JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 5, NUMBER 4 APRIL 1964

Remarks Concerning Reciprocity in Quantum Mechanics*

D. E. BILHORN, L. L. FOLDY, R. M. THALER, AND W. TOBOCMAN†
Case Institute of Technology, Cleveland, Ohio
AND
V. A. MADSEN
Lewis Research Center, Cleveland, Ohio, and *Case Institute of Technology, Cleveland, Ohio*
(Received 22 May 1963)

A new derivation of the reciprocity theorem is given. The general invariance property of the Hamiltonian which leads to symmetry of the Green's function for a quantum mechanical system is exhibited. It is found that reciprocity does not necessarily imply Hermiticity of the Hamiltonian, so that the "complex optical model potential," for example, satisfies the reciprocity relations. The concept of reciprocity is then generalized to include a somewhat wider class of symmetry properties. Some properties of antiunitary transformations are discussed.

I. INTRODUCTION

THE term "reciprocity" usually denotes a symmetry condition on a Green's function.[1] In electrostatics, for example, the symmetry has a simple physical interpretation because the Green's function is just the potential due to a point charge under a specific boundary condition. The reciprocity symmetry ensures that the potential is the same if we reverse source and field point.

In quantum mechanics, "reciprocity" is usually thought of as connected with a symmetry condition on scattering amplitudes.[2] It is readily shown[3] that if the Hamiltonian is both Hermitian and time-reversal-invariant, the system is reciprocal. There are, however, systems that are neither Hermitian nor time-reversal-invariant which nevertheless display reciprocity. For example, the scattering from a complex optical model potential is reciprocal. A more general reciprocity relation which includes both this example and the usual theorem as special cases has been discussed by several workers.[4-6]

In this paper we show that reciprocity symmetry of the Green's function leads directly to reciprocity of the reaction matrix. First, a generalized reciprocity condition on the Green's function is obtained from the invariance properties of the Hamiltonian. Since the Green's function is a wavefunction due to a point source, we see that, when the source is infinitely distant from the interaction region, the wave emerging from the source is plane in the vicinity of the interaction. In the asymptotic region ($r \to \infty$), the scattering amplitude can be identified as the amplitude of the outgoing wavefunction. Thus the reciprocity symmetry relation for the Green's function is shown to lead to a reciprocity condition

* Work performed under the auspices of the U. S. Atomic Energy Commission.
† Supported in part by a grant from the Sloan Foundation.
[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), p. 883.
[2] R. J. Glauber, *Lectures in Theoretical Physics* (Interscience Publishers, Inc., New York, 1958), Vol. 1, pp. 319–321.
[3] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley & Sons, Inc., New York, 1952), pp. 336–339 and pp. 528–530.
[4] S. G. Eckstein and R. H. Dalitz (private communication).
[5] L. C. Biedenharn, Nucl. Phys. **10**, 620 (1959).
[6] F. Coester, Phys. Rev. **89**, 619 (1953).

on the reaction matrix elements. A physical interpretation is, of course, that the reciprocity symmetry assures equality of the scattering amplitude when source and detector are reversed.

In the Appendix an alternative derivation of the reciprocity theorem is given which makes use of formal scattering theory.

I. SYMMETRY OF THE GREEN'S FUNCTION (SPECIAL CASE—POTENTIAL SCATTERING)

Intuitively, we might say that a system is reciprocal if a source located at A produces the same signal at B as would be produced at A by that source located at B. Let us apply this definition straightforwardly to a nonrelativistic one-body quantum mechanical system, characterized by the Hamiltonian

$$H = p^2/2\mu + V(\mathbf{r}), \tag{1}$$

where $p^2/2\mu$ is the kinetic energy operator and $V(\mathbf{r})$, the potential energy, is a multiplicative operator. In order to make the argument as simple as possible, we have taken a spin-independent interaction. Let us further assume that $V(\mathbf{r}) = 0$ for $r > R$. (Clearly, we have made the conventional choice for the location of the origin of our coordinate system.) For convenience, we consider only steady-state conditions, so that we may use the time-independent form of the Schrödinger equation,

$$H\Psi(\mathbf{r}) - E\Psi(\mathbf{r}) = 0. \tag{2}$$

Suppose now that we have a point source at $\mathbf{r} = \mathbf{r}_A$. In that case, the Schrödinger equation becomes

$$H\Psi(\mathbf{r}_A, \mathbf{r}) - E\Psi(\mathbf{r}_A, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_A), \tag{3}$$

where the wavefunction $\Psi(\mathbf{r}_A, \mathbf{r})$ is a function of $\mathbf{r}$ and depends parametrically on the location of the source $\mathbf{r}_A$. Similarly, a source of the same strength located at $\mathbf{r}_B$ will give rise to the Schrödinger equation

$$H\Psi(\mathbf{r}_B, \mathbf{r}) - E\Psi(\mathbf{r}_B, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_B). \tag{4}$$

Equation (3) or (4) is the equation satisfied by the Green's function corresponding to Eq. (2). We shall occasionally refer to $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ as a Green's function.

If we multiply Eq. (3) by $\Psi(\mathbf{r}_B, \mathbf{r})$ and Eq. (4) by $\Psi(\mathbf{r}_A, \mathbf{r})$, take the difference of the resulting two equations, and integrate over all space, we get

$$\int [\Psi(\mathbf{r}_B, \mathbf{r})H\Psi(\mathbf{r}_A, \mathbf{r}) - \Psi(\mathbf{r}_A, \mathbf{r})H\psi(\mathbf{r}_B, \mathbf{r})]\, d\tau = [\Psi(\mathbf{r}_B, \mathbf{r}_A) - \Psi(\mathbf{r}_A, \mathbf{r}_B)]. \tag{5}$$

Thus, if the left-hand side of Eq. (5) vanishes, we have

$$\Psi(\mathbf{r}_B, \mathbf{r}_A) = \Psi(\mathbf{r}_A, \mathbf{r}_B), \tag{6}$$

which is the reciprocity relation. In other words, we may say that reciprocity implies symmetry of the Green's function with respect to interchange of the coordinates, which clearly implies that the wavefunction at $\mathbf{r}_A$ due to a point source at $\mathbf{r}_B$ is equal to the wavefunction at $\mathbf{r}_B$ due to a source of the same strength at $\mathbf{r}_A$.

If the wavefunction obeys homogeneous boundary conditions on some bounding surface S, then the left-hand side of Eq. (5) vanishes. To see this, we write

$$\begin{aligned} L &\equiv \int [\Psi(\mathbf{r}_B, \mathbf{r})H\Psi(\mathbf{r}_A, \mathbf{r}) - \Psi(\mathbf{r}_A, \mathbf{r})H\Psi(\mathbf{r}_B, \mathbf{r})]\, d\tau \\ &= \frac{-\hbar^2}{2\mu} \int [\Psi(\mathbf{r}_B, \mathbf{r})\nabla^2\psi(\mathbf{r}_A, \mathbf{r}) \\ &\quad - \psi(\mathbf{r}_A, \mathbf{r})\nabla^2\psi(\mathbf{r}_B, \mathbf{r})]\, d\tau \\ &= \frac{-\hbar^2}{2\mu} \int_{\substack{\text{bounding} \\ \text{surface}}} [\Psi(\mathbf{r}_B, \mathbf{r})\nabla\Psi(\mathbf{r}_A, \mathbf{r}) \\ &\quad - \Psi(\mathbf{r}_A, \mathbf{r})\nabla\Psi(\mathbf{r}_B, \mathbf{r})]\, dS. \end{aligned} \tag{7}$$

Since homogeneous boundary conditions mean that at the bounding surface

$$\alpha\Psi + \beta(\partial\Psi/\partial n) = 0, \tag{8}$$

where $\partial\Psi/\partial n$ is the derivative in the direction of the outward normal to the surface, it follows immediately that L in Eq. (7) vanishes, and that therefore Eq. (6) holds.

II. RECIPROCITY OF THE SCATTERING AMPLITUDE (SPECIAL CASE—POTENTIAL SCATTERING)

Equation (6) expresses the reciprocity relation in terms of the Green's function. However, from Eq. (6) we can easily obtain a corresponding symmetry property which must be obeyed by the scattering amplitude. The Green's function of interest in the present case must be of the form

$$\Psi(\mathbf{r}_A, \mathbf{r}) = X(e^{ik|\mathbf{r}-\mathbf{r}_A|}/|\mathbf{r} - \mathbf{r}_A|) + \Psi_{\text{scatter}}(\mathbf{r}_A, \mathbf{r}), \tag{9}$$

where $k = (2\mu E/\hbar^2)^{\frac{1}{2}}$ and Ψ_{scatter} is everywhere regular. The singular term represents the spherical wave emanating from the source. Direct integration of Eq. (3) over an infinitesimal volume containing the point $\mathbf{r} = \mathbf{r}_A$ shows us that

$$X = 2\mu/4\pi\hbar^2. \tag{10}$$

Asymptotically, for large values of r, $\Psi_{\text{scatter}}(\mathbf{r}_A, \mathbf{r})$ must have the form of an outgoing spherical wave

emanating from the scattering center (which we have chosen to be at the origin); that is, for $r \gg R$,

$$\Psi_{\text{scatter}} \rightarrow T(-\mathbf{r}_A, \hat{r})(e^{ikr}/r), \tag{11}$$

where $\hat{r}$ is the unit vector $\mathbf{r}/r$. Thus we have for large r

$$\Psi(\mathbf{r}_A, \mathbf{r}) \rightarrow (2\mu/4\pi\hbar^2)\frac{e^{ik|\mathbf{r}-\mathbf{r}_A|}}{|\mathbf{r}-\mathbf{r}_A|} + T(-\mathbf{r}_A, \hat{r})\frac{e^{ikr}}{r} \qquad (r \gg R). \tag{12}$$

The expressions analogous to Eqs. (9)–(12) for the case of the source at $\mathbf{r}_B$ are self-evident.

Now let us choose, for convenience, our bounding surface to be a sphere whose radius ρ is very large compared with R (the range of the force), r_A, and r_B; then we see from Eq. (12) that

$$\left(\frac{1}{\Psi}\frac{\partial\Psi}{\partial r}\right)_{r=\rho} = ik + O\left(\frac{1}{\rho}\right) \rightarrow ik, \tag{13}$$

so that L in Eq. (7) must vanish in the present case, and the reciprocity relation [Eq. (6)] must hold.

In the usual idealization of the scattering experiment, the source is taken to be an infinite distance from the scattering center. We can easily obtain the result for this special case from Eq. (12). If we choose $r_A \gg r \gg R$, Eq. (12) becomes

$$\Psi(\mathbf{r}_A\mathbf{r}) = (2\mu/4\pi\hbar^2)\frac{e^{ikr_A}}{r_A}\left\{e^{-ik\hat{r}_A\cdot\mathbf{r}} + \left[(2\mu/4\pi\hbar^2)\frac{e^{ikr_A}}{r_A}\right]^{-1} T(-\mathbf{r}_A, \hat{r})\frac{e^{ikr}}{r}\right\}. \tag{14}$$

This limiting form of the wavefunction $\Psi(\mathbf{r}_A, \mathbf{r})$ satisfies the same differential equation with the same boundary conditions as the usual wavefunction of time-independent scattering theory except for the over-all factor $2\mu/4\pi\hbar^2(e^{ikr_A}/r_A)$. Thus we have for $r_A \gg R$,

$$[2\mu/4\pi\hbar^2(e^{ikr_A}/r_A)]^{-1}T(-\mathbf{r}_A, \hat{r}) = f(-\hat{r}_A, \hat{r}), \tag{15}$$

where f is the usual scattering amplitude. We note that the validity of Eq. (14) required that $r_A \gg r \gg R$. The restriction $r_A \gg r$, applies only to the source term of Eq. (14), however, and is not required for Eq. (15). We need require that the source term be a plane wave just within the range of the force in order that Eq. (15) hold.

Consider a wavefunction due to another source located at $\mathbf{r}_B$. Asymptotically, it will have the form

$$\Psi(\mathbf{r}_B, \mathbf{r}) \rightarrow (2\mu/4\pi\hbar^2)\frac{e^{ik|\mathbf{r}-\mathbf{r}_B|}}{|\mathbf{r}-\mathbf{r}_B|} + T(-\mathbf{r}_B, \hat{r})\frac{e^{ikr}}{r}, \qquad r \gg R. \tag{16}$$

Using the symmetry relation [Eq. (6)] and the wavefunctions [Eqs. (12) and (16)], we have for $r_A, r_B \gg R$

$$(2\mu/4\pi\hbar^2)\frac{e^{ik|\mathbf{r}_B-\mathbf{r}_A|}}{|\mathbf{r}_B-\mathbf{r}_A|} + T(-\mathbf{r}_A, \hat{r}_B)\frac{e^{ikr_B}}{r_B} = (2\mu/4\pi\hbar^2)\frac{e^{ik|\mathbf{r}_A-\mathbf{r}_B|}}{|\mathbf{r}_A-\mathbf{r}_B|} + T(-\mathbf{r}_B, \hat{r}_A)\frac{e^{ikr_A}}{r_A}, \tag{17}$$

or, by virtue of Eq. (15),

$$f(\hat{k}_i, \hat{k}_f) = f(-\hat{k}_f, -\hat{k}_i), \tag{18}$$

where

$$\mathbf{k}_i = -k(\mathbf{r}_A/r_A), \qquad \mathbf{k}_f = k(\mathbf{r}_B/r_B). \tag{19}$$

The reader will note that the above derivation in no way depended on the Hamiltonian of Eq. (1) being real. On the contrary, the treatment is equally valid for a complex potential. For a Hamiltonian of the form given by Eq. (1), all we needed to assume to obtain our result was that $V(r)$ is a multiplicative operator and that the boundary conditions are such as to make the surface integral in Eq. (7) vanish.

Since a complex potential implies a non-Hermitian Hamiltonian, it is evident that Hermiticity is not necessary in order to obtain reciprocity. This is completely analogous to the well-known result in network theory. It is quite possible for an electrical network to be dissipative but still be reciprocal. Loosely speaking, for an electrical circuit the crucial question is not whether there is energy dissipated in a resistor, but whether there is a rectifier in the circuit. This is exactly analogous to the quantum mechanical case treated above. A non-Hermitian Hamiltonian will not conserve particles, but the reciprocity theorem may hold despite this.

III. RECIPROCITY OF THE GREEN'S FUNCTION

The above discussion suggests that although Hermiticity and time-reversal invariance of the Hamiltonian are sufficient, they are not necessary for reciprocity. We are thus motivated to seek a more general invariance property of the Hamiltonian which will also encompass the results obtained above. Since we seek such a formal result, namely a general invariance property, more abstract arguments than the preceding appear to be called for. Such an argument is presented below.

Let us consider the matrix eigenvalue equation

$$(H - E)\psi = 0, \tag{20}$$

where H is the Hamiltonian matrix and ψ is a column vector. We shall make no distinction between

position coordinates or the various discrete indices which may label the states. Let us now consider the corresponding Green's function equation

$$(H - E)\Psi = 1. \tag{21}$$

To make Eq. (21) more explicit, we may write it as

$$(H_{x'x''} - E\delta_{x'x''})\Psi_{x''x} = \delta_{x'x}. \tag{22}$$

The Einstein summation convention is used in Eq. (22). The summations are, of course, generalized in that we sum over discrete coordinates and integrate over continuous ones. Similarly, the generalized delta function in Eq. (22) is a Kronecker delta for discrete coordinates and a Dirac delta for continuous ones.

