensor analysis, as used here, is inadequate. The universes constructible from solution $U(r)$ (other than the fundamental one) have singular spheres pinched to a single point (the black spots!). In these cases the universe is not a manifold; it is, rather, a space which is locally Euclidean (in topological sense) everywhere except at the pinched points. Therefore, these universes cannot be interpreted strictly within the framework of the theory. Nevertheless, these universes are treated as being physical, partly because they bear analogy to the energy eigenstates of the wave-mechanical systems, and partly due to the fact that Schwarzchild's universe carrying a handle (topologically not Euclidean in the global sense) has been tested experimentally with success.

The following problems arising out of this work have been left open: (i) the uniqueness proof of the solutions $U(x)$ other than the fundamental one; (ii) completeness of these solutions; (iii) the question of existence of solutions $U(x)$ with oscillatory singularity at the origin (either solutions start from zero or have oscillatory singularities, nothing else can happen); (iv) the study of the partial differential equation (4.17); (v) the Rainich problem for combined Klein–Gordon–Maxwell–Einstein field equations; (vi) the second quantization of these combined field equations.

ACKNOWLEDGMENTS

APPENDIX

In this appendix we describe very briefly the method used to compute $U(x)$. It is easily verified that if $y(x) = y_\infty(x)$ is as in Lemma 8, then $y(x)$ satisfies

$$y(x) = \int_0^x tp(t)y^3(t)\, dt + x \int_x^\infty p(t)y^3(t)\, dt.$$

Putting $y(x) = x\omega(x)$, we find that $\omega(x)$ satisfies

$$\omega(x) = x^{-1}\int_0^x t^4 p(t)\omega^3(t)\, dt + \int_x^\infty t^3 p(t)\omega^3(t)\, dt.$$

For convenience we consider instead the nonlinear eigenvalue problem

$$\lambda v(x) = x^{-1}\int_0^x t^4 p(t)v^3(t)\, dt + \int_x^\infty t^3 p(t)v^3(t)\, dt,$$

$$\int_0^\infty t^4 p(t)v^3(t)\, dt = \lambda.$$

This problem can be solved by successive approximation as follows. Beginning with a function $v_0(x)$ which is positive on $[0, \infty)$ and satisfies $v_0(x) = x^{-1} + O(x^{-1})$ as $x \to \infty$, functions $v_n(x)$ with the same properties are defined successively by

$$\lambda_{n+1}v_{n+1}(x) = x^{-1}\int_0^x t^4 p(t)v_n^3(t)\, dt + \int_x^\infty t^3 p(t)v_n^3(t)\, dt,$$

where

$$\lambda_{n+1} = \int_0^\infty t^4 p(t)v_n^3(t)\, dt.$$

This procedure was carried out numerically, on a computer, and was proved to be convergent to a solution pair λ, $v(x)$ of the above eigenvalue problem. $U(x)$ is obtained directly by suitably normalizing $xv(x)$. This method of successive approximation is an adaptation of a method used by Nehari[23] to prove an existence theorem for solutions of nonlinear integral equations. The use of Nehari's method is greatly facilitated in this case by the fact that the operator in question is homogeneous.

---

[23] Z. Nehari, Math. Z. 72, 175 (1959).

Generalized WKB Method with Applications to Problems of Propagation in Nonhomogeneous Media*

E. BAHAR

Electrical Engineering Department, University of Colorado, Boulder, Colorado

(Received 27 February 1967)

A generalized WKB method is derived for the solution of the general second-order differential equation. The problem is reduced to the solution of two coupled first-order differential equations. By an appropriate choice of auxiliary functions, the coupling coefficients may be made sufficiently small to facilitate the solution of the coupled equations. It is shown that these solutions can be used in a range of problems in which the regular WKB solutions fail. These generalized solutions may also be used to derive asymptotic expansions of known functions. Applications of the method to higher-order differential equations are indicated, and solutions to the nonlinear Riccati equation are considered.

1. INTRODUCTION

THIS paper considers the solution of a general second-order differential equation for which no known solutions in closed form exist. The second-order differential equation is transformed into two coupled first-order differential equations. The plane wave type "local" auxiliary solutions are first used to solve the differential equation leading to the regular coupled WKB solutions. Next, a "local" auxiliary wave solution of the Airy integral type is used to facilitate the solution of the coupled equations. These lead to the generalized WKB solutions. In each case physical interpretations of the methods are discussed. From these follow an alternate method to derive the coupling coefficients that are considered as "differential" transmission and reflection coefficients.

Two different iterative methods for solving the coupled equations are discussed. The first yields an infinite series expansion, and the second, an infinite product expansion.

A physical problem for which a rigorous solution exists is considered in Sec. 6 to illustrate the effectiveness of the generalized WKB method. Langer's

---

Putting $y(x) = x\omega(x)$, we find that $\omega(x)$ satisfies

$$\omega(x) = x^{-1}\int_0^x t^4 p(t)\omega^3(t)\,dt + \int_x^\infty t^3 p(t)\omega^3(t)\,dt.$$

For convenience we consider instead the nonlinear eigenvalue problem

$$\lambda v(x) = x^{-1}\int_0^x t^4 p(t)v^3(t)\,dt + \int_x^\infty t^3 p(t)v^3(t)\,dt,$$

$$\int_0^\infty t^4 p(t)v^3(t)\,dt = \lambda.$$

This problem can be solved by successive approximation as follows. Beginning with a function $v_0(x)$ which is positive on $[0, \infty)$ and satisfies $v_0(x) = x^{-1} + O(x^{-1})$ as $x \to \infty$, functions $v_n(x)$ with the same properties are defined successively by

$$\lambda_{n+1}v_{n+1}(x) = x^{-1}\int_0^x t^4 p(t)v_n^3(t)\,dt + \int_x^\infty t^3 p(t)v_n^3(t)\,dt,$$

where

$$\lambda_{n+1} = \int_0^\infty t^4 p(t)v_n^3(t)\,dt.$$

This procedure was carried out numerically, on a computer, and was proved to be convergent to a solution pair λ, $v(x)$ of the above eigenvalue problem. $U(x)$ is obtained directly by suitably normalizing $xv(x)$. This method of successive approximation is an adaptation of a method used by Nehari[23] to prove an existence theorem for solutions of nonlinear integral equations. The use of Nehari's method is greatly facilitated in this case by the fact that the operator in question is homogeneous.

———
[23] Z. Nehari, Math. Z. 72, 175 (1959).

Generalized WKB Method with Applications to Problems of Propagation in Nonhomogeneous Media*

E. BAHAR

Electrical Engineering Department, University of Colorado, Boulder, Colorado

(Received 27 February 1967)

A generalized WKB method is derived for the solution of the general second-order differential equation. The problem is reduced to the solution of two coupled first-order differential equations. By an appropriate choice of auxiliary functions, the coupling coefficients may be made sufficiently small to facilitate the solution of the coupled equations. It is shown that these solutions can be used in a range of problems in which the regular WKB solutions fail. These generalized solutions may also be used to derive asymptotic expansions of known functions. Applications of the method to higher-order differential equations are indicated, and solutions to the nonlinear Riccati equation are considered.

1. INTRODUCTION

THIS paper considers the solution of a general second-order differential equation for which no known solutions in closed form exist. The second-order differential equation is transformed into two coupled first-order differential equations. The plane wave type "local" auxiliary solutions are first used to solve the differential equation leading to the regular coupled WKB solutions. Next, a "local" auxiliary wave solution of the Airy integral type is used to facilitate the solution of the coupled equations. These lead to the generalized WKB solutions. In each case physical interpretations of the methods are discussed. From these follow an alternate method to derive the coupling coefficients that are considered as "differential" transmission and reflection coefficients.

Two different iterative methods for solving the coupled equations are discussed. The first yields an infinite series expansion, and the second, an infinite product expansion.

A physical problem for which a rigorous solution exists is considered in Sec. 6 to illustrate the effectiveness of the generalized WKB method. Langer's
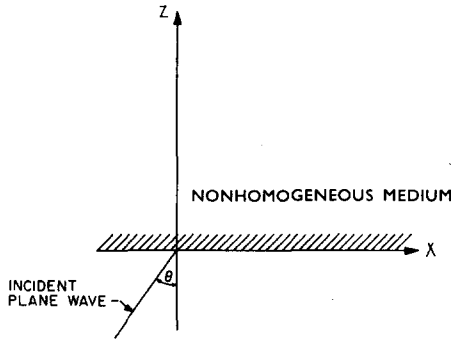
FIG. 1. Plane wave incident from below a nonhomogeneous region with its normal in the x-z plane at an angle θ to the vertical.

solution to this problem is also considered. It is shown that, using the coupled differential equations derived in this paper, it is possible to derive a significant correction factor to Langer's solution. Use of this method to derive asymptotic expansions of known functions for regions in which they are poorly tabulated is inferred. Extension of this method to higher-order differential equations is indicated. Finally, the relationship between the regular and generalized WKB solutions and solutions of the nonlinear Riccati equation are considered.

2. STATEMENT OF THE PROBLEM

In numerous problems in electromagnetic theory it is necessary to solve the second-order differential equation of the form

$$L(\Phi) \equiv \frac{1}{w}\frac{d}{du}\left(p\frac{d\Phi}{du}\right) + k^2 q^2 \Phi = 0, \qquad (2.1)$$

in which Φ, a scalar function of u, is related to the field quantities; w, p, and q are known functions of u depending on the particular problem to be solved; and k is a constant (the wavenumber for free space). For the case in which the medium of propagation is homogeneous, for instance, the solution for Φ can be expressed in terms of a combination of two well-known linearly independent functions that may be identified as outgoing and incoming waves. Furthermore, when the properties of the medium of propagation have certain particular spatial variations, it is also possible to express the exact solution in terms of two known linearly independent functions. But, in general, when the spatial variations of the medium of propagation are arbitrary, the solution for the field quantities cannot be written (in closed form) in terms of two known linearly independent functions. If, in this case, it is attempted to write down the solution in terms of a combination of two known functions, it will be observed that, in general, there will be continuous coupling between these solutions.

As a specific example to the general equation (2.1) treated in the paper, consider the case in which a horizontally polarized plane wave is incident upon a nonmagnetized ionized medium with its normal in the x-z plane at an angle θ to the vertical (see Fig. 1). The ionized medium in the region $z > 0$ is assumed to vary with the z coordinate only; hence the fields are independent of the y coordinate, and, as a consequence of Snell's law, all the field quantities contain a factor $\exp\{-ikSx\}$ in which $S = \sin\theta$. Using Maxwell's equations, it can be shown after certain simplifications that the differential equations satisfied by E_y and H_x in the ionized medium are

$$\partial E_y/\partial z = ik\eta_0 H_x, \quad \partial H_x/\partial z = (ikq^2/\eta_0)E_y,$$
$$\eta_0 \equiv (\mu_0/\epsilon_0)^{\frac{1}{2}}, \qquad (2.2)$$

in which an $\exp\{i\omega t\}$ time dependence is assumed and q is related to the refractive index n through the following equation:

$$q^2 = n^2 - S^2 = 1 - X - S^2 \equiv C^2 - X \quad (2.3)$$

in which

$$X = (\omega_N/\omega)^2 \qquad (2.4)$$

is proportional to the electron density and ω_N is the angular plasma frequency. For simplicity, collisions have been neglected. Eliminating H_x and omitting the common factor $\exp\{-ikSx\}$ in (2.2), the following differential equation for E_y is derived:

$$L(E) \equiv (d^2E_y/dz^2) + k^2 q^2 E_y = 0. \qquad (2.5)$$

The above equation will be treated as a special case of (2.1), and useful physical interpretations of the derived solutions are discussed. The problem of solving (2.1) is frequently encountered in the field of quantum mechanics.

3. TRANSFORMATION OF THE SECOND-ORDER DIFFERENTIAL EQUATION INTO TWO COUPLED FIRST-ORDER DIFFERENTIAL EQUATIONS ·

It has been seen in the previous section how the second-order differential equation (2.5) is derived from two first-order differential equations (2.2); similarly, (2.1) may be represented by the following two equations:

$$\frac{d\Phi}{du} = ik\eta_0\Psi, \quad \frac{1}{w}\frac{d}{du}(p\Psi) = \frac{ik}{\eta_0}q^2\Phi, \qquad (3.1)$$

in which Φ and the above-defined function Ψ correspond to the electric and magnetic field components, respectively. In view of the fact that Φ (and Ψ) are solutions of second-order differential equations, it is

possible to express these functions as follows:

$$\Phi = \Phi_1 + \Phi_2 \quad \text{and} \quad \Psi = (ik\eta_0)^{-1}(G_1\Phi_1 + G_2\Phi_2),$$
(3.2)

in which the independent functions Φ_1 and Φ_2 correspond to the forward and backward propagating waves in homogeneous media, and G_1 and G_2 are two unequal functions that may be chosen arbitrarily. Substituting (3.2) into (3.1), the following equations are obtained:

$$\Phi_1' + \Phi_2' = G_1\Phi_1 + G_2\Phi_2,$$
(3.3)

$$G_1 p\Phi_1' + G_2 p\Phi_2' = -wk^2q^2\Phi_1 - wk^2q^2\Phi_2 - G_1'p\Phi_1$$
$$- G_2'p\Phi_2 - G_1 p'\Phi_1 - G_2 p'\Phi_2, \quad (3.4)$$

in which the primes denote total derivatives with respect to u. Multiply (3.4) by $1/pG_2$ and subtract it from (3.3) to get

$$\Phi_1'\left(1 - \frac{G_1}{G_2}\right) - \left(G_1 + \frac{wk^2q^2}{pG_2} + \frac{G_1'}{G_2} + \frac{G_1p'}{G_2p}\right)\Phi_1$$
$$= \left(G_2 + \frac{wk^2q^2}{pG_2} + \frac{G_2'}{G_2} + \frac{p'}{p}\right)\Phi_2. \quad (3.5)$$

An expression involving Φ_2' may be derived in a similar manner. In the above equation the coefficient of Φ_2 represents the coupling between the functions Φ_1 and Φ_2. Therefore, to facilitate the solution of the coupled first-order differential equation, it is desired to choose the auxiliary functions G_1 and G_2 such that the coupling coefficient is equal to zero if possible. But by setting the coupling coefficient equal to zero, a nonlinear differential equation is obtained. In order to transform it into a linear differential equation, the following substitution is made:

$$G_1 = (d/du)(\ln g_1) = g_1'/g_1,$$
$$G_2 = (d/du)(\ln g_2) = g_2'/g_2.$$
(3.6)

It will be seen later that, with the above substitution, G_1 and G_2 may be related to the "local" propagation coefficient of the two independent functions Φ_1 and Φ_2, yet to be determined. Substituting (3.6) into (3.5), the following equation is obtained:

$$\Phi_1' - \frac{g_1'}{g_1}\Phi_1 + \frac{g_2}{W(g_2, g_1)}\frac{w}{p}L(g_1)\Phi_1$$
$$= \frac{g_1}{W(g_1, g_2)}\frac{w}{p}L(g_2)\Phi_2. \quad (3.7a)$$

Similarly, by interchanging subscripts 1 and 2,

$$\Phi_2' - \frac{g_2'}{g_2}\Phi_2 + \frac{g_1}{W(g_1, g_2)}\frac{w}{p}L(g_2)\Phi_2$$
$$= \frac{g_2}{W(g_2, g_1)}\frac{w}{p}L(g_1)\Phi_1. \quad (3.7b)$$

In (3.7a) and (3.7b), the Wronskian W, defined as

$$W(g_1, g_2) = -W(g_2, g_1) = g_1g_2' - g_1'g_2, \quad (3.8)$$

does not vanish, since g_1 and g_2 are assumed linearly independent functions and L is the same differential operator defined in (2.1). Now define the coupling coefficients

$$C_{12} = -C_{22} = \frac{g_1}{W(g_1, g_2)}\frac{w}{p}L(g_2),$$
$$C_{21} = -C_{11} = \frac{g_2}{W(g_2, g_1)}\frac{w}{p}L(g_1),$$
(3.9a)

and the matrices

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix}, \quad \text{and} \quad G = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix}.$$
(3.9b)

Thus, in matrix notation,

$$\Phi' - G^{-1}G'\Phi = C\Phi. \quad (3.10)$$

Now, provided g_1 and g_2 are chosen such that the coupling coefficients vanish, the solutions of the above equations are simply

$$\Phi_{1,2} = g_{1,2}. \quad (3.11)$$

Obviously this is not the case of particular interest of this paper, for in order to make the coupling coefficients vanish, it is necessary to find two linearly independent solutions of $L(g) = 0$, which are exactly of the same form as (2.1). If such solutions for g do exist, there would be no point to the above transformation (3.7) of the original second-order differential equation (2.1). This paper considers the problem in which $L(g)$ is not equal to zero for any known function g.

Obviously, in order to facilitate the solution of the coupled equations (3.7), g_1 and g_2 must be so chosen that the coupling coefficients are much smaller than the logarithmetic derivatives of g_1 and g_2. Then proceed by solving (3.7), using an iterative method. In the first step solve (3.7) for Φ_1^0 and Φ_2^0, assuming the right-hand side of the equations equal to zero. In the next step substitute Φ_1^0 and Φ_2^0 for Φ_1 and Φ_2, respectively, in the terms on the right-hand side of (3.7), and proceed to solve the resulting first-order nonhomogeneous differential equation. Such a procedure can, of course, be repeated as many times as necessary or feasible. Physically, this iterative process corresponds to the consideration of successive multiple reflections in an inhomogeneous media; hence this method is not suitable if reflections are large.

Another method for solving the coupled equations would be as follows. After making the first obvious

choice for g_1^0 and g_2^0, solve (3.7) for Φ_1^0 and Φ_2^0 after neglecting the terms on the right-hand side as before. Since in each of the above solutions only the cross coupling terms have been neglected (the off-diagonal terms of the matrix C), it is possible that Φ_1^0 and Φ_2^0 are better solutions to (2.1) than are g_1^0 and g_2^0. This can be checked readily in any particular problem by evaluating the new coupling coefficients derived from (3.9a) on substituting $g_1^1 = \Phi_1^0$ and $g_2^1 = \Phi_2^0$. This procedure, which may also be repeated, is employed in the following sections, and the general formulation of this solution is expressed as an infinite product (Sec. 7).

In the next section several procedures for solving (3.7) are discussed in detail for the particular case in which L is the differential operator given in (2.5). In the case of the more general second-order differential equation (2.1), the discussion follows in precisely the same manner.

4. SOLUTION DERIVED IN TERMS OF COUPLED PLANE WAVES—WKB-TYPE SOLUTIONS

In general, one may express the function q^2 in its Taylor series about its value at $z = 0$ (or some other convenient point):

$$q^2(z) = n^2 - S^2 = 1 - S^2 + a_1 z + a_2 z^2 + \cdots$$
$$\equiv C^2 + a_1 z + a_2 z^2 + a_3 z^3, \quad (4.1a)$$

in which a_n are the familiar coefficients of the Taylor series expansion. Consider at first the case in which

$$q^2(z) = C^2 + t(z) \equiv q_0^2 + t, \quad (4.1b)$$

in which $|t/q_0^2| \ll 1$ for all values of z. Note that at the level $z = 0$, n is normalized to be equal to unity, in which case the wavenumber and the sine of the angle of incidence at the level $z = 0$ are k and S, respectively. In this case assume

$$g_{1,2}^0 = \exp\{\mp ikq_0 z\}. \quad (4.2)$$

Hence

$$L(g_{1,2}^0) = k^2 t g_{1,2}, \quad W(g_1^0, g_2^0) = 2ikq_0 g_1^0 g_2^0, \quad (4.3)$$

$$\Phi_{1,2}' \pm ikq_0\left(1 + \frac{t}{2q_0^2}\right)\Phi_{1,2} = \pm \frac{kt}{2iq_0}\Phi_{2,1}. \quad (4.4)$$

The solution of the homogeneous equation is

$$\Phi_{1,2}^0 = \exp\left\{\mp ik\int^z q_0\left(1 + \frac{t}{2q_0^2}\right) dz\right\}$$
$$\approx \exp\left\{\mp ik\int^z q\, dz\right\}. \quad (4.5)$$

The above solution constitutes the well-known phase memory concept. Now $L(\Phi_{1,2}^0) = \mp ik(dq/dz)\Phi_{1,2}^0$; hence this solution is a good approximation, provided that the condition

$$\left|k\frac{dq}{dz}\right| \ll |k^2 q^2| \quad \text{or} \quad \left|\frac{1}{q^2}\frac{dq}{dz}\right| \ll k$$

is satisfied in addition to the above requirement that $|t/q_0^2| \ll 1$ for all values of z.

Using the above approximations for Φ_1 and Φ_2, higher-order approximations may be derived by substituting them into the right-hand sides of (3.7) and solving the nonhomogeneous equation. Alternately, the above solution of (3.7) can be repeated, except that this time let $g_{1,2}^1 = \Phi_{1,2}^0$. Hence,

$$L(g_{1,2}^1) = \mp ik(dq/dz)g_{1,2}^1, \quad W(g_1^1, g_2^1) = 2ikqg_1^1 g_2^1, \quad (4.6)$$

and

$$\Phi_{1,2}' \pm ikq\Phi_{1,2} + \frac{1}{2q}\frac{dq}{dz}\Phi_{1,2} = \frac{1}{2q}\frac{dq}{dz}\Phi_{2,1}. \quad (4.7)$$

The above equations may be readily recognized as the coupled WKB solutions for slowly varying media.[1] Again, on neglecting the right-hand side of the above equations, the approximate solutions obtained are

$$\Phi_{1,2}^1 = q^{-\frac{1}{2}}\exp\left\{\mp ik\int^z q\, dz\right\}. \quad (4.8)$$

These solutions are analogous to the upgoing and downgoing waves. In addition to the phase memory term, the above solutions contain a $q^{-\frac{1}{2}}$ factor that constitutes the condition that if q is real, the power carried by these waves is constant.

Substitution of the above solution into the differential equation (2.5) yields

$$L(\Phi_{1,2}^1) = (\tfrac{3}{4}(q'/q)^2 - \tfrac{1}{2}(q''/q))\Phi_{1,2}^1 \quad (4.9a)$$

and

$$W(\Phi_1^1, \Phi_2^1) = \pm 2ikq\Phi_1^1\Phi_2^1. \quad (4.9b)$$

Now the above solution is a good approximation, provided that $|L(\Phi_{1,2}^1)| \ll |k^2 q^2 \Phi_{1,2}^1|$, from which the following quantitative criteria for a satisfactory solution may be derived[1]:

$$|\tfrac{3}{4}(q'/q^2)^2 - (q''/2q^3)| \ll k^2. \quad (4.10)$$

The above technique may be pursued still further by assuming that $g_{1,2}$ is given by the approximate solution for $\Phi_{1,2}$ in (4.8). In this case the coupling coefficients are all proportional to

$$[\tfrac{3}{4}(q'/q)^2 - \tfrac{1}{2}(q''/q)]/kq. \quad (4.11)$$

---

[1] K. G. Budden, *Radio Waves in the Ionosphere* (Cambridge University Press, London, 1961).

Before proceeding any further with this method, it should be determined, for the particular refractive index profile considered, whether the subsequent coupling coefficients are decreasing. Alternatively, substituting the WKB solutions (4.8) for $g_{1,2}$ in (3.7) and proceeding through the first iterative method discussed in the preceding section, it is possible to derive a series solution in which the higher-order terms correspond to multiple reflections.[2-4] It is obvious that all the above solutions fail for regions in which q is very small or q' is very large. In these regions the reflection process is substantial, and therefore the coupling between the upgoing and downgoing WKB wave solutions is very large. Examination of (4.11) shows that, for the regions in which reflections are large, further iteration of the WKB solutions is of little value.

Before attempting to derive a different set of solutions, it should be pointed out that in this section the functions g_1 and g_2 are chosen to represent the "local" upgoing and downgoing waves that constitute the wave solution if at each level the refractive index is assumed constant. It is interesting to note that the solutions derived in (4.8) may be derived directly by considering the medium of propagation to consist of infinitesimally thin layers in each of which n^2 is considered constant. Let n_1 and n_2 be the refractive index in two such consecutive layers. Now the transmission coefficient for a horizontally polarized wave traveling from medium 1 to medium 2 is

$$T = \frac{2C_1 n_1}{C_1 n_1 + C_2 n_2} = \frac{2C_1 n_1}{C_1 n_1 + (n_2^2 - n_1^2 S_1^2)^{\frac{1}{2}}}, \quad (4.12)$$

in which C and S are the cosine and the sine of the angles of incidence in the medium indicated by the respective subscript. The differential transmission coefficient as $n_2 \to n_1 + \Delta n$ is given by

$$C_{11} = \frac{dT}{dz} = \frac{dT}{dn_2}\frac{dn_2}{dz}\bigg|_{n_2=n_1} = -\frac{n_1 n_1'}{2(C_1 n_1)^2}. \quad (4.13a)$$

Now

$$q_1^2 = n_1^2 - S^2 = n_1^2 - n_1 S_1^2 = n_1^2 C_1^2$$

and

$$q' = 2n_1 n_1'/2q_1.$$

Hence

$$C_{11} = -q_1'/2q_1. \quad (4.13b)$$

Similarly, the reflection coefficient for a wave incident from medium 2 is

$$R = (C_2 n_2 - C_1 n_1)/(C_2 n_2 + C_1 n_1), \quad (4.14)$$

and the differential reflection coefficient (corresponding to coupling from a downward to an upward traveling wave) is given as follows:

$$C_{12} = \frac{dR}{dn_2}\frac{dn_2}{dz}\bigg|_{n_2=n_1} = \frac{n_1 n_1'}{2(C_1 n_1)^2} = \frac{q_1'}{2q_1}. \quad (4.15)$$

The other two coupling coefficients may be derived in a similar manner to obtain precisely the coupled equations (4.7) which yield the WKB solutions. The second limiting quantity C_{12} is defined by Bremmer[2] as the reflection coefficient per unit distance.

In the next section, a generalized WKB solution is derived in which the "local" wave solutions g_1 and g_2 are not those corresponding to homogeneous media. The local wave solutions will be chosen such that they exist in a medium with a linearly varying dielectric coefficient. The refractive index n and its gradient n' at every level will determine the "constant" parameters of the local wave solutions. This is equivalent to considering the medium to consist of infinitesimally thin layers, in each of which the dielectric coefficient varies linearly (rather than remaining constant). Hence the dielectric coefficient for the local wave solutions is given by straight lines tangent to the given dielectric coefficient profile. It will be seen that in this case the coupling between these local wave solutions is very small for the regions in which the above WKB solutions fail.

5. GENERALIZED COUPLED WKB SOLUTIONS

Consider first the case in which the gradient of the dielectric coefficient is slowly varying; then the function q^2 may be expressed as

$$q^2 = C^2 + az + r(z), \quad a \equiv d(n^2)/dz, \quad (5.1)$$

and $|r'/a| \ll 1$ and $r \ll q^2$. The local wavefunctions are assumed to be a solution of

$$g_{1,2}'' + k^2(C^2 + az)g_{1,2} = 0. \quad (5.2)$$

Two linearly independent solutions are the Airy integral functions

$$g_1 = \text{Ai}(\zeta) \quad \text{and} \quad g_2 = \text{Bi}(\zeta), \quad (5.3)$$

in which $\zeta = -(k/a)^{\frac{2}{3}}(az + C^2)$ and the principal root is implied. Ai is the appropriate solution for a wave incident from below, and Bi is the appropriate solution for a wave incident from above. Obviously, for $r \neq 0$, there will be coupling between these two solutions. To determine these coupling terms, substitute g_1 and g_2 into (3.9a):

$$L(g_{1,2}) = k^2 r g_{1,2}, \quad W(g_1, g_2) = W(\text{Ai, Bi})(d\zeta/dz),$$

$$= -(k/a)^{\frac{2}{3}}a/\pi. \quad (5.4)$$

[2] H. Bremmer, Physica 15, 593 (1949).

[3] J. R. Wait, Electromagnetic Waves in Stratified Media (Pergamon Press, Inc., New York, 1962).

[4] D. S. Jones, The Theory of Electromagnetism (Pergamon Press, Inc., New York, 1964).

Hence

$$\Phi'_{1,2} - \left(\frac{g'_{1,2}}{g_{1,2}} \mp \frac{\pi r}{a} g_1 g_2 (k^2 a)^{\frac{2}{3}}\right)\Phi_{1,2}$$

$$= \mp \frac{\pi r}{a} g_1 g_2 (k^2 a)^{\frac{2}{3}} \Phi_{2,1}. \quad (5.5)$$

The solution of the above equation, on neglecting cross coupling, is

$$\Phi_{1,2} = g_{1,2} \exp\left\{\mp (\pi/a)(k^2 a)^{\frac{2}{3}} \int^z g_1 g_2 r \, dz\right\}. \quad (5.6)$$

The solution is satisfactory, provided that $L(\Phi_{1,2}) \ll k^2 q^2 \Phi_{1,2}$. Hence, neglecting the term proportional to $(r/a)^2$,

$$\left|2\pi \frac{g'_{1,2}}{g_{1,2}} \frac{r}{a} g_1 g_2 (k^2 a)^{\frac{2}{3}} + k^2 r + \pi \left(\frac{r}{a} g_1 g_2\right)'(k^2 a)^{\frac{2}{3}}\right|$$

$$\ll |k^2 q^2|. \quad (5.7)$$

The above solution is therefore only a partial improvement over the WKB solution in that it is a good solution if (dq/dz) is large, but it still fails if the curvature of q^2 is finite and q is very small. In order to remove this severe restriction, determine first the "phase memory" concept in terms of the Airy integral function.

Let f_1 and f_2 be the two linearly independent solutions that satisfy

$$(\partial^2 f_{1,2}/\partial z^2) + k^2 [C^2 + \alpha(z - z_0)]f_{1,2} = 0, \quad (5.8)$$

in which α and z_0 are functions of z:

$$\alpha = d(q^2)/dz = 2qq', \quad z_0 = z - [(q^2 - C^2)/\alpha]. \quad (5.9)$$

Obviously α and z_0 are defined such that

$$C^2 + \alpha(z - z_0) = q^2;$$

and if α^C and z_0^C are the respective values of α and z_0 at $z = z_C$, then $C^2 + \alpha^C(z - z_0^C)$ is the straight line tangent to the function q^2 at the point $z = z_C$. Furthermore, in (5.8) the symbol for the partial derivative is used to imply that here α and z_0 are considered independent variables (see Fig. 2). Hence f_1 and f_2 are the wave solutions for a medium in which q^2 is a linear function of z. The solutions of (5.8) are
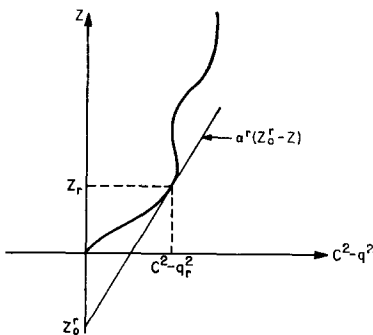

Fig. 2. A geometrical interpretation of the parameters α and z_0 defined in (5.9).

the Airy integral functions Ai (ζ) and Bi (ζ), in which

$$\zeta = -(k/\alpha)^{\frac{2}{3}}[\alpha(z - z_0) + C^2] = -(k/\alpha)^{\frac{2}{3}}q^2 \quad (5.10)$$

and $(k/\alpha)^{\frac{2}{3}} = |(k/\alpha)|^{\frac{2}{3}} \exp \frac{1}{3}i(\arg (k/\alpha)^2)$, implying the principal root. Now the local propagation coefficients in this case are $(\partial/\partial z)$ ln Ai and $(\partial/\partial z)$ ln Bi. Hence the corresponding solutions that constitute the phase memory concept are

$$g_1 = \exp\left\{\int^z \frac{\partial}{\partial z}(\ln \text{Ai}) \, dz\right\}$$

and
$$\quad (5.11)$$

$$g_2 = \exp\left\{\int^z \frac{\partial}{\partial z}(\ln \text{Bi}) \, dz\right\},$$

in which the lower limits of the integrals are arbitrary constants and therefore not specified. Substitute the above functions into (3.9a) to derive the coupling coefficients. Now

$$\frac{1}{g_1}\frac{dg_1}{dz} = \frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z},$$

$$\frac{1}{g_1}\frac{d^2 g_1}{dz^2} = \left(\frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z}\right)^2 + \frac{\left[\frac{d}{d\zeta}\left(\frac{\partial \text{Ai}}{\partial z}\right)\text{Ai} - \frac{d\text{Ai}}{d\zeta}\frac{\partial \text{Ai}}{\partial z}\right]\frac{d\zeta}{dz}}{(\text{Ai})^2} ,$$

and

$$\frac{d}{d\zeta} \rightarrow \frac{1}{-\alpha(k/\alpha)}\frac{\partial}{\partial z}, \quad \frac{d\zeta}{dz} = -\alpha\left(\frac{k}{\alpha}\right)^{\frac{2}{3}}\left[1 - \frac{2}{3\alpha^2}q^2\alpha'\right].$$

Therefore

$$\frac{1}{g_1}\frac{d^2 g_1}{dz^2} = \left(\frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z}\right)^2 + \left[\frac{\partial^2 \text{Ai}/dz^2}{\text{Ai}} - \left(\frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z}\right)^2\right]$$

$$\times \left[1 - \frac{2q^2}{3\alpha^2}\alpha'\right] \quad (5.12)$$

and

$$\frac{1}{g_1}L(g_1) = \frac{2}{3}\left[k^2 q^2 + \left(\frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z}\right)^2\right]\frac{q^2}{\alpha^2}\alpha',$$

$$W_z(g_1, g_2) = \frac{g_1 g_2}{\text{Ai Bi}} W_\zeta(\text{Ai Bi})\frac{\partial \zeta}{\partial z} = -\frac{1}{\pi}\frac{g_1 g_2}{\text{Ai Bi}}\alpha\left(\frac{k}{\alpha}\right)^{\frac{2}{3}},$$

in which the subscripts of W refer to the implied variable of differentiation and $W_\zeta(\text{Ai, Bi}) = 1/\pi$. Therefore

$$C_{11} = -\frac{2}{3}\pi \, \text{Ai Bi} \left(\frac{\alpha}{k}\right)^{\frac{2}{3}}\left[k^2 q^2 + \left(\frac{1}{\text{Ai}}\frac{\partial \text{Ai}}{\partial z}\right)^2\right]\frac{q^2}{\alpha^3}\alpha'.$$

$$(5.13)$$

Now

$$\frac{\partial \text{Ai}}{\partial z} = -\alpha\left(\frac{k}{\alpha}\right)^{\frac{2}{3}}\frac{d \text{Ai}}{d\zeta} = -\alpha\left(\frac{k}{\alpha}\right)^{\frac{2}{3}}\text{Ai}'$$

and

$$\alpha' = 2qq'' + 2(q')^2.$$

Hence

$$C_{11} = -\frac{2}{3}\pi \, \text{Ai Bi} \left(\frac{\alpha}{k}\right)^{\frac{2}{3}}\left[k^2 q^2 + \alpha^2\left(\frac{k}{\alpha}\right)^{\frac{4}{3}}(\text{Ai}'/\text{Ai})^2\right]\frac{q^2}{\alpha^3}\alpha'.$$

$$(5.14)$$

Similarly, interchanging Ai with Bi, it is possible to determine C_{22} directly:

$$C_{22} = -\tfrac{2}{3}\pi \, \text{Ai Bi} \left(\frac{\alpha}{k}\right)^{\tfrac{2}{3}} \left[k^2 q^2 + \alpha^2 \left(\frac{k}{\alpha}\right)^{\tfrac{4}{3}} \left(\frac{\text{Bi}'}{\text{Bi}}\right)^2 \right] \frac{q^2}{\alpha^3} \, \alpha'.$$

$$(5.15)$$

From the above expressions it is obvious that, for $q \to 0$ and $q' \to \infty$, the coupling coefficients are actually zero. Hence, for the region in which the upward and downward going WKB solutions are highly coupled, the above generalized solutions are actually very slightly coupled. Therefore the generalized solutions are particularly adapted for highly reflecting regions in which the regular WKB solutions fail. The following equation, in which cross coupling is neglected, may now be solved for Φ:

$$\Phi_1' - \left(\frac{1}{\text{Ai}} \frac{\partial \, \text{Ai}}{\partial z} + C_{11}\right)\Phi_1 = 0. \quad (5.16)$$

Therefore

$$\Phi_1 = \exp \left\{ \int^z \frac{1}{\text{Ai}} \frac{\partial \, \text{Ai}}{\partial z} \, dz + \int^z C_{11} \, dz \right\},$$

$$= g_1 \exp \int^z C_{11} \, dz, \quad (5.17a)$$

in which g_1 may also be expressed as

$$g_1 = \exp \int^z \frac{d}{d\zeta} (\ln \text{Ai})\left(\frac{\partial \zeta}{\partial z}\frac{dz}{d\zeta}\right) d\zeta, \quad (5.17b)$$

from which it is obvious that if $(d\zeta/dz) = (\partial\zeta/\partial z)$ (α and z_0 are constants), $g_1 = \text{Ai}$. It should be pointed out that this form of the solution is not appropriate for regions in which $q' \to 0$, for in this case the argument of the Airy integral functions approaches infinity. In this region q is very slowly varying, and the generalized WKB solution merges with the ordinary WKB solution. From the above remarks it is then clear that the two solutions (4.8) and (5.17) are complimentary solutions. The first is good for a slowly varying medium and fails in the reflection regions. The second is particularly good in the reflection regions, but is less appropriate in a very slowly varying medium. These comments are borne out by the "illustrative" example of Sec. 6.

Before considering certain generalizations to the above problem, it is interesting to demonstrate how the coupling coefficients (5.14) and (5.15) may be derived directly by considering the corresponding transmission and reflection coefficients at an interface between two media (at $z = z_r$) in each of which the dielectric coefficient varies linearly with height (see Fig. 3). Assume an Airy integral wavefunction incident
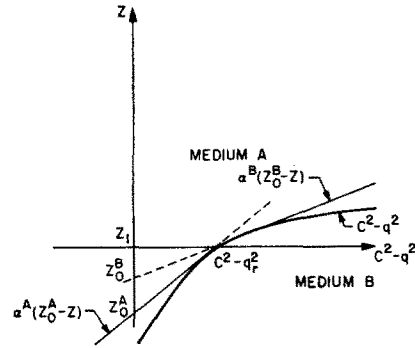

FIG. 3. Substitution of the q^2 profile with straight lines tangent to it at every point.

from medium A. Then, for medium A,

$$E^A = a^A[\text{Ai}\,(\zeta^A)/\text{Ai}\,(\zeta_r^A)] + b^A[\text{Bi}\,(\zeta^A)/\text{Bi}\,(\zeta_r^A)]$$

$$(5.18a)$$

and, for medium B,

$$E^B = a^B[\text{Ai}\,(\zeta^B)/\text{Ai}\,(\zeta_r^B)], \quad (5.18b)$$

in which

$$\zeta^A = -(k/\alpha^A)^{\tfrac{2}{3}}[\alpha^A(z - z_0^A) + C^2] \equiv -(k/\alpha^A)^{\tfrac{2}{3}}(q^A)^2$$

$$(5.18c)$$

with similar expressions for ζ^B, and at the interface $z = z_r$, $\zeta^A = \zeta_r^A$, $\zeta^B = \zeta_r^B$, $q^A = q^B = q(z_r) = q_r$. The continuity of the electric and magnetic field at the interface ($z = z_r$) is given by the following equations:

$$a^A + b^A = a^B,$$

$$\frac{a^A}{\text{Ai}\,(\zeta^A)} \frac{\partial \, \text{Ai}\,(\zeta^A)}{\partial z} + \frac{b^A}{\text{Bi}\,(\zeta^A)} \frac{\partial \, \text{Bi}\,(\zeta^A)}{\partial z}$$

$$= \frac{a^B}{\text{Ai}\,(\zeta^B)} \frac{\partial \, \text{Ai}\,(\zeta^B)}{\partial z}, \quad (z = z_r), \quad (5.19)$$

from which one can determine the transmission coefficient T:

$$T = \frac{a^B}{a^A} = -\left(\frac{\partial \zeta}{\partial z}\right) \frac{W_{\zeta^A}(\text{Ai, Bi})}{\text{Ai}\,(\zeta^A)\,\text{Bi}\,(\zeta^A)}$$

$$\times \left[\left(\frac{1}{\text{Ai}\,(\zeta^B)} \frac{\partial \, \text{Ai}\,(\zeta^B)}{\partial z} - \frac{1}{\text{Bi}\,(\zeta^A)} \frac{\partial \, \text{Bi}\,(\zeta^A)}{\partial z} \right) \right]^{-1},$$

$$(z = z_r). \quad (5.20)$$

To determine the differential transmission coefficient, let $\alpha^B \to \alpha^A + \Delta\alpha$, $z_0^B \to z_0^A + \Delta z_0^A$. Then

$$C_{11} = \frac{dT}{d\alpha^B} \frac{d\alpha^B}{dz} + \frac{dT}{dz_0^B} \frac{dz_0^B}{dz}\bigg|_{\substack{\alpha^B = \alpha^A \\ z_0^B = z_0^A}},$$

$$= \frac{\partial T}{\partial \zeta_r^B} \frac{d\zeta_r^B}{dz}\bigg|_{\zeta_r^B \to \zeta_r^A \to \zeta}. \quad (5.21a)$$

Now, since ζ_r^B is a function of α^B and z_0^B, but not a

function of z,

$$\frac{d\zeta_r^{\rm B}}{dz} = \left(\frac{d\zeta^{\rm B}}{dz} - \frac{\partial\zeta^{\rm B}}{\partial z}\right)\Big|_{\zeta_r^{\rm B}} = \frac{2}{3\alpha^{\rm B}}\left(\frac{k}{\alpha^{\rm B}}\right)^{\frac{2}{3}} q^2(\alpha^{\rm B})'\Big|_{z_r}.$$
(5.21b)

Hence

$$C_{11} = \frac{{\rm Ai}(\zeta)\,{\rm Bi}(\zeta)}{(\partial\zeta/\partial z)W_\zeta({\rm Ai},{\rm Bi})}\left[\frac{1}{\rm Ai}\frac{d}{d\zeta}\frac{\partial\,{\rm Ai}}{\partial z}\right.$$
$$\left. - \frac{1}{({\rm Ai})^2}\frac{\partial\,{\rm Ai}}{\partial z}\frac{d\,{\rm Ai}}{d\zeta}\right]\frac{d\zeta_r}{dz}, \quad z = z_r, \quad (5.22{\rm a})$$

in which the superscripts have been dropped, since $\zeta^{\rm A} \to \zeta^{\rm B}$. Finally, since z_r is an arbitrary reference level, the subscript r can be dropped:

$$C_{11} = -\tfrac{2}{3}\pi\,{\rm Ai}\,(\zeta)\,{\rm Bi}\,(\zeta)\left(\frac{\alpha}{k}\right)^{\frac{2}{3}}$$
$$\times\left[k^2q^2 + \alpha^2\left(\frac{k}{\alpha}\right)^{\frac{4}{3}}\left(\frac{1}{\rm Ai}\frac{d\,{\rm Ai}}{d\zeta}\right)^2\right]\frac{q^2\alpha'}{\alpha^3}, \quad (5.22{\rm b})$$

which is precisely what is derived in (5.14). The derivation of C_{12} follows as in Sec. 4.

6. ILLUSTRATION OF THE GENERALIZED WKB METHOD

As an example of a dielectric coefficient profile with a finite gradient and curvature, consider a lossless nonmagnetized plasma with an exponentially varying electron density. Then

$$q^2 = n^2 - S^2 = C^2 - X = C^2 - Ke^{\beta(z-z_0)},$$
$$= C^2(1 - e^{\beta z}), \quad (6.1)$$

in which K, α, and z_0 are constants and, for convenience, the origin of the height z is chosen where $q^2 = 0$ (see Fig. 4). The differential equation that E_y satisfies is, therefore,

$$(d^2E_y/dz^2) + k^2C^2(1 - e^{\beta z})E_y = 0. \quad (6.2)$$

There exists an exact solution for this equation in terms of the Hankel function. For a wave incident from below, the solution is[1]

$$E_y = H_\nu^{(1)}(\nu u), \quad (6.3)$$

in which $\nu = 2ikC/\beta$, $u = e^{\frac{1}{2}\beta z}$, and β is assumed positive. We now seek the behavior of this solution in the region $|\beta z| < 1$. Using the uniform asymptotic expansion for $|\nu| > 1$ and u in the region around unity, one gets[5]

$$H_\nu^{(1)}(\nu u) \sim D\,{\rm Ai}\,[\exp{(2\pi i/3)}\nu^{\frac{2}{3}}\xi], \quad (6.4)$$

[5] M. Abramowitz and I. A. Stegum, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Department of Commerce, National Bureau of Standards, Washington, D.C., June 1964), Applied Mathematics Series 55.
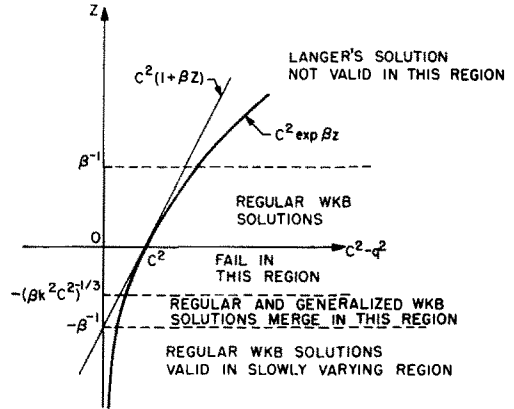

FIG. 4. Illustrative example of the use of the generalized WKB method with $q^2 = C^2[1 - \exp{(\beta z)}]$.

where

$$D = \frac{2^{\frac{1}{3}}\exp{(-\pi i/3)}}{\nu^{\frac{1}{3}}},$$

$$\tfrac{2}{3}\xi^{\frac{3}{2}} = \ln\frac{1 + (1 - u^2)^{\frac{1}{2}}}{u} - (1 - u^2)^{\frac{1}{2}}.$$

Now the expansion for ξ in the region $|\beta z| < 1$, $[u \sim 1 + \frac{1}{2}(\beta z)]$, is

$$\xi \sim (\tfrac{1}{2})^{\frac{2}{3}}(-\beta z),$$

and

$$\exp{(2\pi i/3)}\nu^{\frac{2}{3}}\xi = i^{\frac{4}{3}}(2ikC/\beta)^{\frac{2}{3}}(\tfrac{1}{2})^{\frac{2}{3}}(-\beta z) = (k^2C^2\beta)^{\frac{1}{3}}z.$$

Hence

$$H_\nu^{(1)}(\nu e^{\frac{1}{2}\beta z}) \sim {\rm Ai}\,[(k^2C^2\beta)^{\frac{1}{3}}z]. \quad (6.5)$$

Now in the region considered here, $q^2 \sim -C^2\beta z$ and $(q^2)' \equiv \alpha = -C^2\beta$. The "local" wave solution of the Airy integral function type is, by definition [(5.8) to (5.10)], as follows:

$${\rm Ai}\left[-\left(\frac{k}{\beta C^2}\right)^{\frac{2}{3}}(-C^2\beta z)\right] = {\rm Ai}\,[(k^2C^2\beta)^{\frac{1}{3}}z] = g_1. \quad (6.6)$$

Now, since for small z, $(\partial g_1/\partial z) \approx (dg_1/dz)$, the solution for (6.2), neglecting coupling, is $E_y \approx g_1 = {\rm Ai}\,[(k^2C^2\beta)^{\frac{1}{3}}z]$, which is precisely the uniform asymptotic expansion of the exact solution. Hence the above Airy integral function adequately describes the "local" wave solution in a region of reflection $(q^2 = 0)$ in which the curvature is finite. This is, of course, not surprising, since the coupling coefficients actually vanish at the level for which $q^a = 0$ [(5.14) and (5.15)]. The above solution corresponds to a wave incident from below. Now for $z > 0$, the Airy integral function behaves as an evanescent wave, as should be the case, since $z > 0$ corresponds to the region above the reflection layer.

Consider now the generalized solution for the region $\beta z \gg 1$. Here

$$q^2 \approx -C^2e^{\beta z} \quad \alpha \equiv (q^2)' = -C^2\beta e^{\beta z}. \quad (6.7)$$

The "local" wave solution in this case is given by Ai (ζ), in which

$$\zeta = -(k/\alpha)^{\frac{2}{3}}[\alpha(z - z_0) - C^2] = -(k/\alpha)^{\frac{2}{3}}q^2$$
$$\approx (k/\alpha)^{\frac{2}{3}}C^2 e^{\beta z}. \quad (6.8)$$

Neglecting coupling, it can be shown (Appendix A) that the generalized solution for E_y is (A5)

$$E_y \sim \exp\left(-\tfrac{1}{4}\beta z\right) \exp\left[-(2kC/\beta)e^{\beta z/2}\right]. \quad (6.9)$$

Now the asymptotic expansion of the exact wave solution $H_\nu(v)$ for $v \gg 1$ and $v \gg |\nu|$ is

$$H_\nu^{(1)}(v) \sim (2/\pi v)^{\frac{1}{2}} \exp\left\{iv - \tfrac{1}{2}\nu\pi - \tfrac{1}{2}\pi\right\}, \quad (6.10)$$

in which $v = \nu e^{\frac{1}{2}\beta z} = (i2kC/\beta)e^{\frac{1}{2}\beta z}$. Hence the solution given in (6.9) is equal to the asymptotic expansion of the exact solution for large z.

Finally, consider the case in which $\beta < kC$ and $1 > -\beta z > (\beta/kC)^{\frac{2}{3}}$. It is seen that this corresponds to a region just below the reflection layer. In this region the solution for the electric field may be expressed in terms of the asymptotic expansion of the Airy integral function of negative arguments:

$$E_y \sim \zeta^{-\frac{1}{4}}[\exp\left(-\tfrac{2}{3}\zeta^{\frac{3}{2}}\right) + i\exp\left(\tfrac{2}{3}\zeta^{\frac{3}{2}}\right)], \quad (6.11)$$

in which [(5.10) and (6.6)]

$$\zeta \sim -(k/C^2\beta)^{\frac{2}{3}}q^2 \sim (k^2C^2\beta)^{\frac{1}{3}}z. \quad (6.12)$$

Now in this region

$$ik\int_0^z q\, dz \sim \tfrac{2}{3}\zeta^{\frac{3}{2}} \quad (6.13)$$

and $\zeta^{-\frac{1}{4}}$ is proportional to $q^{-\frac{1}{2}}$; hence the solution may be written as

$$E_y \sim kq^{-\frac{1}{2}}\left[\exp\left(-ik\int_0^z q\, dz\right) + i\exp\left(ik\int_0^z q\, dz\right)\right]. \quad (6.14)$$

Also, for the region under consideration,

$$\frac{1}{k^2}\left|\frac{1}{3}\left(\frac{1}{q^2}\frac{dq}{dz}\right)^2 - \frac{1}{2}\frac{1}{q^3}\frac{d^2q}{dz}\right| < \frac{5}{16}. \quad (6.15)$$

Hence, in view of condition (4.10), it is not surprising that in this region the generalized solution reduces to the regular WKB solution discussed in Sec. 3. The reflection coefficient at a reference level $z = -h$ anywhere below the reflection layer [$\beta h > (\beta/kC)^{\frac{2}{3}}$] can be computed readily from (6.14)[1]:

$$R = i\exp\left(-2ik\int_{-h}^0 q\, dz\right). \quad (6.16)$$

It is interesting to note that, for the case in which q^2 contains higher-order terms in z, for example,

$$q^2 = a_1 z + a_2 z^2 + \cdots. \quad (6.17)$$

Langer has used the following argument η for the

Airy integral function[1–3,6,7]:

$$\eta = \left[\frac{3}{2}i\left(\frac{k}{a_1}\right)^{\frac{1}{3}}\int_0^{\zeta_1} q\, d\zeta_1\right]^{\frac{2}{3}}, \quad (6.18)$$

in which[8]

$$\zeta_1 = -(k/a_1)^{\frac{2}{3}}a_1 z \quad (6.19)$$

such that if the higher-order terms in z are neglected, $\eta \equiv \zeta_1$.

This new variable η is introduced in order to overcome the difficulty that the asymptotic expansion of Ai (ζ_1) does not reduce to the WKB phase integral solutions (4.8) and (6.14) for the cases in which the curvature is not negligible. The asymptotic expansion of Ai (η) obviously does contain the appropriate phase integral term; however, it is not identically the WKB solution, since its coefficient is not $q^{-\frac{1}{2}}$ (except when the curvature is neglected).

Also, it can be shown that $\eta = \zeta + 0(\zeta^2)$. Hence the solution Ai (η) must be restricted to the region $|\eta - \zeta| < |\zeta|$. Indeed, it can be readily shown that Langer's solution is not valid for the case of the exponential profile (6.1) in the region $\beta z \gg 1$. The asymptotic solution for Ai (η) in this case is [Appendix (B2)],

$$\text{Ai}(\eta) \sim \exp\left(-\beta z/12\right)\exp\left[-(2kC/\beta)e^{\beta z/2}\right], \quad \beta z \gg 1. \quad (6.20)$$

In order to fully comprehend the failure of the solution Ai (η) (which strongly resembles the WKB solution), it is necessary to evaluate the coupling coefficient C_{11} for the case in which the auxiliary functions $g_{1,2}$ are chosen to be Ai (η) and Bi (η), respectively. It can be shown that (B6)

$$C_{11} = \frac{\pi}{2}\text{Ai}'(\eta)\text{Bi}(\eta)\frac{d}{dz}\ln\frac{q^2}{\eta},$$
$$= \frac{\pi}{2}\text{Ai}'(\eta)\text{Bi}(\eta)\left[\frac{(q^2)'}{q^2} - \frac{\eta'}{\eta}\right], \quad (6.21)$$

and it is obvious that if curvature is neglected, $C_{11} = 0$. Note that for the case of the exponential profile, $C_{11} = -(\beta/5)$ at $z = 0$ (B8), whereas it vanishes if the generalized WKB solution is used. Furthermore, if the gradient of the profile $(q^2)'$ is large, C_{11} is also large. (Recall C_{11} vanishes in this case if the generalized WKB solution is used.) In particular, for the exponential profile, $C_{11} = -\beta/6$ for $\beta z \gg 1$ (B9). Hence, in this case, the coupling coefficient may not be neglected if Langer's solution is used (except in the region $|\beta z| \ll 1$). Indeed, if the more accurate solution

[6] R. E. Langer, Phys. Rev. **51**, 669 (1937).

[7] C. L. Pekeris, J. Acoust. Soc. Am. **18**, 295 (1946).

[8] There is an obvious error in Ref. 1, in which k appears instead of the dimensionless quantity $(k/a_1)^{\frac{1}{3}}$.

of the coupled equations is used (5.17a), the solution is

$$E_y = g_1 \exp \int^z C_{11}\, dz = \mathrm{Ai}\,(\eta) \exp\left(-\frac{\beta z}{6}\right)$$

$$\sim \exp\left(-\frac{\beta z}{4}\right) \exp\left[-\frac{2kC}{\beta}\, e^{\beta z/2}\right]. \qquad (6.22)$$

The correction factor to Langer's solution $(\exp -\beta z/6)$ is, therefore, sufficient to render the correct asymptotic solution (6.10) to this problem. The above example not only emphasizes that Langer's solution must be limited to the transition region $|\eta - \zeta_1| < |\zeta_1|$, but also points out that even in the transition region Langer's solution is restricted by the requirement that q^2 have a small curvature.[3] Indeed, it is shown (for the particular example discussed in this section) that the coupling coefficient is larger in the transition region $(\beta z \ll 1)$ than outside the transition region where it is shown to fail.

7. DECOUPLED GENERALIZED WKB SOLUTIONS

A straightforward generalization of the decoupled solution (5.17) can now be derived through an iterative process. Let $g_{1,2}^0$ be the zero-order solutions for $\Phi_{1,2}$ [such as (5.11)]; then C_{11}^0 is the coupling coefficient corresponding to the zero-order solution. Thus (3.9a)

$$C_{11}^0 = \frac{w}{p}\, \frac{L(g_1^0)g_2^0}{W(g_1^0,\, g_2^0)}. \qquad (7.1)$$

Then let $g_{1,2}^1$ be the first-order solution given by (5.17a). Thus

$$g_{1,2}^1 = g_{1,2}^0 \exp \int^z C_{11}^0\, dz, \qquad (7.2)$$

and the coupling coefficient corresponding to the first-order solution is

$$C_{11}^1 = \frac{w}{p}\, \frac{L(g_1^1)g_2^1}{W(g_1^1,\, g_2^1)}. \qquad (7.3)$$

In general then, a higher-order solution is given by

$$g_{1,2}^{n+1} = g_{1,2}^n \exp \int^z C_{11}^n\, dz,$$

where C_{11}^n is the coupling coefficient for the nth-order solution

$$C_{11}^n = \frac{w}{p}\, \frac{L(g_1^n)g_2^n}{W(g_1^n,\, g_2^n)}. \qquad (7.4)$$

Hence the general solution may be written in terms of the infinite product

$$\Phi_1 = g_1^0 \prod_{n=0}^{\infty} \exp\left(\int^z C_{11}^n\, dz\right),$$

$$= g_1^0 \exp\left[\sum_{n=0}^{\infty} \int^z C_{11}^n\, dz\right],$$

$$= \lim_{n \to \infty} g_1^n, \qquad (7.5)$$

with a similar expression for Φ_2. Obviously the feasibility of the above solution depends upon the convergence of the series in the argument of the exponential function. It has been pointed out that this depends strongly upon the choice of the zero-order solution $g_{1,2}^0$. Particular attention has been paid to the suitable choice for $g_{1,2}^0$ for the case of the second-order differential equation (2.5). The more general second-order differential equation (2.1) can often be reduced to the form of (2.5), as for the cases in which p and w are polynomials of the independent variable u (e.g., Bessel's functions). If this is not feasible, a direct approach to the solution of the coupled equations (3.7) should be made in the same manner as in Secs. 4 and 5.

As specific applications to (7.5), note that in Sec. 4 it was shown that if $g_{1,2}^0$ are chosen to be the homogeneous plane wave solutions $\exp\{\mp ikq_0z\}$ (4.2), the first-order solutions $g_{1,2}^1$ constitute the well-known phase memory concept (4.5), and the second-order solutions are the regular WKB solutions (4.8). In the previous section it is shown that if g_1^0 is assumed to be Langer's solution, g_1^1 corresponds to the correct asymptotic expansion (6.9) derived from the exact solution (6.3) [or directly from the generalized WKB solution (5.11)].

A brief comparison between the two iterative processes described in Sec. 3 is now made. In the series expansion, the coupling coefficients (corresponding to differential transmission and reflection coefficients) are the same for each successive iteration. Hence the individual terms lead to simple geometric-optical approximations.[2] In the infinite product expansion, the coupling coefficients C_{ii}^n (only diagonal terms needed) in each successive iteration decrease. Thus each set of successive iterations $g_{1,2}^n$ is less coupled and better approximates the full wave solutions.

8. RELATIONSHIP BETWEEN THE REGULAR AND GENERALIZED WKB SOLUTIONS AND SOLUTIONS OF THE RICCATI EQUATION

Now the regular WKB solution is related to the solution of the well-known Riccati nonlinear first-order differential equation of the form[1-4,9]

$$\varphi' + \varphi^2 + k^2q^2 = 0. \qquad (8.1)$$

Hence it is rather interesting to compare the corresponding relationships between the regular and generalized WKB solutions and the solution of the Riccati equation.

The regular WKB solution is based on the following

---

[9] H. Bremmer, *Terrestrial Radio Waves* (Elsevier Publishing Co., Amsterdam, 1949).

approximate solution for φ:

$$\varphi \sim \pm ikq - \tfrac{1}{2}(q'/q) = \pm ikq - \tfrac{1}{4}(d/dz) \ln q^2. \quad (8.2)$$

Obviously this solution is based on the assumption that φ' is much smaller than k^2q^2, which is appropriate for a slowly varying q profile with no critical coupling points. In this case the dependent variable Φ is derived through the relationship

$$\Phi = \exp \int^z \varphi \, dz = (q)^{-\frac{1}{2}} \exp \pm ik \int^z q \, dz. \quad (8.3)$$

However, using the generalized WKB method for a q profile that may contain critical coupling points, the corresponding solution for the Riccati equation, which is apparently new, is

$$\varphi \sim (\partial/\partial z)\big(\ln \mathrm{Ai}\,(\zeta)\big), \quad (8.4)$$

in which ζ, given by (5.10), is a function of z and the "local" parameters (α and z_0) of the q profile (5.9). Now using the above solution (8.4) and (5.12), it can be shown that

$$\varphi' = -(\varphi^2 + k^2q^2) \frac{d\zeta}{dz} \Big/ \frac{\partial\zeta}{\partial z},$$

$$= -(\varphi^2 + k^2q^2)\big(1 - (2q^2/3\alpha^2)\alpha'\big). \quad (8.5)$$

Hence, obviously, the solution derived here for the Riccati equation is restricted only by the condition

$$(2q^2/3\alpha^2)\alpha' \ll 1. \quad (8.6)$$

This condition is satisfied not only in the transition regions in which q is small or α very large, but also for regions in which the curvature of the q profile is small.

9. CONCLUDING REMARKS

The example in Sec. 6 vividly illustrates the strength of the generalized WKB solutions developed in Sec. 4. In these solutions, the "local" wave solutions are chosen to be the Airy integral functions rather than the plane-wave phase memory expression used in the regular WKB solutions. It is shown in the example that the generalized WKB method yields the necessary solution at, above, and below, the reflection level. This was done without introducing any "corrections" in the general formulation of the solution at any of the levels, particularly when the curvature is not negligible. A concise treatment of the WKB method and Langer's method (applicable for small values of q^2) is given by Jones.[4]

For the sake of the illustration, the generalized WKB solutions were compared with the corresponding known asymptotic expansions. This process may be reversed to yield asymptotic expansions of functions in regions that are poorly tabulated.

It should be remembered that while in Sec. 4 the "local" wave solution was chosen to be the Airy

integral function, in certain cases it may be preferable to choose a different "local" solution on which to build the desired solution. The criterion in each case is the minimization of coupling coefficients over the entire range of the independent variable.

The problem discussed in this paper is the general second-order differential equation which is decomposed into two first-order differential equations. It can readily be seen how this method may be generalized to solve certain higher-order differential equations, which may be decomposed into several coupled first-order differential equations. As an example of this kind, one may recall[1] that in a magnetoionic medium the electromagnetic fields satisfy fourth-order differential equations (or four coupled first-order differential equations). When the parameters of the media are constant, these equations may be decoupled into the well-known forward and backward ordinary and extraordinary waves. But when the parameters of the media vary, these are coupled, and straightforward generalization of the method described in this paper may be used to solve the coupled equations.

In the event that it is not possible to choose a "local" solution such that the total coupling (in the entire range of the independent variable z) is not too large and an iterative solution of the coupled equations is not feasible, it is necessary to derive separate solutions in different ranges of the variable z. In this case it is necessary to match the field quantities at the boundaries between the separate regions. This, of course, is related to the solution based on the substitution of a given nonuniform dielectric coefficient profile with a discretely stratified model.[3,10] The difference, of course, between the two methods is that, in the solution developed in this paper, the dielectric coefficient in each layer is considered to vary in precisely the given manner, rather than assumed constant or even assumed to vary linearly. As a consequence, the generalized WKB method would require a significantly smaller number of "layers" and would possibly yield more accurate results. Alternatively, should a numerical method be resorted to in solving a set of coupled first-order differential equations [such as (2.2) and (3.1)], it would be preferable to solve them in terms of the loosely coupled dependent variables Φ_1 and Φ_2 [defined in (3.2) and (3.6)], rather than in terms of the original dependent variables Φ and Ψ. This would reduce the truncation error and permit the choice of a larger "step size" in the numerical analysis. A concise review of a number of numerical methods is given by Budden.[1]

---

[10] Extensive references to this work are given in Ref. 3.

In a recent paper on numerical solution of full wave equations, Inoue and Horowitz[11] strongly emphasize that the step size is critically determined by the assumed expression for the solution and for q in each "slab." Their numerical method is based on the extraction of the phase memory integral from the solution and the assumption that within a subslab the propagation coefficient matrix varies linearly.

Finally, it is of interest to point out that a concept similar to the one used in this paper has been used to derive the solutions of the "coupled" mode equations in nonuniform waveguides.[12,13] It has been shown that for the case in which the nonuniform waveguide is considered to consist of infinitesimal radial (or conical) waveguides, rather than infinitesimal rectangular (or cylindrical) waveguides, the coupling coefficients are substantially smaller.

ACKNOWLEDGMENTS

APPENDIX A

Generalized WKB Solution for (6.2), in which $\beta z \gg 1$

The asymptotic expansion for Ai (ζ) for $-\frac{2}{3}\pi \leq \arg \zeta \leq \frac{2}{3}\pi$ is

$$\text{Ai } (\zeta) \sim \tfrac{1}{2}\pi^{-\frac{1}{2}}\zeta^{-\frac{1}{4}} \exp\left(-\tfrac{2}{3}\zeta^{\frac{3}{2}}\right), \quad \text{(A1)}$$

in which ζ is given by (6.8). Therefore

$$\ln \text{Ai } (\zeta) = \ln \tfrac{1}{2}\pi^{-\frac{1}{2}} - \tfrac{1}{4}\ln \zeta - \tfrac{2}{3}\zeta^{\frac{3}{2}}. \quad \text{(A2)}$$

Now

$$\partial\zeta/\partial z = (k/\alpha)^{\frac{2}{3}}\alpha.$$

Hence

$$\frac{\partial}{\partial z}(\ln \text{Ai}) \sim \alpha\left(\frac{1}{4\zeta} + \zeta^{\frac{1}{2}}\right)\left(\frac{k}{\alpha}\right)^{\frac{2}{3}} = -\left(\frac{\beta}{4} + kCe^{\beta z/2}\right) \quad \text{(A3)}$$

and

$$\ln g_1 \equiv \int \frac{\partial}{\partial z}(\ln \text{Ai})\,dz = -\left(\frac{\beta z}{4} + \frac{2kC}{\beta}e^{\beta z/2}\right). \quad \text{(A4)}$$

Now, neglecting the coupling terms,

$$E_y \sim g_1 = \exp\left(-\tfrac{1}{4}\alpha z\right) \exp\left[-(2kC/\beta)e^{\beta z/2}\right]. \quad \text{(A5)}$$

APPENDIX B

We give asymptotic expansion of Langer's solution Ai (η) and the derivation of the coupling coefficient C_{11} for the case $g_{1,2} = \text{Ai } (\eta)$ and Bi (η), respectively.

For the exponential profile the variable η defined in

[11] Y. Inoue and S. Horowitz, Radio Sci. J. Res. (New Ser.) 1, 957 (1966).
[12] E. Bahar, Radio Sci. J. Res. (New Ser.) 1, 925 (1966).
[13] E. Bahar, Proc. IEE 13, 1741 (1966).

(6.18) is

$$\eta = \left[\tfrac{3}{2}i\left(\frac{k}{a_1}\right)^{\frac{1}{3}}\int_0^z q\frac{d\zeta_1}{dz}\,dz\right]^{\frac{2}{3}},$$

$$= \left[-\tfrac{3}{2}ik\int_0^z q\,dz\right]^{\frac{2}{3}} = \left[\tfrac{3}{2}kC\int_0^z [e^{\beta z} - 1]^{\frac{1}{2}}\,dz\right]^{\frac{2}{3}},$$

$$= \left\{\frac{3kC}{\beta}[(e^{\beta z} - 1)^{\frac{1}{2}} - \tan^{-1}[e^{\beta z} - 1]^{\frac{1}{2}}\right\}^{\frac{2}{3}},$$

$$\sim \left(\frac{3kC}{\beta}\right)^{\frac{2}{3}}e^{\beta z/3}; \quad \beta z \gg 1. \quad \text{(B1)}$$

Hence

$$\text{Ai } (\eta) \sim \eta^{-\frac{1}{4}}\exp\left(-\tfrac{2}{3}\zeta^{\frac{3}{2}}\right)$$
$$\sim \exp -(\beta z/12) \times \exp\left[-(2kC/\beta)e^{\beta z/2}\right], \quad \text{(B2)}$$

and Ai(η) is not in agreement with the correct solution (6.9). With

$$g_1 = \text{Ai } (\eta) \quad \text{(B3)}$$

$$\frac{d^2 g_1}{dz^2} = \frac{d^2 g_1}{d\eta^2}\left(\frac{d\eta}{dz}\right)^2 + \frac{dg_1}{d\eta}\frac{d^2\eta}{dz^2}, \quad \text{(B4)}$$

where

$$\frac{d\eta}{dz} = \frac{d\eta}{d\zeta_1}\frac{d\zeta_1}{dz} = -\frac{ikq}{\eta^{\frac{1}{2}}} \quad \text{and} \quad \frac{d^2\eta}{dz^2} = \frac{d\eta}{dz}\frac{1}{2}\frac{d}{dz}\ln\frac{q^2}{\eta},$$

$$W_z(g_1, g_2) = \frac{d\eta}{dz}W_\eta(g_1, g_2) = \frac{1}{\pi}\frac{d\eta}{dz};$$

$$L(g_1) = \text{Ai}'\frac{d\eta}{dz}\frac{1}{2}\frac{d}{dz}\ln\frac{q^2}{\eta},$$

$$\left.\begin{matrix}\end{matrix}\right\} \quad \text{(B5)}$$

$$C_{11} = \frac{\pi}{2}\text{Ai}'(\eta)\,\text{Bi}(\eta)\frac{d}{dz}\ln\frac{q^2}{\eta}. \quad \text{(B6)}$$

Now for small z

$$q^2 \sim a_1 z\left(1 + \frac{a_2}{a_1}z\right),$$

$$q \sim (a_1 z)^{\frac{1}{2}}\left(1 + \frac{a_2}{2a_1}z\right), \quad \text{(B7)}$$

$$(q^2)' \sim a_1\left(1 + \frac{2a_2}{a_1}z\right),$$

in which

$$a_1 \equiv (q^2)'_{z=0} \quad a_2 \equiv (q^2)''_{z=0},$$

$$\eta^{\frac{3}{2}} \sim -\tfrac{3}{2}ika_1^{\frac{1}{2}}\int_0^z\left(z^{\frac{1}{2}} + \frac{a_2}{2a_1}z^{\frac{3}{2}}\right)\,dz,$$

$$\sim -ika_1^{\frac{1}{2}}z^{\frac{3}{2}}\left(1 + \frac{3}{10}\frac{a_2}{a_1}z\right). \quad \text{(B8)}$$

Hence

$$\frac{d}{dz}\ln\frac{q^2}{\eta} = \frac{(q^2)'}{q^2} + \frac{ikq}{\eta^{\frac{3}{2}}} \sim \frac{4}{5}\frac{a_2}{a_1} = \frac{4}{5}\left[\frac{(q^2)''}{(q^2)'}\right]_{z=0} = \tfrac{4}{5}\beta \quad \text{(B9)}$$

and

$$C_{11} = \tfrac{1}{2}\pi\,\text{Ai}'(0)\,\text{Bi}(0)\tfrac{4}{5}\beta = -\beta/5, \quad z = 0.$$

For large z, $q^2 \sim -C^2 e^{\beta z}$,

$$C_{11} \sim -\tfrac{1}{4}(\beta - \beta/3) = -\tfrac{1}{6}\beta, \quad \beta z \gg 1. \quad \text{(B10)}$$

Recursion Relations for Coulomb Matrix Elements*

SAUL T. EPSTEIN,† JEAN H. EPSTEIN, AND BARBARA KENNEDY†

Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin

(Received 19 September 1966)

It is pointed out that, by using the identity $(\psi, (HW - WH')\psi') = (E - E')(\psi, W\psi')$, it is possible to derive useful relations among physically interesting matrix elements.

IN view of the continuing interest in the calculation of Coulomb matrix elements,[1] we wish to point out that it is possible, in a very simple way,[2] to derive useful recursion relations for such matrix elements. The relations are based on the identity[3]

$$(E - E')(\psi, W\psi') = (\psi, (HW - WH')\psi'), \quad (1)$$

where

$$H\psi = E\psi, \quad H'\psi' = E'\psi', \quad (2)$$

and where at least one of the wavefunctions ψ and ψ' describes a bound state.

For the specific application to Coulomb matrix elements we choose ψ and ψ' to be radial Coulomb wavefunctions, and H and H' to be the corresponding radial Hamiltonians[4]

$$H = \frac{p^2}{2} + \frac{k^2}{2R^2} - \frac{z}{R}, \quad (3)$$

$$H' = \frac{p^2}{2} + \frac{k'^2}{2R^2} - \frac{z'}{R}, \quad (4)$$

with p the radial momentum operator, $k^2 = l(l + 1)$, and $k'^2 = l'(l' + 1)$.

As an example of the sort of results obtainable in this way we quote the following recursion relations, which are derived by straightforward application of Eq. (1)–(4) to $W = R^s$ and $W = pR^s$:

$$(s/i)\{pR^{s-1}\} = (E - E')\{R^s\} + (z - z')\{R^{s-1}\} - \tfrac{1}{2}[s(s - 1) + k^2 - k'^2]\{R^{s-2}\}, \quad (5)$$

$$0 = \frac{(E - E')^2}{s + 2}\{R^{s+2}\} + \frac{(2s + 3)}{(s + 1)(s + 2)} \times (E - E')(z - z')\{R^{s+1}\}$$

$$+ \left[\frac{(z - z')^2}{s + 1} + (s + 1)(E + E') - \frac{(s + 1)}{s(s + 2)}(E - E')(k^2 - k'^2)\right]\{R^s\}$$

$$+ \left[\frac{(z + z')}{2}(2s + 1) - \frac{(2s + 1)(z - z')(k^2 - k'^2)}{2s(s + 1)}\right] \times \{R^{s-1}\}$$

$$+ \left[-sk^2 + \frac{(s(s - 1) + k^2 - k'^2)(s(s + 1) + k^2 - k'^2)}{4s}\right] \times \{R^{s-2}\}, \quad (6)$$

where[5]

$$\{A\} \equiv (\psi, A\psi').$$

The derivation of (5) is straightforward. One simply writes $HW - WH'$ as $HW - WH + W(H - H')$ and then uses the familiar commutation relation $Rp - pR = i$ and the standard rules of commutator algebra[6] to evaluate $HR^s - R^sH$. The derivation of (6) requires one further observation: namely, whenever powers of p higher than the first appear, as they do in evaluating $HpR^s - pR^sH$, then they can be eliminated in favor of first- and zero-order terms. Namely, by using the commutation relations, one can shift all factors of p^2 to the extreme left where, from (1), they act directly on ψ, and hence, from (2) and (3), can be replaced by

$$2E - k^2/R^2 + 2z/R.$$

Once this is done one uses (5) to eliminate the various $\{PR^n\}$ which appear, and the result is (6). Similar results can be derived for other choices of W and for other central potentials.[7]

ACKNOWLEDGMENT

[1] See H. B. Bebb, J. Math. Phys. **7**, 955 (1966), and references therein.

[2] For an application of this method to the calculation of diagonal matrix elements, see J. H. Epstein and S. T. Epstein, Am. J. Phys. **30**, 266 (1962).

[3] For some cautionary remarks concerning the use of such identities, see S. L. Gordon, J. Chem. Phys. **42**, 4184 (1965), Appendix IC.

[4] We let $e = m = \hbar = 1$.

[5] Specializing to the case $z = z'$, $E = E'$, one can, by use of (5) and (6), give a very simple derivation of the Pasternack–Sternheimer theorem [J. Math. Phys. **3**, 1280 (1962)].

[6] See, for example, E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961), p. 161.

[7] Results for $W = R^s e^{\beta R}$, $W = pR^s e^{\beta R}$ (see Ref. 1), $W = R^s e^{\beta R^2}$, and $W = pR^s e^{\beta R^2}$ for hydrogen, and for the isotropic harmonic oscillator, are given in the technical report WIS-TCI-191, available on request from the Reprint Librarian of the Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin.

Analytic Functionals in Quantum Field Theory*

V. Gorgé† AND F. Rohrlich

Department of Physics, Syracuse University, Syracuse, New York

(Received 6 December 1966)

A class of Lorentz invariant generalized functions can be defined as analytic functionals, i.e., as continuous linear functionals which are contour integrals over a suitable space of test functions. These generalized functions include in particular the invariant functions of quantum field theory, but also include "propagators with higher order poles." The analysis shows exactly which of these are well defined, especially in the important special case of zero mass. Various applications to quantum field theory are indicated.

1. INTRODUCTION

THE importance of the Lorentz invariant "functions" $\Delta_\Gamma(x)$ and $D_\Gamma(x)$ in quantum field theory is well appreciated. More recently, in connection with the gauge problem in quantum electrodynamics, similar functions arose which have poles of higher order. These have sometimes led to undefined or ambiguous expressions. Similar quantities arose in the extension of quantum field theory to nonrenormalizable interactions.

While it is recognized that these objects must be defined as distributions in the sense of Schwartz[1] or, more generally, as generalized functions in the sense of Gel'fand and Shilov,[2] the discussion of distribution theory in the physics literature[3,4] does not present the $\Delta_\Gamma(x)$ and $D_\Gamma(x)$ and their higher-order pole generalizations in a form suitable for applications in quantum field theory. The basic mathematical questions have all been presented[1-4] and special attention has also been paid to Lorentz invariant distributions.[5] Our task here is therefore mainly one of application of the theory of generalized functions to the particular functions of our concern and to make the associated mathematical formalism useful for quantum theory.

We found the representation in terms of analytic functionals especially convenient and close to what physicists have been doing on a formal level. After an explanation of these objects in Sec. 2, we present the generalized functions Δ_Γ^n as analytic functionals in Sec. 3. This is followed by an explicit x-space

representation and other properties ≡ of $\Delta_\Gamma^n(x)$ in Sec. 4. The special case $m = 0$, which yields the generalized functions D_Γ^n, is discussed in Sec. 5. The last section is devoted to applications of some of these results to quantum electrodynamics and to asymptotic quantum field theory. Some of the details of the proofs are relegated to two appendixes.

2. ANALYTIC FUNCTIONALS

We consider the set $\mathfrak{D}$ or $\mathfrak{K}$ (in the notation of Ref. 2) of all arbitrarily often differentiable functions of one variable which vanish outside a bounded domain. The set consists of the union (over a) of all complete countably normed spaces $\mathfrak{K}(a)$ of arbitrarily often differentiable functions which vanish outside the bounded domain $2a[g(t) = 0$ for $|t| > a]$. The reader is referred to Ref. 2 for the topology of these spaces since we do not make use of it here.

The Fourier transform $f_a(x)$ of a function in $K(a)$ can be extended to the complex variable $z = x + iy$. $f_a(z)$ is an entire function of slow increase in the sense that

$$|z^n| \cdot |f_a(z)| \leq C_n e^{a|y|} \quad (n = 0, 1, 2, \cdots). \quad (2.1)$$

The set of all $f_a(z)$ is the space $\mathfrak{Z}(a) = \mathfrak{F}\mathfrak{K}(a)$, where $\mathfrak{F}$ is the Fourier transform operation. The union over a of these spaces gives $\mathfrak{Z} = \mathfrak{F}\mathfrak{K}$.

An analytic functional T_Γ on $\mathfrak{Z}$ is a continuous linear functional of the $f \in \mathfrak{Z}$ characterized by a contour integral

$$(T_\Gamma, f) = \int_\Gamma T(\zeta) f(\zeta) \, d\zeta. \quad (2.2)$$

For example, the Dirac δ function can be defined as an analytic functional with support at the point z by

$$(\delta_D(z), f) \equiv \oint \delta(\zeta - z) f(\zeta) \, d\zeta \equiv \frac{1}{2\pi i} \oint \frac{f(\zeta) \, d\zeta}{\zeta - z}, \quad (2.3)$$

so that

$$\delta(\zeta - z) = (1/2\pi i)[1/(\zeta - z)], \quad (2.4)$$

[1] L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1957).

[2] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. I; and *Verallgemeinerte Funktionen* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1962), Vol. II.

[3] L. Gårding and J. L. Lions, Nuovo Cimento (Suppl.) **14**, 9 (1959).

[4] H. J. Bremmermann and L. Durand, III, J. Math. Phys. **2**, 240 (1961), and earlier literature quoted there.

[5] P.-D. Methée, Comm. Math. Helv. **28**, 225 (1954); **32**, 152 (1957).

and the contour is the usual closed path in the positive sense containing the point z. Equation (2.3) implies

$$(\delta_D(z), f) = f(z). \tag{2.5}$$

The nth derivative of $\delta_D(z)$ is also defined:

$$(\delta_D^{(n)}(z), f) \equiv (D^n \delta_D(z), f),$$

$$\equiv \oint \left(\frac{d^n}{d\zeta^n} \delta(\zeta - z) \right) f(\zeta) \, d\zeta,$$

$$= (-1)^n \frac{n!}{2\pi i} \oint \frac{f(\zeta) \, d\zeta}{(\zeta - z)^{n+1}} = (-1)^n f^{(n)}(z). \tag{2.6}$$

The property

$$(D^n T_\Gamma, f) = (-1)^n (T_\Gamma, D^n f) \tag{2.7}$$

of the derivative operator D is, of course, a general property of generalized functions.

A class of analytic functionals of special interest in quantum field theory is closely related to $\delta_D(z)$. We discuss some of these, viz., δ_P, δ_R, and δ_A, first and then generalize to the related Lorentz invariant distributions in the following Sec. 3.

We define for real argument $x = \mathrm{Re}\, z$

$$(\delta_P(x), f) \equiv \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f(\xi) - f(x)}{\xi - x} \, d\xi, \tag{2.8}$$

the contour following the real axis. This integral is just the Cauchy principal value, usually denoted by P preceding the integral. For the purpose of future generalization we define

$$R \int_{-\infty}^{\infty} \frac{f(\xi) \, d\xi}{\xi - x} \equiv \int_{-\infty}^{\infty} \frac{f(\xi) - f(x)}{\xi - x} \, d\xi,$$

$$= \lim_{\epsilon \to +0} \left(\int_{-\infty}^{x-\epsilon} + \int_{x+\epsilon}^{\infty} \right) \frac{f(\xi) \, d\xi}{\xi - x},$$

$$\equiv P \int_{-\infty}^{\infty} \frac{f(\xi) \, d\xi}{\xi - x}. \tag{2.9}$$

We can write symbolically, following the definition (2.8),

$$\delta_P(\xi - x) = (1/2\pi i) R[1/(\xi - x)]. \tag{2.10}$$

By induction one verifies easily that (with symmetric integration about the pole)

$$\frac{d^n}{dx^n} R \int_{-\infty}^{\infty} \frac{f(\xi) \, d\xi}{\xi - x}$$

$$= n! \int_{-\infty}^{\infty} \frac{f(\xi) - \sum_{\nu=0}^{n} \frac{1}{\nu!} (\xi - x)^\nu f^{(\nu)}(x)}{(\xi - x)^{n+1}} \, d\xi. \tag{2.11}$$

If the integral on the right is symbolically written as

$$R \int_{-\infty}^{\infty} \frac{f(\xi) \, d\xi}{(\xi - x)^{n+1}}, \tag{2.12}$$

Eq. (2.11) can be expressed by the formally trivial relation

$$\frac{d^n}{dx^n} R \frac{1}{\xi - x} = n! \, R \frac{1}{(\xi - x)^{n+1}}. \tag{2.13}$$

The integral (2.12) is well defined for all n. The operator R is clearly a generalization of the principal value integral. The latter is defined only for $n = 0$, in which case R and P are identical, according to (2.9). The operation R is sometimes called "regularization." It is not an arbitrary cutoff procedure, but appears here as the natural extension of the Cauchy principal value to $n = 1, 2, \cdots$.

We can now define the generalized function $\delta_P^{(n)}$ by

$$(\delta_P^{(n)}(x), f) \equiv (-1)^n \frac{n!}{2\pi i} R \int_{-\infty}^{\infty} \frac{f(\xi) \, d\xi}{(\xi - x)^{n+1}}$$

$$(n = 0, 1, \cdots). \tag{2.14}$$

It is easy to show, however, that

$$\delta_P^{(n)} = D^n \delta_P \tag{2.15}$$

because

$$(D \delta_P^{(n)}, f) = -(\delta_P^{(n)}, Df),$$

$$= (-1)^{n+1} \frac{n!}{2\pi i} \int_{-\infty}^{\infty} \frac{d\xi}{(\xi - x)^{n+1}}$$

$$\times \left[f'(\xi) - \sum_{\nu=0}^{n} \frac{(\xi - x)^\nu}{\nu!} f^{(\nu+1)}(x) \right],$$

$$= (-1)^{n+1} \frac{(n+1)!}{2\pi i} R \int_{-\infty}^{\infty} \frac{1}{(\xi - x)^{n+2}} f(\xi) \, d\xi,$$

$$= (\delta_P^{(n+1)}, f).$$

Here we use integration by parts and the definition (2.12) for the integral on the right of (2.11). This proves (2.15) by induction.

Closely related are the two analytic functionals

$$(\delta_R, f) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta) \, d\zeta}{\zeta - x}, \tag{2.14_R}$$

$$(\delta_A, f) = \frac{1}{2\pi i} \int_{C_A} \frac{f(\zeta) \, d\zeta}{\zeta - x}, \tag{2.14_A}$$

where the contour $C_R(C_A)$ follows the real axis from $-\infty$ to $x - \epsilon$, describes a semicircle with center at x from $x - \epsilon$ to $x + \epsilon$ going into the upper (lower) half-plane, and continues along the real axis from $x + \epsilon$ to $+\infty$; the limit $\epsilon \to 0$ is then taken. In a well-known way one has

$$\int_{C_{R,A}} \frac{f(\zeta) \, d\zeta}{\zeta - x} = P \int_{-\infty}^{\infty} \frac{f(\zeta) \, d\zeta}{\zeta - x} \mp \frac{1}{2} \oint \frac{f(\zeta) \, d\zeta}{\zeta - x}.$$

Therefore,

$$\delta_{R,A} = \delta_P \mp \tfrac{1}{2} \delta_D. \tag{2.16}$$

The integrals on the right of this decomposition can both be differentiated with respect to x an arbitrary number of times. One can, therefore, define a class of analytic functionals over $\mathfrak{Z}$ by

$$(\delta_{R,A}^{(n)}(x), f) \equiv (-1)^n \frac{n!}{2\pi i} \int_{C_{R,A}} \frac{f(\zeta)\, d\zeta}{(\zeta - x)^{n+1}}, \quad (2.17)$$

and we find

$$\delta_{R,A}^{(n)} = \delta_P^{(n)} \mp \tfrac{1}{2}\delta_D^{(n)}. \quad (2.18)$$

This equation can also be written in terms of the integrands as

$$\frac{1}{(\zeta - x)^{n+1}}\bigg|_{R,A} = R\,\frac{1}{(\zeta - x)^{n+1}} \mp \frac{(-1)^n}{n!}\, i\pi\delta^{(n)}(\zeta - x), \quad (2.19)$$

where $\delta^{(n)}$ is the nth derivative of the δ function (2.4).

Analytic functionals have a Taylor expansion,

$$T_\Gamma(z + h) = \sum_{\nu=0}^{\infty} D^\nu T_\Gamma(z)\frac{h^\nu}{\nu!},$$

where the contour Γ depends, of course, on the argument of T. The analytic functionals $\delta_P^{(n)}(z)$, $\delta_R^{(n)}(z)$, and $\delta_A^{(n)}(z)$ are therefore defined for all z by complex extension, just as $D^n\delta_D(z)$ in (2.6). The contours C_P, C_R, and C_A, which for $z = x$ were defined along the real axis, are now parallel displaced with suitable half-circles around z.

3. "INVARIANT FUNCTIONS" AND RELATED DISTRIBUTIONS

The Lorentz-invariant "functions" used in quantum field theory are actually Lorentz-invariant generalized functions that are defined by a contour integral in the complex plane. They are therefore analytic functionals in the sense discussed in the previous section. More specifically, they are solutions of the homogeneous d'Alembert equation in Minkowski space,

$$(\Box - m^2)\Delta_\Gamma(x) = 0 \quad (3.1)$$

or of the inhomogeneous equation

$$(\Box - m^2)\Delta_\Gamma(x) = -\delta_4(x) \quad (3.2)$$

(with $m \neq 0$ or $m = 0$), but are specified usually as Fourier transforms of a function of $p^2 = \mathbf{p}^2 - (p^0)^2$. The latter is necessarily of the form $(p^2 + m^2)^{-1}$ [or $(p^2)^{-1}$ for $m = 0$] and the contour is specified in the complex p^0 plane. In the notation of (2.2) we have

$$T(p) = 1/(p^2 + m^2) \quad (p^0 \text{ complex}). \quad (3.3)$$

In terms of the differential equations (3.1) or (3.2) the contour is, of course, equivalent to a specification of

asymptotic boundary conditions suitable to characterize the solution uniquely.

We here consider a more general class of analytic functionals which contain the above as a special case. To this end it is convenient to consider separately the two cases $m \neq 0$ and $m = 0$. The latter is postponed to Sec. 5.

The test function space is now the set of functions[6] $\varphi(p) = f(p^0)g(\mathbf{p})$ with $f \in \mathfrak{Z}$ and $g \in S(R^3)$. In the notation of (2.4) we define the generalized function

$$(\Delta_+, \varphi) \equiv \frac{1}{2\pi i} \int \frac{d^3p}{2\omega}\, g(\mathbf{p}) \int_{C_+} \delta(p^0 - \omega)f(p^0)\, dp^0, \quad (3.4)$$

where $\omega = +(\mathbf{p}^2 + m^2)^{\frac{1}{2}}$ and the contour C_+ is a circle containing the point $p^0 = \omega$ and is traversed in the clockwise direction. This object is an analytic functional with parameter ω, viz., $\delta_D(\omega)$, which, when "integrated" over the test function f, yields a tempered distribution:

$$(\Delta_+, \varphi) = (T, g), \quad \text{where } T = -(1/4\pi i\omega)(\delta_D(\omega), f) \quad (3.5)$$

The minus sign arises from the fact that δ_D is defined with the path $-C_+$.

Substitution of (2.4) into (3.4) yields

$$(\Delta_+, \varphi) = \int d^3p \int_{C_+} dp^0 \frac{\varphi(p)}{(2\pi)^2 \cdot 2\omega \cdot (\omega - p^0)},$$

$$= \int d^3p \int_{C_+} dp^0 \tilde{\Delta}_+(p)\varphi(p). \quad (3.6)$$

The notation $\tilde{\Delta}_+(p)$ is used here because, if one formally replaces $\varphi(p)$ by

$$\varphi_x(p) \equiv [1/(2\pi)^2]e^{ip\cdot x} \quad (3.7)$$

(which is *not* in our test function space), one obtains the well-known representation of the tempered distribution $\Delta_+(x)$:

$$(\Delta_+, \varphi_x) = \frac{1}{(2\pi)^2} \int d^3p \int_{C_+} dp^0 \tilde{\Delta}_+(p)e^{ip\cdot x} = \Delta_+(x). \quad (3.8)$$

The Fourier transform is discussed in general at the beginning of Sec. 4.

In the same manner one finds that the distribution defined by

$$(\Delta_-, \varphi) \equiv i \int \frac{d^3p}{4\pi\omega} \int_{C_-} \delta(p^0 + \omega)\varphi(p)\, dp^0,$$

$$= \int d^3p \int_{C_-} dp^0 \tilde{\Delta}_-(p)\varphi(p), \quad (3.9)$$

---

[6] The specification $\varphi \in S(R^3) \otimes \mathfrak{Z}(Z)$ is, of course, the essential point. That $\varphi(p)$ can be factored into $g(\mathbf{p})f(p^0)$ is by no means necessary. It is assumed here only for the sake of clarity of presentation, but the argument clearly carries through without this assumption.

with C_- a circle containing the point $p^0 = -\omega$ and traversed in the clockwise direction, corresponds to the invariant distribution $\Delta_-(x)$. The distributions $\Delta = \Delta_+ + \Delta_-$ and $i\Delta_1 = \Delta_+ - \Delta_-$ then follow due to linearity. This exhausts the four "homogeneous" Δ_Γ functions, i.e., the four well-known Lorentz-invariant solutions of the homogeneous equation (3.1).

The definition of Δ_P, analogous to $\Delta_P(x)$ in terms of an analytic functional, makes use of δ_P defined in (2.10):

$$(\Delta_P, \varphi) = -i \int \frac{d^3p}{4\pi\omega} \int \delta_P(p^0 - \omega)\varphi(p) \, dp^0$$

$$+ i \int \frac{d^3p}{4\pi\omega} \int \delta_P(p^0 + \omega)\varphi(p) \, dp^0. \quad (3.10)$$

The two integrals are *not* Lorentz-invariant separately, but they combine to the Lorentz-invariant expression

$$(\Delta_P, \varphi) = \frac{1}{(2\pi)^2} \int d^3p \, R \int \frac{\varphi(p) \, dp^0}{(2\pi)^2 (p^2 + m^2)},$$

$$= \frac{1}{(2\pi)^2} \int \tilde{\Delta}_P(p)\varphi(p) \, d^4p, \quad (3.11)$$

$$\tilde{\Delta}_P(p) = \frac{1}{(2\pi)^2} R \frac{1}{p^2 + m^2}. \quad (3.12)$$

The meaning of the R operation is that given in the last section, provided the expression $(p^2 + m^2)^{-1}$ is written in its partial fraction expansion [see (3.14) below].

Since all five inhomogeneous Δ_Γ [i.e., solutions of (3.2)] are given in terms of one of them together with the four homogeneous ones, the linearity of the distributions gives us Δ_R, Δ_A, Δ_{1R}, Δ_{1A} trivially, once Δ_P is known.[7]

All Δ_Γ can be expressed in terms of $p^2 + m^2$. Thus, (3.6) and (3.9) could also be written as

$$(\Delta_\pm, \varphi) = \frac{1}{(2\pi)^2} \int d^3p \int_{C_\pm} dp^0 \frac{\varphi(p)}{p^2 + m^2}.$$

We therefore now ask for the representation of the more general class of distributions defined by

$$(\Delta_\Gamma^n, \varphi) = \frac{1}{(2\pi)^4} \int d^3p \int_\Gamma dp^0 \frac{\varphi(p)}{(p^2 + m^2)^{n+1}} \quad (3.13)$$

in terms of the distributions $\delta_D^{(\nu)}(p^2 + m^2)$ and $R[1/(p^2 + m^2)]$. With this definition Δ_Γ^n are the well-known invariant functions Δ_Γ.

The method of derivation consists in reducing this problem to the one-dimensional case discussed in

Sec. 2. This is done by means of the partial fraction expansion

$$\frac{1}{(p^2 + m^2)^n} = \frac{1}{(\omega - p^0)^n(\omega + p^0)^n},$$

$$= \sum_{\nu=0}^{n-1} \binom{n + \nu - 1}{\nu} \frac{1}{(2\omega)^{n+\nu}}$$

$$\times \left[\frac{1}{(\omega - p^0)^{n-\nu}} + \frac{1}{(\omega + p^0)^{n-\nu}} \right]. \quad (3.14)$$

Using

$$\left. \frac{1}{(p^0 \mp \omega)^n} \right|_P = R \frac{1}{(p^0 \mp \omega)^n} \quad (3.15)$$

in the notation of (2.19), we have for the quantity $\tilde{\Delta}_P^n(p)$

$$(2\pi)^2 \tilde{\Delta}_P^n(p) \equiv \left. \frac{1}{(p^2 + m^2)^{n+1}} \right|_P,$$

$$= \sum_{\nu=0}^{n} \binom{n + \nu}{\nu} \frac{1}{(2\omega)^{n+\nu+1}}$$

$$\times \left[R \frac{(-1)^{n-\nu+1}}{(p^0 - \omega)^{n-\nu+1}} + R \frac{1}{(p^0 + \omega)^{n-\nu+1}} \right],$$

$$= R \frac{1}{(p^2 + m^2)^{n+1}}. \quad (3.16)$$

Here we employed (3.14) twice.

For the homogeneous functions (closed contours C_i) we need (2.6), which, for the present variables and in obvious symbolic notation, gives

$$\left. \frac{1}{(p^0 \mp \omega)^{n+1}} \right|_{C_\pm} = - \frac{(-1)^n 2\pi i}{n!} \delta^{(n)}(p^0 \mp \omega). \quad (3.17)$$

The contours C and C_1 are related to C_+ and C_- by $C = C_+ + C_-$ and $C_1 = C_+ - C_-$. Therefore,

$$\left. \frac{1}{(p^2 + m^2)^{n+1}} \right|_{C,C_1} = + \sum_{\nu=0}^{n} \binom{n + \nu}{\nu} \frac{2\pi i}{(n - \nu)!} \frac{1}{(2\omega)^{n+1+\nu}}$$

$$\times [\delta^{(n-\nu)}(p^0 - \omega) \mp (-1)^{n-\nu}\delta^{(n-\nu)}(p^0 + \omega)]. \quad (3.18)$$

In order to express this in terms of $p^2 + m^2$ one needs the auxiliary expansions [$\epsilon(p^0) = \pm 1$ for $p^0 \gtrless 0$]:

$$\delta^{(n)}(p^2 + m^2) = \sum_{\nu=0}^{n} \frac{1}{\nu!} \frac{(n + \nu)!}{(n - \nu)!} \frac{1}{(2\omega)^{n+\nu+1}}$$

$$\times [\delta^{(n-\nu)}(\omega - p^0) + (-1)^{n-\nu}\delta^{(n-\nu)}(p^0 + \omega); \quad (3.19)$$

$$\epsilon(p^0)\delta^{(n)}(p^2 + m^2) = \sum_{\nu=0}^{n} \frac{1}{\nu!} \frac{(n + \nu)!}{(n - \nu)!} \frac{1}{(2\omega)^{n+\nu+1}}$$

$$\times [\delta^{(n-\nu)}(\omega - p^0) - (-1)^{n-\nu}\delta^{(n-\nu)}(p^0 + \omega)]. \quad (3.20)$$

---

[7] For notation adopted here see J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1959), 2nd printing.

TABLE I. The analytic functionals Δ_Γ^n ($n \geq 0$) expressed in terms of $R(p^2 + m^2)^{-\nu}$ and $\delta^{(\nu)}(p^2 + m^2)$.

Δ_Γ^n	Contour	$(2\pi)^2\tilde{\Delta}_\Gamma^n(p)$
$\Delta_\pm^n$	$C_\pm$	$\pm\dfrac{2\pi i}{n!}\theta(\pm p^0)\delta^{(n)}(p^2 + m^2)$
Δ^n	$C = C_+ + C_-$	$\dfrac{2\pi i}{n!}\epsilon(p^0)\delta^{(n)}(p^2 + m^2)$
$i\Delta_1^n$	$C_1 = C_+ - C_-$	$\dfrac{2\pi i}{n!}\delta^{(n)}(p^2 + m^2)$
Δ_P^n	C_P	$R\dfrac{1}{(p^2 + m^2)^{n+1}}$
$\Delta_{R,A}^n$	$C_P \pm \tfrac{1}{2}C$	$R\dfrac{1}{(p^2 + m^2)^{n+1}} \pm \dfrac{i\pi}{n!}\epsilon(p^0)\delta^{(n)}(p^2 + m^2)$
$\Delta_{1R,1A}^n$	$C_P \pm \tfrac{1}{2}C_1$	$R\dfrac{1}{(p^2 + m^2)^{n+1}} \pm \dfrac{i\pi}{n!}\delta^{(n)}(p^2 + m^2)$

These are proved in Appendix A. With their aid we now have

$$\frac{1}{(p^2 + m^2)^{n+1}}\bigg|_C = \frac{2\pi i}{n!}\epsilon(p^0)\delta^{(n)}(p^2 + m^2), \quad (3.21)$$

$$\frac{1}{(p^2 + m^2)^{n+1}}\bigg|_{C_1} = \frac{2\pi i}{n!}\delta^{(n)}(p^2 + m^2). \quad (3.22)$$

The other two homogeneous cases, corresponding to contours C_+ and C_-, then follow as linear combinations of these.

We can thus give $\tilde{\Delta}_\Gamma^n(p)$ for all four homogeneous and all five inhomogeneous cases, the latter being linear combinations of $\tilde{\Delta}_P^n$ and the homogeneous $\tilde{\Delta}_\Gamma^n$. The results are collected in Table I. This table is valid for $n \geq 0$ but can be taken over also for negative integers n, provided one defines $\delta^{(n)} \equiv 0$ ($n < 0$). Then all homogeneous $\Delta_F^{-|n|} = 0$ and all inhomogeneous ones become simply $(p^2 + m^2)^{|n|}$ ($n < 0$).

4. PROPERTIES AND REPRESENTATIONS OF $\Delta_\Gamma^n(x)$

As already mentioned in Sec. 2, $f(p^0) \in \mathfrak{Z}$, for complex p^0 implies that its Fourier transform $\tilde{f}(x^0) \in \mathfrak{D}$, where

$$f(p^0) \equiv \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^\infty \tilde{f}(x^0)e^{-ip^0 x^0}\,dx^0. \quad (4.1)$$

On the other hand, $g(\mathbf{p}) \in S(R^3)$ implies that its Fourier transform $\tilde{g}(x)$ also $\in S(R^3)$. Thus, with[6] $\tilde{\varphi}(x) = \tilde{f}(x^0)\tilde{g}(x)$, one can define

$$\Delta_\Gamma^n(x) \equiv \frac{1}{(2\pi)^2}\int \tilde{\Delta}_\Gamma^n(p)e^{ip\cdot x}\,d^4p, \quad (4.2)$$

and find

$$(\Delta_\Gamma^n(x),\, \varphi(x)) = \frac{1}{(2\pi)^2}\int \tilde{g}(\mathbf{x})e^{i\mathbf{p}\cdot\mathbf{x}}\,d^3x$$

$$\times \int \tilde{f}(x^0)e^{-ip^0 x^0}\tilde{\Delta}_\Gamma(p)\,dx^0\,d^4p,$$

$$= \int g(\mathbf{p})f(p^0)\tilde{\Delta}_\Gamma(p)\,d^4p,$$

$$= (\tilde{\Delta}_\Gamma^n(p),\, \varphi(p)). \quad (4.3)$$

The differential equations satisfied by $\Delta_\Gamma^n(x)$ follow from the well-known equalities [easily established with (3.19)]

$$(p^2 + m^2)^m\delta^{(n)}(p^2 + m^2) = 0 \quad (m > n) \quad (4.4)$$

and

$$(p^2 + m^2)^nR[1/(p^2 + m^2)^n] = 1, \quad (4.5)$$

which can readily be proved. The results tabulated in Table I tell us that the "homogeneous" functions $\Delta_H^n \equiv \{\Delta^n, \Delta_\pm^n, \Delta_1^n\}$ satisfy

$$(\Box - m^2)^{n+1}\Delta_H^n(x) = 0, \quad (4.6)$$

while the "inhomogeneous" functions

$$\Delta_I^n \equiv \{\Delta_P^n, \Delta_{R,A}^n, \Delta_{1R,1A}^n\}$$

satisfy

$$(\Box - m^2)^{n+1}\Delta_I^n(x) = (-1)^{n+1}\delta_4(x). \quad (4.7)$$

A recursion relation for Δ_Γ^n of the same Γ but different n is obtained from the observation that (2.13) permits one to write

$$\left(\frac{d}{dm^2}\right)^nR\frac{1}{p^2 + m^2} = (-1)^n n!\,R\frac{1}{(p^2 + m^2)^{n+1}}. \quad (4.8)$$

Table I and (4.2) combine to yield the fundamental equation

$$\Delta_\Gamma^n(x) = \frac{(-1)^n}{n!}\left(\frac{d}{dm^2}\right)^n\Delta_\Gamma(x), \quad (4.9)$$

and therefore also

$$\Delta_\Gamma^{n+\nu}(x) = (-1)^\nu\frac{n!}{(n+\nu)!}\left(\frac{d}{dm^2}\right)^\nu\Delta_\Gamma^n(x). \quad (4.10)$$

Since $\theta(\pm x^0)$, defined to be 1 for $x^0 \gtrless 0$ and 0 otherwise, commutes with d/dm^2, the usual relations also hold for the Δ_Γ^n:

$$\Delta_{R,A}^n(x) = \pm\theta(\pm x^0)\Delta_P^n(x),$$
$$\Delta_{1R,1A}^n(x) = \Delta_P^n(x) \pm \tfrac{1}{2}i\Delta_1^n(x), \quad (4.11)$$
$$\Delta^n(x) = 2\epsilon(x^0)\Delta_P^n(x).$$

The knowledge of one inhomogeneous function (e.g., Δ_P^n) and two homogeneous functions (e.g., Δ^n and Δ_1^n) therefore completely determines all the others. The last relation, however, shows that Δ_P^n and Δ_1^n

alone suffice. We can now make use of the known representation of the $\Delta_\Gamma(x)$ in terms of cylinder functions. The $\Delta_\Gamma^n(x)$ then follows easily from (4.9). For this purpose one needs the following differential relations for the unmodified cylinder functions $Z_n = \{J_n, N_n, H_n^{(1)}, H_n^{(2)}\}$ and for the modified cylinder function K_n:

$$\left(\frac{d}{z\,dz}\right)^n (z^m Z_m(z)) = z^{m-n} Z_{m-n}(z), \quad (4.12)$$

$$\left(-\frac{d}{z\,dz}\right)^n (z^m K_m(z)) = z^{m-n} K_{m-n}(z), \quad (4.13)$$

$$\left(-\frac{d}{z\,dz}\right)^n \left(\frac{Z_m(z)}{z^m}\right) = \frac{Z_{m+n}(z)}{z^{m+n}}. \quad (4.14)$$

The last relation is also valid for $K_m(z)$; we also note that $Z_{-n} = (-1)^n Z_n$ and $K_{-n} = K_n$.

With the notation $x^2 \equiv x_\mu x^\mu$, $u \equiv +(|x^2|)^{\frac{1}{2}}$, $\delta_{n0} = 1$ for $n = 0$, zero otherwise, one finds

$$\Delta_P^n(x) = \frac{(-1)^n}{n!}\left(\frac{d}{2m\,dm}\right)^n$$
$$\times \left[\frac{1}{4\pi}\delta(u^2) - \theta(-x^2)\frac{m^2}{8\pi}\frac{J_1(mu)}{mu}\right],$$
$$= \frac{\delta_{n0}\delta(u^2)}{4\pi} + \frac{\theta(-x^2)}{16\pi n!}\left(\frac{u}{2m}\right)^{n-1} J_{n-1}(mu). \quad (4.15)$$

In the same way one finds $(n \geq 0)$

$$\Delta_1^n(x) = \frac{1}{8\pi n!}\left(\frac{u}{2m}\right)^{n-1}\begin{cases} -N_{n-1}(mu) \\ \frac{2}{\pi}K_{n-1}(mu) \end{cases} \quad \text{for} \quad \begin{cases} x^2 < 0, \\ x^2 > 0. \end{cases} \quad (4.16)$$

Since $\Delta^n(x) = 2\epsilon(x^0)\Delta_1^n(x)$, according to (4.11), Eqs. (4.15) and (4.16) give two homogeneous and one inhomogeneous Δ_Γ^n, and therefore all Δ_Γ^n by linear combination. Because of their special importance in quantum field theory, we give here the particular combinations Δ_{1R}^n and Δ_{1A}^n explicitly:

$$\Delta_{1R}^n(x) = \frac{\delta_{n0}\delta(u^2)}{4\pi} + \frac{1}{16\pi n!}\left(\frac{u}{2m}\right)^{n-1}$$
$$\times \begin{cases} H_{n-1}^{(2)}(mu) \\ (2i/\pi)K_{n-1}(mu) \end{cases} \quad \text{for} \quad \begin{cases} x^2 < 0, \\ x^2 > 0, \end{cases} \quad (4.17)$$

$$\Delta_{1A}^n(x) = \frac{\delta_{n0}\delta(u^2)}{4\pi} + \frac{1}{16\pi n!}\left(\frac{u}{2m}\right)^{n-1}$$
$$\times \begin{cases} H_{n-1}^{(1)}(mu) \\ -(2i/\pi)K_{n-1}(mu) \end{cases} \quad \text{for} \quad \begin{cases} x^2 < 0, \\ x^2 > 0. \end{cases} \quad (4.18)$$

Equations (4.15)–(4.18) are all valid for all nonnegative integers n.

The special case $\Gamma = C_R$ was first obtained by Bhabha.[8] The function Δ_R^n resulting from the above agrees with his work apart from notation. The functions Δ_{1R}^n and Δ_{1A}^n can be compared with the results of Ref. 2, p. 365. They agree.

5. GENERALIZED FUNCTIONS D_Γ^n

We define

$$D_\Gamma^n = \lim_{m\to 0} \Delta_\Gamma^n. \quad (5.1)$$

The study of this limit is easiest in the x representation. The leading terms of the cylinder functions for $|z| \ll 1$, viz.,

$$J_n(z) \cong (\tfrac{1}{2}z)^n/n!$$

$$N_0(z) \cong \frac{2}{\pi}\ln\frac{\gamma z}{2}; \quad N_n(z) \cong -\frac{1}{\pi}\left(\frac{2}{z}\right)^n \quad (n > 0),$$

$$K_0(z) \cong -\ln\frac{\gamma z}{2}; \quad K_n(z) \cong \frac{1}{2}\left(\frac{2}{z}\right)^n \quad (n > 0), \quad (5.2)$$

where $\ln\gamma = C =$ Euler's constant, can be combined with our results (4.15) and (4.16). One obtains

$$D_P^0(x) = \delta(x_\mu x^\mu)/4\pi,$$
$$D_P^n(x) = [\theta(-x^2)/16\pi(n-1)!\,n!](\tfrac{1}{4}u^2)^{n-1}$$
$$(n > 0), \quad (5.3)$$

$$D_1^0(x) = -1/2\pi^2 x^2. \quad (5.4)$$

The limit for $D_1^n(x)$ $(n > 0)$ does not exist. As a consequence the D_Γ^n for $n = 0$ all exist, but for $n > 0$ only those exist which can be derived from D_P^n, viz., D_P^n, D_R^n, D_A^n, and D^n. The limits D_{1R}^n, D_{1A}^n, D_1^n, $D_\pm^n$ for $n > 0$ do not exist.

This result is in agreement with analyticity considerations. The limit $m \to 0$ leads to an integration over ω [see, e.g., (3.4)] with lower limit zero and corresponds to the coalescence of the two poles $p^0 = \pm\omega$, $\omega = |\mathbf{p}|$ in the analytic functional. If the path of integration is thereby pinched, the expression will not exist. On the other hand, if it is not pinched, analyticity requires its existence. The above results confirm that D_Γ^n exists for $n > 0$ if and only if the contour is not pinched when the poles coalesce. The case $n = 0$ is an exception that works because of the measure $\omega^2\,d\omega$ which vanishes sufficiently fast at $\omega = 0$ (see below).

Consider this problem now in p space. One can, of course, take the limit $m \to 0$ in Table I. But the resultant analytic functionals in p^0 will have a (parametric) dependence on ω which is, in general, not a

[8] H. Bhabha, Phys. Rev. 77, 665 (1950).

distribution. This is due to the high powers of $1/\omega$, resulting from the partial fraction expansion (3.14), and the domain of ω, which now has lower limit $m = 0$. Thus, (3.14) cannot be used in the case $m = 0$.

Nevertheless, despite (3.16), $\tilde{D}_P^n$ exists for all n. The reason lies in the regularity of $R(p^2 + m^2)^n$ as a function of m at $m = 0$:

$$\lim_{m \to 0} R \int_{-\infty}^{\infty} \frac{f(p^0)\, dp^0}{(p^2 + m^2)^n} = R \int_{-\infty}^{\infty} \frac{f(p^0)\, dp^0}{(p^2)^n}$$

exists because (2.14) exists for $x = 0$ for all n.

An apparent paradox arises from the fact proved above that $\tilde{D}^n$ exists for all $n \geq 0$ while $\tilde{D}_1^n$ exists only for $n = 0$. Comparison of (3.19) and (3.20) makes this difficult to understand. An explicit proof is therefore given in Appendix B.

While the partial fraction expansion is thus not applicable for $n > 0$, it does permit one to see that $\tilde{D}_\Gamma^n$ exists for all Γ. With $\omega = |\mathbf{p}|$,

$$(\tilde{D}_{\Gamma,\varphi}^0) = \frac{1}{(2\pi)^2} \int d^3p\, h(\mathbf{p}) \int_\Gamma dp^0 \frac{f(p^0)}{p^2},$$

$$= \frac{1}{(2\pi)^2} \int_0^\infty \frac{\omega^2\, d\omega\, d\Omega\, h(\mathbf{p})}{2\omega}$$

$$\times \int_\Gamma dp^0 f(p^0) \left(\frac{1}{\omega - p^0} + \frac{1}{\omega + p^0} \right). \quad (5.5)$$

The analytic functional can be written, using $\eta = p^0/\omega$,

$$\int_\Gamma d\eta\, f(\omega\eta) \left[\frac{1}{1 - \eta} + \frac{1}{1 + \eta} \right],$$

and is, therefore, nondivergent for $\omega \to 0$. The integral in (5.5) is, therefore, well defined for all Γ.

We remark parenthetically that there are *mathematical* possibilities to define D_Γ^n for *all* Γ. But one must then abandon the definition (5.1) which appears to us natural from the point of view of physics. If one discards (5.1), one can employ a regularization of integrals of the form

$$\int_0^\infty \frac{d\omega}{\omega^n} h(\omega),$$

where $h \in \mathcal{S}$. If distributions of the type $T(\omega) = 0$ $(\omega < 0)$, $T(\omega) = \omega^{-n}$ $(\omega > 0)$ can be defined for all positive integers n, there is no longer any difficulty in employing the partial fraction expansion (3.14) for all Γ. How distributions of this type can be defined was shown by Güttinger and Pfaffelhuber.[9] It introduces an arbitrary parameter which, in physical interpretation where p is the momentum, has the dimensions of a length.

6. APPLICATIONS

In quantum electrodynamics the photon propagator in perturbation expansion is of the form

$$\left(g_{\mu\nu} - c\, \frac{k_\mu k_\nu}{k^2} \right) \frac{1}{k^2} \Big|_{C_{1R}}, \quad (6.1)$$

where the constant c depends on the choice of gauge. The factor $(1/k^2)^2|_{C_{1R}}$ is not defined, as we have just seen, but the presence of the $k_\mu k_\nu$ in the numerator avoids this difficulty. In fact, the argument at the end of the last section shows that even one factor k_μ would suffice to make the integral converge also in the $m \to 0$ limit, since it provides one more factor ω in the numerator.

The existence of $k_\mu/(k^2)^2$ for all contours is also essential in a gauge-independent formulation of quantum electrodynamics proposed recently.[10] There the generalized function ∂_μ^{-1} was defined as

$$(\partial_\mu^{-1}, \varphi)_\Gamma = \pm \frac{1}{4\pi^2 i} \int_\Gamma \partial_\mu^\nu \left(\frac{1}{y^2} \right) \varphi(x \pm y)\, d^4y, \quad (6.2)$$

where Γ is a suitable contour in the complex x^0 plane. This expression thus defines an analytic functional with $\varphi \in \mathcal{D}(R^3) \otimes \mathfrak{Z}$ as before.

It is defined for all Γ in the x^0 plane. The $\Gamma(x^0)$ which correspond to the contours of Table I were given in Ref. 10.

Another application of the results of Secs. 3–6 lies in asymptotic quantum field theory. Nonrenormalizable theories can be treated in perturbation expansion without encountering divergences by use of Δ_Γ^n $(n > 0)$ and the corresponding spin $\frac{1}{2}$ functions $S_\Gamma^n = (\gamma \cdot \partial + m)\Delta_\Gamma^n$. This was shown in a recent paper by Chen.[11]

Finally, this same formulation of field theory is seen to be applicable to particles of zero mass, because only those D_Γ^n and $\gamma \cdot \partial D_\Gamma^n$ occur in the kernel B which are defined for all n. It is B which determines the equations for the S-matrix elements.[12] The fact that the general formalism of asymptotic quantum field theory[12] involves only the analytic functionals $\Delta^n, \Delta_R^n, \Delta_A^n$, but not $\Delta_1^n, \Delta_{1R}^n, \Delta_{1A}^n$, in the equation for the S operator, is to be considered of noteworthy importance. It permits this theory to carry through also in the $m \to 0$ limit for arbitrary n.

---

[9] W. Güttinger, Fortschr. Physik **14**, 483 (1966). We are indebted to Dr. Güttinger for informing us of his work prior to publication.

[10] F. Rohrlich and F. Strocchi, Phys. Rev. **139**, B476 (1965).

[11] T.-W. Chen, Ann. Phys. (N.Y.) **42**, 476 (1967).

[12] For a survey see F. Rohrlich, "Asymptotic Quantum Field Theory" in the symposium volume *Perspectives in Modern Physics*, R. E. Marshak, Ed. (John Wiley & Sons, Inc., New York, 1966).

of New York, Stony Brook, for their hospitality during the period of the Summer Institute when part of this work was carried out. He also acknowledges critical comments from Dr. A. Boehm.

APPENDIX A

We want to prove Eqs. (3.19) and (3.20). One starts by noting that

$$
\begin{aligned}
\delta^{(n)}&(p^2 + m^2)\\
&= \delta^{(n)}(p_0^2 - \omega^2),\\
&= \left(\frac{1}{2p^0}\frac{d}{dp^0}\right)^n \frac{1}{2\omega}[\delta(\omega - p^0) + \delta(p^0 + \omega)],\\
&= \sum_{\nu=0}^{n-1} \frac{(-1)^\nu}{\nu!}\frac{(n + \nu - 1)!}{(n - \nu - 1)!}\frac{1}{2\omega(2p^0)^{n+\nu}}\\
&\quad \times [\delta^{(n-\nu)}(\omega - p^0) + \delta^{(n-\nu)}(p^0 + \omega)]. \quad \text{(A1)}
\end{aligned}
$$

This is easily proven by induction. One then proves the auxiliary result

$$
\begin{aligned}
\frac{\delta^{(m)}(x - \omega)}{x^n} &= \sum_{\nu=0}^{m}\nu!\binom{n + \nu - 1}{\nu}\binom{m}{\nu}\\
&\quad \times \frac{1}{\omega^{n+\nu}}\delta^{(m-\nu)}(x - \omega) \quad \text{(A2)}
\end{aligned}
$$

as follows:

$$
\begin{aligned}
\int &\frac{\delta^{(m)}(x - \omega)}{x^n}f(x)\,dx\\
&= (-1)^m\left(\frac{f(x)}{x^n}\right)^{(m)}_{x=\omega},\\
&= (-1)^m\sum_{\nu=0}^{m}\left(\frac{1}{x^n}\right)^{(\nu)}\binom{m}{\nu}f^{(m-\nu)}(x)\Big|_{x=\omega},\\
&= \sum_{\nu=0}^{m}\nu!\binom{m}{\nu}\binom{n + \nu - 1}{\nu}\frac{1}{\omega^{n+\nu}},\\
&\quad \times \int \delta^{(m-\nu)}(x - \omega)f(x)\,dx.
\end{aligned}
$$

We now apply (A2) to (A1). One obtains a double sum which can be rearranged:

$$
\sum_{\nu=0}^{n-1}\sum_{\nu'=\nu}^{n}F(\nu, \nu') = \sum_{\nu=0}^{n-1}F(\nu, n) + \sum_{\nu'=0}^{n-1}\sum_{\nu=0}^{\nu'}F(\nu, \nu'). \quad \text{(A3)}
$$

The sum over ν in the last term can be carried out, yielding $(n + \nu')/(2^{n+\nu'}\nu'!)$ and leaving

$$
\begin{aligned}
\sum_{\nu'=0}^{n-1}&\frac{1}{2\omega}\frac{(n + \nu' - 1)!}{(n - \nu')!}\frac{n + \nu'}{2^{n+\nu'}\nu'!}\frac{1}{\omega^{n+\nu'}}\\
&\times [\delta^{(n-\nu')}(\omega - p^0) + (-)^{n-\nu'}\delta^{(n-\nu')}(p^0 + \omega)]. \quad \text{(A4)}
\end{aligned}
$$

The first term on the right side of (A3) can also be summed:

$$
\begin{aligned}
\sum_{\nu=0}^{n-1}F(\nu, n) &= n!\binom{2n}{n}\frac{1}{(2\omega)^{2n+1}}\\
&\quad \times [\delta(\omega - p^0) + \delta(p^0 + \omega)].
\end{aligned}
$$

This is identical to a term with $\nu' = n$ in (A4), so that the sum in (A4) can be extended to n, yielding exactly the desired result (3.18).

The proof of (3.19) only requires the observation that

$$
\int \epsilon(p^0)\delta^{(n)}(\omega \mp p^0)f(p^0)\,dp^0 = \pm(-1)^n f^{(n)}(\pm\omega).
$$

APPENDIX B

We want to show that, despite the similarity of the partial fraction expansions (3.19) and (3.20) associated with $\tilde{\Delta}_1^n$ and $\tilde{\Delta}^n$, respectively, the corresponding $m = 0$ limits $\tilde{D}_1^n$ and $\tilde{D}^n$ for $n > 0$ exist only in the latter case but not in the former. The difference lies in the symmetry properties of these objects. $\tilde{D}^n$ has a symmetry which permits it to be written in terms of ω^2 rather than ω:

$$
\begin{aligned}
(\tilde{D}^n, \varphi) &= \left(\left(\frac{d}{dp_0^2}\right)^n\tilde{D}, \varphi\right),\\
&= \int d^3p\,h(\mathbf{p})\int dp^0 f(p^0)\left(\frac{d}{dp_0^2}\right)^n\\
&\quad \times \frac{i}{4\pi\omega}[\delta(p^0 - \omega) - \delta(p^0 + \omega)],\\
&= \frac{i}{4\pi}\int_0^\infty \omega\,d\omega\,\bar{h}(\omega)\int_{-\infty}^\infty dp^0 f(p^0)\\
&\quad \times \left(\frac{d}{dp_0^2}\right)^n[\delta(p^0 - \omega) - \delta(p^0 + \omega)], \quad \text{(B1)}
\end{aligned}
$$

where

$$
\bar{h}(\omega) \equiv \frac{1}{4\pi}\int d\Omega\,h(\mathbf{p}).
$$

Since the expression in square brackets is odd in p^0, only the odd part of $f(p^0)$ will contribute. Writing $\frac{1}{2}[f(p^0) - f(-p^0)] = p^0 g(p_0^2)$, we consider the analytic functional with $T(-p^0) = -T(p^0)$:

$$
\begin{aligned}
\int_{-\infty}^\infty &p^0\,dp^0 g(p_0^2)\frac{d}{dp_0^2}T(p^0)\\
&= \frac{1}{2}\int_{-\infty}^\infty dp^0 g(p_0^2)\frac{d}{dp^0}T(p^0),\\
&= -\frac{1}{2}\int_{-\infty}^\infty T(p^0)\frac{d}{dp^0}g(p_0^2)\,dp^0.
\end{aligned}
$$

If $f(p^0) \in \mathfrak{Z}$, then $f_1(p^0) = [f(p^0) - f(-p^0)]/(4p^0)$ is also $\in \mathfrak{Z}$, and the above equation can be written

$$((d/dp_0^2)T, f) = -(T, f_1'). \qquad (B2)$$

Application of this result n times to the analytic functional in (B1) shows that one obtains, using

$$f_n(p^0) = [f_{n-1}'(p^0) - f_{n-1}'(-p^0)]/(4p^0) \quad (n > 1),$$

$$(\tilde{D}^n, \varphi) = \frac{i}{4\pi} \int_0^\infty \omega \, d\omega \bar{h}(\omega)(-1)^n[f_n'(\omega) - f_n'(-\omega)],$$

$$(B3)$$

which is a well-defined integral since $f_n \in \mathfrak{Z}$.

An analogous argument for D_1^n fails because the corresponding $T(p^0)$ is symmetric rather than antisymmetric.

Canonical Formulation of Relativistic Mechanics

ROBERT NYDEN HILL

Department of Physics, University of Delaware, Newark, Delaware

Instantaneous action-at-a-distance relativistic mechanics of the type considered by Currie and by Hill is cast into a Hamiltonian form wherein the transformations of the inhomogeneous Lorentz group are canonical. The Currie–Jordan–Sudarshan zero-interaction theorem is circumvented by renouncing the demand that physical positions be canonical; the implications for measurement theory of renouncing this demand are discussed. Examples are given.

I. INTRODUCTION

THE fact that the physical world is basically quantum mechanical poses a problem for any classical relativistic action-at-a-distance particle mechanics; it must eventually be quantized. There are two alternative routes which can be followed in attempting to deal with the problem of quantization. One can either try to invent new quantization methods, as was done by Feynman[1] in attempting to quantize the action-at-a-distance version of classical electrodynamics which had been constructed by Wheeler and Feynman,[2] or one can try to formulate a relativistic Hamiltonian particle mechanics and proceed to quantization by conventional methods. The present paper is devoted primarily to the classical aspects of the second alternative; we discuss the construction of a classical relativistic Hamiltonian mechanics within the framework of the canonical representations of the inhomogeneous Lorentz group which preserves world-line invariance.

Before proceeding, it is in order to ask why anyone would want an action-at-a-distance theory (rather than a field theory) in the first place. One obvious answer is that infinite mass corrections can be

trivially avoided by deleting the divergent interaction of a particle with its own field. This, however, is not the only reason. A description of particle motions which employs a field as the intermediary carrying the interaction is most useful when the motion of the sources of the field can be prescribed, at least to a reasonable approximation. This criterion is satisfied only for distant collisions; it fails for close collisions and bound motions, for which action-at-a-distance would seem to be more useful. An additional reason for wanting not only an action-at-a-distance theory, but also an Hamiltonian theory of the same type as the nonrelativistic one, arises when one considers the problem of making relativistic corrections to non-relativistic theories. For example: The Darwin–Breit interaction, valid only for use as a first-order perturbation, is all that is presently known of the relativistic corrections to inter-electronic interactions which must be accounted for in calculating atomic energy levels. An understanding of how further corrections are to be made for such problems requires an understanding of relativistic Hamiltonian mechanics and of approximations thereto.

Single-time[3] relativistic Hamiltonian mechanics has

[1] R. P. Feynman, Phys. Today **19**, 31 (1966); Rev. Mod. Phys. **20**, 267 (1948).

[2] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945); **21**, 425 (1949).

[3] This is in contrast to manifest-covariant, many-time (one time for each particle) theories such as the Wheeler–Feynman theory (Ref. 2) and the recent action-at-a-distance relativistic mechanics of van Dam and Wigner [Phys. Rev. **138**, B1576 (1965); **142**, 838 (1966)], which contains the Wheeler–Feynman theory as a special case.

If $f(p^0) \in \mathfrak{Z}$, then $f_1(p^0) = [f(p^0) - f(-p^0)]/(4p^0)$ is also $\in \mathfrak{Z}$, and the above equation can be written

$$(\!(d/dp_0^2)T, f\!) = -(T, f_1').　　　　(B2)$$

Application of this result n times to the analytic functional in (B1) shows that one obtains, using

$$f_n(p^0) = [f_{n-1}'(p^0) - f_{n-1}'(-p^0)]/(4p^0) \quad (n > 1),$$

$$(\tilde{D}^n, \varphi) = \frac{i}{4\pi} \int_0^\infty \omega \, d\omega \bar{h}(\omega)(-1)^n [f_n'(\omega) - f_n'(-\omega)],$$

$$(B3)$$

which is a well-defined integral since $f_n \in \mathfrak{Z}$.

An analogous argument for D_1^n fails because the corresponding $T(p^0)$ is symmetric rather than anti-symmetric.

Canonical Formulation of Relativistic Mechanics

ROBERT NYDEN HILL

Department of Physics, University of Delaware, Newark, Delaware

(Received 30 December 1966)

Instantaneous action-at-a-distance relativistic mechanics of the type considered by Currie and by Hill is cast into a Hamiltonian form wherein the transformations of the inhomogeneous Lorentz group are canonical. The Currie–Jordan–Sudarshan zero-interaction theorem is circumvented by renouncing the demand that physical positions be canonical; the implications for measurement theory of renouncing this demand are discussed. Examples are given.

I. INTRODUCTION

THE fact that the physical world is basically quantum mechanical poses a problem for any classical relativistic action-at-a-distance particle mechanics; it must eventually be quantized. There are two alternative routes which can be followed in attempting to deal with the problem of quantization. One can either try to invent new quantization methods, as was done by Feynman[1] in attempting to quantize the action-at-a-distance version of classical electrodynamics which had been constructed by Wheeler and Feynman,[2] or one can try to formulate a relativistic Hamiltonian particle mechanics and proceed to quantization by conventional methods. The present paper is devoted primarily to the classical aspects of the second alternative; we discuss the construction of a classical relativistic Hamiltonian mechanics within the framework of the canonical representations of the inhomogeneous Lorentz group which preserves world-line invariance.

Before proceeding, it is in order to ask why anyone would want an action-at-a-distance theory (rather than a field theory) in the first place. One obvious answer is that infinite mass corrections can be

trivially avoided by deleting the divergent interaction of a particle with its own field. This, however, is not the only reason. A description of particle motions which employs a field as the intermediary carrying the interaction is most useful when the motion of the sources of the field can be prescribed, at least to a reasonable approximation. This criterion is satisfied only for distant collisions; it fails for close collisions and bound motions, for which action-at-a-distance would seem to be more useful. An additional reason for wanting not only an action-at-a-distance theory, but also an Hamiltonian theory of the same type as the nonrelativistic one, arises when one considers the problem of making relativistic corrections to non-relativistic theories. For example: The Darwin–Breit interaction, valid only for use as a first-order perturbation, is all that is presently known of the relativistic corrections to inter-electronic interactions which must be accounted for in calculating atomic energy levels. An understanding of how further corrections are to be made for such problems requires an understanding of relativistic Hamiltonian mechanics and of approximations thereto.

Single-time[3] relativistic Hamiltonian mechanics has

[1] R. P. Feynman, Phys. Today 19, 31 (1966); Rev. Mod. Phys. 20, 267 (1948).

[2] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945); 21, 425 (1949).

[3] This is in contrast to manifest-covariant, many-time (one time for each particle) theories such as the Wheeler–Feynman theory (Ref. 2) and the recent action-at-a-distance relativistic mechanics of van Dam and Wigner [Phys. Rev. 138, B1576 (1965); 142, 838 (1966)], which contains the Wheeler–Feynman theory as a special case.

been previously considered by Dirac, Thomas, Bakamjian, and Foldy.[4] Currie, Jordan, and Sudarshan[5,6] have shown that a relativistic two-particle Hamiltonian mechanics in which, (a) physical positions are canonical variables, (b) transformations of the inhomogeneous Lorentz group are canonical transformations, and (c) world-line invariance is demanded, is incompatible with interaction; their zero interaction theorem was extended to three particles by Cannon and Jordan,[7] and to N particles by Leutwyler.[8] The possibility of circumventing their zero-interaction theorem by dropping the requirement that positions be canonical has been pointed out by Kerner.[9] A Newtonian-like approach, yielding a single-time instantaneous action-at-a-distance relativistic mechanics has been made by Currie[10] and by Hill[11]; a simple example of such a theory has been given by Kerner.[12]

When one first contemplates relativistic Hamiltonian mechanics, one is faced with, and perhaps troubled by, the fact that interactions are instantaneous (because the equations of motion are ordinary differential equations). Inasmuch as the very notion of "instantaneous" is not relativistically invariant, how can such an instantaneous action-at-a-distance mechanics be compatible with special relativity? To answer this we first remark that a classical particle dynamics is just a concise description of the possible particle world lines contemplated by the theory. The only requirements placed on such a description of point particles by relativity are that: (a) a set of allowed particle world lines is allowed, no matter which inertial observer's coordinate system is used to describe them (world-line invariance), and (b) that the dynamical equations which provide this concise description—i.e., whose solutions are allowed particle world lines—have the same form in every inertial frame (form invariance of the equations of motion). If differential equations are used to provide this description, conditions (a) and (b) can be satisfied,[10,11] *if* the equations hold for all time so that instantaneous

initial data in one frame can be related to instantaneous initial data in another. When discussing transformations between inertial frames, we will, however, find it convenient to regard the physical states of a system as the solution curves of the equations of motion, rather than as the sets of initial data which imply these solution curves; thus we employ the notion of state sub-specie aeternatis rather than the (non-invariant) notion of instantaneous state frequently used with differential equations.

Previous approaches[4,5] to relativistic Hamiltonian mechanics have begun with the Poisson bracket relations characterizing the inhomogeneous Lorentz group; the task of formulating relativistic particle mechanics was regarded as a matter of constructing a representation of the ten infinitesimal generators of the inhomogeneous Lorentz group in terms of canonical coordinates and momenta. The question of physical interpretation—i.e., of the relation between canonical coordinates and momenta and physical particle positions and velocities—was answered by assuming that physical and canonical coordinates were, or at least could be, the same. The fact that such a physical interpretation is incompatible with world-line invariance was shown by the Currie–Jordan–Sudarshan zero-interaction theorem.[5]

We avoid such difficulties of interpretation by approaching the problem from the more primitive Newtonian level. We assume that an instantaneous action-at-a-distance mechanics of the type considered by Currie and by Hill has been given. The physical interpretation of the Hamiltonian dynamics, i.e., the relation between physical and canonical variables, will then be fixed by the way the Hamiltonian scheme is obtained from the Newtonian-like mechanics. The Currie–Jordan–Sudarshan zero-interaction theorem shows that one cannot hope to pass from the Newtonian level to a Hamiltonian scheme with Lorentz transformations canonical via a Lagrangian, as is done nonrelativistically.[13] Therefore, we approach

[4] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949); L. H. Thomas, Phys. Rev. **85**, 868 (1952); B. Bakamjian and L. H. Thomas, *ibid.* **92**, 1300 (1953); B. Bakamjian, *ibid.* **121**, 1849 (1961); L. L. Foldy, *ibid.* **122**, 275 (1961). The various approaches to special relativistic dynamics before 1965 have been reviewed by P. Havas in *Statistical Mechanics of Equilibrium and Non-Equilibrium*, J. Meizner, Ed. (North-Holland Publishing Company, Amsterdam, 1965), p. 1.
[5] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[6] D. G. Currie, J. Math. Phys. **4**, 1470 (1963).
[7] J. T. Cannon and T. F. Jordan, J. Math. Phys. **5**, 299 (1964).
[8] H. Leutwyler, Nuovo Cimento **37**, 556 (1965).
[9] E. H. Kerner, J. Math. Phys. **6**, 1218 (1965).
[10] D. G. Currie, Phys. Rev. **142**, 817 (1966).
[11] R. N. Hill, Bull. Am. Phys. Soc. **11**, 96 (1966); J. Math. Phys. **8**, 201 (1967).
[12] E. H. Kerner, Phys. Rev. Letters **16**, 667 (1966). See also D. G. Currie and T. F. Jordan, *ibid.* **16**, 1210 (1966).

[13] A Lagrangian may be found, at least in some cases (see preceding footnote), if one is willing to pay for having physical positions canonical by giving up the requirement that the transformations of the inhomogeneous Lorentz group are canonical transformations. However, if the transformations of the inhomogeneous Lorentz group are not canonical, equivalent observers will use canonically inequivalent Hamiltonian formulations. Canonically inequivalent Hamiltonian formulations lead to inequivalent quantum theories if standard quantization methods are applied [P. Havas, Bull. Am. Phys. Soc. **1**, 337 (1956); F. J. Kennedy, Jr., and E. H. Kerner, Am. J. Phys. **33**, 463 (1965; **34**, 271 (1966)]; thus the surrender of the requirement that the transformations of the inhomogeneous Lorentz group be canonical transformations for equivalent observers leads to inequivalent quantum theories for equivalent observers. The alternative we have chosen of surrendering the requirement that physical positions be canonical might appear at first sight to lead to an interference between position, measurements at spacelike separations in a quantum theory; however such an interpretation ignores the description of the measurement process appropriate to an instantaneous interaction theory as we point out in Sec. VI.

the Hamiltonian formulation via the more general route afforded by the Lie–Konigs theorem,[14] as suggested by Kerner.

Section II is devoted to a discussion of the approach to Hamiltonian dynamics via the Lie–Konigs theorem. In Sec. III we consider the infinitesimal invariance transformations of a system of differential equations, and investigate the conditions under which a subgroup of these invariance transformations will be canonical transformations in the Hamiltonian formulation obtained via the procedures of Sec. II.

Section IV applies the general considerations of Secs. II and III to relativistic dynamics. The transformations of the inhomogeneous Lorentz group are written down and the ten infinitesimal generators identified. The separation of external and internal motions by the introduction of a suitable canonical coordinate which is a generalization to relativistic dynamics of the nonrelativistic center of mass is discussed. Section V is devoted to examples. In Sec. VI a new proof, valid in one, two, or three dimensions, of the Currie–Jordan–Sudarshan zero-interaction theorem is presented to clarify the role of the assumption that physical coordinates can be canonical coordinates. The implications of the surrender of this requirement for position measurements in a quantum theory are briefly discussed.

II. THE HAMILTONIZATION OF A SYSTEM OF DIFFERENTIAL EQUATIONS

Lie and Konigs have shown that any even order system of differential equations can be cast into Hamiltonian form.[14] In this section a constructive procedure, brought out by Kerner,[9,15] for effecting this Hamiltonization is outlined. Expressions are found for Poisson brackets in terms of the original physical (rather than the canonical) variables.

We begin with a system of analytic[16] second-order differential equations $a_i = F_i (x_1, \cdots, x_n; v_1, \cdots, v_n; t)$ specifying particle accelerations as functions of position, velocity, and time. These are first rewritten as the analytic first-order system

$$dy_i/dy_0 = h_i(y_0, y_1, \cdots, y_{2n}), \quad i = 1, \cdots, 2n. \quad (1)$$

Here we have in mind that $y_0 = t$, $y_i = x_i$, $h_i = v_i = y_{i+n}$, and $h_{i+n} = F_i (y_1, \cdots, y_n; y_{n+1}, \cdots, y_{2n};$

[14] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Cambridge University Press, New York, 1960), p. 275.
[15] E. H. Kerner, Bull. Math. Biophys. **26**, 333 (1964).
[16] We restrict ourselves to equations for which the F_i are analytic functions in the neighborhood of physically realizable real values of their arguments in order that we might have classical existence and uniqueness theorems on differential equations available.

y_0) for $i = 1, \cdots, n$. We seek to derive Eqs. (1) from a variational principle of the form

$$\delta \int \left[\sum_{i=1}^{2n} U_i(dy_i/dy_0) + U_0 \right] dy_0 = 0, \quad (2)$$

wherein the $y_i(y_0)$, $i = 1, \cdots, 2n$ are to be independently varied. The Euler equations of (2) take the form $\sum_{j=0}^{2n} \Gamma_{ij}(dy_j/dy_0) = 0$ where the matrix Γ_{ij} is defined by

$$\Gamma_{ij} = (\partial U_i/\partial y_j) - (\partial U_j/\partial y_i). \quad (3)$$

In order that it be possible to solve the Euler equations for the derivatives (dy_j/dy_0), the $2n \times 2n$ matrix Γ_{ij} (with $i \neq 0$, $j \neq 0$) must be nonsingular. Solving these Euler equations for dy_j/dy_0 yields the specified equations (1) if the U_i satisfy the differential conditions

$$\sum_{j=0}^{2n} \Gamma_{ij} h_j = 0, \quad i = 0, 1, \cdots, 2n. \quad (4)$$

Here we have allowed $i = 0$ because the equation for $i = 0$ is a consequence of the other $2n$ equations.