We take the transpose of both sides of Eq. (21) and introduce the unitary matrix U, to obtain[7]

$$U\Psi^T U^\dagger U(H^T - E)U^\dagger = 1. \tag{23}$$

Thus, if we define $\tilde{\Psi}$ to be

$$\tilde{\Psi} = U\Psi^T U^\dagger = U\Psi^\dagger{}^* U^\dagger, \tag{24}$$

and $\tilde{H}$ to be

$$\tilde{H} = UH^T U^\dagger = UH^\dagger{}^* U^\dagger, \tag{25}$$

we can rewrite Eq. (23) as

$$\tilde{\Psi}(\tilde{H} - E) = 1. \tag{26}$$

Now we let both sides of Eq. (26) operate on the Green's function Ψ, to get

$$\tilde{\Psi}(\tilde{H} - E)\Psi = \Psi. \tag{27}$$

Thus, if H has the invariance property

$$\tilde{H} = H, \tag{28}$$

then we may insert Eq. (21) into Eq. (27) to obtain the result

$$\tilde{\Psi} = \Psi. \tag{29}$$

This is the general form of the reciprocity theorem expressed in terms of the Green's function.

We see that if the unitary matrix U is taken to be the identity, then the reciprocity theorem reduces to the result that symmetry of the Hamiltonian, namely

$$H = H^T = H^\dagger{}^*, \tag{30}$$

implies symmetry of the Green's function,

$$\Psi = \Psi^T = \Psi^\dagger{}^*, \tag{31}$$

or explicitly, $H_{xx'} = H_{x'x}$ implies that

$$\Psi_{xx'} = \Psi_{x'x}. \tag{32}$$

In a representation in which x and x' are space coordinates, we recognize Eqs. (30)–(32) as the result earlier obtained in Eq. (6). If the Hamiltonian is a symmetric matrix, then the reciprocity relation holds, as is shown in Eqs. (31) and (32), which says that the wavefunction at x due to a point source of unit strength at x' is identically equal to the wavefunction at x' due to a point source of unit strength at x.

IV. SYMMETRY OF THE MATRIX ELEMENTS

Now we take up the general case when U is not unity and determine conditions on the matrix elements of H. Suppose that Ψ^α and Ψ^β are two particular state vectors. Then we may define $H_{\alpha\beta}$ and $\Psi_{\alpha\beta}$ as

$$H_{\alpha\beta} \equiv (\psi^\alpha)^\dagger H\Psi^\beta \equiv (\psi^\alpha)^*_x H_{xx'}(\psi^\beta)_{x'}, \tag{33}$$

and

$$\Psi_{\alpha\beta} \equiv (\psi^\alpha)^\dagger \Psi\psi^\beta \equiv (\psi^\alpha)^*_x \Psi_{xx'}(\psi^\beta)_{x'}. \tag{34}$$

We may rewrite Eq. (33) to read[8]

$$\begin{aligned} H_{\alpha\beta} &= (\psi^\beta)^\dagger{}^* H^\dagger{}^*(\psi^\alpha)^* = (U\psi^{\beta*})^\dagger (UH^\dagger{}^* U^\dagger)(U\psi^{\alpha*}) \\ &= (\psi^{\tilde{\beta}})^\dagger \tilde{H}\psi^{\tilde{\alpha}} = \tilde{H}_{\tilde{\beta}\tilde{\alpha}}, \end{aligned} \tag{35}$$

where U is a unitary matrix, $\tilde{H}$ is as defined in Eq. (25), and

$$\psi^{\tilde{\alpha}} = U\psi^{\alpha*}. \tag{36}$$

Thus if the Hamiltonian obeys the symmetry relation [Eq. (28)], then

$$H_{\alpha\beta} = H_{\tilde{\beta}\tilde{\alpha}}. \tag{37}$$

Similarly, we have

$$\Psi_{\alpha\beta} \equiv (\psi^\alpha)^\dagger \Psi(\psi^\beta) = (\psi^{\tilde{\beta}})^\dagger \tilde{\Psi}\psi^{\tilde{\alpha}} \equiv \tilde{\Psi}_{\tilde{\beta}\tilde{\alpha}}, \tag{38}$$

which according to Eq. (29) yields

$$\Psi_{\alpha\beta} = \Psi_{\tilde{\beta}\tilde{\alpha}}, \tag{39}$$

if $H = \tilde{H}$.

Eq. (36) will be recognized as the general definition of an antiunitary transformation. Familiar examples of antiunitary transformations are the time-reversal and charge-conjugation transformations. The antiunitary transformation which maps the state ψ^α into $\psi^{\tilde{\alpha}}$ as defined in Eq. (36), will map the Hamil-

[7] We use the following notation: A^T denotes the transpose, $A\dagger$ denotes the Hermitian conjugate, and A^* denotes the complex conjugate of A. Only two of the three symbols above are needed, of course, since $A^T = A\dagger^*$.

[8] The first equality in Eq. (35) is a consequence of the fact that ψ is a column vector, so that $(\psi^\alpha)\dagger H(\psi^\beta)$ is a scalar and is hence equal to its transpose.

tonian H into $\bar{H}$, where $\bar{H}$ is defined as

$$\bar{H} = UH^*U^\dagger. \tag{40}$$

We may therefore express the reciprocity theorem as follows:

Given an antiunitary transformation that maps a state ψ^α into $\psi^{\bar{\alpha}}$ and transforms an operator O into $\bar{O}$, then if the Hamiltonian has the invariance property $H = \bar{H}^\dagger$, the Green's function will possess the reciprocity symmetry $\Psi_{\alpha\beta} = \Psi_{\bar{\beta}\bar{\alpha}}$.

Of course, if the Hamiltonian is Hermitian, then reciprocity follows from the invariance of the Hamiltonian under the antiunitary transformation $H = \bar{H}$. This is the more familiar result.

It should be noted that the definition of an antiunitary transformation (unlike a unitary transformation) is not independent of the representation.[6] To see this, we note that for a unitary transformation, we may write

$$\psi = U\phi, \tag{41}$$

where U is unitary and ψ and ϕ are column vectors. Under a change of representation generated by the unitary matrix V, we see that

$$\psi' = U'\phi', \tag{42}$$

where

$$\psi' = V\psi, \tag{43}$$

$$\phi' = V\phi, \tag{44}$$

and

$$U' = VUV^\dagger. \tag{45}$$

On the other hand, an antiunitary transformation

$$\psi = U\phi^*, \tag{46}$$

under a change of representation generated by the unitary matrix V, becomes

$$\psi' = (VUV^{\dagger *})\phi'^* \neq U'\phi'^*, \tag{47}$$

and hence is representation-dependent. Thus in using Eqs. (36) and (40), it must be borne in mind that the choice of representation is of some significance.

However, if in one representation

$$H = \bar{H}^\dagger = UH^{\dagger *}U^\dagger, \tag{48}$$

then in the new representation

$$H' = WH'^{\dagger *}W^\dagger, \tag{49}$$

where

$$W = VUV^{\dagger *}, \tag{50}$$

and the previous results, Eqs. (33)–(40), hold with the unitary matrix U replaced by the unitary matrix W of Eq. (50).

V. RECIPROCITY OF THE REACTION MATRIX

In order to obtain the reciprocity condition on the reaction matrix we first rewrite the results of Sec. IV for the case where the unit matrix is in space–spin coordinates with the space coordinate dependence of the wavefunction shown explicitly. The Green's function equation is

$$(H - E)\Psi(\xi, \xi') = \delta(\xi - \xi')1, \tag{21'}$$

where 1 is the unit matrix in spin space. The reciprocity symmetry condition can be written

$$H^T(\xi', \xi)U^\dagger = U^\dagger H(\xi, \xi'), \tag{28'}$$

$$\Psi^T(\xi', \xi)U^\dagger = U^\dagger\Psi(\xi, \xi'). \tag{29'}$$

The unitary matrix $U^\dagger$ is restricted here to operate only on spin components of H and Ψ as is the transpose operation. The symbol ξ represents the space coordinates of the system.

The complete Hamiltonian can be broken up in a number of ways corresponding to various groupings of the particles:

$$H = H_i + T_i + V_i = H_j + T_j + V_j, \tag{51}$$

where H_i is the Hamiltonian for the internal motion of two groups of particles whose centers of mass are separated by the vector $\mathbf{r}_i$, T_i is the operator for the relative kinetic energy, and V_i is the interaction between them. We assume that H_i, H_j, etc. are Hermitian, although H itself need not be. The wave equation for the internal motion is

$$(H_i - E_t)\phi_{it}(\xi_i) = 0, \tag{52}$$

where ξ_i includes all coordinates ξ except $\mathbf{r}_i$, and t is the state of internal motion.

Multiplying the Green's function [Eq. (21')] on the right by the internal function $\phi_{it}(\xi_i)$ and integrating over ξ_i gives the result

$$(H - E)X_{it}(\xi, \mathbf{r}_i') = \delta(\mathbf{r}_i - \mathbf{r}_i')\phi_{it}(\xi_i), \tag{53}$$

where

$$X_{it}(\xi, \mathbf{r}_i') = \int d\xi_i'\Psi(\xi, \xi')\phi_{it}(\xi_i'). \tag{54}$$

The wavefunction X_{it} plays the same role here as the Green's function $\Psi(\mathbf{r}_A, \mathbf{r})$ introduced earlier. Equation (53) describes the motion of the system with a point source in channel i.

From Eq. (53) we can obtain an integral equation for X_{it} by expanding in the complete orthonormal

set, $e^{i\mathbf{k}\cdot\mathbf{r}_i}\phi'_{it}(\xi_i)$, of eigenfunctions of $H_i + T_i$. For $r'_i \gg r_i \gg R$, the integral equation has the asymptotic form

$$X_{it}(\xi, \mathbf{r}'_i) \to (2\mu_i/4\pi\hbar^2)\frac{\exp[ik_{it}r'_i]}{r'_i} \times \exp[-ik_{it}\hat{r}'_i\cdot\mathbf{r}_i]\phi_{it}(\xi_i) - \sum_{t'}\phi_{it'}(\xi_i)\frac{\exp[ik_{it'}r_i]}{r_i}F^{it,it'}(-\mathbf{r}'_i, \hat{r}_i). \quad (55)$$

Equation (55) has the form of an incident plane wave plus scattered waves in all excited internal states. Because the normalization of the plane wave is not unity, we interpret

$$f^{it,it'} \equiv F^{it,it'}/[(2\mu_i/4\pi\hbar^2)(e^{ik_{it}r_i}/r'_i)] \quad (56)$$

as the scattering amplitude for the process $it \to it'$.

In order to study a reaction from the channel represented by i to a different channel j, we require X_{it} to be expanded in terms of eigenstates of $H_j + T_j$: $e^{i\mathbf{k}\cdot\mathbf{r}_j}\phi_{js}(\xi_j)$. The asymptotic form is

$$X_{it}(\xi, \mathbf{r}'_i) = \sum_{s'}\frac{e^{ik_{js'}r_j}}{r_j}\phi_{js'}(\xi_j) \times \left\{2\mu_j/4\pi\hbar^2\left[\int d\xi''\exp(-ik_{js'}\hat{r}_j\cdot\mathbf{r}''_j)\,\delta(\mathbf{r}''_i - \mathbf{r}'_i) \times \phi^{\dagger}_{js'}(\xi''_j)\phi_{it}(\xi''_i)\right] - F^{it,js'}(-\mathbf{r}'_i, \hat{r}_j)\right\}. \quad (57)$$

The second term in the braces gives the scattered wave in the j channel. The first term, which is due to the source, vanishes for all bound internal states s', if the state t is also bound, since the product $\delta(\mathbf{r}''_i - \mathbf{r}'_i)\phi'^{\dagger}_{js'}(\xi''_j)\phi_{it}(\xi''_i)$ then decays exponentially with $\mathbf{r}''_i$. Again, because the plane-wave term of Eq. (55) does not have unit amplitude, we interpret

$$f^{it,js'} = F^{it,js'}/[(e^{ik_{it}r_i}/r'_i)(2\mu_i/4\pi\hbar^2)] \quad (58)$$

as the reaction amplitude for the process $it \to js'$. Arguments similar to those made in Sec. II show that the amplitudes $f^{it,it'}$ and $f^{it,js'}$ are each independent of r'_i and the quantities $F^{it,it'}$ and $F^{it,js'}$ are therefore dependent on r'_i through the over-all constant $e^{ik_{it}r_i}/r'_i$.

In order to obtain the reciprocity relation for the reaction matrix we multiply both sides of the reciprocity symmetry relation [Eq. (29′)] by a pair of bound-state internal functions—on the left by $\phi^T_{it}(\xi_i)$, and on the right by[9] $\phi_{js}(\xi'_j)$—and integrate over ξ_i and ξ'_j. From the definition of X in Eq. (54), we obtain the result

$$\int d\xi'_j X^T_{it}(\xi', \mathbf{r}_i)U^{\dagger}\phi_{js}(\xi'_j) = \int d\xi_i\phi^T_{it}(\xi_i)U^{\dagger}X_{js}(\xi, \mathbf{r}'_j). \quad (59)$$

Putting in the asymptotic forms for X_{it} and X_{js}, we have for large r_i, r'_j the relation

$$\int d\xi'_j\left\{-\sum_{s'}\frac{e^{ik_{js'}r'_j}}{r'_j}F^{it,js'}(-\hat{r}_i, \mathbf{r}'_j)\phi^T_{js'}(\xi'_j)\right\} \times U^{\dagger}\phi_{js}(\xi'_j) = \int d\xi_i\phi^T_{it}(\xi_i)U^{\dagger} \times \left\{-\sum_{t'}\phi_{it'}(\xi_i)\frac{e^{ik_{it'}r_i}}{r_i}F^{js,it'}(-\mathbf{r}'_j, \hat{r}_i)\right\}. \quad (60)$$

The continuum terms due to the source appearing in Eq. (57) vanish because of orthogonality of the ϕ's. On the left of Eq. (60) we have the integral

$$\int \phi^T_{js'}U^{\dagger}\phi_{js}\,d\xi_{j'} = \int \phi^T_{js'}\phi^*_{js} = \delta_{ss'}, \quad (61)$$

and on the right,

$$\int d\xi_i\phi^T_{it}U^{\dagger}\phi_{it'} = \int d\xi_i\phi^{\dagger}_{it'}\phi_{it'} = \delta_{tt'}. \quad (62)$$

Using Eq. (61) and Eq. (62), we get from Eq. (60)

$$\frac{e^{ik_{it}r_j}}{r_j}F^{it,js}(-\mathbf{r}_i, \hat{r}'_j) = \frac{e^{ik_{it}r_i}}{r'_i}F^{js,it}(-\mathbf{r}'_j, \hat{r}_i), \quad (63)$$

which along with Eq. (58) gives the result

$$\mu_i f^{it,js}(-\hat{r}_i, \hat{r}'_j) = \mu_j f^{js,it}(-\hat{r}'_j, \hat{r}_i). \quad (64)$$

It can be shown easily using the asymptotic form [Eq. (55)] that Eq. (64) holds also for scattering.

We now have a generalized reciprocity relation on the reaction matrix and have shown that it follows directly from the reciprocity symmetry of the Green's function. The conditions under which it holds are that (1) the Hamiltonian has the reciprocity symmetry [Eq. (28)], (2) the channel Hamiltonians are both Hermitian, and (3) the internal states of the two colliding systems before and after the collision are bound.

ACKNOWLEDGMENT

The authors acknowledge useful suggestions by Dr. William Ford, Dr. John Need, and Dr. Howard Volkin.

APPENDIX: DERIVATION OF THE RECIPROCITY THEOREM BY MEANS OF FORMAL SCATTERING THEORY

The reciprocity theorem can be derived readily using the results and methods of formal scattering theory.