Once Eqs. (4) have been solved to yield a set of U_i with nonsingular Γ_{ij} ($i \neq 0$, $j \neq 0$), and hence a variational principle of the form (2), a Hamiltonian formulation can be obtained by solving Pfaff's problem to reduce the differential form $\sum_{i=1}^{2n} U_i dy_i$ to $\sum_{k=1}^{n} P_k dQ_k$ (the fact that this can always be done is the principal result of Pfaff's classic memoir[17]). The Q_k and P_k are the canonical coordinates and momenta. The Hamiltonian is $H = -U_0$, and must be re-expressed in terms of the canonical variables P_k, Q_k obtained by solving Pfaff's problem.

A. Poisson Brackets

If we now think of the P_k, Q_k as functions of the y_i, the equality $\sum_{i=1}^{2n} U_i dy_i = \sum_{k=1}^{n} P_k dQ_k$ implies that

$$U_i = \sum_{k=1}^{n} P_k(\partial Q_k/\partial y_i). \quad (5)$$

By the use of (5) in (3), we discover that the matrix Γ_{ij} ($i \neq 0, j \neq 0$) is actually a Lagrange bracket

$$\Gamma_{ij} = \sum_{k=1}^{n} \left(\frac{\partial P_k}{\partial y_j} \frac{\partial Q_k}{\partial y_i} - \frac{\partial P_k}{\partial y_i} \frac{\partial Q_k}{\partial y_j} \right) = \{y_i, y_j\}. \quad (6)$$

The Lagrange and Poisson brackets of a set of $2n$ independent variables are reciprocal matrices. Hence,

$$\sum_{l=1}^{2n} \Gamma_{il}[y_j, y_l] = \delta_{ij}, \quad i \neq 0, \quad (7)$$

Since Γ_{il} is nonsingular, Eqs. (7) can be inverted to

[17] Pfaff's problem is discussed in considerable detail with an historical summary, by A. R. Forsyth, *Theory of Differential Equations* (Dover Publications, Inc., New York, 1959), Vol. 1.

yield the Poisson brackets $[y_j, y_i]$ as functions of the y_i. With these known, the Poisson brackets of any two functions Φ and Ψ of the y_i can be obtained from

$$[\Phi, \Psi] = \sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] \frac{\partial \Phi}{\partial y_i} \frac{\partial \Psi}{\partial y_j} \qquad (8)$$

without having to solve Pfaff's problem to re-express Φ and Ψ in terms of the Q_k and P_k. We note that (7), and the fact that $h_0 = 1$, can be used to write the solution of (4) for the h_j as

$$h_j = -\sum_{i=1}^{2n} [y_j, y_i] \Gamma_{0i}. \qquad (9)$$

Equation (9) becomes the usual expression $dy_j/dy_0 = [y_j, H]$ if the U_i are independent of y_0.

B. Solution of Pfaff's Problem

The previous results can be used to derive a method for the solution of Pfaff's problem which was first given by Clebsch.[18] For any quantity Φ, it follows from (5) and (8) that

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] U_i \frac{\partial \Phi}{\partial y_j} = \sum_{k=1}^{n} P_k [Q_k, \Phi]. \qquad (10)$$

Clebsch's results now follow by letting $\Phi = Q_l$ in (10), and $\Phi = Q_k, \Psi = Q_l$ in (8) to obtain

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] U_i \frac{\partial Q_l}{\partial y_j} = 0, \qquad (11a)$$

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] \frac{\partial Q_k}{\partial y_i} \frac{\partial Q_l}{\partial y_j} = 0. \qquad (11b)$$

Equations (11) provide a complete set of necessary and sufficient conditions for the determination of the Q_k. The conditions on the P_k follow from a similar application of (8) and (10):

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] U_i \frac{\partial P_l}{\partial y_j} = P_l, \qquad (12a)$$

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] \frac{\partial P_k}{\partial y_i} \frac{\partial P_l}{\partial y_j} = 0, \qquad (12b)$$

$$\sum_{i=1}^{2n} \sum_{j=1}^{2n} [y_i, y_j] \frac{\partial Q_k}{\partial y_i} \frac{\partial P_l}{\partial y_j} = \delta_{k,l}. \qquad (12c)$$

Systematic integration of (11) and (12) produces a solution of Pfaff's problem; different solutions are canonically equivalent.[19]

---

[18] A. R. Forsyth, Ref. 17, pp. 210–214.

[19] The canonical transformations connecting different solutions of Pfaff's problem are known (in the context of classical Hamiltonian dynamics) as Mathieu transformations (E. T. Whittaker, Ref. 14, p. 301). A method, due to Clebsch, for obtaining the general solution of a given Pfaff's problem from any particular solution is discussed by A. R. Forsyth, Ref. 17, pp. 194–197.

It is worth remarking at this point that there is, in general, more than one solution to (4) with non-singular Γ_{ij} (for example, replacement of Γ_{ij} by $C\Gamma_{ij}$, C a constant, produces a new solution). Since the Γ_{ij} are canonical invariants, different solutions yield *canonically inequivalent* Hamiltonizations of the original differential equations.

III. INVARIANCE TRANSFORMATIONS

In this section the structure of the infinitesimal transformations which leave the system (1) invariant is considered. We derive the condition that a particular transformation leaves (1) invariant and the condition that it be canonical. The conserved quantity associated with a canonical transformation is identified and the transformation generated by this conserved quantity is computed. The (canonical) invariant subgroup of transformations which leave individual solutions of (1) invariant is identified and the factor group obtained by decomposing the canonical transformations with respect to this invariant subgroup is shown to be identical with the canonical transformations (generated by the usual Poisson brackets) which leave the independent variable (the time) fixed. An investigation of the conditions under which the action can be made invariant under a subgroup of canonical transformations concludes this section.

A. Notation

We consider infinitesimal transformations

$$y_i \rightarrow y_i' = y_i + \epsilon g_i^\alpha \quad (y_0, y_1, \cdots, y_{2n}) \qquad (13)$$

which leave the system (1) invariant. Here the Greek α indexes the different transformations. The differential operator,

$$L_\alpha = \sum_{i=0}^{2n} g_i^\alpha \partial/\partial y_i, \qquad (14)$$

effects the transformation (13) on the argument of any function; the substantive derivative,

$$D \equiv \sum_{i=0}^{2n} h_i \partial/\partial y_i, \qquad (15)$$

effects an infinitesimal transformation along the solution curves of the system (1). If we permit the Greek letters indexing the transformations to run over the independent transformations generating a subgroup, the commutator of two such transformations is expressible as a sum of these independent infinitesimal transformations:

$$L_\beta g_i^\alpha - L_\alpha g_i^\beta = \sum_\gamma C_{\alpha\beta}^\gamma g_i^\gamma. \qquad (16)$$

The $C^\gamma_{\alpha\beta}$ are the structure constants of the Lie group formed by the subgroup.

B. Conditions for Invariance and Canonicity

The conditions that the transformations (13) leave Eqs. (1) invariant are obtained by inserting the transformation (13) into (1), expanding to first order in ϵ, and demanding that the coefficient of ϵ vanish. The conditions are

$$L_\alpha h_i - D g_i^\alpha + h_i D g_0^\alpha = 0. \quad (17)$$

A transformation is canonical in an Hamiltonian formulation if it adds an exact differential to the form $\sum_{k=1}^n P_k \, dQ_k - H \, dt$ which appears in Hamilton's principle. This is equivalent to adding an exact differential $d\Omega_\alpha$ to the form $\sum_{i=0}^{2n} U_i \, dy_i$ which appears in (2). By computing the effect of the transformation (13) on this form, we find that the condition that (13) be canonically represented is the existence of an Ω_α such that

$$L_\alpha U_i + \sum_{j=0}^{2n} U_j (\partial g_j^\alpha / \partial y_i) = \partial\Omega_\alpha / \partial y_i. \quad (18)$$

We introduce a quantity G_α defined by

$$G_\alpha \equiv \sum_{i=0}^{2n} g_i^\alpha U_i - \Omega_\alpha. \quad (19)$$

The condition (18) that (13) be canonical is then the existence of a G_α such that

$$\sum_{i=0}^{2n} g_i^\alpha \Gamma_{ij} = \partial G_\alpha / \partial y_j. \quad (20)$$

A solution G_α of (20) will exist if and only if the integrability conditions guaranteeing the equality of the mixed partial derivatives $\partial^2 G_\alpha / \partial y_j \partial y_k$ are satisfied. If we differentiate (20) with respect to y_k, require the result to be symmetric under the interchange $j \rightleftarrows k$, and use the integrability conditions

$$\frac{\partial \Gamma_{ij}}{\partial y_k} + \frac{\partial \Gamma_{jk}}{\partial y_i} + \frac{\partial \Gamma_{ki}}{\partial y_j} = 0 \quad (21)$$

on the Γ_{ij} [which guarantee the existence of a set of U_i such that (3) hold], we obtain

$$L_\alpha \Gamma_{jk} + \sum_{i=0}^{2n} \left(\frac{\partial g_i^\alpha}{\partial y_j} \Gamma_{ik} + \frac{\partial g_i^\alpha}{\partial y_k} \Gamma_{ji} \right) = 0, \quad (22)$$

as another statement of the conditions that (13) be canonical. The conditions (17) that the differential equations (1) be invariant under (13) follow easily from (22). By multiplying (22) by h_j, summing on j from 0 to $2n$, and using (4), we obtain

$$\sum_{i=0}^{2n} (D g_i^\alpha - L_\alpha h_i) \Gamma_{ik} = 0.$$

If we split off the $i = 0$ term and use (4), we obtain

$$\sum_{i=1}^{2n} (D g_i^\alpha - L_\alpha h_i - h_i D g_0^\alpha) \Gamma_{ik} = 0$$

from which (17) follows because the $2n \times 2n$ matrix $\Gamma_{ik}(i, k \neq 0)$ is nonsingular. The converse, however, does not hold; (22) is not a necessary consequence of (17).

We are now in a position to see the issues involved in casting a dynamics, originally specified as a set of differential equations in the physical coordinates and velocities, into Hamiltonian form with an M-dimensional subgroup of the invariance transformations canonically represented. We must find a set of Γ_{ij} such that (a) det $\Gamma_{ij} \neq 0$ $(i, j \neq 0)$, (b) (4) and (21) are satisfied, and (c) (22) is satisfied for each of the M-independent transformations specifying the subgroup.

C. Conservation Laws and Generators

If we multiply (20) by h_j, sum on j from 0 to $2n$, and use (4) we discover that G_α is conserved. This constitutes a proof of Noether's theorem in the present context[20]; G_α is the conserved quantity associated with the transformation (13). The association of the conserved quantity with the transformation depends on the way in which the dynamics is cast into Hamiltonian form [i.e., which solution of (4), (21), and (20) is taken]; canonically inequivalent Hamiltonizations [with (13) canonical] can lead to the association of a different conserved quantity with (13).

The conserved quantity associated with an invariance transformation frequently appears as the generator of that transformation. In order to investigate this we compute the Poisson bracket $[y_k, G_\alpha]$ of G_α with one of the variables y_k, $k \neq 0$. If we multiply (20) by $[y_k, y_j]$, sum on j from 1 to $2n$, and use (7), (8), and (9) we obtain

$$[y_k, G_\alpha] = g_k^\alpha - g_0^\alpha h_k. \quad (23)$$

The result (23) is correct for y_0 as well as for the y_k, $k \neq 0$ for which it was computed; it yields $[y_0, G_\alpha] = 0$ consistent with the fact that the Poisson bracket (8) does not transform the independent variable y_0. We note that G_α generates the transformation (13) with which it was associated only if $g_0^\alpha = 0$; if $g_0^\alpha \neq 0$ the formalism compensates for the inability of the Poisson bracket (8) to transform y_0 by shifting the other variables an amount $-\epsilon g_0^\alpha (dy_k / dy_0)$ along the solution curves of (1).

[20] A discussion of Noether's theorem is given by E. L. Hill, Rev. Mod. Phys. 23, 253 (1951). See also R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. 1, pp. 262–266.

D. Transformations Which Change Physical States

The result (23) can be better understood by taking a closer look at the group of invariance transformations of (1), which group we call $\mathcal{G}$. The subgroup of $\mathcal{G}$ that is canonically represented will be denoted by $\mathcal{G}_c$. We first note that there is a subgroup of $\mathcal{G}$, which we call $\mathcal{H}$, which leaves individual solutions invariant, i.e., which carries a solution $y_j = f_j(y_0)$ of (1) into $y'_j = f_j(y'_0)$ with the same f_j. This subgroup $\mathcal{H}$ is generated by transformations of the form

$$y_i \rightarrow y'_i = y_i + \epsilon\Phi(y_0, y_1, \cdots, y_{2n})h_i, \quad (24)$$

where Φ is an arbitrary function. The transformations (24) are always canonical, since (20) is satisfied (with $G = 0$). Furthermore, $\mathcal{H}$ is an invariant (normal) subgroup, since the commutator of a transformation of the form (13) with one of the form (24) is in $\mathcal{H}$: By direct computation, using (17), $\Phi Dg_i^\alpha - L_\alpha(\Phi h_i) = [\Phi Dg_0^\alpha - (L_\alpha\Phi)]h_i$ which is of the form (24). Thus we can decompose $\mathcal{G}$ (or $\mathcal{G}_c$) into cosets relative to $\mathcal{H}$ and consider the factor groups (quotient groups) $\mathcal{G}/\mathcal{H}$ and $\mathcal{G}_c/\mathcal{H}$. Each element of a given coset has the same effect on a solution curve $y_i = f_j(y_0)$ of (1). Thus if we identify physical states with the solution curves (i.e., think of the physical state as a state *sub-specie aeternatis* rather than an instantaneous state), it is the transformations of the factor groups $\mathcal{G}/\mathcal{H}$ and $\mathcal{G}_c/\mathcal{H}$ which change the physical state.

Each coset in the decomposition relative to $\mathcal{H}$ contains one and only one transformation which leaves y_0 fixed; thus, the transformations which leave y_0 fixed provide faithful representations of $\mathcal{G}/\mathcal{H}$ and $\mathcal{G}_c/\mathcal{H}$. We can now see the significance of the result (23): By using the G_α associated with a transformation g_i^α as a generator [with the Poisson bracket (8)] the homomorphic mapping of $\mathcal{G}_c$ onto the faithful representation of $\mathcal{G}_c/\mathcal{H}$ by canonical transformations which leave y_0 fixed is automatically accomplished. Commutators in $\mathcal{G}_c$ are mapped onto Poisson brackets under this homomorphism: By direct computation using (8), (17), (23), and the Jacobi identity we obtain

$$L_\beta g_i^\alpha - L_\alpha g_i^\beta = [y_i, [G_\alpha, G_\beta]] + (L_\beta g_0^\alpha - L_\alpha g_0^\beta)h_i,$$

from which the use of (16) and (23) produces

$$[y_i, [G_\alpha, G_\beta]] = \sum_\gamma C_{\alpha\beta}^\gamma [y_i, G_\gamma]. \quad (25)$$

The preceding analysis helps to show why Hamiltonian dynamics with the usual Poisson brackets, *which do not transform the time*, can comprehend relativistic particle mechanics *despite the lack of an absolute time*: The canonical transformations generated

with these brackets provide a faithful representation of the group $\mathcal{G}_c/\mathcal{H}$ of canonical transformations which change physical states.[21]

E. Transformations Which Leave the Action Invariant

A garden variety canonical transformation adds an exact differential $d\Omega$ to the form $\sum_{i=0}^{2n} U_i\, dy_i$ appearing in the variational principle (2) and changes the action by $\int d\Omega$; only those transformations for which $d\Omega = 0$ leave the action invariant. We now inquire as to the conditions under which a canonical transformation can be found which makes the action ·invariant under a specified subgroup of the canonical transformations. Such invariance of the action is seen in Sec. IV to correspond to the usual notion of manifest invariance for translations and spatial rotations.

We begin by supposing that $\Omega_\alpha \neq 0$ in (18) for at least some of the transformations which generate a subgroup of canonical transformations under which the action is to be made invariant. If there exists a function $\Lambda (y_0, y_1, \cdots y_{2n})$ such that

$$L_\alpha\Lambda = -\Omega_\alpha \quad (26)$$

for each of the independent infinitesimal transformations generating the subgroup, the action can be made invariant under the transformations of the subgroup by adding $d\Lambda$ to the form $\sum_{i=0}^{2n} U_i\, dy_i$. Such an addition effects the canonical transformation

$$U_i \rightarrow U'_i = U_i + \partial\Lambda/\partial y_i. \quad (27)$$

It follows from direct substitution and the use of (26) that U'_i satisfies (18) with zero on the right-hand side.

A solution Λ to (26) exists if and only if the integrability conditions $T_{\alpha\beta} = 0$ are satisfied where $T_{\alpha\beta}$ is defined by

$$T_{\alpha\beta} = L_\beta\Omega_\alpha - L_\alpha\Omega_\beta - \sum_\gamma C_{\alpha\beta}^\gamma \Omega_\gamma. \quad (28)$$

Here the $C_{\alpha\beta}^\gamma$ are the structure constants for the subgroup as defined by (16). $T_{\alpha\beta}$ can be computed by multiplying (18) by g_i^β and summing on i from 0 to $2n$ to obtain $L_\beta\Omega_\alpha$, subtracting the corresponding expression for $L_\alpha\Omega_\beta$, and using (16) and (19) to obtain

$$T_{\alpha\beta} = \sum_{i=0}^{2n} (g_i^\beta L_\alpha U_i - g_i^\alpha L_\beta U_i) + \sum_\gamma C_{\alpha\beta}^\gamma G_\gamma. \quad (29)$$

---

[21] By working only with transformations which change physical states (which is possible because we do not have redundant quantities, such as the separate time coordinates for each particle which appear in manifest covariant formulations) we avoid the constraints whose presence causes difficulties in the quantization of classical Hamiltonian theories [these difficulties are discussed by P. A. M. Dirac in *Lectures on Quantum Mechanics* (Belfer Graduate School of Science Monographs Series, Yeshiva University, New York, 1964)].

The first term of (29) may be rewritten as

$$\sum_{i,j} (g_i^\beta g_j^\alpha - g_i^\alpha g_j^\beta) \partial U_i / \partial y_j = \sum_{i,j} g_i^\beta \Gamma_{ij} g_j^\alpha.$$

The use of (8), (20), and (23) then produces the result

$$T_{\alpha\beta} = -[G_\alpha, G_\beta] + \sum_\gamma C_{\alpha\beta}^\gamma G_\gamma. \tag{30}$$

It follows from (25) that $[G_\alpha, G_\beta] = \sum_\gamma C_{\alpha\beta}^\gamma G_\gamma + d_{\alpha\beta}$ where the $d_{\alpha\beta}$ are constants. Hence, the integrability conditions $T_{\alpha\beta} = 0$ can be satisfied if and only if the $d_{\alpha\beta}$ can be made zero by adding constants to the conserved quantities G_α. Whether or not the $d_{\alpha\beta}$ can all be made zero depends on the structure constants $C_{\alpha\beta}^\gamma$ specifying the subgroup. For the inhomogeneous Lorentz group this can always be accomplished.[22] For the inhomogeneous Galelei group, the $d_{\alpha\beta}$ cannot always all be made zero; this is why $d\Omega \neq 0$ in the usual application of Noether's theorem to extract the center-of-mass theorem of nonrelativistic mechanics.[23]

IV. APPLICATION TO RELATIVISTIC MECHANICS

In this section the g_i^α specifying the transformations of the inhomogeneous Lorentz group are obtained and their commutators computed. The conserved quantities which generate these transformations are identified. The section concludes with the discussion of a suitable generalization of the nonrelativistic center of mass.

A. Notation

It is convenient at this point to change to a vector notation. The formulas of the previous sections continue to hold if we let each of the y_i, $i \neq 0$ become a 3-vector; $y_0 = t$ remains a scalar. The original dynamics is now

$$\mathbf{a}_i = \mathbf{F}_i(\mathbf{x}_1, \cdots, \mathbf{x}_n; \mathbf{v}_1, \cdots, \mathbf{v}_n).$$

The first-order rewrite of the dynamics becomes $d\mathbf{y}_i / dy_0 = \mathbf{h}_i$, the differential form in (2) becomes $\sum_{i=1}^{2n} \mathbf{U} \cdot d\mathbf{y}_i + U_0 \, dy_0$, and each element of Γ_{ij} becomes the dyadic

$$\Gamma_{ij} = \mathbf{U}_i \overleftarrow{(\partial/\partial \mathbf{y}_j)} - \overrightarrow{(\partial/\partial \mathbf{y}_i)} \mathbf{U}_j,$$

where the arrow indicates the direction in which the gradient $\partial/\partial \mathbf{y}$ operates.

B. Transformations of the Inhomogeneous Lorentz Group

The identification of the g_i^α [in accordance with (13)] which specify the transformations of the inhomo-

[22] For a proof of this statement, as well as further references, see Currie, Jordan, and Sudarshan, Ref. 5., Appendix B.
[23] E. L. Hill, Ref. 20, Sec. VII, Part D.

geneous Lorentz group is most easily done from the passive viewpoint. Thus, we define time translation by $y_0 \rightarrow y_0' = y_0 - \epsilon$, $\mathbf{y}_i \rightarrow \mathbf{y}_i' = \mathbf{y}_i$ to obtain

$$g_0^{TT} = -1, \quad \mathbf{g}_i^{TT} = 0, \quad i = 1, 2, \cdots, 2n. \tag{31}$$

Here superscripts denote, in an obvious notation, the different transformations. We note that the generator of time translations in the canonical formalism generates not (31), but instead, the member of the coset (in the decomposition of $\mathcal{G}_c$ relative to $\mathcal{K}$) to which (31) belongs for which $g_0 = 0$, namely $y_0 \rightarrow y_0' = y_0$, $\mathbf{y}_i \rightarrow \mathbf{y}_i' = \mathbf{y}_i + \epsilon \mathbf{h}_i$. The g_i for this *canonical time translation* are therefore

$$g_0^{CTT} = 0, \quad \mathbf{g}_i^{CTT} = \mathbf{h}_i, \quad i = 1, 2, \cdots, 2n. \tag{32}$$

We recall that the first n $\mathbf{y}_i$ are the physical position coordinates; thus, space translation in the direction of the unit vector $\hat{\mathbf{l}}$ is defined by $y_0 \rightarrow y_0' = y_0$, $\mathbf{y}_i \rightarrow \mathbf{y}_i' = \mathbf{y}_i + \epsilon \hat{\mathbf{l}}$, $\mathbf{y}_{i+n} \rightarrow \mathbf{y}_{i+n}' = \mathbf{y}_{i+n}$, $i = 1, 2, \cdots, n$. Hence, with the unit vectors specifying direction indicated in parentheses,

$$g_0^{ST}(\hat{\mathbf{l}}) = 0 \quad \mathbf{g}_i^{ST}(\hat{\mathbf{l}}) = \hat{\mathbf{l}}, \quad \mathbf{g}_{i+n}^{ST}(\hat{\mathbf{l}}) = 0,$$
$$i = 1, 2, \cdots, n. \tag{33}$$

Similarly spatial rotations about an axis specified by the unit vector $\hat{\mathbf{m}}$ are defined by $y_0 \rightarrow y_0' = y_0$, $\mathbf{y}_i \rightarrow \mathbf{y}_i' = \mathbf{y}_i + \epsilon \hat{\mathbf{m}} \times \mathbf{y}_i$, $i = 1, 2, \cdots, 2n$ from which

$$g_0^{SR}(\hat{\mathbf{m}}) = 0, \quad \mathbf{g}_i^{SR}(\hat{\mathbf{m}}) = \hat{\mathbf{m}} \times \mathbf{y}_i, \quad i = 1, 2, \cdots, 2n. \tag{34}$$

We turn next to the pure Lorentz transformations, which we also approach from the passive viewpoint. We consider two frames S and S' which move with a relative velocity $\mathbf{v} = \hat{\mathbf{n}} \tan \beta$. We suppose that in S, a set of particle world lines which form a solution of the dynamical equations (i.e., a physical state) is described by $\mathbf{x}_i = \mathbf{f}_i(t)$ and that this same set of world lines (same physical state) is described as $\mathbf{x}_i' = \mathbf{f}_i'(t')$ in S'. The coset (relative to $\mathcal{K}$) of transformations to be associated with the Lorentz transformation from S to S' consists of all transformations $\mathbf{x}_i \rightarrow \mathbf{x}_i'$, $\mathbf{v}_i \rightarrow \mathbf{v}_i'$, $t \rightarrow t'$ which carry $\mathbf{x}_i = \mathbf{f}_i(t)$ onto $\mathbf{x}_i' = \mathbf{f}_i'(t')$. In Fig. 1, we picture the projection $[\hat{\mathbf{n}} \cdot \mathbf{x}_i = \hat{\mathbf{n}} \cdot \mathbf{f}_i(t)$ in $S]$ of one of these world lines in the direction of relative motion of the two frames. The intersection of the equilateral hyperbola $t^2 - (\hat{\mathbf{n}} \cdot \mathbf{x})^2 = t_0^2$ with the t' axis marks off the time to which the number t_0 is assigned in S'. The element of the coset associated with the Lorentz transformation from S to S' which leaves the independent variable (the time) fixed maps world points corresponding to the same numerical value of the time in their respective frames onto each other.
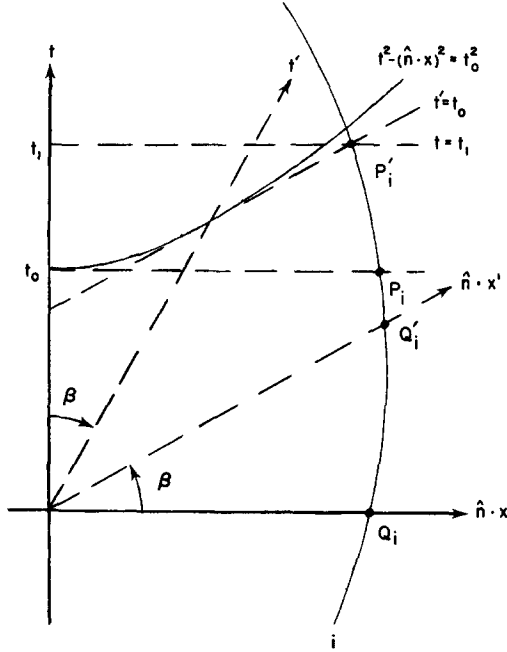
FIG. 1. Lorentz transformation which leaves the time fixed.

Thus the point Q_i [which is assigned coordinates $x_i = f_i(0)$, $t = 0$ in S] is mapped onto Q'_i [which is assigned coordinates $x'_i = f'_i(0)$, $t' = 0$ in S'] under this element of the coset. Similarly, P'_i, with coordinates $x_i = f_i(t_0)$, $t = t_0$ in S is mapped onto P'_i with coordinates $x'_i = f'_i(t_0)$, $t' = t_0$ in S'. Because this transformation involves a shift along the world lines, specification of the *finite* Lorentz transformation would require solution of the dynamical equations to obtain the coordinates $x_i = f_i(t_1)$, $t = t_1$ of P'_i in S. These equations can be trivially solved, however, for the infinitesimal transformation. If now we replace β by the infinitesimal ϵ, so that the relative velocity of the two frames is $\epsilon \hat{n}$, we see that $t_1 - t_0 = \epsilon \hat{n} \cdot x_i(P_i)$. To lowest order in ϵ, then,

$$x'_i(P'_i) = x_i(P_i) + \epsilon[v_i(P_i)\hat{n} \cdot x_i(P_i) - t_0]$$

and

$$v'_i(P'_i) = v_i(P_i) + \epsilon[a_i(P_i)\hat{n} \cdot x_i(P_i)$$
$$- \hat{n} + v_i(P_i)\hat{n} \cdot v_i(P_i)].$$

Expressed in the variables y_0, y_i this reads $y_0 \rightarrow y'_0$,

$$y_i \rightarrow y'_i = y_i + \epsilon[\hat{n}(\hat{n} \cdot y_i) - y_0\hat{n}],$$

$$y_{i+n} \rightarrow y'_{i+n} = y_{i+n} + \epsilon[h_{i+n}(\hat{n} \cdot y_i) - \hat{n} + y_{i+n}(\hat{n} \cdot y_{i+n})],$$
$$i = 1, 2, \cdots, n$$

from which

$$g_0^{\mathrm{LT}}(\hat{n}) = 0, \quad g_i^{\mathrm{LT}}(\hat{n}) = h_i(\hat{n} \cdot y_i) - y_0\hat{n},$$
$$g_{i+n}^{\mathrm{LT}}(\hat{n}) = h_{i+n}(\hat{n} \cdot y_i) - \hat{n} + y_{i+n}(\hat{n} \cdot y_{i+n}), \quad (35)$$
$$i = 1, 2, \cdots, n.$$

To complete the discussion of the Lorentz transformation, we mention briefly another element of the coset (relative to $\mathcal{K}$) which is associated with the Lorentz transformation. For some purposes, it is convenient to have world points on the world line of one of the particles—say the kth—remain unshifted.[24] This transformation is indicated in Fig. 2; in this case the point P_k stays where it is, but the world points on all of the other particle's world-lines shift; thus P_j shifts to P'_j. For the infinitesimal version of this transformation we have

$$t \rightarrow t' = t - \epsilon \hat{n} \cdot x_k,$$
$$x_j \rightarrow x'_j = x_j + \epsilon[v_j\hat{n} \cdot (x_j - x_k) - t\hat{n}],$$

and

$$v_j \rightarrow v'_j = v_j + \epsilon[F_j\hat{n} \cdot (x_j - x_k) - \hat{n} + v_jv_j \cdot \hat{n}].$$

We see that this transformation differs from (35) by an element of $\mathcal{K}$ [the element for which $\Phi = x_k \cdot \hat{n}$ in (24)] in accordance with the discussion of Sec. III.

In order to verify that the transformations (32)–(35) which leave y_0 fixed form a representation of the inhomogeneous Lorentz group, we compute their commutators. We find, with the unit vectors on which the transformations depend indicated in parentheses,

$$L_{\mathrm{ST}}(\hat{i})g_i^{\mathrm{ST}}(\hat{i}') - L_{\mathrm{ST}}(\hat{i}')g_i^{\mathrm{ST}}(\hat{i}) = 0, \quad (36a)$$

$$L_{\mathrm{CTT}}g_i^{\mathrm{ST}}(\hat{i}) - L_{\mathrm{ST}}(\hat{i})g_i^{\mathrm{CTT}} = 0, \quad (36b)$$

$$L_{\mathrm{SR}}(\hat{m})g_i^{\mathrm{SR}}(\hat{m}') - L_{\mathrm{SR}}(\hat{m}')g_i^{\mathrm{SR}}(\hat{m}) = -g_i^{\mathrm{SR}}(\hat{m} \times \hat{m}'),$$
$$(36c)$$

$$L_{\mathrm{SR}}(\hat{m})g_i^{\mathrm{CTT}} - L_{\mathrm{CTT}}g_i^{\mathrm{SR}}(\hat{m}) = 0, \quad (36d)$$

$$L_{\mathrm{SR}}(\hat{m})g_i^{\mathrm{ST}}(\hat{i}) - L_{\mathrm{ST}}(\hat{i})g_i^{\mathrm{SR}}(\hat{m}) = -g_i^{\mathrm{ST}}(\hat{m} \times \hat{i}), \quad (36e)$$

$$L_{\mathrm{LT}}(\hat{n})g_i^{\mathrm{LT}}(\hat{n}') - L_{\mathrm{LT}}(\hat{n}')g_i^{\mathrm{LT}}(\hat{n}) = g_i^{\mathrm{SR}}(\hat{n} \times \hat{n}'), \quad (36f)$$

$$L_{\mathrm{LT}}(\hat{n})g_i^{\mathrm{CTT}} - L_{\mathrm{CTT}}g_i^{\mathrm{LT}}(\hat{n}) = -g_i^{\mathrm{ST}}(\hat{n}), \quad (36g)$$

$$L_{\mathrm{LT}}(\hat{n})g_i^{\mathrm{ST}}(\hat{i}) - L_{\mathrm{ST}}(\hat{i})g_i^{\mathrm{LT}}(\hat{n}) = -(\hat{n} \cdot \hat{i})g_i^{\mathrm{CTT}}, \quad (36h)$$

$$L_{\mathrm{LT}}(\hat{n})g_i^{\mathrm{SR}}(\hat{m}) - L_{\mathrm{SR}}(\hat{m})g_i^{\mathrm{LT}}(\hat{n}) = -g_i^{\mathrm{LT}}(\hat{n} \times \hat{m}).$$
$$(36i)$$

These may be verified by using the definition (14) of L_α and [in the case of (36d, f, g, and i)] the statement (17) of invariance of the original differential equations. These transformations which leave the time fixed are the ones generated by the conserved quantities in the canonical formalism; if we assign the letters H, $\mathbf{P}$, $\mathbf{J}$, $\mathbf{K}$ to the generators of time translations, space translations, spatial rotations, and pure Lorentz

---

[24] This member of the coset was employed by the author (R. N. Hill, Ref. 11) in the derivation of a pair of integro-differential equations for the forces of an instantaneous action-at-a-distance description of two charges interacting via electrodynamic forces.
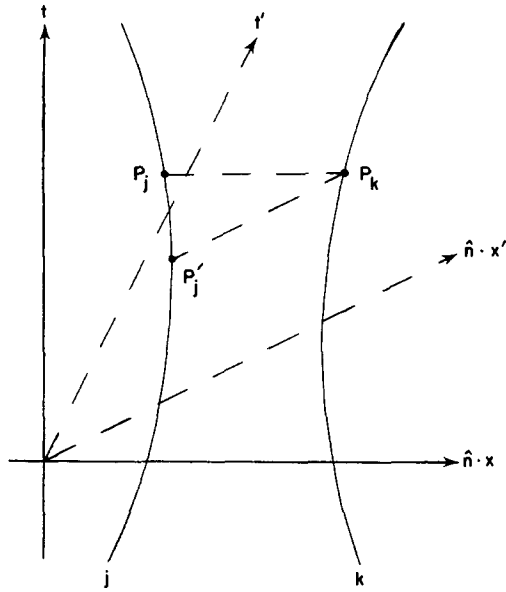
FIG. 2. Lorentz transformation which leaves points on the world line of the kth particle unshifted.

transformations, respectively, we have the Poisson bracket equivalent of (36) (here the Poisson bracket of two vectors **A**, **B** is taken to be the dyadic $[\mathbf{A}, \mathbf{B}] \equiv \sum_{i,j} \hat{\mathbf{n}}_i \hat{\mathbf{n}}_j [A_i, B_j]$ where the $\hat{\mathbf{n}}_i$ are the basis vectors **i**, **j**, **k**, of Cartesian coordinates):

$$[\mathbf{P}, \mathbf{P}] = 0, \tag{37a}$$

$$[\mathbf{P}, H] = 0, \tag{37b}$$

$$[\mathbf{J}, \mathbf{J}] = -\mathbf{I} \times \mathbf{J}, \tag{37c}$$

$$[H, \mathbf{J}] = 0, \tag{37d}$$

$$[\mathbf{P}, \mathbf{J}] = -\mathbf{I} \times \mathbf{P}, \tag{37e}$$

$$[\mathbf{K}, \mathbf{K}] = \mathbf{I} \times \mathbf{J}, \tag{37f}$$

$$[H, \mathbf{K}] = -\mathbf{P}, \tag{37g}$$

$$[\mathbf{P}, \mathbf{K}] = -\mathbf{I}H, \tag{37h}$$

$$[\mathbf{J}, \mathbf{K}] = -\mathbf{I} \times \mathbf{K}. \tag{37i}$$

Here the letters (a)–(i) indicate the correspondence between Eqs. (36) and (37); in Eq. (37) it is assumed that neutral elements have been eliminated by adding suitable constants to the generators, if necessary.

C. Conserved Quantities and Generators

The conserved quantities H, **P**, **J**, and **K** which generate the transformations of the inhomogeneous Lorentz group can be identified from (19) if Ω_α is known. These conserved quantities are all canonical invariants in the sense that their dependence on the physical positions, velocities, and time is unchanged

by adding an exact differential to the form $\sum U_i \, dy_i$ in (2). This is clear from (20) when one notes that the Lagrange bracket Γ_{ij} is a canonical invariant while the relation (4) guarantees that all transformations in the same coset in the decomposition of $\mathcal{G}_c$ relative to $\mathcal{H}$ produce the same $\partial \mathcal{G}_\alpha / \partial y_j$. The Ω_α, U_0, and the U_i, however, are not canonical invariants. In fact, the considerations concerning invariance of the action at the end of Sec. III can be exploited to make all of the Ω_α for the inhomogeneous Lorentz group zero if a set of g_i^α which form a representation of the inhomogeneous Lorentz group (i.e., which satisfy the commutation relations (16) with structure constants $C_{\alpha\beta}^\gamma$ appropriate to the inhomogeneous Lorentz group) is used. The transformations (32)–(35) which leave the time fixed constitute such a set, as evidenced by (36). However, the use of (32) as the representative element of the coset of time translations (for which Ω_{CTT} is to be made zero) would lead to the identification of $\sum_{i=1}^{2n} \mathbf{h}_i \cdot \mathbf{U}_i$ rather than the Hamiltonian-U_0 as the generator of time translations (it would also lead to time-dependent U_0 and $\mathbf{U}_i$). If the Hamiltonian-U_0 is to be the generator of time translations, we must, instead, use the g_i^{TT} of (31) as the representative element of the coset of time translations. The replacement of the g_i^{CTT} of (32) by the g_i^{TT} of (31) will not, however, yield a representation of the inhomogeneous Lorentz group without a compensating modification in the generators of the pure Lorentz transformations, as is clear from (36h). What must be done is to replace $g_0^{LT}(\hat{\mathbf{n}})$ and the $\mathbf{g}_i^{LT}(\hat{\mathbf{n}})$ by another element of the coset to which they belong, i.e., by the "modified Lorentz transformation"

$$\begin{aligned} g_0^{MLT}(\hat{\mathbf{n}}) &= \Phi(\hat{\mathbf{n}}), \\ \mathbf{g}_i^{MLT}(\hat{\mathbf{n}}) &= \mathbf{g}_i^{LT}(\hat{\mathbf{n}}) + \Phi(\hat{\mathbf{n}})\mathbf{h}_i, \end{aligned} \tag{38}$$

where $\Phi(\hat{\mathbf{n}})$ depends *a priori* on y_0, $\hat{\mathbf{n}}$, and the $\mathbf{y}_i$. If we replace L_{CTT}, g_i^{CTT} by L_{TT}, $\mathbf{g}_i^{TT}$ and L_{LT}, $\mathbf{g}_i^{LT}$ by L_{MTL}, g_i^{MLT} in Eq. (36), we obtain [from (f), (g), (h), and (i)] conditions on $\Phi(\hat{\mathbf{n}})$ whose satisfaction guarantees that (31), (33), (34), and (38) form a representation of the inhomogeneous Lorentz group. These conditions, obtained with the aid of (17), are

$$[L_{LT}(\hat{\mathbf{n}}) + \Phi(\hat{\mathbf{n}})D]\Phi(\hat{\mathbf{n}}')$$
$$- [L_{LT}(\hat{\mathbf{n}}') + \Phi(\hat{\mathbf{n}}')D]\Phi(\hat{\mathbf{n}}) = 0, \tag{39f}$$

$$\frac{\partial \Phi(\hat{\mathbf{n}})}{\partial y_0} = 0, \tag{39g}$$

$$L_{ST}(\hat{\mathbf{l}})\Phi(\hat{\mathbf{n}}) = -\hat{\mathbf{n}} \cdot \hat{\mathbf{l}}, \tag{39h}$$

$$L_{SR}(\hat{\mathbf{m}})\Phi(\hat{\mathbf{n}}) = \Phi(\hat{\mathbf{n}} \times \hat{\mathbf{m}}). \tag{39i}$$

[In Eq. (39) conditions are labelled (f)–(i) to indicate the parts of (36) from which they come.] A possible solution of (39) is

$$\Phi(\hat{n}) = -\hat{n} \cdot \mathbf{Q}, \qquad (40)$$

where $\mathbf{Q}$ transforms like a physical coordinate,[25] i.e., $\mathbf{Q}$ satisfies

$$[\mathbf{Q}, \mathbf{K}] = [\mathbf{Q}, H]\mathbf{Q} - \mathbf{I}t, \qquad (41\text{f})$$

$$\partial \mathbf{Q}/\partial t = 0, \qquad (41\text{f, g})$$

$$[\mathbf{Q}, \mathbf{P}] = \mathbf{I}, \qquad (41\text{h})$$

$$[\mathbf{Q}, \mathbf{J}] = -\mathbf{I} \times \mathbf{Q}. \qquad (41\text{i})$$

The letters (t)–(i) indicate which of the conditions (39) each of the conditions (41) is sufficient to guarantee; the Poisson brackets and generators have been used to effect the transformations. The modified Lorentz transformation (38) [with (40) and (41) satisfied] is that member of the coset associated with the Lorentz transformations which leaves points on the invariant world line traced out by $\mathbf{Q}$ unshifted. If $\mathbf{Q}$ were chosen to be the kth physical particle coordinate $\mathbf{x}_k$, we would have the member of the coset indicated in Fig. 2. Subsequently we will find it advantageous to let $\mathbf{Q}$ be the center of inertia.

Since the g_i^α of (31), (33), (34), and (38) [with (40) and (41) satisfied] form a representation of the inhomogeneous Lorentz group, the Ω_α can all be made zero with this set of g_i^α. The conserved quantities H, $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ can then be identified from (19); they are:

$$H = -U_0, \quad \mathbf{P} = \sum_{i=1}^{n} \mathbf{U}_i, \quad \mathbf{J} = \sum_{i=1}^{2n} \mathbf{y}_i \times \mathbf{U}_i,$$

$$\hat{n} \cdot \mathbf{K} = \sum_{i=1}^{2n} \mathbf{g}_i^{\mathrm{LT}}(\hat{n}) \cdot \mathbf{U}_i - \hat{n} \cdot \mathbf{Q}\left[U_0 + \sum_{i=1}^{2n} \mathbf{h}_i \cdot \mathbf{U}_i \right]. \qquad (42)$$

We note that the expressions for H, $\mathbf{P}$, and $\mathbf{J}$ are consistent with the usual expressions for nonrelativistic systems for which $\mathbf{U}_i = \mathbf{P}_i$, $\mathbf{U}_{i+n} = 0$, $i = 1, 2, \cdots, n$, where the $\mathbf{p}_i$ are the mechanical particle momenta. By writing out the conditions (18) for the translations and spatial rotations [with $\Omega_\alpha = 0$

and the g_i^α of (31), (33), and (34)], we obtain:

$$\frac{\partial U_0}{\partial y_0} = \frac{\partial \mathbf{U}_i}{\partial y_0} = 0, \quad \sum_{j=1}^{n} \frac{\partial}{\partial \mathbf{y}_j} U_0 = \sum_{j=1}^{n} \frac{\partial}{\partial \mathbf{y}_j} \mathbf{U}_i = 0,$$

$$\sum_{j=1}^{2n} \left(\mathbf{y}_j \times \frac{\partial}{\partial \mathbf{y}_j} \right) U_0 = 0, \qquad (43)$$

and

$$\sum_{j=1}^{2n} \left(\mathbf{y}_j \times \frac{\partial}{\partial \mathbf{y}_j} \right) \mathbf{U}_i - \mathbf{I} \times \mathbf{U}_i = 0,$$

which are the usual statements of manifest invariance under these transformations, namely that U_0 and the $\mathbf{U}_i$ are independent of time, invariant under space translation, and transform, respectively, as scalar and vector functions of the $\mathbf{y}_i$ under spatial rotations. With respect to the Lorentz transformations, the term "manifest invariance" has been traditionally reserved for formulations employing spinorial and tensorial representations of the inhomogeneous Lorentz group. Since we do not have a full complement of four-vectors, the interpretation of (18) for the effect of the Lorentz transformations (using the g_i^{MLT} with $\Omega_{\mathrm{MLT}} = 0$) on the $\mathbf{U}_i$ is unfamiliar, although it can be used to give an alternative demonstration of the transformation properties (37) of the generators, which do make up a four-vector and an antisymmetric four-tensor. In any event, the requirement that U_0 and the $\mathbf{U}_i$ be so adjusted that the generators take the form (42) may be regarded as the partial prescription of a "standard form" for U_0 and the $\mathbf{U}_i$.

D. Separation of External and Internal Motion

In nonrelativistic mechanics, the motion of the center of mass can be solved for trivially by using the conservation laws. This has the practical advantage of permitting one to ignore external motions and pass immediately to an investigation of the motions of internal coordinates. In order to gain a similar advantage relativistically, we would like to introduce a canonical coordinate involving only the generators H, $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ which would serve as a relativistic generalization of the nonrelativistic mass center. The problem of generalizing the nonrelativistic mass center for an arbitrary closed relativistic system seems to have been first considered by Pryce[26]; an illuminating modern discussion has been given by Fleming.[27] The result of these investigations is that there are three generalizations of interest, which Fleming calls the centers of mass, inertia, and spin. We denote these by $\mathbf{Q}_{\mathrm{c.m.}}$, $\mathbf{Q}_{\mathrm{c.i.}}$, and $\mathbf{Q}_{\mathrm{c.s.}}$, respectively. These three

---

[25] The conditions (41) characterizing the transformation properties of an invariant world line are just the transformations (31), (33), (34), and (35) for physical coordinates expressed with the aid of Poisson brackets and generators. These conditions were first derived by M. H. L. Pryce, Proc. Roy. Soc. (London) A195, 62 (1948); they were used by Currie, Jordan, and Sudarshan as a starting point in the proof of their zero-interaction theorem. The condition (41a) was written without the $-\mathbf{I}t$ by these authors because their $\mathbf{K}$ did not contain the term $-\mathbf{P}t$ (and was not conserved). We have chosen to retain the $-\mathbf{P}t$ in order that $\mathbf{K}$ might properly Lorentz transform all points on a world line rather than just those at $t = 0$; this choice of $\mathbf{K}$ is the choice appropriate to the notion of state subspecie aeternatis which we employ.

[26] M. H. L. Pryce, Ref. 25.

[27] G. N. Fleming, Phys. Rev. 137, B188 (1965).

generalizations can be expressed as:

$$\mathbf{Q}_{c.m.} = H^{-1}(\mathbf{K} + \mathbf{P}t), \tag{44a}$$

$$\mathbf{Q}_{c.i.} = \mathbf{Q}_{c.m.} - [H(H^2 - P^2)]^{-1}\mathbf{P} \times (\mathbf{P} \times \mathbf{K} + \mathbf{J}H), \tag{44b}$$

$$\mathbf{Q}_{c.s.} = \mathbf{Q}_{c.m.} - \frac{\mathbf{P} \times (\mathbf{P} \times \mathbf{K} + \mathbf{J}H)}{H(H^2 - P^2)^{\frac{1}{2}}[H + (H^2 - P^2)^{\frac{1}{2}}]}. \tag{44c}$$

The generator $\mathbf{K}$ can be expressed in terms of any one of these and the other generators as:

$$\mathbf{K} = H\mathbf{Q}_{c.m.} - \mathbf{P}t, \tag{45a}$$

$$\mathbf{K} = H\mathbf{Q}_{c.i.} - \mathbf{P}t - \frac{(\mathbf{J} - \mathbf{Q}_{c.i.} \times \mathbf{P}) \times \mathbf{P}}{H(H^2 - P^2)}, \tag{45b}$$

$$\mathbf{K} = H\mathbf{Q}_{c.s.} - \mathbf{P}t - \frac{(\mathbf{J} - \mathbf{Q}_{c.s.} \times \mathbf{P}) \times \mathbf{P}}{H + (H^2 - P^2)^{\frac{1}{2}}}. \tag{45c}$$

We note that all three of these are equal if the internal angular momentum $\mathbf{S}$ is zero where $\mathbf{S}$ is defined, for any of the three $\mathbf{Q}$'s, as

$$\mathbf{S} = \mathbf{J} - \mathbf{Q} \times \mathbf{P}. \tag{46}$$

The three are also all equal in the center-of-momentum frame (defined by $\mathbf{P} = 0$). It is clear from (44) that all three of the $\mathbf{Q}$'s move with the constant velocity $\mathbf{P}/H$. The transformation properties of these three $\mathbf{Q}$'s follow from their definition (44) and the Poisson bracket relations (37). By using these (and the fact that $\mathbf{K}$ is conserved, which implies $\partial \mathbf{K}/\partial t = -[\mathbf{K}, H] = -\mathbf{P}$), one can show that the center of inertia $\mathbf{Q}_{c.i.}$ transforms like a physical coordinate—i.e., it satisfies the conditions (41), and is in fact the only solution of (41) which reduces to $\mathbf{Q}_{c.m.}$ in the center-of-momentum frame. However, $[\mathbf{Q}_{c.i.}, \mathbf{Q}_{c.i.}] \neq 0$ if $\mathbf{S} \neq 0$, so that $\mathbf{Q}_{c.i.}$ cannot be canonical. Also, it follows from (44) and (37) that the components of the center of spin do have vanishing Poisson brackets [although $\mathbf{Q}_{c.s.}$ does not satisfy (41)]:

$$[\mathbf{Q}_{c.s.}, \mathbf{Q}_{c.s.}] = 0. \tag{47}$$

The center of mass $\mathbf{Q}_{c.m.}$ satisfies neither (41) nor (47) and will not be considered further.

Having summarized the relevant results of Pryce and Fleming, we now proceed to show that $\mathbf{P}$ and the center of spin $\mathbf{Q}_{c.s.}$ can be chosen to be one of the canonical pairs obtained by solving Pfaff's problem, if the $\mathbf{Q}$ in (40) is chosen to be the center of inertia $\mathbf{Q}_{c.i.}$. To begin with, we have (37a) and (47); it can be easily shown from (37) and (44) that $[\mathbf{Q}_{c.s.}, \mathbf{P}] = \mathbf{I}$. Thus $\mathbf{P}$ and $\mathbf{Q}_{c.s.}$ have the correct Poisson brackets with respect to each other; in addition, we must

satisfy the Clebsch conditions (11a) and (12a) in order that the remaining canonical coordinates and momenta can be found. The verification of (12a) for $\mathbf{P}$ is easy; we have

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, \mathbf{P}] = \sum_{i=1}^{n} \mathbf{U}_i = \mathbf{P}.$$

In order to compute the left-hand side of (11a) for $\mathbf{Q}_{c.s.}$, we first use the fact that H, $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ generate the transformations (32)–(35), respectively, to obtain

$$[\mathbf{y}_i, \hat{\mathbf{n}} \cdot (\mathbf{K} + \mathbf{P}t)] = \mathbf{g}_i^{LT}(\hat{\mathbf{n}}) + t\mathbf{g}_i^{ST}(\hat{\mathbf{n}})$$

$$[\mathbf{y}_i, \hat{\mathbf{n}} \cdot \mathbf{P} \times (\mathbf{P} \times \mathbf{K} + \mathbf{J}H)]$$
$$= \mathbf{g}_i^{ST}((\mathbf{P} \times \mathbf{K} + \mathbf{J}H) \times \hat{\mathbf{n}}) + \mathbf{g}_i^{ST}(\mathbf{K} \times (\hat{\mathbf{n}} \times \mathbf{P}))$$
$$+ \mathbf{g}_i^{LT}((\hat{\mathbf{n}} \times \mathbf{P}) \times \mathbf{P}) + (\mathbf{J} \cdot \hat{\mathbf{n}} \times \mathbf{P})\mathbf{h}_i$$
$$+ H\mathbf{g}_i^{SR}(\hat{\mathbf{n}} \times \mathbf{P}). \tag{48}$$

By the use of (42), (48), and $[\mathbf{y}_i, H] = \mathbf{h}_i$, it follows that

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, \hat{\mathbf{n}} \cdot (\mathbf{K} + \mathbf{P}t)]$$
$$= \hat{\mathbf{n}} \cdot \mathbf{K} - \hat{\mathbf{n}} \cdot \mathbf{Q}\left[U_0 + \sum_{i=1}^{2n} \mathbf{h}_i \cdot \mathbf{U}_i\right] + \mathbf{P}t,$$

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, \hat{\mathbf{n}} \cdot \mathbf{P} \times (\mathbf{P} \times \mathbf{K} + \mathbf{J}H)]$$
$$= \left(2H + \sum_{i=1}^{2n} \mathbf{U}_i \cdot \mathbf{h}_i\right)\hat{\mathbf{n}} \cdot (\mathbf{P} \times \mathbf{J}) + 3\hat{\mathbf{n}} \cdot [\mathbf{P} \times (\mathbf{P} \times \mathbf{K})]$$
$$+ \left(U_0 + \sum_{i=j}^{2n} \mathbf{U}_i \cdot \mathbf{h}_i\right)\hat{\mathbf{n}} \cdot [\mathbf{P} \times (\mathbf{P} \times \mathbf{Q})], \tag{49}$$

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, (H^2 - P^2)^{\frac{1}{2}}]$$
$$= H(H^2 - P^2)^{-\frac{1}{2}}\left(\sum_{i=1}^{2n} \mathbf{U}_i \cdot \mathbf{h}_i - P^2\right),$$

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, H] = \sum_{i=1}^{2n} \mathbf{U}_i \cdot \mathbf{h}_i.$$

With the aid of standard vector identities it can be shown from (44) and (49) that

$$\sum_{i=1}^{2n} \mathbf{U}_i \cdot [\mathbf{y}_i, \hat{\mathbf{n}} \cdot \mathbf{Q}_{c.s.}] = H^{-1}\left(U_0 + \sum_{i=1}^{2n} \mathbf{h}_i \cdot \mathbf{U}_i\right)$$
$$\{\hat{\mathbf{n}} + (H^2 - P^2)^{-\frac{1}{2}}[H + (H^2 - P^2)^{\frac{1}{2}}]^{-1}(\hat{\mathbf{n}}P^2 - \hat{\mathbf{n}} \cdot \mathbf{PP})\}$$
$$\cdot (\mathbf{Q} - \mathbf{Q}_{c.i.}). \tag{50}$$

Thus $\mathbf{Q}_{c.s.}$ satisfies (11a), if $\mathbf{Q} = \mathbf{Q}_{c.i.}$ in (40), in which case $\mathbf{P}$ and $\mathbf{Q}_{c.s.}$ can be one of the canonical pairs obtained by solving Pfaff's problem.

In order to complete the discussion, we express the angular momentum $\mathbf{J}$ in terms of the canonical coordinates. We first remark that the fact that the $\mathbf{U}_i$ transform as vector functions [see Eq. (43)] implies that

the $\mathbf{T}_{ij}$ transform as dyadic functions. From this it follows that the Poisson brackets $[\mathbf{y}_i, \mathbf{y}_j]$ reciprocal to the Γ_{ij} transform as dyadic functions also; Clebsch's condition (11a) which reads $\sum_{i,j} \mathbf{U}_i \cdot [\mathbf{y}_i, \mathbf{y}_j] \cdot (\partial/\partial y_j) \mathbf{Q}_k = 0$ then implies that $(\partial/\partial y_j) \mathbf{Q}_k$ transforms as a dyadic function so that the $\mathbf{Q}_k$ can be chosen to transform as vector functions, i.e.,

$$\sum_{i=1}^{2n} (\mathbf{y}_i \times \partial/\partial y_i) \mathbf{Q}_k = \mathbf{I} \times \mathbf{Q}_k. \tag{51}$$

If one inserts $\mathbf{U}_i = \sum_{k=1}^{n} [(\partial/\partial y_i) \mathbf{Q}_k] \cdot \mathbf{P}_k$, which is the three dimensional version of (5), into (42), and uses (51), the angular momentum $\mathbf{J}$ takes the form

$$\mathbf{J} = \sum_k \mathbf{Q}_k \times \mathbf{P}_k. \tag{52}$$

From (52) it follows that the internal angular momentum $\mathbf{S}_{\text{c.s.}} = \mathbf{J} - \mathbf{Q}_{\text{c.s.}} \times \mathbf{P}$ depends only on internal coordinates if $\mathbf{Q}_{\text{c.s.}}$ and $\mathbf{P}$ are taken as a canonical pair. Furthermore, $[\mathbf{Q}_{\text{c.s.}}, (H^2 - P^2)^{\frac{1}{2}}] = 0$ follows immediately from $[\mathbf{Q}_{\text{c.s.}}, H] = \mathbf{P}/H$ and $[\mathbf{Q}_{\text{c.s.}}, \mathbf{P}] = \mathbf{I}$. Also $[\mathbf{P}, (H^2 - P^2)^{\frac{1}{2}}] = 0$, hence, $(H^2 - P^2)^{\frac{1}{2}}$, which is the energy in the center-of-momentum frame, depends only on internal coordinates. The separation into external and internal motions is not as complete as it is nonrelativistically, however, because internal coordinates and momenta are not in general invariant under Lorentz transformation as they are under Galilean transformation nonrelativistically. If we use the form (45c) for $\mathbf{K}$, we find, for an internal coordinate $\dot{\mathbf{q}}_{\text{int}}$,

$$[\mathbf{q}_{\text{int}}, \hat{\mathbf{n}} \cdot \mathbf{K}] = \dot{\mathbf{q}}_{\text{int}} \{ \hat{\mathbf{n}} \cdot \mathbf{Q}_{\text{c.s.}}$$
$$+ [H + (H^2 - P^2)^{\frac{1}{2}}]^{-1} (H^2 - P^2)^{-\frac{1}{2}} \hat{\mathbf{n}} \cdot (\mathbf{S}_{\text{c.s.}} \times \mathbf{P}) \}$$
$$+ [H + (H^2 - P^2)^{\frac{1}{2}}]^{-1} [\mathbf{q}_{\text{int}} (\mathbf{P} \cdot \hat{\mathbf{n}}) - \hat{\mathbf{n}} (\mathbf{q}_{\text{int}} \cdot \mathbf{P})]$$

which will not vanish in general; the same is true of internal momenta.

The forms (45c) and (52) arrived at for $\mathbf{K}$ and $\mathbf{J}$ are in agreement with the result arrived at by Foldy.[28] The difference between the present work and that of Foldy is that here all physical coordinates and velocities transform properly; the connection between physical and canonical variables is fixed by the way in which the Hamiltonian formulation has been obtained, rather than by a postulate as in his work.

V. EXAMPLES

Before proceeding further with general considerations, it seems worthwhile to consider several concrete examples. We begin this section by summa-

rizing the traditional canonical formalism for relativistic free particles. The second example consists of a pair of second-order differential equations whose invariance group consists of the general linear transformations; we interpret these as a description of two one-dimensional particles interacting in a relativistically invariant manner and proceed to find all of the various ways of casting this dynamics into Hamiltonian form with the transformations of the inhomogeneous Lorentz group canonical. The section concludes with another one-dimensional, two-particle example which, although relativistically invariant, does not admit an Hamiltonian formulation with the entire inhomogeneous Lorentz group canonical.

A. Free Particles

The set of ten generators conventionally used[5,28,29] for relativistic free particles are:

$$H = \sum_{i=1}^{n} (p_i^2 + m_i^2)^{\frac{1}{2}},$$

$$\mathbf{P} = \sum_{i=1}^{n} \mathbf{p}_i,$$

$$\mathbf{J} = \sum_{i=1}^{n} \mathbf{x}_i \times \mathbf{p}_i, \tag{53}$$

$$\mathbf{K} = \sum_{i=1}^{n} [\mathbf{x}_i (p_i^2 + m_i^2)^{\frac{1}{2}} - \mathbf{p}_i t],$$

where $\mathbf{x}_i$, $\mathbf{p}_i$ are canonical mates. Here $\mathbf{x}_i$ is the physical position; the mechanical momentum $\mathbf{p}_i$ is related to the physical velocity $\mathbf{v}_i$ by

$$\mathbf{p}_i = m_i \mathbf{v}_i (1 - v_i^2)^{-\frac{1}{2}}. \tag{54}$$

By comparing Hamilton's principle

$$\delta \int \left(\sum_{i=1}^{n} \mathbf{p}_i \cdot d\mathbf{x}_i - H \, dt \right) = 0$$

with (2), we see that

$$\mathbf{U}_i = m_i \mathbf{v}_i (1 - v_i^2)^{-\frac{1}{2}}, \quad \mathbf{U}_{i+n} = 0, \quad i = 1, 2, \cdots, n. \tag{55}$$

The equations of motion are

$$\mathbf{h}_i = \mathbf{v}_i, \quad \mathbf{h}_{i+n} = 0, \quad i = 1, 2, \cdots, n. \tag{56}$$

It can be easily verified that the generators (53) satisfy the Poisson bracket relations (37), and generate the transformations (32–35) of the positions and velocities.

B. A One-Dimensional Example Exhibiting Interaction

A very interesting example of a one-dimensional, two-particle Lorentz invariant dynamics, for which

---

[28] L. L. Foldy, Phys. Rev. **122**, 275 (1961).

[29] Our $\mathbf{K}$ differs from that of Ref. 5 by $-\mathbf{P}t$ (see Ref. 25).

the equations of motion are

$$-a_1 = a_2 = (v_1 - v_2)^2/2(x_1 - x_2), \quad (57)$$

has been given by Kerner.[12] The solutions of (57) are easily found to be

$$x_1 = A + Bt + (|Ct + D|)^{\frac{1}{2}},$$
$$x_2 = A + Bt - (|Ct + D|)^{\frac{1}{2}}. \quad (58)$$

In obtaining (58), a principal value prescription has been used to integrate through the singularity at $x_1 \to x_2$; this gives rise to the absolute value inside the square root. The solutions (58) are parabolas in the $x - t$ plane. It follows from the fact that an arbitrary nonsingular linear transformation

[i.e., $\quad x' = a + bx + ct, \quad t' = d + ex + ft,$
$$ec - bf \neq 0]$$

carries parabolas into parabolas that the differential equations (57) have the general linear group on x and t (of which the inhomogeneous Lorentz group is a subgroup) as invariance group.

The behavior of the integration constants A, B, C, and D under the (finite) transformations of the inhomogeneous Lorentz group can be found by conducting the transformation on the solution (58) and re-arranging the result in the form of Eq. (58). In this way we find that, under the time translation $t \to t' = t - t_0$,

$$A \to A' = A + Bt_0,$$
$$B \to B' = B,$$
$$C \to C' = C, \quad (59)$$
$$D \to D' = D + Ct_0.$$

Under the space translation $x \to x' = x + x_0$,

$$A \to A' = A + x_0,$$
$$B \to B' = B,$$
$$C \to C' = C, \quad (60)$$
$$D \to D' = D.$$

Under the Lorentz transformation,

$$x \to x' = (x - \beta t)(1 - \beta^2)^{-\frac{1}{2}},$$
$$t \to t' = (t - \beta x)(1 - \beta^2)^{-\frac{1}{2}},$$

[where each of the two orbits must be transformed *separately* and rewritten in the form (58)] one finds:

$$A \to A' = (1 - \beta^2)^{\frac{1}{2}}(1 + \beta B)^{-1}[A - \tfrac{1}{2}\beta C(1 + \beta B)^{-1}],$$
$$B \to B' = (B + \beta)(1 + \beta B)^{-1},$$
$$C \to C' = (1 - \beta^2)^{\frac{3}{2}}(1 + \beta B)^3 C,$$
$$D \to D' = (1 - \beta^2)(1 + \beta B)^{-2}[D - \beta AC(1 + \beta B)^{-1}$$
$$+ \tfrac{1}{4}\beta^2 C^2(1 + \beta B)^{-2}]. \quad (61)$$

The integration constants B and C provide the two first integrals

$$B = \tfrac{1}{2}(v_1 + v_2), \quad C = \tfrac{1}{2}(v_1 - v_2)(x_1 - x_2), \quad (62)$$

as can be easily verified from (58). The integration constants A and D provide the two second integrals

$$A = \tfrac{1}{2}(x_1 + x_2) - \tfrac{1}{2}(v_1 + v_2)t,$$
$$D = \tfrac{1}{4}(x_1 - x_2)^2 - (v_1 - v_2)(x_1 - x_2)t. \quad (63)$$

A considerable amount of labor can be saved in the task of casting Eq. (57) into Hamiltonian form if the integrals H, P, and K, which are necessarily functions of the integrals A, B, C, and D of Eqs. (62) and (63), can be found ahead of time. If these are known, Eq. (20) provides algebraic equations for the Γ_{ij} which are more easily solved than the differential equations (22). Since H and P are both first integrals, invariant under both space and time translation, they must be functions of the first integrals B and C. Furthermore, H and P must transform among themselves as a two-vector under pure Lorentz transformation. We note from (61) that B transforms like a velocity while $C(1 - B^2)^{-\frac{3}{2}}$ is invariant under pure Lorentz transformation. Hence, the most general possibility is

$$H = (1 - B^2)^{-\frac{1}{2}}\varphi(\tfrac{1}{2}C(1 - B^2)^{-\frac{3}{2}}),$$
$$P = B(1 - B^2)^{-\frac{1}{2}}\varphi(\tfrac{1}{2}C(1 - B^2)^{-\frac{3}{2}}), \quad (64)$$

where φ is an arbitrary function of the indicated invariant. Under the time translation, $t \to t' = t - t_0$, $K \to K' = K + Pt_0$. Under the space translation, $x \to x' = x + x_0$, $K \to K' = K + Hx_0$. The conserved quantity $\varphi(\tfrac{1}{2}C(1 - B^2)^{-\frac{3}{2}})(1 - B^2)^{-\frac{1}{2}}A$ has this behavior under space and time translation, but is not invariant under Lorentz transformation, as K must be. This can be fixed up by adding on an appropriate function of B and C; hence,

$$K = [A(1 - B^2)^{-\frac{1}{2}} + \tfrac{1}{2}BC(1 - B^2)^{-\frac{3}{2}}]$$
$$\times \varphi(\tfrac{1}{2}C(1 - B^2)^{-\frac{3}{2}}) + \psi(\tfrac{1}{2}C(1 - B^2)^{-\frac{3}{2}}), \quad (65)$$

where φ is as in (64), and ψ is a second arbitrary function. K must be independent of D if it is to transform properly under space and time translations; hence Eq. (65) is the most general possibility for K compatible with Eq. (64). If we interpret one of the quantities in a Poisson bracket as the object being transformed and the other as the generator of the transformation, it is clear that H, P, and K of Eqs. (64) and (65) satisfy the Poisson bracket relations

$$[H, P] = 0, \quad [H, K] = -P, \quad [P, K] = -H, \quad (66)$$

to which (37) reduces in the one-dimensional case.

The computations can be further simplified by using sum and difference variables. Thus we choose

$$y_0 = t, \quad y_1 = \tfrac{1}{2}(x_1 + x_2), \quad y_2 = \tfrac{1}{2}(x_1 - x_2),$$
$$y_3 = \tfrac{1}{2}(v_1 + v_2), \quad y_4 = \tfrac{1}{2}(v_1 - v_2). \quad (67)$$

With these choices, the h_i for this example are

$$h_0 = 1, \quad h_1 = y_3, \quad h_2 = y_4,$$
$$h_3 = 0, \quad \text{and} \quad h_4 = -y_4^2/y_2. \quad (68)$$

The g_i appropriate to the y_i of (67) will be half the sum (or difference) of the g_i of Sec. IV. Thus, for example, application of (35) yields for the Lorentz transformation

$$g_1^{LT} = \tfrac{1}{2}[(v_1 x_1 - t) + (v_2 x_2 - t)] = y_1 y_3 + y_2 y_4 - y_0.$$

The other g_i are computed similarly; therefore

$$g_0^{TT} = -1, \quad g_i^{TT} = 0, \quad i \neq 0;$$
$$g_1^{ST} = 1, \quad g_i^{ST} = 0, \quad i \neq 1;$$
$$g_0^{LT} = 0, \quad g_1^{LT} = y_1 y_3 + y_2 y_4 - y_0, \quad (69)$$
$$g_2^{LT} = y_2 y_3 + y_1 y_4 \quad g_3^{LT} = -1 + y_3^2, \quad \text{and}$$
$$g_4^{LT} = 2y_3 y_4 - (y_1 y_4^2/y_2).$$

The generators H, P, and K of Eqs. (64) and (65) can also be rewritten in terms of the y_i specified in (67); using (62) and (63) we obtain

$$H = (1 - y_3^2)^{-\frac{1}{2}} \varphi(z),$$
$$P = y_3(1 - y_3^2)^{-\frac{1}{2}} \varphi(z), \quad (70)$$
$$K = H[y_1 + (1 - y_3^2)^{\frac{1}{2}} y_3 z + H^{-1}\psi(z)] - Py_0.$$

Here $z \equiv y_2 y_4 (1 - y_3^2)^{-\frac{1}{2}}$. If we now elect to restrict ourselves to sets of U_i for which translational invariance is manifest, i.e., for which $\partial U_i/\partial y_0 = \partial U_i/\partial y_1 = 0$, we have

$$U_0 = -H \quad \text{and} \quad U_1 = P. \quad (71)$$

From these we have immediately seven of the ten Γ_{ij}; thus,

$$\Gamma_{0i} = -\Gamma_{i0} = -\partial H/\partial y_i,$$
$$\Gamma_{1i} = -\Gamma_{i1} = \partial P/\partial y_i. \quad (72)$$

The relations (72) could also be obtained from (20) when written out for the space and time translations, using the g_i^{TT} and g_i^{ST} of (69). The remaining three Γ_{ij} can be obtained by writing out the conditions of Eq. (4) and the conditions of Eq. (20) for the Lorentz transformations; these take the form

$$\Gamma_{0j} + y_3 \Gamma_{1j} + y_4 \Gamma_{2j} - (y_4^2/y_2)\Gamma_{4j} = 0, \quad (73)$$

$$(y_1 y_3 + y_2 y_4 - y_0)\Gamma_{1j} + (y_2 y_3 + y_1 y_4)\Gamma_{2j}$$
$$+ (-1 + y_3^2)\Gamma_{3j} + [2y_3 y_4 - (y_1 y_4^2/y_2)]\Gamma_{4j}$$
$$= \partial K/\partial y_j. \quad (74)$$

For $j = 0, 1$, these are consistent with Eq. (72). For $j = 2, 3, 4$, they provide six equations for the three remaining Γ_{ij}. Of these six, only three are linearly independent; the solutions for the remaining Γ_{ij} are

$$\Gamma_{23} = y_2 y_3 (1 - y_3^2)^{-2}[2 + y_3^2 y_4^2 (1 - y_3^2)^{-2}]\varphi(z)$$
$$+ y_3 y_4 (1 - y_3^2)^{-\frac{3}{2}} \varphi(z) + y_4 (1 - y_3^2)^{-\frac{3}{2}} \psi'(z),$$
$$\Gamma_{24} = y_2 y_4^{-1}(1 - y_3^2)^{-1}\varphi'(z),$$
$$\Gamma_{34} = y_2^2 y_3 y_4^{-1}(1 - y_3^2)^{-2}[1 - y_3^2 y_4^2 (1 - y_3^2)^{-2}]\varphi'(z)$$
$$- y_2 y_3 (1 - y_3^2)^{-\frac{3}{2}} \varphi(z) - y_2 (1 - y_3^2)^{-\frac{3}{2}} \psi'(z). \quad (75)$$

In one dimension there is no angular momentum; hence, the centers of mass, inertia, and spin are identical and may be identified from (70):

$$Q = y_1 + (1 - y_3^2)^{\frac{1}{2}}\{y_3 z + [\psi(z)/\varphi(z)]\}. \quad (76)$$

The Γ_{ij} of (72) and (75) satisfy the integrability conditions (21); hence, (75) can be integrated to obtain U_2, U_3, and U_4. A particular solution is

$$U_2 = \tfrac{1}{2} y_2^{-1} z[y_3^2 + y_4^{-2}(1 - y_3^2)^2]\varphi(z),$$
$$U_3 = \tfrac{1}{2} y_3 (1 - y_3^2)^{-1} z[y_3^2 - y_4^{-2}(1 - y_3^2)^2]\varphi(z)$$
$$\qquad - (1 - y_3^2)^{-1}\psi(z), \quad (77)$$
$$U_4 = \tfrac{1}{2} y_4^{-1} z[y_3^2 - y_4^{-2}(1 - y_3^2)^2]\varphi(z).$$

The U_i of Eq. (77) have been adjusted (by the addition of a gradient) to make the action invariant (in the sense of Sec. III) under the modified Lorentz transformation given by (38) and (40); hence the P of Eq. (70) and the Q of Eq. (76) can be one of the canonical pairs obtained by solving Pfaff's problem with these U_i.

The inversion of the matrix Γ_{ij} ($i, j \neq 0$) to get the Poisson brackets $[y_i, y_j]$ is made easier by the fact that the determinant of an antisymmetric matrix is a perfect square (it is the square of a Pfaffian[30]); thus, $\det \Gamma_{ij} = \Phi^2$ where $\Phi = -\Gamma_{12}\Gamma_{34} + \Gamma_{13}\Gamma_{24} - \Gamma_{14}\Gamma_{23}$. The Poisson brackets are then

$$[y_1, y_2] = -\Phi^{-1}\Gamma_{34}, \quad [y_2, y_3] = -\Phi^{-1}\Gamma_{14},$$
$$[y_1, y_3] = \Phi^{-1}\Gamma_{24}, \quad [y_2, y_4] = \Phi^{-1}\Gamma_{13}, \quad (78)$$
$$[y_1, y_4] = -\Phi^{-1}\Gamma_{23}, \quad [y_3, y_4] = -\Phi^{-1}\Gamma_{12}.$$

By computation,

$$\Phi = y_4^{-2}(1 - y_3^2)^{-1} z \varphi(z)\varphi'(z). \quad (79)$$

The partial-differential operators which appear in Eqs. (11) and (12) can now be computed. For any

---

[30] A. R. Forsyth, Ref. 17, p. 95.

function F of the y_i,

$$\sum_{i=1}^{4}\sum_{j=1}^{4}[y_i,\,y_j]U_i\partial F/\partial y_j$$
$$= \tfrac{1}{2}\Phi^{-1}y_4^{-2}(1-y_3^2)^{-1}[\varphi(z)]^2[-y_3(1-y_3^2)^{\frac{1}{2}}z(\partial F/\partial y_1)$$
$$+ y_2(\partial F/\partial y_2) + y_4(\partial F/\partial y_4)] + \Phi^{-1}(1-y_3^2)^{-2}$$
$$\times \{\tfrac{1}{2}y_3[\varphi(z)]^2 + \varphi(z)\,\psi'(z) - \varphi'(z)\,\psi(z)\}$$
$$\times \{-(y_2/y_4)(\partial F/\partial y_1) + y_3(1-y_3^2)^{-1}$$
$$\times [y_2(\partial F/\partial y_2) - y_4(\partial F/\partial y_4)]\}, \qquad (80)$$

$$[F,P] = \sum_{i=1}^{4}\sum_{j=1}^{4}[y_i,\,y_j]\frac{\partial P}{\partial y_i}\frac{\partial F}{\partial y_j} = \frac{\partial F}{\partial y_1}, \qquad (81)$$

$$[F,Q] = \sum_{i=1}^{4}\sum_{j=1}^{4}[y_i,\,y_j]\frac{\partial Q}{\partial y_i}\frac{\partial F}{\partial y_j}$$
$$= \Phi^{-1}y_4^{-2}z\varphi'(z)\{(2y_3^2-1)z + y_3[\psi(z)/\varphi(z)]\}$$
$$\times (\partial F/\partial y_1) + \Phi^{-1}\{y_4^{-2}y_3(1-y_3^2)^{-\frac{1}{2}}z\varphi'(z)$$
$$\times [1 - y_4^2(1-y_3^2)^{-1}] - (1-y_3^2)^{-\frac{3}{2}}$$
$$\times [\varphi'(z)\psi(z)/\varphi(z)]\}[-y_2(\partial F/\partial y_2)$$
$$+ y_4(\partial F/\partial y_4)] + \Phi^{-1}y_4^{-2}(1-y_3^2)^{-\frac{1}{2}}z\varphi'(z)$$
$$\times [-3y_3y_4(\partial F/\partial y_4) + (1-y_3^2)(\partial F/\partial y_3)]. \qquad (82)$$

An internal canonical coordinate q can now be found by solving (11a) simultaneously with $[q,P] = [q,Q] = 0$. A particular solution is

$$q = \tfrac{1}{2}z[y_4^{-2}(1-y_3^2)^2 - y_3^2] - y_3[\psi(z)/\varphi(z)]. \qquad (83)$$

The canonical mate p to the q of (83) is most easily found from $pdq = \sum_{i=1}^{4}U_i\,dy_i - PdQ$; it is

$$p = \varphi(z). \qquad (84)$$

The generators H and K are now expressible in terms of the canonical variables P, Q, p, q which are related to the physical positions and velocities by (67), (70), (76), (83), and (84); from (70) are obtained

$$H = (p^2 + P^2)^{\frac{1}{2}}, \quad K = HQ - Pt. \qquad (85)$$

Canonically equivalent formulations are obtainable by canonical transformations; different choices of the arbitrary functions φ and ψ give rise to canonically inequivalent formulations.

C. Another Example Exhibiting Interaction

Another example of a one-dimensional, two-particle Lorentz invariant dynamics is specified by the pair of second-order differential equations

$$a_1 = (v_2 - v_1)(1 - v_1^2)[v_2(x_1 - x_2)]^{-1},$$
$$a_2 = (v_2 - v_1)(1 - v_2^2)[v_1(x_1 - x_2)]^{-1}. \qquad (86)$$

Two independent first integrals of Eq. (86) which are

invariant under both space and time translations are given by

$$A = (x_1 - x_2)^2(v_1 + v_2)(v_2 - v_1)^{-1},$$
$$B = (x_1 - x_2)^2(v_1^2 + v_2^2 - 2v_1^2v_2^2)(v_2 - v_1)^{-2}. \qquad (87)$$

By the use of (35) and (86), it follows that for this dynamics

$$L_{\mathrm{LT}} = (v_1x_1 - v_2x_2)(v_1 - v_2)^{-1}(D - \partial/\partial t)$$
$$- [t + v_1v_2(x_1 - x_2)(v_1 - v_2)^{-1}][(\partial/\partial x_1) + (\partial/\partial x_2)].$$

Hence, $L_{\mathrm{LT}}A = L_{\mathrm{LT}}B = 0$, i.e., *both* of the first integrals of (87) are Lorentz invariant.

We now proceed to show that this dynamics *cannot* be cast into Hamiltonian form with all of the transformations of the inhomogeneous Lorentz group canonical. The proof is by contradiction; suppose such an Hamiltonian formulation exists. Then there must exist first integrals P and H, invariant under space and time translations, which form a two-vector under Lorentz transformation. But there are no integrals of Eqs. (86), independent of A and B of Eq. (87), which are invariant under space and time translations. Hence, P and H must be functions of A and B. But, since A and B are both invariant under Lorentz transformation, no pair of functions of them can form a two-vector under Lorentz transformation. Hence, the required P and H do not exist.

The last example of the present section makes it clear that it is not always possible to cast a Lorentz-invariant dynamics into Hamiltonian form if it is required that the transformations of the inhomogeneous Lorentz group be canonical. The second example illustrates the fact that this Hamiltonization can be done in many ways, if it can be done at all. This large number of canonically inequivalent Hamiltonian formulations presents a problem when one comes to construct a quantum theory; which Hamiltonian scheme is to be chosen? Different choices will lead to different quantum theories[31]; it is therefore necessary to resolve this ambiguity before proceeding to quantization.[32]

VI. THE ZERO-INTERACTION THEOREM OF CURRIE, JORDAN, AND SUDARSHAN

The zero-interaction theorem of Currie, Jordan, and Sudarshan[5,6] has been frequently quoted (sometimes without mention of the crucial assumption of

[31] P. Havas, Bull. Am. Phys. Soc. **1**, 337 (1956); F. J. Kennedy and E. H. Kerner, Am. J. Phys. **33**, 463 (1965).

[32] An existence and uniqueness theorem which answers this question for dynamical equations admitting free particle motion asymptotically has been given by R. N. Hill and E. H. Kerner, Phys. Rev. Letters **17**, 1156 (1966). A fuller discussion of this theorem appears in R. N. Hill and E. H. Kerner (to be published).

the identity of physical and canonical coordinates) as evidence of the impossibility of formulating relativistic dynamics within the framework of the canonical representations of the inhomogeneous Lorentz group. Inasmuch as we have formulated just such a dynamics, it seems worthwhile to re-prove the theorem in a way which shows clearly just how we have circumvented its prohibitions. As a byproduct, we are able to clarify the situation in one and two dimensions, which was left quite muddy by previous proofs which were valid only in three dimensions.

We begin with the Poisson bracket rewrite of the transformations (32) and (35)

$$[\mathbf{x}_i, H] = \mathbf{v}_i, \tag{88a}$$

$$[\mathbf{v}_i, H] = \mathbf{F}_i, \tag{88b}$$

$$[\mathbf{x}_i, \hat{\mathbf{n}} \cdot \mathbf{K}] = \mathbf{v}_i(\hat{\mathbf{n}} \cdot \mathbf{x}_i) - \hat{\mathbf{n}}t, \tag{88c}$$

$$[\mathbf{v}_j, \hat{\mathbf{n}} \cdot \mathbf{K}] = \mathbf{F}_j(\hat{\mathbf{n}} \cdot \mathbf{x}_j) - \hat{\mathbf{n}} + \mathbf{v}_j(\hat{\mathbf{n}} \cdot \mathbf{v}_j). \tag{88d}$$

By taking the Poisson bracket of (88c) with $\mathbf{x}_j$, one obtains

$$[\mathbf{x}_j, [\mathbf{x}_i, \hat{\mathbf{n}} \cdot \mathbf{K}]] = (\hat{\mathbf{n}} \cdot \mathbf{x}_i)[\mathbf{x}_j, \mathbf{v}_i] + [\mathbf{x}_j, \mathbf{x}_i] \cdot \hat{\mathbf{n}}\mathbf{v}_i. \tag{89}$$

Interchanging i and j and transposing the dyadics produces

$$[[\mathbf{x}_j, \hat{\mathbf{n}} \cdot \mathbf{K}], \mathbf{x}_i] = (\hat{\mathbf{n}} \cdot \mathbf{x}_j)[\mathbf{v}_j, \mathbf{x}_i] + \mathbf{v}_j\hat{\mathbf{n}} \cdot [\mathbf{x}_j, \mathbf{x}_i]. \tag{90}$$

It follows from (88a), (88b), and the Jacobi identity that

$$[\mathbf{x}_j, \mathbf{v}_i] = -[\mathbf{v}_j, \mathbf{x}_i] + [[\mathbf{x}_j, \mathbf{x}_i], H], \tag{91}$$

$$[\mathbf{F}_j, \mathbf{x}_i] = -[\mathbf{v}_j, \mathbf{v}_i] + [[\mathbf{v}_j, \mathbf{x}_i], H]. \tag{92}$$

The addition of Eqs. (89) and (90) followed by the use of the Jacobi identity and Eq. (91) produces the result

$$\hat{\mathbf{n}} \cdot (\mathbf{x}_i - \mathbf{x}_j)[\mathbf{v}_j, \mathbf{x}_i] = (\hat{\mathbf{n}} \cdot \mathbf{x}_i)[[\mathbf{x}_j, \mathbf{x}_i], H]$$
$$- [[\mathbf{x}_j, \mathbf{x}_i], \hat{\mathbf{n}} \cdot \mathbf{K}] + [\mathbf{x}_j, \mathbf{x}_i] \cdot \hat{\mathbf{n}}\mathbf{v}_i + \mathbf{v}_j\hat{\mathbf{n}} \cdot [\mathbf{x}_j, \mathbf{x}_i]. \tag{93}$$

Similarly, by taking the Poisson bracket of (88c) with $\mathbf{v}_j$ on the left, the Poisson bracket of (88d) with $\mathbf{x}_i$ on the right, adding and using (92) and the Jacobi identity one obtains

$$\hat{\mathbf{n}} \cdot (\mathbf{x}_j - \mathbf{x}_i)[\mathbf{v}_j, \mathbf{v}_i] = (\hat{\mathbf{n}} \cdot \mathbf{x}_j)[[\mathbf{v}_j, \mathbf{x}_i], H]$$
$$- [[\mathbf{v}_j, \mathbf{x}_i], \hat{\mathbf{n}} \cdot \mathbf{K}] + \mathbf{F}_j\hat{\mathbf{n}} \cdot [\mathbf{x}_j, \mathbf{x}_i]$$
$$+ [\mathbf{v}_j, \mathbf{x}_i](\hat{\mathbf{n}} \cdot \mathbf{v}_j) + [\mathbf{v}_j, \mathbf{x}_i] \cdot \hat{\mathbf{n}}\mathbf{v}_i + \mathbf{v}_j\hat{\mathbf{n}} \cdot [\mathbf{v}_j, \mathbf{x}_i]. \tag{94}$$

By taking the Poisson bracket of (88d) with $\mathbf{v}_i$ and performing manipulations similar to those which

produced (93), one obtains

$$\hat{\mathbf{n}} \cdot (\mathbf{x}_j - \mathbf{x}_i)[\mathbf{F}_i, \mathbf{v}_j] = (\hat{\mathbf{n}} \cdot \mathbf{x}_j)[[\mathbf{v}_i, \mathbf{v}_j], H]$$
$$- [[\mathbf{v}_i, \mathbf{v}_j], \hat{\mathbf{n}} \cdot \mathbf{K}] + [\mathbf{v}_i, \mathbf{x}_j] \cdot \hat{\mathbf{n}}\mathbf{F}_j + \mathbf{F}_i\hat{\mathbf{n}} \cdot [\mathbf{x}_i, \mathbf{v}_j]$$
$$+ \hat{\mathbf{n}} \cdot (\mathbf{v}_i + \mathbf{v}_j)[\mathbf{v}_i, \mathbf{v}_j] + [\mathbf{v}_i, \mathbf{v}_j] \cdot \hat{\mathbf{n}}\mathbf{v}_j + \mathbf{v}_i\hat{\mathbf{n}} \cdot [\mathbf{v}_i, \mathbf{v}_j]. \tag{95}$$

Nowhere in the derivation of Eqs. (92)–(95) has the dimensionality of the space been used; they are therefore valid in one or two dimensions as well as in three. They are also quite compatible with interaction; in particular they hold for the Hamiltonian formulation of example (51) obtained in the preceding section.

At this point let us make the crucial *additional assumption* that the physical coordinate can be canonical, which means that $[\mathbf{x}_i, \mathbf{x}_j] = 0$. When this assumption is made, it follows immediately from Eqs. (92)–(95) that, for $i \neq j$,

$$[\mathbf{x}_i, \mathbf{x}_j] = 0, \tag{96a}$$

$$[\mathbf{v}_j, \mathbf{x}_i] = 0, \tag{96b}$$

$$[\mathbf{v}_j, \mathbf{v}_i] = 0, \tag{96c}$$

$$[\mathbf{F}_j, \mathbf{x}_i] = 0, \tag{96d}$$

$$[\mathbf{F}_i, \mathbf{v}_j] = 0. \tag{96e}$$

By the use of Eq. (8), (96a)–(96c) the results (96d) and (96e) take the form:

$$0 = [\mathbf{x}_i, \mathbf{F}_j] = [\mathbf{x}_i, \mathbf{x}_i] \cdot \frac{\partial}{\partial \mathbf{x}_i}\mathbf{F}_j + [\mathbf{x}_i, \mathbf{v}_i] \cdot \frac{\partial}{\partial \mathbf{v}_i}\mathbf{F}_j,$$
$$0 = [\mathbf{v}_i, \mathbf{F}_j] = [\mathbf{v}_i, \mathbf{x}_i] \cdot \frac{\partial}{\partial \mathbf{x}_i}\mathbf{F}_j + [\mathbf{v}_i, \mathbf{v}_i] \cdot \frac{\partial}{\partial \mathbf{v}_i}\mathbf{F}_j. \tag{97}$$

If we now make the further assumption that the dynamics is nondegenerate,[33] so that the positions and velocities form a complete set of dynamical variables, the matrix of coefficients in (97) is nonsingular and we have

$$\frac{\partial}{\partial \mathbf{x}_i}\mathbf{F}_j = \frac{\partial}{\partial \mathbf{v}_i}\mathbf{F}_j = 0, \quad i \neq j, \tag{98}$$

which states that the acceleration of the jth particle is independent of the positions and velocities of the ith particle, i.e., *they do not interact.*

The original proofs of the zero-interaction theorem adopted straight-line motion of the particles as a criterion for the absence of interaction. The connection between this criterion and that of (98) can be obtained by writing out the conditions (17) for the invariance of the equations of motion $d\mathbf{v}_j/dt = \mathbf{F}_j$

---

[33] This somewhat stronger hypothesis was used by Leutwyler (Ref. 8), but was not used by Currie, Jordan, and Sudarshan (Refs. 5 and 6) or by Cannon and Jordan (Ref. 7).

under the Lorentz transformations (35). The conditions read[34]

$$2\mathbf{v}_j\mathbf{F}_j + \mathbf{F}_j\mathbf{v}_j + \sum_{i=1}^{n}\left[(\mathbf{x}_j - \mathbf{x}_i)\left(\mathbf{v}_i \cdot \frac{\partial}{\partial\mathbf{x}_i} + \mathbf{F}_i \cdot \frac{\partial}{\partial\mathbf{v}_i}\right)\right.$$
$$\left. + (1 - \mathbf{v}_i\mathbf{v}_i) \cdot \frac{\partial}{\partial\mathbf{v}_i}\right]\mathbf{F}_j = 0. \quad (99)$$

The only simultaneous solution of (98) and (99) in two or three dimensions is $\mathbf{F}_j = 0$, i.e., straight-line motion. The most general simultaneous solution of Eqs. (98) and (99) in one dimension is $F_j = C_j(1 - v_j^2)^{\frac{3}{2}}$, where the C_j are constants; for $C_j \neq 0$ this describes a particle moving in a constant external field (hyperbolic motion). The only other case admitting curved world lines with the inhomogeneous Lorentz transformations canonical and $[\mathbf{x}_i, \mathbf{x}_j] = 0$ occurs when at least some of the equations are degenerate; Currie's one-dimensional example[6,35] is of this type.

It should now be clear that the Currie–Jordan–Sudarshan zero interaction theorem does not prohibit the formulation of a relativistic dynamics within the framework of the canonical representations of the inhomogeneous Lorentz group. Their theorem can in fact be viewed as a *reductio ad absurdum* proof of the fact that the Poisson brackets $[\mathbf{x}_i, \mathbf{x}_j]$ cannot all vanish in the interaction region of such a relativistic Hamiltonian mechanics. This nonvanishing of $[\mathbf{x}_i, \mathbf{x}_j]$ is, at first sight, somewhat disturbing; because it would seem to imply that particle positions are not simultaneously measureable with arbitrary precision in a quantum theory obtained by the usual Poisson-bracket-to-commutator prescription. It is the present author's contention that such an interpretation of $[\mathbf{x}_i, \mathbf{x}_j] \neq 0$ is not valid, because it ignores the difference between the description of a position measurement in a local field theory and its description in an instantaneous interaction theory.

We discuss this difference at the classical level. For concreteness, let us suppose that the measurement apparatus consists of a sensitized space–time volume, such as a Geiger counter, which has spatial extent Δx and which is sensitized for a time interval Δt as in
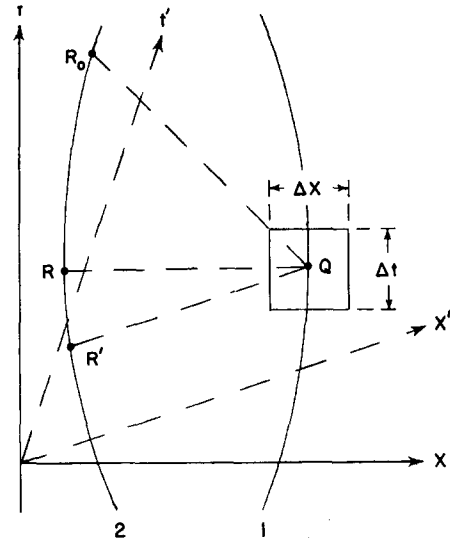

FIG. 3. Measurement of a particle's position.

Fig. 3. As a particle (world line 1 in Fig. 3) passes through the sensitized space–time volume, it in general suffers a perturbation which causes an alteration in its world line. In a local field theory which employs the usual notion of Einstein causality, this perturbation does not effect a change in the motion of a second particle (world line 2 in Fig. 3) until sufficient time has elapsed for a light signal to propagate from the first particle to the second (to the world point R_0 in Fig. 3).

In a conventional field theory, the measurement interaction with the sensitized space–time volume is purely local, and involves only the position (and possibly velocity) of the particle inside the volume (at the world point Q in Fig. 3). The description of a position measurement by such a *local* interaction is, however, incompatible with an instantaneous action-at-a-distance description of interparticle interaction because the world point at which the second particle's motion is first perturbed becomes frame-dependent.[36] Thus, a local perturbation at Q on the first particle's world line would effect the second particle at R if the computation is performed in the unprimed frame and at R' if the computation is performed in the primed frame. The conclusion we draw is that the local measurement interaction must be supplemented by a nonlocal interaction with the second particle which compensates for the instantaneous transmission of

[34] The condition (99) has been given previously by the author [Ref. 11, Eq. (83)].

[35] A two-parameter family of examples of degenerate (first-order) dynamics can be obtained by regarding the invariant first integrals (87) as the specification of a first-order dynamics. Solving these for v_1 and v_2 yields $v_1 = \theta(x_{12})(x_{12}^4 - 2Bx_{12}^2 + A^2)^{\frac{1}{2}}(A + x_{12}^2)^{-1}$, and $v_2 = \theta(x_{12})(x_{12}^4 - 2Bx_{12}^2 + A^2)^{\frac{1}{2}}(A - x_{12}^2)^{-1}$, where $\theta(x_{12})$ is either $+1$ or -1 and changes sign when the square root becomes zero in the course of the motion. Currie's example results from the choice $A = 0$, $B = -\frac{1}{2}e$; all of these degenerate examples can be obtained from a Hamiltonian scheme such as that used by Currie (generators $P = p_1 + p_2$, $H = p_1v_1 + p_2v_2$, and $K = p_1v_1x_1 + p_2v_2x_2$, where p_i and the physical coordinate x_i are canonical mates).

[36] The fact that local external interactions are incompatible with instantaneous interparticle interactions has been recognized by many authors. Various interpretations of this fact have been made by different authors; thus Thomas (Ref. 4) concluded that non-invariant world lines must be employed while Currie (Ref. 6) concluded that external interaction is incompatible with an instantaneous interaction formalism.

the local perturbation in such a way that all observers predict the same alteration in the second particle's motion. A determination of the form of the required nonlocal compensating interaction must await a detailed investigation of the extension of the instantaneous interparticle interaction scheme to include a description of external interactions.[37] If this extension should turn out to be nonunique, questions about measurements made while the particles are interacting might well be unanswerable in the sense that the answers depend on the way the extension is made.

The necessarily nonlocal character of position measurements in an instantaneous interaction theory then raises the following question with respect to the simultaneous measurability (with arbitrary precision) of positions in the corresponding quantum theory; are the operators whose commutation properties determine whether or not simultaneous position measurements interfere with one another the quantum-mechanical counterparts of the classical positions $\mathbf{x}_i$ or the necessarily nonlocal operators involved in the description of the measurement act itself? It is the present author's belief that $[\mathbf{x}_i, \mathbf{x}_j] \neq 0$ should not be interpreted as implying an interference between simultaneous position measurements in an instantaneous interaction quantum theory; it is the commutation properties of the operators involved in the description of the (possibly idealized) measurement interaction which must be looked at in answering this question.

VII. SUMMARY

A relativistic Hamiltonian mechanics has been formulated by beginning at a Newtonian level with an instantaneous action-at-a-distance relativistic mechanics formulated as a system of second-order differential equations in the physical coordinates. The Hamiltonian scheme was obtained by finding a variational principle on the positions and velocities which was reduced to Hamilton's principle by solving Pfaff's problem. An investigation of the invariance

transformations of the dynamics led naturally to the identification of a normal subgroup $\mathcal{K}$ of transformations which left individual physical states unchanged. The factor group of transformations modulo $\mathcal{K}$ was seen to be isomorphic with the subgroup of transformations which left the time fixed; from this we concluded that it was sufficient to consider only transformations which leave the time fixed despite the fact that time is not absolute in special relativity.

We have seen in terms of specific examples that an Hamiltonian scheme with the transformations of the inhomogeneous Lorentz group canonical is not always possible for a Lorentz invariant particle dynamics, and that many canonically inequivalent schemes are possible, if one is possible. The fact that canonically inequivalent schemes lead to inequivalent quantum theories then raises the question of which scheme to use as a suitable basis for formulating a quantum theory. This question has been answered previously[32] for dynamical equations which admit free particle motion asymptotically.

A new proof of the Currie–Jordan–Sudarshan zero-interaction theorem which highlights the importance of the assumption that physical coordinates can be canonical has been constructed; we have suggested that their theorem may be appropriately regarded as a *reductio ad absurdum* proof of the fact that $[\mathbf{x}_i, \mathbf{x}_j] \neq 0$ in the interaction region of a relativistic Hamiltonian dynamics. It has been pointed out that $[\mathbf{x}_i, \mathbf{x}_j] \neq 0$ does not necessarily imply an interference between simultaneous position measurements in a quantum theory corresponding to the present classical theory because such an interpretation neglects the necessarily nonlocal character of an external measurement interaction in a relativistic theory with instantaneous interparticle interactions.

---

[37] We expect that such an extension should be possible by virtue of the fact that a theory all of whose predicted motions are to be calculated by perturbing about the second-order differential equations describing free particles is equivalent to a description of these motions by second-order differential equations.

Expansions in Spherical Harmonics. IV.
Integral Form of the Radial Dependence*

ROBERT A. SACK

*Department of Mathematics, University of Salford, Salford, England† and
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin*

A function $f(r_{AB})Y_L^M(\theta_{AB}, \varphi_{AB})$ of a vector $\mathbf{r}_{AB} = \Sigma\, \mathbf{r}_i$ can be expanded in spherical harmonics $Y_l^m(\theta, \varphi)$ of the directions of the individual vectors. The radial coefficients satisfy simple differential equations which, in three previous papers, were solved in terms of series in r_i^2/r_j^2; these were different in various regions, depending on the relative magnitudes of the r_i. In this paper the solutions are found as multiple integrals over the product of Legendre polynomials and of a function $G(w)$, where w depends linearly on the r_i. The kernel $G(w)$ is independent of the number of constituent vectors, their relative sizes, and the orders of their harmonics; it contains the Heaviside step function $H(w)$ as a factor which takes care of the various regions. The precise form of G can be found from f and L by an integral equation which for $L = 0, 1$ is solved for arbitrary f, and for $L > 1$ for sufficiently large positive powers. The expressions of Milleur, Twerdochlib, and Hirschfelder for the bipolar angle average can be obtained simply by repeated integration of $G(w)$ or directly from the differential equations. For the inverse distance between two points, $G(w)$ becomes Dirac's delta function; the number of integrations is thereby reduced by one. Possible applications of the new approach to the evaluation of molecular many-center integrals are outlined. Some corrections are given for the results of the previous papers in the series.

I. INTRODUCTION

IN a series of three papers[1] the writer has presented a number of expansions for a function of the distance r_{12} between two points Q_1 and Q_2, which are specified by $\mathbf{r}_1$ $(r_1, \theta_1, \varphi_1)$ and $\mathbf{r}_2$ $(r_2, \theta_2, \varphi_2)$ about a common origin O or two distinct origins O_1 and O_2; the directions of the polar axes and of the planes defining $\varphi = 0$ are parallel throughout (cf. Fig. 1). The dependence on each angle, including θ_3 and φ_3 where $\mathbf{r}_3$ $(r_3 = a, \theta_3, \varphi_3)$ is the vector O_1O_2, is given by surface spherical harmonics, expressed either in their unnormalized forms

$$\Omega_l^m(\theta, \varphi) = e^{im\varphi}P_l^m(\cos\theta) \qquad (1a)$$

or normalized forms

$$Y_l^m(\theta, \varphi) = \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!}\right]^{\frac{1}{2}} e^{im\varphi}P_l^m(\cos\theta), \quad (1b)$$

where the associated Legendre functions $P_l^m(u)$ are defined by the standard Rodrigues formula:

$$P_l^m(u) = (-)^m \frac{(1-u^2)^{m/2}}{2^l l!}\left(\frac{d}{du}\right)^{l+m}(u^2-1)^l. \quad (1c)$$

The expansion for a function

$$V = f(\mathbf{r}_{AB}) = f(\mathbf{r}_1 + \mathbf{r}_2 + \cdots + \mathbf{r}_\nu) \qquad (2a)$$

and, more particularly, for an isotropic function

$$V = f(\mathbf{r}_{AB}) = f(|\mathbf{r}_1 + \mathbf{r}_2 + \cdots + \mathbf{r}_\nu|) \qquad (2b)$$

can always be written as[2]

$$V = \sum_{lm}\left[\prod_{i=1}^{\nu}\Omega_{l_i}^{m_i}(\theta_i, \varphi_i)\right]\cdot R(f; \mathbf{l}, \mathbf{m}; \mathbf{r}). \qquad (3)$$

Here the vectors $\mathbf{l}$, $\mathbf{m}$, and $\mathbf{r}$ denote the sets of ν values l_i, m_i, and r_i, respectively; they are not geometric vectors in the three-dimensional space such as $\mathbf{r}_{AB}$. The summation in (3) in general is to be taken over each l_i from 0 to ∞ and each m_i from $-l_i$ to l_i. The only cases hitherto considered in detail have been $\nu = 2$ and $\nu = 3$, which are relevant for the one-center and two-center expansions, respectively.

The basis of the theory developed in I–III was that, since V depends on each x_i only through the linear combination $\sum x_i$, the derivatives $\partial/\partial x_i$ are the same for all i and, correspondingly, for $\partial/\partial y_i$ and $\partial/\partial z_i$. In particular,

$$\nabla_1^2 V = \nabla_2^2 V = \cdots \nabla_\nu^2 V, \qquad (4)$$

which, when substituted in (3), yields for each individual $R(\mathbf{l}, \mathbf{m}; \mathbf{r})$

$$(\partial^2/\partial r_i^2 + (2/r)\partial/\partial r - l_i(l_i+1)/r^2)R = \text{invariant}$$
$$(i = 1, 2\cdots\nu). \quad (5)$$

By solving (5) together with the appropriate boundary conditions for

$$V = r_{AB}^N \quad \text{and} \quad V = r_{AB}^N\Omega_L^M(\theta_{AB}, \varphi_{AB}), \quad (6a, b)$$

† Permanent address.

[1] R. A. Sack, J. Math. Phys. 5, 246, 252, 260 (1964). (Though these papers do not carry the same global title, they are to be considered as Parts I, II, and III of the present series and hereafter are referred to as such.)
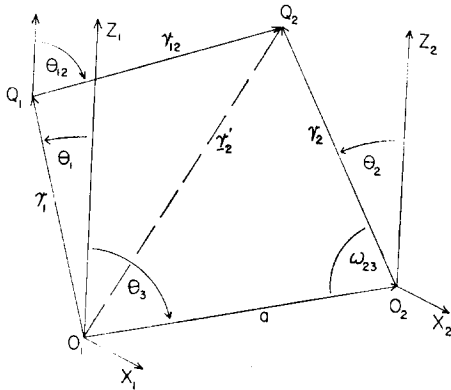
[2] The change of $\mathbf{r}_1$ in $\mathbf{r}_{AB}$ to $-\mathbf{r}_1$ in r_{12} simply introduces a factor ± 1 in the terms in (3) depending on the parity of l_1.

FIG. 1. Polar coordinates for two-center expansion. The angles φ_i are not shown to avoid cluttering up the diagram.

the writer was able to derive explicit expressions for R in terms of hypergeometric functions (Appell functions for $\nu = 3$)[3]; some of the one-center expansions for isotropic V had already been found by Chapman[4] using a different approach. The explicit forms of the radial functions differ according to the relative sizes of the r_i; for $\nu = 2$ there are two regions

$$S_1: r_1 > r_2, \quad S_2: r_2 > r_1; \tag{7a}$$

for $\nu = 3$ there are four regions as first shown by Buehler and Hirschfelder[5]:

$$S_1: r_1 > r_2 + r_3; \quad S_3: r_3 > r_1 + r_2;$$
$$S_2: r_2 > r_1 + r_3; \quad S_4: |r_1 - r_2| < r_3 < r_1 + r_2 \tag{7b}$$

(see Fig. 2). Whereas in the case (6a) the series expressions for R are convergent and reasonably simple in the outer regions S_1, S_2, and S_3, the corresponding series in the overlap region S_0 are, in general, divergent, and the only explicit series obtained in III were for integer $N = -1, 0, 1, 2 \cdots$; even these were not likely to be of great practical use, e.g., for numerical integrations for large l_i, because of the partial cancellation of large terms with a small algebraic sum. In the one-center expansion of isotropic functions of the type (2b), quadratic transformations applied to the hypergeometric functions, in I, led to expressions for R which were symmetric in r_1 and r_2; some of these had already been derived by Chapman[4] and by Fontana[6], but the appearance of powers of $(r_1 + r_2)$ or $(r_1^2 + r_2^2)$ in the denominator seems to preclude their usefulness for most practical purposes. Fontana[6] has also outlined an approach to an analogous symmetric two-center expansion, but, for
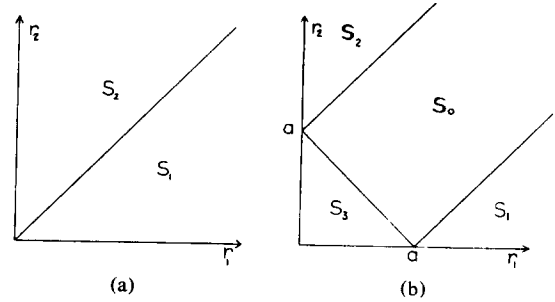

(a)   (b)

FIG. 2. The various regions: a. The one-center case. b. The two-center case.

reasons to be given in a later paper, these expansions would not be absolutely convergent everywhere, quite apart from the complicated analytic form they would take.

A rather different approach to the two-center expansion has been taken independently by Nozawa[7] and by Chiu.[8] These authors are essentially concerned with the solution of the Laplace or the wave equations $\nabla^2 V = \pm k^2 V$ or 0, but their method can be directly applied to any function V. They first break up the vector $\mathbf{r}_{12}$ into $\mathbf{r}_1$ and $\mathbf{r}_2'$ where $\mathbf{r}_2' = O_1 Q_2$ (cf. Fig. 1) and employ the usual one-center expansion, and then similarly re-expand the dependence on $\mathbf{r}_2'$ in terms of $\mathbf{r}_2$ and $\mathbf{r}_3$. As a result they obtain different expressions for R in three regions only:

$$S_1': r_1 > r_2';$$
$$S_2': r_1 < r_2', \quad r_2 > r_3; \tag{8}$$
$$S_3': r_1 < r_2', \quad r_3 > r_2$$

(there is no subdivision of S_1'). The regions S_1, S_2, and S_3 of (7b) are completely contained in their primed counterparts and the expressions obtained are obviously pairwise identical, and, in addition, the overlap region S_0 is split up between the three regions of (8). But as the magnitude of $\mathbf{r}_2'$ depends on the angle ω_{23} between $\mathbf{r}_2$ and $\mathbf{r}_3$, the boundary between S_1' on the one hand and S_2' or S_3' on the other depends on θ_2, φ_2, θ_3, and φ_3; and in the expansion (3) the variables are no longer strictly separated, as the radial coefficients R involve the angles. In consequence, it is no longer possible to use the orthogonality relations for the surface harmonics to carry out the integration over the angles. And any attempt to extend the validity of the expressions for the outer regions into S_0 in such a way that the boundaries are independent of the angles, e.g., by using the formulas for S_i whenever r_i

[3] Bateman Manuscript Project, *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1953), Chaps. 2 and 5.

[4] S. Chapman, Quart. J. Pure Appl. Math. **85**, 16 (1914).

[5] R. J. Buehler and J. O. Hirschfelder, Phys. Rev. **83**, 628 (1951); **85**, 149 (1952).

[6] P. R. Fontana, J. Math. Phys. **2**, 825 (1961).

[7] R. Nozawa, Busseiron-Kenkyu **78**, 35, 75 (1954); J. Math. Phys. **7**, 1841 (1966).

[8] Y. N. Chiu, J. Math. Phys. **5**, 283 (1964).

is the largest of the three vectors, will make the expansion (3) diverge, as was already implicit in the work of Carlson and Rushbrooke[9] on r_{12}^{-1}; though these authors avoid any explicit mention of the overlap region, they specify regions in which the other formulas converge, and these exclude S_0. If additional factors multiplying V enforce convergence after integration over angles, the results obtained in S_0 from formulas valid in S_i are likely to be erroneous. For these reasons the writer considers any expansion of the form (3) which ignores the distinct form of the radial coefficient R in S_0; while not necessarily incorrect, at any rate, it is not very useful for most practical purposes.

Recently a new approach to the two-center expansion has been made by Milleur, Twerdochlib, and Hirschfelder.[10] For an isotropic function $f(r_{12})$ they obtained simple expressions for the angle average $\langle f \rangle = R(f; 0,0; r)$ by explicit integration over angles. The results involve the r_i only as linear combinations $(\pm r_1 \pm r_2 \pm r_3)$, and the functions appearing are obtained from f by integration so that the method is applicable to fractional powers and to piecewise continuous functions to which the series expansions for R derived in III cannot be applied. Two questions posed themselves immediately:

(a) Could the closed form expressions for the angle average be obtained more simply as solutions of the differential equations (5) with the appropriate boundary conditions?

(b) Could the general solution of the equations (5) with arbitrary l_i be given in a form which preserved the linearity in the r_i?

Both these problems were quickly resolved, and the new derivations are presented in Secs. 2 and 3, respectively. It was too much to expect closed expressions for the solutions of (b); hence attempts in this direction were quickly abandoned. Instead, solutions were successfully sought in terms of an integral over a function $G(w)$ where w is a linear function of the r_i. This function was found to be independent of the l_i and of the regions S_i; its exact form is determined by a Volterra-type integral equation involving $V(r)$ and the Heaviside unit step function

$$H(w) = 1, \quad w \geq 0,$$
$$H(w) = 0, \quad w < 0, \tag{9}$$

the derivative of which is Dirac's delta function $\delta(w)$. The explicit solution of this equation was obtained for

functions V given by (2b) or (6b)—in the latter case only for sufficiently large N.

The intervention of the factor $H(w)$ automatically takes care of the different analytic forms of the integrals in the various regions; its influence on the solution and special forms of the results are discussed in Sec. 4, and in Sec. 5 some applications of the new approach for the evaluation of 2-, 3-, and 4-center integrals are outlined.

II. THE ANGLE AVERAGE

The formula derived by Milleur, Twerdochlib, and Hirschfelder[10] for the angle average $\langle f \rangle$, i.e., for the radial coefficient $R(0, 0; r)$, of a spherically symmetric function $f(r_{AB})$ or $f(r_{12})$ with two origins is

In S_0: $\langle f \rangle = (4r_1 r_2 r_3)^{-1}[h(r_1 + r_2 + r_3)$
$\qquad - h(r_1 + r_2 - r_3) - h(r_1 - r_2 + r_3)$
$\qquad - h(r_2 - r_1 + r_3)], \tag{10a}$

In S_1: $\langle f \rangle = (4r_1 r_2 r_3)^{-1}[h(r_1 + r_2 + r_3)$
$\qquad - h(r_1 + r_2 - r_3) - h(r_1 - r_2 + r_3)$
$\qquad + h(r_1 - r_2 - r_3)], \tag{10b}$

where

$$h(w) = \int_0^w v(w - v)f(v)\, dv. \tag{11}$$

The expressions valid in S_2 and S_3 are obtained from (10b) by permutation of the indices. Substitution of this integral in (10) shows that the lower limit of integration is immaterial in S_1, though not in S_0; in consequence a singularity of $f(r_{12})$ at $r_{12} = 0$ will not show up in $\langle f \rangle$ in the outer regions, but may crucially affect the result in the overlap region. Thus, for $f(r_{12}) = r_{12}^N$, (10b) remains meaningful for all N, provided the lower limit in (11) is taken to be $\epsilon > 0$; on the other hand, convergence of the individual terms in (10a) requires $N > -3$. The formulas (10) and (11) have a well-defined meaning for any function $f(r_{12})$ which is integrable for all nonnegative r_{12}; no analyticity of f need be assumed, as was required in III.

The expressions (10) and (11) were obtained by Milleur, Twerdochlib, and Hirschfelder[10] by integrations over the geometric angles. To derive the same results as solutions of differential equations, we must put all the l_i equal to zero in (5). Consider first the one-center case $\nu = 2$ and put

$$T = r_1 r_2 \langle f \rangle. \tag{12}$$

For this function (5) becomes

$$\partial^2 T/\partial r_1^2 = \partial^2 T/\partial r_2^2, \tag{13}$$

which is d'Alembert's equation with the well-known

[9] B. C. Carlson and G. S. Rushbrooke, Proc. Cambridge Phil. Soc. **46**, 215 (1950).

[10] M. B. Milleur, M. Twerdochlib, and J. O. Hirschfelder, J. Chem. Phys. **45**, 13 (1966).

general solution

$$T = g_1(r_1 + r_2) + g_2(r_1 - r_2). \tag{14a}$$

Hence $\langle f \rangle$ must be of the form

$$\langle f \rangle = (r_1 r_2)^{-1}[g_1(r_1 + r_2) + g_2(r_1 - r_2)]. \tag{14b}$$

However, as r_2 tends to zero in the region S_1, $\langle f \rangle$ tends to $f(r_1)$; hence, by L'Hôpital's rule,

$$g_1(r_1) \equiv -g_2(r_1) = \tfrac{1}{2}g(r_1), \tag{15a}$$

$$r_1^{-1} dg(r_1)/dr_1 = f(r_1), \tag{15b}$$

or

$$g(w) = \int_0^w v f(v)\, dv, \tag{16}$$

although again the lower limit of integration is essentially arbitrary. The angle average is thus

$$\text{In } S_1\colon \langle f \rangle = [g(r_1 + r_2) - g(r_1 - r_2)](2r_1 r_2)^{-1}, \tag{17a}$$

$$\text{In } S_2\colon \langle f \rangle = [g(r_1 + r_2) - g(r_2 - r_1)](2r_1 r_2)^{-1}, \tag{17b}$$

with $g(w)$ defined in (16); the second formula follows from the first by symmetry. For the two-center case ($\nu = 3$) we put

$$T = r_1 r_2 r_3 \langle f \rangle, \tag{18}$$

for which (5) becomes

$$\partial^2 T/\partial r_1^2 = \partial^2 T/\partial r_2^2 = \partial^2 T/\partial r_3^2. \tag{19}$$

Considering the three sides of this equation in pairs, we see that each r_i is coupled to the others by addition or subtraction, as in (14a); hence $\langle f \rangle$ is of the form

$$\langle f \rangle = (r_1 r_2 r_3)^{-1}[h_1(r_1 + r_2 + r_3) + h_2(r_1 + r_2 - r_3)$$
$$+ h_3(r_1 - r_2 + r_3) + h_4(r_1 - r_2 - r_3)]. \tag{20}$$

As r_3 tends to zero in S_1, (20) must tend to (17a), and a renewed application of L'Hôpital's rule yields the solution (11) and (10b). The formulas valid in S_2 and S_3 follow from symmetry, and (10a) can be deduced as the only function of the form (20) which smoothly links the known solutions in the outer regions.

Both (16), (17) and (10), (11) can be written in a form independent of the region in which they apply by making use of the Heaviside function $H(w)$:

$$\langle f \rangle = (2r_1 r_2)^{-1} \sum_\sigma (-)^{\sigma_1 + \sigma_2} g(\sigma_1 r_1 + \sigma_2 r_2)$$
$$\times H(\sigma_1 r_1 + \sigma_2 r_2); \tag{21a}$$

$$\langle f \rangle = (4r_1 r_2 r_3)^{-1} \sum_\sigma (-)^{\sigma_1 + \sigma_2 + \sigma_3} h(\sigma_1 r_1 + \sigma_2 r_2 + \sigma_3 r_3)$$
$$\times H(\sigma_1 r_1 + \sigma_2 r_2 + \sigma_3 r_3), \tag{21b}$$

where each σ_i can take the values ± 1 independently. The step function thus takes care of the various regions by eliminating terms of negative argument; there should be no ambiguity on the boundaries of the regions as long as $g(0)$ or $h(0)$ vanish, i.e., as long as $vf(v)$ (or its integral) is integrable at $v = 0$.

The generalization of (10), (11) for the angle

average of a function $f(r_{AB})$, where $\mathbf{r}_{AB}$ is composed of an arbitrary number ν of vectors [cf. (2b)], follows easily by induction:

$$R(0, 0; \mathbf{r}) = 2^{1-\nu}(\pi r_i)^{-1} \sum_\sigma (-)^{\Sigma \sigma_i} g_\nu \left(\sum_1^\nu \sigma_i r_i \right), \tag{22a}$$

where

$$g_1(w) = w f(w) H(w); \quad g_\nu(w) = \int_0^w g_{\nu-1}(v)\, dv. \tag{22b}$$

III. INTEGRAL SOLUTIONS FOR ARBITRARY l_i

A. General Form of the Solutions

The differential equations (5) for the radial coefficients R do not involve the azimuthal quantum numbers m_i, which can, therefore, affect the solution only in the form of a constant factor; hence the functions R can always be decomposed into two factors, one depending on l and m only, the other on l and r as well as on the nature of the functions V to be expanded:

$$R(V; \mathbf{l}, \mathbf{m}; \mathbf{r}) = K''(\mathbf{l}, \mathbf{m}) \times R''(V; \mathbf{l}; \mathbf{r}). \tag{23}$$

As already pointed out in II and III, this partitioning is not unique, as any dependence on the l only may be drawn into either factor.

As mentioned in the introduction, the chief aim of the present investigation was to establish solutions of (5) involving the r_i only in linear combinations, i.e., in the form

$$G(w) = G(r_1 u_1 + r_2 u_2 + \cdots + r_\nu u_\nu) \tag{24}$$

summed or integrated over various values of u_i. As the solution was bound to involve the rotational quantum numbers l_i in some way, the most obvious trial solution was

$$R''(\mathbf{l}, \mathbf{r}) = \int_{-1}^1 \cdots \int_{-1}^1 du_1\, du_2 \cdots du_\nu$$
$$\times G(w) P_{l_1}(u_1) P_{l_2}(u_2) \cdots P_{l_\nu}(u_\nu). \tag{25}$$

Assuming that $G(w)$ has a second derivative everywhere and applying the r_1 operator of (5) to the first integral in (25) only, we obtain

$$\left(\frac{\partial^2}{\partial r_1^2} + \frac{2}{r_1} \frac{\partial}{\partial r_1} - \frac{l_1(l_1 + 1)}{r_1^2} \right) \int_{-1}^1 G(w) P_{l_1}(u_1)\, du_1$$

$$= \int_{-1}^1 \left[u_1^2 G''(w) + \frac{2u_1}{r_1} G'(w) \right.$$
$$\left. - \frac{l_1(l_1 + 1)}{r_1^2} G(w) \right] P_{l_1}(u_1)\, du_1,$$

$$= \int_{-1}^1 G''(w) P_{l_1}(u_1)\, du_1$$
$$+ \int_{-1}^1 \left[\frac{d}{du_1} \frac{(u_1^2 - 1)G'(w)}{r_1} \right.$$
$$\left. - \frac{l_1(l_1 + 1)G(w)}{r_1^2} \right] P_{l_1}(u_1)\, du_1, \tag{26}$$

where the last integral vanishes on integration by parts. The r_1 operator in (5) applied to the trial function R'' of (25) yields thus a similar v-fold integral, with $G''(w)$ replacing $G(w)$; the resulting expression is thus invariant whichever particular operator in (5) is chosen, and (25) is indeed a solution of (5). The general nature of G makes it likely that (25) represents the general solution of (5) except, possibly, for some singular solutions. Although for the purposes of the proof it has been assumed that $G(w)$ possesses second derivatives everywhere, this is not a necessary condition, as even a discontinuous $G(w)$ can be treated as the limit of a sequence of functions with second derivatives.

Having thus established the general nature of the solution of (5), we next have to show that, for the expansion of a function

$$V = f(r_{AB})\Omega_L^M(\theta_{AB}, \varphi_{AB}) \tag{27}$$

[which includes (2b) as a special case] with a suitable choice of the factor K'', the function $G(w)$ is independent of the number of component vectors v and the rotational quantum numbers l_i. Once this has been established, it remains to determine the dependence of G on f and L.

B. Invariance of G

The transformation properties of the spherical harmonics Ω_l^m or Y_l^m of (1) under rotation require that the coefficients of each individual term in (3) involve the azimuthal quantum numbers m only through the integrals (generalized Gaunt's coefficients)

$$I_\Omega\begin{pmatrix} l_1 & l_2 & \cdots & l_s \\ m_1 & m_2 & \cdots & m_s \end{pmatrix} = \int_{-1}^{1} \prod_1^s P_{l_j}^{m_j}(u) \cdot du \tag{28}$$

or

$$I_Y'\begin{pmatrix} l_1 & l_2 & \cdots & l_s \\ m_1 & m_2 & \cdots & m_s \end{pmatrix}$$
$$= \int_0^\pi \int_0^{2\pi} \prod_1^s Y_{l_j}^{m_j}(\theta, \varphi) \cdot \sin\theta \, d\theta \, d\varphi. \tag{29}$$

The latter integrals vanish unless

$$\sum_1^s m_j = 0, \tag{30a}$$

$$\sum_1^s l_j = \text{even} = 2\Lambda, \tag{30b}$$

$$\Lambda - l_j \geq 0 \quad (j = 1, 2 \cdots s). \tag{30c}$$

Here (30b) follows from parity considerations and (30a) and (30c) from the orthogonality relations of the spherical harmonics. The integrals I_Ω in (28) do not necessarily vanish if (30a) is violated; but as the only

integrals of importance are those for which (30a) is valid, we assume the relation must hold. In view of the writer's personal preference for unnormalized harmonics, the derivation will be given in terms of these functions; some of the formulas required in this section are derived in Appendix A, as their presentation here would interrupt the flow of the argument.

If in the expansion of (27) we put for the m-dependent coefficient K'' of R [cf. (3) and (23)]

$$K''(\mathbf{l}, \mathbf{m}) = (-)^M \prod_1^v (l_i + \tfrac{1}{2})$$
$$\times I_\Omega\begin{pmatrix} L & l_1 & l_2 & \cdots & l_v \\ M & -m_1 & -m_2 & \cdots & -m_v \end{pmatrix}, \tag{31}$$

the I_Ω is a consequence of the transformation properties, the additional factors have been chosen for convenience. To show that with this choice of K'' the function $G(w)$ in (24) and (25) is indeed independent of v and $\mathbf{l}$, we note that, by making $r_v = 0$, the only dependence of the integrand of (25) on u_v is through the Legendre polynomial; hence by orthogonality all the integrals (25) vanish, unless $l_v = 0$, in which case the integral is just twice that obtained with r_v and u_v missing; at the same time, (31) has exactly half the value it would have in the absence of $l_v + \tfrac{1}{2}$. Hence for the expansion (3) to be invariant under the addition of an arbitrary number of zero vectors, $G(w)$ must be invariant under the accretion or deletion of an arbitrary number of u_j terms with $l_j = 0$.

More generally, we can show that the invariance of G in (25) ensures the identity of the expansion (3) whether we take two radii (say $\mathbf{r}_1$ and $\mathbf{r}_2$) in the same direction or take a single vector of magnitude $r_1 + r_2$. Collecting only those factors in (3), (25), and (31) which depend on $(\theta_1, \varphi_1) = (\theta_2, \varphi_2)$, l_1 and l_2 and summing over these, we obtain, in the one case, the contribution

$$J_1 = \sum_{l_1 m_1} \sum_{l_2 m_2} \iint \Omega_{l_1}^{m_1}(\theta, \varphi)\Omega_{l_2}^{m_2}(\theta, \varphi)$$
$$\times (l_1 + \tfrac{1}{2})(l_2 + \tfrac{1}{2})I_\Omega\begin{pmatrix} L & l_1 & l_2 & \cdots & l_v \\ M & -m_1 & -m_2 & \cdots & -m_v \end{pmatrix}$$
$$\times P_{l_1}(u_1)P_{l_2}(u_2)G(r_1u_1 + r_2u_2 + \cdots) \, du_1 \, du_2, \tag{32a}$$

and, in the other,

$$J_2 = \sum_{lm} \int \Omega_l^m(\theta, \varphi)(l + \tfrac{1}{2})I_\Omega\begin{pmatrix} L & l & \cdots & l_v \\ M & -m & \cdots & -m_v \end{pmatrix}$$
$$\times P_l(u_1)G[(r_1 + r_2)u_1 + \cdots] \, du_1. \tag{32b}$$

The products of two spherical harmonics in (32a) can be expanded in the usual way as sums of single

harmonics

$$\Omega_{l_1}^{m_1}(\theta, \varphi)\Omega_{l_2}^{m_2}(\theta, \varphi)$$

$$= \sum_l (-)^m I_\Omega \begin{pmatrix} l & l_1 & l_2 \\ -m & m_1 & m_2 \end{pmatrix}(l + \tfrac{1}{2})\Omega_l^m(\theta, \varphi),$$

$$m = m_1 + m_2, \quad (33)$$

whereas the single integral in (32b) can be converted into a double one with the same argument for G as in (32a) by multiplying with the delta function

$$\delta(u_1 - u_2) = \sum (l_2 + \tfrac{1}{2})P_{l_2}(u_1)P_{l_2}(u_2). \quad (34)$$

On substituting (34) in (32b) and expanding the products of Legendre functions of u_1 the same way as in (33), as the factor of

$$(l + \tfrac{1}{2})(l_1 + \tfrac{1}{2})(l_2 + \tfrac{1}{2})\Omega_l^m(\theta, \varphi)P_{l_1}(u_1)P_{l_2}(u_2)G(w)$$

$$(35a)$$

we obtain

$$I_\Omega \begin{pmatrix} L & l & \cdots & l_v \\ M & -m & \cdots & -m_v \end{pmatrix} I_\Omega \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix}. \quad (35b)$$

From (32a) and (33) this factor becomes

$$(-)^m \sum_{m_1} I_\Omega \begin{pmatrix} l & l_1 & l_2 \\ -m & m_1 & m - m_1 \end{pmatrix}$$

$$\times I_\Omega \begin{pmatrix} L & l_1 & l_2 & \cdots \\ M & -m_1 & m_1 - m & \cdots \end{pmatrix} \quad (35c)$$

$$= \sum_{m_1 \lambda} I_\Omega \begin{pmatrix} l & l_1 & l_2 \\ -m & m_1 & m - m_1 \end{pmatrix}(\lambda + \tfrac{1}{2})$$

$$\times I_\Omega \begin{pmatrix} L & \lambda & \cdots \\ M & -m & \cdots \end{pmatrix} I_\Omega \begin{pmatrix} \lambda & l_1 & l_2 \\ m & -m_1 & m_1 - m \end{pmatrix}, \quad (35d)$$

$$= \sum_\lambda \delta_{l\lambda} I_\Omega \begin{pmatrix} L & \lambda & \cdots \\ M & -m & \cdots \end{pmatrix} I_\Omega \begin{pmatrix} l & l_1 & l_2 \\ 0 & 0 & 0 \end{pmatrix} \quad (35e)$$

in view of (A3) and (A6). The expressions (32a) and (32b) are thus identical, provided $G(w)$ does not depend on v or the l. While this is not a rigorous proof that $G(w)$ *must* be independent of these quantities, it makes it more than plausible. A complete proof would have to show that the reduction for v to $v - 1$ vectors gives identical results even when $(\theta_1, \varphi_1) \neq (\theta_2, \varphi_2)$; the writer has been able to derive such a proof, but since it involves quite a number of intermediary lemmas, it is omitted here.

If the spherical harmonics in both (3) and (27) are given in normalized form, we obtain for the factor K_Y'' from (1), (29), and (31)

$$K_Y''(\mathbf{l}, \mathbf{m})$$

$$= (-)^M (2\pi)^{v-1} I_Y' \begin{pmatrix} L & l_1 & l_2 & \cdots & l_v \\ M & -m_1 & -m_2 & \cdots & -m_v \end{pmatrix}. \quad (36)$$

C. Relation of f and G

Having established the invariance of $G(w)$ with v and $\mathbf{l}$, it is an easy matter to find the exact relation of $G(w)$ to f and L in (27). One simply has to put $v = 1$, in which case only one term in the expansion (3) survives in view of the orthogonality relations, and the equation to be solved becomes

$$f(r) = \int_{-1}^{1} G(ru)P_L(u)\, du. \quad (37)$$

Here it should be noticed that r can by definition take only real nonnegative values and $f(r)$ can to some extent be chosen arbitrarily for $r < 0$. The easiest way is to multiply $f(r)$ by the Heaviside function $H(r)$; any ambiguities arising from branch points of $f(r)$ at $r = 0$ are thereby automatically eliminated. Correspondingly, we may choose $G(w) \equiv 0$ for $w < 0$, so that (37) becomes

$$f(r)H(r) = \int_0^1 G_L(ru)P_L(u)\, du, \quad (38)$$

where the suffix L has been added to G to indicate which Legendre polynomial enters into the transform. For $L = 0$ we obtain

$$f(r)H(r) = \int_0^r G_0(w)\left(\frac{dw}{r}\right), \quad (39)$$

with the solution

$$G_0(w) = (d/dw)[wf(w)H(w)]. \quad (40)$$

For $L = 1$ the corresponding equation and solution are

$$f(r)H(r) = \int_0^r G_1(w)(w/r)(dw/r), \quad (41)$$

$$G_1(w) = w^{-1}(d/dw)[w^2f(w)H(w)]. \quad (42)$$

The solutions for the transforms in (38) with $L > 1$ are less straightforward in general, and the only case discussed in the present paper is that of a real power $f(r) = r^N$. It is obvious from (38) that $G_L(w)$ must also be proportional to the same power

$$G_L(w) = C_{LN}w^N H(w), \quad (43)$$

where C_{LN} is the reciprocal of the integral

$$\int_0^1 P_L(u)u^N\, du = \frac{\pi^{\frac{1}{2}}\Gamma(1 + N)2^{-1-N}}{\Gamma(1 + \tfrac{1}{2}N - \tfrac{1}{2}L)\Gamma(\tfrac{3}{2} + \tfrac{1}{2}N + \tfrac{1}{2}L)}$$

$$(44)$$

[cf. (3.12.23) of Ref. 3]. Hence we get

$$C_{LN} = \frac{(1 + N)(3 + N)\cdots(L + N + 1)}{(2 + N - L)(4 + N - L)\cdots N}, \quad L \text{ even,}$$

$$(45a)$$

$$= \frac{(2 + N)(4 + N)\cdots(L + N + 1)}{(2 + N - L)(4 + N - L)\cdots(N - 1)},$$

$$L \text{ odd,} \quad (45b)$$

valid for

$$N > L - 2. \tag{45c}$$

The exceptions are $L = 0$ and $L = 1$ for which the products in the denominators of (45a, b) become empty with the value unity, and hence these formulas are valid provided

$$N > -1, \quad L = 0, \tag{45d}$$

$$N > -2, \quad L = 1, \tag{45e}$$

in agreement with (40) and (42). A more detailed discussion of the solution of (38) in the general case with $L > 1$ will be given in a subsequent paper.

IV. DISCUSSION OF RESULTS

The integral expressions (23)–(25) with (31) and (40), (42), or (38) provide a general solution for the radial factors $R(l, m; r)$ in the expansion (3) of V as defined in (2); to the writer's knowledge, this form of the solution is completely new, apart from one special case mentioned after (48) below. The form of the function (25) is such that the factors R'' can be interpreted as weighted averages of another function G, not of r_{AB} itself, but of its component along a prescribed z direction. However, it should be borne in mind that the quantities u_i occurring in (25) do not represent physical direction cosines, but are simply integration variables; all the dependence on the geometric angles is contained in the spherical harmonics in (3). It is interesting to note that the partitioning of R according to (23) with the object of keeping $G(w)$ invariant leads to factors K'', and hence R'', which agree with the singly-primed factors derived in III (29), (34) for the two-center expansion for $L = 0$, and only differ from those in II (33), (37) for the one-center expansion by the factor $(-)^{\Lambda+L}$; yet the precise partitioning in II had no stronger motivation than keeping the recurrence relations between the R' as simple as possible.

The occurrence of the Heaviside function $H(w)$ as a factor in $G(w)$ in (24), (25), and (38) means that in general the integration is effectively carried out over only half the ν-dimensional hypercube $-1 \leq u_i \leq 1$, the domain on one side of the hyperplane $w = 0$ (which passes through the origin) having zero integrand. If for a particular index i

$$r_i > \sum_{j \neq i} r_j, \tag{46}$$

the surface $u_i = 1$ is everywhere a boundary of the integrated domain, and $u_i = -1$ lies wholly outside it. Thus, if the first integration is carried out over u_i the lower limit is a function of the other u_j's, but the limits of all subsequent integrations are independently
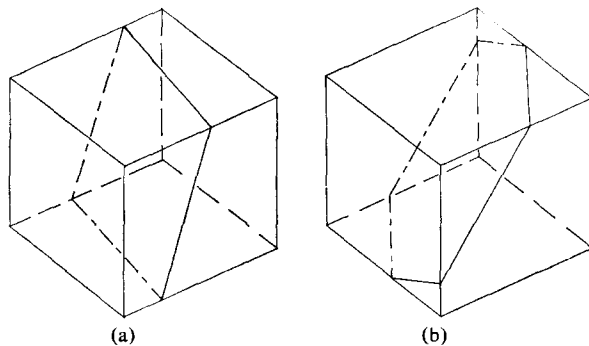

FIG. 3. Position of the plane $w = 0$ in the u cube. (a) $r_i > r_j + r_k$, (b) $|r_1 - r_2| < r_3 < r_1 + r_2$.

-1 and $+1$. If, however, none of the radii satisfies (46), no face of the cube lies entirely on one side of the separating plane, and hence the lower limits of at least two integrations are variable. The Heaviside function thus automatically sorts out the various regions S; a passage of the separating plane through a corner of the cube corresponds to the passage to another region. This is illustrated in Fig. 3 for the case $\nu = 3$.

If the integrand in (25), apart from the factor $H(w)$ of (38), is regular everywhere in the hypercube and is invariant under simultaneous change of sign of all the u_i (an even function), then the integrals over the domains $w \gtrless 0$ are identical. We may therefore drop the factor $H(w)$ and take instead half the integral over the whole hypercube; from this point of view there is no separating plane and the expansion (3) is the same in all regions. Thus, if V is a sum of terms of the form (6b), it follows from (30b), (31), (38), (43), and (45) that, for

$$N = L + 2k, \quad k = 0, 1, 2 \cdots, \tag{47}$$

the same analytic expressions for the $R(l, m; r)$ are valid for all r; this is in agreement with the fact that both the solid harmonics $r_{AB}^L Y_L^M(\theta_{AB}, \varphi_{AB})$ and all positive integral powers of r_{AB}^2 have universally valid expansion coefficients R.[1,6,11]

The evenness of the integrand without analyticity throughout the cube does not ensure the uniformity of the integral through all regions. Thus we obtain from (40) for

$$V = r_{AB}^{-1}, \quad G_0(w) = \delta(w), \tag{48a, b}$$

which is an even function even after multiplication by the $P_{l_i}(u_i)$ compatible with (30b); yet this was the case for which the existence of different regions was first established by Laplace. The identity of the coefficients $r_<^L/r_>^{L+1}$ with the double integral (25), using (23), (31), and (48b), has already been established by Nozawa and Linderberg.[7]

[11] B. Friedman and J. Russek, Quart. Appl. Math. **12**, 13 (1954).

Whenever one of the indices l_i is zero, the integration over the corresponding u_i can be carried out explicitly. In particular, the expressions (21) and (22), and hence the form (10) given by Milleur *et al.*[10] for the angle average $\langle f \rangle$, follow from a repeated integration of (23)–(25) and (40) with all $l_i = 0$ since, in view of (31),

$$K'' \begin{pmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix} = 2^{1-\nu}; \tag{49}$$

the function $G_0(w)$ is thus identical with $g_0(w)$, which would precede g_1 in the recurrence relations (22b).

The integral form of the radial coefficients R, as opposed to the series in I–III, makes possible the expansion of nonanalytic, even discontinuous functions $f(\mathbf{r}_{AB})$ in (2). A discontinuity in f will produce a delta function in $G(w)$ in view of (38)–(42). If $f(\mathbf{r}_{AB})$ diverges at $r_{AB} = 0$ but it is known from other considerations that a required integral over $\mathbf{r}_{AB}$ is convergent, one can introduce a cutoff at $r_{AB} = \epsilon > 0$ and let ϵ tend to zero later; this merely means we replace $H(w)$ by $H(w - \epsilon)$ in (38)–(42). An interesting application of this arises in the expansion of the first-order irregular solid spherical harmonic Υ_1

$$V = \cos \theta_{AB}/r_{AB}^2, \quad G_1(w) = w^{-1}\delta(w), \quad \text{(50a, b)}$$

if the formula (42) were applied uncritically, but this expression is meaningless. On introducing a cutoff, we can put from (42) and (50)

$$G_1(w) = \lim \, [\epsilon^{-1}\delta(w - \epsilon)]; \tag{50c}$$

again the expression has no limit as $\epsilon \to 0$, but we can add any even function $\gamma(w)$ to $G_1(w)$ without affecting the integral (25), as the product of the $P_l(u)$ is odd in view of (30b) and (31). The delta function $\delta(w)$ is such an even function of w; thus we can put

$$G_1(w) = \lim \epsilon^{-1}[\delta(w - \epsilon) - \delta(w)] = -\delta'(w) \tag{50d}$$

by L'Hôpital's rule. Even more elaborate tricks are required in the expansion of the higher Υ_L^M; except for $L = 2$, ϵ cannot be made to tend to zero, at least not for values of the r_i corresponding to the region S_0. Similarly, a cutoff must be introduced in the expansion for $V = r_{AB}^n$ where $n < -1$; after performing the integrations, ϵ may be reduced to zero for $-1 > n > -3$, but not in general for $n \leq -3$ (cf. Milleur *et al.*[10]).

Another possible operation on $G(w)$ follows from (26):

$$V \to G(w) \subset \nabla^2 V \to G''(w). \tag{51}$$

Applying this to (48), we obtain

$$V = \delta^3(r_{AB}) = -\frac{1}{4\pi} \nabla^2 \frac{1}{r_{AB}}; \quad G = -\frac{1}{4\pi} \delta''(w). \tag{52}$$

For the angle average this leads to the expression (22) of Milleur, Twerdochlib, and Hirschfelder[10]; for general l_i (23)–(25) integrate to the formulas given by Tanabe[12] and in III (40); as first pointed out by Milleur *et al.*, the latter formula should be divided by (-8), not only for the angle average, but for all l.

One aspect which awaits fuller investigation is the convergence of the expressions. Two main types of convergence have to be considered: (a) of the individual integrals (25) for all fixed sets r_i; (b) of the sum (3) for all fixed sets $(r_i, \theta_i, \varphi_i)$, the radial functions R being given by (23)–(25) and (31). With regard to (a) it is clear that the integrals converge whenever $G(w)$, considered as a function of a single variable, is everywhere integrable. Difficulties may arise through singularities at the origin and through the introduction of generalized functions; some practical aspects of this have been described in the preceding paragraphs. The point (b) has not been investigated at all, and I can only express my personal opinion (or hope) that the expansion will converge in most practical cases.

Another point arising in this context is the possible interchange of the order of performing the summations in (3) and the integrations in (25). It is easily shown that a summation over all l, m at fixed r, θ, φ, u may diverge; one only has to put $\nu = 2$, $L = 0$, $\theta_1 = 0$, $u_1 = u_2 = 1$, in which case the sum becomes

$$G(r_1 + r_2) \cdot \sum (l + \tfrac{1}{2}) P_l(\cos \theta_2). \tag{53}$$

On the other hand, when integrations over the angles with specific weight factors are carried out, the resulting series at constant u may easily converge; the advantage of this approach is that the integrations over the radii and the angles are thereby completely separate [which they are anyway in the expansion (3)], but in addition the linearity of $G(w)$ of (24), (25) in the r_i remains preserved; an application of this is outlined in the next section.