[9] The wavefunction $\phi_{j\bar{s}}(\xi'_j)$ is the reciprocal state defined in Eq. (36).

Let H be the Hamiltonian of the system. We denote the channel Hamiltonians by H_i, H_j, $\cdots$, and we denote the channel interaction potentials by V_i, V_j, $\cdots$. Thus

$$H = H_i + V_i = H_j + V_j = \cdots . \tag{A1}$$

Next we define the eigenfunctions of the various Hamiltonians

$$(E - H)\psi = 0, \tag{A2}$$

$$(E - H_k)\phi_k = 0. \tag{A3}$$

The corresponding outgoing Green's functions satisfy the equations

$$(E - H)G = 1, \tag{A4}$$

and

$$(E - H_k)G_k = 1. \tag{A5}$$

A well-known result of formal scattering theory is the following expression for the scattering amplitude[10]:

$$T_{ij} = \langle \phi_i | \, V_i + V_i G V_i \, | \phi_i \rangle. \tag{A6}$$

T_{ij} is the amplitude for scattering from state ϕ_j to state ϕ_i. An alternative expression for the scattering amplitude is

$$T_{ij} = \langle \phi_i | \, V_j + V_i G V_j \, | \phi_j \rangle. \tag{A7}$$

The system under discussion will be said to be reciprocal if the scattering amplitude has the property

$$T_{ij} = T_{\tilde{j}\tilde{i}}, \tag{A8}$$

where the state $\phi_{\tilde{\imath}}$ is related to the state ϕ_i by the antiunitary transformation

$$\phi_i = U\phi_i^*, \tag{A9}$$

with

$$U^\dagger = U^{-1}. \tag{A10}$$

To determine under what circumstances the system is reciprocal we make use of the following result:

$$\begin{aligned} \langle \phi_i \, | A | \, \phi_j \rangle &= \langle \phi_{\tilde{\jmath}}^* \, | A^{\dagger *} | \, \phi_{\tilde{\imath}}^* \rangle \\ &= \langle \phi_{\tilde{\jmath}} \, | U A^{\dagger *} U^\dagger | \, \phi_{\tilde{\imath}} \rangle. \end{aligned} \tag{A11}$$

Comparing this expression with Eqs. (A6) and (A7), we see that the system will be reciprocal if

$$\begin{aligned} &U(V_i + V_i G V_j)^{\dagger *} U^\dagger \\ &\quad = V_{\tilde{\imath}} + V_{\tilde{\jmath}} G V_{\tilde{\imath}}, \quad \text{or} \quad V_{\tilde{\jmath}} + V_{\tilde{\jmath}} G V_{\tilde{\imath}}. \end{aligned} \tag{A12}$$

But, by definition we have

$$U(V_i + V_i G V_j)^{\dagger *} U^\dagger = \tilde{V}_i + \tilde{V}_j G \tilde{V}_i, \tag{A13}$$

where

$$\tilde{A} \equiv U A^{\dagger *} U^\dagger. \tag{A14}$$

Thus a sufficient condition for reciprocity is

$$\tilde{V}_i = V_{\tilde{\imath}}, \qquad \tilde{V}_j = V_{\tilde{\jmath}}, \qquad \tilde{G} = G. \tag{A15}$$

The above conditions on the Green's function and the channel interaction potentials are readily translated into requirements on H, H_i, and H_j. From Eqs. (A4) and (A14) we have

$$\tilde{G}(E - \tilde{H}) = 1, \tag{A16}$$

so that

$$\tilde{G}(E - \tilde{H})G = G. \tag{A17}$$

Comparing Eq. (A17) with Eq. (A4), we conclude that $\tilde{H} = H$ implies that $\tilde{G} = G$.

Now, turning to the channel potentials, we may write

$$\begin{aligned} \tilde{V}_i - V_{\tilde{\imath}} &= (\tilde{H} - \tilde{H}_i) - (H - H_{\tilde{\imath}}) \\ &= H_{\tilde{\imath}} - H_{\tilde{\imath}}^\dagger, \end{aligned} \tag{A18}$$

where, in accordance with Eqs. (A3) and (A9), we define

$$H_{\tilde{\imath}} = U H_i^* U^\dagger. \tag{A19}$$

Thus the requirement that $\tilde{V}_i = V_{\tilde{\imath}}$ is equivalent to the requirement[11] that

$$H_{\tilde{\imath}} = H_{\tilde{\imath}}^\dagger. \tag{A20}$$

We conclude that a system is reciprocal with respect to an antiunitary transformation $\phi \rightarrow U\phi^*$ if the Hamiltonian of the system is invariant under the transformation $H \rightarrow UH^{\dagger *}U^\dagger$ and the scattering connects channel states that are eigenstates of Hermitian Hamiltonians, in agreement with the results of Sec. V.

[10] G. F. Chew and M. L. Goldberger, Phys. Rev. **87**, 778 (1952).

[11] It should be noted that the equality of Eqs. (A6) and (A7) already requires the condition Eq. (A20)

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Multiplicative Symmetries in Axiomatic Quantum Field Theory

EDUARD PRUGOVEČKI
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 25 June 1963)

The group of all multiplicative symmetries defined for a finite number of interacting fields is studied in detail. Various theorems are proved connecting the existence of multiplicative symmetries with properties of the Wightman distributions of the fields.

INTRODUCTION

THIS paper represents an attempt at a systematic study of the group of multiplicative symmetries corresponding to any quantum field theory with a finite number of interacting fields. It is assumed that the theory is formulated in the language of Wightman axiomatic field theory.

A multiplicative symmetry, U_M, for a given set of N fields is a unitary operator which commutes with the momentum–energy operator P^μ and which, when applied on any component of any field operator, multiplies it by a complex number of modulus one:

$$U_M\varphi_j(x)U_M^{-1} = \lambda_j\varphi_j(x), \qquad |\lambda_j| = 1$$

(no summation over j).

Here different values of j correspond not only to different spinor fields but also to different components of the same spinor field. (For a more precise definition of a multiplicative symmetry see Sec. 1, Definition 1.)

An example of such a unitary operator is the operator of rotation around any of the spatial axes by the angle 2π. This operator multiplies all the integral-spin field operators with $+1$, and all the half-integer spin field operators with -1. Another, more significant instance, is any gauge group of the first kind—this is just a group of multiplicative symmetries depending on a continuous parameter. Now, the superselection rules for electrical charge and baryon number can be expressed "as the result" of the existence of gauge groups of the first kind, while the first-mentioned example leads to the univalence superselection rule. It is therefore clear that there must be a close relation between the possibility of formulating superselection rules in a given theory of fields and the structure of the group of multiplicative symmetries corresponding to that theory.

In Sec. 1 the results are formulated in the form of theorems followed by proofs; the theorems can be understood independently of their proofs which pretend to full mathematical rigor and therefore are sometimes a little more technical in nature. The main results are contained in Theorems 1, 2, 8, 9, and 10.

Only some remarks of a general nature concerning the theorems of Sec. 1 are contained in the first part of Sec. 2.

In the second part of Sec. 2 the consequences on the results of Sec. 1 of the behavior of a field theory under the restricted Lorentz group $\mathfrak{L}_+^\uparrow$ are taken into account in more detail.

The last part of Sec. 2 represents only an illustration of the possible use of multiplicative symmetries. Of course, the subjects for which multiplicative symmetries can be found useful are not exhausted in this way. In fact, one of the main reasons for studying the group of multiplicative symmetries in such detail was the hope of giving a more precise formulation to some heuristic considerations which occur in the literature and in which the existence of certain multiplicative symmetries is tacitly assumed. It is hoped to present a fuller account of these applications on a later occasion.

1. GENERAL THEOREMS ABOUT MULTIPLICATIVE SYMMETRIES

In this section some general theorems concerning multiplicative symmetries will be proved. The main results are formulated in Theorems 8 and 9. For a mathematically rigorous proof of these theorems, Wightman axioms for a field theory have to be accepted as a starting point. Therefore, though formulations of these axioms can be found in many places,[1,2] a statement of these axioms has been

[1] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That* (W. A. Benjamin Company, Inc., New York, 1964).

[2] A. S. Wightman, "Quelques Problèmes Mathématiques de la Mecanique Relativiste" in *Problèmes Mathématique de la Théorie Relativiste de Champs* (CNRS, Paris, 1959).

given in the appendix in a form which is best suited to the needs of this paper.

We begin by making a few comments on the sense in which the term multiplicative symmetry is used in this paper compared with its usage elsewhere. Namely, some authors[3] refer to these operators as "multiplicative operators," reserving the name "multiplicative symmetry operator" for multiplicative operators commuting with the total Hamiltonian of the system of fields. But, as will be seen very soon (Theorem 1), any multiplicative operator for the *interacting* fields, which satisfies conditions (a) and (b) of Definition 1, commutes with the momentum–energy operators and defines therefore a multiplicative symmetry operator. It has to be mentioned that in Ref. (3) a multiplicative operator is defined by its action on the fields in the interaction picture.[4]

Definition 1. A multiplicative symmetry M is a transformation in Hilbert space $\mathfrak{H}$ (see Axiom I; all axioms referred to in the body of this paper may be found in the Appendix, which can be represented by a unitary operator U_M, satisfying the following three conditions[5]:

(a) $U_M D \subset D$ (see Axiom II),

(b) $U_M \varphi_j(f) U_M^{-1} \Psi = \lambda_j \varphi_j(f)\Psi$, $j = 1, \cdots n$,

$f \in \mathcal{S}_4$, $\Psi \in D$, λ_j are complex numbers,

(c) $U_M \Psi_0 = \Psi_0$ (see Axiom IV and Theorem 1).

The set of all multiplicative symmetries which can be defined for a given theory of fields in which the total number of field components is n, will be denoted by $\mathfrak{U}(\varphi_1, \cdots \varphi_n)$, or, when no ambiguities arise, by $\mathfrak{U}$. It has to be noticed that different values of j can refer to different components of the same field with nonzero spin, and therefore $n = \sum_{r=1}^{N} (2s_r + 1)$ (see Axiom II).

The theorem to follow will show that the condition (c) is not an essential restriction imposed on $\mathfrak{U}(\varphi_1, \cdots \varphi_n)$, but only a convenient agreement which will enable us to define a one-to-one mapping between $\mathfrak{U}$ and a subset $\mathfrak{A}(\mathfrak{U})$ of a real Euclidean space $\mathbf{R}^n$ (see Theorem 2).

Theorem 1.[6] Any unitary operator U which satisfies conditions (a) and (b) of Definition 1 leaves the vacuum state invariant, i.e., $U\Psi_0 = \omega_0\Psi_0 (|\omega_0| = 1)$, and commutes with the unitary representation $U(a, 1) = \exp(iP_\mu a^\mu)$ of the four-dimensional translation group (see Axiom III), i.e., $[U, U(a, 1)]_- = 0$. Therefore $(1/\omega_0)U \in \mathfrak{U}(\varphi_1, \cdots \varphi_n)$.

Proof: By simple manipulations, using the properties (a) and (b) of Definition 1 of U, as well as Axiom III, we show for the translation group $\{a, 1\}$, namely from

$$U(a, 1)\varphi_j(f)U(a, 1)^{-1}\Psi = \varphi_j(f_{\{a,1\}})\Psi, \qquad \Psi \in D,$$

$$f_{\{a,1\}} = f(x - a),$$

that

$$[(U(a, 1)U)^{-1}UU(a, 1), \varphi_j(f)]_-\Psi = 0,$$

$$\Psi \in D, \qquad f \in \mathcal{S}_4.$$

Because of the irreducibility of $\varphi_j(f)$ (see Ref. 1, Chap. 4, Theorem 17.4), we have[7]

$$(U(a, 1)U)^{-1}UU(a, 1) = \omega(a)\mathbf{1}, \qquad |\omega(a)| = 1,$$

where $\omega(a)$ is a complex number. Therefore,

$$U(a, 1) = \omega(a)U^{-1}U(a, 1)U. \tag{1}$$

It is easy to deduce from (1) and from the group property $U(a' + a'', 1) = U(a', 1)U(a'', 1)$ that $\omega(a' + a'') = \omega(a')\omega(a'')$, while Axiom III gives us $\lim_{a' \to a} \omega(a') = \omega(a)$, i.e., the $\omega(a)$'s form a continuous one-dimensional representation of the translation group. As is well known, the general form of such a representation is $\omega(a) = \exp(-ip_\mu a^\mu)$, where p_μ ($\mu = 0, 1, 2, 3$) are some real numbers [that p_μ must be real follows from $|\omega(a)| = 1$]. Then [see Axiom IV and Eq. (1)]

$$\Psi_0 = U(a, 1)\Psi_0 = \omega(a)U^{-1}U(a, 1)U\Psi_0,$$

or

$$U(a, 1)U\Psi_0 = \exp(ip_\mu a^\mu)U\Psi_0.$$

As $U\Psi_0 \in D$ because $\Psi_0 \in D$ (see Axiom IV),

$$P_\mu U\Psi_0 = p_\mu U\Psi_0.$$

But $p_\mu = 0$, $\mu = 0, \cdots 3$ is the only eigenvalue of P_μ, and Ψ_0 is the only eigenstate with this eigenvalue (these are consequences of Axioms III and IV). Therefore there is a complex number ω_0, $|\omega_0| = 1$,

[3] G. Feinberg and S. Weinberg, Nuovo Cimento **14**, 571 (1959).

[4] It is otherwise a well-known result of relativistic quantum field theory, obtained by Haag, that interaction picture, except in trivial cases, is a mathematically meaningless entity.

[5] In all the expressions where j appears twice there is no summation over j; all the summations over j will be explicitly indicated by a summation sign.

[6] The proof of this theorem follows a line of argument due to A. S. Wightman and used very often by him in proving similar theorems about inversions (see Ref. 1, Chap. 3, Theorem 15.1).

[7] When an operator relation is valid for all $\Psi \in \mathfrak{H}$, the dependence upon Ψ will not be explicitly displayed.

such that $U\Psi_0 = \omega_0\Psi_0$. From (1) we finally obtain that, as a consequence of $\omega(a) = 1$, $[U, U(a, 1)]_- = 0$. Q.E.D.

Notice that as a consequence of Theorem 1,

$$[U_M, P^\mu]_-\Psi = 0, \quad \Psi \in D, \quad M \in \mathfrak{U}, \quad \mu = 0, \cdots 3. \tag{2}$$

Definition 2. The set of $\mathbf{k} = \{k_1, \cdots k_n\}$, $k_j = 0, \pm 1, \pm 2, \cdots$ $(j = 1, \cdots n)$ forms a group Z^n in which the operation of group summation is the usual summation,

$$\mathbf{k}' + \mathbf{k}'' = k_1' + k_1'', \cdots k_n' + k_n''.$$

By $\mathfrak{A}^n$ we will denote the factor group R^n/Z^n, where R^n is the n-dimensional space of ordered n-tuples of real numbers regarded as a group with respect to the operation of conventional addition. The general element of $\mathfrak{A}^n$ will be denoted by $\boldsymbol{\alpha} = \{\alpha_1, \cdots \alpha_n\}$, $0 \leq \alpha_j < 1$ $(j = 1, \cdots n)$ and the operation of group addition is

$$\boldsymbol{\alpha}' + \boldsymbol{\alpha}'' = \{\alpha_1' + \alpha_1''(\text{mod } 1), \cdots \alpha_n' + \alpha_n''(\text{mod } 1)\}$$

Note that R^n, Z^n, and $\mathfrak{A}^n$ are all Abelian groups for which it is customary to call the operation of group multiplication just group summation.

Theorem 2. The set $\mathfrak{U}(\varphi_1, \cdots \varphi_n)$ of all multiplicative symmetries which can be defined for a given set of fields $\varphi_1, \cdots \varphi_n$ is an Abelian group with respect to the operation of operator multiplication. An isomorphism between $\mathfrak{U}$ and a subgroup $\mathfrak{A}(\mathfrak{U})$ of $\mathfrak{A}^n$ exists.

The multiplicative symmetry $U \in \mathfrak{U}$ which belongs (uniquely) to an $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$ will be denoted from now on by $U_{\boldsymbol{\alpha}}$.

Proof: The identity operator in $\mathfrak{IC}$ is a trivial case of a multiplicative symmetry with $\lambda_1 = \cdots = \lambda_n = 1$.

By definition to each U_M, $M \in \mathfrak{U}$, belongs a set of numbers $\lambda_1, \cdots \lambda_n$. As a consequence of the unitarity of U_M, $|\lambda_1| = \cdots = |\lambda_n| = 1$, so that we can write $\lambda_j = \exp(2\pi i\alpha_j)$, $0 \leq \alpha_j < 1$ $(j = 1, \cdots n)$. Therefore we can assign uniquely to each M an element $\boldsymbol{\alpha} \in \mathfrak{A}^n$. We will denote by $\mathfrak{A}(\mathfrak{U})$ the set of all such $\boldsymbol{\alpha}$'s. It is a consequence of the irreducibility of the field operators $\varphi_1, \cdots \varphi_n$ and of the requirement (c) in Definition 1, that to each $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$ belongs one and only one $U_{\boldsymbol{\alpha}} \in \mathfrak{U}(\varphi_1, \cdots \varphi_n)$. Namely if $U_{\boldsymbol{\alpha}}\varphi_j(f)U_{\boldsymbol{\alpha}}^{-1}\Psi = \exp(2\pi i\alpha_j)\varphi_j(f)\Psi = \tilde{U}_{\boldsymbol{\alpha}}\varphi_j(f)\tilde{U}_{\boldsymbol{\alpha}}^{-1}\Psi (f \in \mathcal{S}_4, j = 1, \cdots n, \Psi \in D)$, then $[U_{\boldsymbol{\alpha}}\tilde{U}_{\boldsymbol{\alpha}}^{-1}, \varphi_j(f)]_-\Psi = 0$ and this implies $U_{\boldsymbol{\alpha}}\tilde{U}_{\boldsymbol{\alpha}}^{-1} = 1$.

It is an obvious consequence of Definition 1 and of the aforesaid, that, if $\boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathfrak{A}(\mathfrak{U})$, then $U_{\boldsymbol{\alpha}'}U_{\boldsymbol{\alpha}''}$ and $U_{\boldsymbol{\alpha}''}U_{\boldsymbol{\alpha}'}$ are also multiplicative symmetries which belong to $\boldsymbol{\alpha}' + \boldsymbol{\alpha}'' = \boldsymbol{\alpha}'' + \boldsymbol{\alpha}'$. Because of the one-to-one mapping between $\mathfrak{U}$ and $\mathfrak{A}(\mathfrak{U})$, $U_{\boldsymbol{\alpha}'+\boldsymbol{\alpha}''} = U_{\boldsymbol{\alpha}'}U_{\boldsymbol{\alpha}''} = U_{\boldsymbol{\alpha}''}U_{\boldsymbol{\alpha}'}$. Q.E.D.

One of the main attempts of this paper is to establish a close connection between the structure of $\mathfrak{U}$ [or, equivalently, of $\mathfrak{A}(\mathfrak{U})$] and the identical vanishing of a certain class of Wightman distributions.[8] To be able to formulate results in a concise way, we will have to group the set of all Wightman distributions (for given $\varphi_1, \cdots \varphi_n$) into subsets which will be denoted by $\langle\mathbf{k}\rangle$.

Definition 3. For a given set of fields $\varphi_1, \cdots \varphi_n$ and for any $\mathbf{k} \in Z^n$, denote by $\langle\mathbf{k}\rangle$ the set of all Wightman distributions containing the field $\varphi_j(f)$ $k_j' = k_j + k_j''$ times and its Hermitian adjoint $\varphi_j(f)^*k_j''$ times $(k_j', k_j'' = 0, 1, 2, \cdots)$—in case that φ_j is a "complex" field; or alternatively, containing the field φ_j $|k_j|$ times, in case that φ_j is a "real" field $(j = 1, \cdots n; \mathbf{k} = \{k_1, \cdots k_n\})$.

Write, symbolically, $\langle\mathbf{k}\rangle = 0$ if all the elements of $\langle\mathbf{k}\rangle$ are identically vanishing distributions, and $\langle\mathbf{k}\rangle \neq 0$ if there is at least one distribution belonging to $\langle\mathbf{k}\rangle$ which does not vanish identically.

It is useful to know as much as possible about the distributions belonging to a certain $\langle\mathbf{k}\rangle$. The following theorem will be somewhat helpful in this respect.

Theorem 3. If, for a given $\mathbf{k} \in Z^n$, a Wightman distribution, containing the "complex" field φ_j k_j' times and its complex conjugate $k_j'' = k_j' - k_j$ times (j runs over all complex fields), does not vanish identically, then *all* Wightman distributions belonging to $\langle\mathbf{k}\rangle$ and containing φ_j *at least* k_j' times and φ_j^* *at least* k_j'' times do not vanish identically. All the elements of $\langle\mathbf{0}\rangle$, $\mathbf{0} = \{0, \cdots 0\}$, do not vanish identically.

Proof: It is a standard result of the theory of analytic extensions of Wightman functions, that if a Wightman distribution vanishes identically then all the other Wightman distributions, obtained by any permutation of the field operators in the initial one, have to vanish identically. The idea of the proof is that from the identical vanishing of a certain Wightman distribution one concludes (see Ref. 1, Chap. 2, Theorem 9.5) that the corresponding Wightman function is identically equal to zero in

[8] Here and in the following, by the name "Wightman distribution" we will refer to the vacuum expectation value of a product of r field operators regarded as a tempered distribution on the space of functions $f(x_1, \cdots x_r) \in \mathcal{S}_{4r}$, while by a "Wightman function" we will mean the corresponding analytic function in the complex variables $z_\rho = x_\rho - iy_\rho (\rho = 1, \cdots r)$.

the extended tube. As all the extended tubes have real environments in common, the identical vanishing of all permuted Wightman functions and, therefore, of all corresponding distributions is inferred.

To see that the existence of a nonvanishing Wightman distribution containing a *certain* "complex" field φ_i k_i' times and its adjoint φ_i^* k_i'' times implies that there is another nonvanishing Wightman distribution containing φ_i $(k_i'+1)$ times and φ_i^* $(k_i''+1)$ times, one has only to notice that there is a $f \in \mathcal{S}_4$ such that $\langle \varphi_i(f)^* \varphi_i(f) \rangle_0 = ||\varphi_i(f)\Psi_0||^2 \neq 0$ [otherwise the positive-definiteness of $\mathfrak{H}$ (Axiom I) would imply $\varphi_i(f)\Psi_0 = 0$, $f \in \mathcal{S}_4$, and this in turn (see Ref. 1, Chap. 4), that $\varphi_i(f) = \varphi_i(f)^* \equiv 0$, $f \in \mathcal{S}_4$]. The desired result is then obtained by using the cluster decomposition theorem in a similar manner as it will be used in proving the next theorem (see the proof of Theorem 4). By induction one establishes the validity of the first part of Theorem 3.

To prove the assertion about $\langle \mathbf{0} \rangle$, one has only to notice that $\langle \mathbf{1} \rangle_0 = (\Psi_0, \Psi_0)$ as well as $\langle \varphi_i(x)^* \varphi_i(y) \rangle_0$ (if φ_i is complex) belong to $\langle \mathbf{0} \rangle$. The rest of the statement follows from the preceding argument. Q.E.D.

Theorem 4. The set $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ (sometimes denoted simply by $\mathcal{G}$)of all $\mathbf{k} \in Z^n$ such that $\langle \mathbf{k} \rangle \neq 0$ if $\mathbf{k} \in \mathcal{G}(\varphi_1, \cdots \varphi_n)$, is a subgroup of Z^n. For non-identically vanishing fields, $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ is non-empty, containing at least the unit element of Z^n, $\mathbf{0} = \{0, \cdots 0\}$. If a certain field φ_i is real, then the $\mathbf{k}$'s: $\mathbf{k} = \{0, \cdots 0, k_i, 0, \cdots 0\}$, $k_i = \pm 2, \pm 4, \cdots$, also belong to $\mathcal{G}$.

Proof: That $\mathbf{0} \in \mathcal{G}$ has been stated in Theorem 3. It is also easily seen that if $\mathbf{k} \in \mathcal{G}$ then $-\mathbf{k} \in \mathcal{G}$. Namely, if $\langle \mathbf{k} \rangle \neq 0$, then there is a certain non-vanishing Wightman distribution belonging to $\langle \mathbf{k} \rangle$. Its complex-conjugate value is a distribution from $\langle -\mathbf{k} \rangle$ and therefore $\langle -\mathbf{k} \rangle \neq 0$.

To show[9] that $\mathbf{k}'$, $\mathbf{k}'' \in \mathcal{G}$ implies $\mathbf{k}' + \mathbf{k}'' \in \mathcal{G}$, assume, for the sake of simplicity in writing, that

$$\mathfrak{W}_1(\mathbf{x}) = \mathfrak{W}_1(x_1, \cdots x_{k_1'}, \cdots x_{k_1'+\cdots+k_n'}) = \langle \varphi_1(x_1) \cdots \varphi_1(x_{k_1'}) \cdots \varphi_n(x_{k_1'+\cdots k_n'}) \rangle_0 \neq 0,$$

$$\mathfrak{W}_2(\mathbf{y}) = \mathfrak{W}_2(y_1, \cdots y_{k_1''}, \cdots y_{k_1''+\cdots k_n''}) = \langle \varphi_1(y_1) \cdots \varphi_1(y_{k_1''}) \cdots \varphi_n(y_{k_1''+\cdots k_n''}) \rangle_0 \neq 0,$$

where $\mathfrak{W}_1 \in \langle \mathbf{k}' \rangle$, $\mathfrak{W}_2 \in \langle \mathbf{k}'' \rangle$. A straightforward application of the cluster decomposition theorem gives for

$$\mathfrak{W}(\mathbf{x}, \mathbf{y}) = \mathfrak{W}(x_1, \cdots x_{k_1'}, \cdots x_{k_1'+\cdots+k_n'}, \times y_1, \cdots y_{k_1''}, \cdots y_{k_1''+\cdots+k_n''}) = \langle \varphi_1(x_1) \cdots \varphi_n(x_{k_1'+\cdots k_n'}) \times \varphi_1(y_1) \cdots \varphi_n(y_{k_1''+\cdots k_n''}) \rangle_0$$

that $\lim_{a\to\infty} \mathfrak{W}(\mathbf{x}, \mathbf{y} + \mathbf{a}) = \mathfrak{W}_1(\mathbf{x})\mathfrak{W}_2(\mathbf{y}) \not\equiv 0$ when each component of $\mathbf{a}$ tends to infinity in a spacelike direction. As $\mathfrak{W} \in \langle \mathbf{k}' + \mathbf{k}'' \rangle$, this means that $\langle \mathbf{k}' + \mathbf{k}'' \rangle \neq 0$.

The last statement of the theorem concerning real fields is a direct consequence of the definition of $\langle \mathbf{k} \rangle$, of the fact that $\langle \varphi_i(f)\varphi_i(g) \rangle_0 \not\equiv 0$, $f, g \in \mathcal{S}_4$, when φ_i is real, and of the cluster decomposition theorem applied as in the preceding part of the proof. Q.E.D.

Theorem 5. (a) Each $U_\alpha \in \mathfrak{U}$ defines a linear (mod 1) homomorphism of Z^n into $\mathfrak{A}^1$:

$$\mathbf{k} \to \boldsymbol{\alpha}\cdot\mathbf{k} = k_1\alpha_1 + \cdots + k_n\alpha_n \text{ (mod 1)}, \qquad \mathbf{k} \in Z^n, \quad \boldsymbol{\alpha}\cdot\mathbf{k} \in \mathfrak{A}^1, \tag{3}$$

which has the property that the subgroup $\mathcal{G}_\alpha$ of the group Z^n, consisting of all $\mathbf{k}$'s mapped into zero, $\mathcal{G}_\alpha = \{\mathbf{k}: \boldsymbol{\alpha}\cdot\mathbf{k} = 0\} \subset Z^n$, contains $\mathcal{G}(\varphi_1, \cdots \varphi_n)$, i.e., $\mathcal{G}(\varphi_1, \cdots \varphi_n) \subset \mathcal{G}_\alpha$, $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$.

(b) Conversely, each linear (mod 1) homomorphism h of Z^n into $\mathfrak{A}^1$,

$$\mathbf{k} \to \alpha = h(\mathbf{k}), \qquad \mathbf{k} \in Z^n, \qquad \alpha \in \mathfrak{A}^1, \tag{4}$$

for which the subgroup $\mathcal{G}'$ of Z^n consisting of all $\mathbf{k} \in Z^n$ mapped into zero, $\mathcal{G}' = \{\mathbf{k}: h(\mathbf{k})\} = 0$, contains $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ (i.e., $\mathcal{G} \subset \mathcal{G}'$) defines uniquely a multiplicative symmetry U_α belonging to an $\boldsymbol{\alpha} = \{\alpha_1, \cdots \alpha_n\}$, where

$$\alpha_i = h(\mathbf{k}^{(i)}), \quad \mathbf{k}^{(i)} = \{0, \cdots, 0, k_i = 1, 0, \cdots 0\}, \quad j = 1, \cdots n. \tag{5}$$

Proof : (a) It is obvious that (3) defines a homomorphism of Z^n into $\mathfrak{A}^1$. To show that $\mathcal{G}(\varphi_1, \cdots \varphi_n) \subset \mathcal{G}_\alpha$ we have to use only the unitarity of U_α. Namely, if $\mathbf{k} \in \mathcal{G}(\varphi_1, \cdots \varphi_n)$, then there is a $\mathfrak{W} \in \langle \mathbf{k} \rangle$ such that $\mathfrak{W} \not\equiv 0$. Then

$$\mathfrak{W} = (\Psi_0, \varphi_{i_1}(x_1) \cdots \Psi_0) = (\Psi_0, U_\alpha \varphi_{i_1}(x_1) U_\alpha^{-1} \cdots \Psi_0) = \lambda_1^{k_1} \cdots \lambda_n^{k_n} \mathfrak{W}$$

so that $\lambda_1^{k_1} \cdots \lambda_n^{k_n} = \exp[2\pi i(\boldsymbol{\alpha}\cdot\mathbf{k})] = 1$, i.e., $\boldsymbol{\alpha}\cdot\mathbf{k} = 0$. q.e.d.