V. APPLICATION TO TWO-CENTER EXPANSIONS AND FURTHER RESEARCH

The main applications of the theory developed in this paper are likely to concern two-center expansions, corresponding to the case $\nu = 3$ (cf. Fig. 1). This raises the question whether any advantage is gained by identifying the axis O_1O_2 with the Z axis from the start, i.e., by putting $\theta_3 = 0$. Such an approach must be emphatically rejected in the development of the theory. The different radial coefficients R corresponding to given l_1 and l_2 in (3) now depend on m_1 instead of l_3, but their number remains the same; for instance, in the isotropic case $L = 0$ in (27), $|m_1| = |m_2|$ can

---

[12] Y. Tanabe, J. Phys. Soc. Japan **11**, 980 (1956).

take all values between 0 and $l_<$, whereas l_3 runs from $|l_1 - l_2|$ to $l_1 + l_2$ in steps of 2, the number of different values being $l_< + 1$ in either case. On the other hand, one of the sides of the differential equations (5) is lost if θ_3 is kept fixed, and the expressions derived for R are bound to be more complicated than those for fixed l_3. This is borne out by the greater regularity of the coefficients in the expansions for $1/r_{12}$ derived in III than in the formulation by Buehler and Hirschfelder[5]; a generating function which these authors give in their second paper is too cumbersome to be of practical use. Similarly Nozawa, who fixes the direction of O_1O_2,[7] is obliged to define generalized Bessel functions carrying three indices when expanding regular solutions of the wave equation; in the analogous expansion given in II for variable θ_3, the corresponding expressions are merely products of Legendre and Bessel functions which have to be added with the appropriate angular integrals as coefficients.

Once the theory has been established, there are fewer objections to fixing $\theta_3 = 0$. If in the expansion of (27) the radial coefficients corresponding to given values of l_1, m_1, l_2, m_2, and $m_3 = 0$ are summed over all l_3 including those for which R vanishes in view of the conditions (30b, c), the relevant factor becomes [in view of (28) and (34)]

$$\sum_{l_3} (l_3 + \tfrac{1}{2}) I_\Omega \begin{pmatrix} L & l_1 & l_2 & l_3 \\ M & -m_1 & -m_2 & 0 \end{pmatrix} P_{l_3}(u)$$

$$= \int \sum_{l_3} (l_3 + \tfrac{1}{2}) P_L^M(v) P_{l_1}^{-m_1}(v) P_{l_2}^{-m_2}(v) P_{l_3}(v) P_{l_3}(u)\, dv,$$

$$= \int P_L^M(v) P_{l_1}^{-m_1}(v) P_{l_2}^{-m_2}(v) \delta(u_3 - v)\, dv. \tag{54}$$

The complete radial factor R is thus obtained as

$$R(l_1, l_2, \mathbf{m}; \mathbf{r}) = (-)^M (l_1 + \tfrac{1}{2})(l_2 + \tfrac{1}{2}) \delta_{M, m_1 + m_2}$$

$$\times \int_{-1}^1 G_L(r_1 u_1 + r_2 u_2 + r_3 u_3) P_{l_1}(u_1) P_{l_2}(u_2) P_L^M(u_3)$$

$$\times P_{l_1}^{-m_1}(u_3) P_{l_2}^{-m_2}(u_3)\, du_1\, du_2\, du_3, \tag{55}$$

with possibly an additional sign change.[2] From a theoretical point of view it is very doubtful if such a formula could have been derived without going through the procedure of Sec. 3. The expression (55) itself is not too unwieldy, especially for $L = 0$; its main drawback is that the upper indices now appear inside the integral, instead of merely through the constants I_Ω in (28).

The integrals I_Ω in the two-center expansion of an isotropic interaction V as in (2c) involve three factors; in their normalized form (29) they are most easily expressed in terms of Wigner $3j$ symbols, on which there exists an extensive literature regarding both

theory and tabulation (cf. Edmonds[13] or Refs. 3, 5, 10, and 16 of II); an approach to the theory in terms of unnormalized $3j$ symbols, which are integers, was outlined in II and will be further developed in a subsequent paper in this series. Even if $L \neq 0$, the integrals are easily calculated from those with three factors by means of (A3).

The evaluation of integrals over all positions of two particles with interaction

$$\int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 \rho_1(\mathbf{r}_1) \rho_2(\mathbf{r}_2) V(\mathbf{r}_{12}) \tag{56}$$

can be turned, in view of (3), into a summation over l_1, m_1, l_2, m_2, and l_3 of integrals over the radii r_1 and r_2 that involve the radial coefficients R, the separation of origins $r_3 = a$ being kept fixed. As these coefficients themselves are expressed as triple integrals in (23)–(25), the method would involve replacing the sixfold integral (56) by a multiple sum of five-dimensional integrals of the form

$$\int_{-1}^1 \int_{-1}^1 \int_{-1}^1 \int_0^\infty \int_0^\infty d^3\mathbf{u}\, dr_1\, dr_2 G(r_1 u_1 + r_2 u_2 + r_3 u_3)$$

$$\times \chi_1(r_1) \chi_2(r_2) P_{l_1}(u_1) P_{l_2}(u_2) P_{l_3}(u_3), \tag{57}$$

which appears a most uneconomic procedure. However, if the functions $\chi_1(r_1)$ and $\chi_2(r_2)$ either are independent of l_1 and l_2 or else can be broken up into terms which possess this independence, the integrations over r_1 and r_2 could be done first for each point of the cube in $\mathbf{u}$ space, and the remaining three integrations could then be carried out numerically; such an approach would be all the more practicable when the r integrations can be performed analytically.

Fortunately the situation is more favorable in the most important case $V = 1/r_{12}$. In view of (48) the function $G(w)$ in (57) is the delta function $\delta(w)$, and the number of integrations is thereby effectively reduced by one, e.g., any variation of r_1 at constant $\mathbf{u}$ implies a definite linear dependence of r_2 on r_1. However, a further property of the delta function

$$\delta(kw) = \delta(w)/|k| \tag{58}$$

reduces the number of integrations yet further. If the integration over the radii has been carried out at a particular point (u_1, u_2, u_3) in $\mathbf{u}$ space, the corresponding integral at (ku_1, ku_2, ku_3) is simply the original one divided by $|k|$. (This argument cannot be applied if the first integration is over one of the u_i because of the finite limits.) It is therefore sufficient to evaluate the r integrals for points on the surface of the $\mathbf{u}$ cube, and the volume integrals with weight

[13] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1960), 2nd ed.

factors $\Pi P_{l_i}(u_i)$ can be expressed as surface integrals with correspondingly adjusted weights. These surface integrals would have to be evaluated numerically; a Gaussian quadrature scheme could be set up in which the points on the surface of the **u** cube are tabulated for which the r integration is to be performed, together with their weights appropriate to each triple (l_1, l_2, l_3), or (l_1, l_2, m) if the approach of (55) is used.

The most delicate part of such a scheme would be the integration over the radii; it depends on the nature of the functions $\chi(r_1)$ and $\chi(r_2)$—whether these are best performed analytically or numerically. If they are Slater functions or products thereof, analytic methods are appropriate. We may define

$$I_{st} = \int\!\!\!\int_0^\infty \exp\left(-\alpha r_1 - \beta r_2\right) r_1^s r_2^t$$

$$\times \; \delta(r_1 u_1 + r_2 u_2 + a u_3) \, dr_1 \, dr_2,$$

$$= (-\partial/\partial\alpha)^s (-\partial/\partial\beta)^t I_{00}, \qquad (59)$$

where I_{00} can be easily calculated:

$$I_{00} = 0, \qquad\qquad u_1, u_2, u_3 > 0, \quad (60\text{a})$$

$$= \frac{\exp\left(-|u_3|\, a\beta/u_2\right) - \exp\left(-|u_3|\, a\alpha/u_1\right)}{\alpha u_2 - \beta u_1},$$

$$u_1, u_2 > 0, \quad u_3 < 0, \quad (60\text{b})$$

$$= \frac{\exp\left(-\beta a u_3/|u_2|\right)}{\alpha\,|u_2| + \beta u_1}, \quad u_1, u_3 > 0, \quad u_2 < 0, \quad (60\text{c})$$

$$= \frac{\exp\left(-\alpha a u_3/|u_1|\right)}{\alpha u_2 + \beta\,|u_1|}, \quad u_2, u_3 > 0, \quad u_1 < 0. \quad (60\text{d})$$

The derivatives (59) can then be calculated by recurrence relations,[14,15] though care has to be taken to avoid instabilities in the computation of the derivatives of (60b), which can be expressed in terms of confluent hypergeometric functions.[3,14]

The situation is more involved if the functions χ_1 and χ_2 are not just products of powers and exponentials, especially if they contain a factor arising out of the expansion of a Slater orbital about a center other than O_1 or O_2, as in the Barnett–Coulson approach to the evaluation of 3- and 4-center integrals.[16,17] An increment in r_1 will correspond to various increments in r_2, depending on the ratio u_1/u_2. Unless, therefore, χ_1 and χ_2 can be evaluated rapidly for arbitrary values

of their arguments, numerical quadrature would be too time-consuming. It may be that in this case the expansion of an orbital about one center in a complete orthonormal set about another center would be more efficient; with the basis set of Löwdin–Shull functions[18] recently proposed by Smeyers[19] each integral would be a sum of terms (59), which could be evaluated on the basis of (60). It appears, however, from the applications quoted by Smeyers that the convergence is rather slow.

The foregoing discussion is of necessity rather sketchy since no actual calculations have been carried out along these lines; in consequence the writer has no idea how well the new approach would compare with other methods. In view of the importance and the difficulty of calculating 3- and 4- center integrals, no avenue should be left unexplored, and the ideas have therefore been presented as far as they have been thought out to date.

Two other directions for further research are mentioned in conclusion. One concerns the generalization of the expansion (3) to vectors in an arbitrary number of dimensions. The form of the function G, depending on the projection of the vector $\mathbf{r}_{AB}$ onto a fictitious polar axis, can be adapted to these cases without difficulty; the further factors in the integral (25) and in the definition (38) of G would be cosines in the plane and Gegenbauer functions in more than three dimensions (cf. Sec. 3.15 of Ref. 3).

An interesting problem in pure mathematics may be approached from a new direction on the basis of the present research. The solutions for R when V is of the form (6a) were presented in III in terms of Appell functions F_4 and here in terms of the integrals (25), where G is still a power. This suggests the possibility of expressing F_4 in terms of three-dimensional integrals; so far it has not been possible to express this function in terms of simple single or double Euler integrals.[3]

[14] K. Ruedenberg, K. O-Ohata, and D. G. Wilson, J. Math. Phys. 7, 539 (1966).

[15] R. A. Sack, C. C. J. Roothaan, and W. Kolos, J. Math. Phys. 8, 1093 (1967).

[16] M. P. Barnett and C. A. Coulson, Phil. Trans. Roy. Soc. (London) A243, 221 (1951).

[17] F. E. Harris and H. H. Michels, J. Chem. Phys. 43, S165 (1965).

[18] H. Shull and P. O. Löwdin, J. Chem. Phys. 23, 1362 (1953).

[19] Y. G. Smeyers, Theoret. Chim. Acta 4, 452 (1965).

the University of Wisconsin, and also Dr. E. Clementi and Professor K. Ruedenberg for their special encouragement of the ideas expressed in the last section.

APPENDIX A. SOME RELEVANT PROPERTIES OF THE LEGENDRE FUNCTIONS AND THEIR INTEGRALS

The associated Legendre functions P_l^m satisfy the well-known orthogonality condition

$$\int_{-1}^{1} P_l^m(u) P_n^{-m}(u)\, du = (-)^m \delta_{ln}(l + \tfrac{1}{2})^{-1}. \quad (A1)$$

From this relation and the completeness of the functions follows the expansion for the delta function

$$\delta(u - u') = (-)^m \sum_0^\infty (l + \tfrac{1}{2}) P_l^m(u) P_l^{-m}(u'), \quad (A2)$$

valid for arbitrary m; Eq. (34) represents the special case $m = 0$. Similarly with the definition for the integrals (28) and (A1) and (A2), the expression (33) is an identity. The integrals (28) are invariant under permutation of the columns, and the single integral over u can always be turned into a double integral by the insertion of the factor (A2). Carrying out both integrations and summing, we obtain the identity

$$I_\Omega \begin{pmatrix} l_1 & \cdots & l_q & l_{q+1} & \cdots & l_s \\ m_1 & \cdots & m_q & m_{q+1} & \cdots & m_s \end{pmatrix}$$

$$= (-)^m \sum_l (l + \tfrac{1}{2}) I_\Omega \begin{pmatrix} l_1 & \cdots & l_q & l \\ m_1 & \cdots & m_q & -m \end{pmatrix}$$

$$\times I_\Omega \begin{pmatrix} l & l_{q+1} & \cdots & l_s \\ m & m_{q+1} & \cdots & m_s \end{pmatrix};$$

$$m = m_1 + \cdots + m_q. \quad (A3)$$

For the last result we make use temporarily of the normalized harmonics (1b) and their integrals (29). For products of three harmonics these integrals are given in terms of the Wigner $3j$ symbols:

$$I'_Y \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

$$= \left[\frac{\Pi(2l_i + 1)}{4\pi} \right]^{\frac{1}{2}} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}. \quad (A4)$$

The Wigner symbols satisfy the orthogonality relation

[cf. (3.7.8) and (4.6.3) of Ref. 13]

$$\sum_{m_1} \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m - m_1 & -m \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l' \\ m_1 & m - m_1 & -m \end{pmatrix}$$

$$= \frac{\delta_{ll'}}{2l + 1} \delta(l_1\, l_2\, l), \quad (A5)$$

where $\delta(l_1\, l_2\, l_3)$ is unity, provided the triangle conditions (30c) are valid, and zero otherwise [though in this case the sum (30b) may be even or odd]. Reexpressing (1b) and (29) by their unnormalized counterparts (1a) and (28), we obtain from (A4) and (A5)

$$\sum_{m_1} I_\Omega \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m - m_1 & -m \end{pmatrix} I_\Omega \begin{pmatrix} l_1 & l_2 & l' \\ -m_1 & m_1 - m & m \end{pmatrix}$$

$$= \frac{\delta_{ll'}}{l + \tfrac{1}{2}} I_\Omega \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix}. \quad (A6)$$

It is intended in a future publication to establish an exhaustive theory of the integrals I_Ω by analytical methods only, so that no group-theoretical arguments are required to derive results such as (A6).

APPENDIX B. ERRATA TO PARTS I, II, AND III

Part I: J. Math. Phys. 5, 245 (1964).
Abstract: Delete "for $n = -1$ and $n = -2$."
Abstract: In the last line read "or of" for "of of".
(27a): The second line should begin

$$\times F[\tfrac{1}{2}l - \tfrac{1}{4}n, \tfrac{1}{2} - \tfrac{1}{4}n + \tfrac{1}{2}, \cdots$$

(36b): The first line should read

$$= -\frac{(l - 1)!}{(\tfrac{3}{2})_{l-1}} \frac{r_<^l (r_>^2 - r_<^2)^2}{r_>^{l+4}}$$

(49): Read $(2s + 2l + 1)!!$ for $(2s + 2^l + 1)!!$
Part II: J. Math. Phys. 5, 252 (1964).
Line following (47) should begin

"the value of L at constant N".

(57a): Read "r_2^2" for "dr_2^2" in last fraction.
Part III: J. Math. Phys. 5, 260 (1964).
(40b): This should be multiplied by a factor $-\tfrac{1}{8}$ to read

$$_2R'_0(\delta, \mathbf{I}) = \frac{(-)^{l_3}(2l_1 - 1)!!\,(2l_2 - 1)!!}{\pi(2l_3 - 1)!!\, 2^{\Lambda+1}\Lambda!} \cdots$$

p. 266, 2nd line of Sec. (a): read "charge" for "change".

Relation between Creeping Waves and Lateral Waves on a Curved Interface*

BENJAMIN RULF

Courant Institute of Mathematical Sciences, New York University, New York, New York

Diffraction effects at a gently curved interface between two media are investigated. Particular attention is paid to the behavior of the field on the diffracted rays which propagate along the interface into the shadows. It is found that far from the launching point of such a ray the field comprises of a series of modes which decay exponentially, due to the continuous leakage of energy away from the interface. At moderate distances, in the penumbra region, this series is poorly convergent. It can be converted into an integral, which can be evaluated asymptotically there, yielding a field with an algebraic decay. The field is like that diffracted along a plane interface, the so-called lateral wave, and reduces to it when the radius of curvature becomes infinite. The regions of transition from one representation to the other are determined, and uniform asymptotic expressions, valid across those regions, are given. All our results apply to a two-dimensional scalar problem, but results for three dimensions and for vector problems can be derived in a similar way.

I. INTRODUCTION

LET us consider radiation incident from a "slow" medium upon a curved interface with a "fast" medium. Such radiation, incident at angles less than the critical angle, is partly reflected and partly transmitted. A ray incident at the critical angle gives rise to a transmitted ray tangent to the interface. It forms part of the boundary between a region illuminated by transmitted rays and a shadow region (Fig. 1). We study the diffracted field in this shadow region and in the "incident" medium adjacent to it. For simplicity we consider the two-dimensional case, in which the interface is a curve.

We expect the critical ray to produce a surface diffracted ray.[1] This surface ray will shed tangent diffracted rays into the shadow, and shed rays refracted at the critical angle into the "incident" medium. If the interface is curved, this combination of rays is called a "creeping wave," and the corresponding field decays exponentially with distance along the interface. However, if the interface is a plane, there is no shadow and the diffracted field is present only on the "incident" side, where it is called a "lateral wave" or "head wave." The corresponding field decays algebraically with distance like $s^{-\frac{3}{2}}$. We wish to determine the relationship between the two kinds of waves, the creeping and lateral waves, associated with the interface.

[1] J. B. Keller, in *Calculus of Variations and its Applications*, Proc. Symp. Appl. Math. (McGraw-Hill Book Company, Inc., New York, and American Mathematical Society, Providence, Rhode Island, 1958), Vol. 8.

We show that both kinds of waves are present in the case of a curved interface. Far from the point of diffraction, the field is given by a series of creeping waves, one of which is dominant. The series is slowly convergent near the point of diffraction, and can be reexpressed as a lateral wave. We find the distance at which the behavior of the field changes from lateral to creeping wave. In addition we obtain a uniform expression for the field, which reduces to the lateral wave at short distances and to the series of creeping waves at large distances. This expression provides a description of the field in the transition region.

The "fast" and "slow" media are characterized by wavenumbers k_o and k_i, respectively, with

$$k_i > k_o.$$

A time harmonic source Q, with time dependence $e^{-i\omega t}$, is located in the slow medium. The interface is a smooth curve, concave when seen from the source.

At high frequencies (large k_o and k_i) the field can be expressed in terms of rays as follows (Fig. 1):

In region I the leading term of the field arises from a refracted ray QAP_1. The law of refraction (Snell's law) is

$$k_o \sin \beta = k_i \sin \alpha. \tag{1}$$

In region II the leading term arises from a direct ray QP_2 plus a reflected ray QAP_2. In region III there is, besides the reflected ray QCP_3, a diffracted ray $QBEP_3$, whose field may be of the same order of magnitude as (or even larger than) the field of the reflected ray (especially when P_3 lies near the interface and when k_i has a small imaginary part, corresponding to slight loss). Region IV is a shadow. The field at P_4 arises from a diffracted ray $QBDP_4$ and an "evanescent
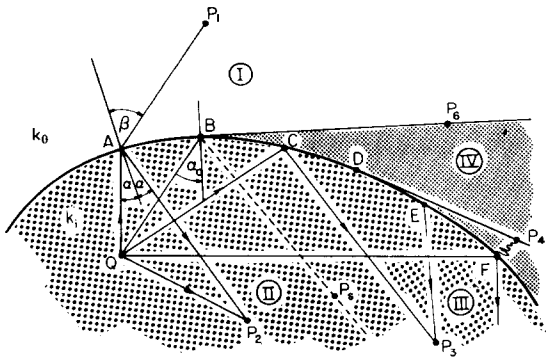
FIG. 1. Ray paths near an interface between two media.

ray" QFP_4, which is usually of a much smaller order of magnitude (unless P_4 lies near the interface and near B). At first glance there are two transition regions: between regions II and III (point P_5), and between regions I and IV (point P_6), in which the field cannot be expressed in simple ray terms, and a uniform asymptotic expansion is necessary.

The geometric theory of diffraction[1] predicts that the ray QB which hits the interface at the critical angle

$$\sin \alpha_o = k_o/k_i \qquad (2)$$

launches diffracted rays which propagate along geodesic lines on the interface. At large distances from B, the decay of the field associated with these rays is exponential due to the continuous shedding of energy tangentially from the interface into the shadow. When the curvature vanishes, the diffracted field does not decay exponentially, since no tangential shedding of energy takes place. There is, however, a continuous shedding of energy back into the slower medium by refraction, giving rise to an algebraic decay of the field. We see that this is also the behavior near B. Thus there is another transition zone between the algebraically decaying field near B and the exponentially decaying field far from B.
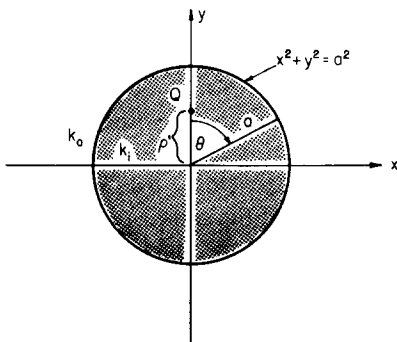

FIG. 2. Line source inside a transparent circular cylinder.

II. ANALYSIS OF A SINGLE MODE

To explain the mechanism of energy transport from source to observation point, particularly in the shadow region, we investigate first the propagation of a single mode, since the total field is a sum of such modes. For the exact analysis of a mode, we assume the interface to be a circle (circular cylinder), since this is the only geometry in which a two-medium problem can be solved by separation of variables (see Fig. 2).

We wish to determine a function u which satisfies the following conditions:

$$(\nabla^2 + k_i^2)u = 0 \quad \text{at} \quad \rho < a, \qquad (3a)$$

$$(\nabla^2 + k_o^2)u = 0 \quad \text{at} \quad \rho > a, \qquad (3b)$$

$$\lim_{\epsilon \to 0} [u(a + \epsilon) - u(a - \epsilon)] = 0, \qquad (3c)$$

$$\lim_{\epsilon \to 0} \left[\frac{\partial u}{\partial \rho}\bigg|_{a+\epsilon} - \frac{\partial u}{\partial \rho}\bigg|_{a-\epsilon} \right] = 0, \qquad (3d)$$

$$\lim_{\rho \to \infty} [\rho^{\frac{1}{2}}(\partial u/\partial \rho - ik_o u)] = 0. \qquad (3e)$$

We choose a mode of the form

$$u = \begin{cases} H_\mu^{(1)}(k_o\rho)e^{i\mu\theta} & \text{for} \quad \rho > a, \\ AJ_\mu(k_i\rho)e^{i\mu\theta} & \text{for} \quad \rho < a. \end{cases} \qquad (4)$$

Because of (3c)

$$A = J_\mu(k_i a)/H_\mu^{(1)}(k_o a), \qquad (5)$$

and because of (3d)

$$\frac{d}{d\rho} H_\mu^{(1)}(k_o\rho)\bigg|_{\rho=a} = A \frac{d}{d\rho} J_\mu(k_i\rho)\bigg|_{\rho=a},$$

which, with the help of (5) yields an equation for the determination of μ:

$$M(\mu) = k_o \frac{H_\mu^{(1)\prime}(k_o a)}{H_\mu^{(1)}(k_o a)} - k_i \frac{J_\mu'(k_i a)}{J_\mu(k_i a)} = 0, \qquad (6)$$

where the prime denotes differentiation with respect to the argument.

Equation (6) has been thoroughly investigated.[2,3] Its asymptotic solution for $k_o a \gg 1$ is

$$\mu_p \sim k_o a[1 + \tau_p e^{i\pi/3} 2^{-\frac{1}{3}}(k_o a)^{-\frac{2}{3}} + O(k_o a)^{-\frac{4}{3}}]$$
$$= k_o a + ia\alpha_p, \qquad (7)$$

where τ_p are the solutions of

$$\frac{A_i'(\tau_p)}{A_i(\tau_p)} = \left(\frac{k_o a}{2}\right)^{\frac{1}{3}} e^{i\pi/6} \left[\left(\frac{k_i}{k_o}\right)^2 - \left(\frac{\mu_p}{k_o a}\right)^2 \right]^{\frac{1}{2}}. \qquad (7a)$$

[2] W. Streifer and R. Kodis, Quart. Appl. Math. **21** (4), 285 (1963); **23** (1), 27 (1965).
[3] Y. M. Chen, Ph.D. thesis, N.Y.U. (1963); J. Math. Phys. **5**, 820 (1964).

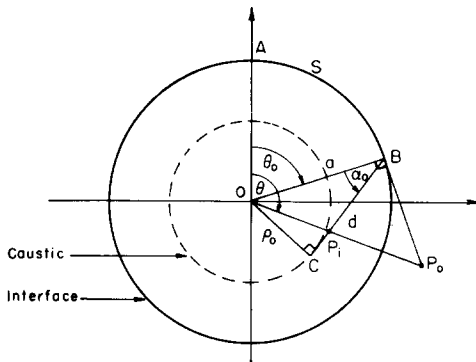FIG. 3. Ray paths for a single mode of a transparent circular cylinder.

Thus, the mode given in (4) may be written as

$$u = \begin{cases} H^{(1)}_{\mu_p}(k_o\rho)e^{i\mu_p\theta} & \text{for } \rho > a, \\ \dfrac{H^{(1)}_{\mu_p}(k_o a)}{J_{\mu_p}(k_i a)} J_{\mu_p}(k_i\rho)e^{i\mu_p\theta} & \text{for } \rho < a. \end{cases} \quad (8)$$

The geometrical interpretation of (8) follows from Fig. 3:

$$\rho = \begin{cases} OP_i \text{ for observation point inside the cylinder,} \\ OP_o \text{ for observation point outside the cylinder,} \end{cases}$$

$$d = BP_i,$$
$$\rho_o = OC = k_o/k_i \quad [\text{from Eq. (2)}],$$
$$s = \text{arc } AB,$$
$$b = BP_o.$$

If P_i is not too near to the caustic (point C), then

$$k_o a < k_i\rho < k_i a,$$

since

$$k_o a = k_i a \sin\alpha_o = k_i\rho_o;$$

thus the Debye approximation for $J_{\mu_p}(k_i\rho)$ and $J_{\mu_p}(k_i a)$ is valid. Also,

$$J_\mu(x) = \tfrac{1}{2}[H^{(1)}_\mu(x) + H^{(2)}_\mu(x)] \sim \tfrac{1}{2}H^{(2)}_\mu(x), \quad (9)$$

with x standing for $k_i a$ or $k_i\rho$, since $H^{(2)}_\mu(x)$ becomes exponentially large, whereas $H^{(1)}_\mu(x)$ becomes exponentially small when k_i has a small positive imaginary part. Thus, for points inside the cylinder [Eqs. (7) and (8)]:

$$u \sim H^{(1)}_{\mu_p}(k_o a)[(k_i^2 a^2 - k_o^2 a^2)/(k_i^2\rho^2 - k_o^2 a^2)]^{\frac{1}{4}}$$

$$\times \exp\left\{ i\left[(k_i^2 a^2 - k_o^2 a^2)^{\frac{1}{2}} - (k_i^2\rho^2 - k_o^2 a^2)^{\frac{1}{2}} \right. \right.$$

$$\left. \left. + (k_o a + ia\alpha_p)\left(\theta + \cos^{-1}\frac{k_o}{k_i} + \cos^{-1}\frac{k_o a}{k_i\rho}\right)\right]\right\}. \quad (10)$$

It is easy to show that

$$k_i d = k_i[a\cos\alpha_o - (\rho^2 - \rho_o^2)^{\frac{1}{2}}],$$
$$= (k_i^2 a^2 - k_o^2 a^2)^{\frac{1}{2}} - (k_i^2\rho^2 - k_o^2 a^2)^{\frac{1}{2}}, \quad (11a)$$

$$\theta_o = \theta + \cos^{-1}(k_o/k_i) - \cos^{-1}(k_o a/k_i\rho), \quad (11b)$$

$$\left(\frac{k_i^2 a^2 - k_o^2 a^2}{k_i^2\rho^2 - k_o^2 a^2}\right)^{\frac{1}{4}} = \left(\frac{a\cos\alpha_o}{a\cos\alpha_o - d}\right)^{\frac{1}{2}}. \quad (11c)$$

For observation points outside the cylinder we get, in a similar way,

$$u \sim (2/\pi i)^{\frac{1}{2}}(k_o^2\rho^2 - k_o^2 a^2)^{-\frac{1}{4}}\exp\left(i\{(k_o^2\rho^2 - k_o^2 a^2)^{\frac{1}{2}} \right.$$
$$\left. + (k_o a + ia\alpha_p)[\theta - \cos^{-1}(k_o a/k_i\rho)]\}\right), \quad (12)$$

with

$$(k_o^2\rho^2 - k_o^2 a^2)^{\frac{1}{4}} = (k_o b)^{\frac{1}{2}}, \quad (13a)$$

$$\theta_o = \theta - \cos^{-1}(k_o a/k_i\rho). \quad (13b)$$

Thus, the asymptotic form of a single mode is

$$u \sim \begin{cases} \left(\dfrac{2}{\pi i}\right)^{\frac{1}{2}} \dfrac{\exp\{ik_o(s + b) - \alpha_p s\}}{(k_o b)^{\frac{1}{2}}}, \\ \\ H^{(1)}_{\mu_p}(k_o a)\left(\dfrac{a\cos\alpha_o}{a\cos\alpha_o - d}\right)^{\frac{1}{2}} \\ \qquad\qquad \times \exp[i(k_i d + k_o s) - \alpha_p s]. \end{cases} \quad (14)$$

The interpretation of (14) in ray terms is obvious[1,4]: A ray which originates at point A (Fig. 3) reaches the observation point P_o in the exterior region by creeping along the interface with an exponential decay α_p to point B, and shedding tangentially from there to P_o. A point P_i in the interior region is reached by creeping along the interface to B as before, and refracting into the interior according to Snell's law [Eqs. (1) and (2)]. Equation (14) also contains the correct geometric divergence or convergence coefficients required by the principle of conservation of energy. Due to the principle of reciprocity, the same formulas would apply for the case where the roles of source and observation points are interchanged. (The diffraction and splitting coefficients[3] associated with this process cannot be determined from the study of a single mode.)

III. ANALYSIS OF A CANONICAL PROBLEM.

The results of the last section are now used in the analysis of a canonical problem, namely, the field of a line source near the interface between a circular cylinder of a slow medium embedded in a faster medium. Consider a unit strength line source located at the point Q in the interior region (Figs. 2, 4). The function u to be determined satisfies Eqs. (3b, c, d, e)

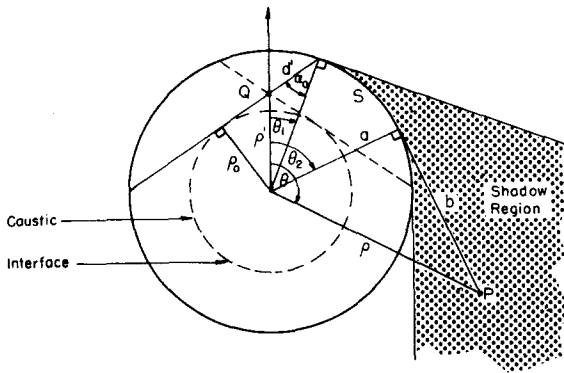[4] J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).

FIG. 4. Path of a diffracted ray from source into shadow region.

as before, and instead of (3a) we require

$$(\nabla^2 + k_i^2)u = -[\delta(\theta)\delta(\rho - \rho')]/\rho \quad \text{at} \quad \rho < a. \quad (15)$$

The techniques of solving this problem are well known.[5,6] The solution in the exterior region ($\rho > a$) is

$$u(\rho, \theta) = \frac{-1}{2\pi a} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{H_\mu^{(1)}(k_o\rho)J_\mu(k_i\rho')e^{i\mu(\theta+2n\pi)}}{J_\mu(k_i a)H_\mu^{(1)}(k_o a)M(\mu)} d\mu,$$

$$= \sum_{n=0}^{\infty} u_n, \quad (16)$$

with $M(\mu)$ given by Eq. (6).

The integrals in (16) may be converted to a sum of residues, yielding

$$u_0(\rho, \theta) = -\frac{i}{a} \sum_{p=1}^{\infty} \frac{H_{\mu_p}^{(1)}(k_o\rho)J_{\mu_p}(k_i\rho')e^{i\mu_p\theta}}{J_{\mu_p}(k_i a)H_{\mu_p}^{(1)}(k_o a)(\partial/\partial\mu)M(\mu)|_{\mu=\mu_p}} \quad (17)$$

and similar expressions for $n = 1, 2, \cdots$. The numbers μ_p are the solutions of Eq. (6), which are given by (7). Thus, using the results of the former section [Eqs. (8) and (14)], we have asymptotically

$$u_0(\rho, \theta) \sim -\frac{1}{a}\left(\frac{2i}{\pi}\right)^{\frac{1}{2}}\left(\frac{a\cos\alpha_o}{a\cos\alpha_o - d'}\right)^{\frac{1}{2}}$$

$$\times \frac{\exp\{i[k_i d' + k_o(s + b)]\}}{(k_o b)^{\frac{1}{2}}}$$

$$\times \sum_{p=1}^{\infty} \frac{e^{-\alpha_p s}}{H_{\mu_p}^{(1)}(k_o a)(\partial/\partial\mu)M(\mu)|_{\mu=\mu_p}} \quad (18)$$

(see Fig. 4). It is seen that for observation points on the interface the expression (17) simplifies to

$$u_0(a, \theta) = -\frac{1}{a}\left(\frac{a\cos\alpha_o}{a\cos\alpha_o - d'}\right)^{\frac{1}{2}}e^{i(k_i d' + k_o s)}$$

$$\times \sum_{p=1}^{\infty} \frac{e^{-\alpha_p s}}{(\partial/\partial\mu)M(\mu)|_{\mu=\mu_p}}. \quad (19)$$

This shows that the diffraction coefficient associated with the tangential shedding of the rays is $D_p = (2i/\pi)^{\frac{1}{2}}[H_{\mu_p}^{(1)}(k_o a)]^{-1}$. It follows from Eq. (7) that for $k_o a \gg 1$, $\mu_p \approx k_o a$. Thus, if (18) converges rapidly (i.e., when s is large) and only the first few terms contribute effectively, we may take the diffraction coefficient to be approximately

$$(\pi/2i)^{\frac{1}{2}}D_p \equiv [H_{\mu_p}^{(1)}(k_o a)]^{-1}$$

$$\approx [H_{k_o a}^{(1)}(k_o a)]^{-1} \sim \left(\frac{k_o a}{2}\right)^{\frac{1}{3}}\left(\frac{3^{\frac{1}{3}}\pi}{\Gamma(\frac{1}{3})}\right) \quad (20)$$

and take it out in front of the summation sign. It is also worth mentioning that the geometric factor $[(a\cos\alpha_o)/(a\cos\alpha_o - d')]^{\frac{1}{2}}$, which accounts for the convergence of the ray tubes towards the caustic, causes the field at the interface to increase when the source point moves farther away from the interface. When the source point is near or on the caustic, the above formulas have to be modified, since the Debye approximations become invalid. For $\rho' < \rho_0$ there will be no shadow region and no critically refracted rays exist. From Eq. (7) it is seen that

$$\alpha_p s \sim \tau_p(\tfrac{1}{2}e^{i\pi/2})^{\frac{1}{3}}(k_o s)(k_o a)^{-\frac{2}{3}}. \quad (21)$$

The series in Eqs. (17) or (18) will converge rapidly only when

$$(k_o s)(k_o a)^{-\frac{2}{3}} > O(1),$$

or equivalently,

$$(k_o s) > O(k_o a)^{\frac{2}{3}}, \quad (22)$$

which means in the deep-shadow region. For $k_o a \gg 1$ (gently curved interface), the penumbra region, in which (18) is useless, extends over a very considerable number of wavelengths. Since the field over a plane interface ($k_o a \to \infty$) is well known,[7] one is led to the assumption that for large $k_o a$ one may find an alternative field representation analogous to the plane interface form.

IV. COMPARISON WITH THE PLANE INTERFACE PROBLEM

Letting $a \to \infty$ in Fig. 2, while leaving the distance $l = a - \rho'$ unchanged, transforms the configuration to that shown in Fig. 5. The field in $y \geq 0$ in this case is given by[7]

$$u(x, y)$$

$$= \frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{\exp\{i[(k_i^2 - t^2)^{\frac{1}{2}}l + tx + (k_o^2 - t^2)^{\frac{1}{2}}y]\}}{(k_i^2 - t^2)^{\frac{1}{2}} + (k_o^2 - t^2)^{\frac{1}{2}}} dt. \quad (23)$$

It can be shown (see Appendix A) that if we let

[5] T. T. Wu, Phys. Rev. **104**, 1201 (1956).

[6] W. Franz, *Theorie der Beugung Elektromagnetischer Wellen* (Springer-Verlag, Berlin, 1957).

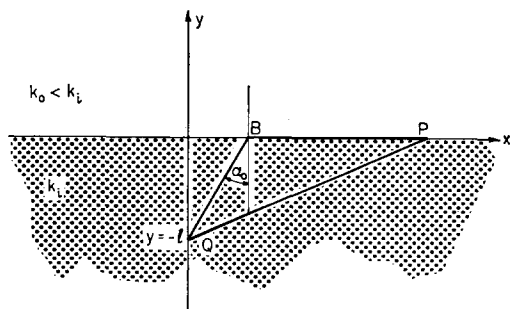[7] L. M. Brekhovskikh, *Waves in Layered Media* (Academic Press Inc., New York, 1960).

FIG. 5. Paths of direct and diffracted rays from source to observation point on a plane interface.

$a \to \infty$, the integral in Eq. (16) which corresponds to $n = 0$ reduces exactly to (23). The change of variables used is

$$\mu = at,$$
$$\eta = \rho - a, \quad \eta' = a - \rho', \qquad (24)$$
$$\xi = a\theta.$$

All the other terms ($n \neq 0$) in (16) correspond to rays which encircle the cylinder n times or undergo n internal refractions before reaching the observation point. For large $k_o a$ these terms make a negligible contribution to the field. In order to investigate the behavior of the field along the interface, we let $y = 0$ and $\rho = a$, and may then write

$$\lim_{a \to \infty} - \frac{1}{2\pi a} \int_{-\infty}^{\infty} \frac{J_\mu(k_i \rho') e^{i\mu\theta}}{J_\mu(k_i a) M(\mu)} d\mu$$
$$= \frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{\exp\{i[(k_i^2 - t^2)^{\frac{1}{2}} l + tx]\}}{(k_i^2 - t^2)^{\frac{1}{2}} + (k_o^2 - t^2)^{\frac{1}{2}}} dt. \quad (25)$$

This integral can be evaluated asymptotically.[1] For observation point P to the right of B (Fig. 5), the two leading terms in the asymptotic evaluation correspond to the contribution from the direct ray QP and the critically refracted lateral ray QBP. When $QP \gg QB$ and the lower medium is slightly lossy, the lateral wave will be the principal part of the field. It is obtained from the integral around the branch cut, arising from the branch point at $t = k_o$, and is given by

$$u_{\text{lat}} \sim \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{k_o^2}{k_i^2 - k_o^2} \frac{\exp\{i[k_i QB + k_o BP - \pi/4]\}}{(k_o BP)^{\frac{3}{2}}}. \quad (26)$$

We may derive an analogous result for a slightly curved interface, starting from Eq. (19). By a slight modification, the field along the interface becomes

$$u_o(a, \theta) \sim - \frac{1}{a} \left(\frac{2i}{\pi}\right)^{\frac{1}{2}} \left(\frac{a \cos \alpha_o}{a \cos \alpha_o - d'}\right)^{\frac{1}{2}} e^{ik_i d'}$$
$$\times \sum_{p=1}^{\infty} \frac{\exp(i\mu s/a)}{(\partial/\partial\mu) M(\mu)\big|_{\mu=\mu_p}}. \quad (27)$$

Since the series in (27) is poorly convergent for small $\mu_p s/a$, we convert it back to an integral:

$$- \frac{1}{a} \sum_{p=1}^{\infty} \frac{\exp(i\mu s/a)}{(\partial/\partial\mu) M(\mu)\big|_{\mu=\mu_p}} = \frac{1}{2\pi} \frac{i}{a} \int_{-\infty}^{\infty} \frac{\exp(i\mu s/a)}{M(\mu)} d\mu,$$
$$= \frac{I(s)}{2\pi}. \quad (28)$$

The path of integration has to encircle all the poles of $M(\mu)$ which lie in the upper half of the μ plane. The arc at infinity does not contribute as long as $s > 0$, which indicates that the result is valid only in the shadow region. It is shown (in Appendix A) that for large $k_o a$

$$I(s) = \frac{i}{a} \int_{-\infty}^{\infty} \frac{\exp(i\mu s/a)}{M(\mu)} d\mu,$$
$$\sim \int_C \frac{e^{its}}{(k_o^2 - t^2)^{\frac{1}{2}} + (k_i^2 - t^2)^{\frac{1}{2}}} dt,$$
$$= \int_{C'} \frac{e^{its}}{(k_o^2 - t^2)^{\frac{1}{2}} + (k_i^2 - t^2)^{\frac{1}{2}}} dt, \quad (29)$$

with the paths C and C' shown in Fig. 6. [The choice of the branch cuts is such as to assure that

$$\text{Im}(k_0^2 - t^2)^{\frac{1}{2}} > 0 \quad \text{and} \quad \text{Im}(k_i^2 - t^2)^{\frac{1}{2}} > 0$$

on the entire upper Riemann sheet. This requirement is necessary to satisfy the radiation condition.] When $k_o s$ is not small (i.e., when s is larger than one wavelength or thereabouts), the integral in (29) may be evaluated asymptotically as follows. Let

$$t = k_o + i k_o z^2. \quad (30)$$

The integrand may be expanded as a power series in z, yielding

$$\int_{C'} \frac{e^{its} dt}{(k_o^2 - t^2)^{\frac{1}{2}} + (k_i^2 - t^2)^{\frac{1}{2}}}$$
$$= \frac{2ik_o e^{ik_o s}}{(k_i^2 - k_o^2)^{\frac{1}{2}}} \int_{C'} \left[z - \frac{(-2i)^{\frac{1}{2}} k_o}{(k_i^2 - k_o^2)^{\frac{1}{2}}} z^2 + O(z^3)\right]$$
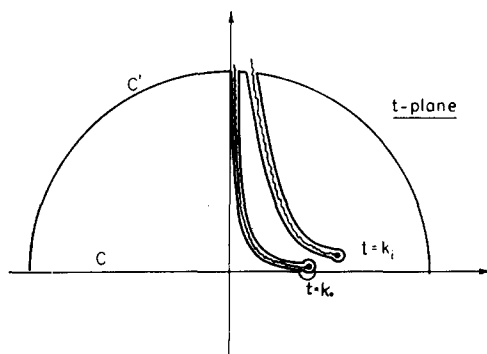$$\times \exp(-k_o s z^2) dz. \quad (31)$$

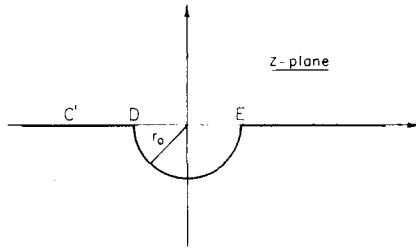

FIG. 6. The path of integration in the complex t plane.

Because of the $\exp(-k_osz^2)$ factor in the integrand of (31), the principal contribution to the integral will arise from that portion of the path where z is small. The path C' as mapped in the z plane is shown in Fig. 7. In the neighborhood of $z = 0$, let $z = r_o e^{i\varphi}$ and

$$\exp(-k_osz^2) = \exp(-k_osr_o^2\cos 2\varphi)$$

$$\times \exp(-ik_osr_o^2\sin 2\varphi) \quad -\pi \le \varphi \le 0. \quad (32)$$

We want the arc DE to contribute significantly to the integral, since we assumed that the principal contribution comes from the neighborhood of $z = 0$. Thus we have to keep the quantity $k_osr_0^2$ small. From (24) and (30) we have

$$r_o = |z| = |(\mu - k_oa)/k_oa|^{\frac{1}{2}}. \quad (33)$$

Thus the condition

$$k_osr_o^2 = O(1)$$

implies

$$|(\mu - k_oa)/k_oa| = O(1/k_os). \quad (34)$$

On the other hand, a necessary condition for (25) and (29) to be valid is (see Appendix A)

$$|\mu - k_oa| > O(k_oa)^{\frac{1}{3}} \quad (35)$$

The relations (34) and (35) can be satisfied simultaneously only if

$$k_os < O(k_oa)^{\frac{2}{3}}. \quad (36)$$

Comparison of (22) and (36) shows exactly where the transition from the creeping-wave representation to the lateral-wave representation occurs.

The integral in (31) can be calculated (asymptotically) in closed form by using the relations

$$\int_{-\infty}^{\infty} z^m \exp(-az^2)\, dz$$

$$= \begin{cases} 0, & \text{for } m = 1, 3, 5, \cdots, \\ \dfrac{1 \cdot 3 \cdot 5 \cdots (m-1)(2\pi)^{\frac{1}{2}}}{(2a)^{\frac{1}{2}(m+1)}}, & \text{for } m = 2, 4, 6, \cdots. \end{cases}$$

Thus we obtain

$$\frac{i}{a}\int_{-\infty}^{\infty}\frac{\exp(i\mu s/a)\, d\mu}{M(\mu)}$$

$$\sim \int_{C'}\frac{e^{its}\, dt}{(k_o^2 - t^2)^{\frac{1}{2}} + (k_i^2 - t^2)^{\frac{1}{2}}}$$

$$\sim \frac{(-2i)^{\frac{3}{2}}(\pi)^{\frac{1}{2}}}{2}\frac{k_o^2}{k_i^2 - k_o^2}\frac{e^{ik_os}}{(k_os)^{\frac{3}{2}}}[1 + O(k_os)^{-1}], \quad (37)$$

and the expression (27) for the field along the interface becomes

$$u_o(a, \theta) \sim \frac{1}{(2\pi)^{\frac{1}{2}}}\frac{k_o^2}{k_i^2 - k_o^2}\left(\frac{a\cos\alpha_o}{a\cos\alpha_o - d'}\right)^{\frac{1}{2}}$$

$$\times \frac{\exp[i(k_id' + k_os)]}{(k_os)^{\frac{3}{2}}}[1 + O(k_os)^{-1}]. \quad (38)$$

Comparison of (26) and (38) shows the complete correspondence when $a \to \infty$. For the transition region

$$k_os \sim (k_oa)^{\frac{2}{3}},$$

we may use the uniform asymptotic expressions for the cylinder functions in $M(\mu)$:

$$\int_{-\infty}^{\infty}\frac{\exp(i\mu s/a)}{M(\mu)}\, d\mu$$

$$\sim -\frac{1}{k_oy}\int_{-\infty}^{\infty}\exp(i\mu s/a)\, d\mu\Bigg/$$

$$\left\{\frac{\text{Ai}'[y(\mu - k_oa)]}{\text{Ai}[y(\mu - k_oa)]} - \frac{i}{y}\left[\left(\frac{k_i}{k_o}\right)^2 - \left(\frac{\mu}{k_oa}\right)^2\right]^{\frac{1}{2}}\right\} \quad (39)$$

with

$$y = \exp(-i\pi/3)(2/k_oa)^{\frac{1}{3}}. \quad (39a)$$

The integral in (39) may be substituted into (28) and (27). It will yield (19) in the region

$$k_os > O(k_oa)^{\frac{2}{3}}$$

and (38) in the region

$$k_os < O(k_oa)^{\frac{2}{3}}.$$

Equation (37) is not valid for $s \le 0$ because the integral cannot be evaluated by a contour integration as shown in Fig. 6. As a matter of fact, $s < 0$ corresponds to the illuminated region (see Fig. 4). The asymptotic evaluation of the integral in Eq. (37) also required s to be large. Thus, for $s \ll 1$ and the transition from $s > 0$ to $s < 0$, we look for a uniform asymptotic representation[8] which connects smoothly the different representations in the illuminated and the shadow regions.

[8] N. Bleistein, "Uniform Asymptotic Expansions of Integrals with Stationary Point near Algebraic Singularity," Commun. Pure Appl. Math. 14, 353 (1966).

The method of Ref. 8 is illustrated by the calculation of $u_o(a, \theta)$ near $\theta = \theta_1$ (Fig. 4), which corresponds to $s = 0$.

$$u_o(a, \theta) = - \frac{1}{2\pi a} \int_{-\infty}^{\infty} \frac{J_\mu(k_i \rho')}{J_\mu(k_i a)} \frac{\exp (i\mu\theta)}{M(\mu)} d\mu,$$

$$\sim \frac{i}{2\pi} \int_{-\infty}^{\infty} \left(\frac{k_i^2 a^2 - \mu^2}{k_i^2 \rho'^2 - \mu^2} \right)^{\frac{1}{4}} \frac{\exp \left(i\{(k_i^2 a^2 - \mu^2)^{\frac{1}{2}} - (k_i^2 \rho'^2 - \mu^2)^{\frac{1}{2}} + \mu[\theta - \cos^{-1}(\mu/k_i a) + \cos^{-1}(\mu/k_i \rho')]\} \right)}{(k_o^2 a^2 - \mu^2)^{\frac{1}{2}} + (k_i^2 a^2 - \mu^2)^{\frac{1}{2}}} d\mu,$$

$$= \frac{i}{2\pi} \int_{-\infty}^{\infty} \left(\frac{k_i^2/k_o^2 - \xi^2}{k_i^2 \rho'^2/k_o^2 a^2 - \xi^2} \right)^{\frac{1}{4}} \frac{\exp \{ik_o a \varphi(\xi)\}}{(1 - \xi^2)^{\frac{1}{2}} + (k_i^2/k_o^2 - \xi^2)^{\frac{1}{2}}} d\xi, \tag{40}$$

with

$$\varphi(\xi) = (k_i^2/k_o^2 - \xi^2)^{\frac{1}{2}} - \left(\frac{k_i^2 \rho'^2}{k_o^2 a^2} - \xi^2 \right)^{\frac{1}{2}} + \xi \left(\theta + \cos^{-1} \frac{k_o a \xi}{k_i \rho'} - \cos^{-1} \frac{k_o \xi}{k_i} \right). \tag{41}$$

Multiplying (40) by

$$[(1 - \xi^2)^{\frac{1}{2}} - (k_i^2/k_o^2 - \xi^2)^{\frac{1}{2}}]/[(1 - \xi^2)^{\frac{1}{2}} - (k_i^2/k_o^2 - \xi^2)^{\frac{1}{2}}]$$

and making the change of variable $\xi = 1 - t$, we obtain

$$u_o(a, \theta) \sim - \frac{i}{2\pi} \frac{k_o^2}{k_i^2 - k_o^2}$$

$$\times \int_{-\infty}^{\infty} \hat{g}(t) \left[(2 - t)^{\frac{1}{2}} t^{\frac{1}{2}} - \left[\frac{k_i^2}{k_o^2} - (1 - t)^2 \right]^{\frac{1}{2}} \right]$$

$$\times \exp \{ik_o a f(t, \alpha)\} \, dt = u_1 + u_2 \tag{42}$$

with

$$\hat{g}(t) = \left[\frac{k_i^2/k_o^2 - (1 - t)^2}{k_i^2 \rho'^2/k_o^2 a^2 - (1 - t)^2} \right]^{\frac{1}{4}} = \frac{g(t)}{(2 - t)^{\frac{1}{2}}} \tag{42a}$$

and

$$f(t, \alpha) = [k_i^2/k_o^2 - (1 - t)^2]^{\frac{1}{2}} - \left[\frac{k_i^2 \rho'^2}{k_o^2 a^2} - (1 - t)^2 \right]^{\frac{1}{2}}$$

$$+ (1 - t) \left[\left(\theta_1 + \frac{s}{a} \right) + \cos^{-1} \frac{k_o a(1 - t)}{k_i \rho'} \right]$$

$$- \cos^{-1} \frac{k_o(1 - t)}{k_i} \right]. \tag{42b}$$

Both u_1 and u_2 have a stationary point at

$$\frac{\partial f}{\partial t} \bigg|_{t=\alpha} = 0 = \cos^{-1} \frac{k_o(1 - \alpha)}{k_i}$$

$$- \cos^{-1} \frac{k_o a(1 - \alpha)}{k_i \rho'} - \left(\theta_1 + \frac{s}{a} \right) \tag{42c}$$

for $s/a \ll 1$; Eq. (42c) can be solved approximately for α, yielding

$$\alpha \approx \left(\frac{s}{a} \right) \frac{[(k_i^2 a^2 - k_o^2 a^2)(k_i^2 \rho'^2 - k_o^2 a^2)]^{\frac{1}{2}}}{k_o a [(k_i^2 a^2 - k_o^2 a^2)^{\frac{1}{2}} - (k_i^2 \rho'^2 - k_o^2 a^2)^{\frac{1}{2}}]} = O\left(\frac{s}{a} \right). \tag{43}$$

We can also show that

$$B = -e^{i\pi/4}(2 |f(\alpha, \alpha) - f(0, \alpha)|)^{\frac{1}{2}} \approx -e^{i\pi/4}(2s/a)^{\frac{1}{2}}, \tag{44}$$

$$k_o a f(0, \alpha) = k_i d' + k_o s. \tag{45}$$

It is seen that u_1 in (42) has a stationary point near a branch point. u_2 in (42) may be evaluated by the saddle-point method, since it has no branch point in the vicinity of the saddle point. It yields exactly half of the direct ray field (QA or QC in Fig. 1) as calculated by geometrical optics. u_1 has to be calculated according to the method given in Ref. 8. The result of that calculation is

$$u_1(a, \theta) \sim \exp [i(k_i d' + k_o s - \pi/4)]$$

$$\times \left[\gamma_o \frac{W_{\frac{1}{2}}[(k_o a)^{\frac{1}{2}} B]}{(k_o a)^{\frac{1}{4}}} - \gamma_1 \frac{W'_{\frac{1}{2}}[(k_o a)^{\frac{1}{2}} B]}{(k_o a)^{\frac{3}{4}}} \right], \tag{46}$$

where

$$W_r(s) = \int t^r \exp [-(t^2/2 + st)] \, dt,$$

$$= (2\pi)^{\frac{1}{2}} e^{i\pi r/2} e^{s^2/4} D_r(is), \tag{46a}$$

$D_r(is)$ being the Weber function[9] of order r.

$$\gamma_o = g(0) \left(\frac{\partial^2 f}{\partial t^2} \bigg|_0 \right)^{-\frac{3}{4}} e^{-i\pi/8}, \tag{46b}$$

$$\gamma_1 = \frac{e^{-i3\pi/8}}{(2s/a)^{\frac{1}{2}}} \left\{ g(0) \left(\frac{\partial^2 f}{\partial t^2} \bigg|_0 \right)^{-\frac{3}{4}} \right.$$

$$\left. - g(\alpha) \left[\frac{\alpha}{-(2s/a)^{\frac{1}{2}}} \right]^{\frac{1}{2}} \left(\frac{\partial^2 f}{\partial t^2} \bigg|_0 \right)^{-\frac{1}{2}} \right\}. \tag{46c}$$

α is defined by (43), $g(t)$ by (42a), and $f(t)$ by (42b). Thus we have [in Eqs. (19), (38), (39), (42), and (46)] asymptotic expressions for the field along the interface anywhere in the shadow region and also in the transition from the shadow to the illuminated region.

[9] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1954).

V. ASYMPTOTIC EXPRESSIONS FOR THE FIELD ABOVE AN INTERFACE

Throughout the following analysis we assume $k_i d' \gg 1$. It is now possible to write down asymptotic expressions for the field in the various regions shown in Fig. 8. Regions A, B, and C are characterized by the fact that

$$k_o(\rho - a) > O(k_o a)^{\frac{1}{3}}.$$

The field in these regions is

$$u_o(\rho, \theta) \sim \frac{1}{2\pi} (2i/\pi)^{\frac{1}{2}} \left(\frac{a \cos \alpha_o}{a \cos \alpha_o - d'}\right)^{\frac{1}{2}}$$

$$\times \frac{\exp [i(k_i d' + k_o b)]}{(k_o b)^{\frac{1}{2}}} D_p I(s), \quad (47)$$

where

$$I(s) = \frac{i}{a} \int_{-\infty}^{\infty} \frac{\exp (i\mu s/a)}{M(\mu)} d\mu. \quad (47a)$$

D_p is the diffraction coefficient, given (asymptotically) by (20). The integral $I(s)$ can be evaluated by (37) in region A, by (39) in region B, and by (28) in region C.

Regions D, E, and F are characterized by the fact that

$$k_o(\rho - a) \leq O(k_o a)^{\frac{1}{3}}.$$

In these regions the Debye approximation is not valid, and we have to replace in (47)

$$D_p (2/\pi i)^{\frac{1}{2}} [e^{ik_o b}/(k_o b)^{\frac{1}{2}}]$$

by the expression

$$\frac{\text{Ai} [k_o(\rho - a)(2/k_o a)^{\frac{1}{3}} e^{-i\pi/3}]}{\text{Ai} (0)}.$$

Thus in regions D, E, and F we have

$$u_o(\rho, \theta) \sim \frac{1}{2\pi} \left(\frac{a \cos \alpha_o}{a \cos \alpha_o - d'}\right)$$

$$\times \frac{\text{Ai} [k_o(\rho - a)(2/k_o a)^{\frac{1}{3}} e^{-i\pi/3}]}{\text{Ai} (0)} e^{ik_i d'} I(s). \quad (48)$$

$I(s)$ is again given by Eq. (47a), and can be evaluated by (37) in region D, by (39) in region E, and by (28) in region F. In the transition region I, between the illuminated and the shadow regions, one may apply the method of the preceding section [Eqs. (40) through (46)] to obtain a uniform asymptotic formula connecting smoothly the field representations in the illuminated region (obtained by geometrical optics) and the shadow region. The corresponding $g(t)$ is more complicated than that of the last section.
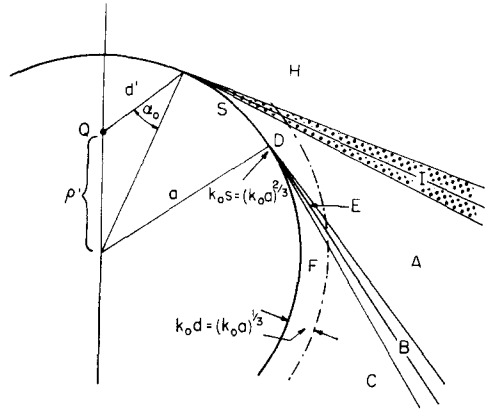

FIG. 8. Regions of validity of the various field representations.

VI. GENERALIZATION TO AN ARBITRARY CONVEX INTERFACE

The geometrical theory of diffraction shows how to generalize our results from a circular interface to an arbitrary convex interface. This can be done as follows:

Let s be arclength along an arbitrary curved interface with $a = a(s)$ the local radius of curvature. Then we replace s in all the above formulas by $\int_0^s ds$ and s/a by $\int_0^s ds/a(s)$. Also α_p must be replaced by [see (21)]

$$\tau_p \left(\frac{ik_o}{2}\right)^{\frac{1}{3}} \int_0^s \frac{ds}{[a(s)]^{\frac{2}{3}}}.$$

The limit $k_o s = (k_o a)^{\frac{2}{3}}$ is replaced by

$$\int_0^s \frac{ds}{[a(s)]^{\frac{2}{3}}} = k_o^{-\frac{1}{3}}.$$

The value of a in the diffraction coefficient D_p should be the radius of curvature at the point of shedding of the diffracted ray. All these principles are discussed in Refs. 1 and 4. Thus we replace (47) by

$$u_o(\rho, \theta) \sim \frac{1}{2\pi} \left(\frac{2i}{\pi}\right)^{\frac{1}{2}} \left[\frac{a(A) \cos \alpha_o}{a(A) \cos \alpha_o - d'}\right]^{\frac{1}{2}}$$

$$\times \frac{e^{i(k_i d' + k_o b)}}{(k_o b)^{\frac{1}{2}}} D_p \hat{I}(s) \quad (49)$$

with $a(A)$ denoting the local radius of curvature of the interface at point A (Fig. 1). D_p may be evaluated by (20), with a replaced by $a(D)$. Likewise

$$\hat{I}(s) = i \int_{-\infty}^{\infty} \left\{\exp \left[i\mu \int_0^s \frac{ds}{a(s)}\right] \Big/ a(s) M(\mu)\right\} d\mu \quad (49a)$$

with $M(\mu)$ given as before by (6), but every a is replaced by $a(s)$. When both source and observation points lie above the interface, a similar analysis of the field in the shadow region ("beyond the horizon") can
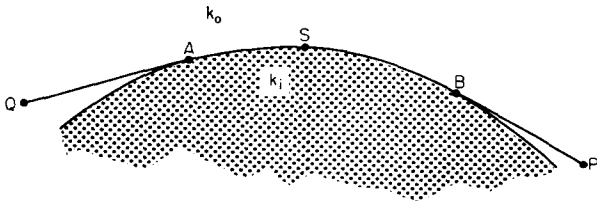
FIG. 9. Path of a diffracted ray when both source and observation point lie in the "fast" medium.

be carried out. For example, the field at $\rho > a$ due to a source at $\rho' > a$ is given by[10]

$$u(\rho, \theta) = \frac{i}{8} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \left\{ H_\mu^{(1)}(k_o\rho') H_\mu^{(2)}(k_o\rho) \right.$$
$$\left. + H_\mu^{(1)}(k_o\rho') H_\mu^{(1)}(k_o\rho) \frac{H_\mu^{(2)}(k_oa)}{H_\mu^{(1)}(k_oa)} R(\mu) \right\} e^{i\mu(\theta+2n\pi)} \, d\mu,$$
(50)

where

$$R(\mu) = \left[k_o \frac{H_\mu^{(2)\prime}(k_oa)}{H_\mu^{(2)}(k_oa)} - k_i \frac{J_\mu'(k_ia)}{J_\mu(k_ia)} \right] [M(\mu)]^{-1}$$
(50a)

and the principal part of the field in the shadow region (Fig. 9) comes from the $n = 0$ term. The appearance of $M(\mu)$ in the denominator of (50) indicates that for small curvature and moderate s [i.e., $1 \ll k_o s < (k_o a)^{\frac{2}{3}}$], the field along the interface (which gives rise to the ray that reaches P) will be proportional to

$$e^{ik_o s} (k_o s)^{-\frac{3}{2}}$$

and the appropriate diffraction coefficients.[3] The existing literature mentions lateral waves only in case that the source is located in the denser medium, since that is the only configuration in which a ray tangential to the interface can be launched on a plane interface. Our analysis shows that this mode of propagation is not restricted to that particular configuration.

ACKNOWLEDGMENTS

APPENDIX

The complex μ plane (Fig. 10) is divided into regions in which different asymptotic expansions for $H_\mu^{(1)}{}^{(2)}(x)$ are valid. The solid and dotted lines separating regions A and B are the loci of the zeros of $H_\mu^{(1)}(x)$ and $H_\mu^{(2)}(x)$, respectively.

Region C is characterized by the relation

$$|\mu - x| \leq x^{\frac{1}{3}}.$$
(A1)

In that region

$$H_\mu^{(1)}(x) \sim (2/\pi) y \, \mathrm{Ai} \, [y(\mu - x)]$$
(A2)

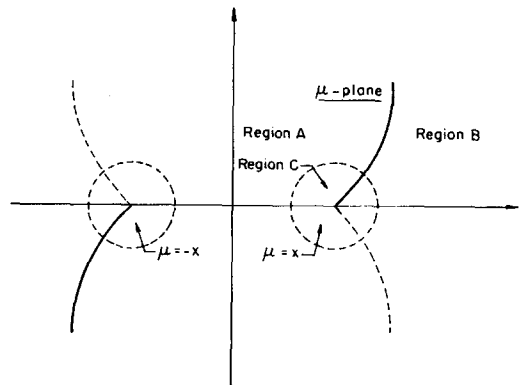[10] W. Streifer and R. Kodis, Quart. Appl. Math. 22, 193 (1964).

FIG. 10. Regions of validity of the various asymptotic expansions of the cylinder functions.

with

$$y = (2/x)^{\frac{1}{3}} e^{-i\pi/3}.$$
(A2a)

In region A

$$H_\mu^{(1)}{}^{(2)}(x)$$

$$\sim (2/\pi)^{\frac{1}{2}} \frac{\exp \{\pm i[(x^2 - \mu^2)^{\frac{1}{2}} - \mu \cos^{-1} \mu / x - \pi/4]\}}{(x^2 - \mu^2)^{\frac{1}{4}}}$$
$$\times [1 + O(x^{-1})].$$
(A3)

In region B

$$H_\mu^{(1)}{}^{(2)}(x)$$

$$\sim \mp i(2/\pi)^{\frac{1}{2}} \frac{\exp \{-(\mu^2 - x^2)^{\frac{1}{2}} + \mu \cosh^{-1} \mu/x\}}{(\mu^2 - x^2)^{\frac{1}{4}}}$$
$$\times [1 + O(x^{-1})].$$
(A4)

For $|x| \gg 1$ and $\mathrm{Im} \, x > 0$, we write

$$J_\mu(x) \sim \tfrac{1}{2} H_\mu^{(2)}(x).$$

When using the change of variables (24), we have for $a \to \infty$

$$(k_i^2 \rho'^2 - \mu^2)^{\frac{1}{2}} = \rho'\{k_i^2 - [a^2 t^2/(a - \eta')^2]^{\frac{1}{2}}\}$$
$$\sim \rho'(k_i^2 - t^2)^{\frac{1}{2}}.$$
(A5a)

Similarly,

$$(k_i^2 \rho^2 - \mu^2)^{\frac{1}{2}} \sim \rho(k_i^2 - t^2)^{\frac{1}{2}}.$$
(A5b)

Thus, from (A3), (A4), and (A5),

$$\frac{H_\mu^{(1)}(k_o\rho) J_\mu(k_i\rho')}{H_\mu^{(1)}(k_oa) J_\mu(k_ia)} e^{i\mu\theta}$$

$$\sim \exp \{i[\eta(k_o^2 - t^2)^{\frac{1}{2}} - \eta'(k_i^2 - t^2)^{\frac{1}{2}} + \xi t]\},$$
(A6)

$$k_o[H_\mu^{(1)\prime}(k_oa)/H_\mu^{(1)}(k_oa)] \sim i(k_o^2 - t^2)^{\frac{1}{2}},$$
(A7)

$$k_i[J_\mu'(k_ia)/J_\mu(k_ia)] \sim -i(k_o^2 - t^2)^{\frac{1}{2}}.$$
(A8)

Equations (A6), (A7), and (A8) are not valid in the neighborhood of $\mu = k_oa$ (region C in Fig. 10); but since the path of integration in (16) and (25) may be deformed so as to avoid this region, we may substitute Eqs. (A6), (A7), and (A8) in these integrals.

Electric Fields in a Semi-Infinite Medium Whose Conductivity Varies Laterally

JAMES E. MANN, JR.
Department of Applied Mathematics and Computer Science,
University of Virginia, Charlottesville, Virginia

(Received 19 September 1966)

The electric field induced in a semi-infinite medium whose conductivity varies laterally is calculated when the inducing field is chosen to approximate a vertically incident magnetic wave which is polarized in the direction of the conductivity variation. The specific form of the variation of conductivity is $\sigma = \sigma_0 + \sigma_1(y/d)^2$, where y is a coordinate parallel to the surface. It is shown for several specific cases that the magnitude of the electric field is less than the electric field in a solid whose conductivity is σ_0. In addition, the electric field is calculated for several values of σ_0/σ_1 at $y = 0$.

I. INTRODUCTION

IN a recent paper, Fischer[1] has given a theoretical discussion of surface impedance when the conductivity has the special form $\sigma = \sigma_0 + \sigma_1 \cos ky$, y being a coordinate parallel to the surface. Aside from the fact that so few theoretical investigations of lateral variations of conductivity are available, investigation of these phenomena is important in certain areas of material science, as pointed out by Fischer.[1] The phenomena associated with such variations may also be of interest in geophysics, especially in geophysical prospecting methods such as AFMAG and the magnetotelluric method.

In this paper, we formulate a problem similar to the one stated by Fischer,[1] but our conductivity function is of an entirely different nature. The conductivity which we use is $\sigma(y) = \sigma_0 + \sigma_1(y/d)^2$. The important differences between this function and $\sigma = \sigma_0 + \sigma_1 \cos ky$ are that it has a single minimum value at $y = 0$ and it is nonperiodic. Like Fischer, we introduce an assumption to render the full wave propagation problem more tractable. Indeed, we assume that current is induced in the solid by a spatially constant magnetic field applied at the surface $z = 0$ (see Fig. 1). Though the assumption of a constant inducing field does away with the concept of wave propagation entirely, the assumption is not so severe as to render the solution useless for understanding physical phenomena.[2] The rest of this paper is devoted to obtaining an expression for the electric field E_x in the x direction (see Fig. 1) when the inducing field $H_y = \text{const} \cdot \exp(i\omega t)$ is applied in the y direction. In addition, we examine the skin depth and surface impedance for this polarization.

II. FORMULATION OF THE PROBLEM

When displacement currents are neglected and $\exp(i\omega t)$ is used to reduce the time dependence, Maxwell's equations take the form

$$\nabla \times \mathbf{H} = \sigma\mathbf{E}, \tag{1}$$

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}. \tag{2}$$

In (1) and (2), $\mathbf{H}$ is the magnetic field, $\mathbf{E}$ the electric field, μ the magnetic permeability, and σ the conductivity; all are in MKS units. By taking the curl of (2) and using (1), we may obtain an equation for $\mathbf{E}$. Thus

$$\nabla \times \nabla \times \mathbf{E} = -i\omega\mu\sigma\mathbf{E}. \tag{3}$$

When all field components are independent of x, Eq. (3) for E_x is

$$\Delta E_x - i\omega\mu\sigma E_x = 0, \tag{4}$$

where

$$\Delta \equiv (\partial^2/\partial y^2) + (\partial^2/\partial z^2).$$

We wish to solve (4) when the inducing field is of the form

$$H_y = \text{const}, \quad \text{on} \quad z = 0. \tag{5}$$

As we shall see, it is impossible to satisfy (5), but a minor modification mitigates the difficulty. We take

$$\sigma = \sigma_0 + \sigma_1(y/d)^2 \tag{6}$$

as the conductivity function, where d is a reference
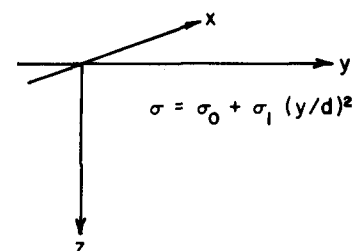
FIG. 1. A conducting half-space. The conductor lies in the region $z \geq 0$.

$$\sigma = \sigma_0 + \sigma_1 \, (y/d)^2$$

[1] G. Fischer, J. Math. Phys. **5**, 1158 (1964).
[2] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1960), pp. 186–188.

length. Using (6), we write (4) as

$$\Delta E_x - i\omega\mu\sigma_1[\delta + (y/d)^2]E_x = 0, \qquad (7)$$

where $\delta = \sigma_0/\sigma_1$. Equation (7) is nondimensionalized by making the following changes of variable:

$$E_x = E_0\psi; \quad y' = y/d,$$
$$z' = z/d; \quad d = (\omega\mu\sigma_1)^{-\frac{1}{2}},$$

where E_0 is a reference value for E_x. In the new variables (we drop the primes on coordinates), Eq. (7) is

$$\Delta\psi - i(\delta + y^2)\psi = 0, \qquad (8)$$

and we choose the constant in (5) in order that

$$-i\omega\mu H_y = \partial E_x/\partial z = E_0/d, \qquad (9)$$

or

$$\partial\psi/\partial z' = 1 \quad \text{on} \quad z = 0. \qquad (10)$$

The first equality of (9) comes directly from (2). We obtain the solution of (8) by separating variables. Thus,

$$Y'' + (b^2 - i\delta - iy^2)Y = 0, \qquad (11)$$

$$X'' - b^2 X = 0, \qquad (12)$$

where $-b^2$ is the separation constant, and Y and X express the y and z variation. By a change of variable, (11) can be transformed to the standard form of the equation whose solutions are the parabolic cylinder functions.[3] If we let

$$\tau = y(2)^{\frac{1}{2}} \cdot \exp{(i\pi/8)}, \qquad (13)$$

Eq. (11) becomes

$$(d^2Y/d\tau^2) - (a + \tfrac{1}{4}\tau^2)Y = 0, \qquad (14)$$

where

$$a = -(b^2 - i\delta)/2 \cdot \exp{(i\pi/4)}. \qquad (15)$$

Equation (14) is exactly the equation of the parabolic cylinder functions. A definite choice of a, and hence of b^2, is made by selecting only the solutions of (11) which are even functions of y and which decay as $y \to \pm\infty$. The only difference between this and the well-known quantum mechanics problems of the parabolic potential is the fact that τ is complex.[4] When the above conditions are imposed, the values of a are restricted to

$$a = -2n - \tfrac{1}{2}, \quad n = 0, 1, 2, 3 \cdots. \qquad (16)$$

When the parameter a is chosen to satisfy (16), the solution of (14) is

$$Y = B_{2n}H_{2n}[\tau/(2)^{\frac{1}{2}}] \exp{[-\tau^2/4]}, \qquad (17)$$

[3] *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C., 1964), Chap. 19, p. 686ff.
[4] R. H. Dicke and J. P. Wittke, *Introduction to Quantum Mechanics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1960), p. 56.

where H_{2n} is the Hermite polynomial of even order. The solution of (12) which decays for large z is

$$X = \exp{(-b_{2n}z)}. \qquad (18)$$

The solution of (8) is, therefore,

$$\psi(y, z) = \sum_{n=0}^{\infty} B_{2n} \exp{[-b_{2n}z - y^2 \exp{(i\pi/4)}/2]}$$
$$\times H_{2n}[y \exp{(i\pi/8)}], \qquad (19)$$

where τ has been eliminated by using (13). The coefficients B_{2n} are determined from (10). That is, the terms B_{2n} are chosen so that

$$-\sum_{n=0}^{\infty} C_{2n} \exp{[-y^2 \exp{(i\pi/4)}/2]}H_{2n}[y \exp{(i\pi/8)}] = 1,$$
$$(20)$$

where $C_{2n} = B_{2n}b_{2n}$. At this point, two difficulties must be overcome. The first, which was mentioned earlier, is that the right-hand side of (20) has an unbounded norm on $(-\infty, \infty)$; this fact makes determination of a convergent series impossible by standard orthogonal function procedures. To rectify this situation, we replace the right-hand side of (20) by $\exp{[-\lambda y^2]}$. By making λ small, we can therefore make $\exp{[-\lambda y^2]}$ arbitrarily close to unity on any finite portion of the y axis and approximate the desired condition.

The second difficulty is that the solutions of (11) are complex-valued functions of the real variable y. However, the Eq. (11) is not Hermitian (the operator is self-adjoint, but not real self-adjoint), and we cannot make use of the orthogonality property of Hermitian operators. This point can be handled as follows. Consider (20) with the substitution

$$\xi = y \exp{(i\pi/8)}, \qquad (21)$$

$$-\sum_{n=0}^{\infty} C_{2n} \exp{[-\xi^2/2]}H_{2n}(\xi) = \exp{[-\lambda\xi^2 \exp{(i\pi/4)}]}.$$
$$(22)$$

Since the right-hand side of (22) is an analytic function of ξ and approaches zero for $\lambda > 0$ as $|\xi| \to \infty$, everywhere within the bow-tie shaped region shown in Fig. 2, we require that (22) be an identity for all complex values of ξ in the region.

Since (22) holds for all ξ in the region described, the equation must hold in particular for real ξ. Hence we may determine C_{2n} using ordinary orthogonality conditions for the Hermite polynomials. Therefore,

$$C_{2n} = -\{\pi^{\frac{1}{2}}2^{2n}(2n)!\}^{-1}$$
$$\times \int_{-\infty}^{\infty} \exp{[-\xi^2(\lambda' + \tfrac{1}{2})]}H_{2n}(\xi) d\xi, \qquad (23)$$
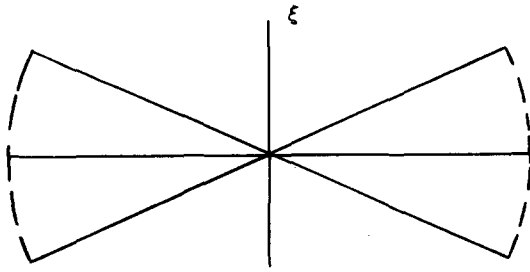
FIG. 2. Region of convergence for (22).

where $\lambda' = \lambda \exp(-i\pi/4)$, and the factor in front of the integral is the square of the normalizing factor for Hermite polynomials. The integral in (23) is evaluated by using the expansion for H_{2n}:

$$H_{2n}(\xi) = \sum_{k=0}^{n} \frac{(-1)^k (2n)!}{k!\,(2n-2k)!} (2\xi)^{2n-2k}. \quad (24)$$

Thus, we obtain

$$C_{2n} = -2^{-2n}(\lambda' + \tfrac{1}{2})^{-\frac{1}{2}} \sum_{k=0}^{n} \frac{(-1)^k}{k!\,(n-k)!} \left(\frac{1}{\lambda' + \tfrac{1}{2}}\right)^{n-k}. \quad (25)$$

Since the series in (25) is part of the binomial expansion, the expression for C_{2n} may be reduced to the simple form

$$C_{2n} = -\frac{(-1)^n}{2^{2n}(\lambda' + \tfrac{1}{2})^{\frac{1}{2}} n!} \left(\frac{\lambda' - \tfrac{1}{2}}{\lambda' + \tfrac{1}{2}}\right)^n. \quad (26)$$

Using (26), (20), and (19), we obtain the following expression for the electric field in the medium:

$$\psi(y, z) = -(\lambda' + \tfrac{1}{2})^{-\frac{1}{2}} \sum_{n=2}^{\infty} \frac{(-1)^n}{2^{2n} b_{2n} n!} \left(\frac{\lambda' - \tfrac{1}{2}}{\lambda' + \tfrac{1}{2}}\right)^n$$
$$\times \exp\left[-b_{2n}z - y^2 \exp(i\pi/4)/2\right]$$
$$\times H_{2n}[y \exp(i\pi/8)]. \quad (27)$$

Again, we emphasize that to obtain a nearly constant magnetic field for a given finite interval of the y axis, we must limit the size of λ and hence the absolute value of λ'. If we define impedance in the usual way as the ratio of the electric field to the orthogonal component of magnetic field evaluated at $z = 0$, we have

$$Z \propto \psi(y, 0), \quad (28)$$

where $Z = E_x/H_y$. We see that the absolute value of the impedance approaches zero as y increases, though the exact manner of this approach requires numerical evaluation of (28). Using formulas given by Erdelyi,[5] we can obtain the asymptotic behavior of the function in the series. For fixed x and large order of the

⁵ A. Erdelyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, Chap. 10.

Hermite polynomials, we have

$$H_{2m}(x) \approx (-1)^m 2^{2m} \exp(x^2/2)\Gamma(m + \tfrac{1}{2})[x(4m)^{\frac{1}{2}}/2]^{\frac{1}{2}}$$
$$\times [J_{-\frac{1}{2}}[x(4m)^{\frac{1}{2}}] + x/[4(4m)^{\frac{1}{2}}]J_{\frac{1}{2}}[x(4m)^{\frac{1}{2}}]], \quad (29)$$

where x is a complex variable with unrestricted argument. Hence, we have

$$H_{2m}[y \exp(i\pi/8)]$$
$$\approx (-1)^m 2^{2m}(\pi)^{\frac{1}{2}} \exp[y^2 \exp(i\pi/4)/2]$$
$$\times \Gamma(m + \tfrac{1}{2}) \exp[|y| \sin(\pi/8)(4m)^{\frac{1}{2}}]$$
$$\times \{\cos[y \cos(\pi/8)(4m)^{\frac{1}{2}}]$$
$$- i \sin[|y| \cos(\pi/8)(4m)^{\frac{1}{2}}] + O(n^{-\frac{1}{2}})\}. \quad (30)$$

Also,

$$b_{2m} \approx (4m)^{\frac{1}{2}} \exp(i\pi/8), \quad (31)$$

where $i\delta$ has been neglected. Using (27), (30), and (31), we can obtain a typical term in the series for large n. Thus

$$\psi(y, z) = -(\lambda' + \tfrac{1}{2})^{-\frac{1}{2}} \sum_{n=0}^{M} \frac{(-1)^n}{2^{2n} b_{2n} n!} (\beta)^n$$
$$\times \exp[-b_{2n}z - y^2 \exp(i\pi/4)/2]$$
$$\times H_{2n}[y \exp(i\pi/8)]$$
$$- \exp(-i\pi/4)(\lambda' + \tfrac{1}{2})^{-\frac{1}{2}} \sum_{n=M+1}^{\infty} A_{2n}(y, z), \quad (32)$$

where

$$A_{2n}(y, z)$$
$$= \exp[-n \log \beta - 2n^{\frac{1}{2}}\{z \exp(i\pi/4) - |y| \sin(\pi/8)\}]$$
$$\times [\cos(y \cos(\pi/8)2n^{\frac{1}{2}}) - i \sin(|y| \cos(\pi/8)2n^{\frac{1}{2}})]$$
$$\times [2^{\frac{1}{2}}m]^{-1}. \quad (33)$$

In (32) and (33),

$$\beta = (\lambda' - \tfrac{1}{2})/(\lambda' + \tfrac{1}{2}), \quad (34)$$

and M is an integer large enough for (30) and (31) to be simultaneously valid. The fact that the terms A_{2n} become exponentially small as $n \to \infty$ in (32) establishes absolute and uniform convergence of the series in (27). The series converges most slowly when $z = 0$. On $z = 0$ we see that the A_{2n} terms do not begin to decrease exponentially until n is large enough for

$$(n)^{\frac{1}{2}} > 2|y| \sin(\pi/8)/\log|\beta|. \quad (35)$$

In addition, we only want to consider y in a range such that

$$\exp(-\lambda y^2) \geq 0.95. \quad (36)$$

Equation (36) implies

$$\lambda < 0.05 y^{-2}, \quad (37)$$

which implies

$$|\beta| \approx 1 - 0.14 y^{-2}. \quad (38)$$

Thus, for large y, a large number of terms in the series of (27) are required if $z = 0$. Using (38) and (35), we see that the number of terms required before the magnitude begins to decrease exponentially is proportioned to y^6. Because of this rapid increase in the number of terms and because of the lack of tabulated values for the Hermite polynomials with complex argument, we examine ψ at $y = 0$ only. At $y = 0$, we use the formula

$$H_{2n}(0) = (-1)^m [(2m)!/m!]. \tag{39}$$

Therefore,

$$\psi(0, z) = -(\lambda' + \tfrac{1}{2})^{-\frac{1}{2}} \sum_{n=0}^{\infty} \frac{\beta^n (2n)!}{2^{2n}(n!)^2 b_{2n}} \exp(-b_{2n}z). \tag{40}$$

This series is also poorly convergent when $z = 0$ as the terms go down like β^n/n with the magnitude of β close to unity. Rapidity of convergence can be greatly improved by noting

$$S = \sum_{n=1}^{\infty} \frac{\beta^n}{n} = -\log(1 - \beta) \tag{41}$$

and adding and subtracting an appropriate multiple of S from (40).

III. RESULTS

The result of calculating $\psi(0, z)$ for several values of δ and $\lambda = 0.001$ is shown in Fig. 3. We note that λ^{-1} is a measure of the distance over which the applied field is constant; in addition δ^{-1} is a measure of how rapidly the conductivity is changing near the origin. The surface impedance at the origin is proportional to $\psi(0, 0)$, and we see how this changes in response to changes in δ. We also can see how the skin-depth changes in response to changes in δ. The electric field becomes virtually independent of δ when $\delta < 0.1$; this fact indicates that local conductivity is not important when the conductivity varies rapidly. In making comparisons of skin depth for the case at hand with the skin depth in a homogeneous conductor, we consider the way in which z was nondimensionalized in (8). Thus we define skin depth z_1 to be that depth for which

$$|\psi(0, 0)|/|\psi(0, z_1)| = e. \tag{42}$$

For a homogeneous conductor, we would have the relation

$$z_1 \delta^{\frac{1}{2}} = 2^{\frac{1}{2}}. \tag{43}$$

In Fig. 3, the broken lines are the graphs of electric field in a homogeneous conductor whose conductivity is δ. The solid lines are the electric field at $y = 0$ in a
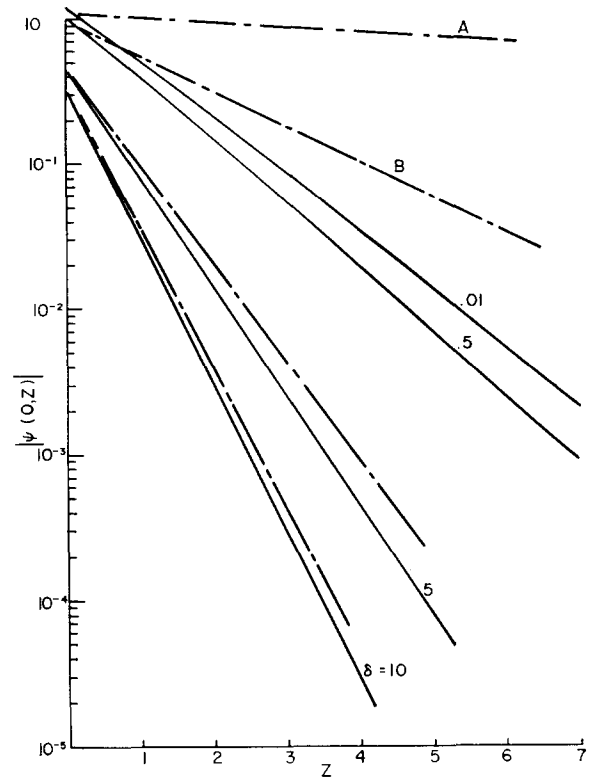
FIG. 3. A graph of the electric field at $y = 0$ versus depth z for several values of δ and $\lambda = 10^{-3}$. The broken lines are graphs of electric field in a homogeneous half-space whose conductivity is δ. The lines marked A and B have correct slope but should have the values of $\psi(0, 0)$ at 10 and 1.414, respectively, and correspond to $\delta = 0.01$ and 0.5.

medium whose conductivity is $\delta + y^2$. We see that when δ is small, there is a great difference between the homogeneous solid and the nonhomogeneous one. The difference manifests itself in two ways. First, the electric field at the surface of the solid is less than the electric field at the surface of the homogeneous solid. Second, the electric field decays more rapidly than in the homogeneous solid. These differences reflect the fact that the conductivity near $y = 0$ is considerably different from δ. When δ is less than 0.1, the solution is virtually independent of δ, indicating the dominant influence of neighboring material.

When δ is large on the other hand, $\psi(0, z)$ is nearly the same as the field in a homogeneous solid.

Degree of Higher-Order Optical Coherence*

C. L. MEHTA

Department of Physics and Astronomy, University of Rochester, Rochester, New York

(Received 27 June 1966)

In this paper we define the normalized coherence function of arbitrary order (m, n), in a manner which seems to be a natural generalization of that defined for the second-order coherence function. Both classical and quantized optical fields are considered and the results are compared. It is shown that for classical fields and also for quantized optical fields having nonnegative definite diagonal coherent state representations of the density operator, the modulus of these normalized coherence functions is bounded by the values 0 and 1. This definition differs from the one recently given by Glauber for quantized optical fields, where the normalized coherence functions may take arbitrarily large values even for fields having nonnegative definite diagonal representations of the density operator. Conditions for "complete coherence," i.e., those under which the modulus of the normalized coherence function attains the limiting value 1, are discussed. Some consequences of stationarity and quasi-monochromaticity are also discussed.

1. INTRODUCTION

IN the analysis of many experiments involving correlation measurements, the "degree of coherence" has played an important role. The degree of coherence has been defined in the past[1] by the equation

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)/\{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0)\Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)\}^{\frac{1}{2}},$$
(1.1)

where

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle V^*(\mathbf{r}_1, t)V(\mathbf{r}_2, t + \tau)\rangle_t \quad (1.2)$$

is the mutual coherence function representing the correlation between disturbances at the two space-time points $\mathbf{r}_1, t$; $\mathbf{r}_2, t + \tau$ in a stationary optical field. Here $V(\mathbf{r}, t)$ is the analytic signal representing the light disturbance[2] at the point $\mathbf{r}$ at time t and $\langle \ \rangle_t$ denotes the time average. The normalization in (1.1) ensures that

$$0 \leq |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \leq 1. \quad (1.3)$$

In recent years it has been recognized that, in order to understand the coherence properties of optical fields other than those due to a thermal light source, it is necessary to consider higher-order coherence functions also. The (m, n)th-order coherence function[3]

is, for example, defined as[4]

$$\Gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \langle V^*(x_1) \cdots V^*(x_m)V(x_1') \cdots V(x_n')\rangle, \quad (1.4)$$

where x_j and x_j' denote the space-time points $\mathbf{r}_j, t_j$ and $\mathbf{r}_j', t_j'$, respectively, and the sharp brackets now refer to an appropriate average which may be the ensemble average or, if the field is stationary, the time average. In quantum description, the (m, n)th-order coherence function is defined as[5]

$$G^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \langle \hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)\hat{A}^{(+)}(x_1') \cdots \hat{A}^{(+)}(x_n')\rangle_q,$$
$$\equiv \text{Tr}\{\hat{\rho}\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)\hat{A}^{(+)}(x_1') \cdots \hat{A}^{(+)}(x_n')\},$$
(1.5)

where $\hat{\rho}$ is the density operator[6] describing the field and $\hat{A}^{(+)}(x)$ and $\hat{A}^{(-)}(x)$ are the positive and negative frequency parts, respectively, of the appropriate field operator (see Footnote 2).

In analogy with the second-order degree of coherence $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$, it is desirable to introduce higher-order normalized coherence functions, which would then provide a quantitative measure of higher-order coherence. Glauber[5,7] has introduced $2n$th-order normalized coherence function[8] for quantized optical fields defined by

$$g^{(n)}(x_1, \cdots, x_n; x_1', \cdots, x_n')$$
$$= \frac{G^{(n,n)}(x_1, \cdots, x_n; x_1', \cdots, x_n')}{\prod_{j=1}^{n} \{G^{(1,1)}(x_j; x_j)G^{(1,1)}(x_j'; x_j')\}^{\frac{1}{2}}}. \quad (1.6)$$

[1] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., Oxford, England, 1965), 3rd ed., p. 501.
[2] For simplicity we consider only scalar fields, but a generalization to vector fields is straightforward; cf. Sec. 6.
[3] For obvious reasons, we introduce a pair of indices (m, n) to indicate the order of coherence. Thus we call $\Gamma^{(m,n)}$ or $G^{(m,n)}$ the (m, n)th-order coherence function. However, in the special case when $m = n$, we call $\Gamma^{(m,m)}$ or $G^{(m,m)}$ the $(2m)$th-order coherence function.

[4] E. Wolf, *Proceedings of the Symposium on Optical Masers* (John Wiley & Sons, Inc., New York, 1963), p. 29.
[5] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).
[6] In this paper all symbols with a circumflex denote operators.
[7] U. M. Titulaer and R. J. Glauber, Phys. Rev. **140**, B676 (1965).
[8] Cf. also L. Mandel and E. Wolf, Rev. Mod. Phys. **37**, 231 (1965), Sec. 4.4.

However, as Glauber already noted, this definition does not guarantee that $|g^{(n)}|$ is bounded between 0 and 1, even for the most common type of optical fields, namely those generated by thermal sources. There is hardly any justification, therefore, in calling a function defined by (1.6) the degree of higher-order coherence. Moreover, no attempt has yet been made to define normalized coherence functions of the arbitrary order (m, n) when $m \neq n$.

It is sometimes asserted[9] that for stationary light beams the correlation function $G^{(m,n)}$ is different from zero only when $m = n$. However, this is not so in general, as can be seen from the discussion given in a recent paper by Kano.[10] If, however, the light is also quasi-monochromatic, the assumption that $G^{(m,n)}$ is different from zero only when $m = n$ is reasonable.

We begin by giving a simple proof of the last statement[11] and then present an alternative definition for the degree of coherence of arbitrary order (m, n) which seems to be a natural generalization of the usual definition for the case $m = n = 1$.

2. STATIONARITY AND QUASI-MONOCHROMATICITY

To see the consequences of stationarity and quasi-monochromaticity,[12] let us expand the field operator $\hat{A}^{(+)}(x)$ in a complete set of plane wave modes:

$$\hat{A}^{(+)}(x) = \frac{1}{L^{\frac{3}{2}}} \sum_{\mathbf{k}} \hat{a}_{\mathbf{k}} \exp\{i(\mathbf{k} \cdot \mathbf{r} - ckt)\}, \quad (2.1)$$

where the summation over $\mathbf{k}$ is restricted to $\mathbf{k}$ values such that

$$k_0 - \Delta k \leq k = |\mathbf{k}| \leq k_0 + \Delta k. \quad (2.2)$$

We may then express $G^{(m,n)}$ in the form

$$G^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \frac{1}{L^{\frac{3}{2}(m+n)}} \sum_{\{\mathbf{k}_i\}} \mathrm{Tr}\, \{\hat{\rho}\, \hat{a}_{\mathbf{k}_1}^\dagger \cdots \hat{a}_{\mathbf{k}_m}^\dagger \, \hat{a}_{\mathbf{k}_1'} \cdots \hat{a}_{\mathbf{k}_n'}\}$$
$$\times \exp\{-i[\mathbf{k}_1 \cdot \mathbf{r}_1 + \cdots + \mathbf{k}_m \cdot \mathbf{r}_m$$
$$\qquad\qquad - \mathbf{k}_1' \cdot \mathbf{r}_1' - \cdots - \mathbf{k}_n' \cdot \mathbf{r}_n']$$
$$\qquad + ic[k_1 t_1 + \cdots + k_m t_m - k_1' t_1' - \cdots - k_n' t_n']\}. \quad (2.3)$$

[9] R. J. Glauber, in *Quantum Optics and Electronics, Les Houches,* 1964, C. DeWitt, A. Blandon, and C. Cohen-Tannoudji, Eds. (Gordon and Breach Science Publishers, Inc., New York, 1965), p. 78; Phys. Rev. **131**, 2785 (1963).

[10] Y. Kano, Ann. Phys. (N.Y.) **30**, 127 (1964).

[11] Cf. C. L. Mehta and L. Mandel, in *Electromagnetic Wave Theory,* Proceedings of a Symposium held at Delft, The Netherlands, September, 1965 (Pergamon Press, Ltd., Oxford, England, 1967), p. 1069.

[12] The main result of this section appropriate for classical coherence functions has been proved in Ref. 11. We therefore consider in this section only coherence functions of the quantized optical fields.

If the field is stationary, $G^{(m,n)}$ must be independent of the origin of time. Thus, if we replace each t_j, t_j' in the expression on the right-hand side of (2.3) by $t_j + t$, $t_j' + t$, we require that the resultant expression be independent of t. On taking the space–Fourier transformation, we then obtain, whenever the coefficient $\mathrm{Tr}\, (\hat{\rho}\hat{a}_{\mathbf{k}_1}^\dagger \cdots \hat{a}_{\mathbf{k}_{n'}})$ is different from zero, the identity

$$\sum_{i=1}^{m} k_i = \sum_{j=1}^{n} k_j'. \quad (2.4)$$

Using (2.2), the left-hand side of Eq. (2.4) is seen to be smaller than or equal to $m(k_0 + \Delta k)$, whereas the right-hand side is greater than or equal to $n(k_0 - \Delta k)$. We thus obtain the inequality

$$mk_0 + m\Delta k \geq nk_0 - n\Delta k, \quad (2.5)$$

which, on simplification, gives the inequality

$$|(m - n)/(m + n)| \leq \Delta k/k_0. \quad (2.6)$$

Inequality (2.6) clearly shows that if the light is also quasi-monochromatic, i.e., if $\Delta k/k_0 \ll 1$, and if $m \neq n$, then $G^{(m,n)}$ will be different from zero only if m and n are very large. As an example, let us assume that $\Delta k/k_0 \sim 10^{-7}$; then the inequality (2.6) can be satisfied for different integral values of m and n only if m and n are of the order of 10^7 or larger. Since coherence functions of such a large order are of no practical interest, we may safely assume that for stationary quasi-monochromatic radiation, it is adequate to consider only even-order coherence functions $G^{(m,m)}$. However, if the field is not quasi-monochromatic, we must in general also consider $G^{(m,n)}$ with unequal indices m and n.

3. DEGREE OF HIGHER-ORDER COHERENCE IN CLASSICAL DESCRIPTION

We define the (m, n)th-order normalized coherence function by the equation

$$\gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \frac{\Gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')}{\left[\prod_{j=1}^{m} \langle\{I(x_j)\}^m\rangle\right]^{1/2m} \left[\prod_{j=1}^{n} \langle\{I(x_j')\}^n\rangle\right]^{1/2n}}. \quad (3.1)$$

Here $I(x) = V^*(x)V(x)$ is the instantaneous intensity at the space–time point x and

$$\langle\{I(x)\}^m\rangle \equiv \Gamma^{(m,m)}(x, \cdots, x; x, \cdots, x). \quad (3.2)$$

The normalization in (3.1) guarantees that

$$0 \leq |\gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')| \leq 1. \quad (3.3)$$

We note in passing that, when all the $2m$ space–time points coincide,

$$\gamma^{(m,m)}(x, \cdots, x; x, \cdots, x) = 1. \quad (3.4)$$

In order to show that (3.3) is true, we make use of a slight generalization of the Hölder–Schwarz inequality[13]: Given a nonnegative definite function $\phi(\{v\})$, we have, for arbitrary functions $f_1(\{v\})$, $f_2(\{v\})$, $\cdots$, $f_N(\{v\})$ of a set of variables $\{v\} \equiv v_1$, v_2, $\cdots$, v_k, the inequality

$$\left| \int \phi(\{v\}) f_1(\{v\}) \cdots f_N(\{v\}) \, d\{v\} \right|$$

$$\leq \prod_{j=1}^{N} \left\{ \int \phi(\{v\}) \, |f_j(\{v\})|^{\mathcal{A}_j} \, d\{v\} \right\}^{\mathcal{A}_j^{-1}}, \quad (3.5)$$

where $\mathcal{A}_j > 0$ $(j = 1, 2, \cdots, N)$ and $\sum_{j=1}^{N} \mathcal{A}_j^{-1} = 1$. Relation (3.5) reduces to an equality if and only if the following conditions are satisfied: $\sum_j \arg f_j(\{v\})$ is a constant, and the modulus of each of the functions $f_j(\{v\})$ is effectively proportional to each other, i.e., there exist some constants α_1, α_2, $\cdots$, α_N, and β such that

$$\frac{\phi(\{v\}) \, |f_1|}{\alpha_1} = \frac{\phi(\{v\}) \, |f_2|}{\alpha_2} = \cdots;$$

and

$$\sum_{j=1}^{N} \arg f_j = \beta. \quad (3.6)$$

If in (3.5) we choose

$$N = m + n;$$

$$f_j = V^*(x_j), \qquad \mathcal{A}_j = 2m, \quad (j = 1, 2, \cdots, m);$$

$$f_{m+k} = V(x_k'), \qquad \mathcal{A}_{m+k} = 2n, \quad (k = 1, 2, \cdots, n),$$

and identify $\phi(\{v\})$ with the weighting factor[14] used in evaluating the averages in the defining equation (1.4), we obtain the inequality

$$|\Gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')|$$

$$\leq \left\{ \prod_{j=1}^{m} \Gamma^{(m,m)}(x_j, \cdots, x_j; x_j, \cdots, x_j) \right\}^{1/2m}$$

$$\times \left\{ \prod_{j=1}^{n} \Gamma^{(n,n)}(x_j', \cdots, x_j'; x_j', \cdots, x_j') \right\}^{1/2n}. \quad (3.7)$$

[13] See, for example, G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities* (Cambridge University Press, Cambridge, England, 1934), p. 140.

[14] If the average in Eq. (1.4) is the ensemble average, then the weighting factor $\phi(\{v\})$ may be regarded as the generalized ensemble distribution (which in classical description is necessarily nonnegative) over the Fourier components $\{v\}$ of $V(x)$:

$$V(x) = \frac{1}{L^{\frac{3}{2}}} \sum_{\mathbf{k}} v_{\mathbf{k}} \exp [i(\mathbf{k} \cdot \mathbf{r} - ckt)].$$

If, however, we are taking time averages in (1.4), so that

$$\Gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} dt \, V^*(\mathbf{r}_1, t_1 + t) \cdots V(\mathbf{r}_n, t_n' + t),$$

then the weighting factor $\phi(\{v\})$ may be regarded as the rectangular function $1/2T$ for $|t| \leq T$ and zero otherwise. The set of variables $\{v\}$ reduces now to a single variable t and the limit $T \to \infty$ is taken at the end of the calculations.

From the relations (3.7) and (3.1) we immediately obtain the required result (3.3).

In analogy with second-order coherence theory, we call $\gamma^{(m,n)}$ the "(complex) degree of (m, n)th-order coherence." The limiting cases of (3.3), namely $|\gamma^{(m,n)}| = 0$ and $|\gamma^{(m,n)}| = 1$, are said to characterize complete "incoherence" and complete "coherence," respectively, of order (m, n). Whenever $m = n$, we call $\gamma^{(m,m)}$ the degree of $(2m)$th-order coherence[3]. The case $|\gamma^{(m,m)}| = 1$ is thus said to characterize complete coherence of order $2m$.

Let us now examine the consequences of complete $(2m)$th-order coherence. It is seen from the definition that whenever $\Gamma^{(m,m)}$ "factorizes" in the form

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$

$$= U_m^*(x_1) \cdots U_m^*(x_m) U_m(x_1') \cdots U_m(x_m'), \quad (3.8)$$

the normalized coherence function $\gamma^{(m,m)}$ is unimodular:

$$|\gamma^{(m,m)}| \equiv 1 \quad (3.9)$$

for all values of its arguments.

One can show that the converse is also true: i.e., (3.9) implies that $\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$ has the form given by (3.8). For this purpose we consider the general conditions under which the relation (3.5) reduces to an equality [i.e., when Eqs. (3.6) holds]. Applying these conditions to the present case of complete $(2m)$th-order coherence, we see that whenever $|\gamma^{(m,m)}| \equiv 1$, the stochastic variable $V(x)$ must satisfy, with unit probability, the relations

$$\left. \begin{aligned} \frac{\phi(\{v\}) \, |V(x_1)|}{\alpha(x_1)} &= \frac{\phi(\{v\}) \, |V(x_2)|}{\alpha(x_2)} = \cdots; \\ \arg V(x) &= \beta(x), \end{aligned} \right\} \quad (3.10)$$

where $\alpha(x)$ and $\beta(x) - \beta(x')$ are not stochastic variables, and $\phi(\{v\})$ is again the weighting factor used in evaluating the averages in (1.4). If we substitute (3.10) in the defining equation for $\Gamma^{(m,m)}$, we find that

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$

$$= \frac{\Gamma^{(m,m)}(x_1, \cdots, x_1; x_1, \cdots, x_1)}{\{\alpha(x_1)\}^{2m}}$$

$$\times \prod_{j=1}^{m} \{\alpha(x_j)\alpha(x_j')e^{-i[\beta(x_j)-\beta(x_j')]}\}. \quad (3.11)$$

Since the left-hand side of (3.11) is symmetric in x_1 and x_2, the function $\{\alpha(x_1)\}^{2m}$ must be proportional to $\Gamma^{(m,m)}(x_1, \cdots, x_1; x_1, \cdots, x_1)$. Without loss of

generality, we can choose the constant of pro-portionality to be unity, and we thus find that $(2m)$th-order coherence implies that

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= U_m^*(x_1) \cdots U_m^*(x_m) U_m(x_1') \cdots U_m(x_m'), \quad (3.12)$$

where

$$U_m(x) = \alpha(x) e^{i\beta(x)}$$
$$= \{\Gamma^{(m,m)}(x, \cdots, x; x, \cdots, x)\}^{1/2m} e^{i\beta(x)}. \quad (3.13)$$

We may summarize the result which we have now established in Theorem 1:

Theorem 1: The necessary and sufficient condition for the case of complete $2m$th-order coherence, i.e., for the validity of the identity $|\gamma^{(m,m)}| \equiv 1$, is that $\Gamma^{(m,m)}$ has the factorized form

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= U_m^*(x_1) \cdots U_m^*(x_m) U_m(x_1') \cdots U_m(x_m').$$

Let us now consider the effect of complete $2m$th-order coherence on $\Gamma^{(n,n)}$ when $n \neq m$. Again if we substitute (3.10) in the defining equation for $\Gamma^{(n,n)}$ and follow a similar argument given above to obtain a factorization theorem, we find that $2m$th-order coherence implies that for *any* n

$$\Gamma^{(n,n)}(x_1, \cdots, x_n; x_1', \cdots, x_n')$$
$$= U_n^*(x_1) \cdots U_n^*(x_n) U_n(x_1') \cdots U_n(x_n'), \quad (3.14)$$

where

$$U_n(x) = \{\Gamma^{(n,n)}(x, \cdots, x; x, \cdots, x)\}^{1/2n} e^{i\beta(x)}. \quad (3.15)$$

We have thus established Theorem 2:

Theorem 2: Complete coherence to any even order $2m$ implies complete coherence to all even orders $2n$.

If the optical field is stationary, it is known that complete second-order coherence implies mono-chromaticity.[15] Hence we conclude that the only field that may be stationary and coherent to all even orders is a monochromatic field.

In the above discussion, we have not considered the case of complete (m, n)th-order coherence when $m \neq n$. For, in most situations of practical interest, namely for quasimonochromatic stationary optical fields, $\Gamma^{(m,n)}$ can for all practical purposes be assumed to be zero when $m \neq n$, as shown in Sec. 2. Further, it can be shown quite generally that if for all $x_1, \cdots,$

$x_m, x_1', \cdots, x_n'$

$$|\gamma^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')| \equiv 1, \quad (3.16)$$

then we also have for all x, x' the identity

$$|\gamma^{(1,1)}(x; x')| \equiv 1.$$

Thus the condition $|\gamma^{(m,n)}| \equiv 1$, $m \neq n$, for stationary fields leads to a contradiction, since this implies $|\gamma^{(1,1)}| = 1$, which in turn implies monochromaticity, and for such fields $\gamma^{(m,n)} = 0$, if $m \neq n$.

4. DEGREE OF HIGHER-ORDER COHERENCE IN QUANTUM DESCRIPTION

In this section we wish to define the degree of higher-order coherence for quantized optical fields. The (m, n)th-order coherence function is now defined by Eq. (1.5). In analogy with (3.1) we define the normalized coherence function of order (m, n) for the quantized optical fields by the equation

$$g^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \frac{G^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')}{\left\{ \prod_{j=1}^{m} \langle :[\hat{I}(x_j)]^m: \rangle_q \right\}^{1/2m} \left\{ \prod_{j=1}^{n} \langle :[\hat{I}(x_j')]^n: \rangle_q \right\}^{1/2n}}. \quad (4.1)$$

Here $\hat{I}(x) = \hat{A}^{(-)}(x)\hat{A}^{(+)}(x)$ is the operator repre-senting the intensity at the space–time point x and colons denote normal ordering, i.e.,

$$:\{\hat{I}(x)\}^m: \equiv \{\hat{A}^{(-)}(x)\}^m \{\hat{A}^{(+)}(x)\}^m$$

and

$$\langle :\{I(x)\}^m: \rangle = G^{(m,m)}(x, \cdots, x; x, \cdots, x). \quad (4.2)$$

It is obvious from the definition that when all the $2m$ space–time points coincide, we have

$$g^{(m,m)}(x, \cdots, x; x, \cdots, x) = 1. \quad (4.3)$$

We will show below that, at least for the fields having a nonnegative definite diagonal coherent state repre-sentation[16] of the density operator $\hat{\rho}$, the inequality

$$0 \leq |g^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')| \leq 1 \quad (4.4)$$

---

[16] The existence of the diagonal representation [Eq. (4.5)] was first observed by E. C. G. Sudarshan [Phys. Rev. Letters **10**, 277 (1963)], who also stressed its *universal validity*. R. J. Glauber [Phys. Rev. **131**, 2766 (1963)] also noticed the possibility of such a repre-sentation in *some special* cases and called it the P-representation, but he still denies its usefulness in the general case. Sudarshan's original formulation of the diagonal representation was somewhat heuristic and a rigorous mathematical meaning to such a representa-tion was given later [cf. C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. **138**, B274 (1965); J. R. Klauder, J. McKenna, and D. G. Currie, J. Math. Phys. **6**, 733 (1965); J. R. Klauder, Phys. Rev. Letters **16**, 534 (1966); C. L. Mehta, *ibid.* **18**, 752 (1967)]. In cases when $\phi(\{v\})$ cannot be defined as an ordinary function, the right-hand side of Eq. (4.6) is to be interpreted as

$$\lim_{N \to \infty} \int \phi_N(\{v\}) V^*(x_1) \cdots V^*(x_m) V(x_1') \cdots V(x_n) d^2\{v\},$$

where the functions $\phi_1(\{v\}), \phi_2(\{v\}), \cdots, \phi_N(\{v\}), \cdots$, can be chosen to be well-behaved functions, such that the sequence of the corre-sponding density operators $\hat{\rho}_N = \int \phi_N(\{v\}) |\{v\}\rangle \langle\{v\}| d^2\{v\}$ converges (in the norm) to $\hat{\rho}$.

---

[15] C. L. Mehta, E. Wolf, and A. P. Balachandran, J. Math. Phys. **7**, 133 (1966).

holds. If we express the density operator $\hat{\rho}$ in the diagonal coherent state representation[17]

$$\hat{\rho} = \int \phi(\{v\}) \, |\{v\}\rangle\langle\{v\}| \, d^2\{v\}, \qquad (4.5)$$

then it is well known (cf. references given in Footnote 16; see also Ref. 17) that the correlation functions $G^{(m,n)}$ are expressible in the form

$$G^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \int \phi(\{v\}) V^*(x_1) \cdots V^*(x_m) V(x_1') \cdots V(x_n') \, d^2\{v\}. \qquad (4.6)$$

Here $|\{v\}\rangle = \prod_k |v_k\rangle$ and $|v_k\rangle$ is the eigenstate of the annihilation operator $\hat{a}_k$ which appears in the plane wave expansion of the field operator $\hat{A}^{(+)}(x)$ [Eq. (2.1)], with the eigenvalue v_k:

$$\hat{a}_k |v_k\rangle = v_k |v_k\rangle, \qquad (4.7)$$

and

$$V(x) = \frac{1}{L^{\frac{3}{2}}} \sum v_k e^{i(k \cdot r - ckt)}. \qquad (4.8)$$

In the case when $\phi(\{v\})$ is nonnegative definite, the quantum mechanical expectation value (1.5) can thus be regarded as an average over a classical ensemble. In this case we can apply the arguments of Sec. 3, and we see immediately that all the results of that section are also valid for the quantum case. In particular, we have the inequality

$$0 \leq |g^{(m,n)}| \leq 1. \qquad (4.9)$$

However, in the general case when $\phi(\{v\})$ is not necessarily nonnegative definite, $|g^{(m,n)}|$ may not be bounded between 0 and 1 and, in fact, may take on arbitrarily large values. We will, however, call $g^{(m,n)}$ the complex degree of (m, n)th-order coherence. Whenever $|g^{(m,n)}|$ exceeds unity, it corresponds to a truly quantum feature of the optical field.

Let us now examine the consequences of complete $(2m)$th-order coherence for the general case when $\phi(\{v\})$ is not necessarily nonnegative definite. It is shown in the Appendix that the coherence function $g^{(m,m)}$ satisfies a nonnegative definiteness condition

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i^* \alpha_j g^{(m,m)}(x_1^{(i)}, \cdots, x_m^{(i)}; x_1^{(j)}, \cdots, x_m^{(j)}) \geq 0, \qquad (4.10)$$

where N is any positive integer, $\alpha_1, \cdots, \alpha_N$ are arbitrary complex constants, and for every i ($i = 1, 2, \cdots, N$), $x_1^{(i)}, \cdots, x_m^{(i)}$ is a set of m arbitrary space–time points. It is also shown in the Appendix,

following an argument similar to the one given in a recent paper,[15] that relation (4.10), when applied to the case of fields having complete $(2m)$th-order coherence, requires that $g^{(m,m)}$ must be of the form

$$g^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= \exp\{-i[f(x_1, \cdots, x_m) - f(x_1', \cdots, x_m')]\}, \qquad (4.11)$$

where f is some function of m space–time points.

We now prove the factorization theorem for the quantum coherence functions. Since the density operator $\hat{\rho}$ is nonnegative definite, we have

$$\mathrm{Tr}\left\{\hat{\rho}\left[\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)\right.\right.$$
$$\left. - \frac{G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} (\hat{A}^{(-)}(x_0))^m\right]$$
$$\times \left[\hat{A}^{(+)}(x_1) \cdots \hat{A}^{(+)}(x_m)\right.$$
$$\left.\left. - \frac{G^{(m,m)}(x_0, \cdots, x_0; x_1, \cdots, x_m)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} (\hat{A}^{(+)}(x_0))^m\right]\right\} \geq 0, \qquad (4.12).$$

with the equality sign holding if and only if

$$\hat{\rho}\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)$$
$$= \frac{G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} \hat{\rho}(\hat{A}^{(-)}(x_0))^m. \qquad (4.13)$$

In (4.12) and (4.13), x_0 is so chosen that

$$G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0) \neq 0.$$

This is always possible in the case of complete $(2m)$th-order coherence. Relation (4.12) can be expressed more simply as an inequality[18]:

$$|G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)|^2$$
$$\leq G^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m)$$
$$\times G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0). \qquad (4.14)$$

If we divide both sides of the inequality (4.14) by a normalization factor, we find that the corresponding inequality also holds for normalized coherence functions:

$$|g^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)|^2$$
$$\leq g^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m)$$
$$\times g^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0). \qquad (4.15)$$

For fields having $(2m)$th-order coherence, both sides of (4.15) are equal to unity, so that relation (4.14) reduces to an equality. In such cases Eq. (4.13) must

[17] L. Mandel and E. Wolf, Ref. 8, p. 246.

[18] Inequality (3.14) of Ref. 5.

be satisfied, and hence we can write

$$\hat{\rho}\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)\hat{A}^{(-)}(x_{m+1}) = \frac{G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} \; \hat{\rho}\{\hat{A}^{(-)}(x_0)\}^m \hat{A}^{(-)}(x_{m+1}). \quad (4.16)$$

Since $\hat{A}^{(-)}(x_0)$ and $\hat{A}^{(-)}(x_{m+1})$ commute, we can use (4.13) once again and obtain the relation

$$\hat{\rho}\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_{m+1}) = \frac{G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)G^{(m,m)}(x_{m+1}, x_0, \cdots, x_0; x_0, \cdots, x_0)}{\{G^{(m,m)}(x_0, \cdots, x_0 ; x_0, \cdots, x_0)\}^2} \; \hat{\rho}\{\hat{A}^{(-)}(x_0)\}^{m+1}.$$

$$(4.17)$$

Now since the left-hand side of (4.17) is symmetric in $x_1, x_2, \cdots, x_{m+1}$, we can rewrite (4.17) in the form

$$\hat{\rho}\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_{m+1})$$
$$= \prod_{j=1}^{m+1} \frac{G^{(m,m)}(x_j, x_0, \cdots, x_0; x_0, \cdots, x_0)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} \; \hat{\rho}\{\hat{A}^{(-)}(x_0)\}^{m+1}.$$

$$(4.18)$$

We deduce from (4.17) and (4.18) that if[19]

$$\hat{\rho}\{\hat{A}^{(-)}(x_0)\}^{m+1} \neq 0,$$

$$G^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)$$
$$= \frac{\prod_{j=1}^{m} G^{(m,m)}(x_j, x_0, \cdots, x_0; x_0, \cdots, x_0)}{[G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)]^{m-1}}. \quad (4.19)$$

Using the defining equation (4.1), we can rewrite (4.19) in terms of the normalized coherence functions. We thus obtain the relation

$$g^{(m,m)}(x_1, \cdots, x_m; x_0, \cdots, x_0)$$
$$= \prod_{j=1}^{m} g^{(m,m)}(x_j, x_0, \cdots, x_0; x_0, \cdots, x_0). \quad (4.20)$$

It now follows immediately from (4.11) and (4.20) that the function $f(x_1, \cdots, x_m)$ appearing in (4.11) must be expressible in the form

$$f(x_1, \cdots, x_m) = \sum_{j=1}^{m} f_1(x_j), \quad (4.21)$$

where $f_1(x_j)$ depends only on the variable x_j. Equations (4.21), (4.11), and (4.1) show that if we

have a complete $(2m)$th-order coherence, we can write $G^{(m,m)}$ in the form

$$G^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= U_m^*(x_1) \cdots U_m^*(x_m)U_m(x_1') \cdots U_m(x_m'), \quad (4.22)$$

where

$$U_m(x) = \{G^{(m,m)}(x, \cdots, x; x, \cdots x)\}^{1/2m}e^{if_1(x)}.$$

$$(4.23)$$

We therefore conclude that for quantized optical fields, also whether or not ϕ is nonnegative definite, the result expressed by Theorem 3 below holds.

Theorem 3: The necessary and sufficient condition for $(2m)$th-order coherence of the quantized optical field, i.e., for the validity of the identity $|g^{(m,m)}| \equiv 1$, is that $G^{(m,m)}$ has the factorized form

$$G^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= U_m^*(x_1) \cdots U_m^*(x_m)U_m(x_1') \cdots U_m(x_m').$$

The singular case of fields having just n photon is excluded.

Let us now obtain a result similar to that expressed by Theorem 2. Using (4.18) and its Hermitian adjoint, we can express

$$G^{(m+1,m+1)}(x_1, \cdots, x_{m+1}; x_1', \cdots, x_{m+1}')$$

in the form

$$G^{(m+1,m+1)}(x_1, \cdots, x_{m+1}; x_1', \cdots, x_{m+1}') = G^{(m+1,m+1)}(x_0, \cdots, x_0; x_0, \cdots, x_0)$$
$$\times \prod_{j=1}^{m+1} \frac{G^{(m,m)}(x_j, x_0, \cdots, x_0; x_0, \cdots, x_0)G^{(m,m)*}(x_j', x_0, \cdots, x_0; x_0, \cdots, x_0)}{\{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)\}^2}$$
$$= \prod_{j=1}^{m+1} U_{m+1}^*(x_j)U_{m+1}(x_j'), \quad (4.24)$$

where $U_{m+1}^*(x_j) = e^{i\alpha(x_0)} \dfrac{G^{(m,m)}(x_j, x_0, \cdots, x_0; x_0, \cdots, x_0)}{G^{(m,m)}(x_0, \cdots, x_0; x_0, \cdots, x_0)} \{G^{(m+1,m+1)}(x_0, \cdots, x_0; x_0, \cdots, x_0)\}^{1/(2m+2)},$

$$(4.25)$$

and $\alpha(x_0)$ is some real function of x_0 to be determined. Making use of the factorization property (4.22) in

---

[19] We exclude the singular case $\hat{\rho}\{\hat{A}^{(-)}(x_0)\}^{m+1} = 0$, which corresponds to fields having precisely m photons.

(4.25) and taking complex conjugates, we obtain the following expression for U_{m+1}:

$$U_{m+1}(x) = \left[\frac{\{G^{(m+1,m+1)}(x_0,\cdots,x_0;x_0,\cdots,x_0)\}^{1/(2m+2)}}{U_m(x_0)} \times e^{-i\alpha(x_0)} \right] U_m(x). \quad (4.26)$$

Since the coefficient on the right-hand side of (4.26) should be independent of x_0, $[-\alpha(x_0)]$ is just the phase of $U_m(x_0)$, so that using (4.24), we deduce from (4.26) that

$$U_{m+1}(x) = \left| \frac{U_{m+1}(x_0)}{U_m(x_0)} \right| U_m(x). \quad (4.27)$$

Equation (4.24) clearly shows that $(2m)$th-order coherence implies $2(m + 1)$th-order coherence, except for the singular case when there are just m photons present in the field. In this case $G^{(m+1,m+1)} \equiv 0$, so that $g^{(m+1,m+1)}$ has the indeterminate form zero divided by zero. By induction, we can extend the argument for any $n \geq m$. We may summarize the result which we have just obtained in Theorem 4.

Theorem 4: Coherence of the quantized field to any even order $2m$ implies coherence to all even orders $2n$ where $n \geq m$, except for the singular case of fields having just n_0 (some finite number $> m$) photons. In this singular case $(2m)$th-order coherence implies $(2n)$th-order coherence, where $m < n \leq n_0$.

The result obtained above should be compared with the corresponding result for fields which Glauber calls coherent to order N. In his definition the fields

for which

$$|g^{(n)}(x_1,\cdots,x_n;x_1',\cdots,x_n')| = 1$$

for *all* $n = 1, 2, \cdots, N$,

where $g^{(n)}$ is the normalized coherence function defined according to (1.6), are called coherent to order N. Thus coherence to any given order in his definition *automatically* implies coherence to all *lower* orders. The situation is quite the opposite in our case. According to the present definition, coherence to any even order can be shown to imply coherence to all *higher* even orders. There is obviously no *a priori* reason to prefer one definition to the other on account of these two different consequences alone. The main reason for the preference for our definition to that of Glauber's is that,in our approach, the normalized coherence functions are bounded by the values 0 and 1 for most of the cases of practical interest, and in those cases for which they exceed unity they correspond to truly quantum-mechanical features of the field. It may also be noted that in the limiting case when all the $2m$ space–time points coincide, the normalized $(2m)$th-order coherence function equals unity [cf. Eqs. (3.4) and (4.3)]. We must emphasize that our approach is not intended to classify "orders of coherence." It is mainly intended to classify fields which have a certain degree of coherence and also to compare the degrees of coherence of two fields, i.e., to give a criterion to decide which one is "more coherent."

It is worthwhile to mention two other possible generalizations of (1.1). If in the defining equation (4.1) we take the normalization constant simply as the product of mth and nth moments of intensity, rather than the normally ordered moments, we obtain

$$\bar{g}^{(m,n)}(x_1,\cdots,x_m;x_1',\cdots,x_n') = \frac{G^{(m,n)}(x_1,\cdots,x_m;x_1',\cdots,x_n')}{\left\{\prod_{j=1}^{m} \langle [\hat{A}^{(-)}(x_j)\hat{A}^{(+)}(x_j)]^m \rangle\right\}^{1/2m} \left\{\prod_{j=1}^{n} \langle [\hat{A}^{(-)}(x_j')\hat{A}^{(+)}(x_j')]^n \rangle\right\}^{1/2n}}. \quad (4.28)$$

It is conjectured that the inequalities

$$0 \leq |\bar{g}^{(m,n)}| < 1, \quad (4.29)$$

and

$$|\bar{g}^{(m,n)}| < |g^{(m,n)}| \quad (4.30)$$

hold for arbitrary fields. However, most of the useful properties such as Eq. (4.3) or the factorization theorem, etc., obeyed by $g^{(m,n)}$ are not obeyed by $\bar{g}^{(m,n)}$. We have therefore not considered this normalization in detail in our present discussion.

Another generalization of (1.1), recently discussed by Sudarshan,[20] is

$$s^{(m,n)}(x_1,\cdots,x_m;x_1',\cdots,x_n') = \frac{G^{(m,n)}(x_1,\cdots,x_m;x_1',\cdots,x_n')}{\{G^{(m,m)}(x_1,\cdots,x_m;x_1,\cdots,x_m)G^{(n,n)}(x_1',\cdots,x_n';x_1',\cdots,x_n')\}^{\frac{1}{2}}}. \quad (4.31)$$

He calls $s^{(m,n)}$ the "coherence index" of order (m, n). Such a normalization and some of the properties of

[20] E. C. G. Sudarshan (to be published).

the corresponding normalized coherence functions were first noted by the present author.[21] In particular, it was shown in the above mentioned reference that for arbitrary radiation fields the inequality

$$0 \leq |s^{(m,n)}| \leq 1 \qquad (4.32)$$

holds. From the general arguments given earlier in this section, it can be shown that the limiting case $|s^{(m,n)}| \equiv 1$ implies the factorization properties of $G^{(m,m)}$ and $G^{(n,n)}$ (except in the singular case when there are precisely n photons present in the field and $m = n$). Thus $|s^{(m,n)}| \equiv 1$ implies also $|g^{(m,m)}| = |g^{(n,n)}| \equiv 1$. For explicit proof and other related properties, we refer the reader to Sudarshan's paper. Although this normalization has the advantage that $|s^{(m,n)}|$ is always bounded by the values 0 and 1, it also leads to the relation

$$s^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m) \equiv 1, \quad (4.33)$$

which holds whether or not the space–time points $x_1, \cdots, x_m$ coincide. Hence the coherence indices defined in this manner are not quantities suitable for comparing the degrees of coherence of different fields.

5. DEGREE OF HIGHER-ORDER COHERENCE FOR GAUSSIAN FIELDS

In this section we derive an expression for the degree of coherence of an optical field that is governed by Gaussian probability distribution. Examples of such fields are blackbody radiation and the fields generated by thermal sources. For such fields the higher-order coherence functions are expressible in terms of second order ones[22] by means of the formulas

$$\Gamma^{(m,n)} \equiv 0, \quad m \neq n, \qquad (5.1a)$$

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= \sum_\pi \Gamma^{(1,1)}(x_1, x_p') \, \Gamma^{(1,1)}(x_2, x_q') \cdots \Gamma^{(1,1)}(x_m; x_r'), \qquad (5.1b)$$

where $\sum_\pi$ denotes summation over all $m!$ permutations $p, q, \cdots, r$ of $1, 2, \cdots, m$. In particular when $x_1 = x_1' = x_2 = x_2' = \cdots = x$, we obtain

$$\Gamma^{(m,m)}(x, \cdots, x; x, \cdots, x) = m! \, \{\Gamma^{(1,1)}(x; x)\}^m. \qquad (5.2)$$

[21] C. L. Mehta, Ph.D. thesis, University of Rochester, 1964.

[22] C. L. Mehta, in *Lectures in Theoretical Physics*, W. E. Brittin, Ed. (University of Colorado Press, Boulder, Colorado, 1965), Vol. VIIc, p. 398; see also I. S. Reed, IRE Trans. Inform. Theory **IT-8**, 194 (1962).

From Eqs. (3.1), (5.1), and (5.2) it then follows that

$$\gamma^{(m,n)} \equiv 0, \quad m \neq n, \qquad (5.3a)$$

$$\gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= \frac{1}{m!} \sum_\pi \gamma^{(1,1)}(x_1, x_p') \gamma^{(1,1)}(x_2, x_q') \cdots \gamma^{(1,1)}(x_m; x_r'). \qquad (5.3b)$$

For the special case of complete coherence we have, according to Theorem 2,

$$\Gamma^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m') = \prod_{j=1}^m U_m^*(x_j) U_m(x_j'). \qquad (5.4)$$

From Eqs. (5.2) and (5.4) it then follows that

$$U_m(x) = (m!)^{1/2m} U_1(x). \qquad (5.5)$$

By Gaussian fields in the quantum mechanical case we understand those fields for which the diagonal representation of the density operator $\phi(\{v\})$ is a multivariate Gaussian distribution. Since in this case $\phi(\{v\})$ is positive definite, the results obtained in this section hold also if the classical coherence functions $\Gamma^{(m,n)}$ and $\gamma^{(m,n)}$ are replaced by the corresponding quantum coherence functions $G^{(m,n)}$ and $g^{(m,n)}$, respectively.

6. SUMMARY AND CONCLUDING REMARKS

We have seen in Sec. 3 that the normalized (m, n)th-order coherence function $\gamma^{(m,n)}$ of the classical field, defined by (3.1), has most of the desirable properties of the "degree of coherence." In particular, it has the following properties:

(1) For all pairs of positive integers m, n, $|\gamma^{(m,n)}| \leq 1$.

(2) A completely coherent field of any even order satisfies the factorization Theorem 1.

(3) Complete coherence to any even order $2m$ implies coherence to all even orders (Theorem 2).

(4) When all the space–time points coincide, we have for all positive integral values of m

$$\gamma^{(m,m)}(x, \cdots, x; x, \cdots, x) = 1. \qquad (6.1)$$

The corresponding normalized quantum coherence function $g^{(m,n)}$ as defined by (4.1) also satisfies all of the above properties in the special case of fields having nonnegative definite diagonal representation. However, in general, $g^{(m,n)}$ does not satisfy all of the above properties. In fact we have:

(1) $|g^{(1,1)}| \leq 1$ always, but, in general, if either m or n (or both) are greater than unity, $|g^{(m,n)}|$ may exceed unity.

(2) A completely coherent field of any even order $2m$ satisfies the factorization Theorem 3 (analogous to Theorem 1 for the classical case).

(3) Coherence to any even order $2m$ implies coherence to all even orders $2n$, where $n \geq m$ (Theorem 4).

(4) When all the space–time points coincide, we have for all positive integral values of m

$$g^{(m,m)}(x, \cdots, x; x, \cdots, x) = 1. \quad (6.2)$$

It is worth mentioning that when the average occupation number of photons in the field is very large (classical limit), it can be shown that the ordering of the operators $A^{(+)}(x)$ and $A^{(-)}(x)$ in defining the coherence function $G^{(m,n)}$ is not important. In this case the coherence function $G^{(m,n)}$ can be approximated by averages of antinormally ordered products:

$$G^{(m,n)}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$\approx \mathrm{Tr}\,\{\hat{\rho}\hat{A}^{(+)}(x_1') \cdots \hat{A}^{(+)}(x_n')\hat{A}^{(-)}(x_1) \cdots \hat{A}^{(-)}(x_m)\}.$$
$$(6.3)$$

It is well known that the phase–space distribution function in such cases is nonnegative definite.[23] Hence, in this limit,

$$0 \leq |g^{(m,n)}| \leq 1. \quad (6.4)$$

We remarked earlier in this paper that for simplicity we considered only scalar fields. A generalization to vector fields is straightforward. Thus the normalized (m, n)th-order coherence tensor for the classical vector fields may be defined as

$$\gamma^{(m,n)}_{i_1, \cdots, i_m; i_1', \cdots, i_n'}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$

$$= \frac{\Gamma^{(m,n)}_{i_1, \cdots, i_m; i_1', \cdots, i_n'}(x_1, \cdots, x_m; x_1', \cdots, x_n')}{\left\{\prod\limits_{j=1}^{m} \Gamma^{(m,m)}_{i_j, \cdots, i_j; i_j, \cdots, i_j}(x_j, \cdots, x_j; x_j, \cdots, x_j)\right\}^{1/2m} \left\{\prod\limits_{j=1}^{n} \Gamma^{(n,n)}_{i_j', \cdots, i_j'; i_j', \cdots, i_j'}(x_j', \cdots, x_j'; x_j', \cdots, x_j')\right\}^{1/2n}}. \quad (6.5)$$

Here

$$\Gamma^{(m,n)}_{i_1, \cdots, i_m; i_1', \cdots, i_n'}(x_1, \cdots, x_m; x_1', \cdots, x_n')$$
$$= \langle V_{i_1}^*(x_1) \cdots V_{i_m}^*(x_m) V_{i_1'}(x_1') \cdots V_{i_n'}(x_n') \rangle \quad (6.6)$$

is the (m, n)th-order coherence tensor (see Ref. 8, p. 244). The subscripts $i_1, \cdots, i_m; i_1', \cdots, i_n'$ label the Cartesian components. The function

$$\Gamma^{(m,m)}_{i_j, \cdots, i_j; i_j, \cdots, i_j}(x_j, \cdots, x_j; x_j, \cdots, x_j)$$
$$= \langle \{V_{i_j}^*(x_j) V_{i_j}(x_j)\}^m \rangle$$

is the mth moment of intensity associated with the i_j component of the field amplitude.

The corresponding normalized coherence tensors for the quantized vector fields are obtained by replacing γ and Γ in (6.5) by g and G, respectively.

It may readily be seen that these normalized coherence tensors also satisfy properties analogous to those for the case of normalized scalar coherence functions.

ACKNOWLEDGMENT

APPENDIX

In this Appendix we derive the nonnegative definiteness condition (4.8) and discuss some of its consequences which were used in the text.

Let

$$\hat{F} = \sum_{i=1}^{N} \beta_i \hat{A}^{(+)}(x_1^{(i)}) \hat{A}^{(+)}(x_2^{(i)}) \cdots \hat{A}^{(+)}(x_m^{(i)}), \quad (A1)$$

where $\hat{A}^{(+)}(x)$ is the positive frequency part of the field operator, N is any positive integer, β_i ($i = $

$1, 2, \cdots, N$) are arbitrary complex parameters, and for each i, $x_1^{(i)}, x_2^{(i)}, \cdots, x_m^{(i)}$ is a set of m arbitrary space–time points. Since the density operator $\hat{\rho}$ is nonnegative definite, we have

$$0 \leq \mathrm{Tr}\,(\hat{\rho}\hat{F}^\dagger \hat{F})$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_i^* \, \mathrm{Tr}\,\{\hat{\rho}\hat{A}^{(-)}(x_1^{(i)}) \cdots \hat{A}^{(-)}(x_m^{(i)})$$
$$\times \hat{A}^{(+)}(x_1^{(j)}) \cdots \hat{A}^{(+)}(x_m^{(j)})\}\beta_j,$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_i^* G^{(m,m)}(x_1^{(i)}, \cdots, x_m^{(i)}; x_1^{(j)}, \cdots, x_m^{(j)})\beta_j.$$
$$(A2)$$

Relations (A2) and (4.1) then give the required inequality (4.8), viz.,

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i^* g^{(m,m)}(x_1^{(i)}, \cdots, x_m^{(i)}; x_1^{(j)}, \cdots, x_m^{(j)})\alpha_j \geq 0,$$
$$(A3)$$

where

$$\alpha_j = \left\{\prod_{l=1}^{m} G^{(m,m)}(x_l^{(j)}, \cdots, x_l^{(j)}; x_l^{(j)}, \cdots, x_l^{(j)})\right\}^{1/2m} \beta_j.$$

Since the parameters β_j in (A1) or (A2) are arbitrary, so are the parameters α_j in (A3).

Let us now discuss some consequences of (A3). For $N = 1$ and $N = 2$, the inequality (A3) expresses the obvious relations

$$g^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m) \geq 0,$$

[23] C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. 138, B274 (1965).

$g^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m)$

$\qquad \times g^{(m,m)}(x_1', \cdots, x_m'; x_1', \cdots, x_m')$

$\qquad \geq |g^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')|^2,$ \qquad (A4)

which are similar to the inequalities (3.12) and (3.14) of Ref. 5. For $N = 3$, the inequality (A3) expresses that, in addition, the determinant

$$\begin{vmatrix} g^{(m,m)}(x^{(1)}; x^{(1)}) & g^{(m,m)}(x^{(1)}; x^{(2)}) & g^{(m,m)}(x^{(1)}; x^{(3)}) \\ g^{(m,m)}(x^{(2)}; x^{(1)}) & g^{(m,m)}(x^{(2)}; x^{(2)}) & g^{(m,m)}(x^{(2)}; x^{(3)}) \\ g^{(m,m)}(x^{(3)}; x^{(1)}) & g^{(m,m)}(x^{(3)}; x^{(2)}) & g^{(m,m)}(x^{(3)}; x^{(3)}) \end{vmatrix}$$

$$\geq 0, \quad (A5)$$

where $x^{(j)}$ now stands for the set of variables $x_1^{(j)}$, $x_2^{(j)}, \cdots, x_m^{(j)}$. In the case of complete $(2m)$th-order coherence, we have $|g^{(m,m)}| = 1$, and hence we can write

$$g^{(m,m)}(x^{(i)}; x^{(j)}) = \{g^{(m,m)}(x^{(j)}; x^{(i)})\}^*,$$

$$= \exp\left[-i\psi(x^{(i)}, x^{(j)})\right], \quad (A6)$$

where ψ is real. It can then be seen, on evaluating the determinant, that (A5) can only be satisfied as

an equality, and we then obtain the relation

$$\psi(x^{(3)}, x^{(2)}) = \psi(x^{(3)}, x^{(1)}) - \psi(x^{(2)}, x^{(1)}). \quad (A7)$$

Setting $x^{(1)} = 0$ (i.e., $x_1^{(1)} = x_2^{(1)} = \cdots = x_m^{(1)} = 0$, which is always permissible by suitable choice of the origin), we finally obtain

$$\psi(x^{(3)}, x^{(2)}) = f(x^{(3)}) - f(x^{(2)}), \quad (A8)$$

where $f(x^{(j)}) = \psi(x^{(j)}, 0)$ is a function of $x^{(j)}$ only. Hence, from (A6) we find that

$$g^{(m,m)}(x^{(i)}; x^{(j)}) = \exp\{-i[f(x^{(i)}) - f(x^{(j)})]\}, \quad (A9)$$

which, when written in full, gives the required relation (4.9): viz.,

$g^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$

$\qquad = \exp\{-i[f(x_1, \cdots, x_m) - f(x_1', \ldots, x_m')]\}.$

$$(A10)$$

It may be noted that the results obtained in this appendix are also true of classical fields and can be obtained in a strictly similar manner.

Energy-Momentum Conservation Implies Translation Invariance: Some Didactic Remarks

DAVID N. WILLIAMS*

Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP n° 2,91, Gif-sur-Yvette, France

We discuss from a rigorous viewpoint two more-or-less familiar cases where energy-momentum conservation implies invariance under space-time translations. First, if a closed linear operator on a Hilbert space has a domain that is invariant under spectral projections belonging to the four-momentum operators, and if it "conserves energy-momentum," it necessarily commutes with the appropriate representation of the translations. (Bounded operators, such as the S matrix, are a special case.) At least for separable spaces, the domain restriction characterizes the closed operators for which the theorem is true. Second, if a bounded bilinear form between momentum states of m and n particles in a Fock space (or more generally, a bounded multilinear form) conserves energy momentum, the corresponding tempered distribution has a conservation delta function at points where the mass shell is a C_∞ manifold; but no derivatives of delta functions can occur. In this connection, we are led to a result that seems to be new: the cluster parameters ("connected amplitudes") of a family of bounded bilinear forms, labeled by (m, n), are also bounded bilinear forms. The two systems, of course, mutually conserve energy momentum.

I. INTRODUCTION

THAT translation invariance implies momentum conservation is a familiar example of the relation between continuous symmetries and conservation

laws, which is classically and elegantly expressed by Noether's theorem.[1] Conversely, to every constant of the motion corresponds the infinitesimal generator of an invariance group of the Hamiltonian or

---

* Present address: The University of Michigan, Ann Arbor, Michigan.