(b) Define on D_0 (see Axiom V) an operator V_α by

$$V_\alpha \mathcal{P}[\varphi_1(f_1), \cdots \varphi_n(f_n)]\Psi_0 = \mathcal{P}[\lambda_1\varphi_1(f_1), \cdots \lambda_n\varphi_n(f_n)]\Psi_0, \quad \lambda_j = \exp(2\pi i\alpha_j), \quad f_j \in \mathcal{S}_4, \quad j = 1, \cdots n, \tag{6}$$

[9] This part of the proof of Theorem 4 uses an idea applied previously by A. S. Wightman to prove the existence of a semigroup defined in a very similar way as $\mathcal{G}$ (private communications).

where α_j are given by (5) and $\mathcal{P}(\varphi_1, \cdots \varphi_n)$ is any polynomial in $\varphi_1, \cdots \varphi_n$, each term of $\mathcal{P}$ containing the fields $\varphi_1, \cdots \varphi_n$ in a certain given order. As it is possible that there are polynomials $\mathcal{P}' \neq \mathcal{P}$ such that, at least for some combination of $f_1, \cdots f_n \in \mathcal{S}_4$, $\mathcal{P}\Psi_0 = \mathcal{P}'\Psi_0$, it is not immediately clear that V_α is an unambiguously defined operator. As this unambiguousity is automatically established[10] if the relation $(V_\alpha\Psi, V_\alpha\Phi) = (\Psi, \Phi)$ is proved for any $\Psi, \Phi \in D_0$ given by means of polynomials acting on the vacuum, we will proceed with the proof of this relation.

Write

$$\Psi = \sum_{k=0}^{r} \sum_{\pi(k)} a_{k\pi}(\varphi_1, \cdots \varphi_n)\Psi_0,$$
$$\Phi = \sum_{k'=0}^{s} \sum_{\pi(k')} b_{k'\pi'}(\varphi_1, \cdots \varphi_n)\Psi_0, \tag{7}$$

where $a_{k\pi}(\varphi_1, \cdots \varphi_n)$ are products of the field operators $\varphi_1(f_1), \cdots \varphi_1(f_{k_1}), \cdots \varphi_n(f_{k_1+\cdots k_n})$ taken in the order of the permutation $\pi(\mathbf{k})$ of $1, \cdots k_1, \cdots (k_1 + \cdots k_n)$ and multiplied by a complex number (similarly for $b_{k'\pi'}$). Then

$$(V_\alpha\Phi, V_\alpha\Psi) = \sum_{k,k'} \sum_{\pi,\pi'} \lambda_1^{k_1} \cdots \lambda_n^{k_n}\bar{\lambda}_1^{k_1'} \cdots \bar{\lambda}_n^{k_n'} \times (\Psi_0, b^*_{k'\pi'}a_{k\pi}\Psi_0). \tag{8}$$

Now, according to the condition $\mathcal{G} \subset \mathcal{G}'$, the product $\lambda_1^{k_1} \cdots \bar{\lambda}_n^{k_n'} = \exp[2\pi i \sum_{j=1}^{n} \alpha_j(k_j - k_j')] = 1$ whenever $(\Psi_0, b^*_{k'\pi'}a_{k\pi}\Psi_0) \not\equiv 0$. Therefore the conclusion that (8) is equal to (Φ, Ψ) follows. Of course, for the sake of the simplicity of notation and exposition, the adjoints of the complex fields have been left out in (7), so that (7) are not expressions of the most general type for an element of D_0; but it is obvious that the same conclusion follows under the more general circumstances.

Now we see that V_α is a bounded operator on D_0 which leaves D_0 invariant. It is a matter of routine to show that from (6) the linearity of V_α follows, and to extend V_α in a unique fashion [remember that $\bar{D}_0 = \mathfrak{H}$, (Axiom V)], to a unitary operator U_α defined on the whole Hilbert space $\mathfrak{H}$. Obviously U_α is the multiplicative symmetry whose existence we have set out to prove. Q.E.D.

Before proceeding with the proof of the next theorem, a question, left open in the course of proving the second part of Theorem 5, must be pointed out.

Namely, it is clear that the unitary operator U_α, whose existence we have proved, leaves D_1 (see Axiom V) as well as D_0 invariant, but because of our scarce knowledge of D, it has not been shown that $U_\alpha D \subset D$ (notice that $D \supset D_1 \supset D_0$). Therefore, strictly speaking, it is not yet quite clear whether U_α is a multiplicative symmetry in the sense of Definition 1. The reason for insisting to define a multiplicative symmetry by the requirement $UD \subset D$ instead of $UD_1 \subset D_1$ is that it is hoped that with a better understanding of the relation of D to D_1, the stronger result will be proved.

Theorem 6. The set $\mathfrak{A}(\mathfrak{U})$ is closed in the customary topology of the n-dimensional Euclidean space[11] $\mathbf{R}^n$; i.e., if we have a sequence $\boldsymbol{\alpha}_\nu \in \mathfrak{A}(\mathfrak{U})$ which tends toward an $\boldsymbol{\alpha} \in \mathfrak{A}^n$ ($\lim_{\nu\to\infty} \boldsymbol{\alpha}_\nu = \boldsymbol{\alpha}$), then $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$. The representation of $\mathfrak{A}(\mathfrak{U})$ given by $\boldsymbol{\alpha} \to U_\alpha$ is continuous, i.e., $||(U_{\alpha_\nu} - U_\alpha)\Phi|| \to 0$ when $\boldsymbol{\alpha}_\nu \to \boldsymbol{\alpha}$ for any $\Phi \in \mathfrak{H}$.

Proof: If $\boldsymbol{\alpha}_\nu \in \mathfrak{A}(\mathfrak{U})$ then, according to the first part of Theorem 5, $\mathcal{G}_{\alpha_\nu} \supset \mathcal{G}$. On the other hand, the second part of the same theorem tells us that an $\boldsymbol{\alpha} \in \mathfrak{A}^n$ belongs to $\mathfrak{A}(\mathfrak{U})$ if $\mathcal{G}_\alpha \supset \mathcal{G}$. If $\boldsymbol{\alpha} = \lim_{\nu\to\infty} \boldsymbol{\alpha}_\nu$ then $\boldsymbol{\alpha}_\nu \cdot \mathbf{k} = 0$ for any $\mathbf{k} \in \mathcal{G}$ (because $\mathcal{G} \subset \mathcal{G}_{\alpha_\nu}$), i.e., $\alpha_{\nu 1}k_1 + \cdots + \alpha_{\nu n}k_n = N(\boldsymbol{\alpha}_\nu, \mathbf{k})$ where, for given $\boldsymbol{\alpha}_\nu$ and $\mathbf{k}$, N is a certain integer. As, for fixed $\mathbf{k}$, $\alpha_{\nu 1}k_1 + \cdots + \alpha_{\nu n}k_n$ is a continuous function in $\boldsymbol{\alpha}$, it follows that there is an $\epsilon > 0$ such that $N(\boldsymbol{\alpha}_\nu, \mathbf{k}) = N(\mathbf{k})$ for all $||\boldsymbol{\alpha}_\nu - \boldsymbol{\alpha}|| < \epsilon$ [$||\boldsymbol{\alpha}|| = (\alpha_1^2 + \cdots + \alpha_n^2)^{\frac{1}{2}}$]. Therefore, $\alpha_1 k_1 + \cdots + \alpha_n k_n = N(\mathbf{k}) = 0 \pmod 1$, i.e., $\boldsymbol{\alpha} \cdot \mathbf{k} = 0$ for $\mathbf{k} \in \mathcal{G}$, so that our assertion $\mathcal{G} \subset \mathcal{G}_\alpha$ is proved and the existence of U_α established.

To prove that the representation $\boldsymbol{\alpha} \to U_\alpha[\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})]$ is continuous is again a matter of routine. First it is shown that $||(U_{\alpha'} - U_\alpha)\Phi|| \to 0$ when $\boldsymbol{\alpha}' \to \boldsymbol{\alpha}[\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in \mathfrak{A}(\mathfrak{U})]$ for $\Phi \in D_0$, and then by standard arguments, based on the density of D_0 in $\mathfrak{H}$, the same is inferred for any $\Phi \in \mathfrak{H}$. Q.E.D.

A few words should be said about the role of the Theorem 5 in proving Theorem 6. Namely, it is obvious that, because of the continuity properties of polynomials, the vectors $U_{\alpha_\nu}\Phi$ form a Cauchy sequence in $\mathfrak{H}$ for any $\Phi \in D_0$ when $\boldsymbol{\alpha}_\nu \to \boldsymbol{\alpha}[\boldsymbol{\alpha}_\nu \in \mathfrak{A}(\mathfrak{U}), \boldsymbol{\alpha} \in \mathfrak{A}^n]$. Using again standard arguments (completeness of $\mathfrak{H}$ and density of D_0 in $\mathfrak{H}$), one can easily establish the existence of a unitary operator U, such that $||(U_{\alpha_\nu} - U)\Phi|| \to 0$ for $\boldsymbol{\alpha}_\nu \to \boldsymbol{\alpha}$, $\Phi \in \mathfrak{H}$. But, except by going through a procedure identical to the one in proving the second part of the Theorem 5, it is not possible to show that U defines a multiplicative symmetry in the sense of the Definition 1,

[10] Take $\Phi = \Psi = (\mathcal{P} - \mathcal{P}')\Psi_0$.

[11] More precisely, in the topology of the torus $\mathbf{R}^n/Z^n$. Theorem 6 then states that $\mathfrak{A}(\mathfrak{U})$ is a topological subgroup of the topological group $\mathfrak{A}^n$.

simply because there is nothing to guarantee that such a U leaves either D or D_1 invariant! And it is obvious that the property $UD \subset D$ (or $UD_1 \subset D_1$) is essential in a precise definition of the concept of multiplicative symmetry, because otherwise the expression $U\varphi_i(f)U^{-1}\Psi$ would not have meaning for all $\Psi \in D$ (or $\Psi \in D_1$); and it is important for the relation $U\varphi_i(f)U^{-1}\Psi = \lambda_i\varphi_i(f)\Psi$ to hold at least for all Ψ in an everywhere dense subset of $\mathfrak{H}$, and not only for the Ψ's in a proper subspace of $\mathfrak{H}$.

Theorem 7. Denote by $\mathfrak{A}_i(\mathfrak{U})$ the subset of $\mathfrak{A}^1$ such that $\alpha \in \mathfrak{A}^1$ if and only if there is at least one $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$ having α, as its j component, $(\alpha_i = \alpha)$. Then $\mathfrak{A}_i(\mathfrak{U})$ is either a finite cyclic subgroup of $\mathfrak{A}^1$ or it is $\mathfrak{A}^1$ itself.

Proof: From the fact that $\mathfrak{A}(\mathfrak{U})$ is a group, it follows in an evident way that $\mathfrak{A}_i(\mathfrak{U})$ has to be a group too. We will show that because $\mathfrak{A}(\mathfrak{U})$ is closed, $\mathfrak{A}_i(\mathfrak{U})$ has to be also closed in the topology of $\mathfrak{A}^1$.

Take an $\alpha \in \mathfrak{A}^1$ which is an accumulation point of an infinite subset of $\mathfrak{A}_i(\mathfrak{U})$, and pick up from this set an infinite sequence of $\alpha^{(\nu)} \to \alpha$ for $\nu \to \infty$. To each $\alpha^{(\nu)} \in \mathfrak{A}_i(\mathfrak{U})$ corresponds, according to the definition of $\mathfrak{A}_i(\mathfrak{U})$, at least one $\boldsymbol{\alpha}^{(\nu)} \in \mathfrak{A}(\mathfrak{U})$. It is obvious that with a careful definition of the neighborhoods of the zero element, all $\mathfrak{A}^n$ $(n = 1, 2, 3, \cdots)$ are closed sets. Therefore, according to the Bolzano–Weierstrass lemma, the $\boldsymbol{\alpha}^{(\nu)} \in \mathfrak{A}(\mathfrak{U})$ have at least one accumulation point $\boldsymbol{\alpha} \in \mathfrak{A}^n$ which belongs also to $\mathfrak{A}(\mathfrak{U})$ because $\mathfrak{A}(\mathfrak{U})$ is closed (Theorem 6). We can choose a subsequence $\boldsymbol{\alpha}^{(\mu)}$ of $\boldsymbol{\alpha}^{(\nu)}$ such that $\boldsymbol{\alpha}^{(\mu)} \to \boldsymbol{\alpha}$. The j component of $\boldsymbol{\alpha}$ is therefore equal to the given $\alpha \in \mathfrak{A}^1$, and as $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$ we conclude $\alpha \in \mathfrak{A}_i(\mathfrak{U})$.

The assertion of the Theorem 7 can be now proved[12] using the group properties and closure properties of $\mathfrak{A}_i(\mathfrak{U})$ (see Ref. 12, p. 139). Q.E.D.

We would like to remark, in addition, that the fact that $\mathfrak{A}_i(\mathfrak{U})$ is closed is essential in proving Theorem 7. (This is otherwise not stressed in Ref. 12.) Namely, there are obviously subgroups of $\mathfrak{A}^1$, like the set of all rational numbers in the interval [0, 1), which are neither of finite order nor identical to $\mathfrak{A}^1$ itself.

In proceeding with preparations for the proof of the next theorem it would seem advisable to apply the general theorems of the Lie group theory. But the elementary theory of Lie groups deals mostly with local properties, while the particularly simple structure of the Abelian group $\mathfrak{A}(\mathfrak{U})$ enables us to get global properties by elementary means. As the method used in proving Theorem 7 yields results formulated in a manner which makes the proof of Theorem 8 particularly simple, this is a further justification of proving Theorem 7 with a direct procedure, though otherwise this theorem could be probably demonstrated by applying the general Lie group theorems to the case of $\mathfrak{A}(\mathfrak{U})$.

For the convenience of the reader we will summarize here the definitions of a few useful concepts appearing in the theory of Abelian groups.[12] If A is an Abelian group we can define unambiguously the symbol νa ($\nu = 0, \pm 1, \pm 2, \cdots$; $a \in A$) in the following way: it is the unit (zero) element when $\nu = 0$ (as in case of Abelian groups, the group multiplication can be called group summation, it is justified to call the unit element just zero and denote it by 0); $\nu a = a + \cdots + a$ (ν times) for $\nu = 1, 2, 3, \cdots$; $\nu a = (-\nu)a^{-1}$, $\nu = -1, -2, -3, \cdots$, or writing $a^{-1} = -a$, $\nu a = (-\nu)(-a)$. A finite system of elements $a_1, \cdots a_p \in A$ is called *linearly independent* if $\nu_1 a_1 + \cdots + \nu_p a_p = 0$ for some integer values for $\nu_1, \cdots \nu_p$ implies $\nu_1 = \cdots = \nu_p = 0$. The maximal number of linearly independent elements of A is called the *rank* of A. A system of *generators* of A is a finite or infinite system of elements $a_1, a_2, \cdots \in A$ such that any $a \in A$ can be written as $a = \nu_1 a_1 + \cdots + \nu_p a_p (p < +\infty)$.

We will now prove two lemmas which will be used in proving Theorems 7 and 8.

Lemma 1. If $\mathbf{k}_1, \cdots \mathbf{k}_p \in Z^n$ $(p \leq n)$ are the generators of the subgroup $\mathfrak{G}_0$ of Z^n, then there exists another system of generators of $\mathfrak{G}_0$ containing p elements $\mathbf{k}'_1, \cdots \mathbf{k}'_p \in \mathfrak{G}_0$ such that in $\mathbf{k}'_r = \{k'_{r1}, \cdots k'_{rn}\}$ $k'_{rs} = 0$ for $s > n - p + r$; i.e., the matrix defined by the $\mathbf{k}'_1, \cdots \mathbf{k}'_p$ has the form

$$\begin{pmatrix} k'_{11}\, k'_{12} \cdots k'_{1n-p+1}\, 0 & \cdots 0 \\ k'_{21}\, k'_{22} \cdots k'_{2n-p+1}\, k'_{2n-p+2} \cdots 0 \\ \cdots \\ k'_{p1}\, k'_{p2} \cdots k'_{pn-p+1}\, k'_{pn-p+2} \cdots k'_{pn} \end{pmatrix}, \qquad (9)$$

where $k_{rn-p+r} \geq 0$ $(r = 1, \cdots p)$.

Proof: The proof is a special case of the proof used to reduce matrices to the canonical form. It is based on the easily demonstrable statement that the so-called elementary operations on a matrix can be performed by multiplying the matrix (which has p rows and n columns) on the left by

[12] L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939).

$p \times p$ and on the right by $n \times n$ matrices with integer elements and whose determinants are ± 1. We will call a square matrix with integer elements and determinant equaling ± 1, an integer-valued unimodular matrix. It is obvious that by multiplying two integer-valued unimodular matrices we obtain again an integer-valued unimodular matrix (they form in fact a group).

Now we will first point out that if we go from the system of generators $\mathbf{k}_1, \cdots \mathbf{k}_p \in \mathcal{G}_0$ of $\mathcal{G}_0$ to a system $\mathbf{k}'_1, \cdots \mathbf{k}'_p \in Z^n$, where $\mathbf{k}'_r = \sum_{s=1}^{p} a_{rs}\mathbf{k}_s$, and if $||a_{rs}||$ is an integer-valued unimodular matrix, then $\mathbf{k}'_1, \cdots \mathbf{k}'_p$ is also a system of generators of $\mathcal{G}_0$. This can be easily seen by writing *any* $\mathbf{k} \in \mathcal{G}_0$ in the form $\mathbf{k} = \sum_{r=1}^{p} \nu_r \mathbf{k}_r$ and noticing that, because $\det ||a_{rs}|| = \pm 1$, the equation $\sum_{s=1}^{p} a_{sr}\nu'_s = \nu_r$ (obtained from $\mathbf{k} = \sum_{r=1}^{p} \nu_r \mathbf{k}_r = \sum_{s=1}^{p} \nu'_s \mathbf{k}'_s$) is always solvable, with integer values for ν'_s $(s = 1, \cdots p)$, for *any* integer values for ν_r.

It is easy to see that we can perform on the matrix

$$\begin{Vmatrix} k_{11}k_{12} \cdots k_{1n} \\ k_{21}k_{22} \cdots k_{2n} \\ \cdots \\ k_{p1}k_{p2} \cdots k_{pn} \end{Vmatrix} \tag{10}$$

any of the elementary operations (e.g., multiplying a row by -1, interchanging two rows, and adding one row to another) by multiplying (10) from the left with integer-valued unimodular matrices; these matrices have all their elements equal to $0, \pm 1$. We assert that by performing these elementary operations a finite number of times on the rows of (10) we arrive at a matrix which has all the elements in the last column, except possibly the one in the last row, equal to zero—this being the first step in reducing (10) to (9). Because if all k_{rn} are not equal to zero then there is a smallest positive one, which we denote by x; we write $k_{rn} = b_r x + c_r$, where b_{ar} and c_{ar} are integers and $0 \leq c_r < x$ $(r = 1, \cdots p)$. If all $c_r = 0$, then by adding (subtracting for $b_r > 0$) the row containing $x\ b_r$ times from the rth row, and interchanging it at the end with the last row, we obtain the desired result. Otherwise there is a smallest positive c_r, lets say c_1. By subtracting the row with $x\ b_1$ times from the first row (adding it if $b_1 < 0$), we have c_1 as the smallest positive element of the last column. If c_1 is not already the common divisor of all the elements of the last column, then by repeating the procedure a finite number of times, we will arrive at that case. That means we are dealing with the previous case when all $c_r = 0$.

Having obtained a matrix which has only zeros in the last column (except on the last place), we can repeat the procedure for the $(n-1)$, $(n-2)$, $\cdots$ columns consecutively, *without affecting* each time the columns which have been already reduced to the desired form. A finite number of steps will bring us to the form (9). As this finite number of steps corresponds to multiplying the matrix (10) from the left with a product of a finite number of integer-valued unimodular matrices (this product being again such a matrix), the lemma is proved. Q.E.D.

We will need, at least partially, the following lemma quoted from Ref. 12 (p. 22, Lemma E):

Lemma 2. Let A be an Abelian group having a system of linearly independent generators, and let B be a subgroup. Then we can select in A a system $a_1, \cdots a_p$ of linearly independent generators such that the elements $\nu_1 a, \cdots \nu_q a_q$ $(q \leq p)$ form a system of generators of the group B, where $\nu_r = 1, 2, 3, \cdots$ $(r = 1, \cdots q)$ and ν_{r+1} is divisible by ν_r $(r = 1, \cdots q - 1)$. (For proof see Ref. 12, p. 23).

Lemma 3. A system of elements $\mathbf{k}_1, \mathbf{k}_2, \cdots \mathbf{k}_p \in Z^n$ is linearly independent in Z^n if and only if $\mathbf{k}_1, \cdots \mathbf{k}_p$, taken as vectors in the real Euclidean space $\mathbf{R}^n$, are linearly independent.

Proof: If $\mathbf{k}_1, \cdots \mathbf{k}_p \in Z^n$ are linearly independent vectors in $\mathbf{R}^n$, then, according to the definition, the only system of *real* numbers $\xi_1, \cdots \xi_p$ which satisfies $\xi_1\mathbf{k}_1 + \cdots + \xi_p\mathbf{k}_p = 0$ is $\xi_1 = \cdots = \xi_p = 0$.

Now let us prove the converse part of the theorem, i.e., show that when $\nu_1\mathbf{k}_1 + \cdots + \nu_p\mathbf{k}_p = 0$ for *integer* values of $\nu_1, \cdots \nu_p$ implies $\nu_1 = \cdots = \nu_p = 0$, then $\xi_1\mathbf{k}_1 + \cdots + \xi_p\mathbf{k}_p = 0$ cannot be satisfied for *any real* values of $\xi_1, \cdots \xi_p$ except for $\xi_1 = \cdots = \xi_p = 0$.

First notice that when we go from a linearly independent (in Z^n) system $\mathbf{k}_1, \cdots \mathbf{k}_p \in Z^n$ to another system of vectors $\mathbf{k}'_1, \cdots \mathbf{k}'_p \in Z^n$, $k'_r = \sum_{s=1}^{p} a_{rs}k_s$, via an integer-valued unimodular matrix $||a_{rs}||$, then $\mathbf{k}'_1, \cdots \mathbf{k}'_p$ are linearly independent in Z^n. Namely, $\sum_{r=1}^{p} \nu_r\mathbf{k}'_r = 0$ implies $\sum_{r=1}^{p} a_{rs}\nu_r = 0$, and as $\det ||a_{rs}|| = \pm 1$, this can be true only for $\nu_1 = \cdots = \nu_p = 0$.

Consider first the case $p = n$. According to Lemma 1, by an integer-valued unimodular matrix transformation we can obtain another system of n linearly independent elements of Z^n, $\mathbf{k}'_1, \cdots \mathbf{k}'_n \in Z^n$, such that the matrix they define has the form

$$\begin{vmatrix} k'_{11} & 0 & 0 & \cdots & 0 \\ k'_{21} & k'_{22} & 0 & \cdots & 0 \\ & & \cdots & & \\ k'_{n1} & k'_{n2} & k'_{n3} & \cdots & k'_{nn} \end{vmatrix} . \tag{11}$$

As $\det ||k_{rs}|| = \pm\det ||k'_{rs}|| = \pm k'_{11}k'_{22} \cdots k'_{nn}$, the $\mathbf{k}_1, \cdots, \mathbf{k}_p$ can be linearly dependent in $\mathbf{R}^n$ only if at least one of the diagonal elements of (11) is zero. Say $k_{ss} = 0$; considering

$$\begin{vmatrix} k'_{11} & 0 & \cdots 0 & \cdots 0 \cdots 0 \\ k'_{21} & k'_{22} & \cdots 0 & \cdots 0 \cdots 0 \\ k'_{s-11} & k'_{s-12} & \cdots k'_{s-1s-1} & 0 \cdots 0 \\ k'_{s1} & k'_{s2} & \cdots k'_{ss-1} & 0 \cdots 0 \end{vmatrix} , \tag{12}$$

we can again, by means of elementary operations applied on the rows of (11) [they obviously do not affect the rows of (11) which don't appear in (12)], obtain a system $\mathbf{k}'_1, \cdots \mathbf{k}'_n$ in which $\mathbf{k}'_1 = 0$. But, as we went over from $\mathbf{k}_1, \cdots \mathbf{k}_n$ to $\mathbf{k}'_1, \cdots \mathbf{k}'_{nn}$ by means of an integer-valued unimodular matrix, the $\mathbf{k}'_1, \cdots \mathbf{k}'_{nn}$ should be linearly independent in Z^n. This is a contradiction, because $\mathbf{k}'_1 = \mathbf{0}$.

For the case $p < n$ the theorem can be proved by a completely similar method, only with a lengthier argument.

It is easy to see that there cannot be p linearly independent elements in Z^n when $p > n$. Namely, by adding to each $\mathbf{k}_r = \{k_{r1}, \cdots k_{rn}\}$ a set of $(p - n)$ zeros, one would obtain a set

$$\mathbf{k}'_r = \{k_{r1}, \cdots k_{rn}, 0, \cdots 0\} \in Z^p (r = 1, \cdots p)$$

of linearly independent elements in Z^p. According to what we proved at the beginning, this would mean that the $\mathbf{k}'_1, \cdots \mathbf{k}'_p$ are linearly independent vectors of $\mathbf{R}^p$, i.e., that the rank of the matrix they determine is p—which is obviously false. Q.E.D.

A word of caution in connection with Lemma 3! Unlike the case of the vector space $\mathbf{R}^n$, not every system of linearly independent elements of Z^n is a generator of Z^n. A simple example, for the case $n = 1$, is the number 2 which generates the subgroup of Z^1 consisting of even integers.

Theorem 8. The group $\mathfrak{A}(\mathfrak{U})$ (for definition see Theorem 2) is the direct product of two of its subgroups, $\mathfrak{C}(\mathfrak{U})$ and $\mathfrak{D}(\mathfrak{U})$.

If the system of generators of $\mathfrak{G}(\varphi_1, \cdots \varphi_n)$ contains $n - m$ $(0 \leq m \leq n)$ linearly independent elements, $\mathfrak{C}(\mathfrak{U})$ is a m-parametric Lie group. Its general element has the form $\boldsymbol{\alpha} = \mathbf{c}_1 t_1 + \cdots + \mathbf{c}_m t_m \pmod 1$, where the t_μ's are some real-valued parameters, $0 \leq t_\mu < 1$, $\mu = 1, \cdots m$, and each $\mathbf{c}_\mu \in \mathbf{R}^n$ $(\mu = 1, \cdots m)$ is a vector different from zero. This group corresponds to a direct product of m gauge groups of the first kind: $\mathfrak{U}^c = \mathfrak{U}_1^c \otimes \cdots \otimes \mathfrak{U}_m^c$.

$\mathfrak{D}(\mathfrak{U})$ is a subgroup of $\mathfrak{A}^n$ of finite order. It defines for $(n - m)$ fields a discrete group $\mathfrak{U}^d$ of multiplicative symmetries.

Proof: According to Lemma 2, $\mathfrak{G}(\varphi_1, \cdots \varphi_n)$ has a system of generators containing $n - m$ linearly independent elements $\mathbf{k}_1, \cdots \mathbf{k}_{n-m}$ $(0 \leq m \leq n)$. Lemma 1 tells us that we can always choose these $\mathbf{k}_1, \cdots \mathbf{k}_{n-m}$ so that the matrix they determine has the form

$$\begin{vmatrix} k_{11} & \cdots k_{1\,m+1} & 0 & \cdots 0 \\ k_{21} & \cdots k_{2\,m+1} & k_{2\,m+2} & \cdots 0 \\ & & \cdots & \\ k_{n-m\,1} & \cdots k_{n-m\,m+1} & k_{n-m\,m+2} & \cdots k_{n-m\,n} \end{vmatrix} , \tag{13}$$

where $k_{rm+r} \geq 0$, $r = 1, \cdots m - n$. On the other hand, Lemma 3 says that $\mathbf{k}_1, \cdots \mathbf{k}_{n-m}$ have to be linearly independent vectors in $\mathbf{R}^{n-m}$, i.e., that the rank of (13) must be $n - m$. By a renumbering of the fields $\varphi_1, \cdots \varphi_n$ we could have always had just

$$\begin{vmatrix} k_{1m+1} & 0 & \cdots 0 \\ k_{2m+1} & k_{2m+2} & \cdots 0 \\ & \cdots & \\ k_{n-mm+1} & k_{n-mm+2} & \cdots k_{n-mn} \end{vmatrix}$$

$$= k_{1m+1}\, k_{2m+2} \cdots k_{n-m}\, k_{n-m} \tag{14}$$

as one of the determinants of (13) which do not vanish; i.e., $k_{rm+r} > 0$, $r = 1, \cdots m - n$.

Now, a sufficient and necessary condition for an $\boldsymbol{\alpha} \in \mathfrak{A}^n$ to belong to $\mathfrak{A}(\mathfrak{U})$ is, according to Theorem 5 and the fact that the set $\mathbf{k}_1, \cdots \mathbf{k}_{n-m}$ is a system of generators of $\mathfrak{G}(\varphi_1, \cdots \varphi_n)$, that $\boldsymbol{\alpha}\cdot\mathbf{k}_1 = \cdots = \boldsymbol{\alpha}\cdot\mathbf{k}_{n-m} = 0$. Written in detail,

$$\begin{aligned} k_{11}\alpha_1 + \cdots + k_{1m+1}\alpha_{m+1} &= q_1(\boldsymbol{\alpha}), \\ k_{21}\alpha_1 + \cdots + k_{2m+1}\alpha_{m+1} + k_{2m+2}\alpha_{m+2} &= q_2(\boldsymbol{\alpha}), \\ &\cdots \\ k_{n-m1}\alpha_1 + \cdots + k_{n-mm+1}\alpha_{m+1} + k_{n-mm+2}\alpha_{m+2} & \\ + \cdots + k_{n-mn}\alpha_n &= q_{n-m}(\boldsymbol{\alpha}), \end{aligned} \tag{15}$$

where $q_1(\boldsymbol{\alpha}), \cdots q_{n-m}(\boldsymbol{\alpha})$ are some integers depending on $\boldsymbol{\alpha}$. We can solve (15) immediately, and we obtain, as an expression for the general element of $\mathfrak{A}(\mathfrak{U})$,

$$\alpha_{m+1} = \frac{q_1}{k_{1m+1}} - \sum_{s=1}^{m} \frac{k_{1s}}{k_{1m+1}} \alpha_s \pmod 1,$$

$$q_1 = 0, 1, \cdots k_{1m+1} - 1,$$

$$\alpha_{m+2} = \frac{q_2}{k_{2m+2}} - \sum_{s=1}^{m+1} \frac{k_{2s}}{k_{2m+2}} \alpha_s (\text{mod } 1),$$
$$q_2 = 0, 1, \cdots k_{2m+2} - 1,$$
$$\cdots$$
$$\alpha_n = \frac{q_{n-m}}{k_{n-mn}} - \sum_{s=1}^{n-1} \frac{k_{n-ms}}{k_{n-mn}} \alpha_s (\text{mod } 1),$$
$$q_{n-m} = 0, 1, \cdots k_{n-m} - 1. \quad (16)$$

Here $\alpha_1, \cdots \alpha_m$ can be arbitrarily varied over the domain $\mathfrak{a}^m$. By taking $\alpha_1 = \cdots = \alpha_s = 0$, we obtain $\mathfrak{D}(\mathfrak{U})$ which is obviously a subgroup of $\mathfrak{a}^n$. Q.E.D.