[1] E. Noether, Nachr. Akad. Wiss. Goettingen, II. Math. Physik. Kl., 235 (1918).

$$g^{(m,m)}(x_1, \cdots, x_m; x_1, \cdots, x_m)$$
$$\times\, g^{(m,m)}(x_1', \cdots, x_m'; x_1', \cdots, x_m')$$
$$\geq |g^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')|^2, \quad \text{(A4)}$$

which are similar to the inequalities (3.12) and (3.14) of Ref. 5. For $N = 3$, the inequality (A3) expresses that, in addition, the determinant

$$\begin{vmatrix} g^{(m,m)}(x^{(1)}; x^{(1)}) & g^{(m,m)}(x^{(1)}; x^{(2)}) & g^{(m,m)}(x^{(1)}; x^{(3)}) \\ g^{(m,m)}(x^{(2)}; x^{(1)}) & g^{(m,m)}(x^{(2)}; x^{(2)}) & g^{(m,m)}(x^{(2)}; x^{(3)}) \\ g^{(m,m)}(x^{(3)}; x^{(1)}) & g^{(m,m)}(x^{(3)}; x^{(2)}) & g^{(m,m)}(x^{(3)}; x^{(3)}) \end{vmatrix}$$
$$\geq 0, \quad \text{(A5)}$$

where $x^{(j)}$ now stands for the set of variables $x_1^{(j)}$, $x_2^{(j)}, \cdots, x_m^{(j)}$. In the case of complete $(2m)$th-order coherence, we have $|g^{(m,m)}| = 1$, and hence we can write

$$g^{(m,m)}(x^{(i)}; x^{(j)}) = \{g^{(m,m)}(x^{(j)}; x^{(i)})\}^*,$$
$$= \exp[-i\psi(x^{(i)}, x^{(j)})], \quad \text{(A6)}$$

where ψ is real. It can then be seen, on evaluating the determinant, that (A5) can only be satisfied as an equality, and we then obtain the relation

$$\psi(x^{(3)}, x^{(2)}) = \psi(x^{(3)}, x^{(1)}) - \psi(x^{(2)}, x^{(1)}). \quad \text{(A7)}$$

Setting $x^{(1)} = 0$ (i.e., $x_1^{(1)} = x_2^{(1)} = \cdots = x_m^{(1)} = 0$, which is always permissible by suitable choice of the origin), we finally obtain

$$\psi(x^{(3)}, x^{(2)}) = f(x^{(3)}) - f(x^{(2)}), \quad \text{(A8)}$$

where $f(x^{(j)}) = \psi(x^{(j)}, 0)$ is a function of $x^{(j)}$ only. Hence, from (A6) we find that

$$g^{(m,m)}(x^{(i)}; x^{(j)}) = \exp\{-i[f(x^{(i)}) - f(x^{(j)})]\}, \quad \text{(A9)}$$

which, when written in full, gives the required relation (4.9): viz.,

$$g^{(m,m)}(x_1, \cdots, x_m; x_1', \cdots, x_m')$$
$$= \exp\{-i[f(x_1, \cdots, x_m) - f(x_1', \ldots, x_m')]\}. \quad \text{(A10)}$$

It may be noted that the results obtained in this appendix are also true of classical fields and can be obtained in a strictly similar manner.

Energy-Momentum Conservation Implies Translation Invariance: Some Didactic Remarks

DAVID N. WILLIAMS*
Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP n° 2,91, Gif-sur-Yvette, France

We discuss from a rigorous viewpoint two more-or-less familiar cases where energy-momentum conservation implies invariance under space-time translations. First, if a closed linear operator on a Hilbert space has a domain that is invariant under spectral projections belonging to the four-momentum operators, and if it "conserves energy-momentum," it necessarily commutes with the appropriate representation of the translations. (Bounded operators, such as the S matrix, are a special case.) At least for separable spaces, the domain restriction characterizes the closed operators for which the theorem is true. Second, if a bounded bilinear form between momentum states of m and n particles in a Fock space (or more generally, a bounded multilinear form) conserves energy momentum, the corresponding tempered distribution has a conservation delta function at points where the mass shell is a C_∞ manifold; but no derivatives of delta functions can occur. In this connection, we are led to a result that seems to be new: the cluster parameters ("connected amplitudes") of a family of bounded bilinear forms, labeled by (m, n), are also bounded bilinear forms. The two systems, of course, mutually conserve energy momentum.

I. INTRODUCTION

THAT translation invariance implies momentum conservation is a familiar example of the relation between continuous symmetries and conservation laws, which is classically and elegantly expressed by Noether's theorem.[1] Conversely, to every constant of the motion corresponds the infinitesimal generator of an invariance group of the Hamiltonian or

---

* Present address: The University of Michigan, Ann Arbor, Michigan.

[1] E. Noether, Nachr. Akad. Wiss. Goettingen, II. Math. Physik. Kl., 235 (1918).

Lagrangian[2]; and for quantum theory in general, we have the following (presumably) well-known formal argument. Let $|p'\rangle$ be an "eigenstate" of the total four-momentum operators P_μ, with eigenvalues p'_μ. Let A be a linear operator that conserves energy momentum, i.e.,

$$\langle p'' | A | p' \rangle = 0 \quad \text{for} \quad p'' \neq p'.$$

In other words, $A|p'\rangle$ is an eigenstate of P with the same eigenvalue p', so it follows that

$$[A, P_\mu] = 0.$$

Hence A is invariant under translation by any space-time four-vector b:

$$[A, e^{iP\cdot b}] = 0.$$

We see in Sec. II that it is a simple exercise to make this argument rigorous under the conditions stated in the abstract, including the case where A is a bounded operator, such as the S matrix.[3]

Actually, the question to what extent the converse of the energy-momentum conservation theorem is true has some relevance for elementary particle physics, where experimental statements are commonly statements about momentum space, involving only macroscopic space-time localization. In this situation the conservation law is verified more directly than the invariance principle.

Some theorists have argued that this matter of practice should be given the status of a matter of principle.[4] Either they deny the operational significance of microscopic space-time[5] or for some more conservative reason they propose to base the theory of strong interactions on momentum space and to treat space-time as a derived concept.[6] Of course, if one advocates this view, he is not thereby prevented from postulating translation invariance, since microscopic displacements could conceivably have a sense,

even if microscopic space-time does not; but it seems more in the spirit of things for those who take the S matrix as the fundamental observable quantity to postulate instead the conservation law.[7]

Whether for reasons of practice or principle, we think there is at least a pedagogical value in spelling out some of the contexts in which energy-momentum conservation implies translation invariance, with the most direct applications being to S-matrix theory. Because the results are to some extent known, and because the proofs as well have very likely occurred to those who have wondered about the question with enough mathematical curiosity, we make no particular claim of originality for our rather straightforward discussion. On the other hand, a rigorous treatment does lead us indirectly to a potentially useful piece of information about the S matrix which is new, as far as we know. Namely, the connected S-matrix elements in momentum space (cluster amplitudes) are not only tempered distributions but kernels of bounded operators.

In Sec. II, we use the spectral theory to formulate the property of energy-momentum conservation for operators on a Hilbert space, and for a certain class of operators we transform the formal argument already given into a proof of the theorem on translation invariance. We discuss to what extent the conditions imposed characterize the operators for which the theorem is true.

In Secs. III and IV, we reformulate and prove the theorem by a different method, for bounded multilinear forms on Cartesian products $(\mathfrak{IC}_{m_1}, \cdots, \mathfrak{IC}_{m_r}, \mathfrak{IC}_{n_1}, \cdots, \mathfrak{IC}_{n_s})$, where $\mathfrak{IC}_m$ is the m-particle subspace of a Fock space. By "multilinear" we mean antilinear on each space $\mathfrak{IC}_{m_i}$ and linear on each $\mathfrak{IC}_{n_j}$. Such forms may correspond to operators between the spaces $\mathfrak{IC}_m$ and $\mathfrak{IC}_n$, where $m = \sum m_i$ and $n = \sum n_j$, but in general they do not. Whether such a general situation has a practical application, we do not know, but the generality costs nothing extra. The second proof deals directly with transition amplitudes in momentum space (tempered distributions), and the idea is to show that energy-momentum conservation is expressed only by delta functions in the transition amplitudes, and not by derivatives of delta functions. This leads at once to translation invariance. We are careful not to write delta functions at points where the mass shell is not a differentiable manifold, because they are not well defined at such points.

In Sec. V we mention that the result extends to the cluster parameters for momentum space amplitudes.

[2] A. Messiah, *Mécanique Quantique* (Dunod Cie., Paris, 1960), Vol. II, Chap. XV.

[3] Although we should not be surprised to learn that the argument in question is known, we have not succeeded in finding it in the literature. For the case of the S matrix, H. P. Stapp [Phys. Rev. **125**, 2139 (1962)] mentions without proof that translation invariance and energy-momentum conservation are equivalent. For bounded operators, the exercise is indeed not only simple but trivial, given the standard results of the spectral theory.

[4] E.g., G. F. Chew, Sci. Progr. (G.B.), **51**, 529 (1963); H. P. Stapp in Ref. 3; E. Lubkin, Nuovo Cimento **32**, 171 (1964).

[5] This strikes us as a radical view because we are not able to imagine all possible theories by means of which the concept could acquire an operational meaning. We do not intend by that a value judgement on the plausibility of theories motivated by such a view.

[6] There have been several interesting attempts in this direction, based on the S matrix. Among them we mention M. L. Goldberger and K. M. Watson, Phys. Rev. **127**, 2284 (1962); M. Froissart, M. L. Goldberger, and K. M. Watson, *ibid.* **131**, 2820 (1963); H. P. Stapp, *ibid.* **139**, B257 (1965); A. Peres, Ann. Phys. (N.Y.) **37**, 179 (1966). The last of these contains a more complete list of references.

[7] This is, for example, the attitude of Stapp. See Ref. 3.

Although this is a trivial fact, we again follow a "didactic" route in an attempt to clarify in what sense it is true. We apply some elementary theorems on Hilbert–Schmidt operators to find that the cluster amplitudes corresponding to a family of bounded *bilinear* forms are themselves kernels of bounded operators between the m- and n- particle Hilbert spaces, which conserve energy momentum if the original amplitudes do.

Finally, in an appendix, we prove that, on a separable Hilbert space, a closed operator commutes with all spectral projections if and only if it commutes with the translations. (The "only if" part is valid for nonseparable spaces as well.) This result is probably known to mathematicians, since it is only a slight generalization of the theorem for bounded operators, but neither the theorem nor its proof seems to be readily accessible to nonspecialists (such as the author).

II. FORMULATION, THEOREM, AND PROOF

What do we mean when we say that an operator conserves energy-momentum? We give ourselves a Hilbert space $\mathcal{H}$ and commuting self-adjoint energy-momentum operators P_μ, $\mu = 0, 1, 2, 3$, defined on a common dense submanifold of $\mathcal{H}$. That a linear operator on $\mathcal{H}$ conserves energy-momentum means at least that its matrix elements do not connect subspaces of $\mathcal{H}$ belonging to disjoint subsets of the spectrum of P_μ.

In other words, let

$$P_\mu = \int_{\mathbf{R}^4} p_\mu \, dE(p)$$

be the simultaneous spectral decomposition of P_μ, where $dE(p)$ is the spectral measure, with support on the spectrum of P_μ.[8] For any Borel set $\Delta \subset \mathbf{R}^4$, consider the projection operator[9]

$$E(\Delta) = \int_\Delta dE(p).$$

The subspace of $\mathcal{H}$ belonging to the part of the spectrum of P_μ contained in Δ is $E(\Delta)\mathcal{H} \equiv \mathcal{H}(\Delta)$. Let A be a linear operator on $\mathcal{H}$ with domain $D(A)$, which we may assume to be dense or not, as we please. Then we say that A conserves energy–momentum if,

whatever be the Borel set Δ or $f \in D(A)$, the condition

$$E(\Delta)f = 0$$

implies that

$$E(\Delta)Af = 0.$$

Thus, if $f \in \mathcal{H}(\Delta') \cap D(A)$ and $g \in \mathcal{H}(\Delta)$, with Δ' and Δ disjoint, we have the minimum requirement just mentioned:

$$\langle g, Af \rangle = 0.$$

This equation is equivalent to the definition; for if $f \in D(A)$ and $E(\Delta)f = 0$, it follows that $f \in \mathcal{H}(\mathbf{R}^4 - \Delta)$, where $\mathbf{R}^4 - \Delta$ is the complement of Δ. Then for any $g \in \mathcal{H}$

$$\langle g, E(\Delta)Af \rangle = \langle E(\Delta)g, Af \rangle = 0;$$

hence

$$E(\Delta)Af = 0.$$

We aim to study under what conditions the fact that A conserves energy momentum implies that it commutes with all spectral projections $E(\Delta)$, and hence with all translations

$$T(b) = \int e^{ip \cdot b} \, dE(p), \quad b \in \mathbf{R}^4.$$

To make sense out of such a statement, we have to know something about the domains of the operators that occur. Following Riesz and Sz.-Nagy,[10] we define the domain of a product $A_1 A_2$ to be the set of all vectors $f \in D(A_2)$ such that $A_2 f \in D(A_1)$. We write $A_1 \subseteq A_2$ if A_2 is an extension of A_1; i.e., $D(A_2) \supset D(A_1)$ and $A_2 f = A_1 f$ for $f \in D(A_1)$. We say that a bounded operator B defined on all of $\mathcal{H}$ commutes with A if $BA \subseteq AB$. We say that A is closed if, whenever both $f_n \in D(A)$ and Af_n are Cauchy sequences in the norm of $\mathcal{H}$, it follows that $\lim f_n = f \in D(A)$ and $\lim Af_n = Af$.

What we actually prove is the following theorem, which perhaps does not characterize the operators for which energy-momentum conservation and translation invariance are equivalent, but which probably comes close enough for practical purposes.

Theorem A: Let A be a closed linear operator on a separable Hilbert space $\mathcal{H}$. Then the following statements are equivalent:

(i) A conserves energy momentum, and $D(A)$ is invariant under spectral projections; i.e., $E(\Delta)D(A) \subset D(A)$ for all Δ;

(ii) $E(\Delta)A \subseteq AE(\Delta)$ for all Δ;

(iii) $T(b)A = AT(b)$ for all b.

If $\mathcal{H}$ is nonseparable, then we still have (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii).

---

[8] Very readable summaries on the "SNAG" theorem are given by R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964), pp. 91–93; R. Jost, *The General Theory of Quantized Fields* (American Mathematical Society, Providence, Rhode Island, 1965), pp. 16–17.

[9] Recall that the Borel sets of $\mathbf{R}^n$ are the smallest family of sets that contains all denumerable unions, intersections, and complements of open sets.

[10] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), Chap. VIII, Secs. 114–116.

The only nontrivial part of the proof is the relation between the statements (ii) and (iii). Because this result belongs properly to the functional calculus of self-adjoint operators, we take it for granted here and reserve the proof for the Appendix. Certainly its formal equivalent is a part of the folklore of quantum mechanics.

We complete the proof of Theorem A by showing the equivalence of (i) and (ii), without assuming that $\mathcal{JC}$ is separable (or even that A is closed). To prove that (i) implies (ii), note that, for $f \in D(A)$ and any Δ,

$$E(\Delta)AE(\mathbf{R}^4 - \Delta)f = 0,$$

from energy-momentum conservation. Because

$$E(\Delta) + E(\mathbf{R}^4 - \Delta) = 1,$$

we have

$$E(\Delta)Af = E(\Delta)AE(\Delta)f,$$
$$AE(\Delta)f = E(\Delta)AE(\Delta)f.$$

Hence $E(\Delta)A \subseteq AE(\Delta)$.

It only remains to show that (ii) implies (i). But that is trivial. First, $E(\Delta)D(A) \subset D(A)$, from the definition of the expression (ii). That A conserves energy momentum follows at once from (ii) and the definition of energy-momentum conservation. Thus, the theorem is proved.

We have not made any restrictions on the spectrum of P_μ. For closed operators and separable spaces, Theorem A says that it is not possible to relax the condition of the invariance of $D(A)$ under spectral projections. Whether the domain requirement is automatically implied in the case of closed operators by energy-momentum conservation as formulated here, we do not know; nor are we inclined to worry about it. The condition that A be closed, or at least have a closure fulfilling the other conditions, seems essential for the proof in the Appendix of the relation between (ii) and (iii); but we do not know whether it can be relaxed. We also do not know whether the statement (iii) $\Rightarrow$ (ii) is true for nonseparable spaces.

At any rate, the conditions of the theorem seem sufficiently general for most practical applications in physics.

III. ALTERNATIVE FORMULATION IN FOCK SPACE

Of course, nothing more has to be said in order to apply the theorem to a Fock space. But in that case, we have constructed another proof, for a certain class of operators and forms, which we think instructive. In the first proof, the nontrivial part was contained in the spectral theory. In the second, the basic mathematical tools are the nuclear theorem for tempered

distributions,[11] plus a theorem of Schwartz on the structure of a distribution with support on a submanifold of some $\mathbf{R}^n$.

For simplicity we put ourselves in the relativistic Fock space $\mathcal{F}$ corresponding to spinless particles with a single mass $M > 0$. The generalization of the discussion to Fock spaces with denumerable numbers of different types of particles with various spins and nonzero masses is trivial. Thus,

$$\mathcal{F} = \bigoplus_{n=0}^{\infty} \mathcal{JC}_n,$$

where for $n \geq 1$,

$$\mathcal{JC}_n = \mathrm{Sym}\left[L_2\left(\frac{d^3p_1}{\omega_1} \cdots \frac{d^3p_n}{\omega_n}, \mathbf{R}^{3n}\right)\right],$$

$$\omega_i \equiv \omega(\mathbf{p}_i) = (M^2 + \mathbf{p}_i^2)^{\frac{1}{2}} > 0,$$

is the symmetrized Hilbert space of momentum-space wavefunctions of n free particles.

Because each $\mathcal{JC}_n$, for $n \geq 1$, is identified with an L_2 space of functions, with a measure that "dominates" Lebesgue measure (and is dominated by it: the zero sets are the same), we can give a meaning to the "support" of a vector $f \in \mathcal{JC}_n$. Namely, let h be any element in the equivalence class of almost everywhere equal functions that corresponds to f; we write $h \in f$. Let supp h be the support of h in the usual sense, i.e., the complement of the largest open set of $\mathbf{R}^{3n}$ on which h vanishes. Then define

$$\mathrm{supp}\, f = \bigcap_{h \in f} \mathrm{supp}\, h.$$

If f can be represented by a continuous function $h \in f$, we have (exercise for the reader)

$$\mathrm{supp}\, f = \mathrm{supp}\, h.$$

First we consider bounded operators on $\mathcal{F}$. To each bounded linear operator B, and to each ordered pair of spaces $(\mathcal{JC}_m, \mathcal{JC}_n)$, we associate the bounded bilinear form

$$B_{mn}(f, g) = \langle f, Bg \rangle,$$

where $f \in \mathcal{JC}_m$ and $g \in \mathcal{JC}_n$. We say that B_{mn} conserves energy momentum if

$$B_{mn}(f, g) = 0$$

for all f and g having supports that nowhere satisfy the energy-momentum conservation equations. More explicitly, let $\mathbf{P} = (\mathbf{p}_1, \cdots, \mathbf{p}_m)$ and $\mathbf{Q} = (\mathbf{q}_1, \cdots, \mathbf{q}_n)$. Then B_{mn} vanishes if

$$(\mathbf{P}, \mathbf{Q}) \in \mathrm{supp}\, f \times \mathrm{supp}\, g$$

---

[11] L. Schwartz, *Théorie des Distributions* (Hermann et Cie., Paris, 1959), Vol. II, Chap. VII.

implies that, for at least one μ,

$$t^\mu(\mathbf{P}, \mathbf{Q}) \equiv \sum_{i=1}^{m} p_i^\mu - \sum_{j=1}^{n} q_j^\mu \neq 0,$$

where all four-vectors are on the positive sheet of the mass hyperboloid; e.g., $p_i^0 = \omega(\mathbf{p}_i)$. We say that B conserves energy-momentum if each B_{mn} does.[12]

It is not difficult to see that this definition is equivalent to the one given before (in the cases where it applies). It is perhaps worth remarking that, for the S matrix, the above statement of energy-momentum conservation for the transition amplitudes is equivalent to the analogous requirement on the observable transition probabilities,[13] as the reader can immediately see for himself.

In the next section, we prove again that if B conserves energy momentum, it commutes with the standard unitary representation of space-time translations defined on $\mathfrak{F}$. We do it by considering the tempered distributions $B_{mn}(\mathbf{P}, \mathbf{Q})$, defined by restricting $B_{mn}(f, g)$ to pairs of functions in the appropriately symmetrized Schwartz spaces[11] ($\mathcal{S}_m$, $\mathcal{S}_n$) of test functions which are C_∞ and decrease at infinity with all derivatives faster than any inverse polynomial. That we get a tempered distribution on the entire subspace of test functions in $\mathcal{S}(\mathbf{R}^{3(m+n)})$ that are symmetric in the first m and last n three-vectors follows from[14]:

(i) the fact that B_{mn} is a bounded bilinear form (after accounting for the antilinearity of the first factor)

$$|B_{mn}(f, g)| \leq C \|f\| \|g\|,$$

where $\|f\|$ indicates the scalar product norm in $\mathfrak{F}$;

(ii) the fact that the topology of $\mathcal{S}_m$ is finer than that induced from the strong topology of $\mathcal{K}_m$;

(iii) the "théorème nucléaire" of Schwartz.[15]

If B conserves energy momentum, the tempered distributions B_{mn} have their supports on the sets where

$$t^\mu(\mathbf{P}, \mathbf{Q}) = 0.$$

---

[12] It will generally be obvious how to take into account the case m or $n = 0$, corresponding to the vacuum with zero energy momentum, so we most often do not mention it explicitly.

[13] I am indebted to D. Iagolnitzer for drawing my attention to this point, as well as to the fact that translation invariance of the probabilities does not imply translation invariance of the amplitudes (although Poincaré invariance does).

[14] This simple proof occurred to several other people before it occurred to the author. In the literature, the question is raised indirectly by E. H. Wichmann and J. H. Crichton [Phys. Rev. **132**, 2788 (1963)], who give a lucid discussion of the cluster decomposition property which assumes that the S-matrix amplitudes are tempered distributions. K. Hepp [Helv. Phys. Acta **37**, 659 (1964)] states it as a fact, without giving the proof. We have found references to the proof in J. R. Taylor, Phys. Rev. **142**, 1236 (1966), and D. Iagolnitzer, "S-Matrix Theory and Phenomenological Space-Time Description," Saclay preprint (to be published).

[15] L. Gårding and J. L. Lions, Nuovo Cimento, Suppl. **14**, 9 (1959).

Our second method of proving the translation invariance of B is to show that, on a sufficiently large space of test functions, B_{mn} factorizes into a product of a delta function for energy-momentum conservation times a "tempered distribution" on the manifold defined by the conservation law. As mentioned in the Introduction, the essential point is to show that derivatives of delta functions cannot occur, because they conflict with the boundedness of B_{mn}, considered as a bilinear form.

Before passing to the theorem and proof, note that, as far as the discussion so far has been concerned, we have never used the fact that the bound C is the same for each bilinear form B_{mn}; we could just as well have a family of positive constants C_{mn} which could be unbounded for large (m, n), corresponding to a class of unbounded operators on $\mathfrak{F}$. (The "number of particles" operator is a simple example.) Actually, we never need to know that B_{mn} is a bounded *bilinear* form; we can do just as well with the weaker statement that it is a bounded *multilinear* form which satisfies

$$|B_{mn}(f_1 \cdots f_m, g_1 \cdots g_n)| \leq C_{mn} \prod_{i,j} \|f_i\| \|g_j\|,$$

with f_i, $g_j \in \mathcal{K}_1$. Such forms can correspond to a larger class of unbounded operators[16]; or, on the other hand, they might not correspond to operators at all, not even between $\mathcal{K}_n$ and $\mathcal{K}_m$.

In fact, the whole discussion goes through for bounded multilinear forms of the type

$$B_{m_1 \cdots m_r, n_1 \cdots n_s}(f_1, \cdots, f_r, g_1, \cdots, g_s),$$

where $f_i \in \mathcal{K}_{m_i}$ and $g_j \in \mathcal{K}_{n_j}$. Energy-momentum conservation is defined in the obvious way, and we still have the reduction to tempered distributions. Although we have in mind no particular situation where such generality might be useful, there is no reason not to state our result for such cases. *A priori*, as we see in Sec. V, we would have said that the cluster amplitudes are an example of bounded multilinear forms on Cartesian products of $\mathcal{K}_1$, if the Hilbert–Schmidt theorems did not tell us that they are really bounded bilinear forms.

IV. THEOREM FOR MULTILINEAR FORMS

The theorem below is stated for forms. If the forms come from closed operators on $\mathfrak{F}$, it extends immediately to the operators, by linearity and

---

[16] The domain specified is translation invariant, but not invariant under spectral projections. It follows from Theorem A, with Theorem B in Sec. IV, that if the operators conserve energy momentum, they are not closed on this domain.

continuity, modulo questions of domain. Certainly there is no problem for bounded operators.

Theorem B: Let $T(b)$ be the unitary representation of space–time translations on $\mathscr{F}$, defined on each $\mathscr{K}_m$ by

$$[T(b)f](\mathbf{P}) = \exp\left(-i\sum_{j=1}^{m} p_j \cdot b\right) f(\mathbf{P}).$$

If $B_{m_1 \cdots m_r, n_1 \cdots n_s}$ is a bounded multilinear form on $(\mathscr{K}_{m_1}, \cdots, \mathscr{K}_{n_s})$ which conserves energy momentum, then

$$B_{m_1 \cdots n_s}[T(b)f_1, \cdots, T(b)g_s] = B_{m_1 \cdots n_s}(f_1, \cdots, g_s).$$

To save writing, the proof is given in detail only for bounded bilinear forms B_{mn}. Very little modification is needed to extend it to multilinear forms, and it will hardly tax the reader to provide it himself.

Consider B_{mn} as a bounded linear transformation B_{mn}: $\mathscr{K}_n \to \mathscr{K}_m$. Because

$$T(b)B_{mn}T(b)^{-1} - B_{mn}$$

is also a bounded (i.e., continuous) linear transformation of $\mathscr{K}_n$ into $\mathscr{K}_m$, it suffices to prove

$$\langle f, T(b)B_{mn}T(b)^{-1}g \rangle = \langle f, B_{mn}g \rangle,$$

where f and g are arbitrary elements of two sets of vectors, each of which spans (by means of finite linear combinations) a dense submanifold of $\mathscr{K}_m$, $\mathscr{K}_n$ respectively. In particular, we always choose $(f, g) \in (\mathcal{S}_m, \mathcal{S}_n)$, with m and n running over the positive integers.

In order to avoid a possible difficulty about defining delta functions and their rth derivatives of the form

$$\prod_{\mu=0}^{3} \delta^{(r_\mu)}[t^\mu(\mathbf{P}, \mathbf{Q})],$$

(where r is a "four-vector" with nonnegative integers as components) at zeros of t where the Jacobian matrix has rank less than four, we make one further restriction on the support of one of the elements, say g, of (f, g). Namely, if $n \geq 2$, we demand that there shall be no $\mathbf{Q} \in \operatorname{supp} g$ for which all corresponding mass hyperboloid four-vectors q_i are collinear; at least two of the four-vectors are to be linearly independent. Hepp[17] has observed that even the smaller subspace of functions in $\mathcal{S}_n$ with supports having *no* two of the corresponding four-vectors collinear ("disjoint velocities") is dense in $\mathscr{K}_n$. The reader may easily verify that, with this restriction, the

Jacobian matrix evaluated for $t = 0$ indeed has rank four.

Now we consider the tempered distribution $B_{mn}(\mathbf{P}, \mathbf{Q})$, restricted to the open set Ω of points $(\mathbf{P}, \mathbf{Q})$ where $\mathbf{Q}$ satisfies the condition just mentioned. The support of B_{mn} in Ω is a C_∞ manifold, which we denote $\operatorname{supp}_\Omega B_{mn}$, of dimension 3 if $m = n = 1$, and of dimension $3(m + n) - 4$ if $m, n \geq 2$. In the latter case, $\operatorname{supp}_\Omega B_{mn}$ is the set of simultaneous zeros of the C_∞ functions $t^\mu(\mathbf{P}, \mathbf{Q})$, which forms a C_∞ manifold by the implicit function theorem.[18] It is not difficult to see that $\operatorname{supp}_\Omega B_{mn}$ can even be covered by a finite number of coordinate neighborhoods.

Some theorems of Schwartz[19] tell us that, on Ω, B_{mn} can be written as a finite sum:

$$B_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_r \prod_{\mu=0}^{3} \delta^{(r_\mu)}[t^\mu(\mathbf{P}, \mathbf{Q})]R_{mn}^{(r)}(\mathbf{P}, \mathbf{Q}),$$

where R_{mn}^r is a "tempered distribution" on $\operatorname{supp}_\Omega B_{mn}$. Of course, if $m = n = 1$, the product in the expression above runs only over $\mu = 1, 2, 3$, and the whole discussion simplifies because the manifold is just R^3.

Finally, we smear with test functions $(f, g) \in (\mathcal{S}_m, \mathcal{S}_n)$ satisfying $\operatorname{supp} f \times \operatorname{supp} g \subset \Omega$. This set of pairs of test functions is invariant under any translation $[T(b)f, T(b)g]$. Since $T(b)$ is unitary, and since B_{mn} is a bounded bilinear form, we have

$$|B_{mn}[T(b)f, T(b)g]| \leq C \|f\| \|g\|,$$

with the right-hand side independent of b.

To see the behavior of the left-hand side, we substitute the decomposition of $B_{mn}(\mathbf{P}, \mathbf{Q})$. Integrating by parts, we get

$$B_{mn}[T(b)f, T(b)g]$$

$$= \int \prod_{i,j} d^3 p_i \, d^3 q_j \sum_r \delta[t(\mathbf{P}, \mathbf{Q})]R_{mn}^{(r)}(\mathbf{P}, \mathbf{Q})$$

$$\times \prod_{\mu=0}^{3} (-1)^{r_\mu} \frac{\partial^{r_\mu}}{(\partial t^\mu)^{r_\mu}}\left[e^{it \cdot b} \frac{f(\mathbf{P})g(\mathbf{Q})}{\prod_{i,j} \omega(\mathbf{p}_i)\omega(\mathbf{q}_j)}\right].$$

Suppose that some derivative of a delta function occurs; that is, there is a term with $r \neq 0$ such that $R_{mn}^{(r)} \neq 0$. Consider the terms of highest homogeneous order in r. Carrying out the differentiation gives a

[18] L. Auslander and R. E. MacKenzie, *Introduction to Differentiable Manifolds* (McGraw-Hill Book Company, Inc., New York, 1963), Chap. II. Our excuse for sketching the proof of these well-known facts about the mass shell is that it takes only a few words, and hopefully makes things clearer. For more details, see K. Hepp, Helv. Phys. Acta 36, 355 (1963); and 37, 55 (1964); H. P. Stapp, "Studies in the Foundations of S-Matrix Theory," University of California, Lawrence Radiation Laboratory Report, UCRL 10843.

[19] L. Schwartz, Ref. 11, Vol. I, pp. 100–103, applied to Ex. 2, p. 114, and using the temperedness of B_{mn}.

[17] K. Hepp, Commun. Math. Phys. 1, 7 (1965).

polynomial in b. We can always find a pair (f, g) in our set (because it is dense) such that the coefficient of at least one $(b^0)^{r_0} (b^1)^{r_1} \cdots (b^3)^{r_3}$ in a term of highest homogeneous degree is nonzero. As a function of b, this term cannot be canceled identically by other terms coming from r of the same or lower order. Thus the left-hand side of our inequality contains a polynomial of nonzero degree, which cannot be bounded as a function of b, conflicting with the right-hand side. We conclude that there are no derivatives of delta functions.

But then the remaining delta function implies that

$$B_{mn}[T(b)f, T(b)g] = B_{mn}(f, g),$$

which is what we set out to prove.

V. CLUSTER AMPLITUDES

To avoid a possible point of confusion, we follow Wichmann and Crichton[20] in emphasizing that a large class of amplitudes, labeled in this case by (m, n), has a cluster parametrization, which is given by a purely combinatorial algorithm, having very little to do with the mathematical nature of the amplitudes involved. The *cluster decomposition property* of the S matrix, for example, is logically independent from the cluster parametrization. The relation between the two is rather one of convenience; the cluster property has an especially simple and useful expression in terms of cluster parameters. That, of course, is why cluster parameters are interesting, but we do not assume here that the cluster property holds, nor, for the moment, that we have energy-momentum conservation. We seek only to determine the general structure of the cluster amplitudes for a family of bounded bilinear forms, in the interest of having as much *relevant* information as possible when we apply the theorem on translation invariance.

To help in defining the cluster amplitudes, we introduce some notation. To each bounded bilinear form B_{mn} we associate a kernel defined by

$$B_{mn}(f, g) = \int d\mathbf{P} \, d\mathbf{Q} B_{mn}(\mathbf{P}, \mathbf{Q}) f(\mathbf{P}) g(\mathbf{Q}),$$

where $d\mathbf{P}$ and $d\mathbf{Q}$ are the invariant measure elements for $\mathcal{K}_m$ and $\mathcal{K}_n$. When f and g are in S_m and S_n, the kernel $B_{mn}(\mathbf{P}, \mathbf{Q})$ is the same as the tempered distribution already considered; but it is also defined as a respectable mathematical object for $(f, g) \in (\mathcal{K}_m, \mathcal{K}_n)$. By the Riesz representation theorem, we may associate L_2 (equivalence classes of) functions $B_{mn}(\mathbf{P}, g) \in \mathcal{K}_m$ and $B_{mn}(f, \mathbf{Q}) \in \mathcal{K}_n$ to any $g \in \mathcal{K}_n$ and $f \in \mathcal{K}_m$;

and we have

$$B_{mn}(f, g) = \int d\mathbf{P} \bar{f}(\mathbf{P}) B_{mn}(\mathbf{P}, g),$$

$$= \int d\mathbf{Q} B_{mn}(f, \mathbf{Q}) g(\mathbf{Q}).$$

In other words, we may "integrate" in either order.

Cluster amplitudes B^c_{mn} for a family of such kernels may be defined recursively on (m, n) as follows[21]:
(i) if m or n is zero,

$$B_{mn} = B^c_{mn};$$

(ii) if m and n are nonzero,

$$B_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_I \prod_i B^c_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i}),$$

where I labels the partitions of the variables $(\mathbf{P}, \mathbf{Q})$ into disjoint sets labeled I_i, each of which contains nonzero numbers m_i and n_i of $\mathbf{p}$'s and $\mathbf{q}$'s. Within each partition, the natural order is preserved. Solving, we may write, for m and n nonzero,

$$B^c_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_I \eta(I) \prod_i B_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i}),$$

where $\eta(I)$ is a numerical factor that does not concern us.

As it stands, B^c_{mn} is well defined as a tempered distribution, which contains an over-all delta function for energy-momentum conservation if the B_{mn} do, and as a bounded multilinear form for finite sums of products of one-particle wavefunctions. In addition, we can prove:

Theorem C: The cluster amplitudes B^c_{mn} for a family of bounded bilinear forms are also bounded bilinear forms. In particular, for m and n nonzero, $B^c_{mn}(\mathbf{P}, \mathbf{Q})$ is the kernel of a bounded linear transformation from $\mathcal{K}_n$ into $\mathcal{K}_m$.

For the proof, we may assume that m and n are nonzero; otherwise the result is trivial. Let us consider what meaning we may assign to $B^c_{mn}(\mathbf{P}, g)$ for $g \in \mathcal{K}_n$. The plan of the proof is to show that:
(i) this expression is well defined as an element of $\mathcal{K}_m$, and $B^c_{mn}(f, g)$, defined in this way for all $(f, g) \in (\mathcal{K}_m, \mathcal{K}_n)$, is a bilinear extension of the form already defined if f or g is a sum of products of one-particle wavefunctions;
(ii) the domain of the adjoint of this linear transformation is all of $\mathcal{K}_m$, so that we know from an

---

[20] See Ref. 14.

[21] By analogy with the definition of truncated Wightman functions due to R. Haag, Phys. Rev. **112**, 669 (1958). Any other consistent choice of momentum-dependent, but measurable phases in this definition would be harmless for our purpose.

extension of the Hellinger–Toeplitz theorem[22] that B^c_{mn} is bounded.

From the definition of the cluster amplitudes, it is enough to look at typical terms of the form

$$K^I(\mathbf{P}, g) = \int d\mathbf{Q} \prod_i B_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i})g(\mathbf{Q}).$$

To define such a term, we first partition $\mathbf{Q}$ into two disjoint parts, $\mathbf{Q}_{I_1}$ and the remaining $n - n_1$ three-vector variables, $(\mathbf{Q}_{I_2}, \cdots, \mathbf{Q}_{I_l})$. Because it is an L_2 function, we may consider

$$g[\mathbf{Q}_{I_1}, (\mathbf{Q}_{I_2}, \cdots)] \equiv g(\mathbf{Q})$$

as the kernel of a Hilbert–Schmidt (H–S) operator from $\mathcal{H}_{n-n_1}$ into $\mathcal{H}_{n_1}$. Standard theorems on H–S operators[23] tell us that the product of the bounded linear operator $B_{m_1 n_1}$ and the H–S "operator" g is an H–S operator, and that

$$K_{m_1, n-n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_2}, \cdots)$$
$$\equiv \int d\mathbf{Q}_{I_1} B_{m_1 n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_1})g[\mathbf{Q}_{I_1}, (\mathbf{Q}_{I_2}, \cdots)]$$

is the kernel of the resultant H–S mapping from $\mathcal{H}_{n-n_1}$ into $\mathcal{H}_{m_1}$. That means precisely that $K_{m_1, n-n_1}$ is in the L_2 space of functions of $3(m_1 + n - n_1)$ variables (always with respect to the invariant measure).

Thus, we may repeat the process, partitioning the variables $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_2}, \cdots, \mathbf{Q}_{I_l})$ into two parts, $\mathbf{Q}_{I_2}$ and the rest, $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_3}, \cdots, \mathbf{Q}_{I_l})$. Then we find that

$$K_{m_1, m_2, n-n_1-n_2}(\mathbf{P}_{I_1}, \mathbf{P}_{I_2}, \mathbf{Q}_{I_3}, \cdots)$$
$$\equiv \int d\mathbf{Q}_{I_2} B_{m_2 n_2}(\mathbf{P}_{I_2}, \mathbf{Q}_{I_2})K_{m_1, n-n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_1}, \cdots)$$

is an H–S kernel from the space corresponding to $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_3}, \cdots, \mathbf{Q}_{I_l})$ into the space corresponding to $\mathbf{P}_{I_2}$, and hence L_2 in the space corresponding to all the variables.

Continuing in this way, we find that $K^I(\mathbf{P}, g)$ is L_2 in the nonsymmetrized space corresponding to $\mathcal{H}_m$. It is clear that we have defined in this way a linear map K^I: $\mathcal{H}_n \to \mathcal{H}_m$ that is an extension of the multilinear form defined trivially for wavefunctions of the type $\prod_i g_i(\mathbf{Q}_{I_i})$.

Consider the adjoint of K^I. By definition, a vector $f \in \mathcal{H}_m$ is in the domain of the adjoint if there exists a vector $h \in \mathcal{H}_n$ such that

$$\langle f, K^I g \rangle = \langle h, g \rangle$$

[22] F. Riesz and Sz.-Nagy, Ref. 10, pp. 305–306.
[23] N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Part II, Chap. XI, Sec. 6.

for all $g \in \mathcal{H}_n$. In our case, we find that such a vector exists for every f, so that the adjoint is everywhere defined. The proof is to show that we can calculate the scalar product on the left-hand side of the above equation by integrating first on the $d\mathbf{P}_{I_i}$ successively in some order, then on $d\mathbf{Q}$. By the same argument as before, the $\mathbf{P}$ integration defines for us a vector h; our only problem is to see that we get the same scalar product.

First[24] we consider the scalar product $\langle f, K^I g \rangle$ as an iterated integral on $\mathbf{P}$ and $\mathbf{Q}$, computed in the order (beginning at the right)

$$\int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l} \int d\mathbf{Q}_{I_l} \cdots \int d\mathbf{Q}_{I_1}.$$

The $\mathbf{Q}$ integrations are defined as already described, and we have used Fubini's theorem to write the $\mathbf{P}$ integration in iterated form. Next we note that after doing the integrations on $d\mathbf{Q}_{I_{l-1}} \cdots d\mathbf{Q}_{I_1}$, we have to integrate the kernel of the bounded operator $B_{m_l n_l}$ with a function that is L_2 in $\mathbf{Q}_{I_l}$ and then with a function that is L_2 in $\mathbf{P}_{I_l}$, for fixed values of the remaining variables. We have already observed that, from the definition of the kernel, we can interchange the order of these two integrations. Thus, we may integrate first on $d\mathbf{P}_{I_l} d\mathbf{Q}_{I_{l-1}} \cdots d\mathbf{Q}_{I_1}$; and by our previous argument, the remaining integrand is L_1 in $(\mathbf{P}_{I_1}, \cdots, \mathbf{P}_{I_{l-1}}, \mathbf{Q}_{I_l})$, being a product of two L_2 functions. By Fubini's theorem, we now see that we get the same scalar product if we do the $d\mathbf{Q}_{I_l}$ integration last, integrating in the order

$$\int d\mathbf{Q}_{I_l} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l} \int d\mathbf{Q}_{I_{l-1}} \cdots \int d\mathbf{Q}_{I_1}.$$

At this stage it is not difficult to verify that the $d\mathbf{P}_{I_l}$ integration can be interchanged successively with each preceding $d\mathbf{Q}_{I_i}$ integration, because the $\mathbf{P}$ and $\mathbf{Q}$ integrations are decoupled for $i \neq l$. Thus we arrive at the sequence of integrations

$$\int d\mathbf{Q}_{I_l} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_{l-1}} \int d\mathbf{Q}_{I_{l-1}} \cdots \int d\mathbf{Q}_{I_1} \int d\mathbf{P}_{I_l}.$$

Reasoning by finite descent, we repeat the whole process; and at last we find that the scalar product can be calculated by integrating in the order

$$\int d\mathbf{Q} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l},$$

where we have used Fubini's theorem for the last time to replace the iterated $\mathbf{Q}$ integrations by a single multiple integration. Therefore, the adjoint of K^I has

[24] The reader who treats the following argument as a recipe for pencil and paper will find it straightforward.

all of $\mathcal{K}_m$ for its domain, and we conclude that K^I is bounded.

The original linear transformation was defined on the dense submanifold of $\mathcal{K}_n$ spanned by wavefunctions of the product form. Thus the extension K^I is unique because it is continuous and, in particular, it does not depend on the order in which we choose to do the original **Q** integrations. We are justified in claiming that K^I is well defined for each I, and that the theorem is proved.

Note that by the same argument the converse of Theorem C is also true. If the B_{mn}^c are bounded bilinear forms, so are the B_{mn}.

Now apply Theorem B. It is clear that if the B_{mn} conserve energy momentum, so do the B_{mn}^c. In that case, the cluster amplitudes are translation invariant. We could reach the same conclusion directly from the translation invariance of B_{mn}.

VI. CONCLUSION

We have verified that energy-momentum conservation implies translation invariance in a fairly general class of theories related to Hilbert space, and in particular for the S matrix. We have also shown that the cluster amplitudes for a family of bounded bilinear forms can be discussed in the same framework, as bounded bilinear forms.

As indicated in the title, our hope in this discussion has been not so much to achieve the virtue of originality as that of clarity. If we have not succeeded in even this modest aspiration, we hope that the reader will agree that it is no reflection on the utility or the simplicity of the mathematical tools that we have chosen.

ACKNOWLEDGMENTS

APPENDIX

Here we prove that $E(\Delta)A \subseteq AE(\Delta)$ for all Borel sets if and only if $T(b)A \subseteq AT(b)$ for all translations.[25] The proof that the commuting of the spectral projections with A implies the commuting of the translations with A is rather easy, given some basic results of measure theory and the fact that A is closed. It is not necessary in this case to assume that $\mathcal{K}$ is separable.

The proof of the converse for separable $\mathcal{K}$ is a little more delicate. Modulo a straightforward reduction, our discussion imitates an argument of Sz.-Nagy,[26] used in the proof of Stone's theorem to show that the spectral projections commute with all bounded operators that commute with all elements of the corresponding continuous, one-parameter, unitary group.

Our basic method of proving the two statements is to show that each operator in one of the two sets, labeled by Borel sets or by four-vectors, can be approximated strongly by finite linear combinations of operators in the other set, and to use the fact[27] that if B_n is a strongly convergent sequence of bounded operators with bounded limit B, and if A is a closed operator such that $B_n A \subseteq AB_n$ for all n, then $BA \subseteq AB$.

By means of the functional calculus for bounded functions of commuting self-adjoint operators (such as P_μ), the approximation of the operators in one class by those of the other can be reduced to that of the approximation of the corresponding functions. Namely, let $h(p)$ be a bounded function on R^4, measurable with respect to the spectral measure, i.e., with respect to all the measures $\langle f, dE(p)g \rangle$; and let $h_n(p)$ be a uniformly bounded sequence of such functions, which converges to $h(p)$ almost everywhere with respect to the spectral measure. Then the corresponding bounded operators

$$h_n = \int h_n(p) \, dE(p)$$

converge strongly to[28]

$$h = \int h(p) \, dE(p).$$

In our case we have to consider two classes of such functions, composed on the one hand of finite linear combinations of characteristic functions of Borel sets,

$$\xi_\Delta(p) = \begin{cases} 1 & \text{if} \quad p \in \Delta, \\ 0 & \text{if} \quad p \notin \Delta, \end{cases}$$

and on the other hand of finite linear combinations of exponentials, $\exp(ib \cdot p)$, i.e., of trigonometric polynomials. These functions are certainly bounded. The characteristic functions are measurable with respect to the spectral measure, because on locally compact Hausdorff spaces such as R^n the Borel sets are measurable with respect to any measure; and continuous functions, such as exponentials, are measurable with respect to any measure on such

[25] Note that $T(b)A \subseteq AT(b)$ for all translations implies by definition that $T(b)D(A) \subset D(A)$, and hence from the group property that $T(b)A = AT(b)$.

[26] F. Riesz and Sz.-Nagy, Ref. 10, p. 383.
[27] Ref. 10, p. 302.
[28] Ref. 10, Sec. 126.

spaces.[29] Our problem is to show that sufficiently many functions in each class can be approximated in the sense described above by functions in the other class.

To emphasize the point at which the separability of $\mathcal{K}$ enters, we divide the work among three lemmas, which together add up to the required results. We do not assume that $\mathcal{K}$ is separable unless we say so explicitly. But we always assume that A is closed.

Lemma 1: If $E(\Delta)A \subseteq AE(\Delta)$ for all Borel sets, then $T(b)A \subseteq AT(b)$ for all b.

Proof: We refer to a basic theorem of measure theory, according to which any measurable function is the limit of an *everywhere*-convergent sequence of simple functions.[30] A simple function is a finite linear combination of characteristic functions of pairwise-disjoint, measurable sets. The sequence can be chosen to be uniformly bounded if the limit function is bounded.[31] Since a continuous function on $\mathbb{R}^n$ is, in particular, Borel-measurable, the result follows from our previous remarks.

Lemma 2: If $T(b)A \subseteq AT(b)$ for all b, then $E(\Delta)A \subseteq AE(\Delta)$ for all compact Δ.

Proof: We have to express ξ_Δ for any compact Δ as the limit of a uniformly bounded, everywhere-convergent sequence of trigonometric polynomials. This can be achieved by the argument of Sz.-Nagy mentioned before.[26] First we take a decreasing sequence $\{U_n\}$ of bounded open neighborhoods of Δ, such that $\bigcap_{n=1}^{\infty} U_n = \Delta$. Applying Urysohn's lemma,[32] we choose a continuous, nonnegative, real function f_n which is unity on Δ, has support in $\bar{U}_n$ (the closure of U_n), and is bounded by unity. Next, we choose an increasing sequence of compact cubes $\square_n < \bar{U}_n$, such that $\bigcup_{n=1}^{\infty} \square_n = \mathbb{R}^4$; and we let g_n be the continuous periodic function defined by f_n in $\square_n$. The uniformly bounded sequence $\{g_n\}$ converges everywhere to ξ_Δ.

Finally, we apply Weierstrass's approximation theorem[33] to approximate g_n uniformly to within $1/n$ by a trigonometric polynomial t_n of the same period. The sequence $\{t_n\}$ is uniformly bounded and converges everywhere to ξ_Δ.

Lemma 3: Let $\mathcal{K}$ be separable. If $E(\Delta)A \subseteq AE(\Delta)$ holds for all compact Δ, it holds for all Borel sets.

Proof: Every Borel set is "summable" with respect to the spectral measure; i.e., $\langle f, E(\Delta)f \rangle$ is finite for all $f \in \mathcal{K}$. According to a basic result of measure theory,[34] if a set is summable with respect to some measure, there is a denumerable family of compact sets $\Delta_n \subset \Delta$ (which can even be chosen pairwise-disjoint) such that the set $R = \Delta - \bigcup_{n=1}^{\infty} \Delta_n$ is a set of zero measure (the difference of two sets is the set of points in the first, not in the second). We want to find a similar family with the property that the remainder R, which is a Borel set in our case, has spectral measure zero, i.e., such that $\langle f, E(R)f \rangle = 0$ for all f.

This equation is true for all vectors in $\mathcal{K}$ if and only if it is true for a dense set in $\mathcal{K}$, because $E(R)$ is a projection, hence bounded, hence continuous. Since $\mathcal{K}$ is separable, we can choose a denumerable dense set of vectors f_i.

Corresponding to each f_i, we choose a decomposition of Δ as above, such that

$$R_i \equiv \Delta - \bigcup_{n=1}^{\infty} \Delta_n^{(i)}$$

has measure zero for the corresponding measure. It follows that

$$R \equiv \bigcap_{i=1}^{\infty} R_i = \Delta - \bigcup_{i=1}^{\infty} \bigcup_{n=1}^{\infty} \Delta_n^{(i)}$$

is a Borel set which satisfies

$$\langle f_i, E(R)f_i \rangle = 0$$

for all f_i, since any subset of a set of zero measure has zero measure, and $R \subset R_i$.

A denumerable union of a denumerable union is still a denumerable union, so by taking all the compact sets in each decomposition of Δ and relabeling them, we get a denumerable family of compact sets $\Delta_n \subset \Delta$ such that

$$E(\Delta) = E(\Delta - R) = E\left(\bigcup_{n=1}^{\infty} \Delta_n\right).$$

Now we have only to note that the characteristic functions of the increasing sequence of compact sets $C_N = \bigcup_{n=1}^{N} \Delta_n$ are uniformly bounded and converge pointwise to $\xi_{\Delta-R}$.

Thus, $E(\Delta) = \lim E(C_N)$, and the lemma follows from the property of closed operators that has been the theme of our discussion.

---

[29] M. A. Naimark, *Normed Rings* (P. Noordhoff Ltd., Groningen, The Netherlands, 1964), Appendix III.

[30] P. R. Halmos, *Measure Theory* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950), p. 86.

[31] Ref. 30, Ex. 2, p. 86.

[32] Ref. 29, p. 28.

[33] Ref. 29, p. 33.

[34] Ref. 29, p. 129.

Angular Coefficients of Atomic Matrix Elements Involving Interelectronic Coordinates

W. H. HEINTZ AND R. L. GIBBS*

Physics Department, Clarkson College of Technology, Potsdam, New York

(Received 20 October 1966)

A method for determining the angular coefficients of atomic matrix elements is illustrated. The angular coefficients of matrix elements for r_{ij}^a, $r_{ij}^a r_{jk}^b$, $r_{ij}^a r_{jk}^b r_{ki}^c$, $r_{ij}^a r_{jk}^b r_{kl}^c$, and $r_{ij}^a r_{jk}^b r_{kl}^c r_{li}^d$ are evaluated using single-particle states of definite angular momentum. The use of tensor operators enables a separation into angular and radial parts. The atomic matrix elements are then expressed as sums over products of n-j symbols and radial integrals. These sums are restricted by the values of the single-particle state angular momenta, and in all cases the effects of single-particle couplings disappear. The calculation does not require the use of a particular coordinate system, as is the case for multiple products of spherical harmonics.

I. INTRODUCTION

RECENTLY much effort has gone into the evaluation of atomic integrals involving functions of the interelectronic coordinates. This effort has been motivated by variational treatments of small atomic systems in which the trial wavefunctions depend explicitly on the interelectronic coordinates in order to account for correlation effects. The matrix elements in question arise from the expectation value of the Hamiltonian and normalization integrals. They contain parameters to be varied in order to minimize the energy value obtained for a particular trial wavefunction. Several schemes have evolved which are capable of handling the integrals for cases in which 2, 3, and 4 electrons are involved. These schemes are well adapted to single-particle states of spherical symmetry, but not to states of $p, d, f, \cdots$ symmetry.

The evaluation of matrix elements involving more than three interelectronic coordinates has been largely developed by Bonham.[1,2] However, as pointed out by Roberts,[3] the method is really limited to single-particle states of spherical symmetry. Special cases have been worked out for matrix elements involving one to three electronic coordinates by resorting to special coordinate systems.[4] Slater's tables[5] are sufficient for the angular coefficients of two electron matrix elements, but general expressions for those involving two or more interelectronic coordinates are required for a procedure such as Öhrn and Nordling's.[6]

The purpose of this paper is to demonstrate the use of irreducible tensor operators in the evaluation of many electron integrals. The matrix elements treated involve the following operators:

$$ r_{ij}^a, \quad r_{ij}^a r_{jk}^b, \quad r_{ij}^a r_{jk}^b r_{ki}^c, \quad r_{ij}^a r_{jk}^b r_{kl}^c, \quad r_{ij}^a r_{jk}^b r_{kl}^c r_{li}^d. $$

More general examples can be handled without further conceptual difficulties. The evaluation is accomplished by expanding the functions of the interelectronic coordinates in terms of tensor operators, recoupling the operators for functions of several interelectronic coordinates, coupling the single-particle states into states of total angular momentum, and finally evaluating the matrix elements. A complete separation of the integrals into sums over products of radial and angular parts is inherent in the procedure.

There are several advantages in the use of tensor operators in this problem. The methods referred to above depend on the use of a particular coordinate system and, for states of other than spherical symmetry, require the use of multiple products of spherical harmonics. The use of tensor operators enables one to construct such products, in terms of recouplings, without the explicit use of a coordinate system. More specifically, the work described here follows closely that of Öhrn and Nordling[6] except that the angular parts in Sec. II of their paper are treated here by means of tensor operators. The radial parts, which depend on the radial part of the trial wavefunctions, can then be handled by the methods referred to in Refs. 1–6.

The main advantage of this method lies in its ability to express readily the angular parts in terms of 3-j symbols. Some of these involve the single-particle angular momenta, but the final expressions do not involve the total angular momenta. The triangular inequalities of the 3-j symbols and parity considerations greatly restrict the sums, whether finite or

* Present address: Louisiana Polytechnic Institute, Ruston, Louisiana.

[1] R. A. Bonham, J. Mol. Spectry **15**, 112 (1965).

[2] R. A. Bonham, J. Mol. Spectry **20**, 197 (1966).

[3] P. J. Roberts, Proc. Phys. Soc. (London) **88**, 53 (1966).

[4] L. Szasz, J. Chem. Phys. **35**, 1072 (1961); J. L. Calais and P. O. Löwdin, J. Mol. Spectry. **8**, 203 (1962); J. Hinze and K. S. Pitzer, J. Chem. Phys. **41**, 3484 (1964); E. A. Burke, J. Math. Phys. **6**, 1691 (1965).

[5] J. C. Slater, *Quantum Theory of Atomic Structure* (McGraw-Hill Book Company, Inc., New York, 1960), Vols. I and II.

[6] Y. Öhrn and J. Nordling, J. Chem. Phys. **39**, 1864 (1963).

infinite. These restrictions enable one to determine systematically those matrix elements that must vanish, those that involve only finite sums, and those that involve infinite sums.

II. TENSOR OPERATORS AND THEIR USE IN EVALUATING ATOMIC MATRIX ELEMENTS

A. Coupling of Single-Particle States

The matrix elements to be evaluated involve products of single-particle states. These are then coupled into states of total angular momentum. The single-particle states are denoted here by $|l_i m_i\rangle$, where l_i is the angular momentum quantum number and m_i is the magnetic quantum number. Product states $|l_i m_i\rangle |l_j m_j\rangle$ are denoted as $|l_i m_i l_j m_j\rangle$ and are abbreviated as $|ij\rangle$. Similar notation is used for products of three- and four-electron states. The coupled states are denoted by $|(l_i l_j)LM\rangle$.

The coupling schemes are illustrated below for products of two-, three-, and four-electron states:

(a) Two electrons

$$|ij\rangle = |l_i m_i l_j m_j\rangle = \sum_L |(l_i l_j)LM\rangle((l_i l_j)LM \mid l_i m_i l_j m_j);$$

(1)

(b) Three electrons

$$|ijk\rangle = |l_i m_i l_j m_j l_k m_k\rangle,$$
$$= \sum_{L, l_{ij}} |(l_i l_j)l_{ij}, l_k LM\rangle((l_{ij}l_k)LM \mid l_{ij}m_{ij}l_k m_k)$$
$$\times ((l_i l_j)l_{ij}m_{ij} \mid l_i m_i l_j m_j);$$

(2)

(c) Four electrons

$$|ijkl\rangle = |l_i m_i l_j m_j l_k m_k l_l m_l\rangle,$$
$$= \sum_{L, l_{il}, l_{jk}} |(l_i l_l)l_{il}, (l_j l_k)l_{jk}, LM\rangle$$
$$\times ((l_i l_l)l_{il}m_{il} \mid l_i m_i l_l m_l)((l_j l_k)l_{jk}m_{jk} \mid l_j m_j l_k m_k)$$
$$\times ((l_{il}l_{jk})LM \mid l_{il}m_{il}l_{jk}m_{jk}).$$

(3)

In each case the sums are limited by the appropriate triangular inequalities. For example, in Eq. (1), L is limited by $|l_i - l_j| \leq L \leq l_i + l_j$ with similar limits on remaining sums of Eqs. (2) and (3). The coefficients in these expressions are the usual Clebsch–Gordan coefficients which may be given in terms of the Wigner 3-j symbols[7]

$$((l_1 l_2)l_{12}m_{12} \mid l_1 m_1 l_2 m_2)$$
$$= (-1)^{-l_1 + l_2 - m_{12}}[l_{12}]^{\frac{1}{2}}\begin{pmatrix} l_1 & l_2 & l_{12} \\ m_1 & m_2 & -m_{21} \end{pmatrix},$$

(4)

where $[l_{12}] = 2l_{12} + 1$.

It should be stressed that the coupling schemes for the single-particle states are not unique, but must be

chosen so as to be consistent with the later couplings of the tensor operators.

B. Definition and Properties of the Tensor Operators

The tensor operators are defined in terms of the normalized spherical harmonics as[8]

$$C_i^{k,q} = (4\pi/[k])^{\frac{1}{2}} Y_{k,q}(\theta_i, \phi_i).$$

(5)

The subscript i refers to the coordinates of the ith electron. The properties of tensor operators needed for this work are reviewed in many places.[7-11]

One may also define mixed tensor operators as

$$\{C_i^n C_j^m\}_Q^K = \sum_{p,q} C_i^{n,p} C_j^{m,q}(npmq \mid (nm)KQ),$$

(6)

where $-n \leq p \leq n$ and $-m \leq q \leq m$. The tensor operators and the mixed operators transform according to irreducible representations of the three-dimensional rotation group.

Of particular importance is the quantity

$$\{C_i^k C_j^k\}_0^0 = (-1)^k[k]^{-\frac{1}{2}} \sum_q (-1)^q C_i^{k,q} C_j^{k,-q},$$

(7)

where $-k \leq q \leq k$. This is related to the scalar product of two tensors as defined in Ref. 11.

The spherical harmonic addition theorem may now be written as

$$P_k(\cos \theta_{ij}) = (-1)^k[k]^{\frac{1}{2}}\{C_i^k C_j^k\}_0^0.$$

(8)

In what follows the value of Q for the mixed tensor operators is not specified if K is zero or if the results are independent of Q.

C. Expansion of the Operators

The operators for which the matrix elements are to be evaluated involve products of powers of the interelectronic coordinates. The operators may be expressed in terms of the tensor operators using the spherical harmonic addition theorem as expressed above. In these expansions it is assumed, though not necessary, that the powers of the interelectronic coordinates are greater than or equal to -1. Several examples are shown below. The procedure follows closely that of Sec. II of the article of Calais and Löwdin in Ref. 4 and Sec. 7.2 of Ref. 7, except that here the results are expressed in terms of the tensor operators.

[7] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).

[8] B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill Book Company, Inc., New York, 1963).
[9] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).
[10] V. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959).
[11] G. Racah, Phys. Rev. **62**, 438 (1942).

Note that

$$r_{12}^a = (r_1^2 + r_2^2 - 2r_1r_2 \cos \theta_{12})^{a/2},$$

$$= r_>^a [1 + (r_<^2/r_>^2) - (2r_</r_>)x]^{a/2}, \quad (9)$$

where $x = \cos \theta_{12}$, $r_>$ is the greater of r_1 and r_2, and $r_<$ is the lesser of r_1 and r_2. It is more convenient to express this as

$$r_{12}^a = \rho_{12}^a(1 + g_{12}^2 - 2g_{12}x)^{a/2}, \quad (10)$$

where $\rho_{12} = r_>$ and $g_{12} = r_</r_>$. The form of the expansion is

$$r_{12}^a = \sum_{p=0}^{\infty}(-1)^p[p]^{\frac{1}{2}}h_a(1, 2, p)\{C_1^pC_2^p\}^0, \quad (11)$$

where

$$h_a(1, 2; p) = \tfrac{1}{2}(2p + 1)\int_{-1}^{1} r_{12}^a P_p(x) \, dx. \quad (12)$$

For example, if $a = -1$,

$$h_{-1}(1, 2; p) = \rho_{12}^{-1}g_{12}^p; \quad (13)$$

or if $a = 1$,

$$h_1(1, 2; p) = \rho_{12}g_{12}^p\left(\frac{g_{12}^2}{2p + 3} - \frac{1}{2p - 1}\right). \quad (14)$$

The procedure can be extended to other functions of r_{12} leading to more complicated radial parts, but to the same angular form expressed in terms of the tensor operators.

For products of functions of two or more of the interelectronic coordinates, one forms the product of the required number of series. For example,

$$r_{12}^a r_{23}^b = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}(-1)^{p+q}([p][q])^{\frac{1}{2}}$$
$$\times h_a(1, 2; p)h_b(2, 3; q)\{C_1^pC_2^p\}^0\{C_2^qC_3^q\}^0, \quad (15)$$

where $h_a(1, 2; p)$ and $h_b(2, 3; q)$ are evaluated as above. All other cases may be generated by taking products in this manner.

D. Matrix Elements of the Mixed Tensor Operators

Before evaluating the matrix elements of the tensor operators, one needs to recouple the tensor operators in accord with the couplings of the single-particle states as given earlier. One may then find the matrix elements of the recoupled tensor operators with respect to states of total angular momentum and finally with respect to the single-particle state angular momenta.

For the examples to be shown in this paper, the following recouplings are sufficient:

$$\{C_1^pC_2^p\}^0\{C_2^qC_3^q\}^0\{C_1^rC_3^r\}^0$$
$$= \sum_{u,v,x} [x][u][v]([p][q][r])^{-\frac{1}{2}}$$
$$\times \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} r & u & p \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{Bmatrix} r & u & p \\ v & q & x \end{Bmatrix}\{\{C_1^uC_2^v\}^xC_3^x\}^0 \quad (16)$$

and

$$\{C_1^pC_2^p\}^0\{C_2^qC_3^q\}^0\{C_3^rC_4^r\}^0\{C_1^sC_4^s\}^0$$
$$= \sum_{u,v,x,y,z} [x]^{\frac{1}{2}}[u][v][y][z]([p][q][r][s])^{-\frac{1}{2}}$$
$$\times \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} r & v & q \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} p & y & s \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} r & z & s \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{Bmatrix} p & u & q \\ v & r & x \end{Bmatrix}\begin{Bmatrix} p & y & s \\ z & r & x \end{Bmatrix}\{\{C_1^uC_4^z\}^x\{C_2^uC_3^v\}^x\}^0, \quad (17)$$

where $\begin{Bmatrix} r & u & p \\ v & q & x \end{Bmatrix}$ of Eq. (16) and the similar forms of Eq. (17) are 6-j symbols.[7,12] The summation indices in these and many of the succeeding equations are limited both by the appropriate triangular inequalities and parity considerations. For example, in Eq. (16), x is limited by $|r - q| \leq x \leq r + q$ and by the requirement that $r + x + q$ be an even integer in order that $\begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix}$ does not vanish. Equations (16) and (17) are verified in Appendix A.

The matrix elements of the tensor operators for states of total orbital angular momentum are now given. The method involves the Wigner–Eckart theorem and follows directly from the procedures in Refs. 7, 8, and 12. The matrix elements of the recoupled tensor operators of Eqs. (16) and (17) are shown below with derivations given in Appendix B:

$$((l_{12}l_3)LM| \{\{C_1^uC_2^v\}^xC_3^x\}^0 |(l_il_k)LM')$$
$$= (-1)^\alpha\delta_{L,L'}\delta_{M,M'}([l_1][l_2][l_3][l_i][l_j][l_k][l_{12}][l_{ij}])^{\frac{1}{2}}$$
$$\times \begin{pmatrix} l_1 & u & l_i \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & v & l_j \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_3 & x & l_k \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{Bmatrix} l_{ij} & l_k & L \\ l_3 & l_{12} & p \end{Bmatrix}\begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}, \quad (18)$$

where $\alpha = x + L + l_{ij} + l_1 + l_2$, and

$$\begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}$$

is a 9-j symbol[7];

$$((l_{14}l_{23})LM| \{\{C_1^yC_4^z\}^x\{C_2^uC_3^v\}^x\}^0 |(l_{il}l_{jk})L'M')$$
$$= (-1)^\beta\delta_{L,L'}\delta_{M,M'}([l_1][l_2][l_3][l_4][l_i][l_j][l_k]$$
$$\times [l_l][l_{14}][l_{23}][l_{il}][l_{jk}][x])^{\frac{1}{2}}$$
$$\times \begin{pmatrix} l_1 & y & l_i \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_3 & v & l_k \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_4 & z & l_l \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{Bmatrix} l_{il} & l_{jk} & L \\ l_{23} & l_{14} & x \end{Bmatrix}\begin{Bmatrix} l_1 & l_i & y \\ l_4 & l_l & z \\ l_{14} & l_{il} & x \end{Bmatrix}\begin{Bmatrix} l_2 & l_j & u \\ l_3 & l_k & v \\ l_{23} & l_{jk} & x \end{Bmatrix}, \quad (19)$$

where $\beta = x + L + l_{il} + l_{23} + l_1 + l_2 + l_3 + l_4$.

---

[12] M. Rootenberg, R. Bivens, N. Metropolis, and J. Wooten, *The 3-j and 6-j Symbols* (Technology Press, Cambridge, Massachusetts, 1959).

Note that in these and the succeeding equations the single-particle state angular momenta, as well as the order of the tensor operators, have been assumed to be integral. This simplifies the phase factors of Eq. (18) and succeeding equations. For the atomic matrix elements the presence of spin is accounted for by multiplying the orbital part of the single-particle states by an appropriate spin function and antisymmetrizing the resulting products. For problems in which the single-particle state angular momenta are assumed to be half odd integers, it is necessary to rederive the phase factors of Eq. (18) and the succeeding equations. This is straightforward and is not done here.

III. ATOMIC MATRIX ELEMENTS

The results of Sec. II are now combined to find the atomic matrix elements for those products of powers of the interelectronic coordinates mentioned in the Introduction. The results are expressed as sums over products of angular coefficients and radial matrix

elements. The angular coefficients are finally expressed in terms of 3-j symbols. The sums for the examples included reduce to at most one restricted infinite sum.

The angular coefficients of $(123| r_{12}^a r_{23}^b r_{31}^c |ijk)$ and $(1234| r_{12}^a r_{23}^b r_{34}^c r_{41}^d |ijkl)$ may be found as follows:

(1) Expand the products of powers of the interelectronic coordinates in terms of the tensor operators as in Sec. IIC;

(2) Express the single-particle product states in terms of states of total orbital angular momentum as in Sec. IIA;

(3) Recouple the mixed tensor operators and evaluate their matrix elements, with respect to states of total angular momenta, as in Sec. IID.

When these steps are combined, one obtains the angular coefficients in terms of n-j symbols. The results for the two matrix elements mentioned above are given below. The other examples referred to in the Introduction are treated as special cases of these two.

$$
(123| r_{12}^a r_{23}^b r_{31}^c |ijk) = \sum_{p,q,r=0}^{\infty} \sum_{u,v,x} \sum_{L,l_{12},l_{ij}} (-1)^{\gamma} [u][v][x][L][l_{12}][l_{ij}]([l_1][l_2][l_3][l_i][l_j][l_k])^{\frac{1}{2}}
$$

$$
\times \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & u & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} r & u & p \\ v & q & x \end{Bmatrix} \begin{pmatrix} l_1 & u & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & v & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_{12} \\ m_1 & m_2 & -m_{12} \end{pmatrix}
$$

$$
\times \begin{pmatrix} l_i & l_j & l_{ij} \\ m_i & m_j & -m_{ij} \end{pmatrix} \begin{pmatrix} l_{12} & l_3 & L \\ m_{12} & m_3 & -M \end{pmatrix} \begin{pmatrix} l_{ij} & l_k & L \\ m_{ij} & m_k & -M \end{pmatrix} \begin{Bmatrix} l_{ij} & l_k & L \\ l_3 & l_{12} & x \end{Bmatrix} \begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}
$$

$$
\times (\gamma_{123}| h_a(1,2;p) h_b(2,3;q) h_c(3,1;r) |\gamma_{ijk}), \tag{20}
$$

where γ_{ijk} refers to the radial part of the single-particle state, $|ijk)$ and

$$
\gamma = p + q + r + l_i + l_j + m_{12} + m_{ij} + l_3 + l_{12} + l_k + x + L.
$$

The factor

$$
(\gamma_{123}| h_a(L,2;p) h_b(2,3;q) h_c(3,1;r) |\gamma_{ijk}),
$$

and the similar factor of Eq. (21) below, is the radial portion of the matrix element:

$$
(1234| r_{12}^a r_{23}^b r_{34}^c r_{41}^d |ijkl) = \sum_{p,q,r,s=0}^{\infty} \sum_{u,v,x,y,z} \sum_{L,l_{14},l_{23},l_{il},l_{kj}} [u][v][x][y][z][L][l_{14}][l_{23}][l_{il}][l_{jk}]([l_1][l_2][l_3][l_4][l_i][l_j][l_k][l_l])^{\frac{1}{2}})
$$

$$
\times \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & v & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & y & s \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & z & s \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} p & u & q \\ v & r & x \end{Bmatrix} \begin{Bmatrix} p & y & s \\ z & r & x \end{Bmatrix} \begin{pmatrix} l_1 & y & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ 0 & 0 & 0 \end{pmatrix}
$$

$$
\times \begin{pmatrix} l_4 & z & l_l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_4 & l_{14} \\ m_1 & m_4 & -m_{14} \end{pmatrix} \begin{pmatrix} l_i & l_l & l_{il} \\ m_i & m_l & -m_{il} \end{pmatrix} \begin{pmatrix} l_2 & l_3 & l_{23} \\ m_2 & m_3 & -m_{23} \end{pmatrix} \begin{pmatrix} l_j & l_k & l_{jk} \\ m_j & m_k & -m_{jk} \end{pmatrix}
$$

$$
\times \begin{pmatrix} l_{14} & l_{23} & L \\ m_{14} & m_{23} & -M \end{pmatrix} \begin{pmatrix} l_{il} & l_{jk} & L \\ m_{il} & m_{jk} & -M \end{pmatrix} \begin{Bmatrix} l_{il} & l_{jk} & L \\ l_{23} & l_{14} & x \end{Bmatrix} \begin{Bmatrix} l_1 & l_i & y \\ l_4 & l_l & z \\ l_{14} & l_{il} & x \end{Bmatrix} \begin{Bmatrix} l_2 & l_j & u \\ l_3 & l_k & v \\ l_{23} & l_{jk} & x \end{Bmatrix}
$$

$$
\times (\gamma_{1234}| h_a(1,2;p) h_b(2,3;q) h_c(3,4;r) h_d(4,1;s) |\gamma_{ijkl}), \tag{21}
$$

where

$$\delta = p + q + r + s + m_{14} + m_{23} + l_{14} + l_i + l_l + m_{il} + l_j + l_k + m_{jk} + l_{jk} + x + L.$$

Equation (20) is derived in Appendix C, and Eq. (21) may be derived in like manner.

These equations can be simplified somewhat. One can perform the sums over the total and intermediate angular momenta and remove the effects of the single-particle couplings. Two equations which facilitate this are

$$\sum_L (1)^L [L] \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_i & l_j & L \\ m_i & m_j & -M \end{pmatrix} \begin{Bmatrix} l_i & l_j & L \\ l_2 & l_1 & x \end{Bmatrix} = (-1)^{x+m_i+m_2} \begin{pmatrix} l_i & l_1 & x \\ m_i & -m_1 & -m_x \end{pmatrix} \begin{pmatrix} l_2 & l_j & x \\ -m_2 & m_j & m_x \end{pmatrix}$$

(22)

and variations of

$$\sum_{l_{12},l_{ij}} (-1)^{l_{12}} [l_{12}][l_{ij}] \begin{pmatrix} l_1 & l_2 & l_{12} \\ m_1 & m_2 & -m_{12} \end{pmatrix} \begin{pmatrix} l_i & l_j & l_{ij} \\ m_i & m_j & -m_{ij} \end{pmatrix} \begin{pmatrix} l_{ij} & l_{12} & x \\ m_{ij} & -m_{12} & -m_x \end{pmatrix} \begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}$$

$$= (1)^{u+l_i+l_2} \begin{pmatrix} l_1 & l_i & u \\ m_1 & -m_i & -m_u \end{pmatrix} \begin{pmatrix} l_2 & l_j & v \\ -m_2 & m_j & m_v \end{pmatrix} \begin{pmatrix} v & x & u \\ m_v & m_x & m_u \end{pmatrix}. \quad (23)$$

Equations (22) and (23) are derived in Appendix D.

The process of summing over L, l_{12}, and l_{ij} in Eq. (20) is accomplished by using Eq. (22) to sum over L and Eq. (23) to sum over l_{12} and l_{ij}. The result is

$$(123| r_{12}^a r_{23}^b r_{31}^c |ijk) = \sum_{p,q,r=0}^{\infty} \sum_{u,v,x} (-1)^\epsilon \times [u][v][x] ([l_1][l_2][l_3][l_i][l_j][l_k])^{\frac{1}{2}} (\gamma_{123}| h_a(1,2;p) h_b(2,3;q) h_c(3,1;r) |\gamma_{ijk})$$

$$\times \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & u & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & u & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & v & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & u & l_i \\ -m_1 & m_u & m_i \end{pmatrix}$$

$$\times \begin{pmatrix} l_2 & v & l_j \\ -m_2 & m_v & m_j \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ -m_3 & m_x & m_k \end{pmatrix} \begin{pmatrix} v & x & u \\ m_v & m_x & m_u \end{pmatrix} \begin{Bmatrix} v & x & u \\ r & p & q \end{Bmatrix}, \quad (24)$$

where $\epsilon = p + q + r + m_1 + m_2 + m_3$. In order to put this into a more symmetrical form, one can eliminate the 6-j symbol using Eq. (D1) to give

$$(123| r_{12}^a r_{23}^b r_{31}^c |ijk) = \sum_{p,q,r=0}^{\infty} \sum_{m_p,m_q,m_r} \sum_{u,v,x} (-1)^A [u][v][x] ([l_1][l_2][l_3][l_i][l_j][l_k])^{\frac{1}{2}}$$

$$\times (\gamma_{123}| h_a(1,2;p) h_b(2,3;q) h_c(3,1;r) |\gamma_{ijk})$$

$$\times \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & u & r \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & v & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & u & l_i \\ -m_1 & m_u & m_i \end{pmatrix}$$

$$\times \begin{pmatrix} l_2 & v & l_j \\ -m_2 & m_v & m_j \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ -m_3 & m_x & m_k \end{pmatrix} \begin{pmatrix} r & x & q \\ -m_r & m_x & m_q \end{pmatrix} \begin{pmatrix} p & u & r \\ -m_p & m_u & m_r \end{pmatrix} \begin{pmatrix} q & v & p \\ -m_q & m_v & m_p \end{pmatrix}, \quad (25)$$

where $A = m_1 + m_2 + m_3 + m_p + m_q + m_r$. The single-particle states restrict the magnetic quantum numbers m_p, m_q, and m_r so that only one is independent. In using Eq. (25) it is necessary to recall the triangular inequalities between u, v, and x from Eq. (24).

To simplify Eq. (21) one uses Eq. (22) once to sum over L and variations of Eq. (23) are used twice to

sum over l_{14}, l_{23}, l_{il}, and l_{jk}. The result is

$$(1234| r_{12}^a r_{23}^b r_{34}^c r_{41}^d |ijkl) = \sum_{p,q,r,s=0}^{\infty} \sum_{u,v,x,y,z} (-1)^\zeta [u][v][x][y][z]([l_1][l_2][l_3][l_4][l_i][l_j][l_k][l_l])^{\frac{1}{2}}$$

$$\times (\gamma_{1234}| h_a(1,2;p)h_b(2,3;q)h_c(3,4;r)h_d(4,1;s) |\gamma_{ijkl}) \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & v & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & y & s \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & z & s \\ 0 & 0 & 0 \end{pmatrix}$$

$$\times \begin{pmatrix} l_1 & y & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_4 & z & l_l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & y & l_i \\ -m_1 & m_y & m_i \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ -m_2 & m_u & m_j \end{pmatrix}$$

$$\times \begin{pmatrix} l_3 & v & l_k \\ -m_3 & m_v & m_k \end{pmatrix} \begin{pmatrix} l_4 & z & l_l \\ -m_4 & m_z & m_l \end{pmatrix} \begin{pmatrix} v & x & u \\ m_v & -m_x & m_u \end{pmatrix} \begin{pmatrix} z & x & y \\ m_z & m_x & m_y \end{pmatrix} \begin{Bmatrix} v & x & u \\ p & q & r \end{Bmatrix} \begin{Bmatrix} z & x & y \\ p & s & r \end{Bmatrix}, \tag{26}$$

where $\zeta = p + q + r + s + x + m_1 + m_4 + m_j + m_k$.

One can now use Eq. (D1), the symmetry properties of the 3-j symbols, and Eq. (D5) to put Eq. (26) in the form

$$(1234| r_{12}^a r_{23}^b r_{34}^c r_{41}^d |ijkl) = \sum_{p,q,r,s=0}^{\infty} \sum_{m_p,m_u,m_r,m_s} \sum_{u,v,y,z} (-1)^B [u][v][y][z]([l_1][l_2][l_3][l_4][l_i][l_j][l_k][l_l])^{\frac{1}{2}}$$

$$\times (\gamma_{1234}| h_a(1,2;p)h_b(2,3;q)h_c(3,4;r)h_d(4,1;s) |\gamma_{ijkl}) \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & v & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & y & s \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & z & s \\ 0 & 0 & 0 \end{pmatrix}$$

$$\times \begin{pmatrix} l_1 & y & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_4 & z & l_l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & u & q \\ m_p & m_r & m_q \end{pmatrix} \begin{pmatrix} r & v & q \\ m_r & -m_v & m_q \end{pmatrix} \begin{pmatrix} p & y & s \\ -m_p & m_y & m_s \end{pmatrix}$$

$$\times \begin{pmatrix} r & z & s \\ m_r & m_z & -m_s \end{pmatrix} \begin{pmatrix} l_1 & y & l_i \\ -m_1 & m_y & m_i \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ -m_2 & m_u & m_j \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ -m_3 & m_v & m_k \end{pmatrix} \begin{pmatrix} l_4 & z & l_l \\ -m_4 & m_z & m_l \end{pmatrix}, \tag{27}$$

where $B = m_1 + m_4 + m_j + m_k + m_r + m_s + m_v$. The single-particle states again restrict the magnetic quantum numbers m_p, m_y, m_r, and m_s so that only one is independent.

In using Eq. (27) it is necessary to recall the triangular inequalities between u, v, and x and between x, y, and z from Eq. (26). This is necessary even though x does not appear in Eq. (27).

In both Eqs. (25) and (27) it is seen that, from the 3-j symbols, the single-particle state angular momenta specify and restrict the angular summation indices u, v, x, y, and z to finite values but not

the radial summation indices p, q, u, and s. However, the triangular inequalities and parity restrictions do reduce the radial summations to one restricted infinite sum. Specific illustrations are given in the next section.

The remaining integrals, referred to in the Introduction, are now shown as special cases of Eq. (27). Two of these could also be obtained from Eq. (25). If in Eq. (27) one sets the exponent of r_{14} equal to zero, the sum over s is restricted to $s = 0$. On evaluating the 3-j symbols containing s, one obtains

$$(1234| r_{12}^a r_{23}^b r_{34}^c |ijkl) = \sum_{p,q,r} \sum_{u,v} (-1)^C [u][v]([l_1][l_2][l_3][l_4][l_i][l_j][l_k][l_l])^{\frac{1}{2}} (\gamma_{1234}| h_a(1,2;p)h_b(2,3;q)h_c(3,4;r) |\gamma_{ijkl})$$

$$\times \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & v & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_4 & r & l_l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & u & q \\ m_p & m_u & m_q \end{pmatrix} \begin{pmatrix} r & v & q \\ m_r & -m_v & m_q \end{pmatrix}$$

$$\times \begin{pmatrix} l_1 & p & l_i \\ -m_1 & m_p & m_i \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ -m_2 & m_u & m_j \end{pmatrix} \begin{pmatrix} l_3 & v & l_k \\ -m_3 & m_v & m_k \end{pmatrix} \begin{pmatrix} l_4 & r & l_l \\ -m_4 & -m_r & m \end{pmatrix}, \tag{28}$$

where $C = m_i + m_j + m_3 + m_4$. In this instance m_p, m_q, and m_r are restricted to a single value by the single-particle states and hence the summation over them is not indicated. Also the indices u, v, p, and r are limited to finite values by the triangular inequali-

ties and they in turn restrict the values that q can assume by the remaining 3-j symbols. Thus all sums in Eq. (28) are finite.

If one sets the exponent of r_{34} equal to zero, the sum over r is restricted to $r = 0$; Eq. (28) then reduces to

$$(123| r_{12}^a r_{23}^b |ijk) = \sum_{p,q} \sum_u (-1)^D [u]([l_1][l_2][l_3][l_i][l_j][l_k])^{\frac{1}{2}} (\gamma_{123}| h_a(1,2;p) h_b(2,3;q) |\gamma_{ijk}) \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix}$$

$$\times \begin{pmatrix} l_2 & u & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & q & l_k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & u & q \\ m_p & m_u & m_q \end{pmatrix} \begin{pmatrix} l_1 & p & l_i \\ -m_1 & m_p & m_i \end{pmatrix} \begin{pmatrix} l_2 & u & l_j \\ -m_2 & m_u & m_j \end{pmatrix} \begin{pmatrix} l_3 & q & l_k \\ -m_3 & m_q & m_k \end{pmatrix}, \quad (29)$$

where $D = m_1 + m_2 + m_3$. Each of the sums is again finite as is true for Eq. (28). Finally, if the exponent of r_{23} is set equal to zero, q is restricted to 0 and Eq. (29) becomes

$$(12| r_{12}^a |ij) = \sum_p (-1)^E ([l_1][l_2][l_i][l_j])^{\frac{1}{2}} (\gamma_{12}| h_a(1,2;p) |\gamma_{ij})$$

$$\times \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & p & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & p & l_i \\ -m_1 & m_p & m_i \end{pmatrix} \begin{pmatrix} l_2 & p & l_j \\ -m_2 & -m_p & m_j \end{pmatrix}, \quad (30)$$

where $E = m_1 + m_j$ and the sum is restricted.

IV. DISCUSSION OF THE RESULTS

The results of the five matrix elements are now discussed. This includes illustrations of the restrictions on the sums.

Consider Eq. (30). As pointed out by Racah[11,13] and Calais and Löwdin,[4] this result for $a = -1$ can be put into the form given in Chap. 6 of Condon and Shortley.[14] The result takes the same form regardless of the value of a. The angular dependence is the same for any function of r_{12} with the difference entering through the radial integral $(\gamma_{12}| h_a(1,2;p) |\gamma_{ij})$, which for $a = -1$ is the Slater integral $R^p(12_{ij})$ as given by Condon and Shortley[14] and Slater.[5] If one now uses

$$C^p(l_1 m_1 l_i m_i) = (-1)^{m_1} ([l_1][l_i])^{\frac{1}{2}}$$

$$\times \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & p & l_i \\ -m_1 & m_p & m_i \end{pmatrix}, \quad (31)$$

Eq. (30) can be written as (see Calais and Löwdin, Ref. 4)

$$(12| r_{12}^a |ij) = \sum_p C^p(l_1 m_1 l_i m_i) C^p(l_j m_j l_2 m_2)$$

$$\times (\gamma_{12}| h_a(1,2;p) |\gamma_{ij}) \delta(m_1 + m_2, m_i + m_j). \quad (32)$$

The sum over p is restricted by the triangular inequalities and parity restrictions of the 3-j symbols.

It is possible to express the angular coefficients of Eqs. (25), (27), (28), and (29) in terms of the

$C^p(l_1 m_1 l_i m_i)$ as well as Eq. (30). There may be some advantage to this form in Eqs. (28), (29), and (30), where all summation indices are finite, as one could perhaps then use Slater's tables.[5] However, in Eqs. (25) and (27), where the indices become infinite, there seems to be no advantage.

To illustrate the restrictions on the various summation indices appearing in Eqs. (25), (27), (28), (29), and (30), consider diagonal matrix elements for f electrons. Also take the magnetic quantum numbers of the single-particle states to be zero. Then the index p in Eq. (30) can only assume even values $0 \leq p \leq 6$, while $m_p = 0$. In Eq. (29) $m_p = m_q = m_u = 0$, while p, q, and u can only assume even values between 0 and 6. However, they are subject to the inequality $|p - q| \leq u \leq p + q$. In Eq. (28) $m_p = m_q = m_r = m_u = m_v = 0$, the indices p, u, v, and r can only assume even values between 0 and 6, and the index q is even in the range of $0 \leq q \leq 12$. Note that these indices are subject to $|p - u| \leq q \leq p + u$ and $|r - v| \leq g \leq r + v$. It is seen that for these three matrix elements it is sufficient to use Eq. (3.7.17) of Ref. 7 or Slater's tables,[5] if one wishes to introduce the $C^p(l_1 m_1 l_i m_i)$ coefficients.

Bonham[1] has conjectured that all cases of this kind lead to finite sums. He characterizes the matrix elements by diagrams with dots representing electrons and lines representing interactions. Equation (30) is represented by
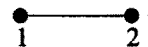
Eq. (29) by
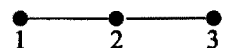
[13] G. Racah, "Group Theory and Spectroscopy," mimeographed notes of 1951 lectures at the Institute for Advanced Study, available as a CERN reprint.

[14] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England, 1953).

and Eq. (28) by
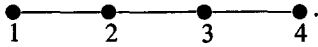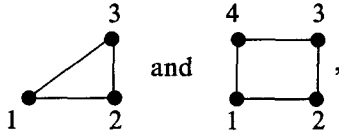

He suggests that, as long as the diagram contains no closed portions, all sums must be finite. This is consistent with the results shown here and with other results worked out by the authors and not shown here.

The two remaining matrix elements of Eqs. (25) and (27) are represented by

respectively. The diagrams are closed and have a more complicated structure insofar as the angular coefficients are concerned. Considering again diagonal matrix elements for f electrons, these equations contain one independent infinite sum. To illustrate, consider Eq. (25). The indices u, v, and x are all even, lie between 0 and 6, and must satisfy the inequality $|u - v| \leq x \leq u + v$. If the magnetic quantum numbers of the single-particle states are all 0, then $m_u = m_v = m_x = 0$. The remaining indices p, q, and r all range from 0 to ∞ subject to the inequalities $|r - x| \leq q \leq r + x$, $|p - u| \leq u \leq p + u$, and $|q - v| \leq p \leq q + v$. Thus if $u = v = x = 0$, one has $p = q = r$, while for $u = 0$, $v = x = 2$, one has $r = p$, $|p - 2| \leq q \leq p + 2$, etc. Similar restrictions would apply to Eq. (27). It is seen that in both Eqs. (25) and (27) there is really a single independent infinite sum due to the triangular inequalities and parity restrictions.

A simpler illustration of Eq. (25) is given by the diagonal matrix element for the configuration $1s^2 2p_0$. Here $u = v = x = 0$, $p = q = r$, $m_p = m_q = m_r$ to give

$$(123| \, r_{12}^a r_{23}^b r_{31}^c \, |ijk) = \sum_p \sum_{m_p} (-1)^{3m_p}[1]$$

$$\times (\gamma_{123}| \, h_a(1, 2; p) h_b(2, 3; p) h_c(3, 1; p) \, |\gamma_{ijk})$$

$$\times \begin{pmatrix} p & 0 & p \\ 0 & 0 & 0 \end{pmatrix}^3 \begin{pmatrix} p & 0 & p \\ -m_p & 0 & m_p \end{pmatrix}^3 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2. \quad (33)$$

On evaluating the 3-j symbols[7] and summing over m_p, this becomes

$$(123| \, r_{12}^a r_{23}^b r_{31}^c \, |ijk) = \sum_p \frac{1}{(2p + 1)^2}$$

$$\times (\gamma_{123}| \, h_a(1, 2; p) h_b(2, 3; p) h_c(3, 1; p) \, |\gamma_{ijk}). \quad (34)$$

Next consider the diagonal matrix element of the configuration $1s^2 2sp_0$ for Eq. (27). This leads to $u = v = y = z = 0$, $p = q = r = s$, and $m_p = -m_q =$

$m_r = m_s$. The matrix element then becomes

$$(1234| \, r_{12}^a r_{23}^b r_{34}^c r_{41}^d \, |ijkl) = \sum_p \sum_{m_p} [1]$$

$$\times (\gamma_{1234}| \, h_a(1, 2; p) h_b(2, 3; p) h_c(3, 4; p) h_d(4, 1; p) \, |\gamma_{ijkl})$$

$$\times \begin{pmatrix} p & 0 & p \\ 0 & 0 & 0 \end{pmatrix}^4 \begin{pmatrix} p & 0 & p \\ m_p & 0 & -m_p \end{pmatrix}^4 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2. \quad (35)$$

On evaluating the 3-j symbols and summing over m_p, the result is

$$(1234| \, r_{12} r_{23} r_{34} r_{41} \, |ijkl) = \sum_p \frac{1}{(2p + 1)^3}$$

$$\times (\gamma_{1234}| \, h_a(1, 2; p) h_b(2, 3; p) h_c(3, 4; p) h_d(4, 1; p) \, |\gamma_{ijkl}). \quad (36)$$

It might be possible to truncate the sums of Eqs. (25) and (27). This depends on the form of the radial integrals. These depend on the power of the interelectronic coordinates and the radial part of the single-particle states. If these radial integrals decrease sufficiently rapidly with increasing values of the indices, it may be possible to obtain sufficient accuracy by terminating the sums. However, a general criterion for this is difficult to realize.

V. CONCLUSIONS

The previous sections illustrate a procedure for the evaluation of the angular coefficients of atomic matrix elements for an arbitrary number of electrons and interelectronic coordinates using single particle states of arbitrary angular momentum. The evaluation is shown for powers of the interelectronic coordinates greater than -1, but in principle any well-behaved function expandable in spherical harmonics can be used.

The tensor operators, when recoupled, completely separate the angular and radial portions. The separation enables one to take advantage of the symmetry inherent in the single-particle product states. This is reflected in the appearance of the 3-j symbols with their triangular inequalities and parity restrictions. The final results are expressed in terms of restricted sums which are easier to use than multiple integral expressions when orbital exponents are varied in energy minimization.

It should also be noted that the effects of couplings of the single-particle states do not appear in the final results. The calculation does not depend on particular coordinate systems, as is the case for multiple products of spherical harmonics. The points mentioned in this paragraph are direct consequences of using the internal symmetry which is independent of the coordinate system.

It is true that the greatest interest has, and will probably continue to be, centered around matrix elements involving one, two, and three interelectronic coordinates. However, there are cases where one needs to go beyond this and include four or more interelectronic coordinates. (See Refs. 1 and 2.) The procedure illustrated here enables one to obtain a complete determination of the angular coefficients for all cases in an efficient manner, regardless of the symmetry of the single-particle states. The authors have evaluated these for several more complicated cases, including that given by Bonham,[2] and in each instance the result can be expressed in the same form. The conjectures put forth by Bonham[1] have been verified for all cases treated.

For those instances which involve one or more closed loops in the diagrams, the expressions contain one or more independent infinite sums. However, as pointed out in Sec. IV, it may be possible to truncate these sums, depending on the form of the radial integrals. Since these radial integrals depend both on the radial parts of the single-particle states and on the function of the radial coordinates introduced from the expansion of the interelectronic coordinates (Sec. IIC), one cannot readily specify a criterion for this truncation.

APPENDIX A. RECOUPLING OF THE MIXED TENSOR OPERATORS

A derivation of Eqs. (16) and (17) is given here. The derivation follows closely the method given in Chap. 4 of Ref. 8. In deriving these, use is made of the following simpler recouplings:

$$\{C_1^p C_1^q\}^x = (-1)^x [x]^{\frac{1}{2}} \begin{pmatrix} p & x & q \\ 0 & 0 & 0 \end{pmatrix} C_1^x, \quad (A1)$$

$$\{C_1^p C_2^p\}^0 \{C_3^q C_4^q\}^0 = \sum_x [x]^{\frac{1}{2}} ([p][q])^{-\frac{1}{2}}$$
$$\times \{\{C_1^p C_4^q\}^x \{C_2^p C_3^q\}^x\}^0, \quad (A2)$$

$$\{C_1^p C_2^p\}^0 \{C_1^q C_2^r\}^x = \sum_{u,v} (-1)^x [u][v][p]^{-\frac{1}{2}}$$
$$\times \begin{pmatrix} q & u & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} q & u & p \\ v & r & x \end{Bmatrix} \{C_1^u C_2^v\}^x, \quad (A3)$$

where $\begin{Bmatrix} q & u & p \\ v & r & x \end{Bmatrix}$ is a 6-j symbol. In (A1), (A2), and (A3) u, v, and x are limited by the appropriate triangular inequalities and parity considerations.

The first of these, Eq. (A1), is the same as Eq. (4.5) of Ref. 8. Since a procedure for deriving it is given there, it is not derived here.

Equation (A2) may be derived as follows: Note that

$$\{C_1^p C_2^p\}^0 \{C_3^q C_4^q\}^0 = \sum_{x,y} ((pq)x, (pq)y, 0| \ (pp)0, (qq)0, 0)$$
$$\times \{\{C_1^p C_4^q\}^x \{C_2^p C_3^q\}^y\}^0, \quad (A4)$$

where $((pq)x, (pq)y, 0 \mid (pp)0, (qq)0, 0)$ are the matrix elements of the transformation from one coupling scheme to the other.[7] These coefficients are related to 9-j symbols by

$$((pq)x, (pq)y, 0 \mid (pp)0, (qq)0, 0) = [x][y]^{\frac{1}{2}} \begin{Bmatrix} p & q & x \\ p & q & y \\ 0 & 0 & 0 \end{Bmatrix}. \quad (A5)$$

In the 9-j symbol x must equal y in order to satisfy the necessary triangular inequality. The 9-j symbol then reduces to a 6-j symbol with one 0 in it[12]:

$$\begin{Bmatrix} p & q & x \\ p & q & x \\ 0 & 0 & 0 \end{Bmatrix} = \frac{(-1)^{p+q+x}}{[x]^{\frac{1}{2}}} \begin{Bmatrix} p & q & x \\ q & p & 0 \end{Bmatrix}. \quad (A6)$$

The 6-j symbol has a simple algebraic value given by[12]

$$\begin{Bmatrix} p & q & x \\ p & q & 0 \end{Bmatrix} = \frac{(-1)^{p+q+x}}{[p][q]^{\frac{1}{2}}}. \quad (A7)$$

On substituting (A5), (A6), and (A7) into (A4), one is led to (A2).

Equation (A3) is derived in a similar manner as follows:

$$\{C_1^p C_2^p\}^0 \{C_1^q C_2^r\}^x = \sum_{u,v} ((pq)u, (pr)v, x \mid (pp)0, (qr)x, x)$$
$$\times \{\{C_1^p C_1^q\}^u \{C_2^p C_2^r\}^v\}^x, \quad (A8)$$

where the transformation coefficients are again related to 9-j symbols, but now with one vanishing argument. This gives

$$((pq)u, (pr)v, x \mid (pp)0, (qr)x, x)$$
$$= ([u][v][x])^{\frac{1}{2}} \begin{Bmatrix} p & q & u \\ p & r & v \\ 0 & x & x \end{Bmatrix}. \quad (A9)$$

As above, the 9-j symbol reduces to a 6-j symbol:

$$\begin{Bmatrix} p & q & u \\ p & r & v \\ 0 & x & x \end{Bmatrix} = \frac{(-1)^{u+r+p+x}}{[p][x]^{\frac{1}{2}}} \begin{Bmatrix} q & u & p \\ v & r & x \end{Bmatrix}. \quad (A10)$$

One now uses (A1) twice on the recoupled operator of (A8) to get

$$\{\{C_1^p C_1^q\}^u \{C_2^p C_2^r\}^v\}^x$$
$$= (-1)^{u+v} ([u][v])^{\frac{1}{2}} \begin{pmatrix} p & u & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p & v & r \\ 0 & 0 & 0 \end{pmatrix} \{C_1^u C_2^v\}^x.$$
$$\quad (A11)$$

If one substitutes (A9), (A10), and (A11) into (A8), the result is (A3).

One may now use (A1), (A2), and (A3) to derive Eqs. (16) and (17). Equation (16) is derived by letting $4 \rightarrow 2$, $2 \rightarrow 3$, and $p \rightarrow r$ in (A2) to give the last two factors on the left side of Eq. (16) as

$$\{C_2^q C_3^q\}^0 \{C_1^r C_3^r\}^0 = \sum_x [x]^{\frac{1}{2}} ([p][q])^{-\frac{1}{2}}$$
$$\times \{\{C_1^r C_2^q\}^x \{C_3^r C_3^q\}^x\}^0. \quad \text{(A12)}$$

On substituting this into $\{C_1^p C_2^p\}^0 \{C_2^q C_3^q\}^0 \{C_1^r C_3^r\}^0$, one is led to

$$\{C_1^p C_2^p\}^0 \{C_2^q C_3^q\}^0 \{C_1^r C_3^r\}^0$$
$$= \sum_x [x]^{\frac{1}{2}} ([q][r])^{-\frac{1}{2}} \{C_1^p C_2^p\}^0 \{\{C_1^r C_2^q\}^x \{C_3^r C_3^q\}^x\}^0. \quad \text{(A13)}$$

The recoupled operator on the right side of (A13) may be simplified by using (A1) and the fact that $\{C_1^p C_2^p\}^0$ is a scalar:

$$\{C_1^p C_2^p\}^0 \{\{C_1^r C_2^q\}^x \{C_3^r C_3^q\}^x\}^0$$
$$= (-1)^x [x]^{\frac{1}{2}} \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix} \{\{C_1^p C_2^p\}^0 \{C_1^r C_2^q\}^x C_3^x\}^0. \quad \text{(A14)}$$

If one now employs (A3) on the recoupled operator on the right of Eq. (A14), the result is

$$\{\{C_1^p C_2^p\}^0 \{C_1^r C_2^q\}^x C_3^x\}^0$$
$$= \sum_{uv} (-1)^x [u][v][p]^{-\frac{1}{2}} \begin{pmatrix} r & u & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} r & u & p \\ v & q & x \end{Bmatrix}$$
$$\times \{\{C_1^u C_2^v\}^x C_3^x\}^0. \quad \text{(A15)}$$

On substituting (A14) and (A15) into (A13), one obtains Eq. (16):

$$\{C_1^p C_2^p\}^0 \{C_2^q C_3^q\}^0 \{C_1^r C_3^r\}^0$$
$$= \sum_{u,v,x} [u][v][x]([p][q][r])^{-\frac{1}{2}}$$
$$\times \begin{pmatrix} r & x & q \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r & u & p \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} q & v & p \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} r & u & p \\ v & q & x \end{Bmatrix}$$
$$\{\{C_1^u C_2^v\}^x C_3^x\}^0. \quad \text{(A16)}$$

Equation (17) may be derived by first using (A2) to recouple the first and third factors of the left side and then using (A3) twice to give the recoupling on the right side.

APPENDIX B. MATRIX ELEMENTS OF THE RECOUPLED TENSOR OPERATORS

The derivations of Eqs. (18) and (19) given below follow directly from the material presented in Sec. 3.6 of Ref. 8. Before showing Eqs. (18) and (19), the matrix element for the mixed tensor operator $\{C_1^p C_2^p\}^0$ is given. Note that, as mentioned in Sec. IIB, this is, to within a factor, the scalar product of two tensors

of degree x^{13}. It's matrix element is

$$((l_1 l_2) LM| \{C_1^p C_2^p\}^0 |(l_i l_j) L'M')$$
$$= (-1)^{p+l_1+l_2+L} \delta_{L,L'} \delta_{M,M'} [p]^{-\frac{1}{2}}$$
$$\times \begin{Bmatrix} l_i & l_j & L \\ l_2 & l_1 & p \end{Bmatrix} (l_1 \| C_1^p \| l_i)(l_2 \| C_2^p \| l_j) \quad \text{(B1)}$$

by using Eqs. (3.33) and (3.36) of Ref. 8. The quantity $(l_1 \| C_1^p \| l_i)$ is the reduced matrix element of the set of operators C_1^p. It's value, given by Eq. (4.4) of Ref. 8, is

$$(l_1 \| C_1^p \| l_i) = (-1)^{l_1} [l_1][l_i])^{\frac{1}{2}} \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix}. \quad \text{(B2)}$$

Substituting into (B1) gives

$$((l_1 l_2) LM| \{C_1^p C_2^p\}^0 |(l_i l_j) L'M')$$
$$= (-1)^{p+l_1+l_i+L} \delta_{L,L'} \delta_{M,M'} ([l_1][l_2][l_i][l_j])^{\frac{1}{2}} [p]^{-\frac{1}{2}}$$
$$\times \begin{pmatrix} l_1 & p & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & p & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} l_i & l_j & L \\ l_2 & l_1 & p \end{Bmatrix}. \quad \text{(B3)}$$

The procedure to be followed for the operators $\{C_1^u C_2^v\}^x C_3^x\}^0$ and $\{\{C_1^u C_4^v\}^x \{C_2^u C_3^v\}^x\}^0$ is quite similar. For the first of these one obtains

$$((l_{12} l_3) LM| \{\{C_1^u C_2^v\}^x C_3^x\}^0 |(l_{ij} l_k) L'M')$$
$$= (-1)^{x+l_{ij}+l_3+L} \delta_{L,L'} \delta_{M M'} [x]^{-\frac{1}{2}} \begin{Bmatrix} l_{ij} & l_k & L \\ l_3 & l_{12} & x \end{Bmatrix}$$
$$\times (l_{12} \| \{C_1^u C_2^v\}^x \| l_{ij})(l_3 \| C_3^x \| l_k). \quad \text{(B4)}$$

The reduced matrix element $(l_{12} \| \{C_1^u C_2^v\}^x \| l_{ij})$ can be expressed as $((l_1 l_2) l_{12} \| \{C_1^u C_2^v\}^x \| (l_i l_j) l_{ij})$ and, from Eq. (3.35) of Ref. 8, becomes

$$((l_1 l_2) l_{12} \| \{C_1^u C_2^v\}^x \| (l_i l_j) l_{ij})$$
$$= ([l_{12}][l_{ij}][x])^{\frac{1}{2}} \begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix} (l_1 \| C_1^u \| l_i)(l_2 \| C_2^v \| l_j). \quad \text{(B5)}$$

If this is substituted into (B4) and (B2) is used three times, one obtains

$$((l_1 l_2) LM| \{\{C_1^u C_2^v\}^x C_3^x\}^0 |(l_i l_j) L'M')$$
$$= (-1)^x \delta_{L,L'} \delta_{M,M'} ([l_1][l_2][l_3][l_i][l_j][l_k][l_{12}][l_{ij}])^{\frac{1}{2}}$$
$$\times \begin{pmatrix} l_1 & u & l_i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_2 & v & l_j \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_3 & x & l_k \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{Bmatrix} l_{ij} & l_k & L \\ l_3 & l_{12} & x \end{Bmatrix} \begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}, \quad \text{(18)}$$

where

$$\alpha = x + L + l_{ij} + l_1 + l_2 \quad \text{and} \quad \begin{Bmatrix} l_1 & l_i & u \\ l_2 & l_j & v \\ l_{12} & l_{ij} & x \end{Bmatrix}$$

is a 9-j symbol.[7]

For the operator $\{\{\mathbf{C}_1^y\mathbf{C}_4^z\}^x\{\mathbf{C}_2^u\mathbf{C}_3^v\}^x\}^0$ (B4) gives

$$((l_{14}l_{23})LM|\{\{\mathbf{C}_1^y\mathbf{C}_4^z\}^x\{\mathbf{C}_2^u\mathbf{C}_3^v\}^x\}^0|(l_{il}l_{jk})L'M')$$

$$= (-1)^{x+l_{il}+l_{23}+L}\delta_{L,L'}\delta_{M,M'}[x]^{-\frac{1}{2}}\begin{Bmatrix}l_{il} & l_{jk} & L \\ l_{23} & l_{14} & x\end{Bmatrix}$$

$$\times (l_{14}\|\{\mathbf{C}_1^y\mathbf{C}_4^z\}^x\|l_{il})(l_{23}\|\{\mathbf{C}_2^u\mathbf{C}_3^v\}^x\|l_{jk}). \quad \text{(B6)}$$

On using (B5) twice and (B2) four times, one obtains

$$((l_{14}l_{23})LM|\{\{\mathbf{C}_1^y\mathbf{C}_4^z\}^x\{\mathbf{C}_2^u\mathbf{C}_3^v\}^x\}^0|(l_{il}l_{jk})L'M')$$

$$= (-1)^\beta \delta_{L,L'}\delta_{M M'}$$

$$\times ([l_1][l_2][l_3][l_4][l_i][l_j][l_k][l_l][l_{14}][l_{23}][l_{il}][l_{jk}][x])^{\frac{1}{2}}$$

$$\times \begin{pmatrix}l_1 & y & l_i \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_2 & u & l_j \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_3 & v & l_k \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_4 & z & l_l \\ 0 & 0 & 0\end{pmatrix}$$

$$\times \begin{Bmatrix}l_{il} & l_{jk} & L \\ l_{23} & l_{24} & x\end{Bmatrix}\begin{Bmatrix}l_1 & l_i & y \\ l_4 & l_l & z \\ l_{14} & l_{il} & x\end{Bmatrix}\begin{Bmatrix}l_2 & l_j & u \\ l_3 & l_k & v \\ l_{23} & l_{jk} & x\end{Bmatrix}, \quad \text{(19)}$$

where $\beta = x + L + l_{il} + l_{23} + l_1 + l_2 + l_3 + l_4$.

APPENDIX C. DERIVATION OF THE ATOMIC MATRIX ELEMENTS

The procedure outlined in Sec. III is illustrated here in deriving Eq. (20). To begin, one expands each power of the product $r_{12}^a r_{23}^b r_{31}^c$ in terms of the tensor operators as illustrated in Sec. IIC. Using Eq. (11) or extending Eq. (15), one obtains

$$r_{12}^a r_{23}^b r_{31}^c = \sum_{p,q,r=0}^{\infty}(-1)^{p+q+r}([p][q][r])^{\frac{1}{2}}$$

$$\times h_a(1,2;p)h_b(2,3;q)h_c(1,3;r)$$

$$\times \{\mathbf{C}_1^p\mathbf{C}_2^p\}^0\{\mathbf{C}_2^q\mathbf{C}_3^q\}^0\{\mathbf{C}_3^r\mathbf{C}_1^r\}^0. \quad \text{(C1)}$$

One may then substitute (C1) into $(123|r_{12}^a r_{23}^b r_{31}^c|ijk)$ and separate the radial from the angular parts:

$$(123|r_{12}^a r_{23}^b r_{31}^c|ijk)$$

$$= \sum_{p,q,r=0}(-1)^{p+q+r}([p][q][r])^{\frac{1}{2}}(\gamma_{123}|h_a h_b h_c|\gamma_{ijk})$$

$$\times (123|\{\mathbf{C}_1^p\mathbf{C}_2^p\}^0\{\mathbf{C}_2^q\mathbf{C}_3^q\}^0\{\mathbf{C}_3^r\mathbf{C}_1^r\}^0|ijk). \quad \text{(C2)}$$

Now substitute Eq. (2) on both sides of the matrix element of the tensor operator, and use Eq. (4). This matrix element then becomes

$$(123|\{\mathbf{C}_1^p\mathbf{C}_2^p\}^0\{\mathbf{C}_2^q\mathbf{C}_3^q\}^0\{\mathbf{C}_3^r\mathbf{C}_1^r\}^0|ijk)$$

$$= \sum_{L,L',l_{12},l_{ij}}([L][L'][l_{12}][l_{ij}])^{\frac{1}{2}}$$

$$\times (-1)^{l_1+l_2+l_i+l_j+m_{12}+m_{ij}+l_3+l_{12}+l_k+l_{ij}+M+M'}$$

$$\times \begin{pmatrix}l_1 & l_2 & l_{12} \\ m_1 & m_2 & -m_{12}\end{pmatrix}\begin{pmatrix}l_i & l_j & l_{ij} \\ m_i & m_j & -m_{ij}\end{pmatrix}$$

$$\times \begin{pmatrix}l_{12} & l_3 & L \\ m_{12} & m_3 & -M\end{pmatrix}\begin{pmatrix}l_{ij} & l_k & L \\ m_{ij} & m_k & -M\end{pmatrix}$$

$$\times ((l_1l_2)l_{12}, l_3, LM|\{\mathbf{C}_1^p\mathbf{C}_2^p\}^0\{\mathbf{C}_2^q\mathbf{C}_3^q\}^0\{\mathbf{C}_3^r\mathbf{C}_1^r\}^0$$

$$\times |(l_il_j)l_{ij}, l_k, L'M'). \quad \text{(C3)}$$

One then uses Eq. (16) to recouple the tensor operator and Eq. (18) to find the matrix element of the recoupled operator. To complete the procedure Eq. (16) and then Eq. (18) are substituted into (C3), which is then substituted into (C2). The result is Eq. (20). Equation (21) may be derived in the same way.

APPENDIX D. DERIVATION OF EQS. (22) AND (23)

Equation (22) is a variation of Eq. (2.19) of Ref. 12, which may be derived from Eq. (2.20) of that reference. Equation (2.20) in the notation used here reads

$$\begin{pmatrix}l_i & l_j & L \\ m_i & m_j & -M\end{pmatrix}\begin{Bmatrix}l_i & l_j & L \\ l_2 & l_1 & x\end{Bmatrix}$$

$$= \sum_{m_1',m_2',m_x}(-1)^{l_2+l_1+x+m_2'+m_1'+m_x}\begin{pmatrix}l_i & l_1 & x \\ m_i & m_1' & -m_x\end{pmatrix}$$

$$\times \begin{pmatrix}l_2 & l_j & x \\ -m_2' & m_j' & m_x\end{pmatrix}\begin{pmatrix}l_2 & l_1 & L \\ m_2' & -m_1' & -M\end{pmatrix}. \quad \text{(D1)}$$

One next multiplies by $\begin{pmatrix}l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix}$, $[L]$, and $(-1)^L$, and sums over L and M:

$$\sum_{L,M}(-1)^L[L]\begin{pmatrix}l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix}$$

$$\times \begin{pmatrix}l_i & l_j & L \\ m_i & m_j & -M\end{pmatrix}\begin{Bmatrix}l_i & l_j & L \\ l_2 & l_1 & x\end{Bmatrix}$$

$$= \sum_{L,M,m_1',m_2',m_x}(-1)^{L+l_2+l_1+x+m_2'+m_1'+m_x}$$

$$\times \begin{pmatrix}l_i & l_1 & x \\ m_i & m_1' & -m_x\end{pmatrix}\begin{pmatrix}l_2 & l_j & x \\ -m_2' & m_j & m_x\end{pmatrix}$$

$$\times \begin{pmatrix}l_2 & l_1 & L \\ m_2' & -m_1' & -M\end{pmatrix}\begin{pmatrix}l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix}. \quad \text{(D2)}$$

The right side of (D2) can be simplified by using the symmetry property[7]

$$\begin{pmatrix}l_2 & l_1 & L \\ m_2' & -m_1' & -M\end{pmatrix}$$

$$= (-1)^{l_1+l_2+L}\begin{pmatrix}l_1 & l_2 & M \\ -m_1' & m_2' & -M\end{pmatrix} \quad \text{(D3)}$$

and the orthogonality relation[12]

$$\sum_{L,M}[L]\begin{pmatrix}l_1 & l_2 & L \\ -m_1' & m_2' & -M\end{pmatrix}\begin{pmatrix}l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix}$$

$$= \delta_{-m_1',m}\delta_{m_2',m_2}. \quad \text{(D4)}$$

Equation (D2) then reads

$$
\begin{aligned}
\sum_{L,M} (-1)^L [L] &\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \\
&\times \begin{pmatrix} l_i & l_j & L \\ m_i & m_j & -M \end{pmatrix} \begin{Bmatrix} l_i & l_j & L \\ l_2 & l_1 & x \end{Bmatrix} \\
= \sum_{m_x} (-1)^{x+m_1+m_2} & \\
&\times \begin{pmatrix} l_i & l_1 & x \\ m_i & -m_1 & -m_x \end{pmatrix} \begin{pmatrix} l_2 & l_j & x \\ -m_2 & m_j & m_x \end{pmatrix}, \quad \text{(D5)}
\end{aligned}
$$

where use has been made of the fact that the single-particle state angular momenta are all integral. For the examples included in this paper the single-particle state angular momenta are assumed fixed. Thus all of the magnetic quantum numbers in (D5) are fixed and the sums over M and m_x can be dropped, giving Eq. (22).

Equation (23) is derived by three applications of Eq. (22). One first expands the 9-j symbol on the left side of Eq. (23) using Eq. (6.43) of Ref. 7:

$$
\begin{Bmatrix} l_1 & l_i & y \\ l_4 & l_l & z \\ l_{14} & l_{il} & x \end{Bmatrix} = \sum_w (-1)^{2w} [w] \begin{Bmatrix} l_1 & l_2 & l_{12} \\ l_{ij} & x & w \end{Bmatrix}
$$
$$
\times \begin{Bmatrix} l_i & l_j & l_{ij} \\ l_2 & w & v \end{Bmatrix} \begin{Bmatrix} u & v & x \\ w & l_1 & l_i \end{Bmatrix}, \quad \text{(D6)}
$$

where $(-1)^{2w} = 1$ since w is an integer. If this is substituted into the left side of Eq. (23),

$$
\begin{aligned}
\sum_{l_{12}, l_{ij} w} (-1)^{12} [l_{12}][l_{jk}][w] &\begin{pmatrix} l_1 & l_2 & l_{12} \\ m_1 & m_2 & -m_{12} \end{pmatrix} \\
\times &\begin{pmatrix} l_i & l_j & l_{ij} \\ m_i & m_j & -m_{ij} \end{pmatrix} \begin{pmatrix} l_{ij} & l_{12} & x \\ m_{ij} & -m_{12} & -m_x \end{pmatrix} \\
\times &\begin{Bmatrix} l_1 & l_2 & l_{12} \\ l_{ij} & x & w \end{Bmatrix} \begin{Bmatrix} l_i & l_j & l_{ij} \\ l_2 & w & v \end{Bmatrix} \begin{Bmatrix} u & v & x \\ w & l_1 & l_i \end{Bmatrix}. \quad \text{(D7)}
\end{aligned}
$$

Using the symmetry property[7]

$$
\begin{pmatrix} l_{ij} & l_{12} & x \\ m_{ij} & -m_{12} & -m_x \end{pmatrix} = \begin{pmatrix} x & l_{ij} & l_{12} \\ -m_x & m_{ij} & -m_{12} \end{pmatrix} \quad \text{(D8)}
$$

on the third 3-j symbol of (D7), one can sum over l_{12} to give

$$
\begin{aligned}
\sum_{l_{ij}, w} (-1)^{w+m_1+m_{ij}} [l_{ij}][w] &\begin{pmatrix} l_i & l_j & l_{ij} \\ m_i & m_j & -m_{ij} \end{pmatrix} \\
\times &\begin{pmatrix} l_1 & x & w \\ m_1 & m_x & -m_w \end{pmatrix} \begin{pmatrix} l_{ij} & l_2 & w \\ -m_{ij} & m_2 & m_w \end{pmatrix} \\
\times &\begin{Bmatrix} l_i & l_j & l_{ij} \\ l_2 & w & v \end{Bmatrix} \begin{Bmatrix} u & v & x \\ w & l_1 & l_i \end{Bmatrix}. \quad \text{(D9)}
\end{aligned}
$$

Repeating this two more times gives the right side of Eq. (23).

High-Energy Phase Shifts Produced by Repulsive Singular Potentials*

A. PALIOV* AND S. ROSENDORFF

Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel

(Received 13 December 1966)

An expansion in inverse powers of the energy (usually fractional powers) of the WKB phase shifts δ_l produced by repulsive potentials, singular at the origin, has been derived. In most cases this expansion is valid for angular momenta $l < l_{max}$, where l_{max} increases with energy. For large l a power series expansion in the "coupling constant" g is developed. The two regions of validity complement each other and sometimes even overlap. The potentials considered have an r^{-p}, $p \geq 1$, singularity and a k^m, $m < 2$, energy dependence. Define $q = p + m - 2$; then for $q > 0$, the case of strong interaction, we get $\delta_l \sim -(gk^q)^{1/p}$, when $k \to \infty$. The constant of proportionality is independent of l. However, the next term in the expansion depends on l. For $q = 0$, the case of intermediate interaction, δ_l becomes independent of energy when $k \to \infty$, but it depends in a complicated way on both g and l. Finally, for $q < 0$, the case of weak interaction, $\delta_l \sim -gk^q(l + \frac{1}{2})^{1-p}$, $p > 1$.

I. INTRODUCTION

IN recent years it has been tried with some success to explain high-energy phenomena, such as $p - p$ and $\pi - p$ scattering, by the optical model[1-3] or the similar potential model[4] approach. In order to apply these models thoroughly, one is interested to know as much as possible about the analytic behavior of the phase shifts $\delta_l(k, g)$ as functions of energy, coupling constants, and angular momenta. Generally, the potentials used are complex functions. The high-energy phase shifts due to potentials, either real or complex, which are not singular at the origin, have been dealt with[5] before. In this paper we discuss the high-energy behavior of the partial-wave scattering phase shifts produced by repulsive potentials which are singular at the origin. Phase shifts produced by singular but complex potentials, like those used in Tiktopoulos's paper, will be discussed in a later communication. Calogero[6] and Bertocchi, Fubini, and Furlan[7] have discussed before the high-energy phase shifts due to a repulsive potential of the kind $V(r) \sim (1/r)p$ with $p > 2$. They found the dominant terms of the expansion of δ_l/k in powers of $k^{-2/p}$, where k is the momentum of the scattering particles. These terms do not depend on the angular momentum l, and therefore their method yields good approxi-

mations to the phase shifts as long as one deals with their numerical calculation only. However, when one seeks the calculation of the scattering amplitude, those terms which depend on l are the most important ones, since it is necessary to sum up all the partial-wave l-dependent amplitudes.

The purpose of the present paper is to derive explicit expressions for the phase shifts at high energy which are valid for all angular momentum states. To be definite, we assume the potential to be of the form

$$V(r, k) \propto g(k) \cdot r^{-p} \cdot F(r), \qquad (1)$$

where $p \geq 1$, the positive function $F(r)$ is either constant or a rapidly decreasing smooth cutoff function with $F(0) \neq 0$, and $g > 0$ is a dimensionless coupling constant which might depend on the energy according to $g(k) = gk^m$. It has been pointed out[8] that, in a Schrödinger potential model, it is reasonable to assume that $V(r, k)$ diverges as $k^2 \to \infty$ at least linearly[9] in k. We therefore pay special attention to the case $m = 1$, which is dealt with in Sec. III. The more general case is discussed in Sec. IV.

The basis of the present calculation of the partial-wave scattering phase shifts is the WKB approximation, which, as has been shown explicitly in Ref. 6, is very good in the high-energy region for potentials of the kind considered here, even for the case of s-wave scattering.

It is well known that the perturbative approach, i.e., a power series expansion of the phase shift in g, breaks down for singular potentials with $p > 2$. The physical reason is simply the fact that, for these potentials with $g < 0$, the scattered particle falls into

* This work is based partly upon the M.S. thesis of the first author submitted to the Department of Physics, Technion-Israel Institute of Technology.

[1] R. Serber, Phys. Rev. Letters 10, 357 (1963); Rev. Mod. Phys. 36, 649 (1964).

[2] A. D. Krisch, *Lectures in Theoretical Physics* (University of Colorado, Boulder, Colorado, 1965), Vol. VII B.

[3] H. H. Aly, D. Lurié, and S. Rosendorff, Phys. Letters 7, 198 (1963).

[4] G. Tiktopoulos, Phys. Rev. 138, B1550 (1965).

[5] S. Rosendorff and S. Tani, Phys. Rev. 128, 457 (1962).

[6] F. Calogero, Phys. Rev. 135, B693 (1964).

[7] L. Bertocchi, S. Fubini, and G. Furlan, Nuovo Cimento 35, 633 (1965).

[8] Y. Nambu and M. Sugawara, Phys. Rev. Letters 10, 304 (1963).

[9] The $m = 1$ case may also be considered as a Schrödinger model with an energy-independent potential in which the relativistic dependence of the mass has been taken care of.

the attractive core. As has been pointed out before,[10] the mathematical reflection of this fact is the divergence of the phase shift as function of g when g approaches the negative axis, $g < 0$. Although the exact phase shift diverges, it is interesting to observe that the phase shift calculated according to the WKB approximation does not diverge under the same circumstances, provided $l > l_{min}$, where l_{min} depends on energy. In other words, for given energy it is always possible to find a whole range of angular momenta for which δ_{WKB} is expandable in a power series of g. This is shown in Sec. II. For small angular momenta $l < l_{max}$, a nonperturbative approach is adopted. For s waves $m = 0$ and $F(r) = 1$, our result coincides with that obtained in Ref. 7.

A simple upper limit on the absolute value of the WKB phase shift is derived in Appendix B. We use throughout this paper $\hbar = c = 1$.

II. THE ρ REPRESENTATION AND THE λ REPRESENTATION

The starting point of our calculation is the WKB approximation:

$$\delta^{WKB}(k, \lambda) = \int_{r_0}^{\infty} \left[k^2 - \frac{\lambda^2}{r^2} - 2MV(r, k) \right]^{\frac{1}{2}} dr$$
$$- \int_{\rho}^{\infty} \left[k^2 - \frac{\lambda^2}{r^2} \right]^{\frac{1}{2}} dr, \quad (2)$$

where M is the mass of the scattering particles, $\lambda = l + \frac{1}{2}$, and $\rho = \lambda/k$ is the classical impact parameter. The lower limit r_0 is the zero of the corresponding integrand. Let us assume now that $m = 1$, and write $MV(r, k) = gkV(r)$. It is then possible for given ρ to expand the above expression in a power series[11] of g/k:

$$\delta = k \sum_{n=1}^{\infty} \left(\frac{g}{k} \right)^n \delta^{(n)}(\rho). \quad (3)$$

We refer to this expression as the ρ representation of the phase shift. Later we discuss expansions in inverse powers (usually fractional powers, but often terms proportional to $\ln k$ appear too) of k where the coefficients depend on λ rather than on ρ. We refer to this as the λ representation of the phase shift. One finds, provided the potential decreases faster than $1/r$ at infinity,

$$\delta^{(1)} = - \int_{\rho}^{\infty} \frac{rV(r)}{(r^2 - \rho^2)^{\frac{1}{2}}} dr, \quad (4)$$

and for $n > 1$,

$$\delta^{(n)} = - \frac{1}{n! \, \rho^{2(n-1)}} \int_{\rho}^{\infty} \frac{r \, dr}{(r^2 - \rho^2)^{\frac{1}{2}}} \left(\frac{dr^3}{dr} \right)^{n-1} V^n(r). \quad (5)$$

The meaning of the operator $(dr^3/dr)^{n-1}$ acting on $V^n(r)$ is

$$\left(\frac{dr^3}{dr} \right)^{n-1} V^n(r)$$
$$= \frac{d}{dr} \left[r^3 \cdots \frac{d}{dr} \left\{ r^3 \frac{d}{dr} \left[r^3 \frac{d}{dr} \left(r^3 V^n(r) \right) \right] \right\} \cdots \right], \quad (6)$$

where d/dr appears $(n - 1)$ times. The derivation of expression (5) is straightforward, although some care has to be taken because of the divergence of the integrand at the lower limit. More detail may be found in Ref. 5, where the first three orders[12] are worked out explicitly. Equation (4) for $\delta^{(1)}$ is the well-known expression for the phase shift as it appears in conventional optical model theory. At first glance the higher-order terms, according to Eq. (3), represent only a small correction to $\delta^{(1)}$ at high energy. Let us remember, however, that in the impact parameter representation of the scattering amplitude one integrates over ρ from zero to infinity. Hence the above statement would be correct provided all the $\delta^{(n)}(\rho)$ were nonsingular functions of ρ. In Ref. 5 it has been shown that this is indeed the case for nonsingular potentials. However, potentials singular at the origin will give rise to $\delta^{(n)}(\rho)$, which are singular at $\rho = 0$. Moreover, the strength of the singularity increases with the order n. We thus conclude that a perturbation expansion of the WKB phase shifts provides an asymptotic expansion in inverse powers of k; for nonsingular potentials this holds for every value of ρ, whereas for singular potentials ($p \geq 1$), this is true for $\rho > \rho_{min}$, where ρ_{min} is determined by the potential and by k.

Let us demonstrate these points by two examples.

Example 1:

$$V(r) = 1/\mu_0 r^2. \quad (7)$$

The parameter μ_0 has been introduced for reason of dimensions.

This function typifies all repulsive singular potentials with $p = 2$ at high energy. We find

$$\delta^{(1)}(\rho) = -\pi/2\mu_0\rho,$$
$$\delta^{(n)}(\rho) = -\frac{\pi}{2} \binom{\frac{1}{2}}{n} \rho \left(\frac{2}{\mu_0\rho^2} \right)^n, \quad n > 1. \quad (8)$$

We thus conclude that for this particular example the ρ representation is valid whenever

$$\rho \geq \rho_{min} = (2g/\mu_0 k)^{\frac{1}{2}}, \quad (9)$$

i.e., ρ_{min} decreases as k increases. For potential

---

[10] F. Calogero and M. B. De Stefano, Phys. Rev. **146**, 1196 (1966). In this paper many references to the problem of scattering on singular potentials are to be found.

[11] In general, we get a power series in gk^{m-2}.

[12] For the relativistic wave equation.

Eq. (7) the phase shift can be calculated in close form (the WKB approximation coincides with the exact expression):

$$\delta = \tfrac{1}{2}\pi k\{\rho - [\rho^2 + (2g/\mu_0 k)]^{\frac{1}{2}}\}. \quad (10)$$

Obviously expansion of the square root in powers of g/k is possible only when condition (9) is satisfied. It is very easy to derive the λ representation for Eq. (10). Introduce $\lambda = \rho k$,

$$\delta = \tfrac{1}{2}\pi\{\lambda - [\lambda^2 + (2gk/\mu_0)]^{\frac{1}{2}}\}, \quad (10')$$

and expand in inverse powers of k:

$$\delta = \frac{\pi}{2}\left(\frac{2gk}{\mu_0}\right)^{\frac{1}{2}}\left[\lambda\left(\frac{\mu_0}{2gk}\right)^{\frac{1}{2}} - \sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}\left(\frac{\mu_0\lambda^2}{2gk}\right)^n\right], \quad (11)$$

which is the λ representation for the potential $g/\mu_0 r^2$.

Equation (11) is valid only if

$$\lambda \leq \lambda_{\max} = (2gk/\mu_0)^{\frac{1}{2}}. \quad (12)$$

For this particular example the validity conditions for the two representations complement each other, as follows from comparison of Eqs. (9) and (12).

Example 2: The repulsive Coulomb potential with cutoff

$$V(r) = \begin{cases} 1/r, & r \leq R, \\ 0, & r > R. \end{cases} \quad (13)$$

Here again it is possible to calculate the WKB phase shift in close form:

$$\delta = \lambda \sin^{-1}\frac{g}{(g^2 + \lambda^2)^{\frac{1}{2}}} + g[\ln(g^2 + \lambda^2)^{\frac{1}{2}} - 1]$$
$$- g \ln 2kR. \quad (14)$$

We realize immediately that the λ representation consists of two terms only: one which is independent of k and one which behaves like $\ln kR$. Hence for this particular potential the λ representation is valid for every value of λ. On the other hand, the ρ representation is valid only for

$$\rho > g/k. \quad (15)$$

III. PHASE SHIFTS DUE TO POTENTIALS WITH FIRST- AND SECOND-ORDER SINGULARITIES

In this section we continue to assume that the potential is proportional to the momentum k, i.e., $m = 1$.

A. Simple Pole Potentials

Most of this part deals with the detailed derivation of the WKB phase shift in the λ representation due to the Yukawa potential,

$$V(r, k) = (gk/Mr)e^{-\mu r}. \quad (16)$$

We define

$$t = \mu/k. \quad (17)$$

Thus, according to Eq. (2), we get

$$\delta(k, \lambda) = \int_{z_1(t)}^{\infty}\left[1 - \frac{\lambda^2}{x^2} - 2g\frac{e^{-tx}}{x}\right]^{\frac{1}{2}}dx$$
$$- \int_{\lambda}^{\infty}\left[1 - \frac{\lambda^2}{x^2}\right]^{\frac{1}{2}}dx, \quad (18)$$

where $z_1(t)$ is the zero of the first integrand and the integration variable $x = kr$ has been introduced. Now, as

$$y(x) = (\lambda^2/x^2) + 2g(e^{-tx}/x) \leq 1 \quad (19)$$

for the whole range of integration, it is possible to expand the integrand of the first integral in a power series[13] of y.

We thus obtain

$$\delta(k, \lambda) = \frac{\pi\lambda}{2} - z_1(t) + \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}\lambda^{2n}$$
$$\times \sum_{s=0}^{n}\binom{n}{s}\left(\frac{2g}{\lambda^2}\right)^s\int_{z_1(t)}^{\infty}\frac{e^{-stx}}{x^{2n-s}}dx. \quad (20)$$

We are interested in the expansion of δ for small values of t. Repeated integration by parts of the integral gives rise to the following expansion, assuming $\nu > 1$:

$$\int_z^{\infty}\frac{e^{-stx}}{x^\nu}dx = \frac{e^{-stz}}{(\nu-1)!}\sum_{r=1}^{\nu-1}(\nu-1-r)!\frac{(-st)^{r-1}}{z^{\nu-r}}$$
$$- \frac{(-st)^{\nu-1}}{(\nu-1)!}E_1(-stz), \quad (21)$$

where the expansion of the exponential integral function is given by

$$E_1(-x) = \ln\gamma_0 x + \sum_{k=1}^{\infty}\frac{(-x)^k}{k\cdot k!}. \quad (22)$$

$\ln\gamma_0 = 0.577\cdots$ is Euler's constant. Furthermore, we need the behavior of $z_1(t)$ for small t. By Eq. (18) we have

$$1 - (\lambda^2/z_1^2) - 2g(e^{-tz_1}/z_1) = 0. \quad (23)$$

Expanding $z_1(t)$ in a power series of t,

$$z_1(t) = \alpha + \beta t + \gamma t^2 + \cdots, \quad (24)$$

and comparing term by term, we find for the first two ones

$$\alpha = g + (g^2 + \lambda^2)^{\frac{1}{2}},$$
$$\beta = -\frac{g}{(g^2 + \lambda^2)^{\frac{1}{2}}}[g + (g^2 + \lambda^2)^{\frac{1}{2}}]. \quad (25)$$

---

[13] It should be noted that the power series which represents the function $(1 - y(z))^{\frac{1}{2}}$ at $|y(z)| < 1$ is absolutely convergent at all boundary points of the circle of convergence. See, e.g., K. Knopp, *Infinite Sequences and Series* (Dover Publications, Inc., New York, 1956), p. 140. Hence the above series is uniformly convergent, and interchange of summation and integration is permissible. Obviously, the series of Eq. (20) converges absolutely.

Now it is easy to obtain, from Eqs. (20), (21), and (22), the first two dominating terms for small t:

$$\delta(k, \lambda) = g \ln (\gamma_0 \alpha t) + \frac{\pi\lambda}{2} - \alpha - \frac{\lambda^2}{2\alpha}$$

$$+ \sum_{n=2}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \lambda^{2n} \sum_{s=0}^{n} \binom{n}{s} \left(\frac{2g}{\lambda^2}\right)^s (2n - s - 1)^{-1} \alpha^{-2n+s+1}$$

$$+ \text{ higher order terms in } t. \tag{26}$$

In order to evaluate the infinite sum we first transform the sum over s into an integral.

The following formula is easily proved:

$$\sum_{s=0}^{n} \binom{n}{s} \frac{u^s}{2n - s - 1} = u^{2n-1} \int_u^\infty \left(\frac{1 + x}{x^2}\right)^n dx. \tag{27}$$

Thus, calling $u = 2g\alpha/\lambda^2$, Eq. (26) becomes

$$\delta(k, \lambda) = g \ln (\gamma_0 \alpha t) + \frac{\pi\lambda}{2} - \alpha - \frac{\lambda^2}{2\alpha}$$

$$+ \frac{\lambda^2}{2g} \lim_{x \to \infty} \left[\int_u^x \left(1 - \frac{1 + x}{x^2} \frac{4g^2}{\lambda^2}\right)^{\frac{1}{2}} dx \right.$$

$$\left. - \sum_{n=0}^{1} (-1)^n \binom{\frac{1}{2}}{n} \int_u^x \left(\frac{1 + x}{x^2} \left(\frac{2g}{\lambda}\right)^2\right)^n dx \right] + \cdots,$$

$$= g \ln \tfrac{1}{2}\gamma_0 t + \lambda \sin^{-1} \frac{g}{(g^2 + \lambda^2)^{\frac{1}{2}}}$$

$$+ g[\ln (g^2 + \lambda^2)^{\frac{1}{2}} - 1] + \cdots. \tag{28}$$

This result proves, as expected, that the dominating term at high energy is identical with the phase shift due to a Coulomb potential. Comparison of the above result with Eq. (14) shows that the "range" of the Yukawa potential is equal to $(\gamma_0 \mu)^{-1}$. The foregoing derivation of the leading term shows that in principle it is not difficult (although the algebra becomes more and more involved) to derive the expansion in t to any order desired. We have calculated the phase shift up to the second order in t:

$$\delta(k, \lambda) = (a_0 + a_1 t + a_2 t^2 + \cdots)$$

$$+ (b_0 + b_1 t + \cdots) \ln \tfrac{1}{2}\gamma_0 t. \tag{29}$$

a_0 and b_0 are given by Eq. (28), and

$$a_1 = -g^2 \ln 2(g^2 + \lambda^2)^{\frac{1}{2}},$$

$$b_1 = -g^2,$$

$$a_2 = \tfrac{1}{2}g\lambda^2 \left(\frac{\lambda^2}{g^2 + \lambda^2} - \frac{3}{2}\right) + \tfrac{9}{4}g^3 \ln 3 \tag{30}$$

$$+ \tfrac{1}{8}g(9g^2 + \lambda^2) \ln (g^2 + \lambda^2)^{\frac{1}{2}},$$

$$b_2 = \tfrac{1}{4}g(9g^2 + \lambda^2).$$

Naturally, the question arises for what values of λ is the above expansion valid. This is a difficult problem, and no attempt is made here to answer this

question rigorously. We might obtain a crude estimate as to what the upper limit on λ is by requiring that in the $\ln t$ expansion the third term should be smaller than the second term. We find

$$\lambda < [(4g/\mu)k]^{\frac{1}{2}}. \tag{31}$$

Essentially the same condition on λ is found by requiring that the a series should be descending. We have also calculated the general term in the $\ln t$ series; it is, $n > \nu \geq 0$,

$$(-1)^{n+1} \binom{\frac{1}{2}}{n} \binom{n}{n - \nu} \frac{(\nu - n)^{n+\nu-1}}{(n + \nu - 1)!} (2g)^{n-\nu}\lambda^{2\nu} t^{n+\nu-1} \ln t. \tag{32}$$

The ratio of two adjacent terms for given $\nu, n \to \infty$, is given by $-2egt$. Thus the high-energy condition is $k > 2e\mu g$. More details are to be found in Appendix A. It has been pointed out in Sec. II that the ρ representation should be used for large values of λ. Comparison in Eq. (3) of the second with the first term again yields the high-energy condition, whereas the third to the second term gives

$$\lambda > \approx (gk/10\mu)^{\frac{1}{2}}. \tag{33}$$

Thus the two conditions on λ, Eqs. (31) and (33), overlap.

The foregoing method is certainly not limited to the Yukawa potential. For example, for the potential

$$V(r) = e^{-(\mu r)^2}/r \tag{34}$$

we find that the leading term of the phase shift is identical with the corresponding one for the Yukawa potential, except that the range of this potential "seen" by a high-energy particle is longer by $(\gamma_0)^{\frac{1}{2}}$ than the corresponding Yukawa range. However, the higher-order terms are different for the two potentials. The dependence on t is identical with Eq. (29):

$$\delta(k, \lambda) = (a_0 + a_1 t + \cdots)$$

$$+ (b_0 + b_1 t + \cdots) \ln \tfrac{1}{2}(\gamma_0)^{\frac{1}{2}} t. \tag{35}$$

We have calculated the following coefficients:

$$a_0 = \lambda \sin^{-1} \frac{g}{(g^2 + \lambda^2)^{\frac{1}{2}}} + g[\ln (g^2 + \lambda^2)^{\frac{1}{2}} - 1],$$

$$b_0 = g,$$

$$b_1 = 0, \tag{36}$$

$$b_2 = -\tfrac{1}{2}g(3g^2 + \lambda^2).$$

B. Potentials with Second-Order Singularity

Next we come to the problem of potentials with second-order singularity $(p = 2)$. Take, for example, the potential

$$V(r, k) = (gk/M\mu_0 r^2)e^{-\mu r}. \tag{37}$$

Then the same procedure which leads to Eq. (20) for the Yukawa potential gives for the phase shift in the present case $(t = \mu/k)$

$$\delta(k, \lambda) = \frac{\pi\lambda}{2} - z_2(t) + \sum_{n=1}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n \lambda^{2n}$$
$$\times \sum_{s=0}^{n} \binom{n}{s} \left(\frac{2g\mu}{\lambda^2 t}\right)^s \int_{z_2(t)}^{\infty} \frac{e^{-stx}}{x^{2n}} dx, \quad (38)$$

where, for simplicity, we call $g/\mu_0 \to g$.