In Theorem 5 we have established that $\mathfrak{G} \subset \mathfrak{G}_\alpha$, $\alpha \in \mathfrak{A}(\mathfrak{U})$. As the intersection of subgroups of a certain group is again a subgroup of the same group, we have that $\bigcap_{\alpha \in \mathfrak{A}(\mathfrak{U})} \mathfrak{G}_\alpha$ is a subgroup of Z^n and $\mathfrak{G}(\varphi_1, \cdots \varphi_n) \subset \bigcap_{\alpha \in \mathfrak{A}(\mathfrak{U})} \mathfrak{G}_\alpha$.

Theorem 9. For a set of fields $\varphi_1, \cdots \varphi_n$,

$$\mathfrak{G}(\varphi_1, \cdots \varphi_n) \subset Z^n$$

(defined in Theorem 4) is identical to the intersection of *all* the subgroups $\mathfrak{G}_\alpha$ belonging to all multiplicative symmetries U_α (for a definition fo $\mathfrak{G}_\alpha$ see Theorem 5). Concisely written, $\mathfrak{G}(\varphi_1, \cdots \varphi_n) \equiv \bigcap_{\alpha \in \mathfrak{A}(\mathfrak{U})} \mathfrak{G}_\alpha$.

Proof: If $\mathbf{k} \in \mathfrak{G}_\alpha$ then, according to the definition of $\mathfrak{G}_\alpha$, $\boldsymbol{\alpha} \cdot \mathbf{k} = 0$, $\alpha \in \mathfrak{A}(\mathfrak{U})$.

We first consider the case $m = 0$. Then $\mathfrak{A}(\mathfrak{U})$ does not contain any Lie groups, and $\mathfrak{G}$ has a generator of n linearly independent elements $\mathbf{k}_1, \cdots \mathbf{k}_n$. By choosing these $\mathbf{k}_1, \cdots \mathbf{k}_n$ in a convenient way, the $\alpha \in \mathfrak{A}(\mathfrak{U})$ fulfill Eq. (16) for this case:

$$\alpha_1 = \frac{q_1}{k_{11}}, \qquad q_1 = 0, 1, \cdots k_{11} - 1,$$
$$\alpha_2 = \frac{q_2}{k_{22}} - \frac{k_{21}}{k_{22}} \alpha_1 (\text{mod } 1),$$
$$q_2 = 0, 1, \cdots k_{22} - 1, \quad (17)$$
$$\cdots$$
$$\alpha_n = \frac{q_n}{k_{nn}} - \frac{k_{n1}}{k_{nn}} \alpha_1 - \cdots - \frac{k_{nn-1}}{k_{nn}} \alpha_{n-1} (\text{mod } 1),$$
$$q_n = 0, 1, \cdots k_{nn} - 1.$$

Assume $\mathbf{k}$ is of the form $\mathbf{k}^{(1)} = \{k_1^{(1)}, 0, \cdots 0\}$. Then taking the $\boldsymbol{\alpha}^{(1)}$ which is obtained from (17) by inserting $q_1 = 1, q_2 = \cdots = q_n = 0$, we conclude from $\boldsymbol{\alpha}^{(1)} \cdot \mathbf{k} = 0$ that $k_1^{(1)}/k_{11} = \nu_1$ (ν_1 is an integer), i.e., $\mathbf{k}^{(1)} = \nu_1 \cdot \mathbf{k}_1$. (Remember that $k_{12} = \cdots = k_{1n} = 0$).

We will prove now the theorem by induction. Assume that it has been shown for all $\mathbf{k}^{(i-1)} \in \bigcap_{\alpha \in \mathfrak{A}(\mathfrak{U})} \mathfrak{G}_\alpha$ of the form $\mathbf{k}^{(i-1)} = \{k_1^{(i-1)}, \cdots k_{i-1}^{(i-1)}, 0, \cdots 0\}$ that they belong to $\mathfrak{G}$, i.e., that they can be expressed in terms of $\mathbf{k}_1, \cdots \mathbf{k}_{i-1}$ as a linear combination with integral coefficients. Take any $\mathbf{k}^{(i)} = \{k_1^{(i)}, \cdots k_{i-1}^{(i)}, k_i^{(i)}, 0, \cdots 0\}$ which satisfies $\mathbf{k}^{(i)} \cdot \boldsymbol{\alpha} = 0$ for all $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$. Choose an $\boldsymbol{\alpha}^{(i)}$ with the following components: $\alpha_1^{(i)} = \cdots = \alpha_{i-1}^{(i)} = 0$ (take $q_1 = \cdots = q_{i-1} = 0$), $\alpha_i^{(i)} = 1/k_{ii}$ (take $q_i = 1$), and the rest of the components calculated from (17) by taking $q_{i+1} = \cdots = q_n = 0$. Then from $\mathbf{k}^{(i)} \cdot \boldsymbol{\alpha}^{(i)} = 0$ we obtain $k_i^{(i)} = \nu_i \cdot k_{ii}$, where ν_i is an integer. Therefore the element $\mathbf{k}^{(i)} - \nu_i \mathbf{k}_i$ has its $j, j+1, \cdots n$ components equal to zero (remember that $k_{ii+1} = \cdots = k_{in} = 0$), and, according to the assumption, it can be written as $\mathbf{k}^{(i)} - \nu_i \mathbf{k}_i = \sum_{r=1}^{i-1} \nu_r k_r$ ($\nu_1, \cdots \nu_{i-1}$ are integers), i.e., $\mathbf{k}^{(i)} = \nu_1 \mathbf{k}_1 + \cdots + \nu_i \mathbf{k}_i$. q.e.d.

In the case that $m > 0$ we proceed in a similar fashion. Assume that there is a $\mathbf{k}^{(1)} \in \bigcap_\alpha \mathfrak{G}_\alpha$ having the form $\mathbf{k}^{(1)} = \{k^{(1)}, \cdots k_{m+1}^{(1)}, 0, \cdots 0\}$. Then $\boldsymbol{\alpha} \cdot \mathbf{k}^{(1)} = 0$, $\boldsymbol{\alpha} \in \mathfrak{A}(\mathfrak{U})$; using (16) we write this explicitly:

$$k_1^{(1)} \alpha_1 + \cdots + k_m^{(1)} \alpha_m + k_{m+1}^{(1)} \left(\frac{q_1}{k_{1m+1}} - \sum_{s=1}^{m} \frac{k_{1s}}{k_{1m+1}} \alpha_s \right)$$
$$= \left(k_1^{(1)} - \frac{k_{m+1}^{(1)}}{k_{1m+1}} k_{11} \right) \alpha_1 + \cdots$$
$$+ \left(k_m^{(1)} - \frac{k_{m+1}^{(1)}}{k_{1m+1}} k_{1m} \right) \alpha_m$$
$$+ \frac{k_{m+1}^{(1)}}{k_{1m+1}} q_1 = \nu(q_1, \alpha_1, \cdots \alpha_m). \quad (18)$$

Here $\nu(q_1, \alpha_1, \cdots \alpha_m)$ can assume only integer values. Take $q_1 = 1$ and consider (18) in a neighborhood of $\alpha_1 = \cdots = \alpha_m = 0$. As the left-hand side of (18) is a continuous function in $\alpha_1, \cdots \alpha_m$, while the right-hand side can take on only discrete values, $\nu(1, \alpha_1, \cdots \alpha_m) = \nu_1 = \text{const}$ for $\alpha_1, \cdots \alpha_m$ in a sufficiently small neighborhood of $\alpha_1 = \cdots = \alpha_m = 0$. But then (18) can be fulfilled for all $\alpha_1, \cdots \alpha_m$ from such a neighborhood only if

$$k_1^{(1)} = \frac{k_{m+1}^{(1)}}{k_{1m+1}} k_{11}, \cdots k_m^{(1)} = \frac{k_{m+1}^{(1)}}{k_{1m+1}} k_{1m}, \qquad \frac{k_{m+1}^{(1)}}{k_{1m+1}} = \nu_1,$$

that is $\mathbf{k}^{(1)} = \nu_1 \mathbf{k}_1$, which is the same result as for the case $m = 0$. The final result can be obtained again by the method of complete induction, using the same arguments as for the case $m = 0$. Q.E.D.

If we are given a set of fields and we determine somehow the group of multiplicative symmetries $\mathfrak{u}$, i.e., we find out $\mathfrak{A}(\mathfrak{U})$, and we would like to find out $\mathfrak{G}(\varphi_1, \cdots \varphi_n)$, we don't have, fortunately, to deal with *all* the $\mathfrak{G}_\alpha$'s. This assertion is based on

the easily demonstrable relation

$$(\mathcal{G}_{\alpha'} \cap \mathcal{G}_{\alpha''}) \subset \mathcal{G}_{\alpha'+\alpha''}, \qquad \alpha', \alpha'' \in \mathfrak{A}(\mathfrak{U}).$$

Therefore we can write

$$\mathcal{G}(\varphi_1, \cdots \varphi_n) = (\bigcap_{\alpha \in \mathfrak{P}(\mathfrak{U})} \mathcal{G}_\alpha) \cap (\bigcap_{\alpha \in \mathfrak{D}(\mathfrak{U})} \mathcal{G}_\alpha)'.$$

Notice that the set $\alpha^{(1)}, \cdots \alpha^{(n-m)} \in \mathfrak{D}(\mathfrak{U})$ corresponding to $q_1 = 1, q_2 = \cdots = q_{n-m} = 0; \cdots;$ $q_1 = \cdots = q_{n-m-1} = 0, q_{n-m} = 1$, respectively, is a system of generators of $\mathfrak{D}(\mathfrak{U})$. [It has been assumed that $k_{rm+r} \geq 2$ for all $r = 1, \cdots n - m$; if this is not the case and there are r_0 of these k_{rm+r}, which are equal to one, the introduced system of generators of $\mathfrak{D}(\mathfrak{U})$ will contain only $(n - m - r_0)$ elements.] Therefore,

$$\bigcap_{\alpha \in \mathfrak{D}(\mathfrak{U})} \mathcal{G}_\alpha \equiv \mathcal{G}_{\alpha^{(1)}} \cap \cdots \cap \mathcal{G}_{\alpha^{(n-m)}}. \tag{19}$$

The continuous group $\mathfrak{C}(\mathfrak{U})$ will not give much trouble! Writing the general element of $\mathfrak{C}(\mathfrak{U})$ in the form $\alpha = \mathbf{c}_1 t_1 + \cdots + \mathbf{c}_m t_m$ (mod 1), the condition that $\mathbf{k} \in \mathfrak{C}(\mathfrak{U})$, namely that $\alpha \cdot \mathbf{k} = 0$ is explicitly written (ν is integer-valued)

$$\sum_{r=1}^{m} (c_{r1}k_1 + \cdots + c_{rn}k_n)t_r = \nu(t_1, \cdots t_m).$$

By an argument already used in the course of proving Theorem 9, we conclude that all $\mathbf{k} \in \bigcap_{\alpha \in \mathfrak{C}} \mathcal{G}_\alpha$ are determined by the set of equations

$$c_{r1}k_1 + \cdots + c_{rn}k_n = 0, \qquad r = 1, \cdots m. \tag{20}$$

Until now we have dealt with the connection between multiplicative symmetries and Wightman distributions. How do these results affect the Hilbert space $\mathfrak{H}$? To deal with this problem we have to introduce some notations very similar to the ones in Definition 3.

Definition 4. For any $\mathbf{k} \in Z^n$, denote by $|\mathbf{k}\rangle$ the set of vectors from D_0, obtained by applying to the vacuum state Ψ_0 a product of field operators, each field having been taken at some arbitrary point $f \in \mathcal{S}_4$; this product of field operators has to contain a complex field φ_i k_i' times, and its adjoint $k_i'' = k_i' - k_i$ times, and it has to contain a real field $\varphi_{i'}$ $k_{i'}$ times.

It is a restatement of the axiom about the cyclicity of Ψ_0 (Axiom V) when we say that the closure of the linear manifold determined by the set $\bigcup_{\mathbf{k} \in Z^n} |\mathbf{k}\rangle$ is the whole $\mathfrak{H}$.

Theorem 10. If $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ does not coincide with Z^n then the Hilbert space $\mathfrak{H}$ is the direct sum of nontrivial mutually orthogonal subspaces $\bar{D}_g$, where g runs over the factor group of Z^n by its (invariant) subgroup $\mathcal{G}(\varphi_1, \cdots \varphi_n)$: $\mathfrak{H} = \bigoplus_g \bar{D}_g$, $g \in Z^n/\mathcal{G}$; D_g is the linear manifold spanned by the vectors of the union $\bigcup_{\mathbf{k} \in g} |\mathbf{k}\rangle$. Each $\bar{D}_g$ is left invariant by $U(a, 1)$: $U(a, 1)\bar{D}_g \subset \bar{D}_g$ (i.e., the energy–momentum operator P^μ takes $D \cap \bar{D}_g$ into itself). All the vectors in $\bar{D}_g$ are eigenvectors [with the same eigenvalue $\exp(2\pi i \alpha \mathbf{k})$, $\mathbf{k} \in g$] of each multiplicative symmetry U_α.

Proof: First notice that, if $\Psi_{\mathbf{k}} \in |\mathbf{k}\rangle$ and $\Psi_{\mathbf{k}'} \in |\mathbf{k}'\rangle$, then $(\Psi_{\mathbf{k}}, \Psi_{\mathbf{k}'})$ is a Wightman distribution and that $(\Psi_{\mathbf{k}}, \Psi_{\mathbf{k}'}) \in \langle \mathbf{k}' - \mathbf{k} \rangle$. Therefore if $\Psi_{\mathbf{k}} \in |\mathbf{k}\rangle$, $\mathbf{k} \in Z^n$, then $||\Psi_{\mathbf{k}}||^2 \in \langle \mathbf{0} \rangle$. As *all* the elements of $\langle \mathbf{0} \rangle$ do not identically vanish (see Theorem 3), each $|\mathbf{k}\rangle$ must contain nonzero vectors; and so does then the linear manifold D_g spanned by all the vectors in $\bigcup_{\mathbf{k} \in g} |\mathbf{k}\rangle$, $g \in Z^n/\mathcal{G}$, which proves that $\bar{D}_g$ is a nontrivial subspace of $\mathfrak{H}$ (except when $\mathcal{G} \equiv Z^n$, in which case $Z^n/\mathcal{G}$ is trivial).

On the other hand, if $\mathbf{k} \in g$, $\mathbf{k}' \in g'(g, g' \in Z^n/\mathcal{G})$ and $g \neq g'$, then $\mathbf{k} - \mathbf{k}' \notin \mathcal{G}(\varphi_1, \cdots \varphi_n)$ so that $\langle \mathbf{k} - \mathbf{k}' \rangle = 0$; which means that all vectors from $\bigcup_{\mathbf{k} \in g} |\mathbf{k}\rangle$ are orthogonal to all vectors from $\bigcup_{\mathbf{k}' \in g'} |\mathbf{k}'\rangle$. Therefore, $\bar{D}_g$ is orthogonal to $\bar{D}_{g'}$ for $g \neq g'$. Obviously $\mathfrak{H} = \bigoplus_{g \in Z^n/\mathcal{G}} \bar{D}_g$.

We notice that each $\Psi_{\mathbf{k}} \in |\mathbf{k}\rangle$ is an eigenvector of each $U_\alpha \in \mathfrak{U}$ having the eigenvalue $\exp(2\pi i \alpha \cdot \mathbf{k})$. All the vectors in $\bigcup_{\mathbf{k} \in g} |\mathbf{k}\rangle$ belong to the same eigenvalue $\exp(2\pi i \alpha \cdot \mathbf{k})$, where $\mathbf{k}$ is any element of $g \in Z^n/\mathcal{G}$. As all U_α are bounded and therefore continuous operators, all the vectors in $\bar{D}_g$ belong to this same eigenvalue.