$z_2(t)$ is obviously determined by

$$1 - (\lambda^2/z_2^2) - (2g/tz_2^2)e^{-tz_2} = 0,$$

from which the expansion for small values of t follows:

$$z_2(t) = \left(\frac{2g\mu}{t}\right)^{\frac{1}{2}} - g\mu + \frac{\lambda^2 + 3(g\mu)^2}{2(2g\mu)^{\frac{1}{2}}} t^{\frac{1}{2}} + \cdots. \quad (39)$$

We use again Eq. (21) for the expansion of the integral. It is easily verified that to a particular term in the t expansion only certain values of the index s contribute (for the first two terms only $s = n$). This is contrary to the Yukawa case where we had to sum over all values of s, $0 \le s \le n$, and thus Eq. (27) had to be used. We have found the following expansion of the phase shift for potential (37) (some details of its derivation are given in Appendix A):

$$\delta(k, \lambda) = \left(\frac{a_0}{t^{\frac{1}{2}}} + a_1 + a_2 t^{\frac{1}{2}} + \cdots\right)$$
$$+ (b_0 + b_1 t + \cdots) \ln t, \quad (40)$$

where

$$\begin{aligned}
a_0 &= -(\tfrac{1}{2}g\mu\pi^2)^{\frac{1}{2}}, \\
b_0 &= -\tfrac{1}{2}g\mu, \\
a_1 &= \tfrac{1}{2}\pi\lambda - \tfrac{1}{2}g\mu \ln (g\mu\gamma_0^2/2e^2), \\
b_1 &= -\tfrac{1}{3}(g\mu)^2.
\end{aligned} \quad (41)$$

We have also calculated the expansion of the phase shift for the potential

$$V(r, k) = (gk/M\mu_0 r^2)e^{-(\mu r)^2}. \quad (42)$$

The result is

$$\delta(k, \lambda) = (a_0/t^{\frac{1}{2}}) + a_1 + a_2 t^{\frac{1}{2}} + \cdots, \quad (43)$$

with

$$\begin{aligned}
a_0 &= -(\tfrac{1}{2}g\mu\pi^2)^{\frac{1}{2}}, \\
a_1 &= \tfrac{1}{2}\pi\lambda + \pi^{\frac{1}{2}}g\mu, \\
a_2 &= -\pi\lambda^2/4(2g\mu)^{\frac{1}{2}},
\end{aligned} \quad (44)$$

where again $g/\mu_0 \to g$ has been used. Note that the expansion of δ for $V = 1/\mu r^2$, Eq. (11), agrees, as it should, with Eqs. (41) and (44), provided we put $\mu = 0$. It should also be mentioned that for these potentials, having the same singularity at the origin, the dominating terms in the high-energy expansion

of δ are identical. Note also that for potential (42) the phase shift has no $\ln t$ dependence. The reason for this is discussed briefly in Sec. IVB.

As to the upper limit on t and λ we find the following. A general term in the $\ln t$ series of Eq. (40) is given by

$$(-1)^n \binom{\frac{1}{2}}{n} \frac{2^{n-\nu-1}}{(2n-1)!} \binom{n}{n-\nu} (n-\nu)^{2n-1}$$
$$\times g^{n-\nu}\lambda^{2\nu} t^{n+\nu-1} \ln t, \quad (45)$$

where $n > \nu \ge 0$. It turns out that the ratio of two adjacent terms for constant ν when $n \to \infty$ is equal to $\tfrac{1}{2}g\mu te^2$. Thus, irrespective of the value of λ and ν, the high-energy condition for potential (37) is

$$k > \tfrac{1}{2}e^2 g\mu^2. \quad (46)$$

For potential (42) there is no $\ln t$ series. However, a condition very similar to (46) is found from (44) by putting $\lambda = 0$. As to $\lambda_{\max}$, no exact condition has been found. A glance at Eqs. (41) and (44) shows that it should be close to that found for the $1/r^2$ potential, i.e.,

$$\lambda < (2gk/\mu_0)^{\frac{1}{2}}. \quad (12)$$

IV. GENERAL CASE

In this section we assume a more general energy-dependence for the potential. In Eq. (2) we put

$$MV(r, k) = gk^m V(r) \quad (47)$$

with[14] $m < 2$ (including negative values). The coupling constant g is dimensionless, and $V(r)$ has the following form:

$$V(r) = (1/\mu_0^q r^p)F(r) \quad (48)$$

with $p > 0$. The parameter μ_0 has been introduced for reasons of dimensions. The power q is given by

$$q = p + m - 2. \quad (49)$$

We classify[4] the interaction according to the following three categories:

$$\begin{aligned}
q &> 0, \quad \text{strong interaction;} \\
q &= 0, \quad \text{intermediate interaction;} \quad (50) \\
q &< 0, \quad \text{weak interaction.}
\end{aligned}$$

In the following two subsections we outline the general behavior of the phase shifts as functions of the energy for the above three cases, with two different "cutoff" functions

$$F(r) = \begin{cases} 1, & p > 1, \\ e^{-\mu r}. \end{cases} \quad (51)$$

---

[14] According to Eq. (3) and Footnote 11, the p representation becomes meaningless for $m \ge 2$.

A. $F(r) = 1$

We first deal with the case of strong interaction.

The lower limit $z_p(t) = kr_0$, where r_0 is the zero of the integrand in Eq. (2), is determined by

$$1 - (\lambda^2/z_p^2) - (2g/t^q z_p^p) = 0,$$

where $t = \mu_0/k$. If we define

$$\tau = \lambda^2 (t^q/2g)^{2/p}, \tag{52}$$

we find for small τ

$$z_p = (\lambda/\tau^{\frac{1}{2}})(1 + a_1\tau + a_2\tau^2 + \cdots) \tag{53}$$

with

$$a_1 = 1/p,$$

$$a_2 = (1/2p^2)(p - 3).$$

It is easy to show that the coefficient of the jth order is

$$a_j = (1/p^j)P_j(p),$$

where $P_j(p)$ is some polynomial in p. We note that the leading term of r_0, which classically is the shortest possible distance between the interacting particles, is determined by g and not by λ. The reason is that, in the present case, the interaction is "stronger" than the centrifugal barrier; thus the latter becomes important only at an energy which is not extremely high; in other words, only the higher-order terms in (53) are λ dependent. The phase shift is calculated by the same method as outlined in Sec. III. We find, taking (53) into account,

$$\delta = \frac{\pi\lambda}{2} - z_p + \sum_{n=1}^{\infty}(-1)^n \binom{\frac{1}{2}}{n}\lambda^{2n}$$

$$\times \sum_{s=0}^{n}\binom{n}{s}\left(\frac{2g}{\lambda^2 t^q}\right)^s \int_{z_p}^{\infty}\frac{dx}{x^{2n+(p-2)s}},$$

$$= \frac{\pi\lambda}{2} - z_p\left[1 - \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}\sum_{s=0}^{n}\binom{n}{s}\tau^{n-s}\right.$$

$$\left.\times\frac{(1 + a_1\tau + \cdots)^{-2n-(p-2)s}}{2n + (p-2)s - 1}\right]. \tag{54}$$

We see immediately that the expression in the parentheses is expandable in an ascending power series in τ. Thus the phase shift is of the following form:

$$\delta = \frac{\pi\lambda}{2} - \frac{\lambda}{\tau^{\frac{1}{2}}}\sum_{i=0}A_i\tau^i, \tag{55}$$

which, by (52), gives the general dependence of δ on k, g, and λ. The leading term is λ independent. We have calculated A_0 and A_1 explicitly; assuming $p > 1$, we

find

$$A_0 = 1 - \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}\frac{1}{pn - 1} = \frac{1}{2}\int_1^{\infty}\frac{p\,dx}{x^p(1 - x^{-p})^{\frac{1}{2}}}$$

$$= \frac{1}{2}B(1 - 1/p, \frac{1}{2}), \tag{56}$$

$$A_1 = a_1\left(1 - \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}\frac{1}{pn - 1}\right)$$

$$+ \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}\frac{a_1 pn}{pn - 1} - \sum_{n=1}^{\infty}(-1)^n\binom{\frac{1}{2}}{n}$$

$$\times\frac{n}{pn - p + 1} = \frac{1}{2p}B(1/p, \frac{1}{2}), \tag{57}$$

where the Euler function $B(\mu, \nu)$ is defined by

$$B(\mu, \nu) = \Gamma(\mu)\Gamma(\nu)/\Gamma(\mu + \nu).$$

Note that the special case $p = 2$, $m = 1$ has been treated in Sec. II. For the special case of an energy-independent potential ($m = 0$) and s wave ($\lambda = \frac{1}{2}$), the above expression has been derived[15] in Ref. 7. Obviously, Eq. (55) is valid for small values of τ only. In order to know the exact radius of convergence one has to evaluate all the coefficients A_i, which is rather a difficult task. A rough estimate of the upper limit on τ or on λ is obtained by comparing the first two terms. Thus

$$\tau < A_0/A_1$$

or

$$\lambda < C_p(g/t^q)^{1/p}, \tag{58}$$

where

$$C_p = 2^{1/p}[pB(1 - 1/p, \frac{1}{2})/B(1/p, \frac{1}{2})]^{\frac{1}{2}}.$$

We find, e.g., $C_3 = 1.70$; $C_{10} = 1.45$.

In order to obtain the high-energy expression for large values of λ, we apply the ρ representation. Inserting Eq. (48) with $F = 1$ into Eqs. (3) and (5), we find

$$\delta^{(n)} = (-1)^n(1/\mu_0^{qn}\rho^{pn-1})D_n, \tag{59}$$

where

$$D_n = \begin{cases} \dfrac{\pi}{2^{(p-1)n}}\dfrac{(pn - 2)!}{n!\left(\dfrac{(p-2)n}{2}\right)!\left(\dfrac{pn - 2}{2}\right)!}, & pn \text{ even}, \\[20pt] 2^{(p-1)n-2}\dfrac{\left(\dfrac{pn - 3}{2}\right)!\left(\dfrac{(p-2)n - 1}{2}\right)!}{n!\,((p-2)n)!}, & \\[15pt] & pn \text{ odd}. \end{cases} \tag{60}$$

---

[15] There is a slight discrepancy between our value of A_1 and the corresponding one in Ref. 7. This is probably due to the fact that we use the WKB approximation, whereas in Ref. 7 an improved WKB approximation has been used.

The ratio of two adjacent terms for $n \to \infty$ becomes

$$\lim_{n \to \infty} \left| \frac{\delta^{(n)}}{\delta^{(n-1)}} \right| = \frac{p^{p/2}}{(p-2)^{(p-2/2)}} \frac{1}{\mu_0^q \rho^p}.$$

Therefore Eq. (3) converges if

$$\lambda > [p/(p-2)]^{\frac{1}{2}}(p-2)^{1/p}(g/t^q)^{1/p}. \tag{61}$$

Comparison with the upper limit, Eq. (58), obtained by the λ representation, shows that the two representations complement each other.

Next we come to the intermediate interaction. Putting $q = 0$ into Eq. (55), we get

$$\delta = \frac{\pi \lambda}{2} - (2g)^{1/p} \sum_{i=0} A_i \left(\frac{\lambda^2}{(2g)^{2/p}} \right)^i, \tag{62}$$

which is an energy-independent ascending power series in λ. That the WKB phase shift for intermediate interaction is independent of energy follows, of course, directly from the original WKB expression, Eq. (2), which yields the following expression:

$$\delta = \frac{\pi \lambda}{2} - z_p + \int_0^{1/z_p} \frac{dx}{x^2} [(1 - \lambda^2 x^2 - 2gx^p)^{\frac{1}{2}} - 1]. \tag{63}$$

The upper limit z_p^{-1} is equal to the zero of the square root. The radius of convergence of (62) is obviously given by Eq. (58) with $q = 0$. In order to obtain an explicit expression of δ beyond that value, we apply again the ρ representation. Its application, it should be noted, is independent of the value of q. Equation (59) together with (60) thus represent a general term of expansion of δ for large λ. The lower limit on λ follows from (61) with $q = 0$. Therefore, the ρ representation is complementary to Eq. (62).

The explicit expression of δ in the case of the weak interaction is obtained from Eqs. (55) and (52) by the transformation $q \to -q$. This, of course, is not a high-energy expansion. For given energy, it is an expansion for small λ, its upper limit being given again by Eq. (58) with $q \to -q$, i.e., the higher the energy, the smaller the radius of convergence. For large values of λ Eqs. (59) and (60) of the ρ representation apply, its lower limit being given by (61) with $q \to -q$. By use of Eq. (3), we see that the nth-order term is proportional to

$$kg^n k^{(m-2)n}(1/\rho^{pn-1}).$$

By the use of $\rho = \lambda/k$, this becomes

$$g^n k^{qn}/\lambda^{pn-1}.$$

As in the present case $q < 0$, this is nothing more than a general term of the λ representation. We thus conclude that, for potentials of the form $1/r^p$ in the case of weak interaction, the two representations coincide.

B. $F(r) = e^{-\mu r}$

The same method as outlined in the foregoing sections applies here too. However, in the present case, the algebra is much more involved than in the case $F \equiv 1$. We give here only the general dependence of the phase shift on $t = \mu/k$. The lower limit of the integral $z_p(t)$ is determined by

$$1 - (\lambda^2/z_p^2) - (2g/z_p^p t^q)e^{-z_p t} = 0,$$

where we have put $g(\mu/\mu_0)^q \to g$.

We find for small t

$$z_p(t) = (2g/t^q)^{1/p} P_1(t^{(p-q)/p}), \tag{64}$$

where P_1 is an ascending power series with $P_1(0) = 1$. Substituting this into the expression of the phase shift

$$\delta = \frac{\pi \lambda}{2} - z_p + \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \lambda^{2n}$$
$$\times \sum_{s=0}^n \binom{n}{s} \left(\frac{2g}{\lambda^2 t^q} \right)^s \int_{z_p}^{\infty} \frac{e^{-stx}}{x^{2n+(p-2)s}} dx, \tag{65}$$

we obtain the following dependence of δ on t:

$$\delta = \frac{\pi \lambda}{2} + t^{-q/p} P_2(t^{(p-q)/p}) P_3(t^{2q/p})$$
$$+ [\ln t + P_4(t^{(p-q)/p})] \sum_{n=1}^{\infty} \sum_{s=1}^n a_{n,s}^{(st.)} t^{2n-ms-1}. \tag{66}$$

The three functions P_i are ascending power series with $P_2(0)P_3(0)P_4(0) \neq 0$. It is easily verified that the term $t^{-q/p} P_2(t^{(p-q)/p}) P_3(0)$ arises from $s = n$ in (65), the rest of the series coming from $s < n$. The logarithmic term and P_4 arise from the Ei function [see Eq. (21)]. We have explicitly calculated only two terms: the leading term

$$P_2(0)P_3(0) = -\tfrac{1}{2}(2g)^{1/p} B(1 - (1/p), \tfrac{1}{2}),$$

which is identical with the leading term of $\delta(V \sim 1/r^p)$, Eq. (55), and the first term of the $\ln t$ series, $a_{1,1}^{(st.)}$,

$$a_{1,1}^{(st.)} = g \frac{(-1)^{p-1}}{p!} (2 - m).$$

More explicit results have been derived in Sec. IIIB for the special case $p = 2$, $m = 1$.

For the sake of completeness, we also give here the results for the intermediate and the weak interactions.

In the intermediate case the lower limit is determined by

$$z_p^p = \lambda^2 z_p^{p-2} + 2ge^{-tz_p}.$$

Thus, for small t,

$$z_p = \alpha + \beta t + \gamma t^2 + \cdots. \tag{67}$$

The leading term α depends on both λ and g. The

physical reason is that for the intermediate case the interaction and the centrifugal barrier are of "equal" strength. Inserting Eq. (67) in Eq. (65) with $q = 0$, we get the general form of the high-energy expansion of the phase shift:

$$\delta = \sum_{j=0} a_j^{(\text{int})} t^j + \sum_{n=1}^{\infty} \sum_{s=1}^{n} a_{n,s}^{(\text{int})} t^{2n-ms-1} \ln t. \quad (68)$$

We have calculated the first term of each series:

$$a_0^{(\text{int})} = \frac{\pi\lambda}{2} - \alpha + \delta_{p,1} g \ln \gamma_0 \alpha$$
$$+ \int_0^{1/\alpha} \frac{dx}{x^2} [(1 - \lambda^2 x^2 - 2gx^p)^{\frac{1}{2}} - 1 + \delta_{p,1} gx], \quad (69)$$

$$a_{1,1}^{(\text{int})} = g(-1)^{p-1}/(p-1)!. \quad (70)$$

As expected, we see that, for $p > 1$, the leading term $a_0^{(\text{int})}$ is identical to the phase shift due to the corresponding potential without cutoff function ($F = 1$), Eq. (63). For $p = 1$, $a_0^{(\text{int})}$ is equal to the phase shift due to the Coulomb potential with range $(\gamma_0\mu)^{-1}$. This has been shown in Sec. IIIA, where the special case of the Yukawa potential $p = m = 1$ has been calculated explicitly.

The expressions Eqs. (66) and (68) for the strong and intermediate interactions, respectively, are valid for small λ. For large λ the ρ representation should be employed. The lower and upper limits on λ for the two representations have not been evaluated.

For the weak interaction the lower limit is determined by

$$z_p^p = \lambda^2 z_p^{p-2} + 2gt^{-q}e^{-tz_p},$$

where $q < 0$. For small t we find

$$z_p = \lambda + t^{-q} P_0(t), \quad (71)$$

where $P_0(t)$ is an ascending power series of t with $P_0(0) = g/\lambda^{p-1}$. From this expression it follows that, for $k \to \infty$, the shortest distance of the interacting particles is equal to λ/k; in other words, in the case of the weak interaction the centrifugal barrier prevails at infinite energy. Inserting Eq. (71) into Eq. (65), we obtain the general dependence of the phase shift on energy:

$$\delta = t^{-q} \sum_{j=0} a_j^{(w)} t^j + \sum_{n=1}^{\infty} \sum_{s=1}^{n} a_{n,s}^{(w)} t^{2n-ms-1} \ln t. \quad (72)$$

Again we have calculated the first term of each series:

$$a_0^{(w)} = -\frac{g}{\lambda^{p-1}} \int_0^1 \frac{x^{p-2} dx}{(1 - x^2)^{\frac{1}{2}}},$$
$$= -\frac{g}{2\lambda^{p-1}} B\left(\frac{p-1}{2}, \frac{1}{2}\right), \quad p > 1,$$

$$a_0^{(w)} = -g \ln (2/\gamma_0\lambda), \quad p = 1, \quad (73)$$

$$a_{1,1}^{(w)} = g[(-1)^{p-1}/(p-1)!], \quad p \geq 1. \quad (74)$$

It is easily verified that for $p > 1$ the leading term $a_0^{(w)} t^{-q}$ coincides with the value obtained from the ρ representation, namely, $gk^{m-1}\delta^{(1)}(\rho)$ [see Eqs. (3) and (4)]. According to the discussion at the end of Sec. IVA, this is just the leading term of the high-energy expansion of the phase shift with $F(r) = 1$.

We thus conclude that for $p > 1$ the behavior of the phase shift, when $k \to \infty$, is solely determined by the strength of the singularity of the potential at the origin and by its dependence on k.

It is easy to be convinced that the expansion Eq. (72) is valid for large values of λ only. However, it is not identical with the ρ representation, as happens for the corresponding case of $F = 1$. This follows simply from the fact that expansion of any expression of $gt^{-q}(e^{-xt}/x^p)$ in powers of g must be different from expansion in powers of t^{-q}. Only when $e^{-xt} \equiv 1$ are the two identical.

Lastly, the appearance or nonappearance of $\ln t$ depends on the behavior of the integral [see Eq. (65)]

$$\int_{tz_p}^{\infty} \frac{(F(x))^s dx}{x^{2n+(p-2)s}}, \quad \begin{aligned} n &= 1, 2, \cdots, \\ s &= 0, 1, \cdots, n, \end{aligned}$$

which, in turn, depends on the value of p and the behavior of $F(x)$. It is easily verified by integration by parts that if p is an even number and if $F(x)$ is an even function of x, no logarithmic dependence of the lower limit tz_p can arise. We thus conclude that the appearance of $\ln t$ is only possible if the potential $F(r)/r^p$ is not an even function of r. A demonstration of the nonappearance of $\ln t$ has been given in Sec. IIIB, where we have calculated the phase shift for the even potential $e^{-(\mu r)^2}/r^2$. Also, if $F(r)$ is a polynomial, no logarithmic dependence on energy can appear, irrespective of the value of p.

APPENDIX A

Here we show in some detail several derivations of the high-energy expansion of the WKB phase shifts discussed in Sec. III.

1. *The Yukawa potential.* In the text we have evaluated the first two terms of the a series, Eq. (29). Higher-order terms are calculated by the same method; however, the algebra rapidly becomes complicated. On the other hand, a general term of the b series is easily derived as follows: Start with Eq. (20); with the help of (21), (22), and (24), we get

$$-\sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \lambda^{2n} \sum_{s=0}^{n} \binom{n}{s} \left(\frac{2g}{\lambda^2}\right)^s \frac{(-st)^{2n-s-1}}{(2n-s-1)!} \ln t. \quad (A1)$$

We are interested in a particular power of t. Call $\nu = n - s$, $n > \nu \geq 0$; then we get Eq. (32). It is now

easy to show that for constant ν the series converges if $2egt < 1$.

2. *The potential* $V(r) = ge^{-\mu r}/\mu_0 r^2$. We start with Eq. (38). The leading term $a_0/t^{\frac{1}{2}}$ in Eq. (40) is obtained as follows: The integral is calculated according to Eq. (21) with $r = 1$. Next put $\exp(-stz_2) \Rightarrow 1$ and use the leading term of $z_2(t)$, i.e., $(2g\mu/t)^{\frac{1}{2}}$, and put $s = n$. This gives for the infinite series of (38)

$$\left(\frac{2g\mu}{t}\right)^{\frac{1}{2}} \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{1}{2n-1}. \qquad (A2)$$

The following is easily shown:

$$\sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{1}{2n-1} = \int_1^{\infty} \left[\left(1 - \frac{1}{x^2}\right)^{\frac{1}{2}} - 1\right] dx,$$

$$= 1 - \frac{\pi}{2}. \qquad (A3)$$

This, together with the leading term of z_2 in (38), yields the leading term $(2g\mu/t)^{\frac{1}{2}}$ of the phase shift. There are five contributions from the infinite series in (38) to the energy-independent term a_1.

(i) In Eq. (21) put $r = 1$ and $\exp(-stz_2) \Rightarrow 1$. In (38) put $s = n$ and take the second term in the expansion of $1/z_2(t)$. Thus

$$g\mu \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n}. \qquad (A4)$$

(ii) The same as in (i), but, instead of expanding $1/z_2(t)$, expand $\exp(-stz_2)$. This gives

$$-2g\mu \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{n}{2n-1}. \qquad (A5)$$

(iii) In (21) put $r = 2$ $\exp(-stz_2) \Rightarrow 1$. In (38) put $s = n$ and take the leading term of $1/z_2$. Note that $n \geq 2$. We get

$$-2g\mu \sum_{n=2}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{n}{(2n-1)(2n-2)}. \qquad (A6)$$

Adding these three contributions, we obtain

$$g\mu \left(\frac{1}{2} - \sum_{n=2}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{1}{n-1}\right) = g\mu \ln 2. \qquad (A7)$$

(iv) The next contribution is derived from the exponential integral term with $(n, s) = (1, 1)$. It is

$$-\frac{1}{2}g\mu \ln(\gamma_0^2 2g\mu). \qquad (A8)$$

(v) The last contribution comes from the first two terms in (38):

$$\frac{1}{2}\pi\lambda + g\mu. \qquad (A9)$$

Adding all contributions, we get a_1 as given in (41). As to the b series in (40), its derivation is very similar to the corresponding Yukawa case.

APPENDIX B

In this appendix, we wish to derive a simple expression for an upper limit on the absolute value of δ_{WKB}. Assume that the repulsive potential $V(r, k)$ satisfies the condition that $r^2 V(r, k)$ is a monotonically decreasing function as r increases. Now expansion of the first integrand in Eq. (2) yields

$$\delta_{\text{WKB}} = \frac{\pi\lambda}{2} - z + \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \lambda^{2n}$$

$$\times \int_z^{\infty} \frac{dx}{x^{2n}} \left(1 + \frac{2M}{\lambda^2}\left(\frac{x}{k}\right)^2 V\left(\frac{x}{k}, k\right)\right)^n, \qquad (B1)$$

where z satisfies the equation

$$1 - \frac{\lambda^2}{z^2} - \frac{2M}{k^2} V\left(\frac{z}{k}, k\right) = 0. \qquad (B2)$$

Note that every term of the infinite series is negative. Thus, by the above condition on V, we get the inequality

$$|\delta_{\text{WKB}}| \leq -\frac{\pi}{2}\lambda + z - \sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \lambda^{2n}$$

$$\times \left(1 + \frac{2M}{\lambda^2}\left(\frac{z}{k}\right)^2 V\left(\frac{z}{k}, k\right)\right)^n \int_z^{\infty} \frac{dx}{x^{2n}}, \qquad (B3)$$

which by (B2) becomes

$$|\delta_{\text{WKB}}| \leq -\frac{\pi}{2}\lambda + z - z\sum_{n=1}^{\infty} (-1)^n \binom{\frac{1}{2}}{n} \frac{1}{2n-1}. \qquad (B4)$$

The last sum is given by (A3). Therefore

$$|\delta_{\text{WKB}}| \leq \frac{1}{2}\pi(z - \lambda). \qquad (B5)$$

Note that this inequality holds for every angular momentum, every energy (as far as the WKB approximation is reliable at lower energies), and is independent of the dependence of the potential on energy. For the special case $V \propto 1/r^2$, (B5) becomes an equality.

Initial-Value Problem for Relativistic Plasma Oscillations*

IAN LERCHE

Enrico Fermi Institute for Nuclear Studies, The University of Chicago, Chicago, Illinois

The solution of the collisionless, relativistic Vlasov–Maxwell set of equations is given for the case where neither ambient electric nor magnetic fields are present and a smeared-out negative charge background preserves over-all space-charge neutrality with a relativistic proton plasma. The method of solution is based upon the eigenfunction expansion method invented by van Kampen. It is shown that, besides the relativistically modified modes of Case and Zelazny, there exists a new discrete set of modes whose phase velocities are greater than the speed of light *in vacuo*. The solution also exhibits the coupling of electrostatic and electromagnetic modes and the existence of complex phase velocities. This initial-value problem (or the problem with the electron and proton roles reversed) is of interest because of the existence of cosmic rays, relativistic electrons emitting synchrotron radiation in nonthermal radio sources, solar noise generation, where electron velocities may approach c (Bailey, 1951), and high-energy electrons in laboratory machines (Post, 1960). In these cases the nonrelativistic treatment is clearly inadequate.

1. INTRODUCTION

THE study of plasma oscillations from the point of view of an initial-value problem has received considerable attention, particularly in the absence of any ambient electric and magnetic fields. The behavior of longitudinal waves in a nonrelativistic plasma, developed by Landau[1] and van Kampen,[2] has been investigated in great detail by Case,[3] who showed the equivalence of van Kampen's normal mode analysis and Landau's Laplace transform treatment. More recently, Zelazny[4] has discussed the behavior of longitudinal and transverse waves in a nonrelativistic plasma, when coupling occurs between the two types of waves, from the point of view of an initial-value problem. He shows that a nonrelativistic generalization of Case's procedure yields a complete set of eigenfunctions of the van Kampen type and he points out the equivalence of this method of solution to a Landau-type approach.

In recent years it has been realized that situations can occur, either naturally or man-made, where the nonrelativistic treatment is inadequate. Thus, in trying to understand the better-than 1% isotropy of cosmic rays (Greisen[5]) we need to consider the oscillations which can occur in a relativistic plasma. Likewise, in the generation of solar noise, it has been suggested (Bailey[6]) that electron velocities may reach some appreciable fraction of the velocity of light.

Again, in high-energy laboratory plasma devices, there is reason to believe that electrons can attain velocities close to that of light (Post[7]). Further, if we attribute nonthermal emission of noise in radio sources to synchrotron radiation, we have a situation where the speed of a particle approaches c.

Therefore, we feel that the problem of oscillations in a relativistic plasma is of considerable physical interest. This problem is closely tied to an initial-value problem since oscillations must be triggered by some initial perturbation either in the electric and magnetic fields or in the plasma distribution function, or both.

In this paper, we propose to examine a simple initial-value problem in order to gain some physical insight into the behavior of relativistic plasmas. The situation to be developed deals with the collisionless, relativistic Vlasov equation for a proton plasma together with the full set of Maxwell's equations from the point of view of an initial perturbation in the distribution function only. There is assumed to be a smeared-out electron charge background which takes no part in any motion and serves to satisfy over-all space-charge neutrality. The method of solution makes use of the normal mode technique of van Kampen,[2] extended by Case[3] and generalized by Zelazny[4] for nonrelativistic plasmas. We show that in order to take the relativistic nature of the problem into account, we have to modify even further the van Kampen[2] technique.

A recent paper by Wang[8] has considered purely

[1] L. Landau, J. Phys. USSR **10**, 25 (1946).
[2] N. G. van Kampen, Physica **21**, 949 (1955).
[3] K. M. Case, Ann. Phys. (N.Y.) **7**, 349 (1959).
[4] R. S. Zelazny, Ann. Phys. (N.Y.) **20**, 261 (1962).
[5] K. Greisen, *Progress in Cosmic Ray Physics* (North-Holland Publishing Company, Amsterdam, 1956), Vol. 3, Chap. 1.
[6] V. A. Bailey, Phys. Rev. **83**, 439 (1951).
[7] R. F. Post, in *Proceedings of the International Summer Course in Plasma Physics* (Danish Atomic Energy Commission, Riso, Denmark, 1960), p. 367.
[8] H. S. C. Wang, "Stationary Wave Theory for Relativistic Plasma Oscillations," Technical Report of the National Bureau of Standards, Boulder, Colorado, 1961.

longitudinal modes in a relativistic plasma from the point of view of an initial-value problem without considering the coupling of these space-charge waves to transverse waves. However, even in the nonrelativistic plasma, it is well known (Kahn[9]) that the coupling is quite strong. It can also be shown (Lerche[10]) that the coupling of stable space-charge waves to transverse waves drastically alters the stability requirements of the pure transverse wave in a relativistic plasma. Thus, the initial-value problem should (and will) be discussed taking the full set of Maxwell's equations into account.

Further, it has been brought to our attention by the referee that Felderhof (1963) has considered in detail the nonrelativistic initial-value problem when the distribution function is isotropic. In this case one can separate out the longitudinal and transverse waves since there is no coupling. Felderhof[11] demonstrates that an isotropic relativistic plasma leads to equations which can be handled in exactly the same manner as the nonrelativistic situation. Again one does not obtain coupling between the longitudinal and transverse waves. In view of the remarks already made and because of the number of situations, both natural and man-made, in which the plasma is known, or suspected, to have an anisotropic distribution function, we feel that the solution of the initial-value problem for anisotropic, relativistic plasmas is not without some interest.

We assume that the equilibrium state of the relativistic proton plasma's distribution function corresponds to the absence of any space-charge density and current density. It is further assumed that no ambient magnetic or electric fields are present. Thus any fields which arise are a direct consequence of the distribution function being perturbed at some instant of time which, for simplicity, we take to be $t = 0$. Also the equilibrium system is taken to be homogeneous in coordinate space. However, we make no statement concerning the variation of the equilibrium distribution function in momentum space.

2. EQUATIONS OF MOTION

The relativistic proton plasma's distribution function F satisfies the Vlasov equation:

$$c^{-1}\frac{\partial F}{\partial t} + \frac{\mathbf{p}}{(1+p^2)^{\frac{1}{2}}}\cdot\frac{\partial F}{\partial \mathbf{x}}$$
$$+ \frac{\epsilon}{mc^2}\left[\mathbf{E} + \frac{(\mathbf{p}\times\mathbf{H})}{(1+p^2)^{\frac{1}{2}}}\right]\cdot\frac{\partial F}{\partial \mathbf{p}} = 0. \quad (1)$$

[9] F. D. Kahn, J. Fluid. Mech. 14, 321 (1962).
[10] I. Lerche, J. Math. Phys. (to be published).
[11] B. U. Felderhof, Physica 29, 293 (1963).

Here ϵ, m are the charge and rest mass of a proton; the normalized momentum $\mathbf{p}$ is defined in terms of the actual momentum $\mathbf{p}'$ through $mc\mathbf{p} = \mathbf{p}'$.

The set of Maxwell's equations can be written

$$\nabla\times\mathbf{H} - c^{-1}\frac{\partial\mathbf{E}}{\partial t} = 4\pi\epsilon\int(F_+ - F_-)\frac{\mathbf{p}}{(1+p^2)^{\frac{1}{2}}}d^3p, \quad (2)$$

$$\nabla\times\mathbf{E} + c^{-1}(\partial\mathbf{H}/\partial t) = 0, \quad (3)$$

$$\nabla\cdot\mathbf{E} = 4\pi\epsilon\int(F_+ - F_-)\,d^3p, \quad (4)$$

$$\nabla\cdot\mathbf{H} = 0, \quad (5)$$

where F_+ and F_- are the proton and electron distribution functions, respectively.

We assume that the relativistic proton plasma equilibrium distribution function $f_0(\mathbf{p})$, which is known and is solely a function of $\mathbf{p}$, is perturbed by a small amount f_1. Then f_1 satisfies the linearized equation

$$c^{-1}\frac{\partial f_1}{\partial t} + \frac{\mathbf{p}}{(1+p^2)^{\frac{1}{2}}}\cdot\frac{\partial f_1}{\partial \mathbf{x}}$$
$$+ \frac{\epsilon}{mc^2}\left[\mathbf{E} + \frac{(\mathbf{p}\times\mathbf{H})}{(1+p^2)^{\frac{1}{2}}}\right]\cdot\frac{\partial f_0}{\partial \mathbf{p}} = 0. \quad (6)$$

Also Eqs. (2) and (4) become, respectively,

$$\nabla\times\mathbf{H} - c^{-1}\frac{\partial\mathbf{E}}{\partial t} = 4\pi\epsilon\int\frac{\mathbf{p}f_1}{(1+p^2)^{\frac{1}{2}}}d^3p, \quad (7)$$

$$\nabla\cdot\mathbf{E} = 4\pi\epsilon\int f_1\,d^3p. \quad (8)$$

We now spatially Fourier transform f_1, $\mathbf{E}$, and $\mathbf{H}$, and we obtain

$$c^{-1}\frac{\partial f_1}{\partial t} + \frac{i(\mathbf{k}\cdot\mathbf{p})}{(1+p^2)^{\frac{1}{2}}}f_1 + \frac{\epsilon}{mc^2}$$
$$\times\left[\mathbf{E} + \frac{(\mathbf{p}\times\mathbf{H})}{(1+p^2)^{\frac{1}{2}}}\right]\cdot\frac{\partial f_0}{\partial \mathbf{p}} = 0, \quad (9)$$

$$i\mathbf{k}\times\mathbf{H} - c^{-1}\frac{\partial\mathbf{E}}{\partial t} = 4\pi\epsilon\int\frac{\mathbf{p}f_1}{(1+p^2)^{\frac{1}{2}}}d^3p, \quad (10)$$

$$\mathbf{k}\cdot\mathbf{H} = 0, \quad (11)$$

$$i\mathbf{k}\cdot\mathbf{E} = 4\pi\epsilon\int f_1\,d^3p, \quad (12)$$

$$i\mathbf{k}\times\mathbf{E} + c^{-1}(\partial\mathbf{H}/\partial t) = 0, \quad (13)$$

as our basic set of equations. A factor $e^{i\mathbf{k}\cdot\mathbf{x}}$ and an implied index $\mathbf{k}$ have been omitted in Eq. (9) through Eq. (13).

3. SET OF EIGENFUNCTIONS

The general approach of van Kampen's eigenfunction expansion method, which we preserve, is to look for solutions of Eqs. (9) through (13) in the form

$$\begin{Bmatrix} f_1(\mathbf{p}, t) \\ \mathbf{E}(t) \\ \mathbf{H}(t) \end{Bmatrix} = \exp\left(-ikc\beta t\right) \begin{Bmatrix} \varphi(\beta, \mathbf{p}) \\ \mathbf{E}_1(\beta) \\ \mathbf{H}_1(\beta) \end{Bmatrix}, \quad (14)$$

where $k = |\mathbf{k}|$.

Inserting Eq. (14) into Eqs. (9) through (13), we obtain

$$i\varphi(\beta, \mathbf{p})\left[\frac{\mathbf{k} \cdot \mathbf{p}}{(1 + p^2)^{\frac{1}{2}}} - k\beta\right] + \frac{\epsilon}{mc^2}$$

$$\times \left[\mathbf{E}_1 + \frac{(\mathbf{p} \times \mathbf{H}_1)}{(1 + p^2)^{\frac{1}{2}}}\right] \cdot \frac{\partial f_0}{\partial \mathbf{p}} = 0, \quad (15)$$

$$i\mathbf{k} \times \mathbf{H}_1 + ik\beta \mathbf{E}_1 = 4\pi\epsilon \int \frac{\varphi \mathbf{p}}{(1 + p^2)^{\frac{1}{2}}} \, d^3p, \quad (16)$$

$$\mathbf{k} \cdot \mathbf{H}_1 = 0, \quad (17)$$

$$i\mathbf{k} \cdot \mathbf{E}_1 = 4\pi\epsilon \int \varphi \, d^3p, \quad (18)$$

$$\mathbf{k} \times \mathbf{E}_1 = k\beta \mathbf{H}_1. \quad (19)$$

It now proves convenient to consider a momentum coordinate system defined by

$$p_\parallel = k^{-1}(\mathbf{k} \cdot \mathbf{p}); \quad \mathbf{p}_\perp = \mathbf{p} - k^{-1}p_\parallel \mathbf{k}. \quad (20)$$

It is also convenient to decompose the electric and magnetic fields into components pointing along, and normal to **k**. Thus, we define

$$E_\parallel = (\mathbf{k} \cdot \mathbf{E}_1)k^{-1}; \quad \mathbf{E}_\perp = \mathbf{E} - \mathbf{k}k^{-1}E_\parallel,$$

$$H_\parallel = (\mathbf{k} \cdot \mathbf{H}_1)k^{-1}; \quad \mathbf{H}_\perp = \mathbf{H} - \mathbf{k}k^{-1}H_\parallel.$$

In terms of these variables we see that Eqs. (15) through (19) become

$$\varphi(\beta, p_\parallel, \mathbf{p}_\perp)\left[\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right] + \frac{\epsilon}{ikmc^2}$$

$$\times \left[\mathbf{E}_1 + \frac{(\mathbf{p} \times \mathbf{H}_\perp)}{(1 + p^2)^{\frac{1}{2}}}\right] \cdot \frac{\partial f_0}{\partial \mathbf{p}} = 0, \quad (21)$$

$$E_\parallel = \frac{4\pi\epsilon}{ik} \int \varphi(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p, \quad (22)$$

$$\mathbf{k} \times \mathbf{H}_\perp + k\beta \mathbf{E}_\perp$$

$$= -i4\pi\epsilon \int \frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}} \varphi(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p, \quad (23)$$

$$\mathbf{k} \times \mathbf{E}_\perp = k\beta \mathbf{H}_\perp, \quad (24)$$

$$H_\parallel = 0. \quad (25)$$

We also note that

$$\beta \int \varphi \, d^3p = \int \frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} \varphi \, d^3p. \quad (26)$$

From Eqs. (23) and (24) we have

$$\mathbf{E}_\perp(\beta) = \frac{4\pi\beta\epsilon}{ik(\beta^2 - 1)} \int \frac{\mathbf{p}_\perp \varphi(\beta, p_\parallel, \mathbf{p}_\perp)}{(1 + p^2)^{\frac{1}{2}}} \, d^3p$$

$$+ \mathcal{E}_\perp(\beta)\delta(\beta^2 - 1), \quad (27)$$

$$\mathbf{H}_\perp(\beta) = \frac{4\pi\epsilon}{ik^2(\beta^2 - 1)} \int \frac{(\mathbf{k} \times \mathbf{p}_\perp)\varphi(\beta, p_\parallel, \mathbf{p}_\perp)}{(1 + p^2)^{\frac{1}{2}}} \, d^3p$$

$$+ \mathbf{h}_\perp(\beta)\delta(\beta^2 - 1), \quad (28)$$

where $\mathcal{E}_\perp(\pm 1)$ and $\mathbf{h}_\perp(\pm 1)$ are so far completely undetermined. However, they are not independent. This can easily be seen by using Eq. (24), and it follows that

$$\mathbf{h}_\perp(\beta) = (k\beta)^{-1}(\mathbf{k} \times \mathcal{E}_\perp(\beta)). \quad (29)$$

Thus, we have the solution of Maxwell's equations in terms of ϕ and $\mathcal{E}_\perp$. Making use of $\mathbf{E}_\perp$, $E_\parallel$, $\mathbf{H}_\perp$, and $H_\parallel$ in Eq. (21), we obtain

$$\varphi\left[\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right] - \frac{4\pi\epsilon^2}{k^2mc^2} \frac{\partial f_0}{\partial p_\parallel} \int \varphi \, d^3p$$

$$+ \left[\frac{\epsilon\delta(\beta^2 - 1)}{ikmc^2\beta} \mathcal{E}_\perp(\beta) - \frac{4\pi\epsilon^2}{k^2mc^2(\beta^2 - 1)} \int \frac{\mathbf{p}_\perp \varphi \, d^3p}{(1 + p^2)^{\frac{1}{2}}}\right]$$

$$\cdot \left[\left(\beta - \frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}}\right) \frac{\partial f_0}{\partial \mathbf{p}_\perp} + \frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}} \frac{\partial f_0}{\partial p_\parallel}\right] = 0. \quad (30)$$

For brevity, we now define

$$\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp)$$

$$= \alpha\left[\left(\beta - \frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}}\right) \frac{\partial f_0}{\partial \mathbf{p}_\perp} + \frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}} \frac{\partial f_0}{\partial p_\parallel}\right], \quad (31a)$$

$$U_\parallel(p_\parallel, \mathbf{p}_\perp) = \alpha(\partial f_0/\partial p_\parallel), \quad (31b)$$

$$\alpha = 4\pi\epsilon^2/mc^2k^2. \quad (31c)$$

Then Eq. (30) can be written as

$$\varphi\left[\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right] - U_\parallel(p_\parallel, \mathbf{p}_\perp) \int \varphi \, d^3p$$

$$- \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp)}{(\beta^2 - 1)} \cdot \int \frac{\mathbf{p}_\perp \varphi \, d^3p}{(1 + p^2)^{\frac{1}{2}}} - \frac{ik\alpha}{4\pi\epsilon\beta}$$

$$\times \mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \mathcal{E}_\perp(\beta)\delta(\beta^2 - 1) = 0. \quad (32)$$

The formal solution to Eq. (32) can be written as

$$\varphi(\beta, p_\parallel, \mathbf{p}_\perp) = \frac{U_\parallel(p_\parallel, \mathbf{p}_\perp)}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}} \int \varphi \, d^3p$$

$$+ \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \int \frac{\mathbf{p}_\perp \varphi \, d^3p}{(1 + p^2)^{\frac{1}{2}}}}{(\beta^2 - 1)\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$

$$+ \frac{ik\alpha \mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \mathbf{\mathcal{E}}_\perp(\beta)\delta(\beta^2 - 1)}{4\pi\epsilon\beta\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$

$$+ \lambda(\beta, \mathbf{p}_\perp)\delta\left(\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right), \quad (33)$$

where so far $\lambda(\beta, \mathbf{p}_\perp)$ is completely arbitrary. We can use the fact that Eq. (33) is the solution to Eq. (32) to set some restrictions on $\lambda(\beta, \mathbf{p}_\perp)$ through the perturbed space-charge density and current density. Thus, if we integrate Eq. (33) with respect to d^3p, we obtain

$$\rho(\beta)\left[1 - \int^* \frac{U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}\right]$$

$$- \left[\frac{\mathbf{J}_\perp(\beta)}{(\beta^2 - 1)} + \frac{ik\alpha\delta(\beta^2 - 1)}{4\pi\epsilon\beta}\mathbf{\mathcal{E}}_\perp(\beta)\right]$$

$$\cdot \int^* \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$

$$= \int \lambda(\beta, \mathbf{p}_\perp)\delta\left(\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right) d^3p, \quad (34)$$

where

$$\rho(\beta) = \int \varphi \, d^3p, \quad (35)$$

$$\mathbf{J}_\perp(\beta) = \int \frac{\mathbf{p}_\perp \varphi}{(1 + p^2)^{\frac{1}{2}}} \, d^3p. \quad (36)$$

The asterisk on the integrals denotes the principle-value integrals in view of the fact that an arbitrary amount of δ function has been added to Eq. (33).

Likewise, if we first multiply Eq. (33) by $\mathbf{p}_\perp/(1 + p^2)^{\frac{1}{2}}$ and then integrate with respect to d^3p, we obtain

$$\mathbf{J}_\perp(\beta) - \rho(\beta)\int^* \frac{\mathbf{p}_\perp U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{[p_\parallel - \beta(1 + p^2)^{\frac{1}{2}}]}$$

$$- \left[\frac{\mathbf{J}_\perp(\beta)}{(\beta^2 - 1)} + \frac{ik\alpha\delta(\beta^2 - 1)}{4\pi\epsilon\beta}\mathbf{\mathcal{E}}_\perp(\beta)\right]$$

$$\cdot \int^* \frac{\mathbf{p}_\perp \mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p}{[p_\parallel - \beta(1 + p^2)^{\frac{1}{2}}]},$$

$$= \int \frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}} \lambda(\beta, \mathbf{p}_\perp)\delta\left(\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right) d^3p. \quad (37)$$

For simplicity we now define

$$A_\parallel(\beta) = 1 - \int^* \frac{U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}, \quad (38a)$$

$$\mathbf{B}_\perp(\beta) = \int^* \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}. \quad (38b)$$

Provided f_0 satisfies Hölder–Lipschitz conditions (we assume it does), we can write

$$\mathbf{B}_\perp(\beta) = \int^* \frac{\mathbf{p}_\perp U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}(1 + p^2)^{\frac{1}{2}}}. \quad (39)$$

We also define the tensor

$$\mathsf{T}(\beta) = \mathsf{I}(\beta^2 - 1) - \int^* \frac{\mathbf{p}_\perp \mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \, d^3p}{[p_\parallel - \beta(1 + p^2)^{\frac{1}{2}}]}. \quad (40)$$

Again using the fact that f_0 satisfies the Hölder–Lipschitz conditions, we have

$$\mathsf{T}(\beta) = \mathsf{I}(\beta^2 - 1) - \alpha\int \frac{f_0}{(1 + p^2)^{\frac{1}{2}}}$$

$$\times \left[\mathsf{I} - \frac{\mathbf{p}_\perp \mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}}\right] d^3p$$

$$- \int^* \frac{\mathbf{p}_\perp \mathbf{p}_\perp U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{(1 + p^2)\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}. \quad (41)$$

Note that I is a second-rank unit tensor in the 2-space of $\mathbf{p}_\perp$.

Now the harmonic mean energy of the plasma is defined by

$$\langle E^{-1}\rangle = (mc^2)^{-1}\int \frac{f_0 \, d^3p}{(1 + p^2)^{\frac{1}{2}}} \equiv \gamma(mc^2)^{-1}, \quad \text{say.} \quad (42a)$$

Also, the tensor velocity dispersion per given energy in the plane normal to $\mathbf{k}$ is just

$$\langle \mathbf{U}_\perp \mathbf{U}_\perp E^{-1}\rangle = m^{-1}\int \frac{f_0 \mathbf{p}_\perp \mathbf{p}_\perp \, d^3p}{(1 + p^2)^{\frac{3}{2}}}, \quad (42b)$$

$$\equiv c^2\langle E^{-1}\rangle\mathsf{W}_\perp, \quad (42c)$$

thus defining $\mathsf{W}_\perp$.

Hence, Eq. (41) can be written as

$$\mathsf{T}(\beta) = \mathsf{I}(\beta^2 - 1 - \gamma\alpha) + \gamma\alpha\mathsf{W}_\perp$$

$$- \int^* \frac{\mathbf{p}_\perp \mathbf{p}_\perp U_\parallel(p_\parallel, \mathbf{p}_\perp) \, d^3p}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}. \quad (43)$$

Thus Eqs. (34) and (37) become

$$\rho(\beta)A_\parallel(\beta) - \mathbf{R}_\perp(\beta) \cdot \mathbf{B}_\perp(\beta)$$

$$- \frac{ik\alpha\delta(\beta^2 - 1)}{4\pi\epsilon\beta}\mathbf{\mathcal{E}}_\perp(\beta) \cdot \mathbf{B}_\perp(\beta)$$

$$= \int \lambda(\beta, \mathbf{p}_\perp)\delta\left(\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right) d^3p, \quad (44)$$

$$-\rho(\beta)\mathbf{B}_\perp(\beta) + \mathbf{R}_\perp(\beta) \cdot \mathsf{T}(\beta)$$

$$- \frac{ik\alpha\delta(\beta^2 - 1)}{4\pi\epsilon\beta}\mathbf{\mathcal{E}}_\perp(\beta) \cdot [\mathsf{I}(\beta^2 - 1) - \mathsf{T}(\beta)]$$

$$= \int \frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}} \lambda(\beta, \mathbf{p}_\perp)\delta\left(\frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}} - \beta\right) d^3p, \quad (45)$$

where

$$\mathbf{R}_\perp(\beta) = \mathbf{J}_\perp(\beta)(\beta^2 - 1)^{-1}.$$

In order to consider various β values, it is convenient to make use of three simple properties of δ functions. These are:

(i) $|\beta| \leq 1$ and β real,

$$\delta\left(\frac{p_\parallel}{(1 + \mathbf{p}^2)^{\frac{1}{2}}} - \beta\right) \equiv \frac{(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}}{(1 - \beta^2)^{\frac{3}{2}}}$$

$$\times \delta\left(p_\parallel - \frac{\mathbf{p}(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}}{(1 - \beta^2)^{\frac{3}{2}}}\right); \quad (46a)$$

(ii) $|\beta| > 1$ and β real,

$$\delta\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\} \equiv 0. \quad (46b)$$

(iii) If β is complex, it is clear that $p_\parallel - \beta(1 + p^2)^{\frac{1}{2}}$ is never zero since $p_\parallel$ and $(1 + p^2)^{\frac{1}{2}}$ are real variables. As a consequence

$$\lambda(\beta, p_\perp) \equiv 0, \quad (46c)$$

for complex β.

We now have the following situation for various ranges of β.

Class 1: For real β lying in the range $-1 \leq \beta \leq 1$ for which $A_\parallel(\beta)$, $\mathbf{B}_\perp(\beta)$, and $T(\beta)$ do not vanish, the following conditions must be satisfied by $\lambda(\beta, \mathbf{p}_\perp)$:

$$\int \frac{\lambda(\beta, \mathbf{p}_\perp)(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}} d^2\mathbf{p}_\perp}{(1 - \beta^2)^{\frac{3}{2}}}$$

$$= \rho(\beta)A_\parallel(\beta) - \mathbf{R}_\perp(\beta) \cdot \mathbf{B}_\perp(\beta)$$

$$- \frac{ik\alpha\delta(\beta^2 - 1)\boldsymbol{\mathcal{E}}_\perp(\beta) \cdot \mathbf{B}_\perp(\beta)}{4\pi\epsilon\beta}, \quad (47)$$

$$\int \mathbf{p}_\perp(1 - \beta^2)^{-1}\lambda(\beta, \mathbf{p}_\perp) d^2\mathbf{p}_\perp$$

$$= -\rho(\beta)\mathbf{B}_\perp(\beta) + \mathbf{R}_\perp(\beta) \cdot T(\beta)$$

$$- \frac{ik\alpha\delta(\beta^2 - 1)}{4\pi\epsilon\beta}\boldsymbol{\mathcal{E}}_\perp(\beta) \cdot [\mathbf{I}(\beta^2 - 1) - T(\beta)]. \quad (48)$$

Class 2: For complex β we obtain a system of homogeneous equations for $\rho(\beta)$ and $\mathbf{R}_\perp(\beta)$:

$$\rho(\beta)A_\parallel(\beta) - \mathbf{R}_\perp(\beta) \cdot \mathbf{B}_\perp(\beta) = 0, \quad (49)$$

$$-\rho(\beta)\mathbf{B}_\perp(\beta) + \mathbf{R}_\perp(\beta) \cdot T(\beta) = 0. \quad (50)$$

Thus, a nontrivial solution exists for complex β if and only if

$$\det \begin{vmatrix} A_\parallel(\beta) & -\mathbf{B}_\perp(\beta) \\ -\mathbf{B}_\perp(\beta) & T(\beta) \end{vmatrix} = 0. \quad (51)$$

We suppose that Eq. (51) is satisfied for M complex values β_j $(j = 1, 2, \cdots, M)$.

We assume, for simplicity only, that the roots of

Eq. (51) are simple. The more general case of multiple roots is described in the Appendix.

The eigenfunction corresponding to β_j is

$$\varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp)$$

$$= \frac{\rho(\beta_j)U_\parallel(p_\parallel, \mathbf{p}_\perp) + \mathbf{U}_\perp(\beta_j, p_\parallel, \mathbf{p}_\perp) \cdot \mathbf{R}_\perp(\beta_j)}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta_j\}}$$

$$(j = 1, 2, \cdots, M). \quad (52)$$

Since $\rho(\beta_j)$ and $\mathbf{R}_\perp(\beta_j)$ are related through Eqs. (49) and (50), there is only one undetermined constant, say $\rho(\beta_j)$, for each $\phi_j(\beta_j)$.

Class 3: For real β and $|\beta| > 1$, we again obtain the system of homogeneous Eqs. (49) and (50). The condition for eigenvalues to exist is the same as Eq. (51) and the eigenfunctions are of the general form Eq. (52). We assume that Eq. (51) is satisfied for some real β_j with $|\beta_j| > 1$, where $j = M + 1, M + 2, \cdots, N$.

Class 4: It may also happen that some real β_j with $|\beta_j| \leq 1$ satisfy Eq. (51). We suppose that this gives rise to β_j $(j = N + 1, N + 2, \cdots, P)$. This situation occurs if and only if

$$U_\parallel(p_\parallel, \mathbf{p}_\perp)|_{p_\parallel = \beta_j(1 + p^2)^{\frac{1}{2}}} = 0,$$

$$(j = N + 1, N + 2, \cdots, P). \quad (53)$$

This condition arises because we must demand that simultaneously

$$\det \begin{vmatrix} A_\parallel^+(\beta_j) & -\mathbf{B}_\perp^+(\beta_j) \\ -\mathbf{B}_\perp^+(\beta_j) & T^+(\beta_j) \end{vmatrix}$$

$$= \det \begin{vmatrix} A_\parallel^-(\beta_j) & -\mathbf{B}_\perp^-(\beta_j) \\ -\mathbf{B}_\perp^-(\beta_j) & T^-(\beta_j) \end{vmatrix} = 0, \quad (54)$$

where the superscripts $+$, $-$ are described in Sec. 4. The corresponding eigenfunctions are of the form Eq. (52).

We have now found a set of eigenfunctions consisting of a continuum subset with real eigenvalues β lying in the range $-1 \leq \beta \leq 1$; a discrete subset with complex eigenvalues β_j $(j = 1, \cdots, M)$; a discrete subset with real eigenvalues β_j $(j = M + 1, \cdots, N)$, and $|\beta_j| > 1$; and a discrete (somewhat singular) subset with real eigenvalues β_j $(j = N + 1, \cdots, P)$ and $|\beta_j| \leq 1$. If this set of eigenfunctions is complete, and we demonstrate that it is, then any, and every, solution of the Vlasov–Maxwell equations can be expanded in terms of this set.

4. PROOF OF COMPLETENESS OF THE EIGENFUNCTION SET

The aim of this paper is the solution of the initial-value problem, and the proof of completeness is intimately related to this problem. Thus, if the set of

eigenfunctions is complete, it follows that the general solution of the spatial Fourier transform of the Vlasov–Maxwell set of equations can be written as

$$f_1(p_\parallel, \mathbf{p}_\perp, \mathbf{k}, t) = \sum_{j=1}^{P} \varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp) e^{-ikc\beta_j t}$$
$$+ \int_{-1}^{+1} \varphi(\beta, p_\parallel, \mathbf{p}_\perp) e^{-ikc\beta t} \, d\beta. \quad (55)$$

When f_1 is given at some initial time, which for the sake of convenience we take to be $t = 0$, we have

$$f_1 (t = 0) = \sum_{j=1}^{P} \varphi_j(\beta_j) + \int_{-1}^{+1} \varphi(\beta, p_\parallel, \mathbf{p}_\perp) \, d\beta. \quad (56)$$

Note $f_1 (t = 0)$ is given. Thus, if we can demonstrate that the unknowns $\lambda(\beta, \mathbf{p}_\perp)$, $\mathcal{E}_\perp(\pm 1)$, and

$$\rho(\beta_j) \, (j = 1, \cdots, P)$$

can be determined, it follows that a solution of the initial-value problem is at one and the same time a proof of completeness of the eigenfunction set. We now demonstrate that it is indeed possible to determine the unknown parameters.

The constants $\mathcal{E}_\perp(\pm 1)$ depend upon a precise statement of the initial-value problem. If we assume that the electric and magnetic fields are a direct consequence of the initial distribution function perturbation, then we require

$$\mathbf{E}_\perp (t = 0) = 0 = \mathbf{H}_\perp (t = 0).$$

We could choose to demand that some initial perturbation in the fields determines the perturbed distribution function $f_1 (t = 0)$. This choice of initial values does not alter the argument; it merely changes the values of $\mathcal{E}_\perp(\pm 1)$. As a consequence, for a proof of completeness, it is sufficient to choose the initial values in the manner described.[12]

Then we have

$$\mathbf{E}_\perp (t = 0) = \int_{-1}^{+1} \mathbf{E}_\perp(\beta) \, d\beta + \sum_{j=1}^{P} \mathbf{E}_\perp(\beta_j) = 0, \quad (57a)$$

$$\mathbf{H}_\perp (t = 0) = \int_{-1}^{+1} \mathbf{H}_\perp(\beta) \, d\beta + \sum_{j=1}^{P} \mathbf{H}_\perp(\beta_j) = 0. \quad (57b)$$

Upon making use of Eqs. (27) and (28), we see that Eq. (57) becomes

$$\frac{4\pi\epsilon}{ik} \left[\sum_{j=1}^{P} \beta_j \mathbf{R}_\perp(\beta_j) + \int_{-1}^{+1} \beta \mathbf{R}_\perp(\beta) \, d\beta \right]$$
$$+ \int_{-1}^{+1} \mathcal{E}_\perp(\beta)\delta(\beta^2 - 1) \, d\beta = 0, \quad (58a)$$

$$\frac{4\pi\epsilon}{ik} \mathbf{k} \times \left[\sum_{j=1}^{P} \mathbf{R}_\perp(\beta_j) + \int_{-1}^{+1} \mathbf{R}_\perp(\beta) \, d\beta \right]$$
$$+ \mathbf{k} \times \int_{-1}^{+1} \mathcal{E}_\perp(\beta)\beta^{-1}\delta(\beta^2 - 1) \, d\beta = 0. \quad (58b)$$

[12] An alternative procedure is to let $\mathbf{E}_\perp (t = 0) = \mathbf{E}_0$, $\mathbf{H}_\perp (t = 0) = \mathbf{H}_0$, but then we would have to connect H_0 and E_0 with the initial value of the perturbed distribution function $f_1 (t = 0)$ through the Vlasov–Maxwell equations at time $t = 0$. The mathematical problem is unchanged but the constants $\mathcal{E}\perp (\pm 1)$ have different values.

From Eq. (58), it is a trivial matter to show that

$$\mathcal{E}_\perp(+1) = \frac{2\pi\epsilon}{ik} \left[\sum_{j=1}^{P} (\beta_j + 1)\mathbf{R}_\perp(\beta_j) \right.$$
$$\left. + \int_{-1}^{+1} (\beta + 1)\mathbf{R}_\perp(\beta) \, d\beta \right], \quad (59)$$

$$\mathcal{E}_\perp(-1) = \frac{2\pi\epsilon}{ik} \left[\sum_{j=1}^{P} (\beta_j - 1)\mathbf{R}_\perp(\beta_j) \right.$$
$$\left. + \int_{-1}^{+1} (\beta - 1)\mathbf{R}_\perp(\beta) \, d\beta \right]. \quad (60)$$

Thus, we now know $\mathcal{E}_\perp(\pm 1)$ in terms of $\rho(\beta)$ and $\mathbf{R}_\perp(\beta)$. Let us now return to the problem of solving Eq. (56).

From Eq. (32) we see that Eq. (56) can be written as

$$U_\parallel(p_\parallel, \mathbf{p}_\perp) \int_{-1}^{+1} \frac{\rho(\beta) \, d\beta}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$
$$+ \int_{-1}^{+1} \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \mathbf{R}_\perp(\beta) \, d\beta}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$
$$+ \frac{ik\alpha}{4\pi\epsilon} \int_{-1}^{+1} \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \mathcal{E}_\perp(\beta)\delta(\beta^2 - 1) \, d\beta}{\{[p_\parallel/(1 + p^2)^{\frac{1}{2}}] - \beta\}}$$
$$+ \int_{-1}^{+1} \lambda(\beta, \mathbf{p}_\perp)\delta\left(\beta - \frac{p_\parallel}{(1 + p^2)^{\frac{1}{2}}}\right) d\beta$$
$$= f_1(p_\parallel, \mathbf{p}_\perp, \mathbf{k}, t = 0) - \sum_{j=1}^{P} \varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp). \quad (61)$$

Setting $v_\parallel = p_\parallel/(1 + p^2)^{\frac{1}{2}}$ and

$$\Phi(p_\parallel, \mathbf{p}_\perp) = f_1(p_\parallel, \mathbf{p}_\perp, \mathbf{k}, 0) - \sum_{j=1}^{P} \varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp)$$
$$- \frac{ik\alpha}{4\pi\epsilon} \left[\frac{\mathbf{U}_\perp(1, p_\parallel, \mathbf{p}_\perp) \cdot \mathcal{E}_\perp(1)}{(v_\parallel - 1)} \right.$$
$$\left. - \frac{\mathbf{U}_\perp(-1, p_\parallel, \mathbf{p}_\perp) \cdot \mathcal{E}_\perp(-1)}{(v_\parallel + 1)} \right], \quad (62)$$

we see that Eq. (61) can be written as

$$\lambda(v_\parallel, \mathbf{p}_\perp) + U_\parallel(p_\parallel, \mathbf{p}_\perp) \int_{-1}^{+1} \frac{\rho(\beta) \, d\beta}{(v_\parallel - \beta)}$$
$$+ \int_{-1}^{+1} \frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp) \cdot \mathbf{R}_\perp(\beta) \, d\beta}{(v_\parallel - \beta)} = \Phi(p_\parallel, \mathbf{p}_\perp). \quad (63)$$

We must solve Eq. (63) together with the restrictive conditions of Class 1, since $v_\parallel$ is real and $|v_\parallel| \leq 1$. In principle, this enables us to determine $\rho(\beta)$, $\mathbf{R}_\perp(\beta)$, $\lambda(\beta, \mathbf{p}_\perp)$, and $\rho(\beta_j)$ $(j = 1, \cdots, P)$. In order to demonstrate this fact, note first of all that

$$\frac{\mathbf{U}_\perp(\beta, p_\parallel, \mathbf{p}_\perp)}{(v_\parallel - \beta)} = -\alpha \frac{\partial f_0}{\partial \mathbf{p}_\perp} + \frac{\mathbf{p}_\perp U_\parallel(p_\parallel, \mathbf{p}_\perp)}{(v_\parallel - \beta)(1 + p^2)^{\frac{1}{2}}}. \quad (64)$$

Use of Eq. (64) enables Eq. (63) to be written as

$$\lambda(v_\parallel, \mathbf{p}_\perp) + U_\parallel(p_\parallel, \mathbf{p}_\perp)\int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$+ U_\parallel(p_\parallel, \mathbf{p}_\perp)\frac{\mathbf{p}_\perp}{(1 + p^2)^{\frac{1}{2}}}\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$= \Phi(p_\parallel, \mathbf{p}_\perp) + \alpha\frac{\partial f_0}{\partial \mathbf{p}_\perp}\cdot\int_{-1}^{*+1}\mathbf{R}_\perp(\beta)\,d\beta,$$

$$\equiv \Psi(p_\parallel, \mathbf{p}_\perp), \quad \text{say.} \tag{65}$$

To solve Eq. (65) together with Eqs. (47) and (48) it is convenient to choose $v_\parallel$ and $\mathbf{p}_\perp$ as a basic triad of coordinates, rather than $p_\parallel$ and $\mathbf{p}_\perp$.

The range of integration of $v_\parallel$ is $-1 \leq v_\parallel \leq 1$, and

$$dp_\parallel = [(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}/(1 - v^2)^{\frac{3}{2}}]\,dv_\parallel. \tag{66}$$

Thus, for example, the number density n of particles per unit volume is

$$n = \int_{-\infty}^{\infty} dp_\parallel \int f_0(p_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp,$$

$$\equiv \int_{-1}^{+1}\frac{dv_\parallel}{(1 - v^2)^{\frac{3}{2}}}\int(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}\tilde{f}_0(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp, \tag{67}$$

where $\tilde{f}_0(v_\parallel, \mathbf{p}_\perp)$ is just $f_0(p_\parallel, \mathbf{p}_\perp)$ expressed with respect to $v_\parallel$ and $\mathbf{p}_\perp$ as basic coordinates.

Note also that

$$p_\parallel = v_\parallel(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}/(1 - v^2)^{\frac{1}{2}}.$$

In terms of $v_\parallel$ and $\mathbf{p}_\perp$ we see that Eq. (65) becomes

$$\lambda(v_\parallel, \mathbf{p}_\perp) + \tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$+ (1 - v_\parallel^2)^{\frac{1}{2}}\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\frac{\mathbf{p}_\perp}{(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}}\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$= \Psi(v_\parallel, \mathbf{p}_\perp). \tag{68}$$

Multiplying Eq. (68) by $(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}$, integrating with respect to $d^2\mathbf{p}_\perp$, and making use of Eq. (47), we obtain

$$\rho(v_\parallel)A_\parallel(v_\parallel) - \mathbf{R}_\perp(v_\parallel)\cdot\mathbf{B}_\perp(v_\parallel)$$

$$+ \int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)}\times\int\frac{(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)^{\frac{3}{2}}}$$

$$+ (1 - v_\parallel^2)^{-1}\int\mathbf{p}_\perp\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$= \frac{ik\alpha\delta(v_\parallel^2 - 1)\mathbf{\mathcal{E}}_\perp(v_\parallel)\cdot\mathbf{B}_\perp(v_\parallel)}{4\pi\epsilon v_\parallel}$$

$$+ \int\frac{\Psi(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)^{\frac{3}{2}}}. \tag{69}$$

Likewise, if we multiply Eq. (68) by $\mathbf{p}_\perp$, integrate with respect to $d^2\mathbf{p}_\perp$, and make use of Eq. (48), we

obtain

$$-\rho(v_\parallel)\mathbf{B}_\perp(v_\parallel) + \mathbf{R}_\perp(v_\parallel)\cdot\mathbf{T}(v_\parallel)$$

$$+ \int\frac{\mathbf{p}_\perp\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)}\int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)} + (1 - v_\parallel^2)^{-\frac{1}{2}}$$

$$\times \int\frac{\mathbf{p}_\perp\mathbf{p}_\perp\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)}{(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}}\,d^2\mathbf{p}_\perp\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$= \frac{ik\alpha\delta(v_\parallel^2 - 1)}{4\pi\epsilon v_\parallel}\mathbf{\mathcal{E}}_\perp(v_\parallel)\cdot[\mathbf{I}(v_\parallel^2 - 1) - \mathbf{T}(v_\parallel)]$$

$$+ \int\frac{\mathbf{p}_\perp\Psi(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)}. \tag{70}$$

We now let

$$\int(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp = (1 - v_\parallel^2)^{\frac{3}{2}}u_1(v_\parallel), \tag{71a}$$

$$\int\mathbf{p}_\perp\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp = (1 - v_\parallel^2)\mathbf{V}_\perp(v_\parallel), \tag{71b}$$

$$\int\frac{\mathbf{p}_\perp\mathbf{p}_\perp}{(1 + \mathbf{p}_\perp^2)^{\frac{1}{2}}}\tilde{U}_\parallel(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp = (1 - v_\parallel^2)^{\frac{1}{2}}\mathbf{W}(v_\parallel) \tag{71c}$$

We note that $v_1(v_\parallel)$, $\mathbf{V}_\perp(v_\parallel)$, and $\mathbf{W}(v_\parallel)$ are known functions of $v_\parallel$ since $\tilde{f}_0(v_\parallel, \mathbf{p}_\perp)$ is specified. We also set

$$F_\parallel(v_\parallel) = \int\frac{\Psi(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)^{\frac{3}{2}}}$$

$$+ \frac{ik\alpha\delta(v_\parallel^2 - 1)\mathbf{\mathcal{E}}_\perp(v_\parallel)\cdot\mathbf{B}_\perp(v_\parallel)}{4\pi\epsilon v_\parallel}, \tag{72a}$$

$$\mathbf{F}_\perp(v_\parallel) = \int\frac{\mathbf{p}_\perp\Psi(v_\parallel, \mathbf{p}_\perp)\,d^2\mathbf{p}_\perp}{(1 - v_\parallel^2)}$$

$$+ \frac{ik\alpha\delta(v_\parallel^2 - 1)\mathbf{\mathcal{E}}_\perp(v_\parallel)}{4\pi\epsilon v_\parallel}\cdot[\mathbf{I}(v_\parallel^2 - 1) - \mathbf{T}(v_\parallel)]. \tag{72b}$$

Then Eqs. (69) and (70) become, respectively,

$$\rho(v_\parallel)A_\parallel(v_\parallel) - \mathbf{R}_\perp(v_\parallel)\cdot\mathbf{B}_\perp(v_\parallel) + u_1(v_\parallel)\int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$+ \mathbf{V}_\perp(v_\parallel)\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)} = F_\parallel(v_\parallel), \tag{73}$$

$$\rho(v_\parallel)\mathbf{B}_\perp(v_\parallel) - \mathbf{R}_\perp(v_\parallel)\cdot\mathbf{T}(v_\parallel) - \mathbf{V}_\perp(v_\parallel)\int_{-1}^{*+1}\frac{\rho(\beta)\,d\beta}{(v_\parallel - \beta)}$$

$$- \mathbf{W}(v_\parallel)\cdot\int_{-1}^{*+1}\frac{\mathbf{R}_\perp(\beta)\,d\beta}{(v_\parallel - \beta)} = -\mathbf{F}_\perp(v_\parallel). \tag{74}$$

In order to solve the singular integral Eqs. (73) and

(74) it proves convenient to define the column vectors

$$\mathbf{Q}(v_{\parallel}) = \begin{bmatrix} \rho(v_{\parallel}) \\ \mathbf{R}_{\perp}(v_{\parallel}) \end{bmatrix}, \qquad (75a)$$

$$\mathbf{F}(v_{\parallel}) = \begin{bmatrix} F_{\parallel}(v_{\parallel}) \\ -\mathbf{F}_{\perp}(v_{\parallel}) \end{bmatrix}. \qquad (75b)$$

We also define the matrices

$$\mathcal{L}(v_{\parallel}) = \begin{bmatrix} A_{\parallel}(v_{\parallel}) & -\mathbf{B}_{\perp}(v_{\parallel}) \\ \mathbf{B}_{\perp}(v_{\parallel}) & -\mathsf{T}(v_{\parallel}) \end{bmatrix}, \qquad (76a)$$

$$\mathcal{M}(v_{\parallel}) = \begin{bmatrix} u_1(v_{\parallel}) & \mathbf{V}_{\perp}(v_{\parallel}) \\ -\mathbf{V}_{\perp}(v_{\parallel}) & -\mathsf{W}(v_{\parallel}) \end{bmatrix}. \qquad (76b)$$

In terms of these quantities, Eqs. (73) and (74) become

$$\mathcal{L}(v_{\parallel})\mathbf{Q}(v_{\parallel}) + \mathcal{M}(v_{\parallel})\int_{-1}^{*+1} \frac{\mathbf{Q}(\beta)\,d\beta}{(v_{\parallel} - \beta)} = \mathbf{F}(v_{\parallel}). \quad (77)$$

Before proceeding with solving Eq. (77), we note that a relation exists between $\mathcal{L}(v_{\parallel})$ and $\mathcal{M}(v_{\parallel})$. We can write

$$\mathcal{L}(v_{\parallel}) = \mathcal{J} - \int_{-1}^{*+1} \frac{\mathcal{M}(\beta)\,d\beta}{(\beta - v_{\parallel})}, \qquad (78)$$

where the constant matrix $\mathcal{J}$, which is known once f_0 is specified, is given by

$$\mathcal{J} = \begin{bmatrix} 1 & 0 \\ 0 & \mathsf{I}(\beta^2 - 1 - \gamma\alpha) + \gamma\alpha\mathsf{W}_{\perp} \end{bmatrix}. \qquad (79)$$

Thus Eq. (77) becomes

$$\mathcal{J}\mathbf{Q}(v_{\parallel}) - \mathbf{Q}(v_{\parallel})\int_{-1}^{*+1} \frac{\mathcal{M}(\beta)\,d\beta}{(\beta - v_{\parallel})}$$
$$+ \mathcal{M}(v_{\parallel})\int_{-1}^{*+1} \frac{\mathbf{Q}(\beta)\,d\beta}{(v_{\parallel} - \beta)} = \mathbf{F}(v_{\parallel}). \quad (80)$$

We now restrict the problem slightly by demanding that $f_0(v_{\parallel}, p)$ tend to zero as $v_{\parallel} \to \pm 1$ in such a manner that $\mathcal{M}(\pm 1) = 0$ and $\mathbf{F}(\pm 1) = 0$, so that $\mathbf{Q}(\pm 1) = 0$. In any physical situation, this will indeed be the case.

Under this restriction, we note that even though the functions $\mathcal{M}(v_{\parallel})$, $\mathbf{F}(v_{\parallel})$, and $\mathbf{Q}(v_{\parallel})$ are defined only for $|v_{\parallel}| \leq 1$, we can extend their values into the domain $|v_{\parallel}| > 1$ by setting

$$\mathbf{Q}(z) = \mathbf{F}(z) = \mathcal{M}(z) = 0; \quad |z| > 1. \qquad (81)$$

Upon so doing, we see that Eq. (80) becomes

$$\mathcal{J}\mathbf{Q}(v_{\parallel}) - \int_{-\infty}^{*\infty} \frac{\mathcal{M}(\beta)\,d\beta}{(\beta - v_{\parallel})}\mathbf{Q}(v_{\parallel})$$
$$+ \mathcal{M}(v_{\parallel})\int_{-\infty}^{*\infty} \frac{\mathbf{Q}(\beta)\,d\beta}{(v_{\parallel} - \beta)} = \mathbf{F}(v_{\parallel}). \quad (82)$$

We can now gainfully employ the method of solution developed by Case[3] and Zelazny[4] to solve Eq. (82) We define

$$\sigma(z) = (2\pi i)^{-1}\int_{-\infty}^{\infty} \frac{\mathcal{M}(x)\,dx}{(x - z)}, \qquad (83a)$$

$$\mathbf{q}(z) = (2\pi i)^{-1}\int_{-\infty}^{\infty} \frac{\mathbf{Q}(x)\,dx}{(x - z)}, \qquad (83b)$$

$$\mathbf{f}(z) = (2\pi i)^{-1}\int_{-\infty}^{\infty} \frac{\mathbf{F}(x)\,dx}{(x - z)}. \qquad (83c)$$

We denote by $+$ and $-$, respectively, the values of these functions as Z tends to $v_{\parallel}$ lying on the real axis in the complex x plane from above or below the real axis. The following Plemelj formulas are valid:

$$\sigma^+(v_{\parallel}) - \sigma^-(v_{\parallel}) = \mathcal{M}(v_{\parallel}), \qquad (84a)$$

$$\pi i[\sigma^+(v_{\parallel}) + \sigma^-(v_{\parallel})] = \int_{-\infty}^{*\infty} \frac{\mathcal{M}(\beta)\,d\beta}{(\beta - v_{\parallel})},$$
$$\equiv \int_{-1}^{*+1} \frac{\mathcal{M}(\beta)\,d\beta}{(\beta - v_{\parallel})}, \qquad (84b)$$

$$\mathbf{q}^+(v_{\parallel}) - \mathbf{q}^-(v_{\parallel}) = \mathbf{Q}(v_{\parallel}), \qquad (84c)$$

$$\pi i[\mathbf{q}^+(v_{\parallel}) + \mathbf{q}^-(v_{\parallel})] = \int_{-\infty}^{*\infty} \frac{\mathbf{Q}(\beta)\,d\beta}{(\beta - v_{\parallel})},$$
$$\equiv \int_{-1}^{*+1} \frac{\mathbf{Q}(\beta)\,d\beta}{(\beta - v_{\parallel})}, \qquad (84d)$$

$$\mathbf{f}^+(v_{\parallel}) - \mathbf{f}^-(v_{\parallel}) = \mathbf{F}(v_{\parallel}), \qquad (84e)$$

$$\pi i[\mathbf{f}^+(v_{\parallel}) + \mathbf{f}^-(v_{\parallel})] = \int_{-\infty}^{*\infty} \frac{\mathbf{F}(\beta)\,d\beta}{(\beta - v_{\parallel})},$$
$$\equiv \int_{-1}^{*+1} \frac{\mathbf{F}(\beta)\,d\beta}{(\beta - v_{\parallel})}. \qquad (84f)$$

We see that with these definitions, Eq. (82) becomes

$$[\mathcal{J} - 2\pi i\sigma^+(v_{\parallel})]\mathbf{q}^+(v_{\parallel}) - \mathbf{f}^+(v_{\parallel})$$
$$= [\mathcal{J} - 2\pi i\sigma^-(v_{\parallel})]\mathbf{q}^-(v_{\parallel}) - \mathbf{f}^-(v_{\parallel}). \quad (85)$$

It follows that

$$[\mathcal{J} - 2\pi i\sigma(z)]\mathbf{q}(z) - \mathbf{f}(z) \equiv \chi(z), \quad \text{say,}$$

is a holomorphic function in the whole of the complex Z plane. As a consequence of Liouville's theorem, it is a constant. By considering $|Z| \to \infty$ and by use of Eq. (81), we see that $\chi(Z) \equiv 0$. Hence

$$\mathbf{q}(z) = [\mathcal{J} - 2\pi i\sigma(z)]^{-1}\mathbf{f}(z) \equiv j(z)\mathbf{f}(z)/\det[\mathcal{L}(z)], \quad (86)$$

where $j(Z)$ is the cotensor of $\mathcal{J} - 2\pi i\sigma(Z)$. We note that the determinant of $\mathcal{L}(Z)$ is just the determinant on the left-hand side of Eq. (51).

Since $\mathbf{q}(Z)$ is everywhere sectionally holomorphic, it follows that, whenever $Z = Z_j$ $(j = 1, \cdots, P)$,

$$\det [\mathfrak{L}(z_j)] = 0, \qquad (87)$$

then we must also have $j(Z_j)\mathbf{f}(Z_j) = 0$, $(j = 1, \cdots, P)$ at the same set Z_j of points.

Thus, we can now find all the arbitrary constants which enter into Eq. (56) and are included in $\mathbf{f}(Z)$, i.e., all the $\rho(\beta_j)$ $(j = 1, \cdots, P)$ can be determined and hence the $\mathbf{R}_\perp(\beta_j)$.

Knowing $\mathbf{q}(Z)$, we see from Eq. (84c) that

$$\mathbf{Q}(v_\parallel) = \frac{j^+(v_\parallel)\mathbf{f}^+(v_\parallel)}{\det [\mathfrak{L}^+(v_\parallel)]} - \frac{j^-(v)\mathbf{f}^-(v_\parallel)}{\det [\mathfrak{L}^-(v_\parallel)]}. \qquad (88)$$

We must now determine $\lambda(\beta, \mathbf{p}_\perp)$ and $\mathcal{E}_\perp(\pm 1)$. Use of Eqs. (88), (59), and (60) enables us, by integrating with respect to β, to obtain two equations containing only the two unknowns $\mathcal{E}_\perp(\pm 1)$. In principle, these equations can then be solved to yield $\mathcal{E}_\perp(\pm 1)$ as an explicit function of f_0.

Knowing $\mathcal{E}_\perp(\pm 1)$, making use of Eq. (84d), (87), and (88), we see from Eq. (63) that it is again possible, in principle, to determine $\lambda(v_\parallel, \mathbf{p}_\perp)$. Thus, our task is now complete. We have demonstrated that all the unknown functions which enter the problem can be determined. As a consequence, the solution of the stated initial-value problem does yield a completeness proof of the set of eigenfunctions of Sec. 3.

5. DISCUSSION

In the absence of any ambient magnetic or electric fields, we have shown that an initial-value problem for a collisionless, relativistic plasma can be solved taking into account the full set of Maxwell's equations. The class of continuum waves which results differs from the continuum class obtained for a nonrelativistic plasma. In the relativistic case, this class is limited to $-1 \leq \beta \leq 1$, while in the nonrelativistic case, the corresponding limitation is $-\infty \leq \beta \leq \infty$. Thus, the nonrelativistic plasma has, in reality, a spurious continuum of waves for β real and $|\beta| > 1$. The techniques employed by van Kampen,[2] Case,[3] and Zelazny[4] for solving the nonrelativistic initial-value problem have to be modified to take account of the finite nature of the velocity of light. The consequence of this fact gives rise to two discrete subsets of eigenfunctions which have real eigenvalues β. The eigenfunctions for these two subsets differ depending upon whether $|\beta| \leq 1$ or $|\beta| > 1$. This particular point is not brought to light by the corresponding nonrelativistic initial-value problem, where the discrete subset of real β values with $|\beta| > 1$ does not arise. It is clear that the physics of the situation demands that these two subsets differ, since for $|\beta| \leq 1$, there are always some particles which can resonate with the wave, while for $|\beta| > 1$, there are no particles which can resonate. (See also Felderhof.[11])

The problem also demonstrates that coupling exists between the space-charge waves and the transverse waves. There are two simple ways to see that such a coupling is to be expected. If a time-varying perturbation space-charge $\rho(\mathbf{X}, t)$ is generated, then a perturbation current density $\mathbf{j}(\mathbf{X}, t)$ will be induced through

$$\partial \rho / \partial t + \nabla \cdot \mathbf{j} = 0. \qquad (89)$$

The other method is to note that a perturbation electrostatic potential $\phi(\mathbf{x}, t)$ will induce a perturbation electromagnetic vector potential $\mathbf{A}(\mathbf{x}, t)$ because of the gauge condition

$$c^{-1}(\partial \varphi / \partial t) + \nabla \cdot \mathbf{A} = 0. \qquad (90)$$

It should be obvious that the two approaches are equivalent. In view of the formidable amount of mathematics involved, no account has been given of the manner in which the coupling perturbs the purely longitudinal and purely transverse waves. A simple discussion of wave coupling, neglecting initial-value behavior, has been given elsewhere (Kahn,[9] nonrelativistic treatment; Lerche,[10] relativistic treatment).

We also wish to point out that no statement can be made concerning completeness from the above analysis in the presence of ambient magnetic and electric fields, since it is well known that such fields increase the number of possible modes.

Finally, we state that the initial-value problem which has been discussed is one of the simplest we have been able to think of. The complications which arise when the system is spatially inhomogeneous, possesses equilibrium currents and space-charge density, has a streaming velocity, or is embedded in ambient fields, are frankly uninviting.

With these remarks in mind, it should be obvious that the above calculation is less than a full resolution of the initial-value problem for a collisionless, relativistic plasma. Its advantage is that it demonstrates several interesting physical points without involving an extremum of complicated mathematics.

APPENDIX

When the roots of Eq. (51) are not simple but are say, of multiplicity m_j $(j = 1, \cdots, R)$ $R \leq P$, it

means that the eigenfunctions are of a type different than Eq. (52).

In particular, we must use as eigenfunctions the type of function[13]

$$f_j \sim (\partial^r/\partial\beta_j^r)[e^{-ikc\beta_j t}\varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp)], \quad (j = 1, \cdots, R).$$
(A1)

Then for every discrete eigenfunction β_j we now have to determine m_j arbitrary constants $S_r(\beta_j)$: $r = 0, \cdots, m_j - 1$.

The general solution equivalent to Eq. (55) becomes

---

[13] G. Backus, J. Math Phys. **1**, 178 (1960).

in this case

$$f_1(p_\parallel, \mathbf{p}_\perp, \mathbf{k}, t) = \sum_{j=1}^{P} \sum_{r=0}^{m_j-1} \frac{\partial^r}{\partial\beta_j^r} [e^{-ikc\beta_j t}\varphi_j(\beta_j, p_\parallel, \mathbf{p}_\perp)]$$
$$+ \int_{-1}^{+1} \varphi(\beta, p_\parallel, \mathbf{p}_\perp) \, d\beta. \quad \text{(A2)}$$

Instead of Eq. (87), we would obtain

$$(\partial^r/\partial\beta_j^r)[j(\beta_j)\mathbf{f}(\beta_j)] = 0,$$
$$(j = 1, \cdots P; \quad r = 0, \cdots, m_j - 1). \quad \text{(A3)}$$

Thus we can determine all the unknowns in the discrete β-spectrum part of Eq. (A2). As a consequence, it is sufficient to treat the roots of Eq. (51) as simple in order to demonstrate the completeness theorem.

Coordinate and Momentum Observables in Axiomatic Quantum Mechanics

STANLEY GUDDER

Department of Mathematics, University of Wisconsin, Madison, Wisconsin

An axiomatic model for quantum mechanics is formulated using physically significant axioms. The model contains a slight strengthening of Mackey's first six axioms, together with two axioms which ensure the existence of coordinate and momentum observables. The symmetries or rigid motions are an essential part of the structure, and a link is constructed between these and the quantum proposition system. Coordinate and momentum observables are defined in terms of abstract coordinate systems and one-parameter groups of motions. It is then shown that as far as the statistical properties of these observables in certain canonical states are concerned, the abstract model may be represented by the usual Hilbert space formulation. Spectral properties of σ-homomorphisms are also investigated. Also included are two appendices of a technical nature: the first considers one-parameter groups of derivables, and the second absolutely continuous σ-homomorphisms.

1. INTRODUCTION

AT the foundations of quantum mechanics there seem to be three basic concepts: states, propositions, and observables. Using these concepts, essentially two approaches to axiomatic quantum mechanics have evolved. The first approach, due to Jordan, Wigner, and von Neumann,[1] takes the observables, while the second approach (of Birkhoff and von Neumann[2]) takes the propositions as its basic axiomatic elements. These two approaches have matured into distinct axiomatic schools of thought. The first approach was developed further by Segal[3] and forms the basis of the quantum field theories of Haag and Wightman.[4] The second approach has been studied to some extent by Bodiou[5] and Mackey[6] and more recently by Piron, Jauch, Emch, Guenin,[7] and others. This paper is concerned with the second approach and is based on the quantum-mechanical framework in Mackey's book.[6] We refer the reader to this reference for details that are not included here.

Mackey's first six axioms postulate that the proposition system L is an orthocomplemented, partially-ordered set with a full strongly-convex set of states, while his seventh axiom states that L is isomorphic to the lattice of all closed subspaces of a Hilbert space. Although the first six axioms are intuitive and reasonable from a physical point of view, it is admitted that the seventh is *ad hoc*. Attempts have been made to add further physically justifiable axioms to these first six so that axiom seven may be deduced from previous ones and thus need not be postulated.[8] However, the physical justifications for many of these additional axioms are extremely doubtful and there seems to be no experimental evidence supporting their existence. For example, almost all additional axioms included the postulate that L is a complete atomic lattice. There seems to be little experimental evidence that L is complete and atomic, while the fact that L is a lattice seems to contradict the Heisenberg uncertainty principle if the propositions of L are interpreted in the natural way. Of course, Mackey's axiom seven also has these defects.

To clarify this last assertion concerning the lattice structure of L, we must examine more closely what is meant by a proposition system. A proposition is usually interpreted as a meaningful statement made about a physical system which can be both verified and refuted by a definite experiment. We now form a proposition system L by adding to these physically significant propositions two "ideal" propositions, the absurd proposition 0 which is never true, and the self-evident proposition 1 which is always true. The absurd proposition corresponds physically to a statement which is refuted by every relevant experiment, while the self-evident proposition is verified by every relevant experiment. Now if a and b are propositions, the natural interpretation of $a \wedge b$ is that

[1] P. Jordan, E. Wigner, and J. von Neumann, Ann. Math. **35**, 29 (1934).

[2] G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).

[3] A bibliography of Segal's work is given in *Mathematical Problems of Relativistic Physics* (American Mathematical Society, Providence, Rhode Island, 1963).

[4] Cf. R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962). A. Wightman, *Les problemes mathematiques de la theorie quantique des champs* (Centre National Recherche Scientifique, Paris, 1959).

[5] G. Bodiou, *Theory dialectique des probabilities* (Gauthier-Villars, Paris, 1964).

[6] G. Mackey, *The Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, Inc., New York, 1963).

[7] Compare with C. Piron, Helv. Phys. Acta **37**, 439 (1964); G. Emch and C. Piron, J. Math. Phys. **4**, 469 (1963); M. Guenin, *ibid.* **7**, 271 (1966); J. Jauch and C. Piron, Helv. Phys. Acta **36**, 827 (1963).

[8] Compare with Ref. 5, the first reference of Ref. 7, and N. Zierler, Pac. J. Math. **12**, 1151 (1961).

$a \wedge b$ is the proposition which is true whenever both a and b are true. The author contends that one should not assume *a priori* that $a \wedge b$ exists as a member of L. Consider, for example, a spinless free particle p in one dimension. Let a be the proposition: p is between x_0 and $x_0 + \Delta x$ cm from the origin. Let b be the proposition: p has momentum between p_0 and $p_0 + \Delta p$ gm cm/sec where $\Delta x \Delta p \ll h$. Now if $a \wedge b$ exists, it must be the absurd proposition since, by the Heisenberg uncertainty principle, there is no experiment capable of verifying $a \wedge b$. But any relevant experiment measuring the position and momentum of p will give position and momentum values between certain limits, depending on the accuracy of the apparatus, and such an experiment will not *always* refute $a \wedge b$. Hence we cannot even say that $a \wedge b$ is the absurd proposition. Thus the author contends that $a \wedge b$ is not a meaningful proposition and should not *a priori* be assumed to be a member of L. It can, of course, happen that $a \wedge b$ does exist as a member of L for some noncompatible propositions a and b, but this should be a deduction of the theory and not a postulate.

Although much literature has been devoted to this second axiomatic approach, until now it does not seem to have been developed to the point where coordinate and momentum observables have been defined. In this paper we shall start with Mackey's first six axioms and add three more which seem to have physical justification. The first (Axiom 1) is physically obvious, while the second and third (Postulates II and III) are needed only to ensure the existence of coordinate and momentum observables. We are not able to deduce Axiom 7 from these mild axioms; however, we are able to do something almost as good. If A is an observable, the only experimentally accessible properties of A are its statistical properties. For example, in the Hilbert space formulation of quantum mechanics, the experimentally meaningful properties of A are given by its average values $\langle \psi, A\psi \rangle$ for different states ψ. Now if L_1 is a proposition system satisfying our physically significant axioms, we show that there is a map $m \rightarrow \hat{m}$ from a subset M_1 of the states of L_1 (we call M_1 the canonical states) into the vectors of a Hilbert space H which preserves strongly-convex combinations, and for any position or momentum observable A there is a self-adjoint operator $\hat{A}$ on H such that the average value of A in the state m is given in the usual way by $\langle \hat{m}, \hat{A}\hat{m} \rangle$. For concrete examples of physical significance M_1 turns out to be a very large set; in fact, we show in one example that $\hat{M}_1$ spans H. We thus show that, as far as the statistical properties of a quantum-mechanical system in a canonical state are concerned, our abstract axiomatic model may be represented by the usual Hilbert space formulation.

In Sec. 2 the axiomatic model is formulated and physical reasons are given for the axioms. Section 3 introduces the concepts of coordinate systems and motions. These concepts are necessary in defining the coordinate and momentum observables. The reader should notice that, although it is easy to define the coordinate observables, it is not at all obvious how the momentum observables should be defined. For this reason some space is devoted to showing that the one-parameter motions are invariants of the motions of the system, and it is, therefore, from these that the momentum observable should be defined. In Sec. 4 the coordinate and momentum observables are defined and the main theorems of this paper are proved. Section 5 derives some properties of σ-homomorphisms which may be useful. In Appendices A and B technical mathematical proofs are given which are needed in Sec. 4. The mathematical techniques in these Appendices are fairly standard but are included for completeness.

The author would like to point out that the present paper deals specifically with Euclidean-like spaces only. We are not concerned here with more complicated spaces; in particular, relativistic theory will not be considered. However, the axiomatic system that is given is quite general, and it is possible that the theory can be extended to these cases.

2. AXIOMATIC MODEL

Let $L = \{a, b, c, \cdots\}$ be an orthocomplemented, partially ordered set with a full strongly-convex set of states $M = \{m, m_1, m_2, \cdots\}$. The elements of L are referred to as *propositions*. The states may be thought of as giving the condition of the quantum-mechanical system, and $m(a)$ may be thought of as the probability that the proposition a is true in the state m. If $a \leq b'$, then we say that a and b are *disjoint* and write $a \perp b$. If $a \perp b$, we denote $a \vee b$ by $a + b$. We say that a, b *split* and write $a \leftrightarrow b$ if there are mutually disjoint propositions a_1, b_1, c such that $a = a_1 + c$ and $b = b_1 + c$. We may think of propositions which split as physically being propositions which are simultaneously verifiable. It is obvious that $a \vee b$ exists if $a \leftrightarrow b$. Now if a, b, c mutually split, physically one would expect $a \leftrightarrow b \vee c$. One can give examples[9] in which this result does not hold. If L satisfies

Axiom: For every a, b, $c \in L$ which mutually split, $a \leftrightarrow b \vee c$

[9] A. Ramsey, J. Math. Mech. **15**, 277 (1966).

then (L, M) is called a *proposition system.* Let Ω be a Hausdorff space. We denote the σ-algebra of subsets of Ω generated by the open sets by $B(\Omega)$. Let (L, M) be a proposition system. A σ-homomorphism $x: B(\Omega) \rightarrow L$ is a map which satisfies (i) $x(\Omega) = 1$; (ii) $x(\Lambda) \perp x(\Gamma)$ if $\Lambda \cap \Gamma = \phi$; (iii) $x(\bigcup \Lambda_i) = \Sigma x(\Lambda_i)$, $i = 1, 2, \cdots$, if $\Lambda_i \cap \Lambda_j = \phi$, $i \neq j$. If Ω is the real line R, then x is called an *observable.* If Ω is the complex plane C, then x is called a *derivable.* Two σ-homomorphisms $x_1: B(\Omega_1) \rightarrow L$, $x_2: B(\Omega_2) \rightarrow L$ are *simultaneous* (written $x_1 \leftrightarrow x_2$) if every element in the range of x_1 splits with every element in the range of x_2. The spectrum $\sigma(x)$ of a σ-homomorphism $x: B(\Omega) \rightarrow L$ is the smallest closed subset $\Lambda \in B(\Omega)$ such that $x(\Lambda) = 1$. A derivable x is unitary if $\sigma(x) \subset \{\lambda : |\lambda| = 1\}$. The spectrum of an observable represents the allowable values that an observable may attain. If $f: \Omega_1 \rightarrow \Omega_2$ is a Borel function and $x: B(\Omega_1) \rightarrow L$ a σ-homomorphism, then $f(x)$: $B(\Omega_2) \rightarrow L$ is the σ-homomorphism $f(x)(\Lambda) = x(f^{-1}(\Lambda))$ for every $\Lambda \in B(\Omega_2)$. It is clear that $x \leftrightarrow f(x)$. If $x: B(\Omega) \rightarrow L$ is a σ-homomorphism, $u: \Omega \rightarrow C$, a Borel function, and $m \in M$, the *expectation* or *average value* $m[u(x)]$ of the derivable $u(x)$ in the state m is

$$m[u(x)] = \int_C u(\omega) m[x(d\omega)].$$

It follows from results in Ref. 10 that a sequence of derivables (x_i) are simultaneous if and only if there are complex Borel functions (u_i) and a derivable x such that $x_i = u_i(x)$, $i = 1, 2, \cdots$. Now let ϕ be a complex n-dimensional Borel function and x_i, $i = 1, 2, \cdots, n$, simultaneous derivables. If $x_i = u_i(x)$, we define the derivable $\phi(x_1, \cdots, x_n)$ as

$$\phi(x_1, \cdots, x_n)(E) = x\{\lambda : \phi(u_1(\lambda), \cdots, u_n(\lambda)) \in E\}$$

for every $E \in B(C)$. In particular, $x_1 x_2(E) = x\{\lambda : u_1(\lambda) u_2(\lambda) \in E\}$. It can be shown[11] that $\phi(x_1, \cdots, x_n)$ is well defined, that is, independent of x and u_i, $i = 1, 2, \cdots, n$.

Let S be the *physical space* corresponding to a laboratory experiment. For example, in the case of a system with a finite number of degrees of freedom, S would be the "configuration space" of our system. Mathematically we only assume that S is a locally compact Hausdorff space with second countability. Let G be the *group of rigid motions* on S. For example, in the finite degrees of freedom case G would be the group generated by the translations, rotations, and reflections in a finite-dimensional Euclidean space.

Mathematically we assume that G is a locally compact group with second countability and that G is a continuous effective transitive transformation group on S. That is, there exists a map from $G \times S$ onto S denoted by $(g, s) \rightarrow gs$, $g \in G$, $s \in S$, such that (i) if $s_1, s_2 \in S$, there is a $g \in G$ such that $s_1 = gs_2$ (transitivity); (ii) for every $g \in G$, $s \rightarrow gs$ is a homeomorphism of S with itself; (iii) $g_1(g_2(s)) = g_1 g_2(s)$ for every $g_1, g_2 \in G$, $s \in S$; (iv) $g(s) = s$ for every $s \in S$ if and only if $g = e$, where e is the identity of G (effectiveness). Note that our finite degrees of freedom example satisfies these conditions.

We are now ready to postulate our axiomatic model.

I. The propositions and states of a quantum-mechanical system form a proposition system (L, M).

II. There is a σ-homomorphism $X: B(S) \rightarrow L$ (called the *position σ-homomorphism*) such that $g \rightarrow m[X(g\Lambda)]$ is continuous for every $m \in M$, $\Lambda \in B(S)$.

III. For every $g \in G$ there is a unitary derivable $x_g: B(C) \rightarrow L$ such that
(i) $g \rightarrow m(x_g)$ is continuous for every $m \in M$;
(ii) if $g_1 g_2 = g_2 g_1$, then $x_{g_1} \leftrightarrow x_{g_2}$ and $x_{g_1 g_2} = x_{g_1} x_{g_2}$;
(iii) $X(\Lambda) = X(g\Lambda)$ if and only if $X(\Lambda) \leftrightarrow x_g(E)$ for every $E \in B(C)$.

Notice that $m(x_g)$ is the expectation of x_g defined earlier and $x_{g_1} x_{g_2}$ is the product of simultaneous derivables which was also defined earlier. The physical justification of I is discussed in the literature. Axiom II is necessary to ensure the existence of coordinate observables. $X(\Lambda)$ may be thought of as the proposition that the position of the system is in a set $\Lambda \subset S$. Axiom II gives a way of representing the position of the system in the proposition system. Axiom III is necessary to ensure the existence of momentum observables and gives a way of exhibiting the group action in the proposition system. Axiom III, condition (ii), states that if g_1 and g_2 commute, then their actions on S can be observed at the same time. That is, the action described by g_1 does not affect the action described by g_2, and vice versa. The rest of Axiom III, condition (ii), states that the action described by the product of g_1 and g_2 is the product of the actions. Axiom III, condition (iii), just gives two ways of saying that the proposition $X(\Lambda)$ is unaffected by the action of g on S. The reader may wonder why x_g is a unitary derivable and not, as would seem more natural, an observable. The fact is that our theory can be carried through with the assumption that x_g is an observable. However, we assume x_g is a unitary derivable because we want our theory to include the

---
[10] V. Varadarajan, Comm. Pure Appl. Math. **15**, 217 (1962).
[11] S. Gudder, Trans. AMS **119**, 428 (1965).

usual quantum-mechanical formulation, in which it is usually assumed that G has a continuous unitary representation on the state space.

The reader should note that Axioms I, II, III, condition (i), and Axiom III, condition (ii), apply to classical systems as well as quantum systems. Thus as far as these axioms are concerned, L might be a Boolean σ-algebra (i.e., all propositions split). However, Axiom III, condition (iii), ushers us into a purely quantum-mechanical realm. Specifically, we use the fact that $X(\Lambda) \leftrightarrow x_g(E)$ for all $E \in B(C)$ to describe the situation that the propositions $X(\Lambda)$ and $X(g\Lambda)$ are identical. Thus for the case in which $X(\Lambda) \neq X(g\Lambda)$, which is, of course, quite common, we must have $x(\Lambda) \leftrightarrow x_g(E)$ for some $E \in B(C)$, and the latter is a purely quantum-mechanical phenomenon. To emphasize this fact we give the following theorem, which shows that L cannot be a Boolean σ-algebra unless the space S is trivial. To avoid pathologies we always assume $0 \neq 1$ in L (i.e., L has more than one element).

Theorem 2.1: Suppose we have a system satisfying Axioms I, II, and III. If L is a Boolean σ-algebra, then S contains only one point.

Proof: Suppose L is a Boolean σ-algebra and s_1 and s_2 are distinct points of S. Since S is Hausdorff, there are disjoint neighborhoods Λ_1, Λ_2 such that $s_1 \in \Lambda_1$ and $s_2 \in \Lambda_2$. By transitivity there is a $g \in G$ such that $s_2 = gs_1$. Since $s \to gs$ is a homeomorphism, there is a neighborhood Λ_3 of s_1 such that $g\Lambda_3 \subset \Lambda_2$. Letting $\Lambda = \Lambda_1 \cap \Lambda_3$, we see that $\Lambda \subset \Lambda_1$ is a neighborhood of s_1, $g\Lambda$ is a neighborhood of s_2, and $g\Lambda \subset \Lambda_2$. Now applying Axiom III, condition (iii), $X(\Lambda) = X(g\Lambda)$. But since Λ and $g\Lambda$ are disjoint sets, we must have $X(\Lambda) \perp X(g\Lambda)$ and hence $X(\Lambda) = 0$. Since S satisfies second countability, it easily follows that $X(S) = 0$, which is impossible.

We shall see later (Theorem 4.1) that as a consequence of Axiom III, condition (iii), the conjugate coordinate and momentum observables are not simultaneous.

It is easily seen that our definition of a proposition system is more general than that given in previous investigations. For example, one can show that the axioms of Emch and Piron (first two references in Ref. 7) imply our axioms concerning L and that our definition of splitting is equivalent to their definition of compatibility in their more restrictive axiomatic system. Also our definition of a state is more general, since Emch and Piron require, in addition to our state axioms, the condition that $m(a) = m(b) = 1$ implies $m(a \wedge b) = 1$.

3. COORDINATE SYSTEMS AND MOTIONS

In this section we develop material preliminary to defining coordinate and momentum observables. A *coordinate system* for physical space S is a collection of real-valued continuous functions $\{f_\alpha : \alpha \in A\}$ on S such that $f_\alpha(s_1) = f_\alpha(s_2)$ for all $\alpha \in A$ implies $s_1 = s_2$. For simplicity, the coordinate systems that we are using here are generalizations of rectangular coordinate systems. If S is a sphere, for example, and we were using angular coordinates, then the coordinate functions would be maps into the unit circle. The arguments which follow may easily be altered to include such situations. For more elaborate spaces and coordinate systems one would probably assume that S is an analytic manifold and G a Lie group. For mathematical simplicity we do not consider such spaces in this paper. If $\{f_\alpha : \alpha \in A\}$ is a coordinate system, a subset G_β, $\beta \in A$, of the group of rigid motions G is a *motion in the β direction* if $f_\alpha(gs) = f_\alpha(s)$ for all $s \in S$, $g \in G_\beta$, and $\alpha \in A$ such that $\alpha \neq \beta$; and $f_\beta(g_1 g_2 s) = f_\beta(g_1 s) + f_\beta(g_2 s) - f_\beta(s)$, for all $s \in S$, $g_1, g_2 \in G_\beta$. $G_\beta \subset G$ is a *group motion* in the β direction if G_β is a closed Abelian group which is a motion in the β direction.

Theorem 3.1: If $\{f_\alpha : \alpha \in A\}$ is a coordinate system and G_β a motion in the β direction, then G_β is contained in a group motion in the β direction.

Proof: We first show G_β is commutative. If g_1, $g_2 \in G_\beta$, then for $\alpha \neq \beta$, $f_\alpha(g_1 g_2 s) = f_\alpha(g_1(g_2 s)) = f_\alpha(g_2 s) = f_\alpha(s) = f_\alpha(g_2 g_1 s)$. Also $f_\beta(g_1 g_2 s) = f_\beta(g_2 g_1 s)$ and hence $g_1 g_2 s = g_2 g_1 s$ for all $s \in S$. Therefore, by effectiveness, $g_1 g_2 = g_2 g_1$. Now $G_\beta \cup \{e\}$ is a motion in the β direction. We next show that $G_\beta \cup G_\beta^{-1} \cup \{e\}$ is a motion in the β direction. If $g \in G_\beta$, then $f_\alpha(s) = f_\alpha(gg^{-1}s) = f_\alpha(g^{-1}s)$ for $\alpha \neq \beta$. If g_1 is also in G_β, then

$$f_\beta(gs) = f_\beta(gg_1 g_1^{-1}s),$$
$$= f_\beta(gg_1^{-1}s) + f_\beta(g_1 g_1^{-1}s) - f_\beta(g_1^{-1}s).$$

Hence

$$f_\beta(gg_1^{-1}s) = f_\beta(gs) + f_\beta(g_1^{-1}s) - f_\beta(s).$$

Also

$$f_\beta(g^{-1}s) = f_\beta(g^{-1}g_1 g_1^{-1}s) = f_\beta(g_1 g^{-1}g_1^{-1}s),$$
$$= f_\beta(g_1 g_1^{-1}s) + f_\beta(g^{-1}g_1^{-1}s) - f_\beta(g_1^{-1}s).$$

Therefore

$$f_\beta(g^{-1}g_1^{-1}s) = f_\beta(g^{-1}s) + f_\beta(g_1^{-1}s) - f_\beta(s),$$

and

$$G_\beta \cup G_\beta^{-1} \cup \{e\}$$

is a motion in the β direction. Now let G_β' be the set of all finite products of elements of $G_\beta \cup G_\beta^{-1} \cup \{e\}$.

Then G'_β is an Abelian subgroup of G. Using induction, the next calculation shows that G'_β is a motion in the β direction. If g_1, $g_2 \in G'_\beta$, then $f_\alpha(g_1 g_2 s) = f_\alpha(g_2 s) = f_\alpha(s)$, $\alpha \neq \beta$. If g_3, g_4 are also in G'_β, then, carrying out the obvious steps, we have

$$f_\beta((g_1 g_2)(g_3 g_4)s) = f_\beta(g_1 g_2 s) + f_\beta(g_3 g_4 s) - f_\beta(s).$$

We now show that the closure of any motion is a motion. Let g_i be a sequence of elements in a motion in the β direction and suppose $g_i \to g$. Then $f_\alpha(gs) = \lim f_\alpha(g_i s) = f_\alpha(s)$, $\alpha \neq \beta$. Now if g_i^0 is a sequence in this same motion and $g_i^0 \to g_0$, we have

$$f_\beta(g g_0 s) = \lim f_\beta(g_i g_i^0 s),$$
$$= \lim f_\beta(g_i s) + \lim f_\beta(g_i^0 s) - f_\beta(s),$$
$$= f_\beta(gs) + f_\beta(g_0 s) - f_\beta(s).$$

Therefore the closure of G'_β is a closed Abelian subgroup of G which is a motion in the β direction containing G_β.

We see that a group motion G_β is a subgroup of G which corresponds physically to a movement of the system in the β direction and must therefore be connected in some sense to the momentum in the β direction. Since there may be many group motions in the β direction, and since the momentum should be uniquely defined, we look for some kind of invariant of the group motion, that is, a property which is independent of which particular group motion we consider. It will turn out that such an invariant is a one-parameter subgroup of the group motion. P_β is a *one-parameter motion in the β direction* if P_β is a motion in the β direction and if there is a continuous map $\lambda \to g_\lambda$ from $(-\infty, \infty)$ onto P_β such that $g_{\lambda+\mu} = g_\lambda g_\mu$ and $f_\beta(s_1 g) - f_\beta(s_1) = f_\beta(s_2 g) - f(s_2)$ for all s_1, $s_2 \in S$, $g \in P_\beta$. The next theorem characterizes one-parameter motions.

Theorem 3.2: A continuous map $\lambda \to g_\lambda$ from $(-\infty, \infty)$ into G is a one-parameter motion in the β direction if and only if there is a real number c such that $f_\beta(g_\lambda s) = c\lambda + f_\beta(s)$ and $f_\alpha(g_\lambda s) = f_\alpha(s)$, $\alpha \neq \beta$, for all $s \in S$.

Proof: To prove necessity, fix $s \in S$ and define the function $h:(-\infty, \infty) \to R$ by $h(\lambda) = f_\beta(g_\lambda s)$. Then

$$h(\lambda + \mu) = f_\beta(g_{\lambda+\mu} s) = f_\beta(g_\lambda g_\mu s),$$
$$= f_\beta(g_\lambda s) + f_\beta(g_\mu s) - f_\beta(s) = h(\lambda) + h(\mu) - f_\beta(s).$$

Since h is continuous, it follows from a well-known theorem in real variables that h must have the form $h(\lambda) = c\lambda + f_\beta(s)$ for some $c \in R$. Of course $f_\alpha(g_\lambda s) = f_\alpha(s)$, $\alpha \neq \beta$, since $\lambda \to g_\lambda$ is a one-parameter motion.

To prove sufficiency, we have $f_\alpha(g_{\lambda+\mu} s) = f_\alpha(s) = f_\alpha(g_\mu s) = f_\alpha(g_\lambda g_\mu s)$ for $\alpha \neq \beta$. Also

$$f_\beta(g_\lambda g_\beta s) = f_\beta(g_\lambda(g_\mu s)) = c\lambda + f_\beta(g_\mu s),$$
$$= c(\lambda + \mu) + f_\beta(s) = f_\beta(g_{\lambda+\mu} s).$$

Therefore $g_{\lambda+\mu} s = g_\lambda g_\mu s$ for all $s \in S$ and, by effectiveness, $g_{\lambda+\mu} = g_\lambda g_\mu$. To show we have a motion,

$$f_\beta(g_\lambda g_\mu s) = f_\beta(g_{\lambda+\mu} s) = c(\lambda + \mu) + f_\beta(s),$$
$$= f_\beta(g_\lambda s) + f_\beta(g_\mu s) - f_\beta(s).$$

We will need the next corollary for our later work.

Corollary 3.3: Let $\{f_\alpha : \alpha \in A\}$ be a coordinate system, P_β a one-parameter motion in the β direction, and X a position σ-homomorphism. Then there is a real number c such that $X[g_\lambda f_\beta^{-1}(E)] = f_\beta(X)(E + c\lambda)$ for every $g_\lambda \in P_\beta$, $E \in B(R)$.

Proof: Let c be the real number found in Theorem 3.2. First note that $f_\beta(s) = f_\beta(g_\lambda g_\lambda^{-1} s) = f_\beta(g_\lambda^{-1} s) + c\lambda$. Therefore $f_\beta(g_\lambda^{-1} s) = f_\beta(s) - c\lambda$. We now show that $g_\lambda f_\beta^{-1}(E) = f_\beta^{-1}(E + c\lambda)$ for every $g_\lambda \in P_\beta$, $E \in B(R)$. If $s \in g_\lambda f_\beta^{-1}(E)$, then $f_\beta(s) = f_\beta(g_\lambda s_1)$ where $s_1 \in f_\beta^{-1}(E)$. Hence $f_\beta(s) = f_\beta(s_1) + c\lambda \in E + c\lambda$ and $s \in f_\beta^{-1}(E + c\lambda)$. If $s \in f_\beta^{-1}(E + c\lambda)$, then $f_\beta(s) \in E + c\lambda$ and $f_\beta(g_\lambda^{-1} s) = f_\beta(s) - c\lambda \in E$. Hence $g_\lambda^{-1} s \in f_\beta^{-1}(E)$ and $s \in g_\lambda f_\beta^{-1}(E)$. We therefore have

$$X[g_\lambda f_\beta^{-1}(E)] = X[f_\beta^{-1}(E + c)] = f_\beta(X)(E + c\lambda).$$

If $G_0 = \{g_\alpha : \alpha \in A\}$ is the motion $g_\alpha = e$ for all $\alpha \in A$, then G_0 is called the *trivial motion*. It is easily seen that a one-parameter motion is the trivial motion if and only if $c = 0$. Two one-parameter motions $P_1 = \{g_\lambda^1\}$, $P_2 = \{g_\lambda^2\}$ are *equivalent* (written $P_1 \sim P_2$) if there is a nonzero real number μ such that $g_\lambda^2 = g_{\mu\lambda}^1$ for all $\lambda \in (-\infty, \infty)$. We easily see that $\sim$ is an equivalence relation. Therefore the set of one-parameter motions is partitioned into equivalence classes, with each one-parameter motion contained in one and only one class. Physically two equivalent one-parameter motions are essentially the same, since one results from the other by a linear change of scale. We now characterize equivalences.

Corollary 3.4: Two nontrivial one-parameter motions are equivalent if and only if they are in the same direction.

Proof: Necessity is obvious. To prove sufficiency, let $\{f_\alpha : \alpha \in A\}$ be a coordinate system and let $P_1 = \{g_\lambda^1\}$, $P_2 = \{g_\lambda^2\}$ be nontrivial one-parameter motions in the β direction. By Theorem 3.2, there are constants $c_i \neq 0$ such that $f_\beta(g_\lambda^i s) = c_i \lambda + f_\beta(s)$, $i = 1, 2$.

Letting $\mu = c_2/c_1$, we have

$$f_\beta(g^1_{\mu\lambda}s) = c_1\mu\lambda + f_\beta(s) = f_\beta(g^2_\lambda s).$$

Also

$$f_\alpha(g^1_{\mu\lambda}s) = f_\alpha(s) = f_\alpha(g^2_\lambda s) \quad \text{for} \quad \alpha \neq \beta.$$

Hence $g^1_{\mu\lambda}s = g^2_\lambda s$ for all $s \in S$ and, by effectiveness, $g^1_{\mu\lambda} = g^2_\lambda$ for all $\lambda \in (-\infty, \infty)$ and $P_1 \sim P_2$.

We thus see that the nontrivial one-parameter motions are invariants which depend only upon the direction. If a group motion admits a nontrivial one-parameter subgroup, then we call it a *conjugate group motion*. If a coordinate function f_β admits a conjugate group motion in the β direction, it is called a *conjugate coordinate function*. Not every group motion is conjugate. Obviously the trivial motion is not conjugate. For a less-trivial example, the group of order three consisting of rotations in the plane of $120°$ is not conjugate. One can even construct examples of coordinate systems which admit no conjugate group motions at all, and hence none of the coordinate functions are conjugate. However, in physical situations of interest (e.g., n-dimensional Euclidean space) it is clear that the usual coordinate functions are conjugate coordinate functions.

4. COORDINATE AND MOMENTUM OBSERVABLES

If $\{f_\alpha : \alpha \in A\}$ is a coordinate system and X the position σ-homomorphism, the *coordinate observables* are the observables $f_\alpha(X)$, $\alpha \in A$. Of course, all coordinate observables are simultaneous. Now suppose f_β is a conjugate coordinate function and P_β the essentially unique nontrivial one-parameter motion corresponding to f_β. Applying Postulate III, there are unitary derivables $x_\lambda = x_{g_\lambda}$, $\lambda \in (-\infty, \infty)$, which satisfy $x_{\lambda+\mu} = x_{g_{\lambda+\mu}} = x_{g_\mu g_\mu} = x_{g_\lambda}x_{g_\mu} = x_\lambda x_\mu$ and $\lambda \to m(x_\lambda)$ is continuous. Thus $\{x_\lambda : \lambda \in (-\infty, \infty)\}$ is a continuous one-parameter group of unitary derivables, and hence $\{x_\lambda\}$ has an infinitesimal generator. (A proof of this fact is given in Appendix A, Theorem I.2.) That is, there exists an observable p_β such that $x_\lambda = e^{i\lambda p_\beta}$, $\lambda \in (-\infty, \infty)$. p_β is called the *momentum observable* in the β direction. $f_\beta(X)$ and p_β are called *conjugate* coordinate and momentum observables. Now it can be shown that p_β is unique up to a multiplicative constant. That is, if p^1_β and p^2_β correspond to two equivalent one-parameter motions, then there is a positive constant c such that $p^2_\beta = cp^1_\beta$. Thus to every conjugate coordinate observable there corresponds essentially one momentum observable. One should note that the coordinate and momentum observables defined here reduce to the usual ones in the Hilbert space formulation of quantum mechanics. The next

theorem proves the important fact that conjugate coordinate and momentum observables are not simultaneous.

Theorem 4.1: If $f_\alpha(X)$ and p_α, $\alpha \in A$, are conjugate coordinate and momentum observables, then $f_\beta(X) \leftrightarrow p_\alpha$, $\beta \neq \alpha$, but $f_\alpha(X) \nleftrightarrow p_\alpha$.

Proof: Let g_λ be a nontrivial one-parameter motion in the α direction, and let x_λ be the corresponding unitary derivables, $\lambda \in (-\infty, \infty)$. Suppose $\beta \neq \alpha$ and $E \in B(R)$. If $s \in f_\beta^{-1}(E)$, then

$$f_\beta(g^{-1}_\lambda s) = f_\beta(s) \in E \quad \text{and} \quad g^{-1}_\lambda s \in f_\beta^{-1}(E).$$

Hence $s \in g_\lambda f_\beta^{-1}(E)$ and thus $f_\beta^{-1}(E) \subset g_\lambda f_\beta^{-1}(E)$ for all $\lambda \in (-\infty, \infty)$. The inclusion in the other direction is trivial and so $f_\beta^{-1}(E) = g_\lambda f_\beta^{-1}(E)$ for all $\lambda \in (-\infty, \infty)$. By Axiom III, condition (iii), we have $f_\beta(X) \leftrightarrow x_\lambda$ for all $\lambda \in (-\infty, \infty)$; and applying Corollary I.6, $f_\beta(X) \leftrightarrow p_\alpha$. Let c be the constant corresponding to g_λ in Theorem 3.2 and suppose that $f_\alpha(X) \leftrightarrow x_\lambda$ for all $\lambda \in (-\infty, \infty)$. By Axiom III, condition (iii), and Corollary 3.3, $f_\alpha(X)(E) = X[g_\lambda f_\alpha^{-1}(E)] = f_\alpha(X)(E + c\lambda)$ for all $\lambda \in (-\infty, \infty)$, $E \in B(R)$. Now $f_\alpha(X)(-\infty, 0] = 1$ since if $f_\alpha(X)(-\infty, 0] < 1$, then $f_\alpha(X)(R) > 1$, which is impossible. Similarly, $f_\alpha(X)(-\infty, -n] = 1$ for $n = 1, 2, \cdots$. Therefore $1 = \bigwedge_n f_\alpha(X)(-\infty, n] = f_\alpha(X)(\bigcap_n (-\infty, -n]) = f_\alpha(X)(\phi) = 0$, a contradiction. Therefore $f_\alpha(X) \nleftrightarrow x_\lambda$ for some $\lambda \in (-\infty, \infty)$ and, by Corollary I.6, $f_\alpha(X) \nleftrightarrow p_\alpha$.

Because of the transitive nature of G on S, it is easily seen that there is a one–one correspondence between points of S and elements of the quotient space G/H, where H is a closed subgroup of G whose elements leave a point of S invariant. If we define Borel sets $B(G/H)$ on G/H in the usual manner (cf. Appendix B), this correspondence preserves Borel sets both ways.[12] Identifying S and G/H, we may think of the position σ-homomorphism as a map $X: B(G/H) \to L$. A σ-finite measure ν on $B(G/H)$ is *quasi-invariant* if $\nu(\Lambda) = 0$ implies $\nu(g\Lambda) = 0$ for every $g \in G$, $\Lambda \in B(G/H)$. Applying Theorem II.3 in Appendix B and identifying S and G/H, there exists a map $m \to \hat{m}$ of M into $L_2(S, \nu)$ such that $m[X(\Lambda)] = \int_\Lambda \hat{m}^2 \, d\nu$. It is shown in Ref. 12 that quasi-invariant measures always exist, and that if ν_1 is any other quasi-invariant measure, then ν and ν_1 have the same null sets and these are precisely the sets whose inverse image under the canonical map have Haar measure zero. It then easily follows that $L_2(S, \nu_1)$ and $L_2(S, \nu)$ are unitarily equivalent. Thus the Hilbert space is independent of

---

[12] G. Mackey, Acta Math. **99**, 265 (1958).

the quasi-invariant measure and is unique up to a unitary equivalence.

Now $\hat{m}$ is a kind of position probability density, and hence $\int \hat{m}(s)\hat{m}(gs)\nu(ds)$ may be thought of as giving the probability that the system has moved from $s \in S$ to gs. Intuitively, one might expect $m(x_g)$ to be this same probability. We say that a state m is *canonical* with respect to the measure ν if

$$m(x_g) = \int \hat{m}(s)\hat{m}(gs)\nu(ds)$$

for every $g \in G$. This definition is not independent of the quasi-invariant measure, so once one has chosen a measure, he must stick to it.

In the examples of quantum-mechanical systems known to the author, there are always an abundance of canonical states. For example, let S be n-dimensional Euclidean space R^n, L the lattice of orthogonal projections on $L_2(S, \nu)$, where ν is n-dimensional Lebesgue measure, and G the group of translations, rotations, and reflections on R^n. For $\Lambda \in B(S)$, define $X(\Lambda)$ to be multiplication by the characteristic function χ_Λ, and let M be the states on L defined in the usual way. Then S, G, L, M satisfy Axioms I and II. For $g \in G$, let U_g be the operator on $L_2(S, \nu)$ defined by $U_g\psi(s) = \psi(gs)$. Then U_g is a unitary operator; now suppose the resolution of the identity for U_g is x_g. With the map $g \to x_g$ our system satisfies Axioms I, II, III. Now let m be a state corresponding to a unit vector ψ_m; i.e., $m(P) = \langle \psi_m, P\psi_m \rangle$ for any $P \in L$. Then by the proof of Theorem II.3, $\hat{\psi}_m(s) = |\psi_m(s)|$. Thus in order for m to be canonical we must have

$$\int \psi_m(s)\overline{\psi_m(gs)}\nu(ds) = \int |\psi_m(s)|\,|\psi_m(gs)|\,\nu(ds)$$

for all $g \in G$. Hence if $\psi_m(s) \geq 0$, $s \in S$, then ψ_m is canonical. We thus see that there is an abundance of canonical states. In fact, $L_2(S, \nu)$ is the linear hull of vectors corresponding to canonical states.

The next theorem shows that the statistical properties of a quantum-mechanical system in a canonical state satisfying Axioms I, II, III are described by operators in a Hilbert space.

Theorem 4.2: Let (L, M) be a proposition system, and let $f_\alpha(X)$, p_α be conjugate coordinate and momentum observables. Then there exists a Hilbert space H, a map $m \to \hat{m}$ from M into H which preserves convex sets in the sense that $[(\Sigma\lambda_i m_i)\hat{m}]^2 = \Sigma\lambda_i(\hat{m}_i)^2$, and self-adjoint operators S_α, T_α such that $m[f_\alpha(X)] = \langle \hat{m}, S_\alpha\hat{m} \rangle$; and if m is canonical, $m(p_\alpha) = \langle \hat{m}, T_\alpha\hat{m} \rangle$ when these expressions exist.

Proof: Let $H = L_2(S, \nu)$ and let $m \to \hat{m}$ be the map which exists according to Theorem II.3. If $m = \Sigma\lambda_i m_i$, then

$$\int_\Lambda (\hat{m})^2\,d\nu = m[X(\Lambda)] = \Sigma\lambda_i m_i[X(\Lambda)],$$
$$= \Sigma\lambda_i \int_\Lambda (\hat{m}_i)^2\,d\nu.$$

Applying the monotone convergence theorem, this last expression equals

$$\int_\Lambda \Sigma\lambda_i(\hat{m}_i)^2\,d\nu,$$

and hence $(\hat{m})^2 = \Sigma\lambda_i(\hat{m}_i)^2$. By Theorem II.3, condition (iii), there are self-adjoint operators S_α such that $m[f_\alpha(X)] = \langle \hat{m}, S_\alpha\hat{m} \rangle$ when these expressions exist. Defining the unitary operators U_g by $U_g\phi(s) = \phi(gs)$ for $g \in G$, $\phi \in H$, we see that $g \to U_g$ is a continuous unitary representation of G. Now let g_λ be the motion associated with p_α and let x_λ be the corresponding unitary derivables. If m is canonical, we have $m(x_\lambda) = \int \hat{m}U_{g_\lambda}\hat{m}\,d\nu$. Letting T_α be the self-adjoint operator which is the infinitesimal generator of the continuous one-parameter group U_{g_λ}, we have

$$m(e^{i\lambda p_\alpha}) = \int \hat{m}e^{i\lambda T_\alpha}\hat{m}\,d\nu.$$

Applying Corollary I.5, we have

$$m(p_\alpha) = \frac{d}{d\lambda}\,m(e^{i\lambda p_\alpha})\big|_0 = \int \hat{m}\,\frac{d}{d\lambda}\,e^{i\lambda T_\alpha}\big|_0\,\hat{m}\,d\nu,$$
$$= \langle \hat{m}, T_\alpha\hat{m} \rangle$$

when $m(p_\alpha)$ exists.

Now the unitary operator U_g is the "natural" operator for representing x_g on the Hilbert space. We thus see that the canonical states are all of the states for which this representation is possible, in the sense that m is canonical if and only if $m(x_g) = \int \hat{m}U_g\hat{m}\,d\nu$ for all $g \in G$.

5. SOME PROPERTIES OF σ-HOMOMORPHISMS

In this section we consider some properties of σ-homomorphisms which might be useful for physical applications. The first theorem gives a way of calculating the spectrum of a function of a σ-homomorphism. If the position σ-homomorphism is known and the coordinate function f is given, this theorem gives us the spectrum of the coordinate observable $f(X)$.

Theorem 5.1: Let $X: B(\Omega_1) \to L$ be a σ-homomorphism, f, g Borel functions from Ω_1 into Ω_2, and

let $\Lambda \in B(\Omega_1)$ satisfy $X(\Lambda) = 1$. (i) If $f(\lambda) = g(\lambda)$ for $\lambda \in \Lambda$, then $f(X) = g(X)$. (ii) $\sigma[f(X)] \subset Clf(\Lambda)$. (iii) If f is continuous, $\sigma[f(X)] = Clf[\sigma(X)]$. In (iv) general, $\sigma[f(X)] = \bigcap \{Clf(\Lambda) : X(\Lambda) = 1\}$.

Proof: (a) $f(X)(\Gamma) = X[f^{-1}(\Gamma)] = X[\Lambda \cap f^{-1}(\Gamma)]$ $= X[\Lambda \cap g^{-1}(\Gamma)] = g(X)(\Gamma)$. (b) We first show $\sigma[f(X)] \subset Clf(\Omega_1)$. Suppose $\lambda \in \sigma[f(X)]$. Then if $\lambda \in U$ and U is open, we have $0 \neq f(X)(U) = X[f^{-1}(U)]$. Therefore, $f^{-1}(U) \neq \phi$ and $U \cap f(\Omega_1) \neq \phi$. Thus $\lambda \in Clf(\Omega_1)$ and $\sigma(f(X)) \subset Clf(\Omega_1)$. Now let $\lambda_0 \in f(\Lambda)$ and define $g(\lambda) = f(\lambda)$, if $\lambda \in \Lambda$, and $g(\lambda) = \lambda_0$ otherwise. By (a) we have $\sigma[f(X)] = \sigma[g(X)] \subset Clg(\Omega_1) \subset Clf(\Lambda)$. (c) Since $f(X)\{\sigma[f(X)]'\} = 0$, then $X\{f^{-1}[\sigma(f(X))']\} = 0$. Since f is continuous, the set $f^{-1}\{\sigma[f(X)]'\}$ is open and hence in $\sigma(X)'$. Therefore $f^{-1}\{\sigma[f(X)]'\} \cap \sigma(X) = \phi$ and $\sigma[f(X)]' \cap f[\sigma(x)] = \phi$. Hence $f[\sigma(X)] \subset \sigma[f(X)]$ and $Clf[\sigma(X)] \subset \sigma[f(X)]$, since the latter is closed. The inclusion in the other direction follows from (b). (d) Suppose

$$\lambda \in \bigcap \{Clf(\Lambda) : X(\Lambda) = 1\}.$$

Suppose $\lambda \in U$ where U is open and $f(X)(U) = 0$. Then $1 = f[X(U')] = X[f^{-1}(U')]$. Therefore

$$\lambda \in Clf[f^{-1}(U')] \subset U',$$

since the latter is closed. But this is a contradiction. Therefore $\lambda \in \sigma[f(X)]$. The inclusion in the other direction follows from (b).

The last theorem was a generalization of Theorem 4.2 and the corollary to Theorem 4.3 (Ref. 11). The next theorem is an existence theorem which states that we can find simultaneous σ-homomorphisms whose spectra are any closed sets we would like.

Theorem 4.2: Let K_α be a nonempty closed subset of the topological space Ω_α, $\alpha \in A$. If L has an infinite number of disjoint propositions, then there exist simultaneous σ-homomorphisms $X_\alpha : B(\Omega_\alpha) \to L$, $\alpha \in A$, such that $K_\alpha = \sigma(X_\alpha)$.

Proof: Fix $\beta \in A$. Since Ω_β has a countable basis for its open sets, K_β has a countable dense subset $\{p_i\}$. Let $\{a_i\}$ be a countable set of mutually disjoint non-absurd propositions for which $\Sigma a_i = 1$. Define the σ-homomorphism $X_\beta : B(\Omega_\beta) \to L$ (it is easily checked that this is a σ-homomorphism) as follows:

$$X_\beta(\Lambda) = \Sigma\{a_i : p_i \in \Lambda\}, \quad \Lambda \in B(\Omega_\beta).$$

Now let $p_i \in U$, where U is open in Ω_β. Then $X_\beta(U) \neq 0$ and hence $\{p_i\} \subset \sigma(X_\beta)$; and since $\sigma(X_\beta)$ is closed, $K_\beta = Cl\{p_i\} \subset \sigma(X_\beta)$. Also $X_\beta(K_\beta') = 0$ and hence $K_\beta' \subset \sigma(X_\beta)'$. Thus $\sigma(X_\beta) \subset K_\beta$ and $K_\beta = \sigma(X_\beta)$. Now for each $\alpha \in A$ define a σ-homomorphism

$X_\alpha : B(\Omega_\alpha) \to L$ with range in $\{a_i\}$ as above. These X_α are all simultaneous and $K_\alpha = \sigma(X_\alpha)$.

Let $X : B(\Omega) \to L$ be a σ-homomorphism and $u : \Omega \to C$ a Borel function. The *variance* $V_m[u(X)]$ of $u(X)$ in the state m is

$$V_m[u(X)] = m[u^2(X)] - (m[u(X)])^2.$$

A state m is an *eigenstate* of X corresponding to the eigenvalue $\omega \in \Omega$ if $m[X(\Lambda)] = 1$ for every $\Lambda \in B(\Omega)$ containing ω. Of course, the eigenvalues of X are in $\sigma(X)$ and are a generalization of the point spectrum of a self-adjoint operator. Physically, an eigenvalue of an observable x corresponding to the eigenstate m is a value which x attains with certainty in the state m. Intuitively, one would expect that if m is an eigenstate for X, then the uncertainty, or variance, of X in the state m is zero. The next theorem shows this and further characterizes eigenstates.

Theorem 5.3: m is an eigenstate of X if and only if $V_m[u(X)] = 0$ for every real Borel function u.

Proof: Let m correspond to the eigenvalue ω_0, and define the Borel set $\Lambda = \{\omega : u(\omega) = u(\omega_0)\}$. Since $\omega_0 \in \Lambda$, we have

$$m[u(X)] = \int_\Lambda u(\omega) m[X(d\omega)] = u(\omega_0).$$

Similarly,

$$m[u^2(X)] = \int_\Lambda u^2(\omega) m[X(du)] = u^2(\omega_0)$$

and

$$V[u(X)] = 0.$$

Conversely, suppose $V_m[u(X)] = 0$ for every real Borel function. If $\Lambda \in B(\Omega)$ and χ_Λ denotes the characteristic function of Λ, then $V_m[\chi_\Lambda(X)] = 0$. Therefore,

$$m[X(\Lambda)] = m[\chi_\Lambda^2(X)] = m[\chi_\Lambda(X)]^2 = m[X(\Lambda)]^2$$

and $m[X(\Lambda)] = 0$ or 1 for every $\Lambda \in B(\Omega)$. Let $\{\Gamma_i : i = 1, 2, \cdots\}$ be the basis sets for which $m[X(\Gamma_i)] = 1$. Now $m[X(\Gamma_1 \cap \Gamma_2)] = 1$, since otherwise

$$m[X(\Gamma_1 \cup \Gamma_2)]$$
$$= m\{X[(\Gamma_1 - \Gamma_1 \cap \Gamma_2) \cup (\Gamma_2 - \Gamma_1 \cap \Gamma_2)]\},$$
$$= m[X(\Gamma_1)] + m[X(\Gamma_2)] = 2,$$

which is impossible. By induction, $m[X(\bigcap_{i=1}^n \Gamma_i)] = 1$ and

$$m\left[X\left(\bigcap_{i=1}^\infty \Gamma_i\right)\right] = \lim_{n \to \infty} m\left[X\left(\bigcap_{i=1}^n \Gamma_i\right)\right] = 1.$$

Let $\omega_0 \in \bigcap_{i=1}^\infty \Gamma_i$. Since Ω is T_1, ω_0 is closed and hence

$\omega_0 \in B(\Omega)$. Suppose $m[X(\omega_0)] = 0$. Then

$$m[X(\Gamma_1 - \{\omega_0\})] = 1.$$

Since $\Gamma_1 - \{\omega_0\}$ is open, it is the countable union of basis sets Λ_i, $i = 1, 2, \cdots$. But at least one of these basis sets must satisfy $m[X(\Lambda_j)] = 1$. But this is a contradiction, since ω_0 is in every such basis set. Therefore, $m[X(\omega_0)] = 1$, and m is an eigenstate.

ACKNOWLEDGMENT

APPENDIX A. ONE-PARAMETER GROUPS OF DERIVABLES

The *identity derivable I* is the unique derivable which satisfies $\sigma(I) = \{1\}$. A derivable x is bounded if $\sigma(x)$ is a bounded set. If x is bounded, we define the *norm* of x to be $|x| = \sup\{|\lambda|: \lambda \in \sigma(x)\}$. It is shown in Ref. 11 that $|x|$ is indeed a norm. A *one-parameter group of derivables* is a family $\{x_t: t \in (-\infty, \infty)\}$ of bounded simultaneous derivables which satisfies $x_0 = I$ and $x_{s+t} = x_s x_t$ for all $s, t \in (-\infty, \infty)$. A one-parameter group of derivables is *continuous* if $m(x_t)$ is a continuous function of t for every state m.

Lemma I.1: If x_t is a one-parameter group of derivables, then

$$\lambda_0 = \inf_{t>0} \log |x_t|/t = \lim_{t \to \infty} \log |x_t|/t < \infty.$$

Proof: Since $\log |x_{s+t}| = \log |x_s x_t| \leq \log(|x_s| |x_t|) = \log |x_s| + \log |x_t|$, the function $t \to \log |x_t|$ is subadditive on $[0, \infty)$, and the result is a well-known property of subadditive functions.

Theorem I.2: x_t is a continuous one-parameter group of derivables if and only if there exists a derivable y such that e^{ty} is bounded and $x_t = e^{ty}$, $t \in (-\infty, \infty)$.

Proof: We first prove necessity. Applying Theorems 3.3 and 4.1 (Ref. 11), there exists a measurable space (Ω, A), a σ-homomorphism $h: A \to L$, and A measurable functions f_t such that $|x_t| = \sup\{|f_t(\omega)|: \omega \in \Omega\}$, $x_t(E) = h[f_t^{-1}(E)]$ for all $E \in B(C)$, and $f_{s+t}(\omega) = f_s(\omega)f_t(\omega)$ for $\omega \in N$ where $h(N') = 0$. Since for fixed $\omega \in N$, $f_{s+t}(\omega) = f_s(\omega)f_t(\omega)$ for s, t rational, we have $f_t(\omega) = e^{g(\omega)t}$, t rational. Now g is an A-measurable function, and hence $E \to y(E) = h[g^{-1}(E)]$

is a derivable. Applying Theorem 3.3 (Ref. 11), we get $x_t(E) = h[f_t^{-1}(E)] = h(e^{ty})^{-1}(E)] = e^{ty}(E)$, and hence $x_t = e^{ty}$ for t rational. Now

$$\text{Re } g(\omega) = \text{Re } \log f_t(\omega)/t = \log |f_t(\omega)|/t$$
$$\leq \log |\sup\{f_t(\omega): \omega \in \Omega\}|/t = \log |x_t|/t,$$

t positive rational. Hence

$$\text{Re } g(\omega) \leq \lim\{|x_t|/t: t \text{ rational}, t \to \infty\} = \lambda_0 \leq \infty.$$

Letting

$$S = \{\lambda: \text{Re } \lambda \leq \lambda_0\},$$

we have $m[y(S)] = m[h(g^{-1}(S))] = 1$. For a fixed $t \in (0, \infty)$, $\lambda \to e^{t\lambda}$ is a bounded function for $\lambda \in S$. Now let t_i be a sequence of positive rationals converging to t, and let m be a state. Then, by continuity and the dominated convergence theorem,

$$m(x_t) = \lim m(x_{t_i}) = \lim m(e^{t_i y}),$$

$$= \lim \int e^{t_i \lambda} m[y(d\lambda)] = \int \lim e^{t_i \lambda} m[y(d\lambda)] = m(e^{ty}).$$

Since x_t and y are simultaneous, $x_t = e^{ty}$. A similar argument is used for negative t. For sufficiency, suppose x_t is bounded, $x_t = e^{ty}$, $t \in (-\infty, \infty)$. Certainly, $x_0 = I$ and

$$x_{s+t}(E) = e^{(s+t)y}(E) = y\{\lambda: e^{(s+t)\lambda} \in E\},$$

$$= y\{\lambda: e^{s\lambda}e^{t\lambda} \in E\} = e^{sy}e^{ty}(E) = x_s x_t(E)$$

for all $E \in B(C)$. Hence $x_{s+t} = x_s x_t$. Now let t be positive and let t_i be a sequence of positive numbers converging to t. Letting $t_0 > t$, for large enough i, we have $t_i \leq t_0$. Since x_{t_0} is bounded, applying Theorem 5.4 (Ref. 11), there is a constant K and a set $E \in B(X)$ such that $y(E) = 1$ and $|\exp(t_0 E)| \leq K$. Hence, for i large $|\exp(t_i E)| < K$. Applying the dominated convergence theorem,

$$\lim m(x_{t_i}) = \lim(e^{t_i y}) = \lim \int_E e^{t_i \lambda} m[y(d\lambda)],$$

$$= \int_E \lim e^{t_i \lambda} m[y(d\lambda)] = m(e^{ty}) = m(x_t).$$

A similar argument holds for negative t. Hence x_t is a continuous one-parameter group.

The derivable y in Theorem I.2 is called an *infinitesimal generator* of the one-parameter group x_t. We have thus shown that a one-parameter group of derivables has an infinitesimal generator if and only if it is continuous. It is not known whether an infinitesimal generator of a continuous one-parameter group

of derivables is unique, although we now derive a kind of uniqueness result.

Theorem I.3: Let x_t be a continuous one-parameter group with an infinitesimal generator y. Then $m(y)$ exists if and only if $\lim_{t \to 0} m[(e^{ty} - I)/t]$ exists; when they exist, $m(y) = \lim_{t \to 0} m[(e^{ty} - I)/t]$.

Proof: Suppose $m(y) = \int \lambda m[y(d\lambda)]$ exists. Since $\lim (e^{t\lambda} - 1)t = \lambda$ pointwise, applying Fatou's lemma our result follows. The converse follows in a similar fashion.

Corollary I.4: Let y_1 and y_2 be infinitesimal generators of a continuous one-parameter group. Then $m(y_1)$ exists if and only if $m(y_2)$ exists; when they exist, $m(y_1) = m(y_2)$.

Corollary I.5: Let x_t be a continuous one-parameter group with an infinitesimal generator y. Then $(d/dt)[m(x_t)]|_0$ exists if and only if $m(y)$ exists. When they exist, $m(y) = (d/dt)[m(x_t)]|_0$.

Proof:

$$(d/dt)[m(x_t)]|_0 = \lim_{t \to 0} [m(x_t) - m(x_0)]/t,$$
$$= \lim [m(x_t) - 1]/t,$$
$$= \lim m[(x_t - 1)/t],$$
$$= \lim m[(e^{ty} - I)/t],$$
$$= m(y).$$

It is easily seen that the derivables of a continuous one-parameter group are unitary if and only if any infinitesimal generator is of the form iy where y is an observable.

Corollary I.6: Let $\{x_\lambda : \lambda \in (-\infty, \infty)\}$ be a continuous one-parameter group of unitary derivables with infinitesimal generator p. An σ-homomorphism $X \leftrightarrow x_\lambda$ for all $\lambda \in (-\infty, \infty)$ if and only if $X \leftrightarrow p$.

Proof: The proof follows, using the methods of Theorem I.2.

APPENDIX B. ABSOLUTELY CONTINUOUS σ-HOMOMORPHISMS

Let G be a locally compact group and let H be a closed subgroup of G. We form the space of left cosets G/H, with the topology induced by the canonical mapping h from G into G/H. The Borel subsets $B(G/H)$ of G/H are generated by the open sets already defined on G/H. We denote the Haar measure on G by μ.

Lemma II.1: If ν is a σ-finite Borel measure on G/H and if $\mu[h^{-1}(\Lambda)] = 0$, $\Lambda \in B(G/H)$, then $\nu(g\Lambda) = 0$ for μ almost every $g \in G$.

Proof: It suffices to assume that ν is finite. Suppose $\mu[h^{-1}(\Lambda)] = 0$. Let χ_Λ be the characteristic function of Λ and apply Fubini's theorem to the Borel function $\chi_\Lambda(g^{-1}x)$ on $G \times (G/H)$:

$$0 = \int \mu[h^{-1}(x)(h^{-1}(\Lambda))^{-1}]\nu(dx),$$
$$= \int \left[\int \chi_\Lambda(g^{-1}x)\mu(dg) \right]\nu(dx),$$
$$= \int \left[\int \chi_\Lambda(g^{-1}x)\nu(dx) \right]\mu(dg),$$
$$= \int \nu(g\Lambda)\mu(dg).$$

Since $g \to \nu(g\Lambda) \geq 0$, we have $\nu(g\Lambda) = 0$ μ almost everywhere.

Let $X : B(G/H) \to L$ be a σ-homomorphism. Then we say that X is *absolutely continuous* with respect to a measure ν on $B(G/H)$ if $\nu(E) = 0$ implies $X(E) = 0$.

Lemma II.2: Let G be a locally compact group, H a closed subgroup of G, and $X : B(G/H) \to L$ a σ-homomorphism such that $g \to m[X(g\Lambda)]$ is continuous for all $m \in M$, $\Lambda \in B(G/H)$. Then X is absolutely continuous with respect to any quasi-invariant measure on G/H.

Proof: Let $\Lambda \in B(G/H)$ satisfy $\mu[h^{-1}(\Lambda)] = 0$. Now if $m \in M$, then $m[X(\cdot)]$ is a finite measure on G/H. Applying Lemma II.1, $m[X(g\Lambda)] = 0$ for μ almost every $g \in G$. Let U be a neighborhood of the identity $e \in G$. Since open sets have positive Haar measure, there is a point $g_U \in U$ such that $m[X(g_U\Lambda)] = 0$. In this way we obtain a generalized sequence g_α, converging to e, for which $m[X(g_\alpha\Lambda)] = 0$. Applying the continuity, we have

$$0 = m[X(g_\alpha\Lambda)] \to m[X(e\Lambda)] = m[X(\Lambda)]$$

and hence $m[X(\Lambda)] = 0$ for every $m \in M$, and thus $X(\Lambda) = 0$. Now Lemma 1.3 (Ref. 12) states that if ν is quasi-invariant on G/H, then $\nu(\Lambda) = 0$ if and only if $\mu[h^{-1}(\Lambda)] = 0$. Hence X is absolutely continuous with respect to ν.

Theorem II.3: Let X be the σ-homomorphism of Lemma II.2, ν a quasi-invariant measure on G/H, and $L_2(\nu)$ the Hilbert space of complex-valued square-integrable functions on G/H. Then there is a map $m \to \hat{m}$ of M into $L_2(\nu)$ and a projection-valued

measure $\hat{X}(\cdot)$ from $B(G/H)$ into the orthogonal projections on $L_2(\nu)$ which satisfies (a) $m[X(\Lambda)] = \langle \hat{m}, \hat{X}(\Lambda)\hat{m} \rangle$ for all $\Lambda \in B(G/H)$. If u is a real Borel function on Ω, there is a self-adjoint operator U on $L_2(\nu)$ with resolution of identity $U(\cdot)$ such that (b) $m[u(X)(E)] = \langle \hat{m}, U(E)\hat{m} \rangle$ for all $E \in B(R)$; (c) when it exists, $m[u(X)] = \langle \hat{m}, U\hat{m} \rangle$.

Proof: For $m \in M$, $m[X(\cdot)]$ is an absolutely continuous measure with respect to ν. Applying the Radon–Nikodym theorem, there is a nonnegative Borel function f on G/H such that $m[X(\Lambda)] = \int_\Lambda f \, d\nu$ for all $\Lambda \in B(G/H)$. Define the function $\hat{m}(\lambda) = +[f(\lambda)]^{\frac{1}{2}}$ to get $m[X(\Lambda)] = \int_\Lambda (\hat{m})^2 \, d\nu$ for all $\Lambda \in B(G/H)$. We now define $\hat{X}(\Lambda): L_2(\nu) \to L_2(\nu)$ by $\hat{X}(\Lambda)\psi(\lambda) = \chi_\Lambda(\lambda)\psi(\lambda)$, where χ_Λ is the characteristic function of Λ. It is easy to check that $\hat{X}(\Lambda)$ is an orthogonal projection on $L_2(\nu)$ and that $\Lambda \to \hat{X}(\Lambda)$ is a projection-valued measure. We then have

$$m[X(\Lambda)] = \int \hat{X}(\Lambda)(\hat{m})^2 \, d\nu = \langle \hat{m}, \hat{X}(\Lambda)\hat{m} \rangle.$$

(b) follows directly from (a). To prove (c) we have

$$m[u(X)] = \int u(\lambda)m[X(d\lambda)] = \int u(\lambda)\int_{d\lambda} (\hat{m})^2 \, d\nu,$$
$$= \int u(\lambda)\int \hat{X}(d\lambda)(\hat{m})^2 \, d\nu = \int u(\lambda)\langle \hat{m}, \hat{X}(d\lambda)\hat{m} \rangle,$$
$$= \left\langle \hat{m}, \int u(\lambda)\hat{X}(d\lambda)\hat{m} \right\rangle = \langle \hat{m}, U\hat{m} \rangle,$$

where U is the self-adjoint operator whose resolution of the identity is $\hat{X}[u^{-1}(\cdot)]$.

SUMMARY

An axiomatic model for quantum mechanics is formulated using physically significant axioms. The model is a slight strengthening of Mackey's first six axioms, together with two axioms which ensure the existence of coordinate and momentum observables. Coordinate and momentum observables are defined. It is then shown that as far as the statistical properties of these observables in certain states are concerned, the abstract axiomatic model may be represented by the usual Hilbert space formulation.

Algebraic Calculation of Nonrelativistic Coulomb Phase Shifts

DANIEL ZWANZIGER*
Courant Institute of Mathematical Sciences, New York University, New York, New York

It is observed that the in and out states of a particle of energy $k^2/2m$, entering or leaving a Coulomb field along the $+z$ direction, are eigenstates of L_3 with eigenvalue 0, and of A_3, the third component of the Runge–Lenz vector, with eigenvalues $\pm\alpha + ik/m$, respectively. From this characterization and the commutation relations of the symmetry group, the phase shifts are easily obtained algebraically.

WE show that the algebra of the symmetry group of the Kepler problem not only determines the energy levels completely, but also the phase shifts.

In 1926 Pauli[1] showed that the bound-state spectrum of the quantum-mechanical nonrelativistic Kepler problem could be obtained from the commutation relations of the angular momentum vector and the Runge–Lenz vector. Later on it was recognized[2] that this algebra generates the O(4) symmetry group for bound-state energy levels and O(1, 3) for continuum levels, and it was realized[3] that the separability of the wave equation in spherical and parabolic coordinates is due to this symmetry. A recent review[4] gives a detailed exposition of the group-theoretical approach.

Let the Hamiltonian be

$$H = \mathbf{p}^2/2m - \alpha/r \tag{1}$$

with the canonical commutation relations

$$[x_i, x_j] = 0, \quad [p_i, p_j] = 0, \quad [p_i, x_j] = -i\delta_{ij}. \tag{2}$$

The angular momentum vector $\mathbf{L}$ and the Runge–Lenz vector $\mathbf{A}$ are defined by

$$\mathbf{L} = \mathbf{x} \times \mathbf{p}, \quad \mathbf{A} = (1/2m)(\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - \alpha\hat{x}, \tag{3}$$

* Work performed under a Ford Foundation grant.
[1] W. Pauli, Z. Physik **36**, 336 (1926).
[2] V. Fock, Z. Physik **98**, 145 (1935).
[3] V. Bargmann, Z. Physik **99**, 576 (1936).

[4] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).

measure $\hat{X}(\cdot)$ from $B(G/H)$ into the orthogonal projections on $L_2(\nu)$ which satisfies (a) $m[X(\Lambda)] = \langle \hat{m}, \hat{X}(\Lambda)\hat{m} \rangle$ for all $\Lambda \in B(G/H)$. If u is a real Borel function on Ω, there is a self-adjoint operator U on $L_2(\nu)$ with resolution of identity $U(\cdot)$ such that (b) $m[u(X)(E)] = \langle \hat{m}, U(E)\hat{m} \rangle$ for all $E \in B(R)$; (c) when it exists, $m[u(X)] = \langle \hat{m}, U\hat{m} \rangle$.

Proof: For $m \in M$, $m[X(\cdot)]$ is an absolutely continuous measure with respect to ν. Applying the Radon–Nikodym theorem, there is a nonnegative Borel function f on G/H such that $m[X(\Lambda)] = \int_\Lambda f \, d\nu$ for all $\Lambda \in B(G/H)$. Define the function $\hat{m}(\lambda) = +[f(\lambda)]^{\frac{1}{2}}$ to get $m[X(\Lambda)] = \int_\Lambda (\hat{m})^2 \, d\nu$ for all $\Lambda \in B(G/H)$. We now define $\hat{X}(\Lambda): L_2(\nu) \to L_2(\nu)$ by $\hat{X}(\Lambda)\psi(\lambda) = \chi_\Lambda(\lambda)\psi(\lambda)$, where χ_Λ is the characteristic function of Λ. It is easy to check that $\hat{X}(\Lambda)$ is an orthogonal projection on $L_2(\nu)$ and that $\Lambda \to \hat{X}(\Lambda)$ is a projection-valued measure. We then have

$$m[X(\Lambda)] = \int \hat{X}(\Lambda)(\hat{m})^2 \, d\nu = \langle \hat{m}, \hat{X}(\Lambda)\hat{m} \rangle.$$

(b) follows directly from (a). To prove (c) we have

$$m[u(X)] = \int u(\lambda) m[X(d\lambda)] = \int u(\lambda) \int_{d\lambda} (\hat{m})^2 \, d\nu,$$
$$= \int u(\lambda) \int \hat{X}(d\lambda)(\hat{m})^2 \, d\nu = \int u(\lambda) \langle \hat{m}, \hat{X}(d\lambda)\hat{m} \rangle,$$
$$= \left\langle \hat{m}, \int u(\lambda) \hat{X}(d\lambda)\hat{m} \right\rangle = \langle \hat{m}, U\hat{m} \rangle,$$

where U is the self-adjoint operator whose resolution of the identity is $\hat{X}[u^{-1}(\cdot)]$.

SUMMARY

An axiomatic model for quantum mechanics is formulated using physically significant axioms. The model is a slight strengthening of Mackey's first six axioms, together with two axioms which ensure the existence of coordinate and momentum observables. Coordinate and momentum observables are defined. It is then shown that as far as the statistical properties of these observables in certain states are concerned, the abstract axiomatic model may be represented by the usual Hilbert space formulation.

Algebraic Calculation of Nonrelativistic Coulomb Phase Shifts

DANIEL ZWANZIGER*
Courant Institute of Mathematical Sciences, New York University, New York, New York

It is observed that the in and out states of a particle of energy $k^2/2m$, entering or leaving a Coulomb field along the $+z$ direction, are eigenstates of L_3 with eigenvalue 0, and of A_3, the third component of the Runge–Lenz vector, with eigenvalues $\pm\alpha + ik/m$, respectively. From this characterization and the commutation relations of the symmetry group, the phase shifts are easily obtained algebraically.

WE show that the algebra of the symmetry group of the Kepler problem not only determines the energy levels completely, but also the phase shifts.

In 1926 Pauli[1] showed that the bound-state spectrum of the quantum-mechanical nonrelativistic Kepler problem could be obtained from the commutation relations of the angular momentum vector and the Runge–Lenz vector. Later on it was recognized[2] that this algebra generates the O(4) symmetry group for bound-state energy levels and O(1, 3) for continuum levels, and it was realized[3] that the

separability of the wave equation in spherical and parabolic coordinates is due to this symmetry. A recent review[4] gives a detailed exposition of the group-theoretical approach.

Let the Hamiltonian be

$$H = \mathbf{p}^2/2m - \alpha/r \tag{1}$$

with the canonical commutation relations

$$[x_i, x_j] = 0, \quad [p_i, p_j] = 0, \quad [p_i, x_j] = -i\delta_{ij}. \tag{2}$$

The angular momentum vector $\mathbf{L}$ and the Runge–Lenz vector $\mathbf{A}$ are defined by

$$\mathbf{L} = \mathbf{x} \times \mathbf{p}, \quad \mathbf{A} = (1/2m)(\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - \alpha\hat{x}, \tag{3}$$

[1] W. Pauli, Z. Physik **36**, 336 (1926).
[2] V. Fock, Z. Physik **98**, 145 (1935).
[3] V. Bargmann, Z. Physik **99**, 576 (1936).

[4] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).

and are found, from (1) and (2), to satisfy

$$[\mathbf{L}, H] = 0, \tag{4a}$$

$$[\mathbf{A}, H] = 0, \tag{4b}$$

$$[L_i, L_j] = i\epsilon_{ijk}L_k, \tag{5a}$$

$$[L_i, A_j] = i\epsilon_{ijk}A_k, \tag{5b}$$

$$[A_i, A_j] = -i\epsilon_{ijk}L_k(2H/m), \tag{5c}$$

$$\mathbf{L} \cdot \mathbf{A} = 0, \tag{6a}$$

$$\mathbf{A}^2 = (\mathbf{L}^2 + 1)(2H/m) + \alpha^2, \tag{6b}$$

as is well known.

Equations (4) show that $\mathbf{L}$ and $\mathbf{A}$ commute with H, and in the remaining relations (5) and (6) no other variables appear. Consequently, we may restrict ourselves to the subspace where H has a definite value and write

$$H = k^2/2m, \quad k > 0, \tag{7}$$

since we are concerned only with the continuum. It is convenient to renormalize $\mathbf{A}$ according to

$$\mathbf{K} \equiv (m/k)\mathbf{A}, \tag{8}$$

so Eqs. (5) and (6) become

$$[L_i, L_j] = i\epsilon_{ijk}L_k, \tag{9a}$$

$$[L_i, K_j] = i\epsilon_{ijk}K_k, \tag{9b}$$

$$[K_i, K_j] = -i\epsilon_{ijk}L_k, \tag{9c}$$

$$\mathbf{K} \cdot \mathbf{L} = 0, \tag{10a}$$

$$\mathbf{L}^2 - \mathbf{K}^2 = -\alpha^2 m^2/k^2 - 1, \tag{10b}$$

or

$$k^2/(m^2\alpha^2) = -(\mathbf{L}^2 - \mathbf{K}^2 + 1)^{-1}. \tag{11}$$

The Lie algebra defined by Eqs. (9) is that of $O(1, 3)$, or the Lorentz group. The representations of this algebra may be found in Naimark.[5] Their derivation may be understood as follows. Choose a basis where $\mathbf{L}^2$ and L_3 are diagonal, say $|l, m\rangle$. Then by virtue of Eq. (9b) and the Wigner–Eckart theorem, the matrix elements of $\mathbf{K}$ are expressed in terms of the reduced matrix elements $K_{l+1, l}$, $K_{l, l}$, $K_{l-1, l}$. Hermiticity of $\mathbf{K}$ says that $K_{l, l}$ is real and expresses $K_{l-1, l}$ in terms of $K^*_{l, l-1}$, the phase of which is fixed by convention.

[5] M. A. Naimark, *Linear Representations of the Lorentz Group* (American Mathematical Society, Providence, Rhode Island, 1957), Chap. 3, Sec. 2, No. 3, Eqs. (51)–(55). Irreducible representations of the Lorentz group are conventionally labeled by (l_0, c), where l_0 is a nonnegative integer or half-integer and c is an arbitrary complex number. The Casimir invariants are $\mathbf{K} \cdot \mathbf{L} = il_0c$ and $\mathbf{L}^2 - \mathbf{K}^2 = l_0^2 + c^2 - 1$. From Eq. (10a) we have $c = 0$ or $l_0 = 0$. If $c = 0$, Eq. (11) gives $k^2(m\alpha)^{-2} = -l_0^{-2} < 0$, which contradicts the assumption of positive energy. Hence $l_0 = 0$ and $c^2 = -\alpha^2m^2/k^2 < 0$, so c is pure imaginary. This defines a unitary representation in the principal series.

Equation (10a) determines $K_{l, l}$, and Eq. (9c) gives a recursion relation for $K_{l+1, l}$ which terminates at some minimum value l_0. The only result which is required is

$$K_3 |l, 0\rangle = il\left[\frac{l^2 + \alpha'^2}{4l^2 - 1}\right]^{\frac{1}{2}} |l - 1, 0\rangle$$
$$- i(l + 1)\left[\frac{(l + 1)^2 + \alpha'^2}{4(l + 1)^2 - 1}\right]^{\frac{1}{2}} |l + 1, 0\rangle, \tag{12}$$

where $\alpha' = \alpha m/k = \alpha/v$.

Let us now consider the characterization of the in (or out) scattering state corresponding to a plane wave entering (or leaving), traveling in the $+z$ direction. Because L_3 and K_3 are commuting constants of the motion, whatever value they have for appropriate wave packets before (or after) scattering, they will have these values at all times and they are suitable labels for the in (or out) states. If $L_3 = -i(\mathbf{x} \times \mathbf{\nabla})_3$ is applied to a free plane wave $\exp(ikz)$, one finds $L_3 \exp(ikz) = 0$, so

$$L_3 |k\hat{z}^{\text{in}}_{\text{out}}\rangle = 0. \tag{13a}$$

Similarly for the first term of $\mathbf{A}$, Eq. (3), one finds,[6] with $\mathbf{p} = -i\mathbf{\nabla}$,

$$(1/2m)(\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p})_3 \exp(ikz) = (ik/m)\exp(ikz).$$

Now consider a wave packet traveling in the $+z$ direction toward (or away from) the origin, and which is asymptotically far away. Then $\hat{x}_3$, the component of the direction $\hat{x}$ along the z axis, will have the value -1 (or $+1$). Hence, since

$$A_3 = (\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p})_3/2m - \alpha\hat{x}_3,$$

we have

$$A_3 |k\hat{z}^{\text{in}}_{\text{out}}\rangle = (\pm\alpha + ik/m) |k\hat{z}^{\text{in}}_{\text{out}}\rangle,$$
$$K_3 |k\hat{z}^{\text{in}}_{\text{out}}\rangle = (\pm\alpha m/k + i) |k\hat{z}^{\text{in}}_{\text{out}}\rangle. \tag{13b}$$

Equations (13) provide an adequate characterization of the in and out states, as we see later.

Consider the expansion of the scattering states into spherical waves. Because of Eq. (13a), only values

[6] At first sight it is surprising that the Hermitian operator on the left-hand side should have an imaginary eigenvalue. However, the plane wave is not an element of the Hilbert space, so there is no contradiction. This operator also has a complete real spectrum. These remarks apply to A_3 and K_3 as well. Use of complex eigenvalues is actually implicit in the standard treatment, in which the wavefunction that asymptotically approaches a plane wave is factored in parabolic coordinates. The separation constant is complex. Equation (13b) states that, although A_3 commutes with the Hamiltonian and is thus a constant of the motion, it yields different eigenvalues when applied to in and out states. Consequently A_3 does not commute with the S operator, as is easily verified by applying it to both sides of Eq. (18). In general, of course, one does not expect every operator that commutes with H to commute with S, because if this were true, then, for a fixed energy, S would be a constant multiple of the identity, and a monochromatic beam would not scatter at all.

$L_3 = m = 0$ contribute, and we have

$$|k\hat{z}^{\text{in}}\rangle = \sum_l (2l + 1)^{\frac{1}{2}} a_l^{\text{in}} |l, 0\rangle, \qquad (14a)$$

$$|k\hat{z}^{\text{out}}\rangle = \sum_l (2l + 1)^{\frac{1}{2}} a_l^{\text{out}} |l, 0\rangle. \qquad (14b)$$

To determine the expansion coefficients a_l we use Eqs. (12) and (13b) to find a_l^{in}; the result for a_l^{out} is obtained by the substitution $\alpha' \to -\alpha'$:

$$K_3 |k\hat{z}^{\text{in}}\rangle = i \sum_l (2l + 1)^{\frac{1}{2}}(1 - i\alpha')a_l^{\text{in}} |l, 0\rangle,$$

$$= i \sum_l (2l + 1)^{\frac{1}{2}} a_l^{\text{in}} \left\{ l \left[\frac{l^2 + \alpha'^2}{(2l + 1)(2l - 1)} \right]^{\frac{1}{2}} |l - 1, 0\rangle \right.$$

$$\left. - (l + 1) \left[\frac{(l + 1)^2 + \alpha'^2}{(2l + 1)(2l + 3)} \right]^{\frac{1}{2}} |l + 1, 0\rangle \right\}.$$

Upon equating coefficients of the independent vectors $|l, 0\rangle$, one obtains the recursion relation

$$a_{l+1}^{\text{in}} = (l + 1)^{-1}[(l + 1)^2 + \alpha'^2]^{-\frac{1}{2}}$$

$$\times [(2l + 1)(1 - i\alpha')a_l^{\text{in}} + l(l^2 + \alpha'^2)^{\frac{1}{2}} a_{l-1}^{\text{in}}].$$

For $l = 0$ this yields

$$a_1^{\text{in}} = [(1 - i\alpha')/(1 + i\alpha')]^{\frac{1}{2}} a_0^{\text{in}}.$$

The general term is easily proven by induction to be

$$a_l^{\text{in}} = [(l - i\alpha')!/(l + i\alpha')!]^{\frac{1}{2}}[(i\alpha')!/(-i\alpha')!]^{\frac{1}{2}} a_0^{\text{in}}.$$

We choose the undetermined over-all phase factor

$$a_0^{\text{in}} = [(-i\alpha')!/(i\alpha')!]^{\frac{1}{2}},$$

so that

$$a_l^{\text{in}} = [(l - i\alpha')!/(l + i\alpha')!]^{\frac{1}{2}}. \qquad (15a)$$

By the substitution $\alpha' \to -\alpha'$,

$$a_l^{\text{out}} = [(l + i\alpha')!/(l - i\alpha')!]^{\frac{1}{2}}, \qquad (15b)$$

which gives the desired expansion coefficients.

The S matrix is defined by

$$S(\mathbf{k}', \mathbf{k}) = \langle \mathbf{k}'^{\text{out}} | \mathbf{k}^{\text{in}} \rangle \qquad (16)$$

and the S operator by

$$S(\mathbf{k}', \mathbf{k}) = \langle \mathbf{k}'^{\text{out}} | S | \mathbf{k}^{\text{out}} \rangle. \qquad (17)$$

So from Eqs. (16) and (17),

$$|k^{\text{in}}\rangle = S |k^{\text{out}}\rangle. \qquad (18)$$

Upon substituting this into Eqs. (14) and recalling that S commutes with $\mathbf{L}$, so that

$$S |l, m\rangle = S_l |l, m\rangle = \exp(2i\delta_l) |l, m\rangle,$$

one finds

$$a_l^{\text{in}} = S_l a_l^{\text{out}}, \qquad (19)$$

$$S_l = \exp(2i\delta_l) = (l - i\alpha')!/(l + i\alpha')!, \quad \alpha' = \alpha m/k. \qquad (20)$$

This is the familiar form for the Coulomb phase shifts. The S_l have been determined up to an arbitrary common energy-dependent phase factor which is fixed by the phase convention between states of different energy. Equation (20) gives the usual one. The scattering amplitude $f(k, \theta)$ is obtained from

$$f(k, \theta) = (2ik)^{-1} \sum (2l + 1) \exp(2i\delta_l) P_l(\cos\theta),$$
$$\theta \neq 0, \qquad (21)$$

which yields

$$f(k, \theta) = \frac{1}{2} \frac{\alpha m}{k^2 \sin^2 \frac{\theta}{2}}$$

$$\times \exp\left[i(\alpha m/k) \ln \sin^2 \frac{\theta}{2} \right] \frac{(-i\alpha m/k)!}{(i\alpha m/k)!}, \qquad (22)$$

as may be verified by projecting with the $P_l(\cos\theta)$ and using Rodrigues's formula. This is the standard expression for the scattering amplitude. It has been obtained purely algebraically.

ACKNOWLEDGMENT

Refractive Index, Attenuation, Dielectric Constant, and Permeability for Waves in a Polarizable Medium*

DAVID J. VEZZETTI

Department of Physics, University of Illinois at Chicago Circle, Chicago, Illinois

AND

JOSEPH B. KELLER

Courant Institute of Mathematical Sciences, New York University, New York, New York

A new method is presented for calculating the refractive index, attenuation, dielectric constant, and permeability for electromagnetic waves in a medium of polarizable particles. It is similar to the method of Yvon and Kirkwood for finding the static dielectric constant. The main merit of the method is that it avoids the statistical hypotheses used in such calculations by Lorentz, Reiche, Hoek, Rosenfeld, and other authors. In addition, it permits the calculations to be continued to any degree of accuracy. We first use the method to obtain the dispersion equation as a power series in the molecular polarizability. The nth term in this series involves the distribution function of $n + 1$ particles. The terms of first and second degree are written out explicitly in terms of the two- and three-particle distribution functions. When terms of second and higher degree are omitted and the result specialized to particles with a scalar electric polarizability and zero magnetic polarizability, the dispersion equation agrees with that of Rosenfeld. When terms of second degree are retained and the static limit considered, the result reduces to that of Yvon. We next use the method to obtain the dispersion equation as a power series in the particle number density, which seems to be new. To obtain it we introduce "pure" n-particle scattering functions, which are analogous to the Ursell functions of statistical mechanics. This permits us to obtain the density expansion directly in a form simpler than is obtained by resumming the polarizability series.

1. INTRODUCTION

A FUNDAMENTAL difficulty in the theoretical calculation of the refractive index, dielectric constant, and permeability of a polarizable medium is that of determining the average electric or magnetic field acting upon a molecule of the medium. This average field is called the effective field, since it is the field which polarizes the molecule. Most authors have employed the hypothesis, first stated explicitly by Lorentz, that the effective electric field equals the average electric field plus one-third the average electric polarization. However, for the case of static fields, Yvon[1] and Kirkwood[2] introduced a systematic method of calculating the effective field which avoids this hypothesis. This method has been used by Green,[3] Brown,[4–6] and de Boer *et al.*[7] Our objective is to present a similar systematic method of calculation for time-harmonic fields. This method is similar to the one introduced by Keller[8,9] to treat waves in discrete

and continuous random media and employed by Karal and Keller.[10,11]

We first use our method to obtain the dispersion equation as a power series in the molecular polarizability. The nth term involves the $(n + 1)$-particle distribution function. The linear and quadratic terms reduce to the result of Yvon[1] in the static case, provided we specialize our result to particles with zero magnetic polarizability and scalar electric polarizability. For such particles and nonstatic fields the linear term alone reduces to the result of Rosenfeld.[12] Thus our result provides the generalization of Yvon's result to time-harmonic fields. It also justifies Rosenfeld's result and shows when it is valid and how to improve it.

We also use our method to obtain the dispersion equation as a power series in the particle number density. This virial expansion can be obtained by resumming the polarizability series. However, we obtain it directly, and in a simpler form, by introducing "pure" n-particle scattering functions. They are analogous to the Ursell functions of statistical mechanics and can probably be used in other problems. The density expansion seems to be new. de Boer *et al.*[7] previously obtained a double series in polarizability and density for the static dielectric constant by expanding the terms in Yvon's polarizability series in powers of the density.

* The research in this paper was supported by the National Science Foundation under Grant No. GP 3668 at New York University.

[1] J. Yvon, Compt. Rend. Acad. Sci. (Paris) **202**, 35 (1936).

[2] J. Kirkwood, J. Chem. Phys. **4**, 592 (1936).

[3] H. S. Green, in *Handbuch der Physik* (Springer-Verlag, Berlin, 1960), Vol. 10.

[4] W. F. Brown, J. Chem. Phys. **18**, 1193 (1950).

[5] W. F. Brown, J. Chem. Phys. **21**, 1121 (1953).

[6] W. F. Brown, in *Handbuch der Physik* (Springer-Verlag, Berlin, 1956), Vol. 17.

[7] J. de Boer, F. Van der Maesen, and C. ten Seldam, Physica **19**, 265 (1953).

[8] J. B. Keller, Proc. Symp. Appl. Math. **13**, 227 (1962).

[9] J. B. Keller, Proc. Symp. Appl. Math. **16**, 145 (1964).

[10] F. C. Karal, Jr., and J. B. Keller, J. Math. Phys. **5**, 537 (1964).

[11] J. B. Keller and F. C. Karal, Jr., J. Math. Phys. **7**, 661 (1966).

[12] L. Rosenfeld, *Theory of Electrons* (North-Holland Publishing Company, Amsterdam, 1951).

Some authors have used, instead of the Lorentz effective field hypothesis, a different one. This hypothesis states that the average field acting on a given fixed particle, when a second particle is also held fixed, is the same as if the second particle were not held fixed. This hypothesis was used by Reiche,[13] Hoek,[14] and Rosenfeld,[12] and in other applications by Foldy[15] and Twersky,[16] while a slightly different version was used by Lax.[17] We have already mentioned that this method yields results which are correct to the first order in the polarizability.

The dispersion equation has one solution for the refractive index n corresponding to a transverse electromagnetic wave, and possibly others corresponding to longitudinal waves. The complex index n determines the phase velocity and attenuation of the fields and the dielectric constant and permeability of the medium.

Our analysis is restricted to molecules without permanent moments and to fields so weak that they do not affect the distribution functions.

2. FORMULATION FOR A GIVEN CONFIGURATION

Let us consider the electric field $E(x \mid x^{(N)})$ and magnetic field $H(x \mid x^{(N)})$ produced in unbounded space by a wave $E_0(x)$, $H_0(x)$ incident upon N polarizable particles without permanent moments located at the points $x_1, \cdots, x_N$. The symbol $x^{(N)}$ represents the collection of all the particle coordinates $x_1, \cdots, x_N$. We assume that the fields and the dipole moments which they induce are time-periodic with angular frequency ω, and we omit the time factor $e^{i\omega t}$. Let $p_i \equiv p_i(x^{(N)})$ and $m_i \equiv m_i(x^{(N)})$ denote, respectively, the induced electric and magnetic dipole moments of particle i located at x_i, $i = 1, \cdots, N$. It follows from Maxwell's equations that the electric and magnetic fields for a given configuration $x^{(N)}$ are, in rationalized MKS units,

$$E(x \mid x^{(N)}) = E_0(x) - k_0^2 \epsilon_0^{-1} \sum_{j=1}^{N} G^{(1)}(x, x_j) \cdot p_j$$

$$+ i\omega\mu_0 \sum_{j=1}^{N} G^{(2)}(x, x_j) \times m_j, \quad (2.1)$$

$$H(x \mid x^{(N)}) = H_0(x) - i\omega \sum_{j=1}^{N} G^{(2)}(x, x_j) \times p_j$$

$$- k_0^2 \sum_{j=1}^{N} G^{(1)}(x, x_j) \cdot m_j. \quad (2.2)$$

[13] F. Reiche, Ann. Physik **50**, 1, 121 (1916).
[14] H. Hoek, Doctoral dissertation, Leiden (1939), Physica **8**, 209 (1941).
[15] L. Foldy, Phys. Rev. **67**, 107 (1945).
[16] V. Twersky, J. Math. Phys. **3**, 700 (1962).
[17] M. Lax, Rev. Mod. Phys. **23**, 287 (1951).

In (2.1) and (2.2) ϵ_0 and μ_0 denote the electric and magnetic inductive capacities of free space, and $k_0 = \omega(\epsilon_0\mu_0)^{\frac{1}{2}}$ is the propagation constant or wavenumber of a wave of angular frequency ω in free space. The functions $G^{(1)}$ and $G^{(2)}$ are, respectively, the dyadic (tensor) and vector Green's functions defined by

$$G^{(1)}(x, y) = G^{(1)}(y - x) = ik_0(6\pi)^{-1} \text{ P.V. } [h_0^{(2)}(k_0 r)I$$

$$- \tfrac{1}{2}h_2^{(2)}(k_0 r)\{I - 3rr^{-2}\}] + I(3k_0^2)^{-1}\delta(r), \quad (2.3)$$

$$G^{(2)}(x, y) = G^{(2)}(y - x) = ik_0^2(4\pi r)^{-1}rh_1^{(2)}(k_0 r). \quad (2.4)$$

In (2.3) and (2.4), $r = y - x$, $r = |r|$, I is the unit dyad or matrix, $h_n^{(2)}(k_0 r)$ is the spherical Hankel function of the second kind of order n, and P.V. indicates that, whenever the term which it precedes is integrated, the integral is to be defined as a principal value integral. This means that the integral is the limit, as ϵ tends to zero, of an integral with respect to y over a domain excluding a sphere of radius ϵ centered at the point x.

In order to write (2.1) and (2.2) and succeeding formulas in a more compact form, we define the 6-component vectors π_i, F, and F_0 by

$$\pi_i = \begin{Bmatrix} p_i \\ m_i \end{Bmatrix}, \quad F(x \mid x^{(N)}) = \begin{Bmatrix} E(x \mid x^{(N)}) \\ H(x \mid x^{(N)}) \end{Bmatrix},$$

$$F_0(x) = \begin{Bmatrix} E_0(x) \\ H_0(x) \end{Bmatrix}. \quad (2.5)$$

Equations (2.1) and (2.2) then become the single 6-vector equation

$$F(x \mid x^{(N)}) = F_0(x) + \sum_{j=1}^{N} \Gamma(x, x_j)\pi_j. \quad (2.6)$$

Here $\Gamma(x, x_j)$ is the 6×6 matrix of the coefficients appearing in (2.1) and (2.2).

We assume that the moments induced in particle i are proportional to the fields $E'(x_i \mid x^{(N)})$, $H'(x_i \mid x^{(N)})$ incident upon it. These fields are given by (2.1) and (2.2) or (2.6) by setting $x = x_i$ and omitting the term with $j = i$. For the sake of generality, we admit the possibility that each moment is proportional to both fields and write

$$p_i = \alpha^{(11)}E'(x_i \mid x^{(N)}) + \alpha^{(12)}H'(x_i \mid x^{(N)}), \quad (2.7)$$

$$m_i = \alpha^{(21)}E'(x_i \mid x^{(N)}) + \alpha^{(22)}H'(x_i \mid x^{(N)}). \quad (2.8)$$

The coefficients $\alpha^{(jk)}$ are the 3×3 polarizability matrices or tensors of a particle, which can be combined into the 6×6 matrix α defined by

$$\alpha = \begin{pmatrix} \alpha^{(11)} & \alpha^{(12)} \\ \alpha^{(21)} & \alpha^{(22)} \end{pmatrix}. \quad (2.9)$$

By using (2.9) and (2.5), we can rewrite (2.7) and (2.8) as

$$\pi_i = \alpha F'(x_i \mid x^{(N)}). \quad (2.10)$$

Upon inserting (2.6) into (2.10), we obtain the following matrix equations, which must be satisfied by the moments $\mathbf{p}_i$ and $\mathbf{m}_i$:

$$\pi_i = \alpha\left\{F_0(\mathbf{x}_i) + \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j)\pi_j\right\}; \quad i = 1, \cdots, N.$$
(2.11)

The delta-function part of the Green's function $G^{(1)}$ does not appear in (2.11), due to the exclusion of the term with $j = i$ from the summation and the fact that $x_i \neq x_j$ for $i \neq j$.

In terms of the moments, the electric and magnetic polarizations $\mathbf{P}(\mathbf{x} \mid \mathbf{x}^{(N)})$ and $\mathbf{M}(\mathbf{x} \mid \mathbf{x}^{(N)})$ are defined by

$$\mathbf{P}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \sum_i \mathbf{p}_i \delta(\mathbf{x} - \mathbf{x}_i),$$
(2.12)

$$\mathbf{M}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \sum_i \mathbf{m}_i \delta(\mathbf{x} - \mathbf{x}_i).$$
(2.13)

By introducing the 6-vector

$$\Pi(\mathbf{x} \mid \mathbf{x}^{(N)}) = \{\mathbf{P}(\mathbf{x} \mid \mathbf{x}^{(N)}), \mathbf{M}(\mathbf{x} \mid \mathbf{x}^{(N)})\},$$

we can write (2.12) and (2.13) as

$$\Pi(\mathbf{x} \mid \mathbf{x}^{(N)}) = \sum_i \pi_i \delta(\mathbf{x} - \mathbf{x}_i).$$
(2.14)

The electric displacement $\mathbf{D}(\mathbf{x} \mid \mathbf{x}^{(N)})$ and magnetic induction $\mathbf{B}(\mathbf{x} \mid \mathbf{x}^{(N)})$ are given by

$$\mathbf{D}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \epsilon_0 \mathbf{E}(\mathbf{x} \mid \mathbf{x}^{(N)}) + \mathbf{P}(\mathbf{x} \mid \mathbf{x}^{(N)}),$$
(2.15)

$$\mathbf{B}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \mu_0[\mathbf{H}(\mathbf{x} \mid \mathbf{x}^{(N)}) + \mathbf{M}(\mathbf{x} \mid \mathbf{x}^{(N)})].$$
(2.16)

Again introducing a 6-vector

$$\mathfrak{D}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \{\mathbf{D}(\mathbf{x} \mid \mathbf{x}^{(N)}), \mathbf{B}(\mathbf{x} \mid \mathbf{x}^{(N)})\},$$

we can write (2.15) and (2.16) in the form

$$\mathfrak{D}(\mathbf{x} \mid \mathbf{x}^{(N)}) = \begin{pmatrix} \epsilon_0 I & 0 \\ 0 & \mu_0 I \end{pmatrix} F(\mathbf{x} \mid \mathbf{x}^{(N)})$$
$$+ \begin{pmatrix} I & 0 \\ 0 & \mu_0 I \end{pmatrix} \Pi(\mathbf{x} \mid \mathbf{x}^{(N)}). \quad (2.17)$$

The problem of determining the fields first requires solving (2.11) for the moments π_i, assuming that the incident field $F_0(\mathbf{x})$, the polarizabilities α, and the particle positions $\mathbf{x}_i$ are given. Then the fields F and $\mathfrak{D}$ are given by (2.6) and (2.17).

3. STATISTICS

We consider now an ensemble of particle configurations and identify the macroscopic fields and polarizations with the ensemble averages of the corresponding quantities for a given configuration. We assume that the N-particle probability distribution for the positions $\mathbf{x}_i$ is a symmetric function of the $\mathbf{x}_i$ denoted by

$$\nu^{(N)}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \nu^{(N)}(\mathbf{x}^{(N)}).$$

Then, as is usual in statistical mechanics, lower-order distribution functions $\nu^{(n)}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ are defined by

$$\nu^{(n)}(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \int \nu^{(N)}(\mathbf{x}_1, \cdots, \mathbf{x}_N) \, d\mathbf{x}_{n+1} \cdots d\mathbf{x}_N.$$
(3.1)

The one-particle density $\rho(\mathbf{x}_i)$ is defined by

$$\rho(\mathbf{x}_i) = N\nu^{(1)}(\mathbf{x}_i).$$
(3.2)

In terms of ρ, the n-particle correlation function $g^{(n)}(\mathbf{x}_i, \cdots, \mathbf{x}_n)$ is defined by

$$\nu^{(n)}(\mathbf{x}_1, \cdots, \mathbf{x}_n)$$
$$= \frac{(N - n)!}{N!} \rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n) g^{(n)}(\mathbf{x}_1, \cdots, \mathbf{x}_n). \quad (3.3)$$

With the aid of $\nu^{(N)}$, the ensemble average of any quantity $A(\mathbf{x} \mid \mathbf{x}^{(N)})$ is defined to be

$$\langle A(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle = \int A(\mathbf{x} \mid \mathbf{x}^{(N)}) \nu^{(N)}(\mathbf{x}^{(N)}) \, d\mathbf{x}^{(N)}. \quad (3.4)$$

A partial average with respect to the positions of all but one particle, say the jth, is written as $\langle A \rangle_j$. It is given by

$$\langle A(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle_j = [\nu^{(1)}(\mathbf{x}_j)]^{-1}$$
$$\times \int A(\mathbf{x} \mid \mathbf{x}^{(N)}) \nu^{(N)}(\mathbf{x}^{(N)}) \, d\mathbf{x}_1 \cdots d\mathbf{x}_{j-1} \, d\mathbf{x}_{j+1} \cdots d\mathbf{x}_N.$$
(3.5)

From (3.4) and (3.5) we see that, for any $A(\mathbf{x} \mid \mathbf{x}^{(N)})$, $\langle A \rangle$ is related to $\langle A \rangle_j$ by

$$\langle A(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle = \int \langle A(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle_j \, \nu^{(1)}(\mathbf{x}_j) \, d\mathbf{x}_j. \quad (3.6)$$

We now use the preceding definitions and relations to compute the ensemble averages of (2.17), (2.14), and (2.6). By using (3.5) and (3.6), we obtain

$$\langle \mathfrak{D}(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle = \begin{pmatrix} \epsilon_0 I & 0 \\ 0 & \mu_0 I \end{pmatrix} \langle F(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle$$
$$+ \begin{pmatrix} I & 0 \\ 0 & \mu_0 I \end{pmatrix} \langle \Pi(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle, \quad (3.7)$$

$$\langle \Pi(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle = \rho(\mathbf{x})[\langle \pi_i \rangle_i]_{\mathbf{x}_i = \mathbf{x}}, \quad (3.8)$$

$$\langle F(\mathbf{x} \mid \mathbf{x}^{(N)}) \rangle = F_0(\mathbf{x}) + \int \Gamma(\mathbf{x}, \mathbf{x}_j)\rho(\mathbf{x}_j)\langle \pi_j \rangle_j \, d\mathbf{x}_j. \quad (3.9)$$

These three equations show that the macroscopic fields and polarizations can be determined from the function $\langle \pi_i \rangle_i$, which is the average moment of particle i when it is held fixed at $\mathbf{x}_i$. Therefore it is this function which we wish to obtain.

4. EQUATION FOR $\langle \pi_i \rangle_i$

In order to determine the function $\langle \pi_i \rangle_i$, we must solve (2.11) for π_i and then average the solution over

all configurations, holding $\mathbf{x}_i$ fixed. It is tempting to average (2.11) first in order to obtain an equation for $\langle \pi_i \rangle_i$, but this yields

$$\langle \pi_i \rangle_i = \alpha \left\{ F_0(\mathbf{x}_i) + \int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \rho(\mathbf{x}_j) g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \langle \pi_j \rangle_{ij}\, d\mathbf{x}_j \right\}. \tag{4.1}$$

We see that this equation contains the additional unknown function $\langle \pi_j \rangle_{ij}$, which is the average of $\pi_j(\mathbf{x}^{(N)})$ with respect to the positions of all particles except those at $\mathbf{x}_i$ and $\mathbf{x}_j$. Thus (4.1) is unsuitable for the determination of $\langle \pi_i \rangle_i$. An attempt to obtain an equation for $\langle \pi_i \rangle_{ij}$ introduces $\langle \pi_j \rangle_{ijk}$, and continuation of this procedure leads to a hierarchy of equations for $\langle \pi_j \rangle_{ijk}$, $\langle \pi_j \rangle_{ijkm}$, etc. To avoid considering this hierarchy, Reiche,[13] Hoek,[14] and Rosenfeld[12] replace $\langle \pi_j \rangle_{ij}$ by $\langle \pi_j \rangle_j$ in (1). This yields an equation for $\langle \pi_i \rangle_i$. We do not use this procedure because it involves an unjustified replacement and gives no indication of when the resulting equation is valid. In addition, it does not provide any way to improve the result.

Instead of averaging (2.11) we proceed differently. First we suppose that (2.11) can be solved for π_i. Then π_i is a linear combination of the quantities $F_0(x_j)$, $j = 1, \cdots, N$, which we write in symbolic form in terms of some linear operator Q as

$$\pi_i = Q F_0. \tag{4.2}$$

Then we average (4.2) with x_i fixed to obtain

$$\langle \pi_i \rangle_i = \langle Q F_0 \rangle_i. \tag{4.3}$$

This type of result, which expresses $\langle \pi_i \rangle_i$ in terms of the incident field, is useful for calculating the average field scattered by a finite collection of particles, such as those constituting a medium of finite size. However, to determine the refractive index, dielectric constant, and permeability, it is advantageous to consider an infinite collection of particles constituting an unbounded medium. Doing so eliminates effects due to the size and shape of the medium. In such a medium there can be freely propagating waves or fields for which $F_0 \equiv 0$. They may be thought of as being produced by sources at infinity. To determine them we consider (4.3) as an integral equation for $F_0(x)$ with $\langle \pi_i \rangle_i$ given. We write its solution in terms of some linear operator L as

$$F_0 = L \langle \pi_i \rangle_i. \tag{4.4}$$

To find the free or natural modes we seek solutions of (4) with $F_0 = 0$:

$$L \langle \pi_i \rangle_i = 0. \tag{4.5}$$

We assume that the medium is statistically homogeneous. This implies that the operator L is translationally invariant. As a consequence its eigenfunctions are plane waves, so we seek a solution of (4.5) of the form

$$\langle \pi_i \rangle_i = A e^{-ink_0 \cdot \mathbf{x}_i}. \tag{4.6}$$

Here A is a constant 6-vector, k_0 is the propagation constant in free space, and n is the complex refractive index, which is to be determined along with A. Upon inserting (4.6) into (4.5) and multiplying the resulting equation on the left by $e^{ink_0 \cdot \mathbf{x}_i}$, we obtain

$$e^{ink_0 \cdot \mathbf{x}_i} L e^{-ink_0 \cdot \mathbf{x}_i} A = 0. \tag{4.7}$$

This is a set of six linear equations for the components of A. They have a nontrivial solution if and only if the coefficient matrix is singular, which implies

$$\det \left[e^{ink_0 \cdot \mathbf{x}_i} L e^{-ink_0 \cdot \mathbf{x}_i} \right] = 0. \tag{4.8}$$

The matrix in (4.8) is independent of $\mathbf{x}_i$ because L is translationally invariant.

Equation (4.8) is the exact dispersion equation for the refractive index n as a function of ω, α, and the statistical properties of the particle distribution. If the medium is statistically isotropic, n is independent of the direction of propagation, i.e., the direction of $\mathbf{k}_0$. When n is found to satisfy (4.8), (4.7) can be solved for A. To make the formal results (4.7) and (4.8) useful, we must calculate L. In the next two sections we present two different ways of obtaining series representations for L. The first yields a series in powers of the polarizability α, and the second yields a series in powers of ρ, the number density of particles.

5. EXPANSION IN THE POLARIZABILITY α

We now solve (2.11) for π_i as a power series in α. This solution can be obtained by iteration. Upon writing the first few terms explicitly and indicating that the rest are $O(\alpha^4)$, we have

$$\pi_i = \alpha F_0(\mathbf{x}_i) + \alpha \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha F_0(\mathbf{x}_j)$$
$$+ \alpha \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \sum_{k \neq j} \Gamma(\mathbf{x}_j, \mathbf{x}_k) \alpha F_0(\mathbf{x}_k) + O(\alpha^4). \tag{5.1}$$

This equation is an explicit form of (4.2). Before averaging this solution, it is convenient to rewrite it in the form

$$\pi_i = \alpha F_0(\mathbf{x}_i) + \alpha \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha F_0(\mathbf{x}_j)$$
$$+ \alpha \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \Gamma(\mathbf{x}_j, \mathbf{x}_i) \alpha F_0(\mathbf{x}_i)$$
$$+ \alpha \sum_{j \neq i} \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \sum_{\substack{k \neq j \\ k \neq i}} \Gamma(\mathbf{x}_j, \mathbf{x}_k) \alpha F_0(\mathbf{x}_k) + O(\alpha^4). \tag{5.2}$$

Now we average (5.2) with $\mathbf{x}_i$ fixed to obtain the

following explicit form of (4.3):

$$\langle \pi_i \rangle_i = \alpha F_0(\mathbf{x}_i)$$

$$+ \alpha \int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \rho(\mathbf{x}_j) g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \alpha F_0(\mathbf{x}_j) \, d\mathbf{x}_j$$

$$+ \alpha \int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \Gamma(\mathbf{x}_j, \mathbf{x}_i) \rho(\mathbf{x}_j) g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \alpha F_0(\mathbf{x}_i) \, d\mathbf{x}_j$$

$$+ \alpha \int\int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \Gamma(\mathbf{x}_j, \mathbf{x}_k) \rho(\mathbf{x}_j) \rho(\mathbf{x}_k)$$

$$\times g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \alpha F_0(\mathbf{x}_k) \, d\mathbf{x}_j \, d\mathbf{x}_k + O(\alpha^4). \qquad (5.3)$$

Next we solve (5.3) for $F_0(\mathbf{x}_i)$ by iteration or successive substitution to obtain

$$\alpha F_0(\mathbf{x}_i) = \langle \pi_i \rangle_i - \alpha \int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \rho(\mathbf{x}_j) g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \langle \pi_j \rangle_j \, d\mathbf{x}_j$$

$$- \alpha \int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \Gamma(\mathbf{x}_j, \mathbf{x}_i) \rho(\mathbf{x}_j) g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) \langle \pi_i \rangle_i \, d\mathbf{x}_j$$

$$- \alpha \int\int \Gamma(\mathbf{x}_i, \mathbf{x}_j) \alpha \Gamma(\mathbf{x}_j, \mathbf{x}_k) \rho(\mathbf{x}_j) \rho(\mathbf{x}_k)$$

$$\times [g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) - g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) g^{(2)}(\mathbf{x}_j, \mathbf{x}_k)]$$

$$\times \langle \pi_k \rangle_k \, d\mathbf{x}_j \, d\mathbf{x}_k + O(\alpha^3). \qquad (5.4)$$

This is an explicit form of (4.4).

For a statistically homogeneous medium we have

$$\rho(\mathbf{x}_i) = \text{const} = \rho, \qquad (5.5a)$$

$$g^{(2)}(\mathbf{x}_i, \mathbf{x}_j) = g^{(2)}(\mathbf{x}_j - \mathbf{x}_i), \qquad (5.5b)$$

$$g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) = g^{(3)}(\mathbf{x}_j - \mathbf{x}_i, \mathbf{x}_k - \mathbf{x}_j); \text{ etc.} \qquad (5.5c)$$

By using (5.5) and the fact that $\Gamma(\mathbf{x}_i, \mathbf{x}_j) = \Gamma(\mathbf{x}_j - \mathbf{x}_i)$, and setting $F_0 = 0$, we can write (5.4) as

$$\langle \pi_i \rangle_i - \alpha\rho \int \Gamma(\mathbf{x}_j - \mathbf{x}_i) g^{(2)}(\mathbf{x}_j - \mathbf{x}_i) \langle \pi_j \rangle_j \, d\mathbf{x}_j$$

$$- \alpha\rho \int \Gamma(\mathbf{x}_j - \mathbf{x}_i) \alpha \Gamma(\mathbf{x}_i - \mathbf{x}_j) g^{(2)}(\mathbf{x}_j - \mathbf{x}_i) \langle \pi_i \rangle_i \, d\mathbf{x}_j$$

$$- \alpha\rho^2 \int\int \Gamma(\mathbf{x}_j - \mathbf{x}_i) \alpha \Gamma(\mathbf{x}_k - \mathbf{x}_j)$$

$$\times [g^{(3)}(\mathbf{x}_j - \mathbf{x}_i, \mathbf{x}_k - \mathbf{x}_j)$$

$$- g^{(2)}(\mathbf{x}_j - \mathbf{x}_i) g^{(2)}(\mathbf{x}_k - \mathbf{x}_j)]$$

$$\times \langle \pi_k \rangle_k \, d\mathbf{x}_j \, d\mathbf{x}_k + O(\alpha^3) = 0. \qquad (5.6)$$

Now we insert the plane wave form (4.6) into (5.6), multiply on the left by $e^{in\mathbf{k}_0 \cdot \mathbf{x}_i}$, and obtain

$$\Big\{ I - \alpha\rho \int \Gamma(\mathbf{R}) g^{(2)}(\mathbf{R}) e^{-in\mathbf{k}_0 \cdot \mathbf{R}} \, d\mathbf{R}$$

$$- \alpha\rho \int \Gamma(\mathbf{R}) \alpha \Gamma(-\mathbf{R}) g^{(2)}(\mathbf{R}) \, d\mathbf{R}$$

$$- \alpha\rho^2 \int\int \Gamma(\mathbf{R}) \alpha \Gamma(\mathbf{S}) [g^{(3)}(\mathbf{R}, \mathbf{S}) - g^{(2)}(\mathbf{R}) g^{(2)}(\mathbf{S})]$$

$$\times e^{-in\mathbf{k}_0 \cdot (\mathbf{R}+\mathbf{S})} \, d\mathbf{R} \, d\mathbf{S} + O(\alpha^3) \Big\} A = 0. \qquad (5.7)$$

Here $\mathbf{R} = \mathbf{x}_j - \mathbf{x}_i$ and $\mathbf{S} = \mathbf{x}_k - \mathbf{x}_j$. This is an explicit form of (4.7).

In order that (5.7) has a solution A which is not zero, the determinant of the expression in braces must vanish:

$$\det \Big[I - \alpha\rho \int \Gamma(\mathbf{R}) g^{(2)}(\mathbf{R}) e^{-in\mathbf{k}_0 \cdot \mathbf{R}} \, d\mathbf{R}$$

$$- \alpha\rho \int \Gamma(\mathbf{R}) \alpha \Gamma(-\mathbf{R}) g^{(2)}(\mathbf{R}) \, d\mathbf{R}$$

$$- \alpha\rho^2 \int\int \Gamma(\mathbf{R}) \alpha \Gamma(\mathbf{S}) [g^{(3)}(\mathbf{R}, \mathbf{S}) - g^{(2)}(\mathbf{R}) g^{(2)}(\mathbf{S})]$$

$$\times e^{-in\mathbf{k}_0 \cdot (\mathbf{R}+\mathbf{S})} \, d\mathbf{R} \, d\mathbf{S} + O(\alpha^3) \Big] = 0. \qquad (5.8)$$

This is an explicit form of the dispersion equation (4.8) as a power series in the polarizability α. Terms beyond those shown explicitly can be found by continuing the iteration processes employed in solving (2.11) and (5.3). This dispersion equation is one of our main results. We solve it in Sec. 7.

The term proportional to α^n in (5.8) corresponds to a wave which is scattered $n-1$ times before arriving at $\mathbf{x}_i$. This term is composed of parts in which s distinct particles are involved, where $s = 2, \cdots, n$, and each such part is proportional to ρ^{s-1}. Thus the term proportional to α^n contains parts proportional to $\rho, \rho^2, \cdots, \rho^{n-1}$. In order to obtain a series in powers of ρ instead of the series in powers of α in (5.8), we could rearrange the series in α by collecting all terms proportional to each power of ρ. Since ρ^n occurs in infinitely many terms in (5.8), we would have to sum an infinite number of terms to obtain its coefficient. This summation can be described by associating a diagram with each term. Each variable $\mathbf{x}_j$ in the term corresponds to a vertex labeled $\mathbf{x}_j$, and each factor $\Gamma(\mathbf{x}_i, \mathbf{x}_j)$ corresponds to a line or bond joining the vertices $\mathbf{x}_i$ and $\mathbf{x}_j$. Then the term of order α^n is obtained by summing contributions from all diagrams with exactly $n-1$ bonds, while the term of order ρ^n is obtained by summing contributions from all diagrams with exactly $n+1$ vertices. Although this method of resummation of the series (5.8) can be carried out, it is unnecessary. We now present a different method which yields the density expansion directly and in a more useful form.

6. EXPANSION IN THE DENSITY ρ

To obtain a density (or virial) expansion of the dispersion equation, we begin with a different way of solving (2.11) for π_i. Let us define $\pi_{ii_1 \ldots i_{s-1}}^s$ to be the exact solution of (2.11) for s particles at $\mathbf{x}_i$, $\mathbf{x}_{i_1}, \cdots, \mathbf{x}_{i_{s-1}}$. This solution is clearly a symmetric function of $i_1 \cdots i_{s-1}$. The subscripts on $\pi_{ii_1 \ldots i_{s-1}}^s$

must be distinct. For $s = 1$ and $s = 2$ we find from (2.11) the explicit expressions

$$\pi_i^1 = \alpha F_0(\mathbf{x}_i), \qquad (6.1)$$

$$\pi_{ij}^2 = [1 - \alpha\Gamma(\mathbf{x}_i, \mathbf{x}_j)\alpha\Gamma(\mathbf{x}_j, \mathbf{x}_i)]^{-1}$$
$$\times [\alpha F_0(\mathbf{x}_i) + \alpha\Gamma(\mathbf{x}_i, \mathbf{x}_j)\alpha F_0(\mathbf{x}_j)]. \quad (6.2)$$

In analogy with the Ursell functions of statistical mechanics, we introduce functions ϕ_i^s which represent all the "pure" s-particle scattering. Each ϕ_i^s is defined to be the sum over $i_1 \cdots i_{s-1}$ of $\pi_{ii_1\cdots i_{s-1}}^s$, minus all the $\pi_{ij_1\cdots j_{t-1}}^t$ with $t < s$ which are contained in $\pi_{ii_1\cdots i_{s-1}}^s$. Thus ϕ_i^s vanishes when any one of the points $\mathbf{x}_{i_1}\cdots \mathbf{x}_{i_{s-1}}$ is at an infinite distance from $\mathbf{x}_i$. The first few are given by

$$\phi_i^1 = \pi_i^1, \qquad (6.3)$$

$$\phi_i^2 = \sum_{\substack{j=1 \\ j\neq i}}^N (\pi_{ij}^2 - \pi_i^1), \qquad (6.4)$$

$$\phi_i^3 = \sum_{\substack{j,k=1 \\ j\neq i,j\neq k,k\neq i}}^N (\pi_{ijk}^3 - \pi_{ij}^2 - \pi_{ik}^2 + \pi_i^1), \qquad (6.5)$$

$$\phi_i^4 = \sum_{\substack{i,j,k,l=1 \\ ijkl\,\text{distinct}}}^N (\pi_{ijkl}^4 - \pi_{ijk}^3 - \pi_{ijl}^3 - \pi_{ikl}^3 $$
$$+ \pi_{ij}^2 + \pi_{ik}^2 + \pi_{il}^2 - \pi_i^1). \quad (6.6)$$

For any s the average of ϕ_i^s with $\mathbf{x}_i$ fixed is given by (6.9) below.

In terms of the ϕ_i^s we can write the desired solution π_i^N as

$$\pi_i \equiv \pi_{ii_1\cdots i_{N-1}}^N = \sum_{s=1}^N \phi_i^s. \qquad (6.7)$$

Equation (6.7) is an identity because the ϕ_i^s are defined so that the coefficient of each $\pi_{ii_1\cdots i_{s-1}}^s$ in the sum is zero if $s < N$ and the coefficient of $\pi_{ii_1\cdots i_{N-1}}^N$ is unity. This is proved in Appendix A for the average of (6.7) with $\mathbf{x}_i$ fixed, which is all that we need, since our derivation is based on this average. The proof of (6.7) itself is similar. Equation (6.7) is a special case of (4.2). Upon averaging (6.7) with x_i fixed and using (3.3) with $\rho = $ const, we obtain

$$\langle \pi_i \rangle_i = \sum_{s=1}^N \langle \phi_i^s \rangle_i = \sum_{s=1}^N \frac{(N-s)!}{(N-1)!} \rho^{s-1}$$
$$\times \int g^{(s)}(\mathbf{x}_i, \mathbf{x}_{i_1}\cdots \mathbf{x}_{i_{s-1}})\phi_i^s\, d\mathbf{x}_{i_1}\cdots d\mathbf{x}_{i_{s-1}}. \quad (6.8)$$

This result (6.8) yields $\langle \pi_i \rangle_i$ as a polynomial in ρ, which becomes a power series if $N = \infty$.

Using the fact that $\pi_{ii_1\cdots i_{s-1}}^s$ is symmetric in its last

$s - 1$ subscript, we obtain, from the definition of ϕ_i^s,

$$\langle \phi_i^s \rangle_i = \binom{N-1}{s-1} \sum_{l=1}^s (-1)^{s+l}\binom{s-1}{l-1}\langle \pi_{ii_1\cdots i_{l-1}}^l \rangle_i. \qquad (6.9)$$

By using (6.9) in (6.8), we get

$$\langle \pi_i \rangle_i = \sum_{s=1}^N \frac{(-1)^s \rho^{s-1}}{(s-1)!} \int\cdots\int \sum_{l=1}^s \left[(-1)^l \binom{s-1}{l-1}\pi_{ii_1\cdots i_{l-1}}^l \right]$$
$$\times g^{(s)}(\mathbf{x}_i, \mathbf{x}_{i_1}\cdots \mathbf{x}_{i_{s-1}})\, d\mathbf{x}_{i_1}\cdots d\mathbf{x}_{i_{s-1}}. \quad (6.10)$$

The first few terms in this expansion are

$$\langle \pi_i \rangle_i = \pi_i^1 + \rho \int [\pi_{ij}^2 - \pi_i^1]g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)\, d\mathbf{x}_j$$
$$+ \frac{\rho}{2} \int [\pi_{ijk}^3 - 2\pi_{ij}^2 + \pi_i^1]$$
$$\times g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)\, d\mathbf{x}_j\, d\mathbf{x}_k + O(\rho^3). \quad (6.11)$$

We now insert into (6.11) the expressions such as (6.1) and (6.2) for the $\pi_{ii_1\cdots i_{s-1}}^s$ in terms of F_0. In writing the result it is convenient to introduce the abbreviation

$$T_{ij} = \alpha\Gamma(\mathbf{x}_i, \mathbf{x}_j). \qquad (6.12)$$

Then we obtain from (6.11)

$$\langle \pi_i \rangle_i = \alpha F_0(\mathbf{x}_i) + \rho \int g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)(1 - T_{ij}T_{ji})^{-1}$$
$$\times T_{ij}[\alpha F_0(\mathbf{x}_j) + T_{ji}\alpha F_0(\mathbf{x}_i)]\, d\mathbf{x}_j$$
$$+ \rho^2 \int g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)[C_{ijk}\alpha F_0(\mathbf{x}_i)$$
$$+ D_{ijk}\alpha F_0(\mathbf{x}_j) + E_{ijk}\alpha F_0(\mathbf{x}_k)]\, d\mathbf{x}_j\, d\mathbf{x}_k + O(\rho^3). \quad (6.13)$$

The matrices C_{ijk}, D_{ijk}, and E_{ijk} are defined by

$$2C_{ijk} = 1 + H_{ijk} - (1 - T_{ij}T_{ji})^{-1} - (1 - T_{ik}T_{ki})^{-1}, \qquad (6.14)$$

$$2D_{ijk} = H_{ijk}(T_{ij} + T_{ik}T_{kj})(1 - T_{jk}T_{kj})^{-1}$$
$$- (1 - T_{ij}T_{ji})^{-1}T_{ij}, \quad (6.15)$$

$$2E_{ijk} = H_{ijk}[(T_{ij} + T_{ik}T_{kj})(1 - T_{jk}T_{kj})^{-1}T_{jk} + T_{ik}]$$
$$- (1 - T_{ik}T_{ki})^{-1}T_{ik}. \quad (6.16)$$

Here H_{ijk} is defined by

$$H_{ijk} = [1 - (1 - T_{ik}T_{ki})^{-1}(T_{ij} + T_{ik}T_{kj})(1 - T_{jk}T_{kj})^{-1}$$
$$\times (T_{ji} + T_{jk}T_{ki})]^{-1}(1 - T_{ik}T_{ki})^{-1}. \quad (6.17)$$

We see that (6.13) is of the form (4.3).

Next we solve (6.13) for $\alpha F_0(\mathbf{x}_i)$ in terms of $\langle \pi_i \rangle_i$

by iteration and obtain the result:

$$\alpha F_0(\mathbf{x}_i) = \langle \pi_i \rangle_i - \rho \int g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)(1 - T_{ij}T_{ji})^{-1}$$

$$\times T_{ij}(T_{ji}\langle \pi_i \rangle_i + \langle \pi_j \rangle_j) \, d\mathbf{x}_j$$

$$+ \rho^2 \int [g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)g^{(2)}(\mathbf{x}_i, \mathbf{x}_k)(1 - T_{ij}T_{ji})^{-1}$$

$$\times T_{ij}T_{ji}(1 - T_{ik}T_{ki})^{-1}T_{ik}T_{ki}$$

$$+ g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)C_{ijk}]\langle \pi_i \rangle_i \, d\mathbf{x}_j \, d\mathbf{x}_k$$

$$+ \rho^2 \int [g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)g^{(2)}(\mathbf{x}_i, \mathbf{x}_k)(1 - T_{ij}T_{ji})^{-1}$$

$$\times T_{ij}(1 - T_{jk}T_{kj})^{-1}T_{jk}T_{kj}$$

$$+ g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)D_{ijk}]\langle \pi_j \rangle_j \, d\mathbf{x}_j \, d\mathbf{x}_k$$

$$+ \rho^2 \int [g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)g^{(2)}(\mathbf{x}_i, \mathbf{x}_k)(1 - T_{ij}T_{ji})^{-1}$$

$$\times T_{ij}T_{ji}(1 - T_{ik}T_{ki})^{-1}T_{ik}$$

$$+ g^{(2)}(\mathbf{x}_i, \mathbf{x}_j)g^{(2)}(\mathbf{x}_j, \mathbf{x}_k)(1 - T_{ij}T_{ji})^{-1}$$

$$\times T_{ij}(1 - T_{jk}T_{kj})^{-1}T_{jk}$$

$$+ g^{(3)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)E_{ijk}]\langle \pi_k \rangle_k \, d\mathbf{x}_j \, d\mathbf{x}_k + O(\rho^3).$$
$$(6.18)$$

This equation is of the form (4.4).

To obtain plane wave solutions of (6.18) we set $F_0 = 0$, assume that the medium is statistically homogeneous, insert for $\langle \pi_i \rangle_i$ the plane wave form (4.6), and multiply on the left by $e^{in\mathbf{k}_0 \cdot \mathbf{x}_i}$. In this way we obtain from (6.18)

$$\left\{ I - \rho \int g^{(2)}(\mathbf{R})[1 - \alpha\Gamma(\mathbf{R})\alpha\Gamma(-\mathbf{R})]^{-1}\alpha\Gamma(\mathbf{R}) \right.$$

$$\left. \times [\alpha\Gamma(\mathbf{R}) + e^{-in\mathbf{k}_0 \cdot \mathbf{R}}] \, d\mathbf{R} + O(\rho^2) \right\} A = 0. \quad (6.19)$$

The condition that (6.19) has a nontrivial solution for A is the vanishing of the determinant

$$\det \left\{ I - \rho \int g^{(2)}(\mathbf{R})[1 - \alpha\Gamma(\mathbf{R})\alpha\Gamma(-\mathbf{R})]^{-1}\alpha\Gamma(\mathbf{R}) \right.$$

$$\left. \times [\alpha\Gamma(\mathbf{R}) + e^{-in\mathbf{k}_0 \cdot \mathbf{R}}] \, d\mathbf{R} + O(\rho^2) \right\} = 0. \quad (6.20)$$

The result (6.20) is the dispersion equation for n, expressed as a power series in the particle number density ρ. This is our second main result.

7. ANALYSIS OF THE DISPERSION EQUATION

We now analyze the dispersion equation (5.8) for an isotropic medium, retaining only the constant term and the term linear in α. In the isotropic case

$$g^{(2)}(\mathbf{R}) = g^{(2)}(R). \quad (7.1)$$

Then (5.8) yields

$$\det \left[I - \alpha\rho \int \Gamma(\mathbf{R})g^{(2)}(R)e^{-in\mathbf{k}_0 \cdot \mathbf{R}} \, d\mathbf{R} \right] = 0. \quad (7.2)$$

To evaluate the integral in (7.2), we choose the z axis along the direction of $\mathbf{k}_0$; and after some straightforward calculation, we can rewrite (7.2) in the following explicit form, provided $n \neq 1$:

$$\det \left\{ I - \alpha\rho \begin{bmatrix} M_1 & M_2 \\ -M_2/\mu_0 & \epsilon_0 M_1 \end{bmatrix} \right\} = 0. \quad (7.3)$$

This is the dispersion equation for n in a homogeneous, isotropic medium, valid to first order in α. The matrices M_1 and M_2 are defined by

$$M_1 = \frac{2}{3\epsilon_0(n^2-1)} \begin{pmatrix} (1+c_0) + \tfrac{1}{2}n^2(1+c_2) & 0 & 0 \\ 0 & (1+c_0) + \tfrac{1}{2}n^2(1+c_2) & 0 \\ 0 & 0 & (1+c_0) - n^2(1+c_2) \end{pmatrix}, \quad (7.4)$$

$$M_2 = \left(\frac{\mu_0}{\epsilon_0}\right)^{\frac{1}{2}} \frac{n}{(n^2-1)}(1+c_1) \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (7.5)$$

The scalars $c_0(n)$, $c_1(n)$, and $c_2(n)$, which depend upon n, are defined by

$$c_m(n) \equiv \frac{k_0^3(n^2-1)}{in^m} \int_0^\infty R^2 \, dR\{g^{(2)}(R) - 1\}$$

$$\times h_m^{(2)}(k_0 R)j_m(nk_0 R), \quad m = 0, 1, 2. \quad (7.6)$$

Here $h_m^{(2)}$ and j_m are, respectively, the spherical Hankel function of the mth order of the second kind and the spherical Bessel function of mth order. When

the particle locations are uncorrelated, $g^{(2)}(R) = 1$, and then (7.6) shows that $c_m(n) = 0$, $m = 0, 1, 2$. Thus the $c_m(n)$ account for the correlations of particle positions.

Let us now specialize (7.3) to the case in which each particle has a scalar electric susceptibility χ_e and a scalar magnetic susceptibility χ_m. Then α is a diagonal matrix given by

$$\alpha = \text{diag} \{\epsilon_0\chi_e, \epsilon_0\chi_e, \epsilon_0\chi_e, \chi_m, \chi_m, \chi_m\}. \quad (7.7)$$

Upon using (7.7) in (7.3), we can evaluate the determinant explicitly and write it as the product of three factors. Equating each of the three factors to zero leads to the following three equations:

$$n^2 = \frac{1 + \frac{2}{3}\rho\chi_e[1 + c_0(n)]}{1 + \frac{2}{3}\rho\chi_e[1 + c_2(n)]}, \qquad (7.8a)$$

$$n^2 = \frac{1 + \frac{2}{3}\rho\chi_m[1 + c_0(n)]}{1 + \frac{2}{3}\rho\chi_m[1 + c_2(n)]}, \qquad (7.8b)$$

$$n^2 = \frac{1}{2}\{G_e + G_m + G_{em}$$
$$\pm [(G_e - G_m)^2 + G_{em}^2 + 2G_{em}(G_e + G_m)]^{\frac{1}{2}}\}. \qquad (7.8c)$$

In (7.8c), G_e, G_m, and G_{em} are defined by

$$G_e = \frac{1 + \frac{2}{3}\rho\chi_e[1 + c_0(n)]}{1 - \frac{1}{3}\rho\chi_e[1 + c_2(n)]}, \qquad (7.9a)$$

$$G_m = \frac{1 + \frac{2}{3}\rho\chi_m[1 + c_0(n)]}{1 - \frac{1}{3}\rho\chi_m[1 + c_2(n)]}, \qquad (7.9b)$$

$$G_{em} = \frac{\rho^2\chi_e\chi_m[1 + c_1(n)]^2}{\{1 - \frac{1}{3}\rho\chi_e[1 + c_2(n)]\}\{1 - \frac{1}{3}\rho\chi_m[1 + c_2(n)]\}}. \qquad (7.9c)$$

Since c_0, c_1, and c_2 depend on n, each of the equations (7.8) must still be solved for n.

We may now insert each of the results (7.8) into the matrix equation (5.7) and solve for the corresponding constant vector A. From each A, we can compute the corresponding macroscopic polarizations in the medium via (3.8) and (4.6). From (7.8a) we obtain in this way

$$\langle \mathbf{M}(\mathbf{x}) \rangle = 0, \qquad (7.10a)$$

$$\langle \mathbf{P}(\mathbf{x}) \rangle = a\mathbf{k}_0 e^{-in\mathbf{k}_0 \cdot \mathbf{x}}. \qquad (7.10b)$$

Here a is an arbitrary constant. Thus (7.8a) corresponds to a longitudinal wave of electric polarization and no magnetization. In a similar fashion, (7.8b) leads to

$$\langle \mathbf{M}(\mathbf{x}) \rangle = a'\mathbf{k}_0 e^{-in\mathbf{k}_0 \cdot \mathbf{x}}, \qquad (7.11a)$$

$$\langle \mathbf{P}(\mathbf{x}) \rangle = 0. \qquad (7.11b)$$

Thus (7.8b) corresponds to a longitudinal wave of magnetization and no electric polarization.

The relation (7.8c) leads to a wave in which both the electric polarization and the magnetization are transverse. They are given by

$$\langle \mathbf{P}(\mathbf{x}) \rangle = (\mathbf{a}'' \times \mathbf{k}_0) e^{-in\mathbf{k}_0 \cdot \mathbf{x}}, \qquad (7.12a)$$

$$\langle \mathbf{M}(\mathbf{x}) \rangle = b\mathbf{k}_0 \times \langle \mathbf{P}(\mathbf{x}) \rangle. \qquad (7.12b)$$

Here $\mathbf{a}''$ is an arbitrary constant vector, not parallel

to $\mathbf{k}_0$, and b is given by

$$b = \frac{n^2\{1 - \frac{1}{3}\rho\chi_e[1 + c_2(n)]\} - \{1 + \frac{2}{3}\rho\chi_e[1 + c_0(n)]\}}{k_0 n \rho \chi_e (\mu_0\epsilon_0)^{\frac{1}{2}}[1 + c_1(n)]} \qquad (7.13)$$

Thus (7.8c) is the dispersion equation for transverse waves, while (7.8a) and (7.8b) are the dispersion equations for longitudinal waves.

To solve any one of the dispersion equations (7.8) for the complex refractive index n, we must use (7.6) to determine the $c_m(n)$ which depend upon the radial distribution function $g^{(2)}(R)$. Let us first consider those cases in which the $c_m(n) = 0$. This is true for uncorrelated particle positions, $g^{(2)}(R) \equiv 1$, and for static fields, $k_0 = 0$, as we see from (7.6). When $c_m(n) = 0$, (7.8a)–(7.8c) with the minus sign all yield $n^2 = 1$, which was excluded in deriving (7.3), from which (7.8) follows. That case must be excluded because the integral in (7.2) diverges when $n^2 = 1$ since $g^{(2)}(R)$ tends to one as R tends to infinity. Thus only (7.8c) with the plus sign remains when $c_m(n) = 0$, and it yields the following explicit expression for n^2:

$$n^2 = \frac{1 + \frac{2}{3}\rho\chi_e}{1 - \frac{1}{3}\rho\chi_e} \cdot \frac{1 + \frac{2}{3}\rho\chi_m}{1 - \frac{1}{3}\rho\chi_m}. \qquad (7.14)$$

When the $c_m(n)$ are not zero, we can solve (7.8c) iteratively by first setting the $c_m(n) = 0$ and obtaining the value (7.14) for n^2. If we call this value n_0^2, we define n_1^2 by using $c_m(n_0)$ on the right side of (7.8c), etc. This method should work well when k_0, multiplied by the correlation length of the particle positions, is small. In gases and liquids this length is of the order of the range of interparticle forces. Therefore when the free space wavelength is large compared to the range of force, the iterative solution should converge rapidly.

8. DIELECTRIC CONSTANT, PERMEABILITY, AND AVERAGE FIELDS

The average electric and magnetic fields can be obtained by averaging (2.1) and (2.2) with respect to the particle locations. Let us average them and also assume that the average electric polarization and magnetization are plane waves of the forms

$$\langle \mathbf{P}(\mathbf{x}) \rangle = \mathbf{A}e^{-in\mathbf{k}_0 \cdot \mathbf{x}}, \qquad (8.1)$$

$$\langle \mathbf{M}(\mathbf{x}) \rangle = \mathbf{B}e^{-in\mathbf{k}_0 \cdot \mathbf{x}}. \qquad (8.2)$$

Then upon setting the external fields equal to zero, we obtain from (2.1) and (2.2)

$$\langle \mathbf{E}(\mathbf{x}) \rangle = [\epsilon_0(n^2 - 1)]^{-1}\{\langle \mathbf{P}(\mathbf{x}) \rangle - \mathbf{n}(\mathbf{n} \cdot \langle \mathbf{P}(\mathbf{x}) \rangle)$$
$$- (\mathbf{n}/c) \times \langle \mathbf{M}(\mathbf{x}) \rangle\}, \qquad (8.3)$$

$$\langle \mathbf{H}(\mathbf{x}) \rangle = (n^2 - 1)^{-1}\{\langle \mathbf{M}(\mathbf{x}) \rangle - \mathbf{n}(\mathbf{n} \cdot \langle \mathbf{M}(\mathbf{x}) \rangle)$$
$$+ cn \times \langle \mathbf{P}(\mathbf{x}) \rangle\}. \qquad (8.4)$$

Here $\mathbf{n} = nk_0^{-1}\mathbf{k}_0$ and $c = (\mu_0\epsilon_0)^{-\frac{1}{2}}$.

From (8.3) and (8.4) we derive expressions for the dielectric constant ϵ and permeability μ of the medium formed by the particles. They are defined by

$$\epsilon\langle\mathbf{E}(\mathbf{x})\rangle = \epsilon_0\langle\mathbf{E}(\mathbf{x})\rangle + \langle\mathbf{P}(\mathbf{x})\rangle, \qquad (8.5)$$

$$\mu\langle\mathbf{H}(\mathbf{x})\rangle = \mu_0\{\langle\mathbf{H}(\mathbf{x})\rangle + \langle\mathbf{M}(\mathbf{x})\rangle\}. \qquad (8.6)$$

For the case of longitudinal electric polarization, characterized by $\mathbf{n}\times\mathbf{A} = 0$ and $\mathbf{B} = 0$, we obtain from (8.3) and (8.4), $\langle\mathbf{E}(\mathbf{x})\rangle = -\epsilon_0^{-1}\langle\mathbf{P}(\mathbf{x})\rangle$ and $\langle\mathbf{H}(\mathbf{x})\rangle = 0$. Thus μ is not defined, while (8.5) yields $\epsilon = 0$. In the case of longitudinal magnetization, characterized by $\mathbf{n}\times\mathbf{B} = 0$ and $\mathbf{A} = 0$, (8.3) and (8.4) yield instead $\langle\mathbf{H}(\mathbf{x})\rangle = -\langle\mathbf{M}(\mathbf{x})\rangle$ and $\langle\mathbf{E}(\mathbf{x})\rangle = 0$. Now ϵ is not defined, while (8.6) yields $\mu = 0$.

In the case of transverse electric polarization and transverse magnetization, for which $n\cdot\mathbf{A} = n\cdot\mathbf{B} = 0$, we obtain from (8.3)–(8.6)

$$\left\{\frac{n^2-1}{\epsilon/\epsilon_0-1} - 1\right\}\langle\mathbf{P}(\mathbf{x})\rangle = -\frac{\mathbf{n}}{c}\times\langle\mathbf{M}(\mathbf{x})\rangle, \quad (8.7)$$

$$\left\{\frac{n^2-1}{\mu/\mu_0-1} - 1\right\}\langle\mathbf{M}(\mathbf{x})\rangle = c\mathbf{n}\times\langle\mathbf{P}(\mathbf{x})\rangle. \quad (8.8)$$

The condition that (8.7) and (8.8) be consistent yields

$$n^2 = \mu\epsilon/\mu_0\epsilon_0. \qquad (8.9)$$

To determine ϵ and μ in the transverse case, we combine (8.7) and (8.8) with (7.12). After a simple calculation we obtain

$$\epsilon/\epsilon_0 = 1 + (n^2-1)(1 + bnk_0/c)^{-1}, \quad (8.10)$$

$$\mu/\mu_0 = 1 + (n^2-1)(1 + cn/bk_0)^{-1}. \quad (8.11)$$

Here b is given by (7.13). We note that (8.10) and (8.11) satisfy (8.9).

Let us now specialize (8.10) and (8.11) to the case in which $\chi_m = 0$. Then (8.11) yields $\mu/\mu_0 = 1$, so that (8.9) becomes $\epsilon/\epsilon_0 = n^2$ and (8.10) leads to

$$\frac{\epsilon}{\epsilon_0} = n^2 = \frac{1 + \frac{2}{3}\rho\chi_e[1 + c_0(n)]}{1 - \frac{1}{3}\rho\chi_e[1 + c_2(n)]}. \qquad (8.12)$$

Another special case is that in which $\chi_e = 0$. Then (8.10) yields $\epsilon/\epsilon_0 = 1$ and (8.9) becomes $\mu/\mu_0 = n^2$. Thus (8.11) becomes

$$\frac{\mu}{\mu_0} = n^2 = \frac{1 + \frac{2}{3}\rho\chi_m[1 + c_0(n)]}{1 - \frac{1}{3}\rho\chi_m[1 + c_2(n)]}. \qquad (8.13)$$

The result (8.12) reduces to the Clausius–Mossotti formula for ϵ when $c_0 = c_2 = 0$, and (8.13) then reduces to the analogous formula for μ. In fact, even without χ_m or χ_e being zero, (8.10) and (8.11) become the Clausius–Mossotti results when the $c_m = 0$.

This occurs when the particle positions are uncorrelated and also in the static case. Thus (8.12) and (8.13) may be considered to be generalizations of the Clausius–Mossotti result from the static case to finite frequencies or from uncorrelated to correlated particle positions. To obtain an improvement over the Clausius–Mossotti result in the static case $k_0 = 0$, we must include the terms of order α^2 in (5.8), which involve $g^{(3)}$ as well as $g^{(2)}$. By doing so and assuming a scalar electric polarizability and zero magnetic polarizability, we obtain the result of Yvon,[1] which has also been derived by Green,[3] Brown,[4–6] and de Boer et al.[7]

All the results in Secs. 7 and 8 could have been obtained directly from (4.1) by setting $\langle\pi_j\rangle_{ij} = \langle\pi_j\rangle_j$ and seeking plane wave solutions of the resulting equation. This is essentially the procedure followed by Hoek[14] and Rosenfeld[12] in deriving (8.12). The advantage of our method of derivation is that it permits us to obtain better results by keeping quadratic and higher powers of the polarizability α in the dispersion equation (5.8). It also shows that the results (8.12) and (8.13) are valid when these higher-order terms in (5.8) are negligible.

APPENDIX. PROOF OF AN IDENTITY

We have asserted that (6.7) is an identity, and we now prove that this is so for the average of (6.7) with $\mathbf{x}_i$ fixed. This proof suffices for our purposes because we used only (6.8), the averaged form of (6.7), in our subsequent derivation. A similar proof can be given for (6.7) itself. Upon averaging (6.7) with $\mathbf{x}_i$ fixed and using (6.9), we obtain

$$\langle\pi_{ii_1\cdots i_{N-1}}^N\rangle_i$$
$$= \sum_{s=1}^N \langle\phi_i^s\rangle_i,$$
$$= \sum_{s=1}^N \binom{N-1}{s-1}\sum_{l=1}^s (-1)^{s+l}\binom{s-1}{l-1}\langle\pi_{ii_1\cdots i_{l-1}}^l\rangle_i. \quad (A1)$$

Upon rearranging the summation and then writing out the binomial coefficients, we obtain

$$\langle\pi_{ii_1\cdots i_{N-1}}^N\rangle_i$$
$$= \sum_{l=1}^N \langle\pi_{ii_1\cdots i_{l-1}}^l\rangle_i \sum_{s=l}^N \binom{N-1}{s-1}(-1)^{s+l}\binom{s-1}{l-1},$$
$$= \sum_{l=1}^N \langle\pi_{ii_1\cdots i_{l-1}}^l\rangle_i \frac{(N-1)!}{(l-1)!}\sum_{s=l}^N \frac{(-1)^{s+l}}{(N-s)!(s-l)!}. \quad (A2)$$

The sum over s can be rewritten and evaluated as

follows:

$$\sum_{s=l}^{N}\frac{(-1)^{s+l}}{(N-s)!\,(s-l)!} = \frac{(-1)^{2l}}{(N-l)!}\sum_{s=l}^{N}\frac{(N-l)!\,(-1)^{s-l}}{(N-s)!\,(s-l)!},$$

$$= \frac{1}{(N-l)!}\sum_{s=l}^{N}(-1)^{s-l}\binom{N-l}{s-l},$$

$$= \frac{1}{(N-l)!}\sum_{k=0}^{N-l}(-1)^{k}\binom{N-l}{k},$$

$$= \delta_{lN}. \qquad (A3)$$

When we substitute the result (A3) into (A2), we obtain

$$\langle \pi_{ii_1\cdots i_{N-1}}^{N}\rangle_i = \sum_{l=1}^{N}\langle \pi_{ii_1\cdots i_{l-1}}^{l}\rangle_i \frac{(N-1)!}{(l-1)!}\delta_{lN},$$

$$= \langle \pi_{ii_1\cdots i_{N-1}}^{N}\rangle_i. \qquad (A4)$$

Thus (A1) and, therefore, (6.8) are proved.

Gauge Field of a Point Charge*

HENDRICUS G. LOOS

*Douglas Advanced Research Laboratories, Huntington Beach, California, and
University of California, Riverside, California*

The problem and treatment of integration ambiguities in the conventionally defined Yang–Mills charges is demonstrated explicitly, using a non-Abelian solution of the Yang–Mills equations for a point charge. The internal holonomy group $\mathcal{H}$ for this solution is noncompact and nonsemisimple, and the solution is not expected to have a direct physical meaning. However, it provides a convenient example showing important and quite unexpected features of gauge theories of the Yang–Mills type, before quantization. It is found that the number of unambiguously definable and comparable charges is less than the dimension of $\mathcal{H}$ and less than the rank of $\mathcal{H}$ as well. If a gauge group $\mathcal{G}$ is present in the conventional manner, i.e., $\mathcal{H} \subseteq \mathcal{G}$, this number of charges is less than the rank of the gauge group. Other interesting features of the solution found are: discreteness of certain components of the gauge field, as a result of regularity conditions together with the condition that the Yang–Mills charge density vanishes outside a sphere of finite radius, and a harmonic oscillation of the other gauge components, while the observable charges are steady. Higher-order charges are all found to be zero. No action principle is used and no *a priori* particle fields are introduced. Use is made of the differential-geometric properties of gauge fields.

I. INTRODUCTION

IN gauge theories of the Yang–Mills type,[1] the *raison d'être* for the gauge potentials is to define a gauge-covariant derivative for internal nonscalar quantities. More precisely, the gauge potentials define an "internal" linear connection, i.e., a specification of what is meant by equivalence of internal vectors at neighboring events. One may choose this internal linear connection *a priori*, once and for all, independent of the details of the physical situation, and it is then only reasonable to choose it as integrable.[2] Then a theory like special relativity results (where the external connection is fixed and integrable), and since the amount of nonintegrability of the linear connection is expressed by the gauge fields,[3] the latter vanish in this case.

The alternate choice, made here and in all Yang–Mills-type gauge theories, is to let the internal connection be somehow dependent on the details of the physics, as in general relativity. Another way of putting this is that the internal connection carries part of the physics (as the Christoffel symbols do in general relativity). Then, in a local theory, one generally expects the internal connection to be nonintegrable and, hence, the gauge fields to be nonvanishing. Since the internal connection expresses a basic relation between internal spaces belonging to adjacent events,[4] it should be considered fundamental. If it truly can carry some of the physics as well, then one is tempted to regard the internal connection as part of the substratum which gives rise to particles and interactions.

Whether or not the internal connection can carry part of the physics is a question which preferably

---

[1] C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).
[2] A linear connection is called integrable if the result of an equivalence displacement of any vector to any event is independent of the displacement path.
[3] The $F_{k\lambda}$ of Yang–Mills (Ref. 1).

---

[4] A separate internal space is assigned to each event in order to account for fields of internal vectors, tensors, etc., over event space.

follows:

$$\sum_{s=l}^{N} \frac{(-1)^{s+l}}{(N-s)!\,(s-l)!} = \frac{(-1)^{2l}}{(N-l)!} \sum_{s=l}^{N} \frac{(N-l)!\,(-1)^{s-l}}{(N-s)!\,(s-l)!},$$

$$= \frac{1}{(N-l)!} \sum_{s=l}^{N} (-1)^{s-l} \binom{N-l}{s-l},$$

$$= \frac{1}{(N-l)!} \sum_{k=0}^{N-l} (-1)^{k} \binom{N-l}{k},$$

$$= \delta_{lN}. \tag{A3}$$

When we substitute the result (A3) into (A2), we obtain

$$\langle \pi^{N}_{ii_1 \cdots i_{N-1}} \rangle_i = \sum_{l=1}^{N} \langle \pi^{l}_{ii_1 \cdots i_{l-1}} \rangle_i \frac{(N-1)!}{(l-1)!} \delta_{lN},$$

$$= \langle \pi^{N}_{ii_1 \cdots i_{N-1}} \rangle_i. \tag{A4}$$

Thus (A1) and, therefore, (6.8) are proved.

Gauge Field of a Point Charge*

Hendricus G. Loos

Douglas Advanced Research Laboratories, Huntington Beach, California, and
University of California, Riverside, California

The problem and treatment of integration ambiguities in the conventionally defined Yang–Mills charges is demonstrated explicitly, using a non-Abelian solution of the Yang–Mills equations for a point charge. The internal holonomy group $\mathcal{H}$ for this solution is noncompact and nonsemisimple, and the solution is not expected to have a direct physical meaning. However, it provides a convenient example showing important and quite unexpected features of gauge theories of the Yang–Mills type, before quantization. It is found that the number of unambiguously definable and comparable charges is less than the dimension of $\mathcal{H}$ and less than the rank of $\mathcal{H}$ as well. If a gauge group $\mathcal{G}$ is present in the conventional manner, i.e., $\mathcal{H} \subseteq \mathcal{G}$, this number of charges is less than the rank of the gauge group. Other interesting features of the solution found are: discreteness of certain components of the gauge field, as a result of regularity conditions together with the condition that the Yang–Mills charge density vanishes outside a sphere of finite radius, and a harmonic oscillation of the other gauge components, while the observable charges are steady. Higher-order charges are all found to be zero. No action principle is used and no *a priori* particle fields are introduced. Use is made of the differential-geometric properties of gauge fields.

I. INTRODUCTION

IN gauge theories of the Yang–Mills type,[1] the *raison d'être* for the gauge potentials is to define a gauge-covariant derivative for internal nonscalar quantities. More precisely, the gauge potentials define an "internal" linear connection, i.e., a specification of what is meant by equivalence of internal vectors at neighboring events. One may choose this internal linear connection *a priori*, once and for all, independent of the details of the physical situation, and it is then only reasonable to choose it as integrable.[2] Then a theory like special relativity results (where the external connection is fixed and integrable), and since the amount of nonintegrability of the linear connection is expressed by the gauge fields,[3] the latter vanish in this case.

The alternate choice, made here and in all Yang–Mills-type gauge theories, is to let the internal connection be somehow dependent on the details of the physics, as in general relativity. Another way of putting this is that the internal connection carries part of the physics (as the Christoffel symbols do in general relativity). Then, in a local theory, one generally expects the internal connection to be nonintegrable and, hence, the gauge fields to be nonvanishing. Since the internal connection expresses a basic relation between internal spaces belonging to adjacent events,[4] it should be considered fundamental. If it truly can carry some of the physics as well, then one is tempted to regard the internal connection as part of the substratum which gives rise to particles and interactions.

Whether or not the internal connection can carry part of the physics is a question which preferably

---

[1] C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).
[2] A linear connection is called integrable if the result of an equivalence displacement of any vector to any event is independent of the displacement path.
[3] The $F_{k\lambda}$ of Yang–Mills (Ref. 1).

---

[4] A separate internal space is assigned to each event in order to account for fields of internal vectors, tensors, etc., over event space.

should be approached by investigating the consequences of the internal connection rather fully, in a fashion which is, as much as possible, independent of the contemporary theoretical framework of particle physics. Along these lines, much can be done without an action principle and without *a priori* introduced particle fields, and it may not be necessary to interpret the gauge fields as vector mesons. It has been shown[5] that the internal connection by itself already gives rise to a number of internal particle labels, whose properties of definability, comparability, and conservation depend on the internal holonomy group $\mathcal{JC}$. This group is defined as the result of equivalence transport of internal vectors around closed loops in event space. The covariant divergence-free currents which give rise to the additive internal particle labels[6] belong to the Lie algebra or the enveloping algebra of $\mathcal{JC}$, so that $\mathcal{JC}$ is the relevant group for the algebra of the currents or charges.

It has been remarked[5] that the conventional procedure of constructing total charges in a Yang–Mills-type gauge theory may not always have physical significance, on account of the ambiguity arising from integration of nonscalar internal quantities. In the present paper we show a solution to the Yang–Mills equations for a point charge, which is non-Abelian and nonspherically symmetric, in contradistinction to the Abelian and spherically symmetric solutions obtained by Ikeda and Miyachi[7] for $O(3)$, and by the writer for any gauge group.[8] This solution is used as an example, showing that the difficulties associated to the integration ambiguity can actually occur. The method of the measuring operators[5] is applied to this case, and it is found that the number of definable and comparable charges is less than the dimension[9] of the internal holonomy group $\mathcal{JC}$, and is even less than the rank of $\mathcal{JC}$. For a gauge group equal to $\mathcal{JC}$ this result is in sharp contrast with that of the conventional procedure, where before quantization there are as many charges as the dimension of the gauge group, while after quantization the number of mutually assignable (i.e., commuting) charges is equal to the rank of the gauge group.

II. GAUGE FIELDS

In order to describe multiplet fields over event space as fields of internal vectors or tensors, we assign to every event x^k, $k = 0, 1, 2, 3$, a separate n-dimen-

sional complex linear vector space (called the internal space at x^k); the components of the multiplet are taken as the components of the vector or tensor in internal space. Base transformations in internal space are often called gauge transformations. In order to make differentiation of an internal vector field $v(x^k)$ an invariant process, a (gauge-)covariant derivative is defined as

$$\nabla_k v = \partial_k v - \Gamma_k v, \qquad (1)$$

where $\partial_k = \partial/\partial x^k$, and $\Gamma_k(x^\lambda)$ are four matrix fields which represent the internal connection. In (1), v is a covariant internal vector; for a contravariant internal vector w and for a mixed internal tensor P one has

$$\nabla_k w = \partial_k w + w \Gamma_k, \qquad (2)$$

$$\nabla_k P = \partial_k P - [\Gamma_k, P]. \qquad (3)$$

Internal vectors $v(x^k)$ and $v(x^k + dx^k)$ are called *equivalent* if at x^k,

$$dx^k \nabla_k v = 0. \qquad (4)$$

If equivalence displacement of internal vectors to a remote event depends on the displacement path, then the internal connection is called nonintegrable. The extent of nonintegrability is expressed by "internal curvature" tensor operator

$$\phi_{k\lambda} = \partial_k \Gamma_\lambda - \partial_\lambda \Gamma_k - [\Gamma_k, \Gamma_\lambda]. \qquad (5)$$

The nonintegrability of the internal connection in the large is expressed by the internal holonomy group $\mathcal{JC}$, which is defined as the result of equivalence displacement of internal vectors around closed loops in event space.[10] $\mathcal{JC}$ is a Lie group $\subseteq GL(n, c)$; for analytic internal connections the Lie algebra of $\mathcal{JC}$ is spanned by $\phi_{k\lambda}$ and its covariant derivatives of all orders.[11] Hence, the current density operator

$$\tilde{J}^k = \tilde{g} \nabla_\lambda \phi^{k\lambda} \qquad (6)$$

belongs to the Lie algebra of $\mathcal{JC}$; $\tilde{g} = |\text{Det } g_{k\lambda}|^{\frac{1}{2}}$, where $g_{k\lambda}$ is the metric tensor in event space. The identity

$$\nabla_k \tilde{J}^k = 0 \qquad (7)$$

can be proved easily. Equations (6) are the generalized Yang–Mills equations.[12] $\mathcal{JC}$ may be any finite-dimensional Lie group; in the original work of Yang and Mills,[1] $\mathcal{JC} \subseteq O(3)$. Throughout this paper, ∇_k is

[5] H. G. Loos, Ann. Phys. (N.Y.) **36**, 486 (1966).
[6] Spin is here not considered an internal label.
[7] M. Ikeda and Y. Miyachi, Progr. Theoret. Phys. (Kyoto) **27**, 474 (1962).
[8] H. G. Loos, Nucl. Phys. **72**, 677 (1965).
[9] No quantization is considered.

[10] The event space is taken as simply connected.
[11] This has been shown for "external" holonomy groups by A. Nyenhuys, Koninkl. Ned. Akad. Wetenschap, Proc. Ser. A **56**, 233 (1953); **57**, 17 (1954).
[12] The gauge fields are the coefficients of expansion of the $\phi_{k\lambda}$ in terms of a basis of the Lie algebra of $\mathcal{JC}$ or of a gauge group $\mathcal{G} \supseteq \mathcal{JC}$. Sometimes the $\phi_{k\lambda}$ themselves are called gauge fields; "internal curvature" and "gauge fields" are then synonymous.

covariant under event-space coordinate transformations as well as under gauge transformations; hence Christoffel symbols may enter (6). Both (6) and (7) hold for the Riemannian event space of general relativity, although in this paper the event space is taken as flat.

III. NON-ABELIAN GAUGE FIELD FOR A POINT CHARGE

Let t, r, θ, φ be inertial spherical coordinates in Minkowskian event space. Consider the internal connection

$$\mathbf{\Gamma}_t = [f(\theta)/r]\mathbf{A}, \quad \mathbf{\Gamma}_r = 0, \quad \mathbf{\Gamma}_\theta = 0, \quad \mathbf{\Gamma}_\varphi = h(\theta)\mathbf{B}, \tag{8}$$

where f and h are real functions of θ only, and $\mathbf{A}$ and $\mathbf{B}$ are constant matrices. With (5) one finds for the internal curvature operator

$$\begin{aligned}
\mathbf{\phi}_{tr} &= (f/r^2)\mathbf{A}, & \mathbf{\phi}_{t\theta} &= -(f'/r)\mathbf{A}, \\
\mathbf{\phi}_{t\varphi} &= -(fh/r)[\mathbf{A}, \mathbf{B}], & \mathbf{\phi}_{\varphi r} &= 0, \\
\mathbf{\phi}_{\varphi\theta} &= -h'\mathbf{B}, & \mathbf{\phi}_{r\theta} &= 0,
\end{aligned} \tag{9}$$

and the Yang-Mills current (6) becomes

$$\begin{aligned}
\mathbf{J}^t &= (r^3)^{-1}(f''\mathbf{A} + f'\mathbf{A} \cot\theta \\
&\quad - (fh^2/\sin^2\theta)[\mathbf{B}, [\mathbf{A}, \mathbf{B}]]), \\
\mathbf{J}^r &= 0, \quad \mathbf{J}^\theta = 0, \\
\mathbf{J}^\varphi &= -(r^4 \sin^2\theta)^{-1}(h''\mathbf{B} - h'\mathbf{B}\cot\theta \\
&\quad - f^2h[\mathbf{A}, [\mathbf{A}, \mathbf{B}]]).
\end{aligned} \tag{10}$$

For a point charge at $r = 0$, $\mathbf{J}^k$ must vanish for $r \neq 0$ and (10) becomes a system of nonlinear partial differential equations for the functions $f(\theta)$ and $h(\theta)$. On physical grounds we require that the matrix elements of the internal holonomy group elements belonging to infinitesimal loops located at constant r are bounded numbers times the area of the loop. This implies that the functions

$$f, f', fh/\sin\theta, \quad \text{and} \quad h'/\sin\theta \quad \text{are bounded on}$$
$$0 < \theta < \pi; \tag{11}$$

we call a solution $f(\theta)$, $h(\theta)$ satisfying (11) a regular solution. The trivial solution $f(\theta) = 0$, $h(\theta) = 0$ is ignored.

For a point charge, Eqs. (10) imply

$$\begin{aligned}
[\mathbf{B}, [\mathbf{A}, \mathbf{B}]] &= \alpha\mathbf{A}, \\
[[\mathbf{A}, \mathbf{B}], \mathbf{A}] &= \beta\mathbf{B},
\end{aligned} \tag{12}$$

where α and β are real numbers. Equations (12) and (8) show that the internal holonomy group $\mathcal{K}$ has a Lie algebra spanned by the operators $\mathbf{A}$, $\mathbf{B}$, and $[\mathbf{A}, \mathbf{B}]$. It is convenient to change to the independent variable

$z = \cos\theta$; for a point charge, Eqs. (10) become

$$\ddot{f} - \frac{2z}{1 - z^2}\dot{f} - \frac{\alpha fh^2}{(1 - z^2)^2} = 0, \tag{13a}$$

$$\ddot{h} + \frac{\beta f^2 h}{1 - z^2} = 0, \tag{13b}$$

where the dot denotes differentiation with respect to z. In terms of the variable z, the regularity conditions (11) read

$$f, f(1 - z^2)^{\frac{1}{2}}, fh(1 - z^2)^{-\frac{1}{2}}, \quad \text{and} \quad h \quad \text{are bounded on}$$
$$-1 < z < 1. \tag{14}$$

The operators $\mathbf{A}$ and $\mathbf{B}$ can be normalized such that $|\alpha|$ and $|\beta|$ become unity if they do not vanish. Then there are the possibilities: Case I, $\alpha = 0$, $|\beta| = 1$; Case II, $|\alpha| = 1$, $|\beta| = 1$; Case III, $|\alpha| = 1$, $\beta = 0$; and Case IV, $\alpha = 0$, $\beta = 0$.

Case I: $\alpha = 0$, $|\beta| = 1$

The only regular solution of (13a) is $f(z) = \text{const}$. For $\beta = -1$, (13b) has a regular nonvanishing solution only if f vanishes; this solution is $h = h_0 + h_1 z$, with constant h_0 and h_1. For $\beta = 1$, (13b) has regular solutions only if

$$f^2 = m(m + 1), \quad m \geq 0, \text{ an integer}; \tag{15}$$

these solutions are

$$h(z) = (1 - z^2)(d/dz)P_m(z), \tag{16}$$

where $P_m(z)$ is the Legendre polynomial of degree m. According to (12) with $\alpha = 0$, $\beta = 1$, $\mathbf{B}$ and $[\mathbf{A}, \mathbf{B}]$ generate an Abelian invariant subgroup of $\mathcal{K}$. It can be shown from (12) that one has $\text{Tr } \mathbf{B}^k = 0$ for any positive integer k, and it follows that $\mathbf{B}$ is nilpotent. Hence, $\exp \omega\mathbf{B}$ is a polynomial, so that $\mathcal{K}$ is noncompact. According to (12) one has

$$\begin{aligned}
\mathbf{A} &= \mathbf{L}_1 + a_2\mathbf{L}_2 + a_3\mathbf{L}_3, \\
\mathbf{B} &= b_2\mathbf{L}_2 + b_3\mathbf{L}_3,
\end{aligned} \tag{17}$$

where $\mathbf{L}_1$, $\mathbf{L}_2$, and $\mathbf{L}_3$ are the generators of $\mathcal{K}$, with the commutation relations

$$\begin{aligned}
[\mathbf{L}_1, \mathbf{L}_2] &= \mathbf{L}_3, \\
[\mathbf{L}_2, \mathbf{L}_3] &= 0, \\
[\mathbf{L}_3, \mathbf{L}_1] &= \mathbf{L}_2.
\end{aligned} \tag{18}$$

Since the Lie algebra is taken over the real numbers, a_2, a_3, b_2, and b_3 are real. If the internal space is a two-dimensional complex linear vector space, there always exists a choice of internal base for which

$$\mathbf{L}_1 = \pm\begin{pmatrix} \gamma + i & 0 \\ 0 & \gamma \end{pmatrix}, \quad \mathbf{L}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$\mathbf{L}_3 = \pm\begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \tag{19}$$

where γ is a complex number. The representation (19) has an invariant subspace, but due to the non-compactness it is not completely reducible. There are no irreducible two-dimensional representations of the Lie algebra of $\mathcal{H}$.

Case II: $|\alpha| = 1$, $|\beta| = 1$

Inspection of the differential equation (13b) near the singularities shows that for any nontrivial regular solution $h(z)$, one must have $h(\pm 1) = 0$, $\dot{h}(\pm 1) \neq 0$. On account of the relative sign of $\dot{h}$ and h, these conditions can be met only if β is positive; hence, $\beta = 1$, unless $h(z) = 0$. Using the information about the behavior of the nontrivial regular solution $h(z)$ near the points $z = \pm 1$, inspection of (13a) reveals that for any nontrivial regular solution $f(z)$, one must have $f(\pm 1) \neq 0$ and $\dot{f}(\pm 1) = 0$. These conditions cannot be met if α is positive, since there must be at least one point z, $-1 < z < 1$, with $\dot{f} = 0$, and there $\ddot{f}$ and f would have the same sign. Hence, $\alpha = -1$ is the only remaining possibility. We have not found any solutions for the resulting nonlinear differential equations (13); however, one would expect solutions to exist and to form a discrete set. With $\alpha = -1$, $\beta = 1$, $\mathcal{H} = O(2, 1)$ [i.e., the noncompact modification of $O(3)$].

Case III: $|\alpha| = 1$, $\beta = 0$

Then, $h(z) = h_0 + h_1 z$ with constant h_0 and h_1. For both constants vanishing, Case II, one finds that $h = 0$. For the remaining cases, h cannot vanish at both singularities $z = \pm 1$ of Eq. (13a). For the singularity at which $h \neq 0$, the indicial equation has imaginary roots for negative α. Since $f(z)$ must be real, one must have $\alpha = 1$. At the singularity with $h \neq 0$, the expansion of a nonzero $f(z)$ has the leading term $f_1(z \pm 1)^{\frac{1}{2}|h(\mp 1)|}$ with constant f_1; the sign must be chosen according to which singularity is under discussion. If h vanishes at one of the singularities, one has, at that point, $f \neq 0$ and $\dot{f} = 0$ for any regular nonzero solution $f(z)$. For $\alpha = 1$, $\ddot{f}$ and f have the same sign at points where $\dot{f} = 0$. This is in contradiction with the behavior of $f(z)$ near the singularities. Hence, there are no regular solutions for Case III, except the semitrivial ones with $h(z) = 0$, $f(z) = $ const, and $h(z) = h_0 + h_1 z$, $f(z) = 0$.

Case IV: $\alpha = 0$, $\beta = 0$

The only regular solutions of (13) are $f = $ const, $h = h_0 + h_1 z$, with constant h_0 and h_1.

IV. CHARGES ASSOCIATED WITH THE YANG–MILLS CURRENT

In the conventional procedure[1] the total charge associated with the gauge fields amounts to the flux

$$\mathbf{Q} = \frac{i}{2} \int_R \tilde{g} \boldsymbol{\phi}^{k\lambda} \, d\tilde{f}_{k\lambda}, \qquad (20)$$

where R is the set of events happening at the same time on a large spherical surface enclosing the physical system, and

$$d\tilde{f}_{k\lambda} = \tfrac{1}{2} \tilde{\eta}_{k\lambda\mu\nu} \, df^{\mu\nu}, \qquad (21)$$

where $df^{\mu\nu}$ is a surface element on R, and $\tilde{\eta}_{k\lambda\mu\nu}$ is the totally antisymmetric tensor Δ-density[13] of weight 1, for which $\tilde{\eta}_{0123} = 1$ in every allowable coordinate system. For inertial spherical coordinates in a Minkowskian event space, and R chosen as the set of events with r and t constant, (20) becomes

$$\mathbf{Q} = -i \int \boldsymbol{\phi}_{tr} r^2 \sin\theta \, d\theta \, d\varphi. \qquad (22)$$

This integral involves summation of internal operators at different events on R, *a process which is generally not gauge-invariant*. One could make it so by stipulating that the internal operators $\boldsymbol{\phi}_{tr} r^2 \sin\theta \, d\theta \, d\varphi$ are to be equivalence-displaced to a common collection event x_1^k on R, prior to integration. However, the result of such a procedure depends on the choice of collection point and displacement paths, unless $\boldsymbol{\phi}_{tr}$ commutes with that part of the internal holonomy group $\mathcal{H}$, belonging to loops on the event set R. This difficulty has been discussed before in general terms,[5] but now we have an explicit example. For the non-Abelian solutions of the point-charge Yang–Mills equations derived in Sec. III, one has

$$[\boldsymbol{\phi}_{tr}, \boldsymbol{\phi}_{\varphi\theta}] \neq 0. \qquad (23)$$

Hence, not all components of the charge (22) can be defined unambiguously. Of course, even in this case, one can formally evaluate expression (20) for a gauge chosen such that

$$\int_S \partial_\lambda (\tilde{g} \boldsymbol{\phi}^{k\lambda}) \, d\tilde{f}_k = 0, \qquad (24)$$

where S is the set of events happening between times t_1 and t_2 on the large spherical surface enclosing the physical system; then one surely has conservation of the charge (20), but this amounts to fixing the succession of gauges in time in such a manner that the charge comes out the same at all times.

---

[13] A Δ-density of weight w is a one-component quantity which, under a coordinate transformation with Jacobian Δ, acquires the factor Δ^{-w}.

An alternate method, which is gauge-invariant and more likely to produce physically meaningful results, has been proposed[5] which uses the concept of "measuring operator"; instead of (20) we introduce definable charges[5]

$$q = \frac{i}{2} \int_R \tilde{g} \, \text{Tr} \, (\Phi^{k\lambda} C) \, df_{k\lambda}, \qquad (25)$$

where Tr denotes the trace, and the internal operator C is subject to the condition

$$dx^k \nabla_k C = 0 \quad \text{on } R, \qquad (26)$$

for every dx^k in R. Equation (26) states that the internal operator C is covariant uniform over R; it is called a measuring operator on R. Its use in (25) is to reduce the integration to that of scalars. Different choices for C satisfying (26) may give different components q of the charge (hypercharge, isospin projection, baryon number). Equation (26) implies that C commutes with the part of $\mathcal{K}$ belonging to loops on R. For the choice of hypersurface R and coordinates as in (22), this amounts to the condition that C commutes with $\Phi_{\theta\varphi}$, $\nabla_\theta \Phi_{\theta\varphi}$, and $\nabla_\varphi \Phi_{\theta\varphi}$. We separately consider the measuring operators and total charges first for the non-Abelian solutions for Cases I and II, and thereafter for the Abelian solutions.

Case I

For a two-dimensional complex internal space, and for the internal base which gives (19), any internal operator C which commutes with $\Phi_{\theta\varphi}$ is represented by a linear combination of

$$C_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad (27)$$

both of these matrices commute with $\nabla_\theta \Phi_{\theta\varphi}$ and $\nabla_\varphi \Phi_{\theta\varphi}$ as well, and satisfy (26) on the instantaneous spherical surface R. The definable charges are

$$q_1 = -4\pi i f \, \text{Tr} \, (A C_1) = 0,$$
$$q_2 = -4\pi i f \, \text{Tr} \, (A C_2) = \pm 4\pi (2\gamma i - 1)[m(m+1)]^{\frac{1}{2}},$$
$$(28)$$

where $m \geq 0$ is an integer; use has been made of (17), (19), and (15). Note that for $\gamma \neq -i/2$, C_2 lies outside the enveloping algebra of $\mathcal{K}$. Although $\mathcal{K}$ is three-dimensional, there is only *one* generally nonvanishing definable charge. For fixed γ this charge is restricted to discrete values; but since γ can be any number, the possible values for q_2 form a continuum.

For comparability[5] of charges at in and out states, one must moreover have

$$\nabla_t C = 0 \quad \text{on } S. \qquad (29)$$

The integrability conditions for (29),

$$[\Phi_{\theta t}, C] = 0,$$
$$[\Phi_{\varphi t}, C] = 0, \qquad (30)$$

on S, are satisfied for the matrices (27). However, (29) requires that

$$\partial_t C = [\Gamma_t, C],$$
$$= [m(m+1)]^{\frac{1}{2}}/r^{-1}[L_1 + a_2 L_2 + a_3 L_3, C] \quad (31)$$

on S. For C_2 of (27) this condition is satisfied, but C_1 of (27) must be given the factor $\exp it[m(m+1)]^{\frac{1}{2}}/r$. Since C_1 gives a vanishing charge (28), this factor does not show up in the results. We see that the charge q_2 is comparable at in and out states, and since it is associated with a nonderivative current, it is conserved as well.[5]

Besides the charges (25), one must consider the dual charges

$$q^* = \frac{i}{2} \int_R \text{Tr} \, (\Phi_{k\lambda} C) \, df^{k\lambda}. \qquad (32)$$

In the electromagnetic case, q^* is the magnetic charge. With (17) and (19), one finds $q^* = 0$ for both measuring operators (27).

At this point, a remark is due about the different representations of the Lie algebra (18) of $\mathcal{K}$. Taking traces generally is a representation-dependent process, and moreover, in higher dimensional representations more measuring operators C may exist. Hence, the number and values of definable charges may depend on the representation. This is to be expected if one realizes that different representations of the Lie algebra (18) belong to physically different cases. The dimension of internal space is physically important, as is the dimension of event space, which must be specified in addition to the abstract Lorentz group. Inequivalent representations of the same dimension stand for different physical situations as well: to loops in event space correspond two different sets of internal linear transformations which cannot be identified by executing a base transformation. We have calculated through the case of a three-dimensional complex internal space. There are two inequivalent representations, B is nilpotent of index 2 in one, and of index 3 in the other. In the first case, there exist 3 independent generally-nonvanishing definable charges; but among these only one is comparable between in and out states. For the second representation ($B^3 = 0$) there is only one definable charge, which is also comparable. The dual charge (32) vanishes regardless of the representation because h vanishes at $\theta = 0$ and $\theta = \pi$, on account of (16).

Case II

If non-Abelian solutions for this case exist, they must belong to $O(2, 1)$. However, for these solutions there are no definable nonvanishing charges at all for any finite-dimensional representation. This is due to the fact that, for an internal operator C which commutes with B, Tr (AC) must vanish: from (12) one has

$$\text{Tr } (AC) = -\text{Tr } (B[A, B]C) + \text{Tr } ([A, B]BC),$$
$$= -\text{Tr } (B[A, B]C) + \text{Tr } ([A, B]CB) = 0.$$
$$(33)$$

The dual charges vanish because $h = 0$ at $\theta = 0$ and $\theta = \pi$.

Semi-Trivial Abelian Solutions

For the first solution, $h = 0$, $f = $ const, $\phi_{tr} = fA/r^2$ is the only nonvanishing component of $\phi_{k\lambda}$, and $\mathcal{K}$ is a one-parameter group. This solution is essentially the same as that of Ikeda and Miyachi[7] and a subsequent generalization[8] of their work. Since $\phi_{\theta\varphi}$, $\phi_{\theta t}$, and $\phi_{\varphi t}$ vanish, there is no algebraic restriction on the measuring operator; there are as many definable and comparable charges as the dimension of $\mathcal{K}$ (i.e., one).

For the second solution one has $f = 0$, $h = h_0 + h_1 z$ with constant h_0 and h_1; the only nonvanishing internal curvature tensor component is $\phi_{\varphi\theta} = h_1 B \sin \theta$. Since $\phi_{tr} = 0$, there is only a dual charge; it is definable and comparable, and it is restricted to discrete values by an argument[8] similar to that of Dirac for magnetic monopoles.[14]

The third Abelian solution, $f = $ const, $h = h_0 + h_1 z$, is a linear combination of the first two.

V. CHARGES ASSOCIATED WITH HIGHER-ORDER CURRENTS

The current density (6) is not the only one which can be constructed from the internal curvature and which is covariant divergence-free. It has been shown that any scalar density (or Δ-density) of unit weight

$$\tilde{\mathcal{L}}(g_{k\lambda}, \phi_{\mu\nu}, \nabla_\rho \phi_{\mu\nu}, \cdots),$$

which can be constructed from the metric tensor, the internal curvature, and its covariant derivatives to some order, can be used to find a covariant divergence-free current density.[5] If $\tilde{\mathcal{L}}$ does not involve covariant derivatives of $\phi_{k\lambda}$, the resulting current is called *nonderivative*; it is then simply

$$\tilde{J}^k = \nabla_\lambda \tilde{P}^{[k\lambda]},$$
$$(34)$$

where

$$\tilde{P}^{k\lambda} = \frac{\partial \tilde{\mathcal{L}}}{\partial \phi_{k\lambda}}.$$
$$(35)$$

For the nonderivative case, one has for $\tilde{\mathcal{L}}$ the possibilities

$$\tilde{\mathcal{L}}_1 = \tfrac{1}{2}\tilde{g} \text{ Tr } (\phi_k^\lambda \phi_\lambda^k), \quad \tilde{\mathcal{L}}_2 = \tfrac{1}{3}\tilde{g} \text{ Tr } (\phi_k^\lambda \phi_\lambda^\mu \phi_\mu^k), \text{ etc.,} \quad (36)$$

and those obtained by replacing a number of $\phi_{\mu\nu}$'s in these traces by their duals

$$\phi_{k\lambda}^* = \tfrac{1}{2}\tilde{g}\tilde{\eta}_{k\lambda\mu\nu}\phi^{\mu\nu},$$
$$(37)$$

and further, those densities obtained from all the foregoing ones by commutation of factors. The Yang–Mills current (6) is derivable from $\tilde{\mathcal{L}}_1$ of (36). We briefly consider the charges associated with the "higher-order currents" $\tilde{J}_2^k$, etc., derived from the scalar densities mentioned, for the internal curvature fields obtained in Sec. III. Instead of the charge (25), one gets, for inertial spherical coordinates and for R the instantaneous events on the sphere,

$$q = i \int \text{Tr } (P^{tr}C)r^2 \sin \theta \, d\theta \, d\varphi.$$
$$(38)$$

Since for the internal curvature field with a pointlike Yang–Mills charge, some of the higher-order charges may be distributed, the limit $r \to \infty$ of (38) has to be taken in order to capture the whole charge. P^{tr} must be of order $O(r^{-2})$ if this charge is nonvanishing. Inspection shows that for $\tilde{\mathcal{L}}_2$, $\tilde{\mathcal{L}}_3$, etc., of (36), P^{tr} is of order $O(r^{-l})$ with $l > 2$, and the same is true when some of the $\phi_{k\lambda}$'s are replaced by their duals, or when factors are commuted. Hence, for the solutions obtained in Sec. III, the higher-order nonderivative currents do not give rise to charges.

Of the derivative currents, we have only investigated the current[15] derivable from

$$\tilde{\mathcal{L}} = \tilde{g} \text{ Tr } ((\nabla_k \phi_{\mu\lambda})\nabla^k \phi^{\mu\lambda})$$

for Case I and a two-dimensional complex internal space, and we found zero charges.

VI. CAN THE CONDITIONS ON THE MEASURING OPERATOR BE RELAXED?

Conditions (26) and (29) for the operator C are not the only ones that lead to conserved charges associated with the Yang–Mills current, or with other nonderivative currents (34). All one needs is an internal operator field C on S such that

$$0 = \int_S \text{Tr } \nabla_\lambda (\tilde{P}^{[k\lambda]}C) \, d\tilde{f}_k;$$
$$(39)$$

---

[14] P. A. M. Dirac, Phys. Rev. 74, 817 (1948).

[15] For derivative currents, the relation between $\tilde{J}^k$ and $\tilde{\mathcal{L}}$ is different from (34) and (35) (Ref. 5).

then the charge

$$q = \int_R \mathrm{Tr}\,(\tilde{\mathbf{P}}^{k\lambda}\mathbf{C})\,d\tilde{f}_{k\lambda} \qquad (40)$$

is conserved.

One should guard here against too much arbitrariness in the field $\mathbf{C}$ over S; the resulting charge is only useful if it can be evaluated unambiguously for different physical cases. In view of the purpose of introducing internal space in the first place, we want to limit the number of possible different fields $\mathbf{C}$ over S to no more than the number of elements of $GL(n, c)$. This excludes fields $\mathbf{C}$ over S which are not uniquely determined by the operator $\mathbf{C}(x_0^k)$ at one event x_0^k on S. In other words, we demand that the field $\mathbf{C}$ over S is restricted by partial-differential equations which have a unique solution over S, when $\mathbf{C}$ is specified at a single event x_0^k. Moreover, we consider with suspicion prescriptions of continuation of $\mathbf{C}$ over S [starting with a given $\mathbf{C}(x_0^k)$], which do not *always* lead to a unique continuation, i.e., for which uniqueness depends on the physical situation at hand. This is because defining the field $\mathbf{C}$ over S, together with the use of $\mathbf{C}$ in the calculation of the charge, is considered to correspond (classically) to physical measurement of that kind of charge. The reason why we restrain ourselves from dismissing the conditionally unique continuation prescriptions altogether is the lepton situation. No internal labels associated with the strong interactions are assigned to leptons because this would have to be arbitrary and useless; this may be a case of conditional assignability.

With this in mind, let us consider conditions on $\mathbf{C}$ such that (39) is true whenever on S the current density (34) vanishes or vanishes asymptotically for $r \to \infty$. The weakest condition which brings this about is

$$\tilde{\eta}_{k\mu\rho\nu}u^\mu v^\rho w^\nu\,\mathrm{Tr}\,(\tilde{\mathbf{P}}^{[k\lambda]}\nabla_\lambda\mathbf{C}) = 0, \qquad (41)$$

where u^μ, v^ρ, and w^ν are vectors tangent to S. Since this is essentially a single scalar equation for $\mathbf{C}$, it can never serve to continue $\mathbf{C}$ over S uniquely from a given $\mathbf{C}(x_0^k)$. Neither can

$$\mathrm{Tr}\,(\tilde{\mathbf{P}}^{[k\lambda]}\nabla_\lambda\mathbf{C}) = 0, \qquad (42)$$

nor

$$\tilde{\eta}_{k\mu\rho\nu}u^\mu v^\rho w^\nu\tilde{\mathbf{P}}^{[k\lambda]}\nabla_\lambda\mathbf{C} = 0, \qquad (43)$$

but

$$\tilde{\mathbf{P}}^{[k\lambda]}\nabla_\lambda\mathbf{C} = 0 \qquad (44)$$

may suffice if the rank of $\tilde{\mathbf{P}}^{[k\lambda]}$ is sufficiently large. Whether or not this is the case depends on the physical situation; for instance, for the non-Abelian solution of Case I and a two-dimensional complex internal space, (44) does not force unique continuation

of $\mathbf{C}$ over S, from a given initial $\mathbf{C}(x_0^k)$. Hence, the choice (44) for the conditions defining the field $\mathbf{C}$ over S falls in the category of conditionally unambiguous continuations of $\mathbf{C}$. These conditions can only be strengthened further by dropping the transvection with $\tilde{\mathbf{P}}^{[k\lambda]}$ in (44); one then arrives[16] at the condition (26) which *always* forces unique continuation of $\mathbf{C}$ over S, from a chosen initial value $\mathbf{C}(x_0^k)$, at an event x_0^k on S.

VII. PULSATING INTERNAL CURVATURE

Solutions of the type (8) with $\beta \neq 0$ describe a nonsteady internal curvature. This becomes clear by calculating the covariant time derivative of $\boldsymbol{\phi}_{t\varphi}$:

$$\nabla_t\boldsymbol{\phi}_{t\varphi} = \frac{f^2 h}{r^2}\,[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] = -\frac{f^2 h\beta}{r^2}\,\mathbf{B}, \qquad (45)$$

using (8), (9), and (12).

For the non-Abelian solution of Case I and a two-dimensional complex internal space, (45) implies

$$\nabla_t\boldsymbol{\phi}_{t\varphi} = \mp \frac{i}{r}\,[m(m+1)]^{\frac{1}{2}}\boldsymbol{\phi}_{t\varphi}, \qquad (46)$$

which shows that $\boldsymbol{\phi}_{t\varphi}$ is oscillating harmonically with a frequency $r^{-1}[m(m+1)]^{\frac{1}{2}}$. The same thing happens to $\boldsymbol{\phi}_{\varphi\theta}$, but $\boldsymbol{\phi}_{tr}$ and $\boldsymbol{\phi}_{t\theta}$ are steady. This oscillation is not immediately noticeable in (8) or (9), because for this form of the solutions, the internal bases are not covariant-constant along world lines r, θ, φ constant; one could say that the internal bases are oscillating. This point may be clarified further by a gauge transformation such that the new internal bases along world lines r, θ, φ constant are equivalent. This is achieved by an internal base transformation $\mathbf{U}$ which makes $\boldsymbol{\Gamma}_t'$ vanish:

$$0 = \boldsymbol{\Gamma}_t' = \mathbf{U}^{-1}(\boldsymbol{\Gamma}_t\mathbf{U} - \partial_t\mathbf{U}). \qquad (47)$$

This implies with (8) that

$$\partial_t\mathbf{U} = \frac{f}{r}\,\mathbf{A}\mathbf{U}; \qquad (48)$$

a solution is

$$\mathbf{U} = \exp\,(fr^{-1}\mathbf{A}t). \qquad (49)$$

In the new base, the internal curvature tensor operator is expressed as

$$\boldsymbol{\phi}_{k\lambda}' = [\exp\,(-fr^{-1}\mathbf{A}t)]\boldsymbol{\phi}_{k\lambda}[\exp\,(fr^{-1}\mathbf{A}t)]. \qquad (50)$$

For the non-Abelian solution of Case I and with a two-dimensional complex internal space, both $\boldsymbol{\phi}_{t\varphi}$

---

[16] Dropping in (44) the transvection with $\tilde{\mathbf{P}}^{[k\lambda]}$ gives $\nabla_\lambda\mathbf{C} = 0$; but since only derivatives in the hypersurface S are needed, $v^\lambda\nabla_\lambda\mathbf{C} = 0$, v^λ tangent to S, suffices.

and $\boldsymbol{\phi}_{\varphi\theta}$ are proportional to the matrix $\mathbf{B}$; one has

$$[\mathbf{A}, \mathbf{B}] = i\mathbf{B}, \quad f = \pm [m(m+1)]^{\frac{1}{2}},$$

and

$$\left[\exp\left(-\frac{f}{r}\mathbf{A}t\right)\right]\mathbf{B}\left[\exp\left(\frac{f}{r}\mathbf{A}t\right)\right]$$

$$= \mathbf{B}\{\exp[-fr^{-1}t(\mathbf{A}+i)]\}\exp fr^{-1}\mathbf{A}t,$$

$$= \mathbf{B}\exp(-ifr^{-1}t). \qquad (51)$$

Hence,

$$\boldsymbol{\phi}'_{t\varphi} = \boldsymbol{\phi}'_{t\varphi}(0)\exp \mp ir^{-1}t[m(m+1)]^{\frac{1}{2}},$$
$$\boldsymbol{\phi}'_{\varphi\theta} = \boldsymbol{\phi}'_{\varphi\theta}(0)\exp \mp ir^{-1}t[m(m+1)]^{\frac{1}{2}}, \qquad (52)$$

where (0) denotes zero time. It is interesting to note that, whereas the internal curvature is pulsating, the concomitant observables are constant in time.[17] This situation is very much like a stationary state in quantum mechanics.

VIII. DISCUSSION

The main purpose for discussing these non-Abelian solutions of the point-charge Yang–Mills equations is to show that the integration ambiguities for the conventionally defined total charges can really occur, and that the number of unambiguously definable charges can be less than the dimension and the rank of the internal holonomy group, and therefore, less than the dimension and the rank of the gauge group. This result is derived for Yang–Mills-type gauge theory, simply by being careful about integration of nonscalar internal quantities. Investigation of weakened conditions for the definition of measuring operators shows the only candidate to be condition (44); but one can object against it on the grounds that it leads to a conditionally nonambiguous definition of charge. The writer presently does not consider this objection quite strong enough for complete dismissal. It can be expected that the reduction in the number of definable and comparable charges by the internal curvature will survive quantization.

The solution of the Yang–Mills equations used to demonstrate these aspects of internal curvature is probably not of direct physical interest; there are not enough definable and comparable charges. The form of the solution (8) is extremely simple in comparison with more general solutions, and the internal holonomy group is noncompact and, for the non-Abelian

solution of Case I, nonsemisimple as well. However, the features of non-Abelian gauge fields shown here may be expected to occur as well for other more general solutions with different holonomy groups and more charges; they are chiefly consequences of the non-Abelian nature of $\mathcal{K}$.

In the conventional Yang–Mills-type gauge theory, the internal connection is restricted such that $\mathcal{K}$ is a subgroup of the gauge group $\mathcal{G}$, which is defined as the internal symmetry group of the Lagrangian. Here no action principle is used at all, and $\mathcal{G}$ does not occur. If an action principle is introduced, but no *a priori* particle fields, then the Lagrangian must be a function of concomitants of the internal connection and $g_{k\lambda}$. But Lagrangian densities like[18] $\tilde{g}\,\mathrm{Tr}\,(\boldsymbol{\phi}_k^\lambda\,\boldsymbol{\phi}_\lambda^k)$ do not require introduction of an internal metric. $\mathcal{G}$ is then just $GL(n, c)$, a group which is already present as a consequence of choosing the internal space as an n-dimensional complex linear vector space. Hence, in such a case, $\mathcal{G}$ need not be specified either. An internal metric would be required if the Lagrangian were a function of internal vectors, committed by the internal connection, for instance, as internal eigenvectors of $\boldsymbol{\phi}_{k\lambda}$ or $\boldsymbol{\phi}_k^\lambda\boldsymbol{\phi}_\lambda^k$.

The discreteness of the solutions considered here is due to the nonlinearity of the Yang–Mills equations for non-Abelian $\mathcal{K}$ and the regularity conditions (11), together with the restriction of the Yang–Mills charge to a point charge. A charge of finite extent still gives discrete exterior solutions, as long as outside a certain finite sphere the Yang–Mills current density vanishes. The resulting "quantization" of the charge (28) is the wrong one for an additive quantum number, and moreover, continuity is reinstalled by the continuous values which the constant γ can assume. However, it is interesting to find that the classical Yang–Mills equation for a point charge is capable of producing a quantum-mechanical-like effect at all. The same comment applies to the harmonic oscillation found for certain components of the internal curvature, whereas the observables are steady.

---

[17] This is true since these observables are independent of the internal base and an internal base exists for which the $\boldsymbol{\Gamma}_k$ are not dependent on time; in (8) one has such a base.

---

[18] We presently do not wish to enter into a discussion about which Lagrangians are physically appropriate.

Three-Particle States and Green's Function

JEANNE-YVONNE PASQUIER

Laboratoire de Physique Théorique et Hautes Energies, Orsay, France

The six-dimensional spherical harmonics are specified in connection with SO_6 and particular subgroups isomorphic to SO_3 and SO_2; three-particle states are then labeled by the grand-angular momentum n, the usual angular momentum j, its projection on a fixed axis m, an integer μ related to SO_2, and a degeneracy number ω. Also, we derive the plane-wave and free Green's function expansions in terms of these spherical harmonics.

I. INTRODUCTION

OUR purpose here is to study a particular classification of three-particle states. Many of the results we derive are known—in particular, the spherical coordinates and the related three-particle quantum numbers. These coordinates and quantum numbers have generally been introduced according to different methods: group-theoretical considerations,[1,2] algebraic calculations,[3,4] or both.[5] A special class of spherical harmonics in six dimensions (those involved here) has also been investigated before[6]; on the other hand, the use of a symmetrical orthonormal basis of eigenfunctions has been made in scattering[7] or bound-state[3,8] problems in some simple cases (S waves in particular). However, these papers are actually disconnected, in the sense that no connection has been made between these coordinates and quantum numbers and the usual functions of quantum mechanics (for instance, the free Green's function, which is a powerful tool for scattering as well as for bound-state problems). To our knowledge, only group-theoretical considerations are able to furnish easily such a connection; more precisely, only the introduction of the spherical harmonics, labeled with the good physical quantum numbers, i.e., the decomposition of the SO_6 group with respect to the physical subgroups, in particular SO_3, may furnish us an expansion of the free Green's function in terms of these quantum numbers. We must mention that such an "expansion" has been investigated before[7] only in the case of S waves.

In the present paper we shall use group-theoretical tools as much as possible. For the sake of completeness, we introduce in Sec. II the spherical coordinates for the three particles which are closely connected to the invariance subgroups of the problem. A first expansion of the plane wave and of the free Green's function is obtained in Sec. III; this expansion exhibits only the "grand-angular momentum" first introduced by Smith.[5] The spherical harmonics are introduced in Sec. IV as vectors of an irreducible representation of SO_6; in Sec. V and in the Appendices, this representation is reduced with respect to the invariance subgroups of Sec. II. Finally, we write the full expansion of the free Green's function in Sec. IV.

II. GEOMETRY AND INVARIANCES

As mentioned above, we first introduce the spherical coordinates of a three-particle state, insisting particularly on the symmetries and invariances of the state. The three particles are supposed to be in their center-of-mass system, so that the system possesses six degrees of freedom; hence the associated spherical harmonics will be invariant under the group SO_6. The principal purpose of this section is then to look for the subgroups of SO_6 under which the system of three free particles is invariant.

To be clearer and more concise, we study only the case of three particles of unit mass; the generalization to the case of unequal masses is easy but tedious. Here and there, we put between brackets some remarks on the general case abbreviated as u.m. (read unequal masses).

The center-of-mass restriction yields

$$\sum_{i=1}^{3} X_i = \sum_{i=1}^{3} Y_i = \sum_{i=1}^{3} Z_i = 0, \qquad (2.1)$$

where (X_i, Y_i, Z_i) are the Cartesian coordinates of particle i in a given orthonormal coordinate system $\mathbb{C}$; let us call M_i the position of particle i at a given time t and 0 the center-of-gravity of the triangle $M_1 M_2 M_3$. In order to take Eq. (2.1) into account, we introduce a fictitious three-dimensional space $\mathcal{F}$ of Ω origin

[1] A. J. Dragt, J. Math. Phys. 6, 533 (1965).
[2] J. M. Lévy-Leblond and M. Lévy-Nahas, J. Math. Phys. 6, 1571 (1965).
[3] W. Zickendraht, Ann. Phys. (N.Y.) 35, 18 (1965).
[4] Yu. A. Simonov, Yadernaya Fiz. 3, 630 (1966) [English transl.: Soviet J. Nucl. Phys. 3, 461 (1966)].
[5] F. T. Smith, Phys. Rev. 120, 1058 (1960).
[6] M. A. B. Beg and H. Ruegg, J. Math. Phys. 6, 677 (1965).
[7] V. Gallina, P. Nata, L. Bianchi, and G. Viano, Nuovo Cimento 24, 835 (1962).
[8] A. M. Badalyan and Yu. A. Simonov, preprint (Moscow, 1965).

and three points in it, whose Cartesian coordinates are

$$P:(X_1, X_2, X_3),$$
$$Q:(Y_1, Y_2, Y_3), \qquad (2.2)$$
$$R:(Z_1, Z_2, Z_3).$$

[u.m.: for instance, P has coordinates (m_1X_1, m_2X_2, m_3X_3).] Equation (2.1) implies that P, Q, R move in a fixed plane $\mathfrak{T}$ of $\mathfrak{F}$.

Let us notice that the positions of P, Q, R are related to the reference system $\mathcal{C}$. Had we taken another system $\mathcal{C}'$ instead of $\mathcal{C}$, P, Q, R would have been changed into P', Q', R', still in $\mathfrak{T}$, according to

$$\begin{pmatrix} \Omega P' \\ \Omega Q' \\ \Omega R' \end{pmatrix} = \mathcal{R}_r \begin{pmatrix} \Omega P \\ \Omega Q \\ \Omega R \end{pmatrix}, \qquad (2.3)$$

where[9] $\mathcal{R}_r$ is the 3×3 rotation matrix which transforms $\mathcal{C}$ into $\mathcal{C}'$, namely,

$$\begin{pmatrix} X_i' \\ Y_i' \\ Z_i' \end{pmatrix} = \mathcal{R}_r \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}.$$

We now choose for $\mathcal{C}$ a fixed (in time) system of axes $(0X, 0Y, 0Z)$, and $\mathcal{R}_r$ is the rotation which transforms $\mathcal{C}$ into the usual reference system $(0x, 0y, 0z)$ whose axes are the principal axes of inertia of the triangle $M_1M_2M_3$, $0z$ being normal to its plane. According to Eq. (2.2), to the axes $(0x, 0y, 0z)$ correspond three points in $\mathfrak{T}$; let us call the first two A and B, the third one being just Ω. With this particular choice, Eq. (2.3) reads now

$$\begin{pmatrix} \Omega A \\ \Omega B \\ 0 \end{pmatrix} = \mathcal{R}_r \begin{pmatrix} \Omega P \\ \Omega Q \\ \Omega R \end{pmatrix}. \qquad (2.3')$$

In addition, we have $\Omega A \cdot \Omega B = 0$. (u.m.: ΩA and ΩB are two conjugate directions of an ellipse.) Notice also that ΩA^2 and ΩB^2 are the two principal momenta of inertia of the three-particle system.

In the following, what we call the spherical coordinates of a three-particle configuration are defined by

(i) the three Euler angles of $\mathcal{R}_r$ (external coordinates),

(ii) polar coordinates of A and B in $\mathfrak{T}$ with respect to an axis $\Omega \xi$ (internal coordinates).

---

[9] The index r refers to the position of the particles; see index k below.

More precisely, we set

$$\Omega A = r \cos y,$$
$$\Omega B = r \sin y,$$
$$x = \text{polar angle of } A = (\Omega \xi, \Omega A).$$

The ranges of x and y are easily seen to be $[0, \pi]$ and $[0, \tfrac{1}{4}\pi]$, respectively. Indeed, a change of x into $x + \pi$ or y into $y + \tfrac{1}{4}\pi$ induces a new definition of the axis $(0x, 0y, 0z)$ (symmetry with respect to 0 in the first case, exchange of $0x$ and $0y$ in the second case). Notice that $r^2 = r_1^2 + r_2^2 + r_3^2$ where $r_i = 0M_i$.

We may formally write, identifying a two-vector in $\mathfrak{T}$ with the 2×1 column matrix of its components with respect to two perpendicular axes $\Omega \xi$ and $\Omega \eta$,

$$\Omega A = R_2(x) \begin{pmatrix} r \cos y \\ 0 \end{pmatrix}, \qquad (2.4a)$$

$$\Omega B = R_2(x) \begin{pmatrix} 0 \\ r \sin y \end{pmatrix}, \qquad (2.4b)$$

where $R_2(x)$ is the usual 2×2 rotation matrix

$$R_2(x) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}. \qquad (2.5)$$

The third null vector of the left-hand side of Eq. (2.3') may be written as

$$0 = R_2(x) \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \qquad (2.4c)$$

In conclusion, the six coordinates needed to fix a three-particle configuration will be r and five angles (three from $\mathcal{R}_r$, x and y) or six Cartesian coordinates, which are chosen as $(\xi_P, \xi_Q, \xi_R, \eta_P, \eta_Q, \eta_R)$ in this order, where, for instance, (ξ_P, η_P) are the Cartesian coordinates of P in $\mathfrak{T}$ with respect to $(\Omega \xi, \Omega \eta)$. Such a six-vector, or the column matrix of its six Cartesian coordinates, is denoted R. From Eq. (2.3') and (2.4) we have

$$\mathbf{R} = [\mathcal{R}_r^{-1} + \mathcal{R}_r^{-1}][R_2(x) \otimes I_3]\mathbf{R}_0, \qquad (2.6)$$

where each bracket is a 6×6 matrix, $+$ means the direct sum of matrices, $\otimes$ the tensorial product, I_3 is the unit 3×3 matrix, and $\mathbf{R}_0$ has the components $(r \cos y, 0, 0, 0, r \sin y, 0)$. The six lengths of R and $\mathbf{R}_0$ are both r. A caret denotes a unit six-vector: $\mathbf{R} = r\hat{\mathbf{R}}$. Notice also that the 6×6 matrices $[\mathcal{R}_r^{-1} + \mathcal{R}_r^{-1}]$ and $[R_2(x) \otimes I_3]$ are two commutative matrices of SO_6.

Let us notice that the condition $\sum \mathbf{r}_i = 0$ is so far irrelevant. We may as well set $\sum \mathbf{r}_i = \mathbf{R}$; the plane $\mathfrak{T}$

will not be fixed in time, but the coordinates are unchanged.

In the same way, we may define spherical coordinates for the set of the three momenta of the particles ($\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$). The associated six-vector is denoted by K, its six-length by k, the three-dimensional rotation by $\mathfrak{R}_k$; ξ and η correspond to the preceding x and y. We write similarly,

$$K = [\mathfrak{R}_k^{-1} + \mathfrak{R}_k^{-1}][R_2(\xi) \otimes I_3]K_0, \qquad (2.7)$$

where K_0 has the components ($k \cos \eta$, 0, 0, 0, $k \sin \eta$, 0). This parametrization of the momenta does not depend upon the value of the masses. Notice that k^2 is the total energy of the three free particles.

As mentioned in the Introduction, these coordinates are not new. In momentum space, our set is the same as in Refs. 1 and 2, except ξ, which stands for their $\frac{1}{2}\varphi$; the same holds for the position coordinates of Ref. 1. In Ref. 3, the position coordinates are "y," "α," "β," $\mathfrak{R}$ with

$$\text{"}y\text{"} = (\tfrac{2}{3})^{\frac{1}{2}}r, \quad \text{"}\alpha\text{"} = \tfrac{1}{2}\pi - 2y, \quad \text{"}\beta\text{"} = 2x.$$

Let us now study the invariances of the problem. We have, of course, the rotational invariance, apparent in (2.6) in the factor $[\mathfrak{R}_r^{-1} + \mathfrak{R}_r^{-1}]$, which is related below to the subgroup of SO_6, isomorphic to SO_3, the matrices of which are

$$\begin{pmatrix} \mathfrak{R}_r & 0 \\ 0 & \mathfrak{R}_r \end{pmatrix}.$$

On the other hand, the axis $\Omega\xi$ in $\mathfrak{T}$ is arbitrary, so that x is defined up to a constant. This invariance is related to the permutation group of the three particles[1,2,4] and corresponds to a subgroup of SO_6, isomorphic to SO_2; its matrices read

$$R_2(x) \otimes I_3.$$

Let us recall that these two subgroups of SO_6 are two commutative subgroups.

III. FREE HAMILTONIAN—GRAND-ANGULAR MOMENTUM

We do not write explicitly the free Hamiltonian in function of the spherical variables because its exact structure is of no interest for what follows; it can be found elsewhere.[2,3] We just notice that it may be split into two parts, namely,

$$H_0 = -\Delta_6 = -\frac{2}{3r^5}\frac{\partial}{\partial r}r^5\frac{\partial}{\partial r} - \frac{2}{3r^2}\delta_5, \quad (3.1)$$

where δ_5 does not act on r; Δ_6 is the six-dimensional Laplacian. The eigenfunctions of δ_5, often called the square of the grand-angular momentum,[5] have been

extensively studied in the literature; all we need by now is that its eigenvalues be $-n(n+4)$ with n integer ≥ 0.

The eigenvalue problem

$$H_0\psi = E\psi = k^2\psi$$

may be solved in two ways. First, we consider H_0 as the sum of three Laplacians in three dimensions; thus we get the plane wave

$$i\mathbf{k}_1 \cdot \mathbf{r}_1 + i\mathbf{k}_2 \cdot \mathbf{r}_2 + i\mathbf{k}_3 \cdot \mathbf{r}_3, \qquad (3.2)$$

with $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ or $\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 = 0$. The scalar product of the exponent is nothing but the six-dimensional Euclidian scalar product of the above vectors K and R, which in matrix language reads[10] K^tR; we can also write it as

$$K^tR = K \odot R = kr \cos_6 \theta. \qquad (3.3)$$

The second type of solutions is found with the help of Eq. (3.1). They are given by the product of an eigenfunction of δ_5 with a solution of the remaining part of H_0. Namely, ψ can be written as a superposition of ψ_n of the form

$$\psi_n = r^{-2}J_{n+2}(kr)F_n \qquad (3.4)$$

with

$$\delta_5 F_n = -n(n+4)F_n.$$

The relation between the plane wave and this last class of solutions is given by the Neumann series[11]

$$\exp(iK \odot R) = 4\sum_n (n+2)i^n J_{n+2}(kr)(kr)^{-2}C_n^2(\hat{K} \odot \hat{R}), \qquad (3.5)$$

where $C_n^2(t)$ is the Gegenbauer polynomial of degree n. The above expansion is a generalization of the well-known expansion in Legendre polynomials.

Equation (3.5) is the first step in expanding the plane wave. The free Green's function may also be written as an expansion in Gegenbauer polynomials. Indeed, we have, in Dirac notation,

$$\langle R| G_0(s) |R'\rangle = G(R, R', s) = \int \langle R | K\rangle \frac{d_6K}{k^2 - s} \langle K | R'\rangle, \qquad (3.6)$$

where, as usual, we have set

$$\langle R | K\rangle = (2\pi)^{-3} \exp(iK \odot R).$$

d_6K is the six-dimensional phase space, which can be evaluated in functions of the symmetrical variables introduced above. We have

$$d_6K = k^5 dk \, d\Omega(\hat{K}),$$

---

[10] K^t means the transposed matrix of K.
[11] *Higher Transcendental Functions*, A. Erdelyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, Formula 7.10.5.

where $d\Omega(\hat{K}) = d\mathcal{R}_x \, d\xi \sin 4\eta \, d\eta$ is the element of the unit sphere in six dimensions.

Now, each plane wave in Eq. (3.6) admits an expansion in Gegenbauer polynomials of the type (3.5). According to the orthogonality relation[12] on the unit sphere $\Omega(\hat{K})$,

$$\int_{\Omega(\hat{K})} C_n^2(\hat{R} \odot \hat{K}) C_m^2(\hat{K} \odot \hat{R}') \, d\Omega(\hat{K})$$

$$= \delta_{nm} \frac{2\pi^3}{n+2} C_n^2(\hat{R} \odot \hat{R}'), \quad (3.7)$$

we find

$$G(\mathbf{R}, \mathbf{R}', s) = \sum_n (n+2) G_n(r, r', s) C_n^2(\hat{R} \odot \hat{R}'), \quad (3.8)$$

with

$$G_n(r, r', s) = \frac{1}{2\pi^3} \frac{1}{(rr')^{\frac{3}{2}}} g_{n+\frac{3}{2}}(r, r', s), \quad (3.9)$$

$$g_l(r, r', s) = \frac{i\pi}{2} \frac{1}{(rr')^{\frac{1}{2}}} J_{l+\frac{1}{2}}(r_< s^{\frac{1}{2}}) H_{l+\frac{1}{2}}^{(1)}(r_> s^{\frac{1}{2}}) \quad (3.10)$$

being the well-known two-particle free Green's function.[13] Notice, however, that $g_l(r, r', s)$ is here involved for unphysical values of the two-particle angular momentum.

IV. SPHERICAL HARMONICS

To be complete, our plane-wave expansion must exhibit the dependence upon the five angles of $\hat{K}$ and of $\hat{R}$ involved in the Gegenbauer polynomial of Eq. (3.5). Indeed, this polynomial is well known, and $K \odot R$ is easy to compute; but this does not give directly the angular momentum dependence, for example, which is of great importance. To exhibit the angular momentum and other needed quantum numbers, we use the following expansion of the Gegenbauer polynomials in a function of a complete orthonormal set of spherical harmonics $S_n^l(\hat{R})$[12]:

$$C_n^2(\hat{K} \odot \hat{R}) = \frac{2\pi^3}{n+2} \sum_{l=1}^{h(n)} S_n^{l*}(\hat{K}) S_n^l(\hat{R}), \quad (4.1)$$

where $h(n) = \frac{1}{12}(n+1)(n+2)^2(n+3)$ is the total number of independent spherical harmonics of degree n. The spherical harmonics satisfy

$$\delta_5 S_n^l(\hat{R}) = -n(n+4) S_n^l(\hat{R}).$$

Let A be a proper orthogonal transformation in our

[12] See Ref. 11, Vol. II, Chap. XI.
[13] To perform the $k^5 \, dk$ integration indicated in Eq. (3.6) so as to obtain Eqs (3.9) and (3.10), see, for example, A. Sommerfeld, *Partial Differential Equations in Physics* (Academic Press Inc., New York, 1949), Sec. 28.

six-dimensional space, i.e., $A \in SO_6$; then

$$S_n^l(A\hat{R}) = \sum_{l'=1}^{h(n)} g_{ll'}^{(n)}(A) S_n^l(\hat{R}), \quad (4.2)$$

where the matrix $g^{(n)}(A)$ belongs to the irreducible representation of SO_6 of degree $h(n)$. Also

$$\sum_{l=1}^{h(n)} S_n^{l*}(A\hat{K}) S_n^l(A\hat{R}) = \sum_{l=1}^{h(n)} S_n^{l*}(\hat{K}) S_n^l(\hat{R}), \quad (4.3)$$

which is nothing but the invariance of the scalar product $\hat{K} \odot \hat{R}$ under the transformation A.

Now, what we are looking for is the form of $S_n^l(\hat{R})$ in a function of the spherical variables $\mathcal{R}_r$, x, and y, or equivalently, in a function of the good quantum numbers of the free three-particle states, i.e., angular momentum related to the invariance under SO_3, and "μ" related to SO_2. The decomposition (2.6) of R is of great interest for this task because it immediately exhibits the two transformations we need, namely, $\mathcal{R}_r$ and $R_2(x)$, which belong to SO_3 and SO_2, respectively.

The principal problem is then to decompose the representation $g^{(n)}$ of SO_6 with respect to the irreducible representation of SO_3 and SO_2. We do it in Sec. V. Here we are going to exploit the rotational invariance of the six-dimensional scalar product $\hat{K} \odot \hat{R}$ [cf. Eq. (4.3)]. As mentioned above, $\hat{K} \odot \hat{R}$ is just the matrix product $\hat{K}^t \hat{R} = \hat{R}^t \hat{K}$. From Eq. (2.6) and (2.7), and from the commutativity of the two subgroups SO_3 and SO_2 of SO_6, we have

$$\hat{K} \odot \hat{R} = \hat{K}_0^t A \hat{R}_0,$$

where A is, now and in what follows, given by

$$A = (\mathcal{R} + \mathring{\mathcal{R}})[R_2(x - \xi) \otimes I_3],$$

$$= [R_2(x - \xi) \otimes I_3](\mathcal{R} + \mathring{\mathcal{R}}), \quad (4.4)$$

with

$$\mathcal{R} = \mathcal{R}_k \mathcal{R}_r^{-1}. \quad (4.5)$$

Hence $\hat{K} \odot \hat{R} = \hat{K}_0 \odot (A\hat{R}_0)$, so that Eq. (4.1) may as well be written as

$$C_n^2(\hat{K} \odot \hat{R}) = \frac{2\pi^3}{n+2} \sum_{l=1}^{h(n)} S_n^{l*}(\hat{K}_0) S_n^l(A\hat{R}_0),$$

$$= \frac{2\pi^3}{n+2} \sum_{l,l'=1}^{h(n)} S_n^{l*}(\eta) g_{ll'}^{(n)}(A) S_n^{l'}(y), \quad (4.6)$$

where we have set

$$S_n^l(\hat{R}_0) = S_n^l(y) \quad (4.7)$$

to exhibit the fact that R_0 depends only on y. Similarly,

$$S_n^l(\hat{K}_0) = S_n^l(\eta).$$

So far, the functions $S_n^l(\hat{R})$ are greatly undetermined—they just form a "complete orthonormal set of spherical harmonics." There are, of course, several such sets, obtained from one of them by suitable linear combinations. In the following, we choose the $S_n^l(\hat{R})$ as eigenfunctions of J^2, J_z, and $-i\partial/\partial x$, with eigenvalues $j(j+1)$, m, and μ, which is possible because each of these operators commutes with the two others and with δ_5. Hence, l may be replaced by a set of four integers (j, m, μ, ω); the last one refers to the remaining degeneracy,[1,14] if any.

V. REDUCTION OF $g^{(n)}(A)$

$g^{(n)}(A)$ belongs to the irreducible representation of SO_6 of degree $h(n)$, the highest weight of which is $(n, 0, 0)$. If A is conjugate to

$$R_2(\tau_1) \dotplus R_2(\tau_2) \dotplus R_2(\tau_3), \tag{5.1}$$

we have[15]

$$\mathrm{Tr}\, g^{(n)}(A) = \frac{\begin{vmatrix} \cos(n+2)\tau_1 & \cos\tau_1 & 1 \\ \cos(n+2)\tau_2 & \cos\tau_2 & 1 \\ \cos(n+2)\tau_3 & \cos\tau_3 & 1 \end{vmatrix}}{\begin{vmatrix} \cos 2\tau_1 & \cos\tau_1 & 1 \\ \cos 2\tau_2 & \cos\tau_2 & 1 \\ \cos 2\tau_3 & \cos\tau_3 & 1 \end{vmatrix}}. \tag{5.2}$$

On the other hand, A belongs to a particular subgroup G of SO_6; thus $g^{(n)}(A)$ must reduce with respect to the irreducible representations $h_\lambda^{(n)}(A)$ of this subgroup G, according to

$$g^{(n)}(A) = \bigoplus_\lambda \gamma^\lambda h_\lambda^{(n)}(A), \tag{5.3}$$

where

$$\gamma^\lambda = \int_G \mathrm{Tr}\, g^{(n)}(A)^* \,\mathrm{Tr}\, h_\lambda^{(n)}(A)\, dg, \tag{5.4}$$

dg being the normalized elementary volume of the subgroup G.

The particular choice of $S_n^l(R)$ we have made at the end of Sec. IV and the form of A induce remarkable properties of $g^{(n)}(A)$. It is the (commutative) product of two matrices, one being exactly the direct sum of irreducible representations D^j of SO_3, the other a diagonal matrix whose elements are the one-dimensional representations $e^{i\mu(x-\xi)}$ of SO_2. Furthermore, the commutativity of the two associated subgroups of SO_6 enables us to write

$$g^{(n)}(\mathcal{R} \dotplus \mathcal{R}) = \bigoplus_{j=0}^{j\,\max} \alpha_j^{(n)} D^j(\mathcal{R}) \tag{5.5}$$

and

$$g^{(n)}[R_2(x-\xi) \otimes I_3] = \bigoplus_{j=0}^{j\,\max} \bigoplus_{\varphi=1}^{\varphi=\alpha_j^{(n)}} [e^{i\mu_j,\varphi(x-\xi)} I_{2j+1}], \tag{5.6}$$

where I_{2j+1} is the $(2j+1) \times (2j+1)$ unit matrix, the second formula being a consequence of Schur's lemma. More explicitly, for a given value of j, we have at most $\alpha_j^{(n)}$ different values of μ. The allowed values of j and μ for a given n and the corresponding degeneracies $\alpha_j^{(n)}$ and $\beta_\mu^{(n)}$ may be calculated with the help of Eq. (5.4) and (5.2); we must, however, notice that the first subgroup is characterized by $\tau_1 = \tau_2 = \varphi$, $\tau_3 = 0$ and the second by $\tau_1 = \tau_2 = \tau_3 = x$. We also need the values of dg and the traces of the irreducible representations of the preceding subgroups. In the first case, dg is proportional to $\sin^2(\frac{1}{2}\varphi)\,d\varphi$ and $\mathrm{Tr}\, D^j(\mathcal{R}) = \sin(j+\frac{1}{2})\varphi/(\sin\varphi/2)$. In the second case, dg is proportional to dx, and the trace is the matrix itself $e^{i\mu x}$. In Appendix A, we give a resumé of these calculations. It is interesting to notice the following properties of these numbers:

$$\alpha_0^{(n)} = 0 \quad \text{if} \quad n \text{ odd},$$
$$\alpha_j^{(n)} = 0 \quad \text{if} \quad j > n,$$
$$\alpha_j^{(n)} = 0 \quad \text{if} \quad [\tfrac{1}{2}(n+1)] = [\tfrac{1}{2}j],$$

except[16] when n and j are both even, in which case $\alpha_j^{(n)} = n+1 = j+1$;

$$\beta_\mu^{(n)} = \beta_{-\mu}^{(n)},$$
$$\beta_\mu^{(n)} \neq 0 \quad \text{only if } \mu \text{ and } n \text{ have the same parity,}$$
$$\beta_\mu^{(n)} = 0 \quad \text{if} \quad |\mu| > n.$$

However, the most interesting problem at this point is not the numbers $\alpha_j^{(n)}$ and $\beta_\mu^{(n)}$, but the degeneracy when j and μ are given, i.e., when we are dealing with the whole matrix A. We have

$$g^{(n)}(A) = \sum_{j,\mu} \gamma_{j,\mu}^{(n)}[D^j(\mathcal{R}) \cdot e^{i\mu(x-\xi)}]$$

and the γ's may be evaluated from Eq. (5.4). The calculations are somewhat tedious and are indicated in Appendix B. Let us give here some properties of these numbers:

$$\gamma_{j,\mu}^{(n)} \neq 0 \quad \text{only if } \mu \text{ and } n \text{ have the same parity,}$$
$$\gamma_{j,\mu}^{(n)} = 0 \quad \text{if} \quad n < j \quad \text{or} \quad n < |\mu|,$$
$$\gamma_{j,\mu}^{(n)} = \gamma_{j,-\mu}^{(n)},$$
$$\gamma_{j,\mu}^{(n)} \leq [\tfrac{1}{2}j] + 1 \quad \text{if } n \text{ even,}$$
$$\gamma_{j,\mu}^{(n)} \leq [\tfrac{1}{2}(j+1)] \quad \text{if } n \text{ odd.}$$

[14] For simplicity, ω is here an integer which runs from 1 to $\gamma_{j,\mu}^{(n)}$ (see Sec. V). Thus our ω is not the same as the ω of Ref. 1.
[15] See, for example, H. Boerner, *Representation of Groups* (North-Holland Publishing Company, Amsterdam, 1963).
[16] $[x]$ means the greater integer less than x.

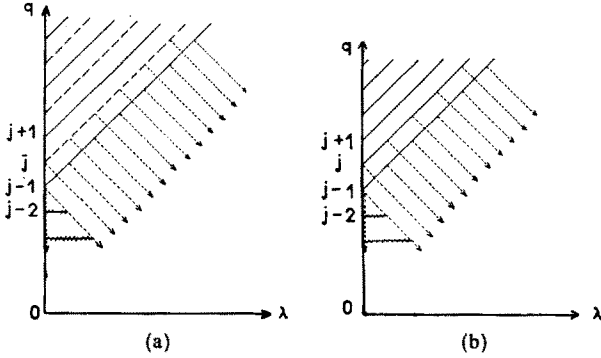FIG. 1. Numbers $\gamma_{j,\mu}^{(n)}$ for a given value of j, in a function of $[\frac{1}{2}n] = q$ and $[\frac{1}{2}\mu] = \lambda$. (a) n and μ even. ------ $\gamma_{j,\mu}^{(n)} = [\frac{1}{2}j] + 1$. ——— $\gamma_{j,\mu}^{(n)} = [\frac{1}{2}j]$. ----→ $\gamma_{j,\mu}^{(n)}$ decreases by one unit at each point of the lattice, in the sense given by the arrow, from its maximum value to 0. /////$\gamma_{j,\mu}^{(n)}$ = const. (b) n and μ odd. ——— $\gamma_{j,\mu}^{(n)} = [\frac{1}{2}(j+1)]$. ----→, ///// the same signification as in case (a).

The last two conditions have an interesting consequence: If j is equal to 0 or 1, $\gamma_{j,\mu}^{(n)}$ is never greater than 1, i.e., n and μ are sufficient to determine a three-particle S or P state. We give the exact value of $\gamma_{j,\mu}^{(n)}$ in Fig. 1.

VI. PLANE-WAVE AND GREEN'S FUNCTION EXPANSION

According to Eqs. (5.5) and (5.6), we may write

$$g_{ll'}^{(n)}(A) = \mathfrak{D}_{mm'}^j(\mathfrak{R})e^{i\mu(x-\xi)},$$

where $l = (j, m, \mu, \omega)$, $l' = (j, m', \mu, \omega)$,

$$g_{ll'}^{(n)}(A) = 0 \quad \text{otherwise.}$$

Hence Eq. (4.6) reads now

$$C_n^2(\hat{K} \odot \hat{R})$$

$$= \frac{2\pi^3}{n+2} \sum_{jmm'\mu\omega} S_n^{jm\mu\omega*}(\eta)\mathfrak{D}_{mm'}^j(\mathfrak{R})e^{i\mu(x-\xi)}S_n^{jm'\mu\omega}(y),$$

where $\sum_{jmm'\mu\omega}$ stands for $\sum_j \sum_{mm'} \sum_\mu \sum_\omega$ in that order and is to be taken over all possible values of the indices; for instance, m goes from $-j$ to $+j$ and ω from 1 to $\gamma_{j,\mu}^{(n)}$; to be precise, we might have given indices to μ and ω, but this would have made the formulas too heavy. We keep this convention below, i.e., the summations are to be taken in the specified order and only over the possible values of the indices, once the preceding indices are fixed.

We are now able to complete the expansions (3.5) and (3.8) of the plane wave and of the free Green's function. We get

$$8\pi^3\langle R \mid K \rangle = e^{iK \odot R},$$

$$= \sum_{jmm'} \mathfrak{D}_{mm'}^j(\mathfrak{R}_k\mathfrak{R}_r^{-1}) \sum_n i^n J_{n+2}(kr)(kr)^{-2}$$

$$\times \sum_\mu e^{i\mu(x-\xi)} \sum_\omega S_n^{jm\mu\omega*}(\eta)S_n^{jm'\mu\omega}(y),$$

(6.1)

$$\langle R \mid G_0(s) \mid R' \rangle = \frac{i\pi}{2} \sum_{jmm'} \mathfrak{D}_{mm'}^j(\mathfrak{R}_r\mathfrak{R}_{r'}^{-1})$$

$$\times \sum_n (rr')^{-2}H_{n+2}^{(1)}[r_> s^{\frac{1}{2}}]J_{n+2}[r_< s^{\frac{1}{2}}] \sum_\mu e^{i\mu(x-x')}$$

$$\times \sum_\omega S_n^{jm\mu\omega*}(y')S_n^{jm'\mu\omega}(y).$$

(6.2)

Our task will now be achieved after we have evaluated the functions $S_n^l(y)$. What we know about them are the following properties:

$$S_n^l(y) = S_n^l(\cos y, 0, 0, 0, \sin y, 0).$$

(6.3)

$S_n^l(\hat{R})$ is a homogeneous polynomial of degree n in the (six) components of $\hat{R}$:

$$S_n^{jm\mu\omega}(\hat{R}) = \sum_{m'} e^{i\mu x}\mathfrak{D}_{mm'}^j(\mathfrak{R}_r^{-1})S_n^{jm'\mu\omega}(y),$$

(6.4)

$$\delta_5 S_n^l(\hat{R}) = -n(n+4)S_n^l(\hat{R}),$$

(6.5)

$$J^2 S_n^l(\hat{R}) = j(j+1)S_n^l(\hat{R}),$$

(6.6)

$$J_Z S_n^l(\hat{R}) = mS_n^l(\hat{R}),$$

(6.7)

$$-i(\partial/\partial x)S_n^l(\hat{R}) = \mu S_n^l(\hat{R}).$$

(6.8)

l stands for the set (j, m, μ, ω). Furthermore, the $S_n^l(\hat{R})$ must satisfy the orthogonality relation

$$\int_{\Omega(\hat{R})} S_n^{l*}(\hat{R})S_m^{l'}(\hat{R}) \, d\Omega(\hat{R}) = \delta_{nm}\delta_{ll'}.$$

(6.9)

In order to take Eqs. (6.6) to (6.8) into account, it is necessary to know δ_5 in a function of the operators J^2, J_Z, $-i\partial/\partial x$. This problem has been solved for particular values of j ($j = 0, 1, 2,$[1,3] $j = 0, 1,$[2] $j = 0$[4]). For $j = 0$, the result is that $S_n^l(y)$ is a Jacobi polynomial.

Let us conclude by some remarks. Equations (6.1) and (6.2) have been written in such a way that the angular momentum dependence of the plane wave and of Green's function appears immediately. In order to get the projection of one of these functions on a wave of definite angular momentum, it is now sufficient to take the coefficient of the $\mathfrak{D}^j$ function in the expansions (6.1) and (6.2). For $j = 0$ the result is already known,[7] but it seems to be new for greater values of j.

VII. CONCLUSION

Equations (6.1) and (6.2) may now be the starting point of the analysis of the wavefunction of three spinless particles, interacting via some potential. They allow us to immediately transform the Schrödinger equation into an integral equation of the Lippmann-Schwinger type, and thus to study the analytic properties of the wavefunction. This is our present purpose and will be the subject of future publications.

APPENDIX A. DEGENERACIES $\alpha_j^{(n)}$ AND $\beta_\mu^{(n}$

I. $\alpha_j^{(n)}$

A rotation $\mathcal{R}$ of SO_3 may be characterized by a single angle φ. The character of the representation of degree $2j + 1$ is

$$\text{Tr } D^j(\mathcal{R}) = \chi_j(\varphi) = \frac{\sin(j + \tfrac{1}{2})\varphi}{\sin \varphi/2}, \quad (A1)$$

and the elementary volume is proportional to $\sin^2(\varphi/2)\,d\varphi$. On the other hand, according to Eq. (4.2) and the fact that our subgroup SO_3 of SO_6 is characterized by $\tau_1 = \tau_2 = \varphi$ and $\tau_3 = 0$, we have

$$\chi_n(\mathcal{R} + \mathcal{R}) = \chi_n(\varphi) = \lim_{\varphi'=\varphi} \begin{vmatrix} \cos(n+2)\varphi' & \cos \varphi' & 1 \\ \cos(n+2)\varphi & \cos \varphi & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

$$\times \left(\begin{vmatrix} \cos 2\varphi' & \cos \varphi' & 1 \\ \cos 2\varphi & \cos \varphi & 1 \\ 1 & 1 & 1 \end{vmatrix} \right)^{-1}. \quad (A2)$$

For the sake of simplicity, we express everything in functions of Tchebichef polynomials[17] of $\cos \varphi/2 = \Phi$:

$$T_n(\Phi) = \cos(n\varphi/2),$$
$$U_n(\Phi) = [\sin(n+1)\varphi/2]/(\sin \varphi/2).$$

Equations (A1) and (A2) may be expressed as

$$\chi_j(\varphi) = U_{2j}(\Phi), \quad (A3)$$

$$\chi_n(\varphi) = \frac{1}{8\Phi} \frac{d}{d\Phi} \frac{1 - T_{2n+4}(\Phi)}{2(1 - \Phi^2)},$$

$$= \frac{1}{8\Phi} \sum_{m=0}^{n+1} U'_{2m}(\Phi). \quad (A4)$$

Furthermore, the orthogonality relation between the characters of SO_3 yields

$$\int_0^{2\pi} U_{2j}(\Phi) U_{2J}(\Phi) \sin^2(\varphi/2)\,d\varphi = \pi \delta_{jJ}, \quad (A5)$$

and the numbers $\alpha_j^{(n)}$ are given by

$$\alpha_j^{(n)} = \frac{1}{\pi} \int_0^{2\pi} \chi_n(\varphi)\chi_j(\varphi) \sin^2(\varphi/2)\,d\varphi. \quad (A6)$$

Now, our problem is just reduced to express $U'_{2m}(\Phi)/\Phi$ as a linear combination of $U_p(\Phi)$; this expansion is summed up in the following formulas,

---

[17] See Ref. 11, Vol. II, Sec. X, 11.

the derivation of which is straightforward:

$$U'_{2m}(\Phi)/(\Phi) = 4 \sum_{p=1}^{m} p U_{2p-1}(\Phi)/\Phi,$$

$$U_{2p-1}(\Phi)/\Phi = 2 \sum_{k=1}^{p} (-)^{p-k} U_{2k-2}.$$

From these expressions and Eq. (A5), we get

$$\alpha_j^{(n)} = (-)^{j+1} \sum_{p=j+1}^{n+1} (-)^p p(n - p + 2). \quad (A7)$$

In particular,

$$\alpha_0^{(0)} = 1,$$
$$\alpha_0^{(1)} = 0, \quad \alpha_1^{(1)} = 2,$$
$$\alpha_0^{(2)} = 2, \quad \alpha_1^{(2)} = 1, \quad \alpha_2^{(2)} = 3,$$
$$\alpha_0^{(3)} = 0, \quad \alpha_1^{(3)} = 4, \quad \alpha_2^{(3)} = 2, \quad \alpha_3^{(3)} = 4.$$

These results are in agreement with those of Ref. 1.

Equation (A7) may be simplified, according to the parity of j and n. We have

$$n < j: \quad \alpha_j^{(n)} = 0,$$
$$n = 2q, \quad j = 2r, \quad \alpha_j^{(n)} = q + r + 1 + 2r(q - r),$$
$$j = 2r + 1, \quad \alpha_j^{(n)} = (2r + 1)(q - r),$$
$$n = 2q + 1, \quad j = 2r, \quad \alpha_j^{(n)} = 2r(q + 1 - r),$$
$$j = 2r + 1, \quad \alpha_j^{(n)} = 2(r + 1)(q + 1 - r).$$

II. $\beta_\mu^{(n)}$

A rotation of SO_2 depends on one angle x. All the representations are of degree 1 and the characters are $e^{i\mu x}$, μ integer; dg is proportional to dx.

Now, the subgroup SO_2 of SO_6 we need is characterized by $\tau_1 = \tau_2 = \tau_3 = x$, and the traces of the corresponding representations are got from Eq. (5.2) by a limiting procedure. Let us write them as

$$\chi_n[R_2(x) \otimes I_3] = \chi_n(x), \quad (A8)$$
$$= \lim_{\tau_1=\tau_2=\tau_3=x} \chi_n[R_2(\tau_1) + R_2(\tau_2) + R_2(\tau_3)].$$

We again use here Tchebichef polynomials of $X = \cos x$. Hence we may write

$$\chi_n(x) = \tfrac{1}{4} T''_{n+2}(X) = \frac{n + 2}{4} U'_{n+1}(X). \quad (A9)$$

The degeneracies $\beta_\mu^{(n)}$ are proportional to the integral

$$\int_0^{2\pi} \chi_n(x) e^{i\mu x}\,dx = \int_0^{2\pi} \chi_n(x) e^{-i\mu x}\,dx,$$

$$= \int_0^{2\pi} \chi_n(x) \cos \mu x\,dx.$$

Hence $\beta_\mu^{(n)} = \beta_{-\mu}^{(n)}$ and we may suppose $\mu \geq 0$ in the following. On the other hand, we may replace the

trace $e^{i\mu x}$ by $\cos \mu x = T_\mu(X)$. The orthogonality relation between the traces now reads

$$\int_0^{2\pi} T_m(X) T_n(X)\, dx = \pi \epsilon_m \delta_{mn}, \qquad \text{(A10)}$$

where $\epsilon_0 = 2$, $\epsilon_m = 1$ if $m > 0$. Finally, we have

$$\beta_\mu^{(n)} = \frac{1}{2\pi} \int_0^{2\pi} \epsilon_\mu \chi_n(x) \cos \mu x\, dx. \qquad \text{(A11)}$$

Thus the problem is reduced to the expansion of $U_{n+1}'(X)$ in functions of $T_\mu(X)$. This expansion follows from formula (40) of Ref. 17. We just give the result:

$$
\begin{aligned}
n < \mu, &\qquad \beta_\mu^{(n)} = 0, \\
n = 2q, \quad \mu = 2\lambda, &\qquad \beta_\mu^{(n)} = (q+1) \\
&\qquad \times (q+1-\lambda)(q+1+\lambda), \\
\mu = 2\lambda + 1, &\qquad \beta_\mu^{(n)} = 0, \\
n = 2q+1, \quad \mu = 2\lambda, &\qquad \beta_\mu^{(n)} = 0, \\
\mu = 2\lambda + 1, &\qquad \beta_\mu^{(n)} = (2q+3) \\
&\qquad \times (q+\lambda+2)(q+1-\lambda)/2.
\end{aligned}
$$

APPENDIX B. DEGENERACY $\gamma_{j,\mu}^{(n)}$

We use the same notation as in Appendix A. We are here concerned with a transformation A which belongs to a subgroup G of SO_6, isomorphic to the direct product of SO_3 and SO_2. The characters of G are just

$$\chi_j(\varphi) \chi_\mu(x) = U_{2j}(\Phi) e^{i\mu x}.$$

According to the symmetry $\mu \to -\mu$, this expression may be replaced in what follows by $U_{2j}(\Phi) T_\mu(X)$. The elementary volume of the group G is proportional to $\sin^2(\varphi/2)\, d\varphi\, dx$.

Considered as an element of SO_6, A is conjugate to $R_2(x + \varphi) \dotplus R_2(x - \varphi) \dotplus R_2(x)$, so that

$$\chi_n(X, \Phi) = \mathrm{Tr}\, g^{(n)}(A),$$

$$= \begin{vmatrix} \cos(n+2)(x+\varphi) & \cos(x+\varphi) & 1 \\ \cos(n+2)(x-\varphi) & \cos(x-\varphi) & 1 \\ \cos(n+2)x & \cos x & 1 \end{vmatrix}$$

$$\times \left(\begin{vmatrix} \cos 2(x+\varphi) & \cos(x+\varphi) & 1 \\ \cos 2(x-\varphi) & \cos(x-\varphi) & 1 \\ \cos 2x & \cos x & 1 \end{vmatrix} \right)^{-1}.$$

After some straightforward calculations, this expression reduces to

$$\chi_n(X, \Phi) = U_{n+1}(\Phi) \Delta_{n+2}(X, \Phi)/4\Phi(\Phi^2 - X^2),$$

where

$$
\begin{aligned}
\Delta_{n+2}(X, \Phi) &= X U_{n+1}(X) T_{n+2}(\Phi) - \Phi U_{n+1}(\Phi) T_{n+2}(X), \\
&= U_{n+2}(X) T_{n+2}(\Phi) - U_{n+2}(\Phi) T_{n+2}(X).
\end{aligned}
$$

By induction, we get

$$\Delta_{n+2}(X, \Phi) = 2(\Phi^2 - X^2) \sum_{k=0}^{[\frac{1}{2}n]} U_{n-2k}(X) U_{n-2k}(\Phi).$$

The degeneracy $\gamma_{j,\mu}^{(n)}$ is given by

$$2\pi^2 \gamma_{j,\mu}^{(n)} = \int_0^{2\pi} dx \int_0^{2\pi} \sin^2(\varphi/2)\, d\varphi\, \chi_n(X, \Phi)$$

$$\times U_{2j}(\Phi) \epsilon_\mu T_\mu(X).$$

According to Eq. (A5) and (A10), $\gamma_{j,\mu}^{(n)}$ is just the coefficient of $U_{2j}(\Phi) \epsilon_\mu T_\mu(X)$ in the expansion of $\chi_n(X, \Phi)$. It is finally possible to write this expansion as follows:

$$
\begin{aligned}
\chi_n(X, \Phi) = &\sum_{J=0}^{J=[\frac{1}{2}n]} U_{2J} \sum_{m=0}^{m=J} \tau_{n-2m} \\
&\times \left\{ \left[\frac{m+1}{2} \right] + \frac{1+(-)^m}{2} \frac{1+(-)^{n+J}}{2} \right\} \\
&+ \sum_{J=0}^{J=[\frac{1}{2}n]-1} U_{2J} \sum_{m=J+1}^{m=[\frac{1}{2}n]} \tau_{n-2m} \\
&\times \left\{ \left[\frac{J+1}{2} \right] + (-)^J \frac{1+(-)^n}{2} \frac{1+(-)^m}{2} \right\} \\
&+ \sum_{J=[\frac{1}{2}n]+1}^{J=n} U_{2J} \sum_{m=0}^{m=n-J} \tau_{n-2m} \\
&\times \left\{ \left[\frac{m+1}{2} \right] + \frac{1+(-)^m}{2} \frac{1+(-)^{n+J}}{2} \right\} \\
&+ \sum_{J=[\frac{1}{2}n]+1}^{J=n} U_{2J} \sum_{m=n-J+1}^{m=[\frac{1}{2}n]} \tau_{n-2m} \left\{ \left[\frac{n-J}{2} \right] + 1 \right\},
\end{aligned}
$$

where τ_r stands for $\epsilon_r T_r$, the arguments of U_{2J} and τ_r being Φ and X, respectively. The corresponding degeneracies are exactly the expressions between braces, and are represented in Fig. 1.

Eigenvalue Problem for Lagrangian Systems. II*

E. M. Barston

Courant Institute of Mathematical Sciences, New York University, New York, New York

The quadratic Lagrangian eigenvalue problem $[\omega^2 I - \omega i A - H]\xi = 0$ for H and iA, completely continuous Hermitian operators in a Hilbert space E, $H \geq 0$, is investigated. The problem is reduced to an equivalent linear eigenvalue problem for a single completely continuous Hermitian operator in the Hilbert space $E \times E$, and existence and convergence theorems for the eigenvectors and variational properties of the eigenvalues for the original quadratic problem are easily obtained from standard theorems. The general solution of the associated time-dependent problem $\ddot{\xi} + A\dot{\xi} + H\xi = 0$ is obtained under the further restriction that E be finite dimensional. Necessary and sufficient conditions for stability are given.

I. INTRODUCTION

THIS paper continues the discussion presented earlier[1] of the eigenvalue problem

$$H_\omega \xi \equiv [\omega^2 I - \omega i A - H]\xi = 0 \qquad (1)$$

and the associated time-dependent problem

$$\ddot{\eta} + A\dot{\eta} + H\eta(t) = 0, \quad t \in (0, \infty), \qquad (2)$$

with the boundary conditions $\dot{\eta}(0+) = \dot{x}$, $\eta(0+) = x$. The operators H and iA are linear Hermitian operators on and into an inner product space E, $x \in E$, $\dot{x} \in E$, $\eta(t) \in E$ for $0 < t < \infty$; and the eigenvector ξ of Eq. (1) corresponding to the eigenvalue ω (a real or complex number) is a nonzero element of E. Equation (1) is obtained from Eq. (2) by considering solutions of the form $\eta(t) = e^{i\omega t}\xi$.

We confine our attention in this paper primarily to the case where H and iA are completely continuous, $H \geq 0$, and E is a Hilbert space. Under these circumstances, the eigenvalue problem of Eq. (1) can be reduced to an equivalent linear eigenvalue problem for a single completely continuous Hermitian operator on the Hilbert space $E^2 \equiv E \times E$, and existence and convergence theorems for the eigenvectors of Eq. (1) and variational properties of the eigenvalues follow immediately from standard theorems. Section II of this paper is devoted to a discussion of the equivalent problem and some of its consequences.

Section III contains a uniqueness theorem for the solutions of Eq. (2), and the general solution of Eq. (2) is obtained under the additional restriction that E be finite dimensional. This generalizes a result obtained earlier[1] for the case $H > 0$. We find that the general solution of Eq. (2) for $H \geq 0$ is given by a sum of eigenvectors of Eq. (1) plus a solution of the form $t\eta + \delta$, where $H\eta = 0$. Necessary and sufficient

conditions on A and H are given for stability ($\eta = 0$).

The results presented here and in Ref. 1 have numerous applications to problems of small oscillations about states of steady motion in fluid mechanics and plasma physics. In general, these problems involve differential (unbounded) operators. The question of converting these differential equations to Hilbert-space problems of the form of Eq. (1) involving completely continuous operators will be discussed in Part III of this series, along with applications to specific problems.

II. EIGENVALUES, EIGENVECTORS, AND THE EQUIVALENT PROBLEM

We begin with the introduction of some notation. The range and null space of an operator T is denoted by R_T and N_T, respectively, and $\bar{S}$ represents the closure of the set S. Let E be an inner product space with inner product $(\ ,\ )$. The product space $E^2 \equiv E \times E$ is the inner product space consisting of all 2-vectors η of the form $\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$, where $\eta_k \in E$, $k = 1, 2$, with inner product $(\ ,\ )_2$ defined by $(\eta, \zeta)_2 \equiv (\eta_1, \zeta_1) + (\eta_2, \zeta_2)$. If E is a Hilbert space, so is E^2. Finally, we note that the adjoint of a linear operator W on E^2, represented by the 2×2 matrix (W_{ij}) $i, j = 1, 2$, where the W_{ij} are bounded linear operators on E, is given by (W^*_{ji}), the asterisk denoting the adjoint.

Theorem I: Let $H^{\frac{1}{2}}$ be a linear operator defined on and into E with the property $(H^{\frac{1}{2}})^2 = H$. Suppose that

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \in E^2$$

is an eigenvector of

$$W \equiv \begin{pmatrix} 0 & -iH^{\frac{1}{2}} \\ iH^{\frac{1}{2}} & iA \end{pmatrix}$$

[1] E. M. Barston, J. Math. Phys. **8**, 523 (1967).

with eigenvalue ω, i.e., $\omega\eta = W\eta$, $\eta \neq 0$. If $\omega \neq 0$, it is both necessary and sufficient that

$$\eta_1 = -i\omega^{-1}H^{\frac{1}{2}}\xi,$$
$$\eta_2 = \xi, \quad H_\omega\xi = 0, \quad \xi \neq 0. \tag{3}$$

If $\omega = 0$, then $H\eta_2 = 0$ and $\eta_2 = 0$ implies $\eta_1 \neq 0$, $H\eta_1 = 0$; conversely, $H\xi = 0$ and $\xi \neq 0$ implies that $\eta \equiv \begin{pmatrix}\xi\\0\end{pmatrix}$ is an eigenvector of W with eigenvalue 0, if $H^{\frac{1}{2}}$ is Hermitian.

Proof: $W\eta = \omega\eta$ holds if and only if

$$\omega\eta_1 = -iH^{\frac{1}{2}}\eta_2, \tag{4}$$
$$\omega\eta_2 = iH^{\frac{1}{2}}\eta_1 + iA\eta_2. \tag{5}$$

Multiplying Eq. (5) by ω and substituting Eq. (4) into the result yields $H_\omega\eta_2 = 0$. Equation (4), $\omega \neq 0$, and $\eta \neq 0$ imply that $\eta_2 \neq 0$. Thus Eqs. (4) and (5) hold for $\omega \neq 0$, $\eta \neq 0$ only if Eqs. (3) hold. Conversely, $\omega \neq 0$ and Eqs. (3) imply Eq. (4), and Eq. (5) follows from $\omega^{-1}H_\omega\eta_2 = 0$ and Eq. (4). Suppose $\omega = 0$. Then Eq. (5), $\eta \neq 0$, and $\eta_2 = 0$ imply $\eta_1 \neq 0$ and $H\eta_1 = 0$. $H\xi = 0$, $\xi \neq 0$, and $H^{\frac{1}{2}}$ Hermitian imply $H^{\frac{1}{2}}\xi = 0$, since $(H^{\frac{1}{2}}\xi, H^{\frac{1}{2}}\xi) = (\xi, H\xi) = 0$, so that $W\begin{pmatrix}\xi\\0\end{pmatrix} = 0$.

Theorem II: Let H and iA be completely continuous Hermitian operators on the Hilbert space E (dim $E \leq \infty$), $H \geq 0$, $\|A\| + \|H\| > 0$. Then there exists a unique positive completely continuous linear Hermitian operator $H^{\frac{1}{2}}$ with domain E and range in E with the property $(H^{\frac{1}{2}})^2 = H$, and a nonempty sequence [finite or infinite, with at most $2(\dim E)$ elements] of eigenvectors $\{\xi_k\}$ ($k = 1, 2, 3, \cdots$) of Eq. (1) with real nonzero eigenvalues ω_k having the following properties:

(A) $H_{\omega_k}\xi_k = 0$, $k = 1, 2, 3, \cdots$. (6)

(B) $\{|\omega_k|\}$ is a monotone nonincreasing sequence, and $\lim_{k\to\infty} \omega_k = 0$ if the sequence $\{\xi_k\}$ is infinite.

(C) dim $F(\omega) < \infty$, where $F(\omega)$ is the linear manifold of eigenvectors ξ_k with eigenvalue $\omega_k = \omega$.

(D) $\omega_k\omega_l(\xi_k, \xi_l) + (\xi_k, H\xi_l) = \omega_k^2\delta_{kl}$
$$k, l = 1, 2, 3, \cdots. \tag{7}$$

(E) Let $x \in E$, $\dot{x} \in S(x)$, where

$$S(x) = \{\dot{x} \mid \dot{x} = iH^{\frac{1}{2}}g + iAx, g \in E\}.$$

Then

$$\left\|H^{\frac{1}{2}}\left[x - \sum_{k=1}^{n}\alpha_k\xi_k\right]\right\| \leq |\omega_{n+1}| [\|g\|^2 + \|x\|^2]^{\frac{1}{2}}, \tag{8}$$

$$\left\|\dot{x} - \sum_{k=1}^{n}\alpha_k\omega_k\xi_k\right\| \leq |\omega_{n+1}| [\|g\|^2 + \|x\|^2]^{\frac{1}{2}}, \tag{9}$$

where

$$\alpha_k \equiv \frac{\omega_k(\xi_k, \dot{x}) + (\xi_k, Hx)}{\omega_k^2(\xi_k, \xi_k) + (\xi_k, H\xi_k)}. \tag{10}$$

(F) For $x \in E$, $\dot{x} \in \overline{S(x)}$ [note that $H > 0$ implies $\overline{S(x)} = E$], we have

$$H^{\frac{1}{2}}x = \sum_k \alpha_k H^{\frac{1}{2}}\xi_k, \tag{11}$$

$$\dot{x} = \sum_k \alpha_k\omega_k\xi_k. \tag{12}$$

(G) Let $x \in E$, $\dot{x} \in \overline{S(x)}$, and β_k ($k = 1, 2, \cdots, n$) be any n complex numbers. Then

$$\left\|\dot{x} - \sum_{k=1}^{n}\beta_k\omega_k\xi_k\right\|^2 + \left\|x - \sum_{k=1}^{n}\beta_k\xi_k\right\|_H^2$$
$$= \left\|\dot{x} - \sum_{k=1}^{n}\alpha_k\omega_k\xi_k\right\|^2 + \left\|x - \sum_{k=1}^{n}\alpha_k\xi_k\right\|_H^2$$
$$+ \left\|\sum_{k=1}^{n}(\alpha_k - \beta_k)\omega_k\xi_k\right\|^2 + \left\|\sum_{k=1}^{n}(\alpha_k - \beta_k)\xi_k\right\|_H^2$$
$$\geq \left\|\dot{x} - \sum_{k=1}^{n}\alpha_k\omega_k\xi_k\right\|^2 + \left\|x - \sum_{k=1}^{n}\alpha_k\xi_k\right\|_H^2,$$
$$= (\dot{x}, \dot{x}) + (x, Hx) - \sum_{k=1}^{n}|\alpha_k|^2 \omega_k^2, \tag{13}$$

where the α_k are given by Eq. (10) and $\|y\|_H^2 \equiv (y, Hy)$.

(H) For $x \in E$, $\dot{x} \in \overline{S(x)}$, we have

$$(\dot{x}, \dot{x}) + (x, Hx) = \sum_k |\alpha_k|^2 \omega_k^2. \tag{14}$$

Proof: Since H is a positive linear Hermitian operator on E, there exists a unique positive Hermitian linear operator $H^{\frac{1}{2}}$ with domain E and range in E such that $(H^{\frac{1}{2}})^2 = H$.[2] The complete continuity of $H^{\frac{1}{2}}$ follows from $(H^{\frac{1}{2}})^2 = H$ and the fact that H is completely continuous. The linear operator W defined in Theorem I on the Hilbert space $E \times E$ is therefore completely continuous (each of its elements are completely continuous operators) and Hermitian, and $\|A\| + \|H\| > 0$ implies $\|W\| > 0$. Therefore there exists a nonempty sequence [finite or infinite, containing at most dim $(E \times E) = 2(\dim E)$ elements] of orthonormal elements $\{\eta_k\}$ in $E \times E$ ($k = 1, 2, 3, \cdots$) with the following properties[3]:

(a) $\quad W\eta_k = \omega_k\eta_k, \quad (\eta_k, \eta_l)_2 = \delta_{kl},$
$$k = 1, 2, 3, \cdots. \tag{15}$$

(b) The ω_k's are real and nonzero, $\{|\omega_k|\}$ is a monotone nonincreasing sequence, and $\lim_{k\to\infty} \omega_k = 0$ if the sequence $\{\eta_k\}$ is infinite.

[2] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 265.
[3] Reference 2, pp. 227–234.

(c) $\dim G(\omega) < \infty$, where $G(\omega)$ is the space of all eigenvectors η_k with eigenvalue $\omega_k = \omega$.

(d) For each $\eta \in E \times E$, we have

$$\left\| W\eta - \sum_{k=1}^{n} (\eta_k, W\eta)_2 \eta_k \right\|_2 \leq |\omega_{n+1}| \, \|\eta\|_2. \quad (16)$$

(e) For every $\eta \in \bar{R}_W$, we have

$$\eta = \sum_k (\eta_k, \eta)_2 \eta_k. \quad (17)$$

Define $\xi_k \equiv \eta_{k,2}$, where

$$\eta_k = \begin{pmatrix} \eta_{k,1} \\ \eta_{k,2} \end{pmatrix}, \quad k = 1, 2, 3, \cdots.$$

Statements (A), (B), and (C) follow immediately from Theorem I and (a), (b), and (c). Theorem I implies

$$\eta_{k,1} = -i\omega_k^{-1} H^{\frac{1}{2}} \xi_k, \quad k = 1, 2, 3, \cdots. \quad (18)$$

Thus,

$$\delta_{kl} = (\eta_k, \eta_l)_2 = (\eta_{k,1}, \eta_{l,1}) + (\eta_{k,2}, \eta_{l,2}),$$
$$= \omega_k^{-1}\omega_l^{-1}(\xi_k, H\xi_l) + (\xi_k, \xi_l),$$

from which we obtain (D). Let $x \in E$, $\dot{x} \in S(x)$, so that $\dot{x} = iH^{\frac{1}{2}}g + iAx$ for some $g \in E$. Let

$$f \equiv \begin{pmatrix} -iH^{\frac{1}{2}}x \\ \dot{x} \end{pmatrix}.$$

Then $f = W\eta$ for $\eta = \begin{pmatrix} g \\ x \end{pmatrix}$, and Eq. (16) gives

$$\left\| -iH^{\frac{1}{2}}x - \sum_{k=1}^{n} (\eta_k, f)_2 \eta_{k,1} \right\| \leq |\omega_{n+1}| \, \|\eta\|_2 \quad (19)$$

and

$$\left\| \dot{x} - \sum_{k=1}^{n} (\eta_k, f)_2 \eta_{k,2} \right\| \leq |\omega_{n+1}| \, \|\eta\|_2. \quad (20)$$

Equation (18) gives

$$(\eta_k, f)_2 = (-i\omega_k^{-1} H^{\frac{1}{2}} \xi_k, -iH^{\frac{1}{2}}x) + (\xi_k, \dot{x}),$$
$$= \omega_k^{-1}(\xi_k, Hx) + (\xi_k, \dot{x}) = \alpha_k \omega_k, \quad (21)$$

so that

$$(\eta_k, f)_2 \eta_{k,1} = -i\alpha_k H^{\frac{1}{2}} \xi_k \quad (22)$$

and

$$(\eta_k, f)_2 \eta_{k,2} = \alpha_k \omega_k \xi_k. \quad (23)$$

Statement (E) follows at once from Eqs. (19), (20), (22), and (23). Suppose $x \in E$, $\dot{x} \in \overline{S(x)}$. Then

$$f \equiv \begin{pmatrix} -iH^{\frac{1}{2}}x \\ \dot{x} \end{pmatrix} \in \bar{R}_W,$$

and statement (F) follows from (e) and Eqs. (22) and (23). Let β_k $(k = 1, 2, \cdots, n)$ be any n complex

numbers. Since the $\{\eta_k\}$ is orthonormal, we have

$$\left\| f - \sum_{k=1}^{n} \beta_k \omega_k \eta_k \right\|_2^2$$

$$= \left\| f - \sum_{k=1}^{n} \gamma_k \eta_k \right\|_2^2 + \left\| \sum_{k=1}^{n} (\gamma_k - \beta_k \omega_k) \eta_k \right\|_2^2$$

$$\geq \left\| f - \sum_{k=1}^{n} \gamma_k \eta_k \right\|_2^2 = \|f\|_2^2 - \sum_{k=1}^{n} |\gamma_k|^2, \quad (24)$$

where

$$\gamma_k = (\eta_k, f)_2 = \alpha_k \omega_k \quad (25)$$

by Eq. (21). Equations (18) and (22) through (25) imply (G). Statement (H) follows immediately from (F) and (G). This completes the proof.

Under the hypothesis of Theorem II, the eigenvalues $\omega_k^{\pm}$ may be estimated by the standard extremal methods[3] used to obtain the eigenvalues of the completely continuous Hermitian operator W. In general, these methods will not be convenient, due to the difficulty of computing $H^{\frac{1}{2}}$. The following extremal characterizations of the eigenvalues avoid this problem.

Theorem III: Let the hypothesis of Theorem II hold, and suppose that the positive and negative eigenvalues of Eq. (1) are respectively arranged in the nonincreasing and nondecreasing sequences $\omega_1^+ \geq \omega_2^+ \geq \cdots$, $\omega_1^- \leq \omega_2^- \leq \cdots$. Denote the corresponding eigenvectors by $\xi_k^{\pm}$, $k = 1, 2, 3, \cdots$, where the $\xi_k^{\pm}$ satisfy the appropriate form of Eq. (7). Then

$$\omega_n^+ = \max_{u \in V_n^+} \frac{2 \operatorname{Re} (u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)}, \quad (26)$$

$$\omega_n^- = \min_{u \in V_n^-} \frac{2 \operatorname{Re} (u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)}, \quad (27)$$

where

$$V_n^{\pm} \equiv \{u \mid u \in E^2, \, \omega_k^{\pm}(\xi_k^{\pm}, u_2) + (\xi_k^{\pm}, Hu_1) = 0,$$
$$k = 1, 2, \cdots, n-1\}.$$

We also have

$$\omega_n^+ = \max_{\xi \in U_n^+} F_\xi^+, \quad \omega_n^- = \min_{\xi \in U_n^-} F_\xi^-, \quad (28)$$

where

$$F_\xi^{\pm} = \frac{1}{2} \left\{ \frac{(\xi, iA\xi)}{(\xi, \xi)} \pm \left[\frac{(\xi, iA\xi)^2}{(\xi, \xi)^2} + 4 \frac{(\xi, H\xi)}{(\xi, \xi)} \right]^{\frac{1}{2}} \right\} \quad (29)$$

and

$$U_n^{\pm} \equiv \{\xi \mid \xi \in E, \, (\xi, [H + \omega_k^{\pm} F_\xi^{\pm} I] \xi_k^{\pm}) = 0,$$
$$k = 1, 2, \cdots, n-1\}.$$

Furthermore,

$$\omega_n^+ \geq \sup_{\substack{(\xi, \xi_k^+)=0 \\ k=1,2,\cdots,n-1}} \frac{(\xi, iA\xi)}{(\xi, \xi)},$$

$$\omega_n^- \leq \inf_{\substack{(\xi, \xi_k^-)=0 \\ k=1,2,\cdots,n-1}} \frac{(\xi, iA\xi)}{(\xi, \xi)}, \quad (30)$$

and the number of positive (negative) eigenvalues $\omega_k^+(\omega_k^-)$ is at least as great as the number of positive (negative) eigenvalues of iA (counted as often as their degeneracy).

Proof: Theorem II holds; the operator W of Theorem I is completely continuous and Hermitian on the Hilbert space E^2; and by Theorem I the eigenvalues $\omega_k^\pm$ are precisely the nonzero eigenvalues of the operator W, with associated orthonormal eigenvectors $\eta_k^\pm$ given by

$$\eta_k^\pm = \begin{pmatrix} -i[\omega_k^\pm]^{-1}H^{\frac{1}{2}}\xi_k^\pm \\ \xi_k^\pm \end{pmatrix}, \quad k = 1, 2, 3, \cdots. \quad (31)$$

We have[4]

$$\omega_n^+ = \max_{\eta \in P_n^+} \frac{(\eta, W\eta)_2}{(\eta, \eta)_2}, \quad \omega_n^- = \min_{\eta \in P_n^-} \frac{(\eta, W\eta)_2}{(\eta, \eta)_2}, \quad (32)$$

where

$$P_n^\pm = \{\eta \mid \eta \in E^2, (\eta, \eta_k^\pm)_2 = 0, k = 1, 2, \cdots, n - 1\}$$

and the maximum (minimum) is achieved for $\eta_n^+(\eta_n^-)$. Let $u \in V_n^\pm$. For

$$\eta = \begin{pmatrix} -iH^{\frac{1}{2}}u_1 \\ u_2 \end{pmatrix}, \quad (33)$$

we have $\eta \in P_n^\pm$, so that

$$\omega_n^+ \geq \frac{(\eta, W\eta)_2}{(\eta, \eta)_2} \quad \text{or} \quad \omega_n^- \leq \frac{(\eta, W\eta)_2}{(\eta, \eta)_2}. \quad (34)$$

Since

$$\frac{(\eta, W\eta)_2}{(\eta, \eta)_2} = \frac{2\,\text{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)}, \quad (35)$$

$$\omega_n^\pm = \frac{(\eta_n^\pm, W\eta_n^\pm)_2}{(\eta_n^\pm, \eta_n^\pm)_2}, \quad (36)$$

and $\eta_n^\pm$ is given by Eq. (33) for $u_1 = [\omega_n^\pm]^{-1}\xi_n^\pm$, $u_2 = \xi_n^\pm$, where $u \in V_n^\pm$; Eqs. (26) and (27) follow. Equations (28) are obtained by restricting u to be of the form

$$u = \begin{pmatrix} [F_\xi^\pm]^{-1}\xi \\ \xi \end{pmatrix}$$

in Eqs. (26) and (27), and noting that $\xi \in U^\pm$ implies $u \in V_n^\pm$, $\omega_k^\pm = F_{\xi_k^\pm}^\pm$, and that for $u_1 = [F_\xi^\pm]^{-1}\xi$, $u_2 = \xi$, we have

$$\frac{2\,\text{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)} = F_\xi^\pm. \quad (37)$$

Let $\xi \in E$, $(\xi, \xi_k^+) = 0$ or $(\xi, \xi_k^-) = 0$, $k = 1, 2, \cdots$, $n - 1$. Then

$$\zeta = \begin{pmatrix} 0 \\ \xi \end{pmatrix} \in E^2$$

satisfies $(\eta_k^\pm, \zeta)_2 = 0$, $k = 1, 2, \cdots, n - 1$, i.e., $\zeta \in P_n^\pm$. Now

$$\frac{(\zeta, W\zeta)_2}{(\zeta, \zeta)_2} = \frac{(\xi, iA\xi)}{(\xi, \xi)}, \quad (38)$$

so that Eqs. (32) imply Eqs. (30).

Finally, suppose we have m positive (negative) eigenvalues $\omega_k^+(\omega_k^-)$ with associated eigenvectors $\eta_k^\pm$ given by (31), and n positive (negative) eigenvalues $\lambda_k^+(\lambda_k^-)$ of iA with associated orthonormal eigenvectors $\psi_k^\pm$, and suppose $m < n$. The existence of a further eigenvalue ω_{m+1}^+ or ω_{m+1}^- is assured by Eqs. (32), provided we demonstrate the existence of a vector $\zeta^+(\zeta^-) \in P_{m+1}^+(P_{m+1}^-)$ for which

$$\frac{(\zeta^+, W\zeta^+)_2}{(\zeta^+, \zeta^+)_2} > 0, \quad \frac{(\zeta^-, W\zeta^-)_2}{(\zeta^-, \zeta^-)_2} < 0. \quad (39)$$

But $m < n$ implies that there exists a nonzero vector $y^\pm$ in the n-dimensional space spanned by $\{\psi_k^\pm\}_{k=1}^n$ such that $(y^\pm, \xi_k^\pm) = 0$, $k = 1, 2, \cdots, m$. For $\zeta^\pm \equiv \begin{pmatrix} 0 \\ y^\pm \end{pmatrix} \in E^2$, we have $\zeta^\pm \in P_{m+1}^\pm$, and

$$\frac{(\zeta^+, W\zeta^+)_2}{(\zeta^+, \zeta^+)_2} = \frac{(y^+, iAy^+)}{(y^+, y^+)} = \frac{\sum_{k=1}^n \lambda_k^+ |\alpha_k^+|^2}{\sum_{k=1}^n |\alpha_k^+|^2} > 0, \quad (40)$$

$$\frac{(\zeta^-, W\zeta^-)_2}{(\zeta^-, \zeta^-)_2} = \frac{(y^-, iAy^-)}{(y^-, y^-)} = \frac{\sum_{k=1}^n \lambda_k^- |\alpha_k^-|^2}{\sum_{k=1}^n |\alpha_k^-|^2} < 0, \quad (41)$$

where $y^\pm = \sum_{k=1}^n \alpha_k^\pm \psi_k^\pm$.

Theorem IV: Under the hypothesis of Theorem III we have the following maximum–minimum principles:

$$\omega_n^+ = \min_{\substack{\psi_k \in Q \\ k=1,2,\cdots,n-1}} \max_{\substack{(u,\psi_k)_2=0 \\ k=1,2,\cdots,n-1}} \frac{2\,\text{Re}\,(u_1, Hu_2)+(u_2, iAu_2)}{(u_1, Hu_1)+(u_2, u_2)}, \quad (42)$$

$$\omega_n^- = \max_{\substack{\psi_k \in Q \\ k=1,2,\cdots,n-1}} \min_{\substack{(u,\psi_k)_2=0 \\ k=1,2,\cdots,n-1}}$$
$$\times \frac{2\,\text{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)}, \quad (43)$$

where

$$Q \equiv \{\psi \mid \psi \in E^2, \psi_1 \in R_{H^{\frac{1}{2}}}\}$$

or

$$Q \equiv \{\psi \mid \psi \in E^2, \psi_1 \in R_H\}.$$

Proof: For the particular choice of

$$\psi_k = \begin{pmatrix} H\xi_k^+ \\ \omega_k^+\xi_k^+ \end{pmatrix}, \quad k = 1, 2, \cdots, n - 1,$$

---

[4] Reference 2, p. 237.

we have $(u, \psi_k)_2 = 0$ for $k = 1, 2, \cdots, n - 1$, if and only if $u \in V_n^+$. Equation (26) then implies

$$\omega_n^+ \geq \min_{\substack{\psi_k \in Q \\ k=1,2,\cdots,n-1}} \max_{\substack{(u,\psi_k)_2=0 \\ k=1,2,\cdots,n-1}}$$

$$\times \frac{2\,\mathrm{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)}. \quad (44)$$

Let $\{\psi_k\}_{k=1}^{n-1}$ be any $n - 1$ vectors in Q. Then $\psi_k = T\Psi_k$ for some $\Psi_k \in E^2$, where

$$T \equiv \begin{pmatrix} iH^{\frac{1}{2}} & 0 \\ 0 & I \end{pmatrix},$$

for $k = 1, 2, \cdots, n - 1$. There exists a nonzero vector

$$\eta = \sum_{k=1}^{n} \alpha_k \eta_k^+$$

in the n-dimensional subspace of E^2 spanned by the n-orthonormal eigenvectors $\{\eta_k^+\}_{k=1}^{n}$ of W such that $(\eta, \Psi_k)_2 = 0$ for $k = 1, 2, \cdots, n - 1$. It follows from Eq. (31) that η can be written in the form of Eq. (33), with

$$u_1 = \sum_{k=1}^{n} \alpha_k (\omega_k^+)^{-1} \xi_k^+, \quad u_2 = \sum_{k=1}^{n} \alpha_k \xi_k^+,$$

and for this u we have

$$(u, \psi_k)_2 = (u, T\Psi_k)_2 = (T^*u, \Psi_k)_2 = (\eta, \Psi_k)_2 = 0,$$
$$k = 1, 2, \cdots, n - 1. \quad (45)$$

Equation (35) gives

$$\frac{2\,\mathrm{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)} = \frac{(\eta, W\eta)_2}{(\eta, \eta)_2}$$

$$= \frac{\sum_{k=1}^{n} |\alpha_k|^2 \omega_k^+}{\sum_{k=1}^{n} |\alpha_k|^2} \geq \omega_n^+, \quad (46)$$

so that

$$\max_{\substack{(u,\psi_k)_2=0 \\ k=1,2,\cdots,n-1}} \frac{2\,\mathrm{Re}\,(u_1, Hu_2) + (u_2, iAu_2)}{(u_1, Hu_1) + (u_2, u_2)} \geq \omega_n^+. \quad (47)$$

Equation (42) follows immediately from Eqs. (44) and (47). The proof of Eq. (43) is analogous.

III. GENERAL SOLUTION FOR A FINITE-DIMENSIONAL SPACE

Theorem V: Let E be an inner product space and H and iA linear Hermitian operators with domain and range in E, and let H be bounded on D_H. Let S denote the set of all $\xi(t)$ satisfying the conditions $\xi(t) \in D_H$ for $t \in (O, T)$ where $T \leq \infty$; $\xi(t)$ is twice differentiable on (O, T), $\dot{\xi}(t) \in D_A$, and $\ddot{\xi}(t) \in E$ for $t \in (O, T)$. Then for each $\xi(t) \in S$ satisfying

$$\ddot{\xi} + A\dot{\xi} + H\xi = 0, \quad t \in (O, T), \quad (48)$$

we have

$$(\dot{\xi}, \dot{\xi}) + (\xi, H\xi) = \text{constant on } (O, T). \quad (49)$$

If $H \geq 0$ on D_H, then there exists at most one $\xi(t) \in S$ satisfying Eq. (48) and the boundary conditions

$$\lim_{t \to 0+} \xi(t) = x, \quad \lim_{t \to 0+} \dot{\xi}(t) = \dot{x}, \quad x, \dot{x} \in E. \quad (50)$$

Proof: Let K be a bounded linear Hermitian operator with $D_K \subset E$, $R_K \subset E$. Suppose $f(t) \in D_K$ for $t \in (O, T)$, $f'(t)$ exists on (O, T), and $f(t) \in E$. Now

$$\frac{d}{dt}(f, Kf) = \lim_{\Delta t \to 0} \left\{ \frac{(f', Kf') - (f, Kf)}{\Delta t} \right\},$$

$$= \lim_{\Delta t \to 0} \left\{ \left(\frac{[f' - f]}{\Delta t}, Kf' \right) + \left(Kf, \frac{[f' - f]}{\Delta t} \right) \right\},$$

$$= (\dot{f}, Kf) + (Kf, \dot{f}), \quad (51)$$

where $f' \equiv f(t + \Delta t)$, since $f' \to f$ and $(f' - f)/\Delta t \to \dot{f}$ as $\Delta t \to 0$. (By the existence of $\dot{f}$,

$$\lim_{\Delta t \to 0} \left\| \frac{f' - f}{\Delta t} - \dot{f} \right\| = 0$$

is implied.) Hence for $\xi(t) \in S$,

$$\frac{d}{dt}\{(\dot{\xi}, \dot{\xi}) + (\xi, H\xi)\}$$

$$= (\ddot{\xi}, \dot{\xi}) + (\dot{\xi}, \ddot{\xi}) + (\dot{\xi}, H\xi) + (H\xi, \dot{\xi}),$$

$$= (\ddot{\xi} + H\xi, \dot{\xi}) + (\dot{\xi}, \ddot{\xi} + H\xi),$$

$$= (-A\dot{\xi}, \dot{\xi}) - (\dot{\xi}, A\dot{\xi}) = 0,$$

where we have used Eq. (48). Thus Eq. (49) holds. Let $\xi_1(t)$ and $\xi_2(t)$ both be in S and satisfy the boundary conditions of Eq. (50). Then $\eta(t) \equiv \xi_1(t) - \xi_2(t) \in S$, and we have

$$\lim_{t \to 0+} \eta(t) = 0, \quad \lim_{t \to 0+} \dot{\eta}(t) = 0. \quad (52)$$

Equations (49) and (52) imply $(\dot{\eta}, \dot{\eta}) + (\eta, H\eta) \equiv 0$ on (O, T), so that $H \geq 0$ leads to $\dot{\eta} = 0$ on (O, T). But then

$$\frac{d}{dt}\|\eta\|^2 = \frac{d}{dt}(\eta, \eta) = (\dot{\eta}, \eta) + (\eta, \dot{\eta}) = 0,$$

and we obtain $\|\eta\| = \text{constant}$ on (O, T). Equation (52) implies $\|\eta\| = 0$, so that $\xi_1(t) = \xi_2(t)$.

Lemma: Let H and iA be Hermitian operators defined on and into a finite-dimensional unitary space E. Suppose $H \geq 0$, and let P be the projection operator onto N_H. Then for each $y \in N_H$, there exists $\delta \in E$ and a unique $\eta \in N_H$ such that

$$A\eta + H\delta = 0 \quad \text{and} \quad y = \eta + PA\delta. \quad (53)$$

Proof: Define $B \equiv PAP$ on N_H. Then N_H is an invariant subspace for B, and $N_H = R_B \oplus N_B$. Let $u \in N_B$. Then $PAu = 0$, so that $Au \in R_H$, since $E = R_H \oplus N_H$. Therefore there exists $\psi \in E$ such that $Au + H\psi = 0$. Let $\Delta \equiv (I - P)\psi$. Then $Au + H\Delta = 0$, and Δ is unique. [$Au + H\Delta_1 = 0$ and $Au + H\Delta_2 = 0$ implies $H(\Delta_1 - \Delta_2) = 0$ so that $\Delta_1 = \Delta_2$ for $\Delta_1, \Delta_2 \in R_H$.] Thus we may define the operator K on N_B by $Ku \equiv u + P_0 A \Delta$, where P_0 is the projection operator onto N_B, and we have $Au + H\Delta = 0$, $\Delta \in R_H$. K is clearly linear, and N_B is an invariant subspace for K. Suppose $u \in N_B$ and $Ku = 0$. Then

$$0 = (u, Ku) = (u, u) + (u, P_0 A \Delta),$$
$$= (u, u) - (Au, \Delta) = (u, u) + (\Delta, H\Delta).$$

But $H \geq 0$ then implies $u = 0$. Thus $R_K = N_B$, so that for each $y \in N_H$ there exists an $\eta \in N_B$ and a $\Delta \in R_H$ such that

$$P_0 y = \eta + P_0 A \Delta \quad \text{and} \quad A\eta + H\Delta = 0. \quad (54)$$

Now $(I - P_0)[y - PA\Delta] \in R_B$, so that there exists a $\Delta_0 \in N_H$ for which

$$PA\Delta_0 = (I - P_0)[y - PA\Delta],$$
$$= y - P_0 y - PA\Delta + P_0 A \Delta. \quad (55)$$

Combining (54) and (55), we obtain $y = \eta + PA \times (\Delta + \Delta_0)$. Setting $\delta \equiv \Delta + \Delta_0$, we have $y = \eta + PA\delta$ and $A\eta + H\delta = 0$. The uniqueness of η follows easily from Eqs. (53). Suppose $y = \eta_1 + PA\delta_1 = \eta_2 + PA\delta_2$ for $\eta_1, \eta_2 \in N_H$, and $A\eta_1 + H\delta_1 = 0 = A\eta_2 + H\delta_2$. Let $\eta \equiv \eta_1 - \eta_2$, $\delta \equiv \delta_1 - \delta_2$. Then $0 = \eta + PA\delta$ and $A\eta + H\delta = 0$, so that

$$0 = (\eta, \eta) + (\eta, PA\delta) = (\eta, \eta) - (A\eta, \delta),$$
$$= (\eta, \eta) + (\delta, H\delta).$$

Therefore $H \geq 0$ implies $\eta = \eta_1 - \eta_2 = 0$.

Theorem VI: Let H and iA be Hermitian operators defined on and into a finite-dimensional unitary space E, and suppose $H \geq 0$. Then the time-dependent system

$$\ddot{\xi} + A\dot{\xi} + H\xi(t) = 0, \quad 0 \leq t < \infty, \quad (56)$$

with the initial conditions $\xi(0) = x$, $\dot{\xi}(0) = \dot{x}$, $x, \dot{x} \in E$ admits of the unique solution

$$\xi(t) = \sum_k \alpha_k \xi_k e^{i\omega_k t} + t\eta + \delta + \psi, \quad (57)$$

where the ξ_k are the members of the finite (at most 2 dim E elements) sequence of eigenvectors of Eq. (1) obtained in Theorem II with nonzero eigenvalues ω_k. The numbers α_k are given by

$$\alpha_k = \omega_k^{-2}\{(\xi_k, Hx) + \omega_k(\xi_k, -i\dot{x})\}, \quad (58)$$

$\delta \in E$, $\psi \in N_H$, and $\eta \in N_H$ is uniquely determined by x and $\dot{x}$.

Given initial conditions $x, \dot{x} \in E$, the solution will be stable (i.e., $\eta = 0$) if and only if $P(\dot{x} + Ax) \in PA(N_H)$, where P is the projection operator onto N_H. The system is stable for arbitrary initial conditions if and only if $PA(N_H) = N_H$, or if and only if $\dim N_A \cap N_H = 0$ and $\dim A(N_H) \cap R_H = 0$.

Proof: If $\|A\| + \|H\| = 0$, then $\xi(t) = t\dot{x} + x$. Suppose $\|A\| + \|H\| > 0$. Then the hypothesis of Theorem II is satisfied; and since $H \geq 0$ and $\dim E < \infty$, $R_H = R_{H^{\frac{1}{2}}}$, $N_H = N_{H^{\frac{1}{2}}}$, and $E = R_H \oplus N_H$. Let $x, \dot{x}$ be given. By the previous lemma, there exists a unique $\eta \in N_H$ and a $\delta \in E$ such that

$$A\eta + H\delta = 0, \quad P(\dot{x} + Ax) = \eta + PA\delta. \quad (59)$$

Therefore $\dot{x} - \eta + A(x - \delta) \in R_H = R_{H^{\frac{1}{2}}}$, so that there is a $g \in E$ for which

$$-H^{\frac{1}{2}}g = \dot{x} - \eta + A(x - \delta). \quad (60)$$

We define

$$y \equiv x - \delta, \quad \dot{y} \equiv i(\eta - \dot{x}). \quad (61)$$

Then $\dot{y} = (iH^{\frac{1}{2}}g + iAy) \in S(y)$ [see Theorem II(E)] and Theorem II(F) and Eqs. (7) and (10) lead to

$$H^{\frac{1}{2}}\left(y - \sum_k \alpha_k \xi_k\right) = 0, \quad (62)$$

$$\dot{y} = \sum_k \alpha_k \omega_k \xi_k, \quad (63)$$

where

$$\alpha_k = \omega_k^{-2}\{(\xi_k, Hy) + \omega_k(\xi_k, \dot{y})\}. \quad (64)$$

Set $\psi \equiv y - \sum_k \alpha_k \xi_k$. Then by Eq. (62), $\psi \in N_H$ and we have

$$x = \sum_k \alpha_k \xi_k + \delta + \psi. \quad (65)$$

Equations (61) and (63) yield

$$\dot{x} = i\sum_k \alpha_k \omega_k \xi_k + \eta. \quad (66)$$

Thus Eq. (57) satisfies the required initial conditions. Since the ξ_k are all eigenvectors of Eq. (1) and $H\psi = 0$, Eq. (57) will satisfy (56), provided $\zeta(t) \equiv t\eta + \delta$ satisfies (56). But this follows from $H\eta = 0$ and $A\eta + H\delta = 0$. Thus Eq. (57) is a solution of Eq. (56) satisfying the required initial conditions; by Theorem V, it is the unique solution. Now Eqs. (61) and (64) give

$$\alpha_k = \omega_k^{-2}\{(\xi_k, Hx) + \omega_k(\xi_k, -i\dot{x})\} + \omega_k^{-3}\beta_k, \quad (67)$$

where

$$\beta_k \equiv \omega_k^2(\xi_k, i\eta) - \omega_k(\xi_k, H\delta).$$

$A\eta + H\delta = 0$ implies

$$\beta_k = (\omega_k^2 \xi_k, i\eta) + (\omega_k \xi_k, A\eta),$$
$$= i(\omega_k^2 \xi_k - \omega_k i A \xi_k, \eta),$$
$$= i(H\xi_k, \eta) = i(\xi_k, H\eta) = 0, \qquad (68)$$

since $\eta \in N_H$. Thus Eq. (58) is valid.

Equations (59) and $\eta = 0$ imply $P(\dot{x} + Ax) = PA\delta$ for $\delta \in N_H$, i.e., $P(\dot{x} + Ax) \in PA(N_H)$. Conversely, suppose $P(\dot{x} + Ax) = PA\delta$ for $\delta \in N_H$. Then Eqs. (59) hold for $\eta = 0$; and by the previous lemma, this η is unique. Thus $\eta = 0$ if and only if $P(\dot{x} + Ax) \in PA(N_H)$. The system is therefore stable for arbitrary x and $\dot{x}$ if and only if $PA(N_H) = N_H$. Now $N_H = PA(N_H) \oplus N$, where

$$N = \{x \mid x \in N_H, PAx = 0\} = \{x \mid x \in N_H, Ax \in R_H\},$$

so that $PA(N_H) = N_H$ if and only if $\dim N = 0$. Since $A(N) = A(N_H) \cap R_H$, $N_A \cap N_H \subset N$, and $\dim A(N) = 0$ imply $N = N_A \cap N_H$, we see that $PA(N_H) = N_H$ if and only if the subspaces $A(N_H) \cap R_H$ and $N_A \cap N_H$ both have zero dimension.

Boundary-Value Problems for the Linearized and Weakly Nonlinear Boltzmann Equation

YOUNG-PING PAO

Division of Applied Mathematics, Brown University, Providence, Rhode Island

The boundary-value problems of the linearized and nonlinear Boltzmann equations for a gas contained between two parallel plates are considered. First, the existence and uniqueness for the solution to the linear problem are proved for some function spaces which are convenient to the nonlinear analysis; the proof is based on a previous result due to Cercignani, who used a different function space. Then the existence and uniqueness for the solution to the nonlinear problem can be proved for small boundary data by using the implicit function theorem in functional analysis. With certain classes of boundary data, it is shown, for linear as well as nonlinear problems, that all the moments of the distribution function are continuous in the position variable. The solution to the linearized Boltzmann equation is shown to approximate the solution to the nonlinear Boltzmann equation in the limit of small boundary data. The results apply to hard-sphere molecules and hard-potential molecules, i.e., angle-cutoff power-law molecules with force exponents greater than 5.

I. INTRODUCTION

THE present paper is concerned with the boundary-value problems of the linearized and nonlinear Boltzmann equations for a gas contained between two parallel plates. We first prove, for the linearized problems, the existence and uniqueness in the normed spaces E and V (cf. Sec. III). One of the advantages in using the spaces E and V lies in the possibility of formulating the nonlinear problems in terms of the implicit function theorem in functional analysis. As a consequence, we prove the existence and uniqueness for the nonlinear problems under the condition that the boundary data are small in the norm of E.

The results in this paper require no restriction on the distance between the plates and, therefore, are valid for arbitrary Knudsen numbers. We consider hard-sphere molecules and hard-potential molecules, i.e., angle-cutoff power-law molecules with force exponent $s > 5$. The boundary conditions are given for the distribution function of the re-emitted molecules at the plates. The proof of the existence and uniqueness for the solution to the linear problem is based on a previous result due to Cercignani,[1] who proved the existence and uniqueness for a different function space Ω (cf. the next section).

II. LINEAR PROBLEMS

If we denote the distribution function F in terms of its perturbation $f(c, x)$ from the base Maxwellian f_0 as

$$F = f_0 + f_0^{\frac{1}{2}}f,$$

the linearized Boltzmann equation can be written as

$$c_x(\partial f/\partial x) = Kf - v(c)f, \qquad (1)$$

where the explicit form of the operator K and the function $v(c)$ can be found in Grad.[2] The boundary conditions are given at the plates by

$$f(c, d) = g_-(c), \quad (c_x < 0), \qquad (2a)$$

$$f(c, -d) = g_+(c), \quad (c_x > 0), \qquad (2b)$$

and the physical quantities, such as density, velocity, temperature, stress, etc., are given by the moments of $f_0^{\frac{1}{2}}f$.

In a recent paper,[1] Cercignani introduces the operator H defined by

$$Hf = Kf + \lambda v(c)f,$$

where the positive number λ is so chosen that H is positive definite on the Hilbert space of functions square-integrable in c. Equation (1) can be formally integrated along the characteristics to give

$$f = g + UHf, \qquad (3)$$

where

$$g = g_\pm \exp\left[-(\lambda + 1)v(d + x \operatorname{sgn} c_x)/|c_x|\right], \qquad (4)$$

$$Uh = \frac{1}{c_x}\int_{-d\operatorname{sgn} c_x}^{x} dy\, h(c, y)\exp\left[-\frac{(\lambda + 1)v\,|x - y|}{|c_x|}\right], \qquad (5)$$

and the $\pm$ signs in (4) refer to $c_x > 0$ and $c_x < 0$, respectively. Using an Hilbert space, which is denoted here by Ω, with the inner product

$$(\alpha, \beta) = \int_{-d}^{d} dx \int dc\, \rho(c)\alpha(c, x)\beta(c, x),$$

$$\rho(c) = [(\lambda + 1)^2 v^2(c) + \pi^2 c_x^2/d^2]^{\frac{1}{2}},$$

[1] C. Cercignani, J. Math. Phys. **8**, 1653 (1967).

[2] H. Grad, *Rarefied Gas Dynamics*, J. A. Laurmann, Ed. (Academic Press Inc., New York, 1963), Vol. I.

Cercignani has been able to prove that the operator UH in Eq. (3) is a contracting mapping on Ω, and consequently arrived at the following theorem.

Theorem: Equation (3) has one and only one solution in Ω for any given g in Ω, and $\|f\|_\Omega < \alpha_1 \|g\|_\Omega$ for some positive constant α_1, where $\| \ \|_\Omega$ is the norm in Ω.

However, the space Ω is not suitable to the study of nonlinear problems, and in what follows we shall find some other integral equation [Eq. (6)] which is equivalent to Eq. (3) and which is more convenient to work with in connection with some more suitable function spaces.

First let us examine if the solution to Eq. (3) also satisfies the original linearized Boltzmann equation (1). To this end we note that Kf is square-integrable in c and x if f is in Ω (see Ref. 2, p. 43). Hence Hf is an integrable function of c and x if c is restricted to a finite sphere S of radius R.

Then, according to Fubini's theorem, Hf is integrable in x for almost all c in S, which, in conjunction with Eq. (3), shows that f is absolutely continuous in x for almost all c in S. We can then differentiate Eq. (3) with respect to x and conclude that f satisfies Eq. (1) for almost all c and x if c is in S. But since the radius R of the sphere S is arbitrary, we assert that f satisfies (1) almost everywhere on the product space of the whole ranges of c and x. It is easy to see that f also satisfies the boundary conditions (2) for almost all c.

On the other hand, since f is absolutely continuous in x for almost all c, we can integrate Eq. (1) along the characteristics, now using K instead of H. The result is

$$f = g_1 + U_1 Kf, \qquad (6)$$

where g_1 and U_1 are given by (4) and (5) with λ taken to be zero.

We thus conclude that if $f \in \Omega$ and f satisfies Eq. (3) almost everywhere, then f satisfies Eq. (6) almost everywhere. Conversely, if $f_1 \in \Omega$ and f_1 is a solution to Eq. (6), we can reverse the direction of the above argument and assert that f_1 also satisfies Eq. (3) almost everywhere. Hence we have established a certain equivalence between Eqs. (3) and (6), and as a consequence we obtain the following result from the previously stated existence and uniqueness theorem of Cercignani.

Theorem 1: If $g \in \Omega$, then Eq. (6) has a unique solution in Ω and $\|f\|_\Omega < \alpha_1 \|g\|_\Omega$.

This theorem will now be our point of departure, and it should be mentioned that in this theorem, g is used instead of g_1.

III. EXISTENCE AND UNIQUENESS FOR SPACES V AND E

Working with Eq. (6), it is possible to establish existence and uniqueness for the linear problems in two function spaces V and E which suit the need of our analyses concerning nonlinear problems and the continuity of moments of the distribution function. We define V and E by the norms

$$\|f\|_V^2 = \max_{-d \leq x \leq d} \int dc \, \rho(c) f^2(c, x),$$
$$\|f\|_E = \max_{x,c} (1 + c^2) |f(c, x)|. \qquad (7)$$

First consider the case where the inhomogeneous function g_1 in Eq. (6) belongs to V. This implies that $g \in \Omega$, and consequently Eq. (6) has a unique solution f in Ω, according to Theorem 1. We now prove that this solution f, in fact, belongs to V. To this end, we shall obtain an estimate on the quantity

$$\|U_1 Kf\|_V^2 = \max_x \int dc \, \rho(c) c_x^{-2}$$
$$\times \left\{ \int_{-d \, \text{sgn} \, c_x}^x dy \, [Kf] \exp\left[-\frac{\nu |x - y|}{|c_x|} \right] \right\}^2.$$

Using the inequalities

$$[Kf](c, y)[Kf](c, \eta) \leq \tfrac{1}{2}\{[Kf]^2(c, y) + [Kf]^2(c, \eta)\},$$
$$\left| \frac{1}{c_x} \int_{-d \, \text{sgn} \, c_x}^x dy \, \exp\left[-\nu |x - y| \, |c_x|\right] \right| \leq \frac{1}{\nu(c)},$$

we obtain the result

$$\|U_1 Kf\|_V^2 \leq \int dc \, \rho(c)[c_x \nu]^{-1}$$
$$\times \int_{-d \, \text{sgn} \, c_x}^x dy \, [Kf]^2 \exp\left[-\frac{\nu |x - y|}{|c_x|} \right]$$
$$\leq \int_{-d}^d dy \int dc_1 \int dc_2 \, I(c_1, c_2; t) f(c_1, y) f(c_2, y), \qquad (8)$$

where $t = |x - y|$ and the function I is given in terms of the kernel $K(c, c_1)$ of the integral operator K by

$$I = \int dc \, \frac{\rho(c)}{|c_x| \, \nu(c)} K(c, c_1) K(c, c_2) \exp\left[-\frac{\nu(c) t}{|c_x|} \right]. \qquad (9)$$

For the function I, it is possible to obtain the following estimate:

$$S(t) = \int I^2(c_1, c_2; t)[\rho(c_1)\rho(c_2)]^{-1} \, dc_1 \, dc_2$$
$$\leq [a + b \ln t]^2, \qquad (10)$$

where a and b are positive constants. A detailed proof of estimate (10) is given in the Appendix. The term $\ln t$ is, of course, connected with the singularity at $c_x = 0$ of the integrand in the right-hand term of (9).

With the estimate in (10), we obtain from inequality (8)

$$\|U_1 K f\|_V^2 < \int_{-d}^{d} dy \, [a + b \ln |x - y|] \, \|f\|_V^2, \quad (11)$$

on using the Schwartz inequality for the double integral over c_1 and c_2. On the other hand, we have from Eq. (6)

$$\|f\|_V^2 \le \|g_1\|_V^2 + \|U_1 K f\|_V^2,$$

which, together with (11), yields

$$\|f\|_V^2 < \|g_1\|_V^2 + \int_{-d}^{d} dy(a + b \ln |x - y|)$$
$$\times \left\{ \|g_1\|_V^2 + \int_{-d}^{d} d\eta(a + b \ln |y - \eta|) \, \|f\|_V^2 \right\}.$$

From the fact that $f \in \Omega$ and $g_1 \in E$, it is not difficult to see that, in the last inequality, the order of integrations can be interchanged to give

$$\|f\|_V^2 < \alpha_2 \|g_1\|_V^2 + \alpha_3 \|f\|_\Omega^2$$

for some positive constants α_2 and α_3, and this shows that $f \in V$ if $g_1 \in V$. Furthermore, we know from Theorem 1 that $\|f\|_\Omega < \alpha_1 \|g\|_\Omega$, and it is easily seen that $\|g\|_\Omega < 2d \|g_1\|_V$. We thus have $\|f\|_V < \beta \|g_1\|_V$, where the positive constant β is independent of g_1. From these results and the fact that $f \in V$ implies $f \in \Omega$, we obtain the following theorem.

Theorem 2: If $g_1 \in V$, then Eq. (6) has a unique solution f in V and $\|f\|_V < \beta \|g_1\|_V$.

Now using Theorem 2 as a new starting point, we can prove a similar result for the function space E, which is defined in the beginning of this section.

Consider the case $g_1 \in E$, which obviously implies that $g_1 \in V$, and consequently Eq. (6) has a unique solution f in V. We shall prove that this solution, in fact, belongs to E. For this purpose, it is useful to note the following norms and inequalities given by Grad (cf. Ref. 2, pp. 43–44):

$$N_1[f] = \max_c |f|,$$
$$N_3^{(r)}[f] = \max_c (1 + c^2)^{\frac{1}{2}r} |f|,$$
$$N_1[Kf] < m_1 \left(\int f^2 \, dc \right)^{\frac{1}{2}},$$
$$N_3^{(r+1)}[f] < m_2 N_3^{(r)}[f].$$

Applying the last two inequalities to Eq. (6), we have

$$|f| < |g_1| + m_3 \|g\|_V,$$
$$(1 + c^2)^{\frac{1}{2}} |f| < (1 + c^2)^{\frac{1}{2}} |g_1| + m_4[\max_c |f|],$$

$$(1 + c^2) |f| < (1 + c^2) |g_1| + m_4[\max_c (1 + c^2)^{\frac{1}{2}} |f|],$$

from which we obtain, by successive substitutions, $\|f\|_E < m \|g_1\|_E$ for some positive number m. Therefore we arrive at the following assertion.

Theorem 3: If $g_1 \in E$, then Eq. (6) has a unique solution f in E and $\|f\|_E < m \|g_1\|_E$.

It is quite clear that analogous results can be obtained for spaces similar to E, e.g., spaces with the norm

$$\|f\| = \max_{x,c} (1 + c^2)^{\frac{1}{2}r} |f|$$

for any $r \geq 2$.

IV. CONTINUITY OF ALL MOMENTS

With the result in Theorem 3, it is not difficult to show that all the moments of the distribution function are continuous in x if the boundary data $g_\pm(c)$ belong to E. First let us consider the difference

$$\Delta(x_1, x_2) = \int dc \, |f(c, x_1) - f(c, x_2)|^2.$$

Note that $g_\pm(c) \in E$ implies $g_1 \in E$ and, therefore, $f \in E$. Consequently, $(1 + c^2)^{\frac{1}{2}} Kf$ is bounded (say by A_1); and for $x_1 > x_2$, we obtain from Eq. (6)

$$\Delta(x_1, x_2) \le \int dc \, |g_1(c, x_1) - g_1(c, x_2)|^2$$
$$+ A_1 \int_{c_x > 0} dc(1 + c^2)^{-3} \left| c_x^{-1} \int_{-d}^{x} dy \left[\exp\left(-\frac{\nu |x_2 - y|}{|c_x|} \right) \right. \right.$$
$$\left. \left. - \exp\left(-\frac{\nu |x_1 - y|}{|c_x|} \right) \right] \right|^2$$
$$+ A_1 \int_{c_x > 0} dc(1 + c^2)^{-3} \left| c_x^{-1} \int_{x_2}^{x_1} dy \exp\left(-\frac{\nu |x_1 - y|}{|c_x|} \right) \right|^2$$
$$+ A_1 \int_{c_x < 0} dc(1 + c^2)^{-3} \left| c_x^{-1} \int_{x_1}^{d} dy \left[\exp\left(-\frac{\nu |x_1 - y|}{|c_x|} \right) \right. \right.$$
$$\left. \left. - \exp\left(-\frac{\nu |x_2 - y|}{|c_x|} \right) \right] \right|^2$$
$$+ A_1 \int_{c_x < 0} dc(1 + c^2)^{-3} \left| c_x^{-1} \int_{x_2}^{x_1} dy \exp\left(-\frac{\nu |x_2 - y|}{|c_x|} \right) \right|^2.$$

Using dominated convergence, it is not difficult to see that all the integrals above tend to zero as $x_2 \to x_1$. Similar results hold for $x_2 > x_1$. Therefore, we have

$\Delta(x_1, x_2) \to 0$ as $|x_2 - x_1| \to 0$. Finally, we note that

$$\left| \int dc\, P(c) f_0^{\frac{1}{2}} [f(c, x_1) - f(c, x_2)] \right|^2$$

$$\leq \Delta(x_1, x_2) \int dc\, P^2(c) f_0,$$

and arrive at the following conclusion.

Theorem 4: If $g_{\pm}(c) \in E$, then all the moments $\bar{P}(x)$ given by

$$\bar{P}(x) = \int dc\, P(c) f_0^{\frac{1}{2}} f$$

are continuous in x for all polynomials $P(c)$.

This theorem implies that all the physical quantities such as density, velocity, temperature, stresses, flux, etc., are continuous functions of x. It should be mentioned that the assertion in Theorem 4 remains valid under the condition that $g_{\pm}(c) \in V$; the proof is quite similar to that of Theorem 4.

V. NONLINEAR BOLTZMANN EQUATION: EXISTENCE AND UNIQUENESS

In this section we prove the existence and uniqueness for the boundary-value problems of the nonlinear Boltzmann equation under certain conditions on the boundary data. With the result in Theorem 3, the proof is a simple application of the implicit function theorem using some of the nonlinear estimates obtained by Grad.[3] Following the notation in Ref. 3, we write the nonlinear Boltzmann equation as

$$c_x(\partial f / \partial x) = -\nu f + Kf + \nu\Gamma(f, f); \quad (12)$$

and for the nonlinear operator Γ, the following estimate

$$N_3^{(r)}[\Gamma(f, g)] < \mu N_3^{(r)}[f] N_3^{(r)}[g] \quad (13)$$

is obtained by Grad,[3] where μ is a positive constant.

We now consider the boundary-value problem of Eq. (12) for a gas contained between two parallel plates. The boundary conditions have the same form as (2a) and (2b). Equation (12) can again be formally integrated along the characteristics to yield

$$f = g_1 + U_1 Kf + \nu U_1 \Gamma(f, f). \quad (14)$$

We assume that $g_1 \in E$ and introduce the nonlinear operator Λ as

$$\Lambda f = f - U_1 Kf - \nu U_1 \Gamma(f, f), \quad (15)$$

in terms of which Eq. (14) becomes $\Lambda f - g_1 = 0$. The Frechet differential of Λ at f may be computed

by first noting that

$$\Gamma(f + h, f + h) - \Gamma(f, f) = 2\Gamma(f, h) + \Gamma(h, h). \quad (16)$$

On the other hand, we can apply estimate (13) to U_1 to obtain

$$\|\nu U_1 \Gamma(f, g)\|_E < \mu \|f\|_E \|g\|_E. \quad (17)$$

In view of the last inequality and Eq. (16), we see that the Frechet differential D of Λ at f is

$$D(f, h) = (I - U_1 K)h - 2\nu U_1 \Gamma(f, h),$$

where I is the identity operator and h is the increment of f. Inequality (17) clearly shows that $D(f, h)$ is continuous in f. Furthermore, the Frechet differential D is equal to $I - U_1 K$ at $f = 0$, and consequently Theorem 3 indicates that D has a bounded inverse at $f = 0$. Summarizing all these results, we conclude that Λ satisfies all the hypotheses of the implicit function theorem, e.g., the version given by Liusternik and Sobolev.[4] Therefore, we arrive at the following assertion.

Theorem 5: There exists a $\tau > 0$ such that the nonlinear Equation (14) has a unique solution f in E for all g_1 which satisfies the condition $\|g_1\|_E < \tau$. Moreover, $\|f\|_E \to 0$ as $\|g_1\|_E \to 0$.

Since the above theorem is a local result in the sense that g_1 is restricted to a small neighborhood of the origin of E, the nonlinear equation (14) should be understood as weakly nonlinear. We point out here that Theorem 5 also gives a justification for the linearized Boltzmann equation. In fact, this justification may be stated in the following precise form.

Theorem 6: Let f and f^* be, respectively, the solutions to Eq. (14) and Eq. (6) with the same g_1. Then, as $\|g_1\|_E \to 0$, we have

$$\|f - f^*\|_E / \|f\|_E \to 0.$$

The proof is quite simple. We consider the difference

$$(I - U_1 K)(f - f^*) = \nu U_1 \Gamma(f, f),$$

which, according to Theorem 3 and inequality (17), yields

$$\|f - f^*\|_E < \mu \|f\|_E^2$$

and completes the theorem.

In closing this section, we point out that for the nonlinear solution, just as is the case for linear problems, all the moments of the distribution function are

[3] H. Grad, New York University, Courant Institute Report MF-41, 1964, p. 22; presented at the American Mathematical Society Symposium on Application of Partial Differential Equations in Mathematical Physics, 1964.

[4] L. A. Liusternik and V. J. Sobolev, *Elements of Functional Analysis* (Frederick Ungar Publishing Company, New York, 1964), p. 194.

continuous in x; the proof is exactly the same as that of Theorem 4 and is omitted. It is also mentioned that the solution to Eq. (14) satisfies the original Boltzmann equation (12) almost everywhere, and it also satisfies the boundary conditions (2a) and (2b) for almost all $\mathbf{c}$.

VI. CONCLUDING REMARKS

Throughout this paper we have assumed that $g_{\pm}(\mathbf{c})$ are known functions (which is also assumed in Ref. 1), although, in general, $g_{\pm}(\mathbf{c})$ do depend on the solution of the boundary-value problem. To give a justification for this assumption, we consider the case of diffusive re-emission which represents, in many problems, a reasonably realistic boundary condition. In this case, $g_{\pm}(\mathbf{c})$ have an explicit dependence on $\mathbf{c}$ but contain a number of parameters. These parameters are related to some of the moments of the distribution function at the plates and are pure constants. We then solve the boundary-value problem, linear or nonlinear, and the solution will also contain these unknown constants. Substituting the solution into the moment integrals gives rise to a system of transcendental algebraic equations which can be used to determine the unknown constants.

All the results we obtained in this paper apply to the hard-sphere molecules and hard-potential power-law molecules. We expect all the results to be also valid for the case $s = 5$, for which a more refined method of estimate than that used in the Appendix will be needed.

Finally it is noted that the present analysis should also be applicable to the problems of a gas contained between two concentric cylinders or concentric spheres, e.g., the cylindrical Couette problem or the heat-transfer problems for concentric cylinders or spheres.

ACKNOWLEDGMENTS

APPENDIX. PROOF FOR INEQUALITY (10)

We write $S(t)$ in the form of

$$S(t) = \int d\mathbf{c}_1 \, d\mathbf{c}_2 \, d\mathbf{c}_3 \, d\mathbf{c}_4 \, |R K_{13} K_{14} K_{23} K_{24}|, \quad \text{(A1)}$$

where

$$K_{ij} = K(\mathbf{c}_i, \mathbf{c}_j)$$

and

$$R = [c_{3x} c_{4x} \nu(c_3)\nu(c_4)\rho(\mathbf{c}_1)\rho(\mathbf{c}_2)]^{-1}\rho(\mathbf{c}_3)\rho(\mathbf{c}_4)$$
$$\times \exp\left[-t\nu(c_3)\,|c_{3x}|^{-1} - t\nu(c_4)\,|c_{4x}|^{-1}\right].$$

For ν and K_{ij} we use the following estimate:

$$b_0 < b_0(1 + c)^\gamma < \nu(c) < b_1(1 + c)^\gamma, \quad \text{(A2)}$$

$$|K_{ij}| < [p_{ij} + q_{ij}] \exp\left(-\tfrac{1}{8}v_{ij}^2\right), \quad \text{(A3)}$$

$$p_{ij} = a_1(v_{ij} + v_{ij}^{-1}) \exp\left[-\tfrac{1}{4}(c_i^2 + c_j^2)\right], \quad \text{(A4)}$$

$$q_{ij} = (a_2/v_{ij}) \exp\left[-\tfrac{1}{8}v_i^2 - \tfrac{1}{4}\zeta_{ij}^2\right], \quad \text{(A5)}$$

where $v_{ij} = |\mathbf{c}_i - \mathbf{c}_j|$, $\zeta_{ij}^2 = (c_i^2 - c_j^2)^2/v_{ij}^2$, which can be found in Ref. 2. (Our velocity $\mathbf{c}$ is $\sqrt{2}$ times that used in Ref. 2.) We note that $0 < \gamma \le 1$. If we apply the inequality

$$v_{13}^2 + v_{32}^2 + v_{14}^2 + v_{24}^2$$
$$\ge v_{34}^2 + v_{12}^2 + |\mathbf{c}_3 + \mathbf{c}_4 - \mathbf{c}_1 - \mathbf{c}_2|^2$$
$$\ge v_{34}^2$$

to (A1), we obtain

$$S(t) \le \int d\mathbf{c}_1 \, d\mathbf{c}_2 \, d\mathbf{c}_3 \, d\mathbf{c}_4 \, |R| \, [\Pi(p_{ij} + q_{ij})] \exp\left(-\tfrac{1}{8}v_{34}^2\right),$$
$$\text{(A6)}$$

where the product Π is taken over $i = 1, 2$ and $j = 3$, 4. Let us now consider the term S_1 in the last integral, which involves the product of four q_{ij}'s, i.e.,

$$S_1(t) = \int d\mathbf{c}_1 \, d\mathbf{c}_2 \, d\mathbf{c}_3 \, d\mathbf{c}_4 \, |R| \, q_{13} q_{14} q_{23} q_{24} \exp\left(-\tfrac{1}{8}v_{34}^2\right).$$

Using Schwartz's inequality on the integrals over $\mathbf{c}_1$ and $\mathbf{c}_2$, the last equation becomes

$$S_1(t) \le \int d\mathbf{c}_3 \, d\mathbf{c}_4 \, \frac{T(\mathbf{c}_3)T(\mathbf{c}_4)}{|c_{3x} c_{4x} \nu(c_3)\nu(c_4)|}$$
$$\times \exp\left[-\tfrac{1}{8}v_{34}^2 - tb_0(|c_{3x}|^{-1} + |c_{4x}|^{-1})\right],$$

where

$$T(c_i) = \rho(\mathbf{c}_i) \int d\mathbf{c}_j \, \frac{q_{ij}^2}{\rho(\mathbf{c}_j)}.$$

For the function $T(\mathbf{c}_i)$ we can obtain two types of estimates, using the inequality

$$\int d\mathbf{c}_j v_{ij}^{-1}(1 + c_j)^{-\alpha} \exp\left(-\beta_1 v_{ij}^2 - \beta_2 \zeta_{ij}^2\right)$$
$$< \mu(1 + c_i)^{-1-\alpha}, \quad \text{(A7)}$$

where β_1, β_2, and α can be any positive constants. The last inequality is due to Carlemann.[5] Applying this inequality to $T(\mathbf{c}_i)$, we obtain one estimate

$$T(\mathbf{c}_i)/\rho(\mathbf{c}_i) < \mu(1 + c_i)^{-1-\gamma} \quad \text{(A8)}$$

[5] T. Carlemann, *Problèmes Mathématiques dans la Théorie Cinétique des Gaz* (Almqvist & Wiksells boktryckeri AB, Uppsala, Sweden, 1957), pp. 71–74.

by noting that $\rho(\mathbf{c}) > b_0(1 + c)^\gamma$. To obtain the second type of estimate for T, we note that $\rho(\mathbf{c}) > b_0 + \pi |c_x|/d$, and consequently

$$\frac{T(\mathbf{c}_i)}{\rho(\mathbf{c}_i)} < \int d\mathbf{c}_j \frac{q_{ij}^2}{(b_0 + \pi |c_{jx}|/d)}$$

$$< \exp\left(-\tfrac{1}{32} c_{ix}^2\right) \int_{D_1} d\mathbf{c}_j \, \tilde{q}_{ij}^2$$

$$+ \left(b_0 + \frac{\pi |c_{ix}|}{2d}\right)^{-1} \int_{D_2} d\mathbf{c} \, q_{ij}^2, \quad (A9)$$

where D_1 and D_2 are, respectively, the regions in the $\mathbf{c}_j$ space for $|c_{jx}| < \tfrac{1}{2} |c_{ix}|$ and $|c_{jx}| > \tfrac{1}{2} |c_{ix}|$ and

$$\tilde{q}_{ij}^2 = (a_2/v_{ij}) \exp\left[-\tfrac{1}{16} v_{ij}^2 - \tfrac{1}{4}\zeta_{ij}^2\right].$$

Now inequality (A7) applies to the two integrals in (A9), and as a result

$$T(\mathbf{c}_i)/\rho(\mathbf{c}_i) < \mu_1(1 + c_i)^{-1}(1 + |c_{ix}|)^{-1}. \quad (A10)$$

Using (A8) and (A10) together, we find

$$\frac{T(\mathbf{c}_3)}{\nu(\mathbf{c}_3)} < (\lambda + 1)\left[\frac{T(\mathbf{c}_3)}{\rho(\mathbf{c}_3)}\right] + \frac{|c_{3x}| \, T(\mathbf{c}_3)}{[\nu(\mathbf{c}_3)\rho(\mathbf{c}_3)]}$$

$$< B_0(1 + c_3)^{-1-\gamma},$$

and the same inequality holds for $T(\mathbf{c}_4)/\nu(\mathbf{c}_4)$. Using these results, we obtain for $S_1(t)$

$$S_1(t) < C \int d\mathbf{c}_3 (1 + c_{3y}^2)^{-\tau}(1 + c_{3z}^2)^{-\tau} |c_{3x}|^{-1}$$

$$\times G_1(c_{3x}, t) G_2(c_{3y}) G_2(c_{3z}) \exp\left(-tb_0/|c_{3x}|\right), \quad (A11)$$

where $\tau = (1 + \gamma)/4$ and we have used the fact that $(1 + c_y^2)(1 + c_z^2) < (1 + c)^4$. The functions G_1 and G_2 are given by

$$G_1(\xi, t) = \int_{-\infty}^{\infty} d\eta \, |\eta|^{-1} \exp\left[-\tfrac{1}{8}(\xi - \eta)^2 - \frac{tb_0}{|\eta|}\right],$$

$$G_2(\xi) = \int_{-\infty}^{\infty} d\eta \, (1 + \eta^2)^{-\tau} \exp\left[-\tfrac{1}{8}(\xi - \eta)^2\right].$$

To estimate G_1 we consider the cases $|\xi| > 1$ and $|\xi| < 1$ separately. For $|\xi| < 1$, we have

$$G_1(\xi, t) < \int_{-2}^{2} d\eta \, |\eta|^{-1} \exp\left[-\frac{tb_0}{|\eta|}\right] + \tau_1$$

$$< (\tau_2 + \tau_3 \ln t);$$

and for $|\xi| > 1$

$$G_1(\xi, t) < \exp\left(-\tfrac{1}{16}\xi^2\right)\int_{-\frac{1}{2}}^{\frac{1}{2}} |y|^{-1} \exp\left(-\frac{t}{|y\xi|}\right) dy$$

$$+ 2\int_{-\infty}^{\infty} \exp\left[-\tfrac{1}{8}\xi^2 1(-y)^2\right] dy$$

$$< (\tau_4 + \tau_5 \ln t)[1 + |\xi|]^{-1}.$$

Combination of these results yields

$$G_1(\xi, t) < [a_3 + a_4 \ln t][1 + |\xi|]^{-1}.$$

For G_2, the estimate can be made in a simpler manner, and the result is

$$G_2(\xi) < a_5[1 + \xi^2]^{-\tau}.$$

If we substitute the estimates for G_1 and G_2 into (A11), the c_{3y} and c_{3z} integrals can be seen to converge. We therefore have

$$S_1(t) < A_0[a_3 + a_4 \ln t]$$

$$\times \int dc_{3x} |c_{3x}|^{-1}(1 + |c_{3x}|)^{-1} \exp\left(-\frac{tb_0}{|c_{3x}|}\right),$$

and the last integral may be estimated in a manner similar to what we did for G_1. Finally, we obtain

$$S_1(t) < [A_1 + B_1 \ln t]^2$$

for some positive constants A_1 and B_1. This completes the estimate for S_1. For the remaining terms in the right-hand side of (A6), we can obtain the same estimates and the proof is similar, but even somewhat simpler, on account of the fact that the properties of p_{ij} are nicer than those of q_{ij}.

Representations of the Orthogonal Group. I. Lowering and Raising Operators of the Orthogonal Group and Matrix Elements of the Generators

M. K. F. WONG

Research Institute for Natural Sciences, Woodstock College, Woodstock, Maryland

(Received 15 March 1967)

A parallel method to that of Pang and Hecht for the construction of normalized lowering and raising operators for the orthogonal group $O(n) \supset O(n-1) \supset \cdots \supset O(2)$ is presented. The generators are defined in a slightly different way from those of Pang and Hecht, and the lowering and raising operators are constructed without using graphs. The Gel'fand–Zetlin matrix elements of the infinitesimal generators of $O(n)$ have also been obtained.

INTRODUCTION

RECENTLY Pang and Hecht[1] have obtained normalized lowering and raising operators for the orthogonal group in the canonical group chain $O(n) \supset O(n-1) \supset \cdots \supset O(2)$. Their method differs from that of Nagel and Moshinsky[2] in that they obtain these operators with the aid of graphs.

We present here a parallel method which is more closely connected with the method of Nagel and Moshinsky who first obtained the lowering (raising) operators of the unitary group, though the problem is far more lengthy and complicated in the case of the orthogonal group. The generators we use agree with those of Gel'fand and Zetlin,[3] but differ from those of Pang and Hecht. As a result, at least in appearance, their diagonal generator is the negative of ours, and their raising (lowering) generators become, in appearance, our lowering (raising) generators, respectively.

Possible application of the orthogonal group to physical problems has been discussed in Pang and Hecht's paper. In particular the groups $O(5)$ and $O(8)$ seem to be of interest in nuclear physics.[4] For further references, one can consult the papers listed in Ref. 4.

The main purpose of this paper is to present the general expression of the lowering (raising) operators of $O(n)$ in algebraic form without using graphs. We have found that the concept of primitive roots helps a great deal in simplifying the definition, construction, and proof of these operators. As a result, our proof, though still lengthy, is in principle quite straight-

forward, since it resembles very closely the proof of Nagel and Moshinsky in the case of the unitary group.

In the evaluation of the normalization constants, the method used is essentially that of Pang and Hecht, with one simplification. We have found it not necessary to sum up the graphs. By introducing a simple theorem (Step 7, Sec. I D), we have been able to obtain the normalization constants quite easily. The same remark also applies to the method of Nagel and Moshinsky.

Finally, mainly as a check, we have obtained the Gel'fand–Zetlin matrix elements of the infinitesimal generators of $O(n)$, which, of course, includes a special case of the Wigner coefficients of $O(n)$.

I. THE REPRESENTATIONS OF $O(n)$

A. Generators of $O(n)$

The infinitesimal generators of $O(n)$ are the set of skew symmetric, Hermitian operators J_{ij} with the commutation relations

$$[J_{ab}, J_{cd}] = i[\delta_{ad}J_{cb} + \delta_{ac}J_{bd} + \delta_{bc}J_{da} + \delta_{bd}J_{ac}]. \quad (1)$$

They can be expressed as differential operators

$$J_{ij} = -i\left(x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i}\right).$$

The fact that they are skew symmetric and Hermitian means that they are related to each other as follows:

$$J_{ab} = -J_{ba}, \quad (2)$$

$$J_{ab}^{+} = J_{ab}. \quad (3)$$

We shall classify the generators of $O(n)$ according to the Cartan canonical basis.

(i) For $O(2k+1)$

Type 1: $H_\alpha \equiv J_{2\alpha}^{2\alpha-1}$, $\alpha = 1, 2, \cdots, k$ (note, we have found it convenient to write the second subscript on top, thus making it a superscript).

[1] S. C. Pang and K. T. Hecht, J. Math. Phys. **8**, 1233 (1967).

[2] J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).

[3] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk. USSR **71**, 1017 (1950).

[4] B. H. Flowers and Szpikowski, Proc. Phys. Soc. (London) **84**, 193 (1964); J. C. Parikh, Nucl. Phys. **63**, 214 (1965); J. N. Ginocchio, *ibid.* **74**, 321 (1965); M. Ichimura, Progr. Theoret. Phys. (Kyoto) **32**, 757 (1964); **33**, 215 (1965); K. T. Hecht, Phys. Rev. **139**, B794 (1965); J. D. Louck, Los Alamos Scientific Laboratory Reports LA 2451 (1960).

Type 2:

$$E^p_{2k+1} \equiv \frac{1}{\sqrt{2}}(J^{2p}_{2k+1} + iJ^{2p-1}_{2k+1}) \quad p = 1, 2 \cdots k,$$

$$F^p_{2k+1} \equiv \frac{1}{\sqrt{2}}(J^{2p}_{2k+1} - iJ^{2p-1}_{2k+1}) \quad p = 1, 2 \cdots k.$$

Type 3:

$$A^q_p \equiv \tfrac{1}{2}(J^{2q}_{2p-1} + J^{2q-1}_{2p} + iJ^{2q-1}_{2p-1} - iJ^{2q}_{2p}),$$

$$B^q_p \equiv \tfrac{1}{2}(J^{2q}_{2p-1} + J^{2q-1}_{2p} - iJ^{2q-1}_{2p-1} + iJ^{2q}_{2p}),$$

$$C^q_p \equiv \tfrac{1}{2}(-J^{2q}_{2p-1} + J^{2q-1}_{2p} + iJ^{2q-1}_{2p-1} + iJ^{2q}_{2p}), \quad (4)$$

$$D^q_p \equiv \tfrac{1}{2}(-J^{2q}_{2p-1} + J^{2q-1}_{2p} - iJ^{2q-1}_{2p-1} - iJ^{2q}_{2p}),$$

$$p > q,$$

$$q = 1, 2 \cdots k - 1,$$

$$p = 1, 2, \cdots k - 1, k$$

N.B. $C^k_k = D^k_k = J^{2k-1}_{2k}.$

(ii) For $O(2k)$

Type 1: $H_\alpha = J^{2\alpha-1}_{2\alpha}, \ \alpha = 1, 2, \cdots k.$

Type 2: $A^q_p, B^q_p, C^q_p, D^q_p.$ Same as in Eq. (4). These generators have the following properties:

$$[H_\alpha, H_\beta] = 0, \quad (5)$$

$$[H_\alpha, A^q_p] = (\delta_{\alpha p} + \delta_{\alpha q})A^q_p, \quad (6)$$

$$[H_\alpha, B^q_p] = (-\delta_{\alpha p} - \delta_{\alpha q})B^q_p, \quad (7)$$

$$[H_\alpha, C^q_p] = (\delta_{\alpha p} - \delta_{\alpha q})C^q_p, \quad (8)$$

$$[H_\alpha, D^q_p] = (-\delta_{\alpha p} + \delta_{\alpha q})D^q_p, \quad (9)$$

$$[H_\alpha, E^p_{2k+1}] = \delta_{\alpha p}E^p_{2k+1}, \quad (10)$$

$$[H_\alpha, F^p_{2k+1}] = -\delta_{\alpha p}F^p_{2k+1}, \quad (11)$$

$$[A^q_p, B^q_p] = H_p + H_q, \quad \text{etc.,} \quad (12)$$

$$[D^{q'}_{p'}, C^q_p] = \delta_{pp'}\begin{cases} C^q_{q'} & (q' > q) \\ D^{q'}_q & (q > q') \end{cases}$$
$$- \delta_{qq'}\begin{cases} C^{p'}_p & (p > p') \\ D^p_{p'} & (p' > p) \end{cases}, \quad (13)$$

$$[D^{q'}_{p'}, B^q_p] = \delta_{qq'}\begin{cases} -B^{p'}_p & (p > p') \\ B^p_{p'} & (p' > p) \end{cases}$$
$$- \delta_{q'p}B^q_{p'}, \quad (p' > q), \quad (14)$$

$$[E^p_{2k+1}, F^q_{2k+1}] = \begin{cases} C^q_p & (p > q) \\ D^p_q & (q > p) \end{cases}, \quad (15)$$

$$A^{q^+}_p = B^q_p \quad C^{q^+}_p = D^q_p \quad E^{p^+}_{2k+1} = F^p_{2k+1}, \quad (16)$$

$$\sqrt{2}\, F^p_{2k-1} = B^p_k - C^p_k \quad \sqrt{2}\, E^p_{2k-1} = A^p_k - D^p_k. \quad (17)$$

Other commutation relations can be obtained from Eq. (1). Note that our definition of H_α is opposite to that of Pang and Hecht. This is because we want to agree with Gel'fand and Zetlin's definition of H_α, and to obtain the consistent result that in the generators the first (lower) subscript is always greater than the second (top) subscript (or superscript).

B. The Gel'fand Basis

This is defined in the same way as in Pang and Hecht and will not be repeated here. The branching rules are

$$m_{2p+1,i} \geq m_{2p,i} \geq m_{2p+1,i+1}, \quad (i = 1, 2, \cdots p);$$

$$m_{2p,i} \geq m_{2p-1,i} \geq m_{2p,i+1}, \quad (i = 1, 2, \cdots p - 1);$$

$$m_{2p+1,p} \geq |m_{2p,p}|.$$

The branching rules for the canonical decomposition of $O(n) \supset O(n-1) \supset \cdots \supset O(2)$ have been proved by Boerner.[5]

C. The Lowering and Raising Operators

It is easily seen that A^q_p, D^q_p, E^p_{2k+1} correspond to the roots $e_q + e_p, e_q - e_p, e_q$, respectively, and are therefore the raising generators, while B^q_p, C^q_p, F^q_{2k+1} correspond to the roots $-e_q - e_p, -e_q + e_p, -e_q$, respectively, and are therefore the lowering generators.

However, the definition, construction, and proof of the lowering (raising) operators can be greatly simplified by the introduction of the concept of "primitive roots." These are defined as follows. In $O(2k)$ there are k primitive roots such that any other raising generator (or positive root) can be obtained from combining the primitive roots. Likewise in $O(2k + 1)$ there are also k primitive roots.

These primitive roots are

$$O(2k): \quad D^1_2, D^2_3, \cdots D^{k-1}_k, A^{k-1}_k, \quad (18)$$

$$O(2k + 1): \quad D^1_2, D^2_3, \cdots D^{k-1}_k, E^k_{2k+1}. \quad (19)$$

For example, in $O(2k)$ if we wish to obtain A^{k-2}_{k-1}, we can combine first D^{k-2}_{k-1} and D^{k-1}_k to obtain D^{k-2}_k, since

$$[D^{k-2}_{k-1}, D^{k-1}_k] = D^{k-2}_k.$$

Then we can combine D^{k-2}_k and A^{k-1}_k to obtain A^{k-2}_{k-1}, since

$$[A^{k-1}_k, D^{k-2}_k] = A^{k-2}_{k-1}.$$

The other positive roots or raising generators are obtained by induction.

In $O(2k + 1)$ if we wish to obtain A^{k-1}_k, we can combine first E^k_{2k+1} and D^{k-1}_k to obtain E^{k-1}_{2k+1}, since

$$[D^{k-1}_k, E^k_{2k+1}] = E^{k-1}_{2k+1}.$$

―――――――
[5] H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963), p. 252.

Then we can combine E_{2k+1}^{k-1} and E_{2k+1}^k to obtain A_{k-1}^k, since

$$[E_{2k+1}^{k-1}, E_{2k+1}^k] = A_k^{k-1}. \qquad (27')$$

The other positive roots can again be obtained by induction.

With the help of the primitive roots we now define the lowering (raising) operators as follows.

For $O(2k)$

1. Lowering operator L_{2k}^p, $p = 1, 2, \cdots k - 1$.

$$[J_{2\alpha}^{2\alpha-1}, L_{2k}^p] = -\delta_{\alpha p} L_{2k}^p, \quad 0 \le \alpha \le k - 1, \qquad (20)$$

$$[D_{p'+1}^{p'}, L_{2k}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p' = 1, 2, \cdots k - 2, \qquad (21)$$

$$[E_{2k-1}^{k-1}, L_{2k}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0. \qquad (22)$$

2. Raising operator R_{2k}^p, $p = 1, 2, \cdots k - 1$.

$$[J_{2\alpha}^{2\alpha-1}, R_{2k}^p] = \delta_{\alpha p} R_{2k}^p, \quad 0 < \alpha \le k - 1, \qquad (23)$$

$$[D_{p'+1}^{p'}, R_{2k}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p' = 1, 2, \cdots k - 2, \qquad (24)$$

$$[E_{2k-1}^{k-1}, R_{2k}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0. \qquad (25)$$

3. "Zero-step" operator $O_{2k,k}$.

$$[J_{2\alpha}^{2\alpha-1}, O_{2k,k}] = 0, \quad 0 < \alpha \le k - 1, \qquad (26)$$

$$[D_{p'+1}^{p'}, O_{2k,k}]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p' = 1, 2, \cdots k - 2, \qquad (27)$$

$$[E_{2k-1}^{k-1}, O_{2k,k}]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \qquad (28)$$

$$[R_{2k}^p, O_{2k,k}]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p = 1, 2, \cdots k - 1. \qquad (29)$$

For $O(2k + 1)$

1. Lowering operator L_{2k+1}^p, $p = 1, 2, \cdots k$.

$$[J_{2\alpha}^{2\alpha-1}, L_{2k+1}^p] = -\delta_{\alpha p} L_{2k+1}^p, \quad \alpha = 1, 2, \cdots k, \qquad (30)$$

$$[D_{p'+1}^{p'}, L_{2k+1}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p' = 1, 2, \cdots k - 1, \qquad (31)$$

$$[A_k^{k-1}, L_{2k+1}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0. \qquad (32)$$

2. Raising operator R_{2k+1}^p, $p = 1, 2, \cdots k$.

$$[J_{2\alpha}^{2\alpha-1}, R_{2k+1}^p] = \delta_{\alpha p} R_{2k+1}^p, \qquad (33)$$

$$[D_{p'+1}^{p'}, R_{2k+1}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0, \quad p' = 1, 2, \cdots k - 1, \qquad (34)$$

$$[A_k^{k-1}, R_{2k+1}^p]\left|{m_{n,p} \atop m_{n-1,p}}\right\rangle = 0. \qquad (35)$$

We have obtained the lowering (raising) operators without using graphs. They are:

(i) For $O(2k)$

$$L_{2k}^m = \left\{ C_k^m + B_k^m + \sum_{p=1}^{k-m-1}\sum_{\mu_p > \mu_{p-1}\cdots > \mu_1 = m+1}^{k-1} C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}}(C_k^{\mu_p} + B_k^{\mu_p})\prod_{i=1}^p \delta_{m\mu_i}^{-1} \right.$$

$$+ \sqrt{2}\sum_{p=1}^{k-m-1}\sum_{\mu_p > \mu_{p-1}\cdots > \mu_1 = m+1}^{k-1} C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}}(-F_{2k-1}^{\mu_p})H_k \prod_{i=1}^p \delta_{m\mu_i}^{-1} C_m^{-1} + \sqrt{2}(-F_{2k-1}^m)H_k C_m^{-1}$$

$$+ \sum_{p=1}^{k-1} B_p^m[-(A_k^p + D_k^p)]b_{mp}^{-1} + \sum_{q=1}^{k-1}\sum_{p=1}^{k-m-1}\sum_{\mu_p > \cdots \mu_1 = m+1}^{k-1} C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}} B_q^{\mu_p}[-(A_k^q + D_k^q)]\prod_{i=1}^p \delta_{m\mu_i}^{-1} b_{mq}^{-1}$$

$$+ \sum_{p'=2}^{k-1}\sum_{v_{p'} > \cdots v_1=1}^{k-1}\sum_{p=1}^{k-m-1}\sum_{\mu_p > \cdots \mu_1 = m+1}^{k-1} C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_p}^{\mu_{p-1}} B_{v_{p'}}^{\mu_p}(-C_{v_{p'}}^{v_{p'}-1}) \cdots (-C_{v_2}^{v_1})[-(A_k^{v_1} + D_k^{v_1})]$$

$$\times \prod_{i=1}^p \delta_{m\mu_i}^{-1} \prod_{i'=1}^{p'} b_{mv_{i'}}^{-1} + \sum_{q=1}^{k-1}(-F_{2k-1}^m)(-F_{2k-1}^q)[-(A_k^q + D_k^q)]b_{mq}^{-1} C_m^{-1}$$

$$+ \sum_{q=1}^{k-1}\sum_{p=1}^{k-m-1}\sum_{\mu_p > \cdots \mu_1 = m+1}^{k-1} C_{\mu_1}^m \cdots C_{\mu_p}^{\mu_{p-1}}(-F_{2k-1}^{\mu_p})(-F_{2k-1}^q)[-(A_k^q + D_k^q)]\prod_{i=1}^p \delta_{m\mu_i}^{-1} b_{mq}^{-1} C_m^{-1}$$

$$+ \sum_{p=2}^{k-1}\sum_{\mu_p > \cdots > \mu_1=1}^{k-1} B_{\mu_p}^m(-C_{\mu_p}^{\mu_p-1})(-C_{\mu_{p-1}}^{\mu_p-2}) \cdots (-C_{\mu_2}^{\mu_1})[-(A_k^{\mu_1} + D_k^{\mu_1})]\prod_{i=1}^p b_{m\mu_i}^{-1}$$

$$+\sum_{p=2}^{k-1}\sum_{\mu_p>\cdots>\mu_1=1}^{k-1}(-F_{2k-1}^m)(-F_{2k-1}^{\mu_p})(-C_{\mu_p}^{\mu_p-1})\cdots(-C_{\mu_2}^{\mu_1})[-(A_k^{\mu_1}+D_k^{\mu_1})]\prod_{i=1}^p b_{m\mu_i}^{-1}C_m^{-1}$$

$$+\sum_{p'=2}^{k-1}\sum_{\nu_{p'}>\cdots>\nu_1=1}^{k-1}\sum_{p=1}^{k-m-1}\sum_{\mu_p>\cdots>\mu_1=m+1}^{k-1}C_{\mu_1}^m\cdots C_{\mu_p}^{\mu_p-1}(-F_{2k-1}^{\mu_p})(-F_{2k-1}^{\nu_{p'}})(-C_{\nu_{p'}}^{\nu_{p'}-1})(-C_{\nu_{p'}-1}^{\nu_{p'}-2})\cdots(-C_{\nu_2}^{\nu_1})$$

$$\times\,[-(A_k^{\nu_1}+D_k^{\nu_1})]\prod_{i=1}^p \mathcal{E}_{m\mu_i}^{-1}\prod_{i'=1}^{p'}b_{m\nu_{i'}}^{-1}C_m^{-1}\Big\}\prod_{l=m+1}^{k-1}\mathcal{E}_{ml}\prod_{\mu=1}^{k-1}b_{m\mu}C_m,\quad(36)$$

where

$$\mathcal{E}_{ml}\equiv H_m-H_l+l-m=a_m-a_l,$$
$$b_{m\mu}\equiv a_m+a_\mu-1=a_{m\mu}-1,$$
$$a_m\equiv H_m+k-m,$$
$$C_m\equiv a_m-1.$$

Note that according to definition, B_a^b should be written in such a way that $a>b$. If, however, the formula requires a term B_b^a $(a>b)$, then we should use Eq. (2) and change it to $-B_a^b$:

$$R_{2k}^m=\sum_{p=1}^{m-1}\sum_{\mu_p>\mu_{p-1}\cdots>\mu_1=1}^{m-1}C_m^{\mu_p}C_{\mu_p}^{\mu_{p-1}}\cdots C_{\mu_2}^{\mu_1}(A_k^{\mu_1}+D_k^{\mu_1})\prod_{i=1}^p\mathcal{E}_{m\mu_i}^{-1}\prod_{l=1}^{m-1}\mathcal{E}_{ml}+(A_k^m+D_k^m)\prod_{l=1}^{m-1}\mathcal{E}_{ml},\quad(37)$$

$$O_{2k,k}=2H_k\prod_{\alpha=1}^{k-1}a_\alpha+\sum_{p=1}^{k-1}\sqrt2\,F_{2k-1}^p(A_k^p+D_k^p)a_p^{-1}\prod_{\alpha=1}^{k-1}a_\alpha$$

$$+\sum_{p=2}^{k-1}\sum_{\mu_p>\cdots>\mu_1=1}^{k-1}\sqrt2\,F_{2k-1}^{\mu_p}(-C_{\mu_p}^{\mu_p-1})\cdots(-C_{\mu_2}^{\mu_1})(A_k^{\mu_1}+D_k^{\mu_1})\prod_{i=1}^p a_{\mu_i}^{-1}\prod_{\alpha=1}^{k-1}a_\alpha.\quad(38)$$

(ii) For $O(2k+1)$

$$L_{2k+1}^m=i\Big\{F_{2k+1}^m+\sum_{q=1}^k B_q^m E_{2k+1}^q a_{mq}^{-1}+\sum_{p=1}^{k-m}\sum_{\mu_p>\cdots>\mu_1=m+1}^k C_{\mu_1}^m\cdots C_{\mu_p}^{\mu_p-1}F_{2k+1}^{\mu_p}\prod_{i=1}^p\mathcal{E}_{m\mu_i}^{-1}$$

$$+\sum_{q=1}^k\sum_{p=1}^{k-m}\sum_{\mu_p>\cdots>\mu_1=m+1}^k C_{\mu_1}^m\cdots C_{\mu_p}^{\mu_p-1}B_q^{\mu_p}E_{2k+1}^q\prod_{i=1}^p\mathcal{E}_{m\mu_i}^{-1}a_{mq}^{-1}+\sum_{p=2}^k\sum_{\mu_p>\cdots>\mu_1=1}^k B_{\mu_p}^m(-C_{\mu_p}^{\mu_p-1})\cdots(-C_{\mu_2}^{\mu_1})E_{2k+1}^{\mu_1}$$

$$\times\prod_{i=1}^p a_{m\mu_i}^{-1}+\sum_{p'=2}^k\sum_{\nu_{p'}>\cdots>\nu_1=1}^k\sum_{p=1}^{k-m}\sum_{\mu_p>\cdots>\mu_1=m+1}^k C_{\mu_1}^m\cdots C_{\mu_p}^{\mu_p-1}B_{\nu_{p'}}^{\mu_p}(-C_{\nu_{p'}}^{\nu_{p'}-1})\cdots(-C_{\nu_2}^{\nu_1})$$

$$\times\,E_{2k+1}^{\nu_1}\prod_{i=1}^p\mathcal{E}_{m\mu_i}^{-1}\prod_{i'=1}^{p'}a_{m\nu_{i'}}^{-1}\Big\}\prod_{l=m+1}^k\mathcal{E}_{ml}\prod_{\mu=1}^k a_{m\mu},\quad(39)$$

where $a_{m\mu}=a_m+a_\mu$

$$R_{2k+1}^m=i\Big\{E_{2k+1}^m+\sum_{p=1}^{m-1}\sum_{\mu_p>\cdots>\mu_1=1}^{m-1}C_m^{\mu_p}C_{\mu_p}^{\mu_p-1}\cdots C_{\mu_2}^{\mu_1}E_{2k+1}^{\mu_1}\prod_{i=1}^p\mathcal{E}_{m\mu_i}^{-1}\Big\}\prod_{l=1}^{m-1}\mathcal{E}_{ml}.\quad(40)$$

First, a few words about how these operators are obtained. Take the case of L_{2k}^m. We start with the simplest lowering operator, in this case the generators C_k^m and B_k^m. They evidently satisfy Eq. (20). However, when they are commuted with $D_{p'+1}^{p'}$, we find it necessary to introduce the next operators in line in the form of

$$\sum_p\sum_{\mu_p>\cdots>\mu_1}C_m^m C_{\mu_1}^m C_{\mu_2}^{\mu_1}\cdots C_{\mu_p}^{\mu_p-1}(C_k^{\mu_p}+B_k^{\mu_p})$$

in order that Eq. (21) may be satisfied. This operator then in turn makes it necessary to introduce the next operator in line, and so on, until we reach the last operator in line. The diagonal operators $\mathcal{E}_{ml}$, $a_{m\mu}$ and C_m are so introduced in order to satisfy Eqs. (20)–(35).

These lowering and raising operators expressed in general form, though substantially the same as those of Pang and Hecht, are of some practical help, we hope, since one can write them down immediately without having to draw individual graphs, which increase rapidly in number as k increases.

The proof that these operators indeed satisfy Eqs. (20)–(35) is essentially the same as given by Nagel and Moshinsky.[2] Our $D_{p',+1}^{p'}$ is similar to their $C_{\lambda}^{\lambda+1}$, and our $\mathcal{E}_{ml}$ is exactly the same as theirs. Some typical examples are given in Appendix A.

D. Normalization

The normalization constants $N_{q_m-1}^{q_m}$ and $N_{q_m+1}^{q_m}$ are defined in the same way as in Nagel and Moshinsky.[2] We obtain:

(i) For $O(2k)$

$$(N_{q_m-1}^{q_m})^2 = \left(N \begin{array}{ccccccc} q_1 \cdots q_{m-1} & q_m & q_{m+1} \cdots q_{k-1} & \\ h_1 \cdots h_{m-1} & h_m & h_{m+1} \cdots h_{k-1} & h_k \\ q_1 \cdots q_{m-1} & q_m - 1 & q_{m+1} \cdots q_{k-1} & \end{array} \right)^2$$

$$= 2(h_m - q_m + 1)(h_m + q_m + 2k - 2m - 1)(h_k + q_m + k - m - 1)$$

$$\times (q_m - h_k + k - m - 1)(q_m + k - m - 1)(2q_m + 2k - 2m - 1)$$

$$\times \prod_{\mu=1}^{m-1} \frac{h_\mu - q_m + m - \mu + 1}{q_\mu - q_m + m - \mu + 1}(q_\mu + q_m + 2k - \mu - m - 1)(q_m + h_\mu + 2k - m - \mu - 1)$$

$$\times \prod_{\mu=m+1}^{k-1} (q_m - q_\mu + \mu - m)(q_m - h_\mu + \mu - m - 1)$$

$$\times (q_\mu + q_m + 2k - \mu - m - 1)(q_m + h_\mu + 2k - m - \mu - 1), \tag{41}$$

where according to the notation of Pang and Hecht,

$$q_\alpha = m_{2k-1,\alpha},$$

$$h_\alpha = m_{2k,\alpha},$$

$$(N_{q_m+1}^{q_m})^2 = 2(q_m + h_m + 2k - 2m)(h_m - q_m)(h_k + q_m + k - m)(q_m - h_k + k - m)(q_m + k - m)^{-1}$$

$$\times (2q_m + 2k - 2m + 1)^{-1} \prod_{\mu=1}^{m-1} (q_\mu - q_m + m - \mu)(h_\mu - q_m - \mu + m) \frac{(h_\mu + q_m + 2k - m - \mu)}{(q_\mu + q_m + 2k - m - \mu)}$$

$$\times \prod_{\mu=m+1}^{k-1} \frac{(q_m - h_\mu + \mu - m)(q_m + h_\mu + 2k - m - \mu)}{(q_m - q_\mu + \mu - m + 1)(q_m + q_\mu + 2k - m - \mu)}. \tag{42}$$

(ii) For $O(2k + 1)$

$$(N_{q_m-1}^{q_m})^2 = \tfrac{1}{2}(2q + 2k - 2m)^2(h_m - q_m + 1)(h_m + q_m + 2k - 2m)$$

$$\times \prod_{\mu=1}^{m-1} \frac{(h_\mu - q_m + m - \mu + 1)}{(q_\mu - q_m + m - \mu + 1)}(q_m + q_\mu + 2k - m - \mu)(q_m + h_\mu + 2k - m - \mu)$$

$$\times \prod_{\mu=m+1}^{k} (q_m - q_\mu + \mu - m)(q_m + q_\mu + 2k - m - \mu)$$

$$\times (q_m - h_\mu + \mu - m - 1)(q_m + h_\mu + 2k - m - \mu), \tag{43}$$

$$(N^{q_m}_{q_m+1})^2 = \tfrac{1}{2}(h_m - q_m)(h_m + q_m + 2k - 2m + 1)$$

$$\times \prod_{\mu=1}^{m-1} (h_\mu - q_m + m - \mu)(q_\mu - q_m + m - \mu) \frac{(q_m + h_\mu + 2k - m - \mu + 1)}{(q_m + q_\mu + 2k - m - \mu + 1)}$$

$$\times \prod_{\mu=m+1}^{k} \frac{(q_m - h_\mu + \mu - m)(q_m + h_\mu + 2k - m - \mu + 1)}{(q_m - q_\mu + \mu - m + 1)(q_m + q_\mu + 2k - m - \mu + 1)} . \tag{44}$$

The method used to obtain the normalization constants is essentially a combination of the methods of Pang and Hecht and Nagel and Moshinsky.[6] Before giving an outline of the steps used to obtain the normalization constants, we follow Nagel and Moshinsky and give some preliminary relations first.

(iii) For $O(2k)$

$$\left\langle \begin{matrix} h \\ q \end{matrix} \right| L^m_{2k} = \left(L^{m+}_{2k} \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \right)^+ = \left\langle \begin{matrix} h \\ q \end{matrix} \right| \prod_{l=m+1}^{k-1} \delta_{ml} b_{ml} \prod_{\mu=1}^{m-1} b_{m\mu}(a_{mm} - 1)C_m \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \left\langle \begin{matrix} h \\ q \end{matrix} \right| (B^m_k + C^m_k), \tag{45}$$

$$\left\langle \begin{matrix} h \\ q \end{matrix} \right| R^m_{2k} = \left\langle \begin{matrix} h \\ q \end{matrix} \right| \prod_{l=1}^{m-1} \delta_{ml} \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \left\langle \begin{matrix} h \\ q \end{matrix} \right| (A^m_k + D^m_k), \tag{46}$$

$$N^{q_m}_{q_m+1} = \langle q_m | \frac{\prod_{l=1}^{m-1} \delta_{ml}}{\prod_{p=m+1}^{k-1} (\delta_{mp} + 1)a_{mp} \prod_{\mu=1}^{m-1} a_{m\mu}(a_{mm} + 1)} | q_m \rangle N^{q_m+1}_{q_m}. \tag{47}$$

Equation (47) comes from comparing (45) and (46), taking N to be real. From (45) and (46), we can easily prove that

$$\langle q_m + 1, q_{m'} - 1| [R^m_{2k}, L^{m'}_{2k}] | q_m, q_{m'} \rangle = 0 \quad m \neq m'. \tag{48}$$

We then obtain

$$(N^{q_\mu - \delta_{\mu m} - \delta_{\mu m'}}_{q_\mu - \delta_{\mu m} - \delta_{\mu m'}})^2 = \langle q_m | \frac{(a_{mm'} - 2)(\delta_{mm'} + 1)}{(a_{mm'} - 1)\delta_{mm'}} | q_m \rangle (N^{q_\mu}_{q - \delta_{\mu m'}})^2, \tag{49}$$

$$(N^{q_\mu - \delta_{\mu m'}}_{q_\mu - \delta_{\mu m'} - \delta_{\mu m'}})^2 = \langle q_m | \frac{(a_{mm'} - 2)(\delta_{mm'} + 1)}{(a_{mm'} - 1)\delta_{mm'}} | q_m \rangle (N^{q_\mu}_{q_\mu - \delta_{\mu m}})^2, \tag{50}$$

where $m' > m$.

$$N^2 \begin{pmatrix} q_1 \cdots q_{m-1} & q_m & q_{m+1} \cdots q_{k-1} \\ q_1 \cdots q_{m-1} & q_m - 1 & q_{m+1} \cdots q_{k-1} \end{pmatrix} = \prod_{\mu=1}^{m-1} \frac{(q_\mu + q_m + 2k - \mu - m - 1)(h_\mu - q_m + m - \mu + 1)}{(h_\mu + q_m + 2k - \mu - m - 1)(q_\mu - q_m + m - \mu + 1)}$$

$$\times \prod_{\mu=m+1}^{k-1} \frac{(q_m + q_\mu + 2k - m - \mu - 1)(q_m - q_\mu + \mu - m)}{(q_m + h_\mu + 2k - m - \mu - 1)(q_m - h_\mu + \mu - m)} N^2 \begin{pmatrix} h_1 \cdots h_{m-1} & q_m & h_{m+1} \cdots h_{k-1} \\ h_1 \cdots h_{m-1} & q_m - 1 & h_{m+1} \cdots h_{k-1} \end{pmatrix},$$

$$\text{where} \quad h_\alpha \geq q_\alpha. \tag{51}$$

The problem is now reduced to finding

$$N^2 \begin{pmatrix} h_1 & \cdots & q_m & \cdots & h_{k-1} \\ h_1 & \cdots & q_m - 1 & \cdots & h_{k-1} \end{pmatrix}$$

which shall be denoted as

$$N^2 \begin{pmatrix} q_m, & h \\ q_m - 1, & h \end{pmatrix}.$$

[6] J. G. Nagel and M. Moshinsky, Rev. Mex. Fis. 14, 29 (1965).

We now follow Pang and Hecht and obtain the normalization constants through the following steps:

Step 1. The Casimir Invariant C_{2k},

$$C_{2k} = \sum_{i>j}^{2k} J_{ij}^2 = \sum_{i=1}^{k-1}(B_k^i + C_k^i)(A_k^i + D_k^i) + \sum_{0<i<j}^{k-1}(B_j^i + C_j^i)(A_j^i + D_j^i)$$

$$\times \sum_{0<i<j}^{k} 2F_{2j-1}^i E_{2j-1}^i + \sum_{i=1}^{k}(J_{2i}^{2i-1})^2 + \sum_{i=1}^{k-1}(2k-2i)J_{2i}^{2i-1}. \quad (52)$$

Step 2.

$$\langle i|\,(B_k^p + C_k^p)(A_k^p + D_k^p)\,|i\rangle = \langle i|\,\delta_{ip}^{-1}\prod_{\alpha=i+1}^{p-1}(1 + \delta_{i\alpha}^{-1})(B_k^i + C_k^i)(A_k^i + D_k^i)\,|i\rangle \quad (53)$$

where $p > i$.

Step 3.

$$\langle i|\,\sum_{j\geq i}^{k-1}(B_k^j + C_k^j)(A_k^j + D_k^j)\,|i\rangle = \langle i|\,\prod_{\alpha=i+1}^{k-1}(1 + \delta_{i\alpha}^{-1})(B_k^i + C_k^i)(A_k^i + D_k^i)\,|i\rangle. \quad (54)$$

Step 4.

$$\langle i|\,O_{2k,k}O_{2k,k}\,|i\rangle = \langle i|\,\prod_{\alpha=1}^{k-1} a_\alpha^2\,|i\rangle\langle i|\,4H_k^2 - 2a_i^{-1}\prod_{\alpha=i+1}^{k-1}(1 + \delta_{i\alpha}^{-1})(B_k^i + C_k^i)(A_k^i + D_k^i)\,|i\rangle. \quad (55)$$

Step 5.

$$\langle i|\,\prod_{\alpha=i+1}^{k-1}(1 + \delta_{i\alpha}^{-1})(B_k^i + C_k^i)(A_k^i + D_k^i)\,|i\rangle + \langle i|\,H_k^2\,|i\rangle + \langle i|\,H_i^2\,|i\rangle + \langle i|\,(2k-2i)H_i\,|i\rangle$$

$$= m_{2k,i}^2 + m_{2k,k}^2 + (2k-2i)m_{2k,i}. \quad (56)$$

Step 6.

$$2(m_{2k,i} + k - i)^2 m_{2k\,k}^2 = \langle i|\,(H_i + k - i)^2\,|i\rangle\langle i|\,2H_k^2\,|i\rangle - \langle i|\,(H_i + k - i)\,|i\rangle$$

$$\times \langle i|\,\prod_{\alpha=i+1}^{k-1}(1 + \delta_{i\alpha}^{-1})(B_k^i + C_k^i)(A_k^i + D_k^i)\,|i\rangle. \quad (57)$$

The proof for Steps 1 to 6 is the same as in Pang and Hecht.

Step 7.

$$\langle m|(B_k^m + C_k^m)(A_k^m + D_k^m)\,|m\rangle = \langle m|\,\frac{N_{q_m,h}^{q_m+1,h}N_{q_m+1,h}^{q_m,h}}{(a_{mm}+1)\prod_{p=m+1}^{k-1}a_{mp}(\delta_{mp}+1)\prod_{\mu=1}^{m-1}\delta_{m\mu}a_{m\mu}a_m}\,|m\rangle$$

$$= \langle m|\,\frac{(N_{q_m,h}^{q_m+1,h})^2}{\prod_{p=m+1}^{k-1}a_{mp}^2(\delta_{mp}+1)^2\prod_{\mu=1}^{m-1}a_{m\mu}a_m^2(a_{mm}+1)^2}\,|m\rangle. \quad (58)$$

Step 7 has not been used by Pang and Hecht and is introduced here to eliminate the summation of graphs. The proof is quite simple:

$$\langle m|\,(B_k^m + C_k^m)(A_k^m + D_k^m)\,|m\rangle = \sum_i \langle m|\,(B_k^m + C_k^m)\,|i\rangle\langle i|\,(A_k^m + D_k^m)\,|m\rangle$$

$$= \sum_i \langle m|\,\frac{L_{2k}^m}{\prod_{p=m+1}^{k-1}\delta_{mp}\prod_{\mu=1}^{k-1}b_{m\mu}C_m}\,|i\rangle\langle i|\,\frac{R_{2k}^m}{\prod_{p=1}^{m-1}\delta_{mp}}\,|m\rangle. \quad (58a)$$

Because of Eqs. (20)–(35), the only intermediate state that survives is $|i\rangle = |m+1\rangle$. The reason is that the primitive roots of $O(2k-1)$, with which R_{2k}^m and L_{2k}^m commute, contain also the primitive roots of $O(x)$, $2 \leq x < 2k - 1$. We have seen, for example, how D_k^{k-1} and E_{2k+1}^k in $O(2k+1)$ combine to give A_k^{k-1}, which

is a primitive root of $O(2k)$. Hence any other intermediate state that is not $|m + 1\rangle$ and that gives nonzero matrix elements for R_{2k}^m and L_{2k}^m will be incompatible with Eqs. (20)–(35). Since the canonical decomposition of $O(n) \supset O(n - 1) \supset \cdots \supset O(2)$ is complete, these properties apply to a complete set of intermediate states. Incidentally, this remark also applies to the unitary group. Thus, Step 7 can be applied to Nagel and Moshinsky's work to obtain their second recursion relation Eq. (6.6) in Ref. 6, without the use and proof of Eqs. (4.2a′, b′, a″ and b″).

From (56), (57), and (58), we obtain

$$(N_{q_m-1,h}^{q_m,h})^2 = 2(h_m - q_m + 1)(h_m + q_m + 2k - 2m - 1)(h_k + q_m + k - m - 1)(q_m - h_k + k - m - 1)$$

$$\times (2q_m + 2k - 2m - 1)(q_m + k - m - 1) \prod_{\mu=1}^{m-1} (q_m + h_\mu + 2k - m - \mu - 1)^2$$

$$\times \prod_{\mu=m+1}^{k-1} (q_m + h_\mu + 2k - m - \mu - 1)^2 (q_m - h_\mu + \mu - m)(q_m - h_\mu + \mu - m - 1). \tag{59}$$

From (59) and (51) we obtain (41). From (41) and (47) we obtain (42).

(iv) For $O(2k + 1)$

$$\left\langle \begin{matrix} h \\ q \end{matrix} \right| L_{2k+1}^m = i \left\langle \begin{matrix} h \\ q \end{matrix} \right| \prod_{p=m+1}^{k} \varepsilon_{mp} a_{mp} \prod_{\mu=1}^{m-1} a_{m\mu} a_{mm} \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \left\langle \begin{matrix} h \\ q \end{matrix} \right| F_{2k+1}^m, \tag{60}$$

$$\left\langle \begin{matrix} h \\ q \end{matrix} \right| R_{2k+1}^m = i \left\langle \begin{matrix} h \\ q \end{matrix} \right| \prod_{p=1}^{m-1} \varepsilon_{mp} \left| \begin{matrix} h \\ q \end{matrix} \right\rangle \left\langle \begin{matrix} h \\ q \end{matrix} \right| E_{2k+1}^m, \tag{61}$$

$$N_{q_m+1}^{q_m} = \left\langle q_m \right| \frac{-\prod_{p=1}^{m-1} \varepsilon_{mp}}{\prod_{p=m+1}^{k} (1 + \varepsilon_{mp})(a_{mp} + 1) \prod_{\mu=1}^{m-1} (a_{m\mu} + 1)(a_{mm} + 2)} \left| q_m \right\rangle N_{q_m}^{q_m+1}, \tag{62}$$

$$\langle q_m + 1, q_{m'} - 1 | [R_{2k+1}^m, L_{2k+1}^{m'}] | q_m, q_{m'} \rangle = 0 \quad m \neq m', \tag{63}$$

$$(N_{q_\mu - \delta_{\mu m} - \delta_{\mu m'}}^{q_\mu - \delta_{\mu m}})^2 = \left\langle q_m \right| \frac{(\varepsilon_{mm'} + 1)(a_{mm'} - 1)}{\varepsilon_{mm'} a_{mm'}} \left| q_m \right\rangle (N_{q_\mu - \delta_{\mu m}}^{q_\mu})^2 \tag{64}$$

$$(N_{q_\mu - \delta_{\mu m} - \delta_{\mu m'}}^{q_\mu - \delta_{\mu m'}})^2 = \left\langle q_m \right| \frac{(\varepsilon_{mm'} + 1)(a_{mm'} - 1)}{\varepsilon_{mm'} a_{mm'}} \left| q_m \right\rangle (N_{q_\mu - \delta_{\mu m}}^{q_\mu})^2, \tag{65}$$

where $m' > m$.

$$(N_{q_m-1}^{q_m})^2 = (N_{q_m-1,h}^{q_m,h})^2 \prod_{\mu=1}^{m-1} \frac{(h_\mu - q_m + m - \mu + 1)(q_m + q_\mu + 2k - m - \mu)}{(q_\mu - q_m + m - \mu + 1)(q_m + h_\mu + 2k - m - \mu)}$$

$$\times \prod_{\mu=m+1}^{k} \frac{(q_m - q_\mu + \mu - m)(q_m + q_\mu + 2k - m - \mu)}{(q_m - h_\mu + \mu - m)(q_m + h_\mu + 2k - m - \mu)}, \quad (h_\alpha \geq q_\alpha). \tag{66}$$

$(N_{q_m-1,h}^{q_m,h})^2$, which appears on the right-hand side of Eq. (66), is obtained through the following steps:

Step 1. The Casimir Invariant C_{2k+1},

$$C_{2k+1} = \sum_{i>j}^{2k+1} J_{ij}^2 = 2 \sum_{i=1}^{k} F_{2k+1}^i E_{2k+1}^i + \sum_{0<i<j}^{k} (B_j^i + C_j^i)(A_j^i + D_j^i)$$

$$+ \sum_{0<i<j}^{k} 2F_{2j-1}^i E_{2j-1}^i + \sum_{i=1}^{k} H_i^2 + \sum_{i=1}^{k} (2k - 2i + 1)H_i. \tag{67}$$

Step 2.

$$\langle i| F_{2k+1}^{p} E_{2k+1}^{p} |i\rangle = \langle i| \, \mathcal{E}_{ip}^{-1} \prod_{\alpha=i+1}^{p} (1 + \mathcal{E}_{i\alpha}^{-1}) F_{2k+1}^{i} E_{2k+1}^{i} |i\rangle. \quad (p > i). \tag{68}$$

Step 3.

$$\langle i| \sum_{j\geq i}^{k} F_{2k+1}^{j} E_{2k+1}^{j} |i\rangle = \langle i| \prod_{\alpha=i+1}^{k} (1 + \mathcal{E}_{i\alpha}^{-1}) F_{2k+1}^{i} E_{2k+1}^{i} |i\rangle. \tag{69}$$

Step 4.

$$\langle i| 2\sum_{j\geq i}^{k} F_{2k+1}^{j} E_{2k+1}^{j} |i\rangle = \langle i| (m_{2k+1,i} - H_i)(m_{2k+1,i} + H_i + 2k - 2i + 1) |i\rangle. \tag{70}$$

Step 5.

$$\langle m| F_{2k+1}^{m} E_{2k+1}^{m} |m\rangle = \langle m| \frac{(N_{q_m,h}^{q_m+1,h})^2}{\prod_{p=m+1}^{k} (\mathcal{E}_{mp} + 1)^2 (a_{mp} + 1)^2 \prod_{\mu=1}^{m-1} (a_{m\mu} + 1)^2 (a_{mm} + 2)^2} |m\rangle. \tag{71}$$

Step 6.

$$(N_{q_m-1,h}^{q_m,h})^2 = \tfrac{1}{2}(2q_m + 2k - 2m)^2 (h_m - q_m + 1)(h_m + q_m + 2k - 2m) \prod_{\mu=1}^{m-1} (q_m + h_\mu + 2k - m - \mu)^2$$

$$\times \prod_{\mu=m+1}^{k} (q_m - h_\mu + \mu - m)(q_m - h_\mu + \mu - m - 1)(q_m + h_\mu + 2k - m - \mu)^2. \tag{72}$$

From (72) and (66), we obtain (43). From (43) and (62), we obtain (44).

E. Matrix Elements of $J_{n,n-1}$

This follows almost exactly the same way as in Pang and Hecht.

$$\textit{1. Evaluation of} \quad \left\langle \begin{matrix} m_{n,i} \\ m'_{n-1,i} \\ m_{n-1,i} \end{matrix} \right| J_{n,n-1} \left| \begin{matrix} m_{n,i} \\ m_{n-1,i} \\ m_{n-1,i} \end{matrix} \right\rangle$$

(a) For $n = 2k$

In place of Eq. (5.3) in Pang and Hecht we have

$$\left\{ O_{2k,k} - \left[2H_k \prod_{\alpha=1}^{k-1} a_\alpha^{(2k)} + \prod_{\alpha=1}^{k-1} L_{2k-1}^{\alpha} R_{2k}^{\alpha} h_i' \right] \right\} \left| \begin{matrix} m_{2k,i} \\ m_{2k-1,i} \end{matrix} \right\rangle = 0. \tag{73}$$

From

$$[E_{2k-1}^{j}, O_{2k,k}] \left| \begin{matrix} m_{2k,i} \\ m_{2k-1,i} \end{matrix} \right\rangle = 0 \quad (j = 1, 2, \cdots k - 1), \tag{74}$$

we obtain

$$h_j' = \frac{i\sqrt{2} \left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right| \prod_{\alpha=1}^{k-1} a_\alpha^{(2k)} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle}{N^{(2k-1)} \begin{matrix} (h_j = m_{2k-1,j} + 1) \\ q_j(= m_{2k-1,j}) \\ q_j + 1 \end{matrix} \quad N^{(2k-1)} \begin{matrix} (h_j = m_{2k-1,j} + 1) \\ q_j + 1 \\ q_j(= m_{2k-1,j}) \end{matrix}},$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} + 1 \\ m_{2k-1,j} \end{matrix} \right| J_{2k}^{2k-1} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = \frac{-iN^{(2k)} \begin{matrix} (h_j = m_{2k,j}) \\ q_j(= m_{2k-1,j}) \\ q_j + 1 \end{matrix}}{\sqrt{2} N^{(2k-1)} \begin{matrix} (h_j = m_{2k-1,j} + 1) \\ q_j(= m_{2k-1,j}) \\ q_j + 1 \end{matrix}},$$

$$= \frac{-i}{(m_{2k-1,j} + k - j)(2m_{2k-1,j} + 2k - 2j + 1)^{\frac{1}{2}}} \Bigg| (m_{2k-1,j} + m_{2k,j} + 2k - 2j)$$

$$\times (m_{2k,j} - m_{2k-1,j})(m_{2k-1,j} - m_{2k,k} + k - j)(m_{2k,k} + m_{2k-1,j} + \cdot k - j)$$

$$\times \prod_{\mu=1}^{j-1} \frac{(m_{2k,\mu} - m_{2k-1,j} + j - \mu)(m_{2k,\mu} + m_{2k-1,j} + 2k - \mu - j)}{(m_{2k-1,\mu} - m_{2k-1,j} + j - \mu)(m_{2k-1,\mu} + m_{2k-1,j} + 2k - \mu - j)}$$

$$\times \prod_{\mu=j+1}^{k-1} \frac{(m_{2k-1,j} - m_{2k,\mu} + \mu - j)(m_{2k-1,j} + m_{2k,\mu} + 2k - \mu - j)}{(m_{2k-1,j} - m_{2k-1,\mu} + \mu - j)(m_{2k-1,j} + m_{2k-1,\mu} + 2k - \mu - j)} \Bigg|^{\frac{1}{2}}. \tag{75}$$

From Eq. (73),

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right| J_{2k}^{2k-1} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = m_{2k,k} \prod_{\alpha=1}^{k-1} \frac{(m_{2k,\alpha} + k - \alpha)}{(m_{2k-1,\alpha} + k - \alpha)}. \tag{76}$$

(b) For $n = 2k + 1$

In place of Eq. (5.11) of Pang and Hecht we have

$$\left\{ J_{2k+1}^{2k} - \left[\sum_{\alpha=1}^{k-1} L_{2k}^{\alpha} R_{2k+1}^{\alpha} h_{\alpha}' + R_{2k+1}^{k} h_{k}' + L_{2k+1}^{k} h_{-k}' \right] \right\} \left| \begin{matrix} m_{2k+1,i} \\ m_{2k,i} \end{matrix} \right\rangle = 0. \tag{77}$$

Commuting the left-hand side of Eq. (77) with $A_k^j + D_k^j$, we obtain h_j'. Commuting Eq. (77) with H_k, we obtain h_k' and h_{-k}',

$$h_j' = \frac{-i\sqrt{2}}{N^{(2k)} \begin{matrix} (h_j = m_{2k,j} + 1) \\ q_j(= m_{2k,j}) \\ q_j + 1 \end{matrix} \quad N^{(2k)} \begin{matrix} (h_j = m_{2k,j} + 1) \\ q_j + 1 \\ q_j(= m_{2k,j}) \end{matrix}} \quad j = 1, 2, \cdots k - 1, \tag{78}$$

$$h_k' = \frac{-i}{\sqrt{2} \prod_{p=1}^{k-1} (m_{2k,k} - m_{2k,p} + p - k)}, \tag{79}$$

$$h_{-k}' = \frac{-i}{\sqrt{2} \prod_{\mu=1}^{k-1} (m_{2k,k} + m_{2k,\mu} + k - \mu)}. \tag{80}$$

From Eq. (77) we obtain

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} + 1 \\ m_{2k,j} \end{matrix} \right| J_{2k+1}^{2k} \left| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k,j} \end{matrix} \right\rangle = \frac{-i}{2} \Bigg| \frac{(m_{2k+1,j} - m_{2k,j})(m_{2k+1,j} + m_{2k,j} + 2k - 2j + 1)(m_{2k,j} + k - j)}{(m_{2k,j} - m_{2k,k} + k - j + 1)(m_{2k,j} + m_{2k,k} + k - j)}$$

$$\times \prod_{\mu=1}^{j-1} \frac{(m_{2k+1,\mu} - m_{2k,j} + j - \mu)(m_{2k,j} + m_{2k+1,\mu} + 2k - \mu - j + 1)}{(m_{2k,\mu} - m_{2k,j} + j - \mu)(m_{2k,j} + m_{2k,\mu} - \mu - j + 1)}$$

$$\times \prod_{\mu=j+1}^{k} \frac{(m_{2k,j} - m_{2k+1,\mu} + \mu - j)(m_{2k,j} + m_{2k+1,\mu} + 2k - \mu - j + 1)}{(m_{2k,j} - m_{2k,\mu} + \mu - j)(m_{2k,j} + m_{2k,\mu} + 2k - \mu - j + 1)} \Bigg|^{\frac{1}{2}}$$

$$(j = 1, 2, \cdots k). \tag{81}$$

2. Evaluation of $\left\langle \begin{matrix} m'_{n-1,i} \\ m_{n-2,i} \end{matrix} \right| V^{(n-1)} \left| \begin{matrix} m_{n-1,i} \\ m_{n-2,i} \end{matrix} \right\rangle \left(\begin{matrix} \beta_i \equiv m_{n-1,i} \\ \gamma_i \equiv m_{n-2,i} \end{matrix} \right)$

This follows exactly the same way as in Pang and Hecht. We obtain:

(a) For $n = 2k$

$$\left\langle \begin{matrix} \beta_1 \cdots \beta_{j-1} & \beta_j + 1 & \beta_{j+1} \cdots \beta_{k-1} \\ \gamma_1 \cdots \gamma_{j-1} & \gamma_j & \gamma_{j+1} \cdots \gamma_{k-1} \end{matrix} \right| V^{(2k-1)} \left| \begin{matrix} \beta_1 \cdots \beta_j \cdots \beta_{k-1} \\ \gamma_1 \cdots \gamma_j \cdots \gamma_{k-1} \end{matrix} \right\rangle$$

$$= \left| \prod_{i=1}^{k-1} \frac{(\beta_j - \gamma_i + i - j + 1)(\beta_j + \gamma_i + 2k - i - j - 1)}{(\beta_j - \beta_i + i - j + 1)(\beta_j + \beta_i + 2k - i - j - 1)} \right|^{\frac{1}{2}}, \quad (82)$$

$$\left\langle \begin{matrix} \beta_1 \cdots \beta_i \cdots \beta_{k-1} \\ \gamma_1 \cdots \gamma_i \cdots \gamma_{k-1} \end{matrix} \right| V^{(2k-1)} \left| \begin{matrix} \beta_1 \cdots \beta_i \cdots \beta_{k-1} \\ \gamma_1 \cdots \gamma_i \cdots \gamma_{k-1} \end{matrix} \right\rangle = \prod_{i=1}^{k-1} \frac{\gamma_i + k - i - 1}{\beta_i + k - i - 1}. \quad (83)$$

(b) For $n = 2k + 1$

$$\left\langle \begin{matrix} \beta_1 \cdots \beta_{j-1} & \beta_j + 1 & \beta_{j+1} \cdots \beta_{k-1} & \beta_k \\ \gamma_1 \cdots \gamma_{j-1} & \gamma_j & \gamma_{j+1} \cdots \gamma_{k-1} \end{matrix} \right| V^{(2k)} \left| \begin{matrix} \beta_1 \cdots \beta_j \cdots \beta_{k-1} & \beta_k \\ \gamma_1 \cdots \gamma_j \cdots \gamma_{k-1} \end{matrix} \right\rangle$$

$$= \left| \prod_{i=1}^{k-1} \frac{(\beta_j - \gamma_i + i - j + 1)(\beta_j + \gamma_i + 2k - i - j)}{(\beta_j - \beta_i + i - j + 1)(\beta_j + \beta_i + 2k - i - j)} \right|^{\frac{1}{2}}. \quad (84)$$

Finally, we also obtain the same results as in Pang and Hecht for the matrix elements of J_n^{n-1}, i.e., their Eqs. (5.44), (5.45), and (5.46). With

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

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

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

$$\equiv A_{2k}^{2k-1} = -i \left| \frac{\prod_{a=1}^{k-1} (l_{2k-2,a}^2 - l_{2k-1,j}^2) \prod_{b=1}^{k} (l_{2k,b}^2 - l_{2k-1,j}^2)}{l_{2k-1,j}^2 (4l_{2k-1,j}^2 - 1) \prod_{a \neq j}^{k-1} (l_{2k-1,a}^2 - l_{2k-1,j}^2)[(l_{2k-1,a} - 1)^2 - l_{2k-1,j}^2]} \right|^{\frac{1}{2}}. \quad (85)$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \right| J_{2k}^{2k-1} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,i} \end{matrix} \right\rangle \equiv B_{2k}^{2k-1} = \frac{\prod_{a=1}^{k-1} l_{2k-2,a} \prod_{a=1}^{k} l_{2k,a}}{\prod_{a=1}^{k-1} l_{2k-1,a} (l_{2k-1,a} - 1)}. \quad (86)$$

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

$$\equiv C_{2k+1}^{2k} = \frac{-i}{2} \left| \frac{\prod_{a=1}^{k-1} (l_{2k-1,a} - l_{2k,j} - 1)(l_{2k-1,a} + l_{2k,j}) \prod_{b=1}^{k} (l_{2k+1,b} - l_{2k,j} - 1)(l_{2k+1,b} + l_{2k,j})}{\prod_{a \neq j}^{k} (l_{2k,a}^2 - l_{2k,j}^2)[l_{2k,a}^2 - (l_{2k,j} + 1)^2]} \right|^{\frac{1}{2}}. \quad (87)$$

From these equations, we see that A_{2k}^{*2k-1} can be obtained from A_{2k}^{2k-1} by changing $m_{2k-1,j} + 1$ to $m_{2k-1,j}$, and i to $-i$. Similarly, C_{2k+1}^{*2k} can be obtained from C_{2k+1}^{2k} by changing $m_{2k,j} + 1$ to $m_{2k,j}$ and i to $-i$. Moreover, matrix elements of the other generators can be expressed in terms of A, B, C in Eqs. (85), (86), (87), and the commutation relations in Eq. (1). Thus for example,

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} + 1 \\ m_{2k-1,i} + 1 \end{matrix} \middle| J_{2k+1}^{2k-1} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k-1,i} \end{matrix} \right\rangle = iC_{2k+1}^{2k}A_{2k}^{2k-1} \left| \frac{(l_{2k,j} + l_{2k-1} + 1)}{(l_{2k,j}^2 - l_{2k-1,i}^2)(l_{2k-1,i}^2 - l_{2k,j} - 1)} \right|^{\frac{1}{2}}, \tag{88}$$

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} + 1 \\ m_{2k-1,i} - 1 \end{matrix} \middle| J_{2k+1}^{2k-1} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k-1,i} \end{matrix} \right\rangle = iC_{2k+1}^{2k}A_{2k}^{*2k-1} \left| \frac{(l_{2k-1,i} - l_{2k,j} - 2)}{(l_{2k-1,i} + l_{2k,j})(l_{2k,j}^2 - l_{2k-1,i}^2)} \right|^{\frac{1}{2}}, \tag{89}$$

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} + 1 \\ m_{2k-1,j} \end{matrix} \middle| J_{2k+1}^{2k-1} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = -iC_{2k+1}^{2k}B_{2k}^{2k-1} \frac{1}{l_{2k,j}}, \tag{90}$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} + 1 \\ m_{2k-2,i} + 1 \end{matrix} \middle| J_{2k}^{2k-2} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,i} \end{matrix} \right\rangle = -iA_{2k}^{2k-1}C_{2k-1}^{2k-2} \left| \frac{(l_{2k-2,i} + l_{2k-1,j} + 1)}{(l_{2k-2,i}^2 - l_{2k-1,j}^2)(l_{2k-1,j} - l_{2k-2,i} - 1)} \right|^{\frac{1}{2}}, \tag{91}$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} + 1 \\ m_{2k-2,i} - 1 \end{matrix} \middle| J_{2k}^{2k-2} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,i} \end{matrix} \right\rangle = iA_{2k}^{2k-1}C_{2k-1}^{*2k-2} \left| \frac{(l_{2k-1,j} - l_{2k-2,i} + 1)}{(l_{2k-2,i}^2 - l_{2k-1,j}^2)(l_{2k-1,j} + l_{2k-1,i} - 1)} \right|^{\frac{1}{2}}, \tag{92}$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,i} + 1 \end{matrix} \middle| J_{2k}^{2k-2} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,i} \end{matrix} \right\rangle = iB_{2k}^{2k-1}C_{2k-1}^{2k-2} \frac{1}{l_{2k-2,i}}. \tag{93}$$

Matrix elements of the other generators can then be written down by inspection.

In conclusion, we list some of the differences between our method and that of Pang and Hecht.

1. We have made use of the concept of primitive roots in the definition, construction, and proof of the lowering and raising operators of $O(n)$, thus making a closer contact with the work of Nagel and Moshinsky in their treatment of the unitary group.

2. We have obtained the lowering (raising) operators of $O(n)$ without using graphs. There are some minor differences between our operators and those of Pang and Hecht, mainly because we define H_α as $J_{2\alpha,2\alpha-1}$, while Pang and Hecht define H_α as $J_{2\alpha-1,2\alpha}$.

3. We have obtained the normalization constants by essentially combining the methods of Nagel and Moshinsky and Pang and Hecht. Thus we have found, for example, that the summation of graphs is not necessary.

In all other aspects we follow closely the method of Pang and Hecht, especially in the evaluation of the matrix elements of $J_{n,n-1}$.

ACKNOWLEDGMENT

APPENDIX

Proof that the Lowering and Raising Operators Satisfy Eqs. (20)—(35)

The proof, though lengthy, is in principle straightforward. The method is as follows. Take a particular primitive root, say D_j^{j-1}, and write down all the terms in the lowering (raising) operator that do not commute with it, and sum them up. The sum must be zero.

A simple example is to compute the result of commuting D_{j+1}^j $(m < j \le k - 2)$ through the second term in line of L_{2k}^m. This gives exactly the same results as in Nagel and Moshinsky's proof for the lowering operator of the unitary group.

A more instructive example, typical for the orthogonal group, is to compute the result of commuting D_j^{j-1} $(m+1 < j \leq k-1)$ through the 6th and 7th terms in line of L_{2k}^m. There are four terms, which after D_j^{j-1} has been commuted through, leave the form

$$\sum_{i'=0}^{k} \sum_{v_{i'} > \cdots > v_i=1}^{k-1} \sum_{i=0}^{j-m-1} \sum_{\mu_i > \cdots > \mu_1=m+1}^{j-2} C_{\mu_1}^m C_{\mu_2}^{\mu_1} \cdots C_{\mu_i}^{\mu_{i-1}} C_{j-1}^{\mu_i} B_j^{j-1}(-C_j^{v_{i'}})(-C_{v_{i'}}^{v_{i'}-1}) \cdots (-C_{v_2}^{v_1})[-(A_k^{v_1} + D_k^{v_1})]$$

$$\times \prod_{x=1}^{i} \mathcal{E}_{m\mu_x}^{-1} \prod_{y=1}^{i'} b_{mv_y}^{-1} \prod_{l=m+1}^{k-1} \mathcal{E}_{ml} \prod_{\mu=1}^{k-1} b_{m\mu} C_m. \quad \text{(A1)}$$

These four terms are

$$\cdots C_j^{\mu} B_{j-1}^j (-C_{j-1}^{v_i'}) \cdots \mathcal{E}_{mj}^{-1} b_{mj-1}^{-1} \cdots, \quad \text{(A2)}$$

$$\cdots C_{j-1}^{\mu_i} B_j^{j-1}(-C_j^{v_{i'}}) \cdots \mathcal{E}_{mj-1}^{-1} b_{mj}^{-1} \cdots, \quad \text{(A3)}$$

$$\cdots C_{j-1}^{\mu_i} B_j^{j-1}(-C_j^{j-1})(-C_{j-1}^{v_i'}) \cdots \mathcal{E}_{mj-1}^{-1} b_{mj}^{-1} b_{mj-1}^{-1} \cdots, \quad \text{(A4)}$$

$$\cdots C_{j-1}^{\mu_i} C_j^{j-1} B_{j-1}^j (-C_j^{v_i'}) \cdots \mathcal{E}_{mj-1}^{-1} \mathcal{E}_{mj}^{-1} b_{mj-1}^{-1} \cdots. \quad \text{(A5)}$$

After commuting D_j^{j-1} through, remembering that $B_{j-1}^j = -B_j^{j-1}$, we obtain from Eqs. (A2)–(A5) the common factor in (A1) multiplied by the sum

$$(-\mathcal{E}_{m,j-1}^{-1} b_{m,j-1}^{-1} \mathcal{E}_{m,j}^{-1} b_{m,j}^{-1})(\mathcal{E}_{m,j-1} b_{m,j} - \mathcal{E}_{m,j} b_{m,j-1} + \mathcal{E}_{m,j} \mathcal{E}_{j-1,j} + b_{m,j} \mathcal{E}_{j-1,j})$$

$$= (-\mathcal{E}_{m,j-1}^{-1} b_{m,j-1}^{-1} \mathcal{E}_{m,j}^{-1} b_{m,j}^{-1})(b_{m,j} \mathcal{E}_{m,j} - \mathcal{E}_{m,j} b_{m,j}) = 0.$$

Other cases are similar to the example given above.

Geometrical View of Lagrange Multipliers in Mechanics

J. R. Gaskill, Jr.*

Bell Telephone Laboratories, Holmdel, New Jersey

AND

M. Arenstein

Stevens Institute of Technology, Hoboken, New Jersey

A set of constraints on a mechanical system may be viewed as a set of surfaces in the phase space of the problem. To be satisfied simultaneously by the system, these constraints must intersect forming a differentiable manifold. Lagrange multipliers are eigenvalues which adjust the magnitudes of gradient vectors in the manifold's tangent bundle. A Jacobian transformation on the constraints must therefore exist and be nonsingular over the domain of equilibrium.

1. INTRODUCTION

MOST physical problems of interest deal with systems which are subject to constraints. These constraints are usually implicit ones and are initially used in setting up the equations describing the system. In mechanics, statistical mechanics, and elsewhere, the constraints, at times are dealt with explicitly and the technique of Lagrange multipliers is used in examining them. An investigation of the geometric role played by Lagrange multipliers leads to a requirement which must be satisfied by the set of constraints acting on a system. In particular, this is shown to hold for discrete mechanical systems subject to holonomic and certain nonholonomic constraints.

Expressed in general terms, a set of constraints on a system is representable as a set of surfaces in a generalized coordinate space. To be simultaneously satisfied by the system, the surfaces must intersect forming a differentiable manifold in the coordinate space. The vector space, formed in the manifold's tangent bundle is then the basis of the associated phase space. Lagrange multipliers are eigenvalues which adjust the magnitudes of gradient vectors in the tangent bundle so that a static or dynamic balance is achieved. To guarantee the desired behavior in the tangent vector space and the existence of the intersection itself, a nonsingular Jacobian transformation on the set of constraints must exist at the point of static equilibrium or over a domain of points corresponding to the range of dynamic equilibrium.

Attention is first directed to the case of equilibrium for an n-dimensional system subject to a single constraint only; the required geometric configuration

is established. Then the more general problem of an n-dimensional system subject to a set of k simultaneous constraints is treated. The requirement that a nonsingular Jacobian on the set of holonomic constraint equations exist over the domain of equilibrium then becomes evident. This requirement leads to a necessary condition governing a certain class of nonholonomically constrained discrete systems. It appears that this criteria might be extended to cover a more general class of constrained systems including continuous systems.

2. MECHANICAL SYSTEM UNDER A SINGLE CONSTRAINT

Consider the problem of a particle or system having n degrees of freedom constrained holonomically by

$$\varphi(q_1, q_2, \cdots, q_n) = 0 \tag{2.1}$$

and acted upon by a set of active forces having the resultant

$$\mathbf{Q}^a = \sum_{i=1}^{n} \hat{u}_i Q_i(q_1, q_2, \cdots, q_n), \tag{2.2}$$

where $\{\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_n\}$ are the basis vectors of the space. For static equilibrium, the virtual work, to the first order of small quantities is zero, i.e.,

$$\mathbf{Q}^a \cdot \delta \mathbf{U} = 0 = \delta W. \tag{2.3}$$

The constraint, $\varphi = 0$, may be interpreted to mean that the particle or system is constrained to move only on the surface in coordinate space described by $\varphi = 0$. Then any virtual displacements about equilibrium are confined to a hyperplane (in 3 dimensions— a plane) tangent to the surface $\varphi = 0$. Thus the tangent hyperplane is normal to the gradient of the surface $\varphi = 0$ at the point of equilibrium and therefore

* Present address: Hughes Aircraft Company, Culver City, California.

the gradient is normal to all possible virtual displacements. Thus,

$$\nabla\varphi \cdot \delta\mathbf{U} = 0. \qquad (2.4)$$

Then the requirement that $\delta W = 0$ means that $\mathbf{Q}^a$ is normal to $\delta\mathbf{U}$ and since $\delta\mathbf{U}$ is normal to $\nabla\varphi$, $\mathbf{Q}^a$ must be colinear with (and thus linearly dependent on) $\nabla\varphi$. Thus,

$$\mathbf{Q}^a = \lambda'\nabla\varphi \quad \text{or} \quad \mathbf{Q}^a + \lambda\nabla\varphi = 0. \qquad (2.5)$$

The Lagrange multiplier λ, is now seen to be an eigenvalue which adjusts the magnitude of the eigenvector along $\nabla\varphi$.

In Eq. (2.5) $\mathbf{Q}^a$ is the resultant active force; $\lambda\nabla\varphi$ is also dimensionally a force and is clearly the reaction force of the constraining surface. Accordingly, $\lambda\nabla\varphi$ is called the reaction force of constraint.[1] Then in each direction,

$$Q_i^a + \lambda(\partial\varphi/\partial q_i) = 0; \quad i = 1, 2, \cdots, n. \qquad (2.6)$$

For the case of dynamic equilibrium, suppose that $\mathbf{Q}^a$ is partially conservative so that

$$\mathbf{Q}^a = -\nabla V + \mathbf{Q}_{nc}^a, \qquad (2.7)$$

in which $\mathbf{Q}_{nc}^a$ is the nonconservative force resultant (zero for conservative systems). Then using the Lagrangian formulation, the equation of equilibrium becomes

$$\left(\sum_{i=1}^{n} \hat{u}_i\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{q_i}\right]\right) - \mathbf{Q}_{nc}^a - \lambda\nabla\varphi = 0 \qquad (2.8)$$

or along each direction

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} - Q_{inc}^a - \lambda\frac{\partial\varphi}{\partial q_i} = 0;$$

$$i = 1, 2, \cdots, n. \qquad (2.9)$$

This presupposes that the space is such that Eq. (2.9) may hold simultaneously over all n directions. That is, does the phase space of $\{q_i, \dot{q}_i \mid i \in [1, n: \text{integers}]\}$ exist?

It has been pointed out[2] that if the coordinate space of a discrete classical problem is a manifold M, then its tangent bundle "forms" the associated phase space. The explicit requirement that the coordinate space be a differentiable manifold, needed here to

guarantee that Eq. (2.9) is valid, becomes more salient in the case of multiple constraint.

3. MULTIPLE CONSTRAINT

Consider again the problem of a system of n degrees of freedom acted upon by a set of forces resultant in

$$\mathbf{Q}^a = \sum_{i=1}^{n} \hat{u}_i Q_i(q_1, q_2, \cdots, q_n).$$

In this case the motion is subject to a set of holonomic constraints

$$\{\varphi_j(q_1, q_2, \cdots, q_n) = 0 \mid j \in [1, k: \text{integers}; k < n]\}. \qquad (3.1)$$

In coordinate space, each $\varphi_j = 0$ describes a surface on which the system is constrained to move. To satisfy the set (3.1) simultaneously, the motion is restricted to some $\Phi = 0$ which is the intersection of all the $\varphi_j = 0$. That is motion is restricted to

$$\Phi = \bigcap_{j=1}^{K} \varphi_j = 0. \qquad (3.2)$$

Then by the preceding argument, for static equilibrium,

$$\mathbf{Q}^a + \lambda'\nabla\Phi = 0. \qquad (3.3)$$

If Φ is a manifold in a neighborhood of the equilibrium point, or a series of connectable neighborhoods for dynamic equilibrium, then the tangent bundle, vector space, on $\Phi = 0$ will be so behaved that $\nabla\Phi$ will lie in the "volume" formed by and be a linear combination of the gradients of the individual component surfaces, $\{\varphi_j = 0\}$.[3] That is,

$$\nabla\Phi = \sum_{j=1}^{K} \lambda'\nabla\varphi_j. \qquad (3.4)$$

Thus, Eq. (3.3) becomes $\mathbf{Q}^a + \Lambda'(\sum_{j=1}^{K} \lambda'\nabla\varphi_j) = 0$, and letting

$$\Lambda'\lambda_j' = \lambda_j, \quad \mathbf{Q}^a + \sum_{j=1}^{K} \lambda_j\nabla\varphi_j = 0. \qquad (3.5)$$

For dynamic equilibrium, again using the Lagrangian formulation,

$$\left(\sum_{i=1}^{n} \hat{u}_i\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right]\right) - \mathbf{Q}_{nc} - \sum_{j=1}^{K} \lambda_j\nabla\varphi_j = 0. \qquad (3.6)$$

[1] J. L. Lagrange, *Méchanique Analytique* (Courcier, Paris, 1811), Vol. I, Sec. IV, p. 76.
[2] R. Herman, J. Math. Phys. 6, 1768 (1965).

[3] N. Haaser, J. LaSalle, and J. Sullivan, *Intermediate Analysis* (Blaisdell Publishing Company, New York, 1964), pp. 242–247.

In matrix form, Eq. (3.6) becomes

$$\begin{bmatrix} \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_1}\right) - \dfrac{\partial L}{\partial q_1} - Q_{1nc} \\ \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_2}\right) - \dfrac{\partial L}{\partial q_2} - Q_{2nc} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_K}\right) - \dfrac{\partial L}{\partial q_K} - Q_{Knc} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_n}\right) - \dfrac{\partial L}{\partial q_n} - Q_{nnc} \end{bmatrix} - \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial q_1} & \dfrac{\partial \varphi_2}{\partial q_1} & \cdots & \dfrac{\partial \varphi_K}{\partial q_1} \\ \dfrac{\partial \varphi_1}{\partial q_2} & \dfrac{\partial \varphi_2}{\partial q_2} & \cdots & \dfrac{\partial \varphi_K}{\partial q_2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ \dfrac{\partial \varphi_1}{\partial q_K} & \dfrac{\partial \varphi_2}{\partial q_K} & \cdots & \dfrac{\partial \varphi_K}{\partial q_K} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ \dfrac{\partial \varphi_1}{\partial q_n} & \dfrac{\partial \varphi_2}{\partial q_n} & \cdots & \dfrac{\partial \varphi_K}{\partial q_n} \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_K \end{bmatrix} = 0, \qquad (3.7)$$

where $\Phi = \bigcap_{j=1}^{K} \varphi_j = 0$, and the intersection of each $\varphi_j = 0$ must be a manifold.

For any general

$$\Phi = \bigcap_j^K \varphi_j = 0$$

to be a manifold in a neighborhood of some point P_0, each surface, $\varphi_j = 0$, must have a continuous partial derivative at P_0 and in addition the Jacobian transformation (from the surfaces to the tangent vector space) must be nonsingular. That is,

$$\Delta_{KK} = \begin{vmatrix} \dfrac{\partial \varphi_1}{\partial q_1} & \dfrac{\partial \varphi_1}{\partial q_2} & \cdots & \dfrac{\partial \varphi_1}{\partial q_K} \\ \dfrac{\partial \varphi_2}{\partial q_1} & \dfrac{\partial \varphi_2}{\partial q_2} & \cdots & \dfrac{\partial \varphi_2}{\partial q_K} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ \dfrac{\partial \varphi_K}{\partial q_1} & \dfrac{\partial \varphi_K}{\partial q_2} & \cdots & \dfrac{\partial \varphi_K}{\partial q_K} \end{vmatrix}_{P_0} \neq 0. \quad (3.8)$$

This transformation defines the mapping of the variables which become dependent in the intersection to the remaining independent set by mapping both to the manifold formed by the intersection.[4]

The rectangular matrix on the right of Eq. (3.7) contains (the transposes of) all possible Jacobian transformations on the constraint surfaces. It may be partitioned to form Kth order Jacobian transformations. For Eq. (3.7) to be valid, there must exist

[within the rectangular matrix on the right of Eq. (3.7)] some nonsingular Kth order Jacobian. The variables included, are the ones which become dependent in the intersection and the remaining variables ($n - k$ of them) are independent.

Conversely, if all $\Delta_{KK} = 0$, for all K-dimensional sets, the point P_0 is singular and the intersection $\Phi = 0$ may not exist, or if it exists will not be well defined, e.g., discontinuous or nondifferentiable.

4. NONHOLONOMICALLY CONSTRAINED CASES

There are problems[5] of interest in which the surfaces $\{\varphi_j = 0\}$ are not known explicitly, but in which the constraints may be expressed as

$$\left(\sum_{j=1}^{n} a_{lj}(q_1, q_2, \cdots, q_n, t)\, dq_j\right)$$
$$+ a_{lt}(q_1, q_2, \cdots, q_n, t)\, dt = 0$$
$$\text{for} \quad l = 1, 2, \cdots, K. \quad (4.1)$$

From previous arguments, it is clear that the equilibrium points can lie only on the intersection of some set of constraint surfaces and that the forces of constraint are expressible in the tangent bundle of intersection as

$$\sum_{j=1}^{K} \lambda_j \nabla \varphi_j.$$

Thus, each a_{lj} in Eq. (4.1) must be some

$$\left.\frac{\partial \varphi_l}{\partial q_j}\right|_{P_0}.$$

Therefore, to represent a valid set of constraints,

[4] M. E. Monroe, *Modern Multidimensional Calculus*, (Addison-Wesley Publishing Company, Reading, Massachusetts, 1963), pp. 156, 157, 161–171, 180–182.

[5] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1950), pp. 40–43.

there must be nonsingular Kth-order transformations within

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{n1} & & & a_{nK} \end{bmatrix};$$

although variation is taken with time fixed, this must hold at each instant over the domain of time which is applicable to the problem.

It would seem possible to extend these principles to nonholonomic constraints expressed as inequalities for discrete systems and similarly to systems defined in terms of density functions in as much as similar things are done in modern control theory.

Two-Body Orbitals for One-Dimensional Fermion Gas with Application to Repulsive δ-Function Interactions

E. KAPUY

Research Group for Theoretical Physics, The Hungarian Academy of Sciences, Budapest, Hungary

AND

N. H. MARCH

Department of Physics, The University, Sheffield, England

(Received 13 January 1967)

Explicit strongly orthogonal two-body orbitals (geminals) are constructed from free-particle Wannier functions for the lowest singlet state of a one-dimensional homogeneous fermion gas. As an application of the method, the ground-state energy is evaluated for repulsive delta function interactions for a range of the coupling constant. For weak coupling, the present method yields a lower energy than that of the Overhauser state. Using second-order perturbation corrections as a means of comparison, it is found that, in the high-density limit, the strongly orthogonal geminal product gives only a fifth of the total correlation energy. In the strong coupling regime, the results are more favorable but it proves difficult to determine the asymptotic behavior.

1. INTRODUCTION

WHILE some definite progress has been made in the search for a valid variational method based on low-order density matrices[1-5] rather than on a many-body wavefunction, the necessary and sufficient conditions for a trial second-order density matrix (which determines the total energy when only two-body interactions are involved) are in a form which, with present techniques, render them rather unsuitable for practical applications.[6] The variational validity can at present be established with certainty only if one starts directly with trial many-body wavefunctions which allow the density matrices, and hence the energy, to be calculated without further approximation.

Young and March[3] constructed a two-particle density matrix from suitable two-body orbitals. We outline, in Sec. 2, an alternative scheme in which strongly orthogonal two-body functions (referred to subsequently as geminals) are employed in constructing the antisymmetric total wavefunction, while in Sec. 3 we apply this method to a homogeneous one-dimensional fermion gas, building the geminals from free-particle Wannier functions.

To obtain detailed results for physically interesting quantities such as energy, density, and momentum distribution, we have worked out, in Sec. 4, results for the case of repulsive δ-function interactions between the fermions. Unfortunately, the exact solution of the problem is not yet known.[7] The results are compared with the Overhauser spin density wave state and, in the weak coupling limit, with second-order perturbation corrections.

[1] J. E. Mayer, Phys. Rev. **100**, 1579 (1955).
[2] R. H. Tredgold, Phys. Rev. **105**, 1421 (1957).
[3] W. H. Young and N. H. March, Proc. Roy. Soc. (London) A**256**, 62 (1960).
[4] B. C. Carlson and J. M. Keller, Phys. Rev. **121**, 659 (1961).
[5] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
[6] C. Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).
[7] J. B. McGuire, J. Math. Phys. **5**, 622 (1964); **6**, 432 (1965).

there must be nonsingular Kth-order transformations within

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{n1} & & & a_{nK} \end{bmatrix} ;$$

although variation is taken with time fixed, this must hold at each instant over the domain of time which is applicable to the problem.

It would seem possible to extend these principles to nonholonomic constraints expressed as inequalities for discrete systems and similarly to systems defined in terms of density functions in as much as similar things are done in modern control theory.

Two-Body Orbitals for One-Dimensional Fermion Gas with Application to Repulsive δ-Function Interactions

E. KAPUY

Research Group for Theoretical Physics, The Hungarian Academy of Sciences, Budapest, Hungary

AND

N. H. MARCH

Department of Physics, The University, Sheffield, England

(Received 13 January 1967)

Explicit strongly orthogonal two-body orbitals (geminals) are constructed from free-particle Wannier functions for the lowest singlet state of a one-dimensional homogeneous fermion gas. As an application of the method, the ground-state energy is evaluated for repulsive delta function interactions for a range of the coupling constant. For weak coupling, the present method yields a lower energy than that of the Overhauser state. Using second-order perturbation corrections as a means of comparison, it is found that, in the high-density limit, the strongly orthogonal geminal product gives only a fifth of the total correlation energy. In the strong coupling regime, the results are more favorable but it proves difficult to determine the asymptotic behavior.

1. INTRODUCTION

WHILE some definite progress has been made in the search for a valid variational method based on low-order density matrices[1-5] rather than on a many-body wavefunction, the necessary and sufficient conditions for a trial second-order density matrix (which determines the total energy when only two-body interactions are involved) are in a form which, with present techniques, render them rather unsuitable for practical applications.[6] The variational validity can at present be established with certainty only if one starts directly with trial many-body wavefunctions which allow the density matrices, and hence the energy, to be calculated without further approximation.

Young and March[3] constructed a two-particle density matrix from suitable two-body orbitals. We outline, in Sec. 2, an alternative scheme in which strongly orthogonal two-body functions (referred to subsequently as geminals) are employed in constructing the antisymmetric total wavefunction, while in Sec. 3 we apply this method to a homogeneous one-dimensional fermion gas, building the geminals from free-particle Wannier functions.

To obtain detailed results for physically interesting quantities such as energy, density, and momentum distribution, we have worked out, in Sec. 4, results for the case of repulsive δ-function interactions between the fermions. Unfortunately, the exact solution of the problem is not yet known.[7] The results are compared with the Overhauser spin density wave state and, in the weak coupling limit, with second-order perturbation corrections.

[1] J. E. Mayer, Phys. Rev. **100**, 1579 (1955).
[2] R. H. Tredgold, Phys. Rev. **105**, 1421 (1957).
[3] W. H. Young and N. H. March, Proc. Roy. Soc. (London) A256, 62 (1960).
[4] B. C. Carlson and J. M. Keller, Phys. Rev. **121**, 659 (1961).
[5] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
[6] C. Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).

[7] J. B. McGuire, J. Math. Phys. **5**, 622 (1964); **6**, 432 (1965).

2. STRONGLY ORTHOGONAL GEMINALS

We take a set of geminals $\psi_I(1, 2)$, $I = 1, 2, \cdots, N$, which satisfy the strong orthogonality conditions[8]

$$\int \psi_I^*(1, 2)\psi_J(1, 3) \, d1 = 0, \quad J \neq I. \quad (2.1)$$

Here $1, 2, \cdots$, refers to both space and spin coordinates, while ψ_I is assumed to be normalized to unity, and antisymmetrical such that

$$\psi_I(1, 2) = -\psi_I(2, 1). \quad (2.2)$$

Then we may construct a normalized antisymmetric $2N$-body wavefunction of the form

$$\Psi = \left[\frac{2^N}{(2N)!}\right]^{\frac{1}{2}}$$
$$\times \sum_P (-1)^P P\psi_1(1, 2)\psi_2(3, 4)\cdots\psi_N(2N - 1, 2N), \quad (2.3)$$

where P indicates permutations which interchange the particles between the geminals. This wavefunction takes into account the simultaneous correlations of N pairs including the corresponding unlinked clusters such that even if we pass to an infinite system $N \to \infty$, we do not reproduce the Hartree–Fock approximation unless the two-body functions are simple products of one-electron spin orbitals.

If we confine ourselves to two-body interactions, we can write the Hamiltonian formally as

$$H = H(0) + \sum_{\alpha=1}^{2N} H(\alpha) + \sum_{\beta>\alpha=1}^{2N} H(\alpha, \beta). \quad (2.4)$$

The expectation value of H is then given by[8-13]

$$E = H(0) + \sum_{I=1}^N \int \psi_I^*(1, 2)$$
$$\times [H(1) + H(2) + H(1, 2)]\psi_I(1, 2) \, d1 \, d2$$
$$+ 2 \sum_{\substack{I,J=1 \\ (J \neq I)}}^N \int d1 \, d2 \, d3 \, d4 H(1, 3)[1 - P_{13}]$$
$$\times \psi_I^*(1', 2)\psi_I(1, 2)\psi_J^*(3', 4)\psi_J(3, 4). \quad (2.5)$$

The corresponding expressions for the first- and second-order density matrices γ and Γ are given by[11-13]

$$\gamma(1', 1) = \sum_{I=1}^N \gamma_I(1', 1) = 2 \sum_{I=1}^N \int \psi_I^*(1', 2)\psi_I(1, 2) \, d2, \quad (2.6)$$

$$\Gamma(1', 2', 1, 2)$$
$$= \sum_{I=1}^N \psi_I^*(1', 2')\psi_I(1, 2)$$
$$+ \tfrac{1}{4} \sum_{\substack{I,J=1 \\ (J \neq I)}}^N \{\gamma_I(1', 1)\gamma_J(2', 2) + \gamma_I(2', 2)\gamma_J(1', 1)$$
$$- \gamma_I(2', 1)\gamma_J(1', 2) - \gamma_I(1', 2)\gamma_J(2', 1)\}. \quad (2.7)$$

3. ONE-DIMENSIONAL HOMOGENEOUS MANY-FERMION PROBLEM

We specialize now to the case of $2N$ fermions contained in a box of length L and interacting via a two-body potential $V(x_1, x_2)$. Neglecting the interactions, the orbitals are obviously given by

$$\phi_k = L^{-\frac{1}{2}} \exp(ikx), \quad (3.1)$$

where $k = 2\pi n/L$, $n = 0$, ± 1, ± 2, etc. Assuming that N is odd, the orbitals described by $n = 0$, $\pm 1, \cdots, \pm\frac{1}{2}(N - 1)$ are doubly occupied in the ground state, and the others are empty.

In order to form a more suitable localized basis than the plane waves (3.1), we introduce a lattice (with spacing $x_0 = L/N$) and free-particle Wannier functions corresponding to the occupied plane-wave orbitals. It can be shown that, in the case of δ-function interactions considered in Sec. 4 below, they are the "optimum localized" functions in the sense of Edmiston and Ruedenberg.[14] When $N \to \infty$, $L \to \infty$ such that $N/L = \frac{1}{2}\rho$ remains finite, they are given by

$$w_1^I(x) = (k_0\pi)^{-\frac{1}{2}} \frac{\sin k_0(x - Ix_0)}{x - Ix_0}, \quad I = 0, \pm 1, \cdots \quad (3.2)$$

where $k_0 = \pi/x_0 = \frac{1}{2}\pi\rho$ is the Fermi wavenumber. These Wannier functions are equivalent to the wane-wave orbital (3.1) of the first "band," with $-k_0 \leq k \leq k_0$. For higher bands, the Wannier functions are given by

$$w_i^I(x) = (k_0\pi)^{-\frac{1}{2}}$$
$$\times \left[\frac{\sin k_0 i(x - Ix_0) - \sin k_0(i - 1)(x - Ix_0)}{x - Ix_0}\right],$$
$$i = 1, 2, \cdots, \quad (3.3)$$

and satisfy the orthogonality condition

$$\int w_i^I(x)w_j^J(x) \, dx = \delta_{IJ}\delta_{ij}. \quad (3.4)$$

[8] A. C. Hurley, J. E. Lennard-Jones, and J. A. Pople, Proc. Roy. Soc. (London) A220, 496 (1953).

[9] J. M. Parks and R. G. Parr, J. Chem. Phys. 28, 335 (1958).

[10] R. G. Parr, *Quantum Theory of Molecular Electronic Structure* (W. A. Benjamin, Inc., New York, 1963).

[11] E. Kapuy, Acta Phys. Acad. Sci. Hung. 9, 237 (1958).

[12] R. McWeeny, Proc. Roy. Soc. (London) A253, 242 (1959).

[13] R. McWeeny, Rev. Mod. Phys. 32, 335 (1960).

[14] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. 35, 457 (1963).

If we require that all the geminals be singlets and equivalent to one another, then the most general two-body function may be written in the form

$$\psi_I = 2^{-\frac{1}{2}} \sum_{i,i'} C_{ii'} w_i I(x_1) w_{i'} I(x_2) [\alpha(1)\beta(2) - \alpha(2)\beta(1)].$$
(3.5)

Here the $C_{ii'}$'s are symmetrical in the indices ii' and satisfy the normalization condition

$$\sum_{i,i'} C_{ii'}^2 = 1,$$
(3.6)

while α and β are the usual spin wavefunctions. With this choice of the coefficients, the total wavefunction constructed from the two-body functions (3.5) is invariant under translations of any integral multiple of x_0.

In the following considerations the basis w_i^I is regarded as fixed and only the coefficients $C_{ii'}$ are optimized.

The main advantage of this model is that any product of Wannier functions belonging to the same lattice site

$$w_i^I(x_1) w_j^I(x_2) \cdots w_i^I(x_m)$$

can be summed up over the lattice sites. Thus, the Wannier functions can be written in the form

$$w_i^I(x) = (4k_0\pi)^{-\frac{1}{2}} \int_{(i-1)k_0}^{ik_0} [e^{ik(x-Ix_0)} + e^{-ik(x-Ix_0)}] \, dk.$$
(3.7)

It is then easy to show that

$$\sum_{I=-\infty}^{+\infty} e^{-ikx_0 I} = \frac{2\pi}{x_0} \sum_{M=-\infty}^{+\infty} \delta(k - 2Mk_0),$$

where M is an integer: $0, \pm 1, \pm 2, \cdots, \pm\infty$. By using this formula, we can carry out the summations for the density matrices in (2.6) and (2.7).

The spinless first-order density matrix has the form

$$\rho(x', x) = 2 \sum_I \sum_{i,i',j} C_{i'j} C_{ij} w_i^I(x') w_i^I(x).$$
(3.8)

There are three different cases:
(1) $i' = i$ (diagonal terms)

$$\sum_I w_i^I(x') w_i^I(x)$$

$$= \frac{1}{\pi} \frac{\sin k_0 i(x' - x) - \sin k_0(i - 1)(x' - x)}{x' - x};$$
(3.9)

(2) $i' > i$ (both are even or odd)

$$\sum_{I-} w_i^I(x') w_i^I(x) = \frac{1}{\pi} \left\{ \frac{\sin k_0[i'(x' - x) - (i' - i)x']}{x' - x} \right.$$

$$\left. - \frac{\sin k_0[(i' - 1)(x' - x) - (i' - i)x']}{x' - x} \right\};$$
(3.10)

(3) $i' > i$ (one is even and the other is odd)

$$\sum_I w_{i'}^I(x') w_i^I(x)$$

$$= \frac{1}{\pi} \left\{ \frac{\sin k_0[i'(x' - x) - (i' + i - 1)x']}{x' - x} \right.$$

$$\left. - \frac{\sin k_0[(i' - 1)(x' - x) - (i' + i - 1)x']}{x' - x} \right\}.$$
(3.11)

The corresponding densities $(x' \to x)$ are as follows:

$$\sum_I w_i^I(x) w_i^I(x) = \frac{1}{x_0},$$
(3.9a)

$$\sum_I w_{i'}^I(x) w_i^I(x) = \frac{1}{x_0} \cos (i' - i) k_0 x,$$
(3.10a)

$$\sum_I w_{i'}^I(x) w_i^I(x) = \frac{1}{x_0} \cos (i' + i - 1) k_0 x.$$
(3.11a)

We observe that only the diagonal terms give a homogeneous density. The contribution of the off-diagonal elements is inhomogeneous through the factor $\cos mk_0 x$, where $m = +2, +4, +6, \cdots$, etc.

We can show similarly that the second- and higher-order density matrices are also inhomogeneous.

The momentum distribution can be calculated by using the formula

$$P(k) = (2\pi)^{-1} \int \rho(x', x) e^{-ik(x'-x)} \, dx' \, dx.$$

The general form of the latter for interacting particles in this approach is displayed in Fig. 1. The
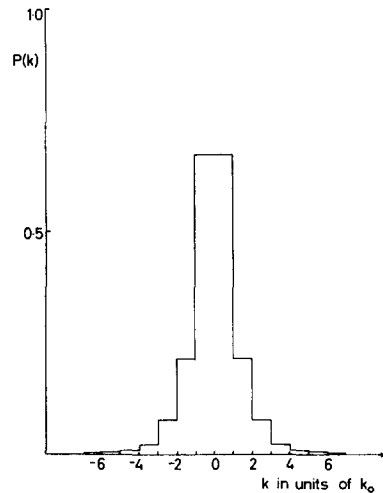

FIG. 1. Momentum distribution $\left(\text{normalized such that}\right.$

$$\int_{-\infty}^{+\infty} P(k) \, dk = 1 \Big)$$

for an interacting fermion gas described by strongly orthogonal geminals based on free-particle Wannier functions.

discontinuities at values of $\pm k_0$, $\pm 2k_0$, $\pm 3k_0$, $\cdots$ of the curve follow from the peculiarities of the basis.

4. CASE OF δ-FUNCTION INTERACTION

As the wavefunction (3.5) is invariant under translations by Mx_0, where M is an integer, it is sufficient to calculate only one term, say the Ith, of the energy expression (2.5). Dividing this part by two we get ϵ, the energy per particle. Thus, we need only sum up the first-order density matrix over the infinite lattice. Since

$$2 \sum_{J \neq I} \sum_{i',i,j} C_{i'j} C_{ij} w_{i'}^J(x') w_i^J(x)$$

$$= \rho(x', x) - 2 \sum_{i',i,j} C_{i'j} C_{ij} w_{i'}^I(x') w_i^I(x),$$

and the summation of $\rho(x', x)$ can be carried out explicitly (see Sec. 3), the expression for ϵ no longer contains summations over lattice sites.

The δ-function interaction $V_0 \delta(x_1 - x_2)$ further simplifies the treatment because nonvanishing interaction exists only between pairs of particles having antiparallel spins.

Introducing the following abbreviations:

$$\int \rho(x, x) w_i^I(x) w_j^I(x)\, dx = \rho_{ij},$$

$$\int w_i^I(x) w_j^I(x) w_k^I(x) w_l^I(x)\, dx = (ijkl),$$

and

$$\eta = V_0/k_0 \pi,$$

we find the following energy expression:

$$\epsilon = \tfrac{1}{2}\left\{ \sum_{i,j} C_{ij}^2 [i^2 - i + \tfrac{1}{3}] + \eta\left[\frac{\pi}{k_0} \sum_{i,j,k,l} C_{ij} C_{kl}(ijkl)\right.\right.$$

$$+ \frac{\pi}{2k_0} \sum_{i,j,k} C_{ij} C_{kj} \rho_{ik}$$

$$\left.\left. - \frac{\pi}{k_0} \sum_{\substack{i,j,k,l \\ m,n}} C_{ij} C_{kj} C_{lm} C_{nm}(ikln)\right]\right\}. \quad (4.1)$$
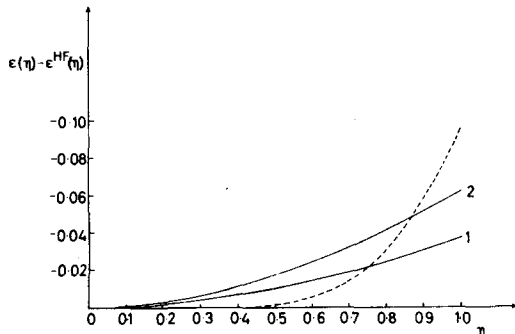
FIG. 2. Energies relative to that of the normal state versus η: (1) strongly orthogonal geminal product in first approximation of Sec. 4B. ———— (2) strongly orthogonal geminal product in second approximation of Sec. 4C; – – – Overhauser state.

In this case we obtain ϵ, the energy per particle, in units of k_0^2.

By minimizing ϵ with respect to the coefficients C_{ij}, for several values of η, $0 \leq \eta \leq \infty$, we can calculate the energy as a function of V_0 and k_0 for the intervals $0 \leq k_0 \leq \infty$ and $0 \leq V_0 \leq \infty$, respectively.

A. Hartree–Fock Approximation

Using the geminal of zeroth order

$$2^{-\frac{1}{2}} w_1^I(x_1) w_1^I(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)], \quad (4.2)$$

we find for the Hartree–Fock normal state the energy

$$\epsilon^{\mathrm{HF}}(\eta) = \tfrac{1}{6} + \tfrac{1}{2}\eta. \quad (4.3)$$

In the strong coupling limit ($\eta \to \infty$), ϵ^{HF} tends to $\tfrac{1}{2}\eta$.

B. First Approximation

We take the first approximate trial function in the form

$$2^{-\frac{1}{2}} \sum_{i,j=1}^{2} C_{ij} w_i^I(x_1) w_j^I(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]. \quad (4.4)$$

As was to be expected, the singly excited configuration gives only an unimportant improvement when $\eta < 1$. This means that then the density $\rho(x', x)$ remains homogeneous, because the off-diagonal elements in the first-order density matrix vanish. $\epsilon_1(\eta) - \epsilon^{\mathrm{HF}}(\eta)$, in the interval $0 \leq \eta \leq 1$, is displayed in Fig. 2. As $\eta \to \infty$, the contribution of the singly excited configuration increases. In the strong coupling limit, $\epsilon_1(\eta)$ tends to 0.3136η which is equivalent to an improvement of 37% on the energy, and the density becomes slightly inhomogeneous

$$\rho_1(x, x) = (1/x_0)(2 + 0.002 \cos 2k_0 x).$$

C. Second Approximation

We then take as the second approximation the trial geminal

$$2^{-\frac{1}{2}} \sum_{i,j=1}^{3} C_{ij}\, w_i^I(x_1) w_j^I(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]. \quad (4.5)$$

When $0 \leq \eta \leq 1$, the contribution from the singly excited configurations is small but the density is still inhomogeneous. In particular the presence of the term $C_{23}[w_2^I(x_1) w_3^I(x_2) + w_2^I(x_2) w_3^I(x_1)]$ gives a non-negligible $\cos 4k_0 x$ component. In this interval, $\epsilon_2(\eta) - \epsilon^{\mathrm{HF}}(\eta)$ is displayed in Fig. 2. As $\eta \to \infty$ we obtain $\epsilon_2(\eta) = 0.2223\eta$ (an improvement of 55% on the energy). At the same time, we find for the density

$$\rho_2(x, x) = (1/x_0)(2 + 0.576 \cos 2k_0 x$$

$$+ 0.264 \cos 4k_0 x).$$

This means that the density is contracted towards the lattice sites. It seems probable that the contraction will be enhanced in the higher approximations. The convergence is rather slow. By using this method it seems to be very difficult (if not impossible) to determine the correct asymptotic behavior as $\eta \to \infty$.

For comparison the energy of the corresponding Overhauser state[15] $\epsilon^0(\eta)$ has been calculated. This is always lower than the normal (HF) state and $\epsilon^0(\eta) - \epsilon^{HF}(\eta)$ is displayed in Fig. 2. We see from this that, in the interval $0 \leq \eta \leq 0.87$, the singlet state described by the trial function (4.5) is lower in energy than the Overhauser state. Using higher approximations the intersection of the two curves is shifted towards larger values of η.

D. Calculation of Exact Second-Order Energy Corrections in High-Density Limit Using Perturbation Theory

As we have seen above, the convergence on increasing the number of configurations is rather slow. Therefore, it seemed of interest to sum up over all possible configurations corresponding to the geminal product ground state. This sum can be carried out exactly in this model at the high-density limit. Geminal (4.2) is the zeroth-order solution of the Hamiltonian operator

$$H^I(\eta) = -\frac{1}{2}\left[\frac{d^2}{dx_1^2} + \frac{d^2}{dx_2^2}\right] + \eta\{\pi k_0\delta(x_1 - x_2) + 2k_0^2$$
$$-\pi k_0[(w_1^I(x_1))^2 + (w_2^I(x_2))^2]\}. \quad (4.6)$$

We obtain for the second-order correction per particle $\epsilon^{(2)}$

$$\epsilon^{(2)} = \frac{1}{2}\sum_\mu \frac{H_{0\mu}^I(\eta)^2}{H_{00}^I(\eta) - H_{\mu\mu}^I(\eta)}. \quad (4.7)$$

Here μ means all possible configurations of the following form:

$$2^{-\frac{1}{2}}w_i^I(x_1)w_j^I(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)], \quad i, j > 1.$$

Only those configurations have nonzero contributions for which $|i - j| < 3$. Summing up over all terms we get

$$\epsilon^{(2)} = -0.1272\eta^2. \quad (4.8)$$

Our first and second approximations (4.4) and (4.5) at the high-density limit give 34 and 58% of this contribution, respectively.

We can carry out exactly the summation over all interpair correlations of order η^2 which conform to the strong orthogonality conditions[16]. Here, only the

[15] A. W. Overhauser, Phys. Rev. Letters **4**, 414, 462 (1960).
[16] E. Kapuy, Theoret. Chim. Acta Berl. **6**, 281 (1966).

singly excited singlet and triplet geminals contribute:

$$\psi_{Ii}^S = 2^{-1}[w_1^I(x_1)w_i^I(x_2) + w_1^I(x_2)w_i^I(x_1)]$$
$$\times [\alpha(1)\beta(2) - \alpha(2)\beta(1)],$$
$$\psi_{Ii}^T = 2^{-1}[w_1^I(x_1)w_i^I(x_2) - w_1^I(x_2)w_i^I(x_1)]$$
$$\times \begin{cases} \alpha(1)\alpha(2), \\ \beta(1)\beta(2), \\ 2^{-\frac{1}{2}}[\alpha(1)\beta(2) + \alpha(2)\beta(1)]. \end{cases}$$

We then find for the correction per particle, after lengthy but elementary manipulations, the result

$$-0.0270\eta^2, \quad (4.9)$$

which is slightly more than 20% of $\epsilon^{(2)}$.

Using free-particle orbitals (3.1), we can calculate exactly the total second-order perturbation correction per particle $\epsilon_T^{(2)}$ to the (singlet) normal state. We have

$$\epsilon_T^{(2)} = -\eta^2\pi^2/12 = -0.8225\eta^2.$$

The correction given by (4.8) plus (4.9) therefore accounts for about 19% of $\epsilon_T^{(2)}$. This is due to the neglect of all configurations which violate the strong orthogonality conditions.

E. Estimate of Effect of Configurations Containing Basis Functions of Other Lattice Sites

Using the above second-order perturbation calculation, it is possible to estimate the effect of Wannier functions centered on other lattice sites in the weak coupling limit. It is equivalent to dropping the strong orthogonality conditions (2.1) imposed on the geminals of the product (2.3). If we restrict ourselves to the first two bands, the only configuration conforming to the strong orthogonality conditions is

$$2^{-\frac{1}{2}}w_2^I(x_1)w_2^I(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)].$$

It gives the correction per particle as $-0.0434\eta^2$. The other possible configurations all violate the strong orthogonality conditions. They are the following:

$$2^{-\frac{1}{2}}w_2^J(x_1)w_2^K(x_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)], \quad J, K \neq I,$$
$$2^{-1}[w_2^I(x_1)w_2^J(x_2) + w_2^I(x_2)w_2^J(x_1)][\alpha(1)\beta(2) - \alpha(2)\beta(1)],$$
$$J \neq I.$$

The lower bound of their contribution to the second-order correction can be calculated exactly. We find for the correction per particle $-0.0567\eta^2$. (When only the nearest neighbors $J = I \pm 1$ are taken into account, $-0.0324\eta^2$ is obtained.) It means that the strongly orthogonal contribution, though appreciable, is only about 43% of the total contribution.

5. CONCLUSION

For the homogeneous fermion system with infinitely short-range interaction, we have shown that, for weak coupling, $\eta < 0.9$, our method based on geminals built from Wannier functions always leads to a lower energy than Overhauser's state. However, in the strong coupling regime an infinite number of terms in our trial geminals would be required to determine the asymptotic behavior. In the weak coupling limit, the method gives only a fraction of the total correlation energy due to the fact that the model is "weakly localizable." More favorable results can, of course, be expected for systems which consist of spatially localized fermion pairs.

It would, we believe, be of interest to apply this method to a three-dimensional problem with realistic interactions. However, in this case, the completely filled first band is usually not equivalent to the Hartree–Fock ground state. Thus, it would probably not be worthwhile to employ Wannier functions with, say, cubic symmetry. Instead, a more fruitful approach might be to use the spherically averaged Wannier functions as March and Young[17] did.

[17] N. H. March and W. H. Young, Phil. Mag. **4**, 384 (1959).

Decompositions of Discrete Most Degenerate Representations of
$SO_0(p, q)$ when Restricted to Representations of $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$

J. Niederle*

International Atomic Energy Agency, International Centre for Theoretical Physics,
Trieste, Italy

(Received 20 February 1967)

The decomposition of any most degenerate unitary irreducible representation (single-valued of discrete principal series) of an arbitrary noncompact rotation group $SO_0(p,q)$ ($p \geq q > 1$) when restricted to the unitary irreducible representations of its maximal noncompact rotation subgroup $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$ is derived, and characteristic features of the decomposition are discussed.

1. INTRODUCTION

SOME physical considerations (e.g., the study of the scattering amplitude of two particles in the t channel) lead to the questions: What is the decomposition of a given irreducible representation of a noncompact group when restricted to irreducible representations of its noncompact subgroup, and, in particular, what irreducible representation of a subgroup appears and how many times. Such problems have been studied to the best of our knowledge only for irreducible representations of the Lie algebra of noncompact unitary groups by using Gel'fand–Tsetlin patterns[1] and for irreducible unitary representations of the de Sitter group[2] $SO(2, 3)$. (For related problems see also Ref. 3.) In our work we decompose a most degenerate unitary irreducible representation (single-valued of a discrete principal series) of an arbitrary noncompact rotation group[4] $SO_0(p,q)$ ($p \geq q > 1$) into the unitary irreducible representations of its maximal noncompact rotation subgroup $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$.

In order to find the desired decomposition, it is sufficient to know how the carrier space of the unitary irreducible representation of $SO_0(p,q)$ decomposes into subspaces each of which is a carrier space of a unitary irreducible representation of $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$. For this purpose we use the approach developed in Refs. 5–7. We restrict ourselves to the

decomposition of the discrete series of the most degenerate representations of $SO_0(p, q)$ ($p \geq q > 1$), since they seem to be of physical interest. That is why we define the representation of $SO_0(p, q)$ on a Hilbert space $\mathcal{H}(X)$ of square integrable functions, the domain of which is a homogeneous space X of rank one under the action of $SO_0(p, q)$. From the definition of a group representation we derive the representation of the corresponding Lie algebra $\mathcal{R}(p, q)$ in terms of differential operators acting on some linear manifold $\mathcal{D}(X)$, which is dense in $\mathcal{H}(X)$ (for more details see Sec. 2). The Gårding theorem[8] guarantees that every representation of the algebra $\mathcal{R}(p, q)$ induces a representation of the group $SO_0(p, q)$. For the most degenerate representation of $SO_0(p, q)$ the ring of invariant operators of the Lie algebra $\mathcal{R}(p, q)$ is generated by only one invariant operator,[9] which is the Laplace–Beltrami operator (see Ref. 10). Analogously to the works[5,6] we can use the generalized Fourier transforms of functions $f \in \mathcal{D}(X)$ with respect to eigenfunctions of the Laplace–Beltrami operator as a basis for the natural carrier space of a discrete and continuous series of representations of $\mathcal{R}(p, q)$. These representations and representations on $\mathcal{D}(X)$ are unitarily equivalent—for more details see Sec. 3. We have only to prove that any infinitesimal representation of $SO_0(p, q)$ acting on such a carrier space (actually its subspace) is Hermitian and irreducible and that every carrier space of an infinitesimal Hermitian irreducible representation of $SO_0(p, q)$ is also a carrier space (after completion) of the corresponding global unitary irreducible representation of $SO_0(p, q)$. This is a brief description of how to construct in general a discrete and a continuous series of the most degenerate irreducible unitary representation of the $SO_0(p, q)$ group by using the method developed in Refs. 5–7. A specification

* On leave of absence from the Institute of Physics of the Czechoslovak Academy of Sciences, Prague.

[1] I. M. Gel'fand and M. I. Graev, "Irreducible Representations of Lie Algebras of $U(p, q)$ Groups" (in Russian), presented at Spring School of Theoretical Physics, Yalta, (1966) 15 April–5 May.

[2] N. T. Evans, J. Math. Phys. 8, 170 (1967).

[3] E. P. Wigner, Ann. Math. 40, 149 (1939); N. Ya. Vilenkin and Ya. A. Smorodinskii, Zh. Eksperim. i Teor. Fiz. 46, 1793 (1964) [English transl.: Soviet Phys.—JETP 19, 1209 (1964)]; F. T. Hadjioannou, CERN Preprint TH. 612 (1965).

[4] We denote a component of the unity of the group $SO(p, q)$ by $SO_0(p, q)$.

[5] R. Rączka, N. Limić, and J. Niederle, J. Math. Phys. 7, 1861 (1966).

[6] N. Limić, J. Niederle, and R. Rączka, J. Math. Phys. 7, 2026 (1966).

[7] N. Limić, J. Niederle, and R. Rączka, J. Math. Phys. 8, 1079 (1967).

[8] K. Maurin, *Metody Hilbertova Prostranstva* (in Russian) (MIR, Moscow, 1965). Chap. X.

[9] I. M. Gel'fand, Am. Math. Soc. Transl. Ser. 2 37, 31 (1964).

[10] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press, Inc., New York, 1962), Chap. X, Sec. 2; I. M. Gel'fand and M. I. Graev, Trudi Moscow Math. Soc. 8, 321 (1959).

of this method for our problems consists of three things: (i) We have to choose a coordinate system on the homogeneous space X in which not only the Laplace–Beltrami operator related to $\mathfrak{R}(p, q)$ is diagonal, but also, that related to $\mathfrak{R}(p, q - 1)$ or $\mathfrak{R}(p - 1, q)$. (ii) We have to construct a basis for the carrier space of a unitary irreducible representation of $SO_0(p, q)$ using the Fourier transforms of functions $f \in \mathfrak{D}(X)$ with respect to the common eigenfunctions of both Laplace–Beltrami operators. (iii) Since we are interested in discrete series of representations of $SO_0(p, q)$, the eigenfunctions of the corresponding Laplace–Beltrami operator have to belong to its discrete spectrum. Then the desired decomposition of the carrier space of a discrete unitary irreducible representation of $SO_0(p, q)$ into the carrier spaces of unitary irreducible representations of $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$ follows automatically.

In Sec. 2 we specify our homogeneous spaces which are of rank one under the action of $SO_0(p, q)$ and define the representations of $SO_0(p, q)$ group and of its corresponding Lie algebra $\mathfrak{R}(p, q)$. Moreover, we solve the eigenvalue problems for the Laplace–Beltrami operators related to our homogeneous spaces and construct the carrier spaces of discrete unitary irreducible representations of $SO_0(p, q)$ and their decomposition as well. Section 3 contains the proof of irreducibility of our representations of $\mathfrak{R}(p, q)$ and the proof that from irreducibility of an infinitesimal Hermitian representation of $SO_0(p, q)$ follows irreducibility of the corresponding global one. Finally, in Sec. 4 we briefly summarize the main results.

2. DISCRETE MOST DEGENERATE UNITARY IRREDUCIBLE REPRESENTATIONS OF $SO_0(p, q)$ AND THEIR DECOMPOSITIONS

A. Homogeneous Spaces X

There have been considered three homogeneous spaces X of rank one under the action of the noncompact rotation group $SO_0(p, q)$ Refs. (4–6, 11):

$$SO_0(p, q)/SO_0(p - 1, q), \quad SO_0(p, q)/SO_0(p, q - 1),$$

$$SO_0(p, q)/T^{p+q-2} \; \boxed{S} \; SO_0(p - 1, q - 1).$$

They can be represented by the hyperboloids H_q^p and H_q^p and by the cone C_q^p, respectively. They are embedded in the $(p + q)$—dimensional pseudo-Euclidean

---

[11] We denote the group of translations in the $(p + q - 2)$-dimensional pseudo-Euclidean space R_{p+q-2}^{q-1} by T^{p+q-1}.

space R_{p+q}^q—and determined by the equation

$$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$$

$$= \begin{cases} 1 & \text{for} \quad H_q^p \quad p \geq q \\ 0 & \text{for} \quad C_q^p \\ -1 & \text{for} \quad H_p^q \quad p \geq q. \end{cases}$$

Since with the cone C_q^p and with the hyperboloid H_p^1 we can relate only continuous series of representations of $SO_0(p, q)$ and $SO_0(p, 1)$, respectively,[7] we consider later only the hyperboloids H_q^p and H_p^q ($p \geq q > 1$) as our homogeneous spaces X.

B. Definition of the Representation of $SO_0(p, q)$ and $R(p, q)$

If $d\mu(\Omega)$ is the Riemannian left-invariant measure on X and $\mathcal{K}(X)$ is the Hilbert space of $L^2(\mu)$ type, then the quasi-regular representation of $SO_0(p, q)$ we define as:

$$SO_0(p, q) \ni g \to (U_g f)(\Omega) = f(g^{-1}\Omega), \quad f \in \mathcal{K}(X).$$
$$(2.1)$$

Such a quasi-regular representation is unitary but not irreducible (see Sec. 3D).

The Lie algebra $\mathfrak{R}(p, q)$ can be expressed in the form of operators x_{ij} of a compact and a noncompact type (l_{ij} and b_{rs}, respectively) satisfying the commutation relations,

$$[l_{ij}, l_{rs}]_- = -\delta_{ir}l_{js} + \delta_{is}l_{jr} + \delta_{jr}l_{is} - \delta_{js}l_{ir},$$
$$[l_{ij}, b_{rs}]_- = -\delta_{ir}b_{js} - \delta_{is}b_{jr} + \delta_{jr}b_{is} + \delta_{js}b_{ir}, \quad (2.2)$$
$$[b_{ij}, b_{rs}]_- = +\delta_{ir}l_{js} + \delta_{is}l_{jr} + \delta_{jr}l_{is} + \delta_{js}l_{ir}.$$

However, if the representation of the $SO_0(p, q)$ group is given by Eq. (2.1), then the corresponding Lie algebra $\mathfrak{R}(p, q)$ can be represented by differential operators X_{ij} of the form

$$L_{ij} = x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i}, \quad \begin{matrix} i, j = 1, \cdots, p \\ \text{or} \\ i, j = p + 1, \cdots, p + q, \end{matrix}$$

$$B_{rs} = x_r \frac{\partial}{\partial x_s} + x_s \frac{\partial}{\partial x_r}, \quad \begin{matrix} r = 1, \cdots, p \\ s = p + 1, \cdots, p + q \end{matrix} \quad (2.3)$$

or vice versa.

The operators L_{ij}, B_{rs} are unbounded operators in the Hilbert space $\mathcal{K}(X)$. Therefore, since we consider the representation of the generators and of their polynomials (e.g., the Casimir operator) we must restrict their domain to some dense linear manifold in $\mathcal{K}(X)$. We take their domain as the linear manifold $\mathfrak{D}(X)$ determined by vectors $f \in \mathcal{K}(X)$ of the form

$$f(\Omega) = P(x_1, \cdots, x_{p+q}) \cdot e^{-\sum_{i=1}^{p+q}(x_i)^2}, \quad f \in \mathfrak{D}(X), \quad (2.4)$$

where $P(x_1, \cdots, x_{p+q})$ is an arbitrary polynomial in $x_1, \cdots, x_{p+q}$ variables. Such a domain $\mathfrak{D}(X)$ is the common invariant domain for all the generators L_{ij}, B_{rs} and is a dense linear manifold in $\mathcal{K}(X)$ (for the proof see, example, Ref. 7).

C. Eigenvalue Problem for the Laplace–Beltrami Operator $\Delta(X)$

The Laplace–Beltrami operator Δ corresponding to a discrete representation of the $SO_0(p, q)$ group is related to either the hyperboloid H_q^p $(p \geq q > 1)$ or to the hyperboloid H_p^q $(p > q > 1)$. (The hyperboloid H_1^p, we do not consider.)

1. Laplace–Beltrami Operator $\Delta(H_q^p)$, $(p \geq q > 1)$

If we introduce the coordinate system on the hyperboloid H_q^p as

$$x_i = x_i' \cosh \eta, \quad i = 1, \cdots, p + q - 1,$$
$$x_{p+q} = \sinh \eta, \quad \eta \in (-\infty, \infty), \quad (2.5)$$

where x_i', $i = 1, \cdots, p + q - 1$, are coordinates on the hyperboloid H_{q-1}^p given in Eqs. 3 and 4 of Ref. 5, then, using the same procedure as in Refs. 5, 6, and 12, we find that the Laplace–Beltrami operator $\Delta(H_q^p)$ is given by

$$\Delta(H_q^p) = \frac{-1}{\cosh^{p+q-2} \eta} \frac{\partial}{\partial \eta} \cosh^{p+q-2} \eta \frac{\partial}{\partial \eta} + \frac{\Delta(H_{q-1}^p)}{\cosh^2 \eta},$$
$$\eta \in (-\infty, \infty). \quad (2.6)$$

Here, $\Delta(H_{q-1}^p)$ is the Laplace–Beltrami operator related to the hyperboloid H_{q-1}^p, which is explicitly written in Eqs. (3 and 11) of Ref. 5. As is shown in Refs. 5–7, the spectrum of $\Delta(H_{q-1}^p)$ is of the form $-\sigma(\sigma + p + q - 3)$ and consists of the discrete spectrum[13]

$$PS[\Delta(H_{q-1}^p)]: \quad \sigma = l, \quad l = -\{\tfrac{1}{2}(p + q - 5)\},$$
$$-\{\tfrac{1}{2}(p + q - 5)\} + 1, \cdots,$$

and the continuous spectrum

$$CS[\Delta(H_{q-1}^p)]: \quad \sigma = i\lambda - \tfrac{1}{2}(p + q - 3), \quad \lambda \in [0, \infty).$$

The explicit forms of the corresponding eigenfunctions of $\Delta(H_{q-1}^p)$ depend essentially on q and are reviewed in the Appendix [Eqs. (A1), (A10), and (A12)].

If we represent the eigenfunctions of $\Delta(H_q^p)$ as a product of the eigenfunctions of $\Delta(H_{q-1}^p)$ and a function $\Psi^{Q,\sigma}(\eta)$, we obtain the following differential equation for the latter function:

$$\left[\frac{-1}{\cosh^{p+q-2} \eta} \frac{d}{d\eta} \cosh^{p+q-2} \eta \frac{d}{d\eta} \right.$$
$$\left. - \frac{\sigma(\sigma + p + q - 3)}{\cosh^2 \eta} - Q \right] \Psi^{Q,\sigma}(\eta) = 0,$$
$$\eta \in (-\infty, \infty). \quad (2.7)$$

Using the transformation

$$\Psi^{Q,\sigma}(\eta) = \cosh^{\frac{1}{2}(2-p+q)} \eta \cdot \psi^{Q,\sigma}(\eta),$$

we derive for $\psi^{Q,\sigma}(\eta)$ the differential equation of a type which has been treated by Titchmarsh.[14] Therefore, we immediately know that both independent solutions $_{1,2}\Psi^{Q,\sigma}(\eta)$ enter into the eigenfunction expansion associated with the differential operator of Eq. (2.7). Moreover, (analogously to Ref. 7) the spectrum Q in (2.7) can be written in the form $Q = -\Sigma(\Sigma + p + q - 2)$, and if σ is from $PS[\Delta(H_{q-1}^p)]$, consists of the discrete spectrum[12]

$$PS[\Delta(H_q^p)]: \quad \Sigma = L, \quad L = -\{\tfrac{1}{2}(p + q - 4)\},$$
$$-\{\tfrac{1}{2}(p + q - 4)\} + 1, -\{\tfrac{1}{2}(p + q - 4)\} + 2, \cdots,$$

as well as the continuous one $CS[\Delta(H_q^p)]: \Sigma = i\Lambda - \tfrac{1}{2}(p + q - 2)$, $\Lambda \in [0, \infty)$, whereas for σ from $CS[\Delta(H_{q-1}^p)]$ it consists only of the continuous spectrum

$$CS[\Delta(H_q^p)]: \quad \Sigma = i\Lambda - \tfrac{1}{2}(p + q - 2), \quad \Lambda \in [0, \infty).$$

Since we are interested in discrete series of representations of $SO_0(p, q)$, we only give here the form of the eigenfunctions of the Laplace–Beltrami operator $\Delta(H_q^p)$ belonging to the discrete spectrum.[15] For $p \geq q > 2$, we have[13]

$$_{(1)}Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{L, l, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[(q-1)/2]}}(\Omega)$$
$$= {}_{(1)}V_l^L(\eta) \cdot Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{l, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[(q-1)/2]}}(\theta, \omega, \tilde{\omega}),$$
$$(2.8)$$

with $L - l = -(2n + 2), \quad n = 0, 1, 2, \cdots,$

$$_{(2)}Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{L, l, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[(q-1)/2]}}(\Omega)$$
$$= {}_{(2)}V_l^L(\eta) \cdot Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{l, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[(q-1)/2]}}(\theta, \omega, \tilde{\omega})$$
$$(2.9)$$

with $L - l = -(2n + 1), \quad n = 0, 1, 2, \cdots,$

---

[12] J. Fischer, J. Niederle, and R. Rączka, J. Math. Phys. 7, 816 (1966).

[13] Here and elsewhere $[x]$ and $\{x\}$ denote the nearest smaller or higher integer than x, respectively.

[14] E. C. Titchmarsh, *Eigenfunction Expansions* (Clarendon Press, Oxford, England, 1962), Pt. I, Sec. 4.19.

[15] The complete set of functions contains, of course, the eigenfunctions of $\Delta(H_q^p)$ belonging to both the discrete and continuous spectra of $\Delta(H_q^p)$. The proof of the completeness of these functions is given in Ref. 7 in another parametrization.

and for $p \geq q = 2$:

$$_{(1,\beta)}Y^{L,l,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega)$$

$$= {}_{(1)}V^L_l(\eta) \cdot {}_{(\beta)}Y^{l,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\theta,\omega), \quad \beta = 1,2,$$
(2.10)

with $L - l = -(2n + 2)$, $n = 0,1,2,\cdots$,

$$_{(2,\beta)}Y^{L,l,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega)$$

$$= {}_{(2)}V^L_l(\eta) \cdot {}_{(\beta)}Y^{l,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{\{p/2\}}}(\theta,\omega), \quad \beta = 1,2,$$
(2.11)

with $L - l = -(2n + 1)$, $n = 0,1,2,\cdots$.

The Y functions on the right-hand sides of Eqs. (2.8)–(2.11) are eigenfunctions of $\Delta(H^p_{q-1})$. The explicit form of all functions in Eqs. (2.8)–(2.11) is given in the Appendix [see (A1), (A10), (A12), and (B1), (B3)].

2. Laplace–Beltrami Operator $\Delta(H^p_q)$, $(p > q > 1)$

The coordinate system on $H^q_p(p > q > 1)$ is introduced in the same way as in 1., Eq. (2.5), but the x'_i are now coordinates on the hyperboloid H^q_{p-1}, which is given in Ref. 5. Hence the Laplace–Beltrami operator $\Delta(H^q_p)$ in this coordinate system has the form

$$\Delta(H^q_p) = \frac{-1}{\cosh^{p+q-2}\eta} \frac{\partial}{\partial\eta} \cosh^{p+q-2}\eta \frac{\partial}{\partial\eta} + \frac{\Delta(H^q_{p-1})}{\cosh^2\eta},$$

$$\eta \in (-\infty, \infty). \quad (2.12)$$

The Laplace–Beltrami operator $\Delta(H^q_{p-1})$ has been investigated in Refs. 5–7, where it has been shown that its spectrum is $-\sigma(\sigma + p + q - 3)$, which for $q > 1$ consists of the discrete part,[12]

$$PS[\Delta(H^q_{p-1})]: \quad \sigma = l, \quad l = -\{\tfrac{1}{2}(p + q - 5)\},$$
$$-\{\tfrac{1}{2}(p + q - 5)\} + 1, \cdots,$$

and the continuous part,

$$CS[\Delta(H^q_{p-1})]: \quad \sigma = i\lambda - \tfrac{1}{2}(p + q - 3), \quad \lambda \in [0, \infty).$$

The eigenfunctions of $\Delta(H^p_{q-1})$ $(p > q > 1)$ we obtain from the eigenfunctions of $\Delta(H^p_{q-1})$ $(p \geq q - 1)$, expressed in the Appendix (A1) and (A10) and (A12) changing $q \rightleftarrows p$, removing the "tilde" from any variable $\vartheta^{\{\cdot\}}$, $\varphi^{[\cdot]}$ or index $l_{\{\cdot\}}$, $m_{[\cdot]}$ which previously had it and at the same time placing a tilde over any variable $\vartheta^{\{\cdot\}}$, $\varphi^{[\cdot]}$ or index $l_{\{\cdot\}}$, $m_{[\cdot]}$ which previously did not have it.

Using the same arguments as in the previous case 1., we obtain the form of the spectrum of $\Delta(H^q_p)$ as $-\Sigma(\Sigma + p + q - 2)$, and for $p > q > 1$, we conclude that it consists of the discrete part,

$$PS[\Delta(H^q_p)]: \quad \Sigma = L, \quad L = -\{\tfrac{1}{2}(p + q - 2)\},$$
$$-\{\tfrac{1}{2}(p + q - 2)\} + 1, \cdots$$

as well as the continuous one,

$$\Sigma = i\Lambda - \tfrac{1}{2}(p + q - 2), \quad \Lambda \in [0, \infty).$$

The eigenfunctions of $\Delta(H^q_p)$ are given as a product of the V function defined in (B1), (B3) and the corresponding eigenfunction of $\Delta(H^q_{p-1})$, analogously to Eqs. (2.7)–(2.10).

D. Discrete Most Degenerate Unitary Irreducible Representation of $SO(p, q)$ and Its Decomposition

The Hilbert space $\mathcal{H}(X)$ of $L^2(\mu)$ type (domain X is our hyperboloid H^p_q or H^q_p) is the carrier space of a discrete and a continuous series of the most degenerate representations of $SO_0(p, q)$. These representations are unitary but not irreducible.[16] In order to find the carrier space of irreducible representations, it is more convenient (for the continuous series it is even necessary) to use, instead of $\mathcal{H}(X)$, its generalized Fourier transform—the Hilbert space[17] $\mathcal{H}(S)$—and then to decompose $\mathcal{H}(S)$ into its subspaces, each of which is a carrier space of an irreducible representation of $SO_0(p, q)$. Since here we are only interested in a discrete series of representations, we only decompose the subspace of $\mathcal{H}(S)$, say $\mathcal{H}_1(S)$, which is a carrier space of a discrete series of representations of $SO_0(p, q)$ into a direct sum of Hilbert spaces, each of which is a carrier space of a discrete irreducible unitary representation of $SO_0(p, q)$. The situation in any particular case of our homogeneous space X is as follows.

1. Hyperboloid H^p_q, $p \geq q > 2$ or $p > q = 2$

The discrete irreducible unitary most degenerate representations of $SO_0(p, q)$ related to H^p_q $(p \geq q > 2$ or $p > q = 2)$ are classified by means of the discrete spectrum of $\Delta(H^p_q)$. Before discussing the decomposition of $\mathcal{H}_1(S)$ into its irreducible subspaces under the action of $SO_0(p, q)$, we have to determine a unitary space $\mathfrak{D}^L$ and then the discrete Hermitian irreducible representation of $\mathcal{R}(p, q)$ and discrete unitary irreducible representation of $SO_0(p, q)$.

For this purpose let us first define our notation. We denote by

$$_{(\gamma)}Y^{L,l,l_{\{p/2\}},l_{\tilde{l}}}_{m_{[\cdot]},\tilde{m}_{[\cdot]}}(\Omega) := {}_{(\gamma)}Y^{L,l,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{(q-1)/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{\{(q-1)/2\}}}(\Omega),$$

where

$$(\gamma) = \begin{cases} (\alpha), & \alpha = 1,2 \quad \text{if} \quad q > 2 \\ (\alpha,\beta), & \alpha,\beta = 1,2 \quad \text{if} \quad q = 2 \end{cases}$$

---

and Y functions are the eigenfunctions of $\Delta(H_q^p)$ given in Eqs. (2.7)–(2.10), and by

$$_{(\gamma)}\chi^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]} := \langle _{(\gamma)}Y^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}, f\rangle$$

$$\equiv \int_{H_q^p} {}_{(\gamma)}\overline{Y^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}}(\Omega) \cdot f(\Omega) \, d\mu(\Omega),$$

where $f \in \mathfrak{D}(H_q^p)$ is defined in Eq. (2.4) with the Ω given by Eq. (2.5) and finally we denote by η_L the set of values of initial $\gamma, l, l_{\{\cdot\}}, \tilde{l}_{\{\cdot\}}, m_{[\cdot]}, \tilde{m}_{[\cdot]}$ restricted by those conditions given by (A3), (A9), (A11), (A13), (B2), and (B4), which are applicable in the particular case considered.

The unitary space $\mathfrak{D}^L$ of l^2 type is determined by sequences

$$\chi^L := \{_{(\gamma)}\chi^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}; \gamma, l, l_{\{\cdot\}}, \tilde{l}_{\{\cdot\}}, m_{[\cdot]}, \tilde{m}_{[\cdot]} \in \eta_L\}.$$

The scalar product and the norm in $\mathfrak{D}^L$ are defined by

$$(\chi^L, \psi^L)_L = \sum_{\eta_L} {}_{(\gamma)}\overline{\chi^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}} \cdot \psi^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}, \quad \chi^L, \psi^L \in \mathfrak{D}^L,$$

and by

$$\|\chi^L\|_L = \sum_{\eta_L} |_{(\gamma)}\chi^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}|^2.$$

The completion of the unitary space $\mathfrak{D}^L$ with respect to the norm $\|\cdot\|_L$ is the Hilbert space $\mathcal{H}^L$.

The discrete unitary irreducible representation of the $SO_0(p, q)$ group and discrete Hermitian irreducible representation of the corresponding Lie algebra $\mathfrak{R}(p, q)$ are defined by

$$SO_0(p, q) \ni g \to U_g^L \chi^L$$

$$:= \{\langle _{(\gamma)}Y^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}, U_g f\rangle;$$

$$\gamma, l, l_{\{\cdot\}}, \tilde{l}_{\{\cdot\}}, m_{[\cdot]}, \tilde{m}_{[\cdot]} \in \eta_L, \quad f \in \mathfrak{D}(H_q^p)\} \in \mathcal{H}^L;$$

$$(2.13)$$

$$\mathfrak{R}(p, q) \ni x_{ij} \to X_{ij}\chi^L$$

$$:= \{\langle _{(\gamma)}Y^{L,l,l\{\cdot\},l\{\cdot\}}_{m[\cdot],\tilde{m}[\cdot]}, X_{ij} f\rangle;$$

$$\gamma, l, l_{\{\cdot\}}, \tilde{l}_{\{\cdot\}}, m_{[\cdot]}, \tilde{m}_{[\cdot]} \in \eta_L, \quad f \in \mathfrak{D}(H_q^p)\} \in \mathfrak{D}^L,$$

$$(2.14)$$

where $(U_g f)$ is determined in (2.1) and X_{ij} in (2.3) with the parametrization on the hyperboloid H_q^p given in Eq. (2.5).

The discrete quasi-regular representation in Eq. (2.1) decomposes into the irreducible representations in the following way:

$$_1U_g = \sum_{L=-\{(p+q-4)/2\}}^{\infty} U_g^L$$

on the Hilbert space

$$_1\mathcal{H}(S) = \sum_{L=-\{(p+q-4)/2\}}^{\infty} \oplus \mathcal{H}^L. \quad (2.15)$$

The proof of irreducibility of U_g^L is presented in Sec. 4. The decomposition of U_g^L with respect to irreducible unitary representations of $SO_0(p, q - 1)$ has the form

$$U_g^L = \sum_{l=L+1}^{\infty} U_g^{L,l}$$

on the Hilbert space

$$\mathcal{H}^L = \sum_{l=L+1}^{\infty} \oplus \mathcal{H}^{L,l}. \quad (2.16)$$

Here, $\mathcal{H}^{L,l}$ are subspaces of $\mathcal{H}^L$ containing the vectors with a fixed eigenvalue l. As is shown in Ref. 5 such spaces are carrier spaces of discrete unitary irreducible representations of the $SO_0(p, q - 1)$ group.

2. Hyperboloid H_p^q, $p > q > 2$

The discrete irreducible unitary representations of $SO_0(p, q)$ are classified now by using the discrete spectra of $\Delta(H_p^q)$. We can again determine the unitary space $\mathfrak{D}^L$ and the discrete irreducible representations of $SO_0(p, q)$ and $\mathfrak{R}(p, q)$ analogously to the previous case 1.

The discrete quasi-regular representation in Eq. (2.1) decomposes now into the discrete irreducible unitary representations in the following way:

$$_1U_g = \sum_{L=-\{(p+q-4)/2\}}^{\infty} U_g^L$$

on the Hilbert space

$$_1\mathcal{H}(S) = \sum_{L=-\{(p+q-4)/2\}}^{\infty} \oplus \mathcal{H}^L. \quad (2.17)$$

The decomposition of $_1U_g^L$ has the form

$$U_g^L = \sum_{l=L+1}^{\infty} U_g^{L,l}$$

on the Hilbert space

$$\mathcal{H}^L = \sum_{l=L+1}^{\infty} \oplus \mathcal{H}^{L,l}, \quad (2.18)$$

where $\mathcal{H}^{L,l}$ are subspaces of $\mathcal{H}^L$ containing the vectors with a fixed eigenvalue l. The irreducibility of U_g^L under the action of $SO_0(p, q)$ is discussed in the next section and the proof of the irreducibility of $U_g^{L,l}$ under the action of $SO_0(p - 1, q)$ is treated in Ref. 5.

3. Hyperboloid H_p^q, $p \geq q = 2$

The classification of the discrete irreducible unitary representation is now given by means of the spectra of two operators $\Delta(H_p^q)$ and $\hat{t}$. The representation of the operator $\hat{t}$ is defined as

$$\hat{T}_{(\alpha)}Y^{L,l,l_2,\cdots,l\{(p-1)/2\}}_{m_1,\cdots,m_{[(p-1)/2]},\tilde{m}_1}(\Omega)$$

$$= (\text{sign } \tilde{m}_1) \cdot {}_{(\alpha)}Y^{L,l,l_2,\cdots,l\{(p-1)/2\}}_{m_1,\cdots,m_{[(p-1)/2]},\tilde{m}_1}(\Omega). \quad (2.19)$$

Denoting by ϵ the doublet of fixed eigenvalues of the operators $\Delta(H_p^q)$ and $\hat{T}$ and by η_ϵ, a subset of

η_L for which $_{(\alpha)}Y^{L,l,\{\cdot\}}_{m[\cdot],\tilde{m}_1}(\Omega)$ belongs to fixed "eigenspace" corresponding to the definite eigenvalue of $\hat{T}$, analogously to Sec. 2Dl, we can define the unitary space $\mathfrak{D}^\epsilon$ and the irreducible representations of $SO_0(p, 2)$, and of $\mathfrak{R}(p, 2)$. The decomposition of the discrete quasi-regular representation of $SO_0(p, q)$ is

$$_1U_g = \sum_{L=-\{(p-2)/2\}}^{\infty} (U_g^{L,+} + U_g^{L,-})$$

on the Hilbert space

$$_1\mathfrak{R}(S) = \sum_{L=-\{(p-2)/2\}}^{\infty} \oplus \mathfrak{K}^{L,+} \oplus \sum_{L=-\{(p-2)/2\}}^{\infty} \oplus \mathfrak{K}^{L,-},$$

$$(2.20)$$

and the decomposition of $U_g^{L,\pm}$ with respect to the discrete irreducible unitary representation of $SO_0(p - 1, 2)$ has the form

$$U_g^{L,\pm} = \sum_{l=L+1}^{\infty} U_g^{L,\pm,l}$$

on the Hilbert space

$$\mathfrak{K}^{L,\pm} = \sum_{l=L+1}^{\infty} \oplus \mathfrak{K}^{L,\pm,l}, \qquad (2.21)$$

where $\mathfrak{K}^{L,\pm,l}$ are subspaces of $\mathfrak{K}^{L,\pm}$ consisting of vectors with a fixed value of l. As is proved in Ref. 5, any $\mathfrak{K}^{L,\pm,l}$ is a carrier space of irreducible discrete unitary representation of $SO_0(p - 1, 2)$. For irreducibility of $U_g^{L,\pm}$ see Sec. 3.

3. IRREDUCIBILITY

Let us consider first the representation of the Lie algebra $\mathfrak{R}(p, q)$ of the group $SO_0(p, q)$. We call the representation of Lie algebra (2.14) irreducible on a common invariant domain $\mathfrak{D}^L$ if for any two vectors, say χ^L, $\psi^L \in \mathfrak{D}^L$, in the enveloping algebra such an operator $\hat{A}^L$ exists that $(\chi^L, \hat{A}\psi^L)_L \neq 0$.

A. The Representation Related to the Hyperboloid H_q^p, $p \geq q > 1$

The subalgebra $\mathfrak{R}(p, q - 1)$ (consisting of L_{ij} $i, j = 1, \cdots, p$ or $i, j = p + 1, \cdots, p + q - 1$ and B_{rs} $r = 1, \cdots, p$, $s = p + 1, \cdots, p + q - 1$ or vice versa) with any operator $B_{i,p+q}$ or $L_{r,p+q}$ $(i = 1, \cdots p$; $r = p + 1, \cdots, p + q - 1)$ generates the whole algebra $\mathfrak{R}(p, q)$.

In order to prove irreducibility of Eq. (2.4), we have to show that the carrier space $\mathfrak{D}^L$ of the Hermitian representation of $\mathfrak{R}(p, q)$ has no invariant subspaces under the action of $\mathfrak{R}(p, q)$. In Refs. 5 and 7, it is proved that each space $\mathfrak{D}^{L,l}$ is a carrier space of the discrete irreducible Hermitian representation of $\mathfrak{R}(p, q - 1)$. Therefore, it is sufficient to check that there always exists one generator in $\mathfrak{R}(p, q)$ and some

vector belonging to definite $\mathfrak{D}^{L,l}$, such that for any $l \in PS[\Delta(H_{q-1}^p)]$, the generator maps this vector into vectors belonging to all possible neighboring spaces of $\mathfrak{D}^{L,l}$, that is, to $\mathfrak{D}^{L,l+1}$ and $\mathfrak{D}^{L,l-1}$. Such a generator, of course, cannot belong to $\mathfrak{R}(p, q - 1)$, and hence, if it exists in $\mathfrak{R}(p, q)$, it can be an arbitrary operator from the set $B_{i,p+q}$ or $L_{r,p+q}$ $(i = 1, \cdots, p$; $r = p + 1, \cdots, p + q - 1)$ as any one of these generators together with $\mathfrak{R}(p, q - 1)$ create the whole algebra of $\mathfrak{R}(p, q)$.

Let us take the simplest operator from this set, $B_{p,p+q}$. In our parametrization (2.5) it has the form

$$B_{p,p+q} = \sin \varphi^{[p/2]} \cdot \cos \vartheta^{\{p/2\}} \cosh \theta \frac{\partial}{\partial \eta}$$

$$- \sin \varphi^{[p/2]} \cdot \cos \vartheta^{\{p/2\}} \sinh \theta \cdot \tanh \eta \frac{\partial}{\partial \theta}$$

$$- \frac{\sin \varphi^{[p/2]} \cdot \sin \vartheta^{\{p/2\}} \cdot \tanh \eta}{\cosh \theta} \frac{\partial}{\partial \vartheta^{\{p/2\}}}$$

$$+ \frac{\cos \varphi^{[p/2]} \cdot \tanh \eta}{\cos \vartheta^{\{p/2\}} \cosh \theta} \frac{\partial}{\partial \varphi^{[p/2]}}$$

$$\text{if} \quad p \text{ is even,} \quad (3.1)$$

$$B_{p,p+q} = \cos \vartheta^{\{p/2\}} \cosh \theta \frac{\partial}{\partial \eta}$$

$$- \sinh \theta \cdot \cos \vartheta^{\{p/2\}} \tanh \frac{\partial}{\partial \theta}$$

$$- \frac{\sin \vartheta^{\{p/2\}} \tanh \eta}{\cosh \theta} \frac{\partial}{\partial \vartheta^{\{p/2\}}}$$

$$\text{if} \quad p \text{ is odd.} \quad (3.2)$$

For $q > 2$ we denote by

$$_{(\alpha)}\Psi^{L,l,l\{p/2\}}_{m\{p/2\}}$$

the vector in $\mathfrak{D}^{L,l}$, the only nonvanishing component of which is equal to

$$_{(\alpha)}\chi^{L,l,l_2,\cdots,l\{p/2\},l_2,\cdots,l\{(q-1)/2\}}_{m_1,\cdots,m\{p/2\},\tilde{m}_1,\cdots,\tilde{m}\{(q-1)/2\}}$$

with indices

$$m_2 = \cdots = m_{[p/2]} = \tilde{m}_2 = \cdots = \tilde{m}_{[(q-1)/2]}$$
$$= l_2 = \cdots l_{\{(q-1)/2\}} = 0$$

and with

$$l_{\{p/2\}} = \max (l + q + 1; 2),$$

$$l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 0,$$

(for p even, p odd), if $l_{\{p/2\}}$ must be even according to (A3), but with $l_{\{p/2\}} = \max (l + q + 1; 3)$ and

$$\begin{cases} l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 1 \text{ for } p \text{ even} \\ l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 0 \text{ for } p \text{ odd} \end{cases}$$

if $l_{\{p/2\}}$ must be odd according to (A3).

For $q = 2$, we denote

$$_{(\alpha,\beta)}\Psi^{L,l,l_{\{p/2\}}}_{m_{[p/2]}}$$

the vector in $\mathfrak{D}^{L,l}$, the only nonvanishing component of which is equal to

$$_{(\alpha,\beta)}\chi^{L,l,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}$$

with indices $m_2 = \cdots = m_{[p/2]} = 0$ and with

$$l_{\{p/2\}} = \max\left(\begin{cases} l+3 & \text{for} \quad \beta = 1 \\ l+2 & \text{for} \quad \beta = 2 \end{cases} ; 2\right).$$

$l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 0$ (for p even or odd) if $l_{\{p/2\}}$ must be even according to (A11) or (A13), but

$$l_{\{p/2\}} = \max\left(\begin{cases} l+3 & \text{for} \quad \beta = 1 \\ l+2 & \text{for} \quad \beta = 2 \end{cases} ; 3\right)$$

and

$$\begin{cases} l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 1 & \text{for} \quad p \quad \text{even,} \\ l_{\{p/2\}-1} = \cdots = l_2 = m_1 = 0 & \text{for} \quad p \quad \text{odd} \end{cases}$$

if $l_{\{p/2\}}$ must be odd according to (A11) or (A13).

Let us show now that the generator $B_{p,p+q}$, and the vector

$$_{(\gamma)}\Psi^{L,l,l_{\{p/2\}}}_{m_{\{p/2\}}} \in \mathfrak{D}^{L,l}, \quad (\gamma) = (\alpha), (\alpha, \beta),$$

have the desired properties for our proof of irreducibility.

After a lengthy calculation, which involves the use of the Clebsch–Gordan coefficients and formulae for the hypergeometric functions from Ref. 18 (Chap. 28, Vol. I), we obtain the following result:

For $p \geq q > 2$

$$B_{p,p+q\,(\alpha)}\Psi^{L,l,l_{\{p/2\}}}_0 = C(\delta_{\alpha 1} + \delta_{\alpha 2}E) \cdot \left[\frac{^{(1+\delta_{\alpha 1})}N(l-1; l_{\{p/2\}}+1)}{_{(\alpha)}N(l; l_{\{p/2\}})}\right]^{\frac{1}{2}} \cdot {}_{(1+\delta_{\alpha 1})}\Phi^{L,l-1,l_{\{p/2\}}+1}$$

$$+ (D + q - 1) \cdot (\delta_{\alpha 1} + \delta_{\alpha 2}F) \cdot \left[\frac{^{(1+\delta_{\alpha 1})}N(l+1; l_{\{p/2\}}+1)}{_{(\alpha)}N(l; l_{\{p/2\}})}\right]^{\frac{1}{2}} \cdot {}_{(1+\delta_{\alpha 1})}\Phi^{L,l+1,l_{\{p/2\}}+1}$$

$$+ D \cdot (\delta_{\alpha 1} + \delta_{\alpha 2}E) \cdot \left[\frac{^{(1+\delta_{\alpha 1})}N(l-1; l_{\{p/2\}}-1)}{_{(\alpha)}N(l, l_{\{p/2\}})}\right]^{\frac{1}{2}} \cdot {}_{(1+\delta_{\alpha 1})}\Phi^{L,l-1,l_{\{p/2\}}-1}$$

$$+ (C + q - 1) \cdot (\delta_{\alpha 1} + \delta_{\alpha 2}F) \cdot \left[\frac{^{(1+\delta_{\alpha 1})}N(l+1; l_{\{p/2\}}-1)}{_{(\alpha)}N(l; l_{\{p/2\}})}\right]^{\frac{1}{2}} \cdot {}_{(1+\delta_{\alpha 1})}\Phi^{L,l+1,l_{\{p/2\}}-1}, \quad (3.3)$$

where $\alpha = 1, 2$, $_{(\alpha)}N(l; l_{\{p/2\}}) = N_{\{p/2\}}(l_{\{p/2\}}) \cdot {}_{(\alpha)}N(l) \cdot N(l; l_{\{p/2\}}))$, and $N_{\{p/2\}}(l_{\{p/2\}})$, $_{(\alpha)}N(l)$, $N(l; l_{\{p/2\}})$ are defined in (A8), (B3), and (A5), respectively, and δ function is the usual Kronecker delta. The coefficients are defined as

$$C = l - l_{\{p/2\}}, \qquad D = l + l_{\{p/2\}} + p - 2,$$
$$E = \tfrac{1}{4}(L - l + 1) \cdot (L + l + p + q - 3), \qquad F = \tfrac{1}{4}(L - l)(L + l + p + q - 2), \qquad (3.4)$$

and

$$_{(\alpha)}\Phi^{L,l,l_{\{p/2\}}\pm 1} = \begin{cases} A(\pm) \cdot ({}_{(\alpha)}\Psi^{L,l,l_{\{p/2\}}\pm 1}_1 - {}_{(\alpha)}\Psi^{L,l,l_{\{p/2\}}\pm 1}_{-1}) & \text{if} \quad p \quad \text{is even} \\ B(\pm) \cdot {}_{(\alpha)}\Psi^{L,l,l_{\{p/2\}}\pm 1}_0 & \text{if} \quad p \quad \text{is odd,} \end{cases} \qquad (3.5)$$

where

$$A(\pm) = \frac{[(l_{\{p/2\}} - l_{\{p/2\}-1} + 1 \pm 1)(l_{\{p/2\}} + l_{\{p/2\}-1} + p - 3 \pm 1)]^{\frac{1}{2}}}{i(2l + p + q - 3)(2 - p - 2l_{\{p/2\}})},$$

$$B(\pm) = \frac{2[(l_{\{p/2\}} - l_{\{p/2\}-1} + \tfrac{1}{2} \pm \tfrac{1}{2})(l_{\{p/2\}} + l_{\{p/2\}-1} + p - \tfrac{5}{2} \pm \tfrac{1}{2})]^{\frac{1}{2}}}{(2l + p + q - 3)(2 - p - 2l_{\{p/2\}})}. \qquad (3.6)$$

[18] Bateman Manuscript Project, *Higher Transcendental Functions* A. Erdélyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1953), Vols. I, II.

For $p > q = 2$ we obtain

$$B_{p\ p+q\ (\alpha,\beta)}\Psi_0^{L,l,l_{\{p/2\}}} = (\delta_{\alpha 1} + \delta_{\alpha 2}E)(C - \delta_{\beta 1}) \cdot \left[\frac{{}^{(1+\delta_{\alpha 1;\beta)}}N(l-1;l_{\{p/2\}}+1)}{{}_{(\alpha,\beta)}N(l;l_{\{p/2\}})}\right]^{\frac{1}{2}}_{(1+\delta_{\alpha 1;\beta)}} \cdot \Phi^{L,l-1,l_{\{p/2\}}+1}$$

$$+ (\delta_{\alpha 1} + \delta_{\alpha 2}F)(D+1+\delta_{\beta 1}) \cdot \left[\frac{{}^{(1+\delta_{\alpha 1;\beta)}}N(l+1;l_{\{p/2\}}+1)}{{}_{(\alpha,\beta)}N(l;l_{\{p/2\}})}\right]^{\frac{1}{2}}_{(1+\delta_{\alpha 1,\beta)}} \cdot \Phi^{L,l+1,l_{\{p/2\}}+1}$$

$$+ (\delta_{\alpha 1} + \delta_{\alpha 2}E)(D - \delta_{\beta 1}) \cdot \left[\frac{{}^{(1+\delta_{\alpha 1;\beta)}}N(l-1;l_{\{p/2\}}-1)}{{}_{(\alpha,\beta)}N(l;l_{\{p/2\}})}\right]_{(1+\delta_{\alpha 1,\beta)}} \cdot \Phi^{L,l-1,l_{\{p/2\}}-1}$$

$$+ (\delta_{\alpha 1} + \delta_{\alpha 2}F)(C+1+\delta_{\beta 1}) \cdot \left[\frac{{}^{(1+\delta_{\alpha 1;\beta)}}N(l+1;l_{\{p/2\}}-1)}{{}_{(\alpha,\beta)}N(l;l_{\{p/2\}})}\right]_{(1+\delta_{\alpha 1,\beta)}} \cdot \Phi^{L,l+1,l_{\{p/2\}}-1}, \quad (3.7)$$

where all symbols are defined as in the previous case, but with $q = 2$ and with

$$_{(\alpha,\beta)}\Psi_{m_{[p/2]}}^{L,l,l_{\{p/2\}}}$$

and $_{(\beta)}N(l, l_{\{p/2\}})$ instead of

$$_{(\alpha)}\Psi_{m_{\{p/2\}}}^{L,l,l_{\{p/2\}}}$$

and $N(l, l_{\{p/2\}})$, respectively. The normalization factors $_{(\beta)}N(l, l_{\{p/2\}})$, $\beta = 1, 2$ are given in (A14) and (A15), respectively.

For $p = q = 2$ the expression for

$$B_{2,4\ (\alpha,\beta)}\Psi_{\tilde{m}_1}^{L,l}$$

is obtained from (3.7) putting $|\tilde{m}_1|$ instead of $l_{\{p/2\}}$ and $p = q = 2$. [Here we use $\tilde{m}_1$ instead of m_1 in accordance with Eqs. (2.19)]. Now the vectors

$$_{(\alpha,\beta)}\Psi_{\tilde{m}_1}^{L,l}$$

belong to $\mathfrak{D}^{L,\pm,l}$ and we see that operators B cannot "connect" vectors belonging to $\mathfrak{D}^{L,+,l}$ and $\mathfrak{D}^{L,-,l}$ and thus the eigenvalue of the Laplace–Beltrami operator $\Delta(H_2^2)$ is not sufficient to specify completely the irreducible representation of $\mathfrak{R}(2,2)$.

Analyzing the coefficients in (3.3) and (3.7), we find that for our vector

$$_{(\gamma)}\Psi_{m_{[p/2]}^{L,l,l_{\{p/2\}}}}, \quad (\gamma) = (\alpha), (\alpha,\beta),$$

can vanish only in accordance with the fact that we are dealing with a representation. For instance, if $q > 2$ and $\alpha = 2$, then for $l = L + 1$, it is $E = 0$ which guarantees that we cannot go lower than the minimum value of l determined by (A13).

Thus we prove that the unitary space $\mathfrak{D}^L(\mathfrak{D}^{L,\pm})$ is a carrier space of the discrete irreducible Hermitian representation of $\mathfrak{R}(p,q)(\mathfrak{R}(2,2))$.

From our construction of the representations of algebra $\mathfrak{R}(p,q)$ (see Sec. 2B), it follows that any of them can be integrated to the representation of the group $SO_0(p,q)$.[8] A representation of the group

obtained in such a way is unitary.[16] Therefore, if the representation of $\mathfrak{R}(p,q)$ is irreducible on $\mathfrak{D}^L$, the corresponding representation of $SO_0(p,q)$ must be irreducible on $\mathcal{K}^L$. To see it, let us use the proof by contradiction. Let us suppose that our representation U_g^L of $SO_0(p,q)$ is not irreducible. Then there has to exist an operator, say $\hat{C}$, such that $U_g^L\hat{C} - \hat{C}U_g^L = 0$, and $\mathcal{K}^L$ is an "eigenspace" of $\hat{C}$ corresponding to at least two different eigenvalues. However, due to the unitarity of U_g^L and due to the fact that $\mathfrak{D}^L$ is the common invariant domain of $\hat{C}X_{ij}$, we conclude that $X_{ij}\hat{C} - \hat{C}X_{ij} = 0$ on $\mathfrak{D}^L$, which contradicts our proof of irreducibility.

B. The Representation Related to the Hyperboloid H_p^q, $p > q > 1$

The proof of irreducibility is completely analogous to that in Case 1 of Sec. 2c, where the roles of

$$l_{\{p/2\}}, \quad m_{[p/2]}, \quad \vartheta^{\{p/2\}}, \quad \varphi^{[p/2]}$$

are now played by

$$\tilde{l}_{\{q/2\}}, \quad \tilde{m}_{[q/2]}, \quad \tilde{\vartheta}^{\{q/2\}}, \quad \tilde{\varphi}^{[q/2]},$$

respectively, and therefore we omit it.

4. CONCLUSION

The decomposition of a discrete most degenerate unitary irreducible representation of an arbitrary noncompact rotation group, $SO_0(p,q)$, $(p \geq q > 1)$, into discrete unitary irreducible representations of its maximal noncompact rotation subgroup $SO_0(p-1,q)$ or $SO_0(p,q-1)$ is explicitly given in Eqs. (2.16), (2.18), and (2.21).

From the decomposition we can conclude that:

(i) The most degenerate representation of $SO_0(p,q)$ decomposes only into the most degenerate representations of $SO_0(p,q-1)$ or $SO_0(p-1,q)$. (ii) Any representation of the subgroup appears in the decomposition, at most, once. In particular, in the

decomposition of discrete series of representations of U_q^L of $SO_0(p, q)$ appear only irreducible representations $U_q^{L,l}$ of $SO_0(p, q - 1)$ or $SO_0(p - 1, q)$ which satisfy $l \geq L + 1$.

From the proof of irreducibility in Sec. 3, we see that sometimes the eigenvalue of the Laplace–Beltrami operator is not sufficient to classify discrete most degenerate irreducible unitary representations and that we have to add a new operator $\hat{t}$ from the commutant to distinguish them. Moreover, we also remark that a carrier space of the discrete irreducible unitary representations of $SO_0(p, q)$ can be constructed by using only the lowest vector (i.e., the vector with $l, l_2, \cdots, '_2, \cdots, m_1, \cdots, \tilde{m}_{[(q-1)/2]}$ having the minimum possible value) and the action of the generators of the corresponding Lie algebra $\Re(p, q)$.

ACKNOWLEDGMENTS

APPENDIX A

The eigenfunctions of the Laplace–Beltrami operator $\Delta(H_{q-1}^p)$ create a complete set of functions with respect to the measure induced by the coordinate system on the hyperboloid H_{q-1}^p [see Sec. 2E, A of Ref. 7].

(a) For $p \geq q > 2$, they are of the form of Refs. 7 and 13:

$$Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{l, l_2, \cdots, l_{\{p/2\}}, l_2, \cdots, l_{\{(q-1)/2\}}}(\theta, \omega, \tilde{\omega})$$

$$= V_{l_{\{p/2\}}, l_{\{(q-1)/2\}}}^{l}(\theta) \cdot Y_{m_1, \cdots, m_{[p/2]}}^{l_2, \cdots, l_{\{p/2\}}}(\omega)$$

$$\cdot Y_{\tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{l_2, \cdots, l_{\{(q-1)/2\}}}(\tilde{\omega}), \quad \text{(A1)}$$

$$Y_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{\lambda, l_2, \cdots, l_{\{p/2\}}, l_2, \cdots, l_{\{(q-1)/2\}}}(\theta, \omega, \tilde{\omega})$$

$$= V_{l_{\{p/2\}}, l_{\{(q-1)/2\}}}^{\lambda}(\theta) \cdot Y_{m_1, \cdots, m_{[p/2]}}^{l_2, \cdots, l_{\{p/2\}}}(\omega)$$

$$\cdot Y_{\tilde{m}_1, \cdots, \tilde{m}_{[(q-1)/2]}}^{l_2, \cdots, l_{\{(q-1)/2\}}}(\tilde{\omega}). \quad \text{(A2)}$$

The variable λ is independent of indices $l_{\{\cdot\}}$, $m_{[\cdot]}$, $l_{\{\cdot\}}$, $\tilde{m}_{[\cdot]}$, whereas l has the following dependence:

$$l - |l_{\{p/2\}}| + |l_{\{(q-1)/2\}}| = -q + 1 - 2n,$$

$$\text{and} \quad n = 0, 1, 2, \cdots. \quad \text{(A3)}$$

Here,

$$V_{l_{\{p/2\}}, l_{\{(q-1)/2\}}}^{l}(\theta) = \frac{\tanh^{|l_{\{(q-1)/2\}}|}\theta}{[N(l, l_{\{p/2\}}, l_{\{(q-1)/2\}})]^{\frac{1}{2}}} \cosh^{-(l+p+q-3)}\theta$$

$$\cdot {}_2F_1\left(\frac{|l_{\{p/2\}}| + |l_{\{(q-1)/2\}}| + l + p + q - 3}{2}, \frac{l - |l_{\{p/2\}}| + |l_{\{(q-1)/2\}}| + q - 1}{2}; |l_{\{(q-1)/2\}}| + \frac{q-1}{2}; \tanh^2\theta\right)$$

$$\text{with} \quad \theta \in [0, \infty), \quad \text{(A4)}$$

and

$$N = \frac{\Gamma(\frac{1}{2}(|l_{\{p/2\}}| - |l_{\{(q-1)/2\}}| - l - q + 3)) \cdot \Gamma^2\left(|l_{\{(q-1)/2\}}| + \frac{q-1}{2}\right) \cdot \Gamma(\frac{1}{2}(|l_{\{p/2\}}| - |l_{\{(q-1)/2\}}| + l + p))}{(2l + p + q - 3) \cdot \Gamma(\frac{1}{2}(|l_{\{p/2\}}| + |l_{\{(q-1)/2\}}| + l + p + q - 3)) \cdot \Gamma(\frac{1}{2}(|l_{\{p/2\}}| + |l_{\{(q-1)/2\}}| - l))} \cdot \quad \text{(A5)}$$

is obtained from

$$V_{l_{\{p/2\}}, l_{\{(q-1)/2\}}}^{\lambda}(\theta)$$

$$V_{l_{\{p/2\}}, l_{\{(q-1)/2\}}}^{l}(\theta)$$

by putting $l = i\lambda - \frac{1}{2}(p + q - 3)$ in the right-hand side of (A4) everywhere. The Y functions have the explicit form

$$Y_{m_1, \cdots, m_{[p/2]}}^{l_2, \cdots, l_{\{p/2\}}}(\omega)$$

$$= \begin{cases} (N_r)^{-\frac{1}{2}} \prod_{k=2}^r \sin^{2-k}\vartheta^k \cdot d_{M_k; M_{k'}}^{J_k}(2\vartheta^k) \cdot \prod_{k=1}^r \exp(im_k\varphi^k), & p = 2r, \\ (N_{r+1})^{-\frac{1}{2}} \sin^{1-r}\vartheta^{r+1} d_{M_{r+1}; 0}^{J_{r+q}}(\vartheta^{r+1}) \cdot \prod_{k=2}^r \sin^{2-k}\vartheta^k \cdot d_{M_k; M_{k'}}^{J_k}(2\vartheta^k) \cdot \prod_{k=1}^r \exp(im_k\varphi^k), & p = 2r + 1, \end{cases} \quad \text{(A6)}$$

where the d functions are defined as in Ref. 19, and the indices and normalization factors by

$$J_k = \tfrac{1}{2}(l_k + k - 2),$$
$$M_k = \tfrac{1}{2}(m_k + l_{k-1} + k - 2),$$
$$M'_k = \tfrac{1}{2}(m_k - l_{k-1} - k + 2), \qquad \text{(A7)}$$
$$k = 2, 3, \cdots, r, \quad (l_1 \equiv m_4),$$

$$J_{r+1} = l_{r+1} + r - 1, \quad M_{r+1} = l_r + r - 1,$$

$$N_r = 2\pi^r \prod_{k=2}^{r} \frac{1}{l_k + k - 1},$$

$$N_{r+1} = \frac{4\pi^r}{2(l_{r+1} + r) - 1} \cdot \prod_{k=2}^{r} \frac{1}{l_k + k - 1}. \qquad \text{(A8)}$$

The indices l_k are nonnegative integers and m_k are integers. They are restricted by the conditions

$$|m_2| + |m_1| = l_2 - 2n_2, \quad |m_3| + l_2 = l_3 - 2n_3, \cdots,$$
$$|m_r| + l_{r-1} = l_r - 2n_r,$$

$$n_k = 0, 1, \cdots, \{l_k/2\}, \quad k = 2, 3, \cdots, r,$$
$$l_r = l_{r+1} - n_{r+1}, \quad n_{r+1} = 0, 1, \cdots, l_{r+1}. \qquad \text{(A9)}$$

The ranges of angles are the following: $\theta^{r+1} \in [0, \pi)$, $\theta^k \in [0, \pi/2) \, k = 2, \cdots r$, $\varphi^p \in [0, 2\pi) \, l = 1, \cdots, r$. The Y functions with tilde indices and variables have completely analogous expressions.

(b) For $p \geq q = 2$ the eigenfunctions of $\Delta(H^p_{q-1})$ are given by Refs. 7 and 13:

$$_{(1)}Y^{l, l_2, \cdots, l_{\{p/2\}}}_{m_1, \cdots, m_{[p/2]}}(\theta, \omega) = \frac{-2 \tanh \theta}{(_{(1)}N)^{\frac{1}{2}}} \cosh^{-(l+p-1)} \theta$$
$$\cdot \, _2F_1\left(\frac{l + l_{\{p/2\}} + p}{2}, \frac{l - l_{\{p/2\}} + 2}{2}; \frac{3}{2}; \tanh^2 \theta\right)$$
$$\cdot \, Y^{l_2, \cdots, l_{\{p/2\}}}_{m_1, \cdots, m_{[p/2]}}(\omega) \qquad \text{(A10)}$$

with the restriction

$$l - |l_{\{p/2\}}| = -(2n + 2), \quad n = 0, 1, 2, \cdots, \quad \text{(A11)}$$

$$_{(2)}Y^{l, l_2, \cdots, l_{\{p/2\}}}_{m_1, \cdots, m_{[p/2]}}(\theta, \omega) = \frac{\cosh^{-(l+p-1)} \theta}{(_{(2)}N)^{\frac{1}{2}}} \theta$$
$$\times \, _2F_1\left(\frac{l + l_{\{p/2\}} + p - 1}{2}, \frac{l - l_{\{p/2\}} + 1}{2}; \frac{1}{2}; \tanh^2 \theta\right)$$
$$\cdot \, Y^{l_2, \cdots, l_{\{p/2\}}}_{m_1, \cdots, m_{[p/2]}}(\omega) \qquad \text{(A12)}$$

with the restriction

$$l - |l_{\{p/2\}}| = -(2n + 1), \quad n = 0, 1, 2, \cdots, \quad \text{(A13)}$$

and by

$$_{(\alpha)}Y^{\lambda, l_2, \cdots, l_{\{p/2\}}}_{m_1, \cdots, m_{[p/2]}}(\theta, \omega)$$

$\alpha = 1, 2$, which we obtain by putting $l = i\lambda - (p-1)/2$ on the right-hand side of (A10) and (A13), respectively. [λ is independent of indices $l_{\{p/2\}}$ so that (A11), (A13) do not hold.]

The angle θ is now from the interval $(-\infty, \infty)$, the Y functions are defined in (A6) and the normalization

19 M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1961).

factors are determined as

$$_{(1)}N(l, l_{\{p/2\}}) = \frac{2\pi}{(2l + p - 1)}$$
$$\times \frac{\Gamma(\tfrac{1}{2}(l_{\{p/2\}} - l))\Gamma(\tfrac{1}{2}(l_{\{p/2\}} + l + p - 1))}{\Gamma(\tfrac{1}{2}(l_{\{p/2\}} + l + p))\Gamma(\tfrac{1}{2}(l_{\{p/2\}} - l + 1))} \qquad \text{(A14)}$$

and

$$_{(2)}N(l, l_{\{p/2\}}) = \frac{2\pi}{(2l + p - 1)}$$
$$\times \frac{\Gamma(\tfrac{1}{2}(l_{\{p/2\}} - l + 1))\Gamma(\tfrac{1}{2}(l_{\{p/2\}} + l + p))}{\Gamma(\tfrac{1}{2}(l_{\{p/2\}} + l + p - 1))\Gamma(\tfrac{1}{2}(l_{\{p/2\}} - l))}. \qquad \text{(A15)}$$

APPENDIX B

For the Laplace–Beltrami operator $\Delta(H^p_q)$ given in (2.6), the V functions appearing in the expressions (2.8) and (2.9) are defined as

$$_{(1)}V^L_l(\eta) = \frac{-2 \tanh \eta}{[_{(1)}N(l)]^{\frac{1}{2}}} \cosh^{-(L+p+q-2)} \eta$$
$$\cdot \, _2F_1\left(\frac{L + l + p + q - 1}{2}, \frac{L - l + 2}{2}; \frac{3}{2}; \tanh^2 \eta\right) \qquad \text{(B1)}$$

with the condition

$$L - l = -(2n + 2), \quad n = 0, 1, 2, \cdots \quad \text{(B2)}$$

and

$$_{(2)}V^L_l(\eta) = \frac{1}{[_{(2)}N(l)]^{\frac{1}{2}}} \cosh^{-(L+p+q-2)} \eta$$
$$\times \, _2F_1\left(\frac{L + l + p + q - 2}{2}, \frac{L - l + 1}{2}; \frac{1}{2}; \tanh^2 \eta\right) \qquad \text{(B3)}$$

with the condition

$$L - l = -(2n + 1), \quad n = 0, 1, 2, \cdots, \quad \text{(B4)}$$

where the range of angle η is $(-\infty, \infty)$ and the normalization factors, $_{(1)}N(l)$ and $_{(2)}N(l)$, are obtained from the expressions (A14) and (A15), respectively, by changing $l_{\{p/2\}}$, l, p into l, L, $p + q - 1$, respectively.

The functions $_{(1)}V^L_l(\eta)$ and $_{(2)}V^L_l(\eta)$ can be expressed in terms of the Legendre polynomials or the Gegenbauer polynomials. $_{(1)}V^L_l(\eta)$, $_{(2)}V^L_l(\eta)$ [in terms of the Gegenbauer polynomials Ref. 18 Vol. II, p. 176] have, the form:

$$_{(\alpha)}V^L_l(\eta) = \frac{(-1)^{\frac{1}{2}(l-L+1-\alpha)}}{(_{(\alpha)}M)^{\frac{1}{2}}} \cosh^{-(L+p-1)} \eta \cdot C^{L+p/2}_{l-L-1}(\eta),$$
$$\alpha = 1, 2$$

where

$$_{(1)}M = \frac{_{(1)}N \cdot \Gamma^2(\tfrac{1}{2}(L + l + p))}{\Gamma^2(L + \tfrac{1}{2}p)\Gamma^2(\tfrac{1}{2}(l - L))},$$

$$_{(2)}M = \frac{_{(2)}N \cdot 4 \cdot \Gamma^2(\tfrac{1}{2}(L + l + p + 1))}{(L + l + p - 1)^2 \cdot \Gamma^2(L + \tfrac{1}{2}p) \cdot \Gamma^2(\tfrac{1}{2}(l - L + 1))}$$

with $_{(1)}N$, $_{(2)}N$ defined in (B3).