Now, as a consequence of Theorem 9, if $g \neq g'$, there is at least one $\alpha \in \mathfrak{A}(\mathfrak{U})$ such that the eigenvalue of U_α in $\bar{D}_g$ differs from its eigenvalue in $\bar{D}_{g'}$. Otherwise $\mathbf{k} \cdot \alpha = \mathbf{k}' \cdot \alpha (\mathbf{k} \in g, \mathbf{k}' \in g')$ for all $\alpha \in \mathfrak{A}(\mathfrak{U})$ and that would imply (Theorem 9) that $\mathbf{k} - \mathbf{k}' \in \mathcal{G}$, i.e., that $g = g'$, which is a contradiction.

As, according to Theorem 1, $U(a, 1)$ commutes with all U_α, the aforesaid implies that $U(a, 1)\bar{D}_g \subset \bar{D}_g$. Q.E.D.

2. IMPLICATIONS OF THE RESULTS OF THE FIRST SECTION

A. Some General Remarks

Throughout the first section, for the sake of generality, the real fields have been considered along with the complex ones. We would like now to remark that when, for instance, φ_1 is a real

field, $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ contains, according to Theorem 4, the element $\{2, 0, \cdots 0\}$; as a consequence, $\mathfrak{A}_1(\mathfrak{U})$[13] may contain only $\alpha = 0, \frac{1}{2}$. This is an obvious result, which otherwise follows immediately when we insert the Hermiticity condition in Definition 1(b), of Sec. 1 to obtain $\lambda_1 = \bar{\lambda}_1$.

The application of the exposed theory on the case of free complex fields does not give any result which is not obvious beforehand. Because of the particularly simple structure of $\mathfrak{K}$ in this case [$\mathfrak{K}$ being just a direct product of the Hilbert spaces $\mathfrak{K}^{(i)}$ for the complex free fields $\varphi_1, \cdots \varphi_n$ and each $\mathfrak{K}^{(i)}$ the direct sum of one, two, etc. particle (antiparticle) subspaces], the group $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ for complex free fields consists only of the zero element. Therefore $\mathfrak{A}(\mathfrak{U})$ coincides with $\mathfrak{A}^n$, and $Z^n/\mathcal{G}$ is identical with Z^n. The complex free fields have all the possible multiplicative symmetries!

One might consider as reasonable to expect that the identical vanishing of Wightman distributions can be always ascribed to the existence of multiplicative symmetries. Because of Theorem 9, it would be sufficient (and, of course, necessary) for this to be true to establish that *all* the elements of an $\langle \mathbf{k} \rangle \neq 0$ (i.e., $\mathbf{k} \in \mathcal{G}$) do not vanish identically. Now, Theorem 3 tells us that when $\mathbf{k} \in \mathcal{G}$ then all elements of $\langle \mathbf{k} \rangle$ of *sufficiently* high order do not vanish identically. On the other hand, there is reason to believe that in some cases this is not true for elements of the lowest order belonging to a $\langle \mathbf{k} \rangle \neq 0$. To prove this assertion rigorously would mean to give a theory formulated in the language of Wightman axioms in which there is at least one $\langle \mathbf{k} \rangle \neq 0$ such that some of the Wightman distributions of the lowest order belonging to $\langle \mathbf{k} \rangle$ are vanishing identically. Unfortunately, at present there are not known theories formulated in a mathematically precise language, except the free fields and generalized free fields, which do not offer an illustration of the desired type. But there can be found among the present mathematically nonrigorously formulated field theories such ones which, when treated by perturbation methods, yield the desired kind of example. As any future theory formulated in the mathematically rigorous language of Wightman axioms is expected and hoped to have many common (desirable) features with the already existing field theories, it can be expected that they will share some of the less significant features. Therefore we have an indication, even though not a rigorous proof, that the hope that the systematic identical vanishing of Wightman distributions cannot be *always* accounted for by the existence of some multiplicative symmetries.

[13] See Theorem 7 for definition of $\mathfrak{A}_i(\mathfrak{U})$.

B. Multiplicative Symmetries and the Restricted Lorentz Group

In Sec. 1, while studying the properties of $\mathcal{G}(\varphi^{(1)}, \cdots \varphi^{(N)})$,[14] there was only one axiom left, whose possible influence on $\mathcal{G}$ has been not completely taken into account; that is Axiom III. Therefore, in this subsection we intend to formulate two theorems and make a few comments on the relation between multiplicative symmetries and the behavior of field components under the restricted Lorentz group in a given quantum field theory which obeys the Wightman axioms. A slight enlargement of our terminology, contained in Definitions 1 and 2, will be useful.

Definition 1. A multiplicative symmetry $U \in \mathfrak{U}(\varphi^{(1)}, \cdots \varphi^{(N)})$, given by the relations

$$U\varphi_r^{(\nu)}(f)U^{-1} = \lambda_r^{(\nu)}\varphi_r^{(\nu)}(f), \qquad \nu = 1, \cdots N,$$
$$r = -s_\nu, -s_\nu + 1, \cdots s_\nu,$$

will be called a *proper* multiplicative symmetry if

$$\lambda_{-s_\nu}^{(\nu)} = \lambda_{-s_\nu+1}^{(\nu)} = \cdots = \lambda_{s_\nu}^{(\nu)} = \lambda^{(\nu)}, \quad \nu = 1, \cdots N.$$

All the multiplicative symmetries not satisfying this condition will be called *improper*.

We can now formulate a theorem which is easy to prove using the same technique as for the proof of Theorem 1 in Sec. 1.

Theorem 1. A necessary and sufficient condition for a multiplicative symmetry U_α to commute with $U(0, \Lambda)$, $\Lambda \in \mathfrak{L}_+^\uparrow$, is that U_α is a proper multiplicative symmetry. Therefore, for quantum field theories in which $M^{\mu\nu}$ ($\mu, \nu = 0, 1, 2, 3$) correspond to observables, the supersymmetry group[15] $\mathfrak{S}$ consists only of proper multiplicative symmetries.

The question can be now raised whether improper multiplicative symmetries can exist at all in a field theory obeying Wightman axioms. An example which is trivial but sufficient to assure us that the answer to the question asked is affirmative, is the case of "complex" free fields, which possess all possible, proper and improper, multiplicative symmetries.

Definition 2. A multiplicative symmetry $U_\alpha \in$

[14] See Sec. 1, Theorem 4. As in this section we are interested in the behavior of $\mathcal{G}$ in relation to Lorentz rotations, the explicit notation of different components of the same spinor field is suppressed, i.e., $\varphi^{(1)} = \{\varphi_{-s_1}(1), \cdots \varphi_{+s_1}(1)\}$, etc.

[15] See Part C, this section.

$\mathfrak{U}(\varphi^{(1)}, \cdots \varphi^{(N)})$ will be called *trivial with respect* to the field $\varphi^{(\nu_0)}$ if

$$\alpha^{(\nu_0)}_{-s_{\nu_0}} = \cdots = \alpha^{(\nu_0)}_{+s_{\nu_0}} = 0,$$

i.e.

$$\lambda^{(\nu_0)}_{-s_{\nu_0}} = \cdots = \lambda^{(\nu_0)}_{+s_{\nu_0}} = 1.$$

It will be called a proper multiplicative symmetry *with respect* to the field $\varphi^{(\nu_1)}$ if

$$\alpha^{(\nu_1)}_{-s_{\nu_1}} = \cdots = \alpha^{(\nu_1)}_{-s_{\nu_1}}, \quad \text{i.e.} \quad \lambda^{(\nu_1)}_{-s_{\nu_1}} = \cdots = \lambda^{(\nu_1)}_{+s_{\nu_1}}.$$

Theorem 2. If the vacuum expectation value of one of the components of the field $\varphi^{(\nu)}$ does not vanish identically, say $\langle\varphi^{(\nu)}_{s_\nu}(f)\rangle_0 \not\equiv 0$, $f \in \mathfrak{S}_4$, then the vacuum expectation values of the rest of the components of $\varphi^{(\nu)}$ are not identically equal to zero, i.e., $\langle\varphi^{(\nu)}_r(f)\rangle_0 \not\equiv 0$, $r = -s_\nu, -s_\nu + 1, \cdots s_\nu$, $f \in \mathfrak{S}_4$. Therefore $\mathfrak{U}(\varphi^{(1)}, \cdots \varphi^{(N)})$ contains only multiplicative symmetries which are trivial with respect to the field $\varphi^{(\nu)}$.

Proof: The proof is very simple and makes use only of the irreducibility of $S^{(\nu)}(A)$ (see Axiom III).

Take a $(2s_\nu + 1)$-dimensional Hilbert space $L^{(\nu)}$ in which, for an orthonormal basis e_r $(r = -s_\nu, \cdots s_\nu)$, we define the finite dimensional representation of $\mathfrak{L}^\uparrow_+$: $S^{(\nu)}(A)e_r = \sum_{r'=-s}^{s} S^{(\nu)}_{r'r}(A)e_{r'}$. Denote by $L^{(\nu)}_0$ the subspace spanned by all vectors

$$\sum_{r=-s_\nu}^{s_\nu} \langle\varphi^{(\nu)}_r(f)\rangle_0 e_r, \quad f \in \mathfrak{S}_4.$$

Because of the relation

$$\langle\varphi^{(\nu)}_r(f_{(0,A^{-1})})\rangle_0 = \sum_{r'=-s_\nu}^{s_\nu} S^{(\nu)}_{rr'}(A)\langle\varphi^{(\nu)}_{r'}(f)\rangle_0,$$

derivable from Axiom III, we have that $L^{(\nu)}_0$ is left invariant by $S^{(\nu)}(A)$, $A \in SL(2, C)$. As $S^{(\nu)}(A)$ is an irreducible representation of $SL(2, C)$, we conclude that $L^{(\nu)}_0$ is a trivial subspace of $L^{(\nu)}$. But as, according to the assumption, there is a $f_0 \in \mathfrak{S}_4$ such that $\langle\varphi^{(\nu)}_{s_\nu}(f_0)\rangle_0 \not\equiv 0$, $L^{(\nu)}_0$ cannot coincide with the null vector, i.e., $L^{(\nu)}_0 = L^{(\nu)}$. This relation cannot be true if for some integer r, $-s_\nu \leq r \leq s_\nu$, we have for all $f \in \mathfrak{S}_4$ that $\langle\varphi^{(\nu)}_r(f)\rangle_0 = 0$. q.e.d.

The last part of the theorem is an obvious consequence of

$$\langle\varphi^{(\nu)}_r(f)\rangle_0 = (U^{-1}_\alpha \Psi_0, \varphi^{(\nu)}_r(f)U^{-1}\Psi_0)$$
$$= \exp(2\pi i\alpha^{(\nu)}_r)\langle\varphi^{(\nu)}_r(f)\rangle_0, \qquad r = -s_\nu, \cdots s_\nu.$$

It follows, of course, also from Theorem 5 of Sec. 1. Q.E.D.

In connection with Theorem 3, it is interesting to remark that it has become a convention[16] to take $\langle\varphi^{(\nu)}_r(f)\rangle_0 \equiv 0$, $\nu = 1, \cdots N$, $r = -s_\nu, \cdots s_\nu$, $f \in \mathfrak{S}_4$. Namely, if we have a field theory which does not fulfill this requirement, it is always possible to introduce new fields $\varphi'^{(\nu)}_r(f)$, defined by

$$\varphi'^{(\nu)}_r(f) = \varphi^{(\nu)}_r(f) - \langle\varphi^{(\nu)}_r(f)\rangle_0,$$

which conform to the convention. The change from $\varphi^{(\nu)}$ to $\varphi'^{(\nu)}$ is regarded as trivial, probably because in all present field theories, which are formulated in terms of equations of motion and canonical commutation relations, it leaves both unaffected. But, as Theorem 3 shows, a significant change in the group of multiplicative symmetries might occur: the theory of $\varphi'^{(\nu)}$ fields *can be* richer in multiplicative symmetries than the theory with $\varphi^{(\nu)}$ fields.

It would be natural to expect that, by using the relativistic invariance of a considered field theory, one would be able to infer, from the fact that some Wightman distributions vanish identically, that a larger class of these distributions are also identically equal to zero. In other words, it looks as if the theorems in Sec. 1 do not exhaust all the possible restrictions on $\mathcal{G}(\varphi_1, \cdots \varphi_n)$ which are derivable from the invariance of the theory under $\mathfrak{L}^\uparrow_+$. It could be expected that, by using a generalization of the method employed in proving Theorem 2, one could prove similar theorems for Wightman distributions containing two or more field components.

The reason why such a method fails is that Wightman distributions containing two or more field components undergo, in their spinor indices, transformations which are direct products of irreducible Lorentz representations; i.e., these transformations are in general *reducible*. To inquire what is the effect of the condition $\mathfrak{W}_p(f) \equiv 0$, $f \in \mathfrak{S}_{4p}$, valid for *some* Wightman distributions of the pth order, on the behavior of the other Wightman distributions of the same order, one has to proceed in the following way:

Let us say that an irreducible representation of $\mathfrak{L}^\uparrow_+$ occurs only once in the direct product $S^{(\nu_1)} \otimes \cdots \otimes S^{(\nu_p)}$, of the irreducible representations $S^{(\nu_1)}(A)$, $\cdots$ $S^{(\nu_p)}(A)$, according to which the p-order Wightman distributions $\mathfrak{W}_p$ transform in *their spin indices*. Then one has to find the conditions which define the set $\mathfrak{S}^{(0)} \subset \mathfrak{S}_{4n}$ of all $f \in \mathfrak{S}_{4n}$ for which all $\mathfrak{W}(f)$ of the pth order can be regarded, in their spin indices, to be vectors in the subspace of the space in which $S^{(\nu_1)} \otimes \cdots \otimes S^{(\nu_p)}$ is acting, which is left invariant by the mentioned irreducible representation. Applying a reasoning analogous to the one used in proving Theorem 2,

[16] See, e.g., R. Haag and B. Schroer, J. Math. Phys. 3, 248 (1962).

one can infer from the relation $\mathfrak{W}_p(f) \equiv 0, f \in \mathcal{S}^{(0)}$, valid for a certain subclass Wightman distribution of the pth order, that the same relation holds for all the p-order Wightman distributions.

A situation, when the procedure which has just been described is applicable, is encountered in the case when the spin of all fields of the considered theory is not greater than $\frac{1}{2}$. But already in this simplest case the conditions defining an $\mathcal{S}^{(0)}$ with the mentioned properties are of a complicated and physically uninteresting nature. They may be of some interest in case of certain given theories where these conditions may acquire a simpler form. Therefore the conclusion is that the invariance of the theory under $\mathfrak{L}_+^\uparrow$ does not supply us with any additional theorems which are of completely *general* nature and for whose proof *only* the validity of Wightman axioms is required.

C. Multiplicative Symmetries as Supersymmetries

In the last few years the application of von Neumann's achievements in the theory of W^* algebras on the algebra of observables has resulted in a rigorously formulated theory of the structure of the algebra determined by all observables a quantum-mechanical theory.[17] As a result of this approach, one can formulate the following theorem:

The W^* algebra of all observables, $\mathfrak{N}$, determines, uniquely up to permutations in the index λ, a decomposition

$$\mathfrak{K} = \int^{\oplus} \mathfrak{K}_\lambda \sqrt{d\sigma(\lambda)}$$

of the Hilbert space $\mathfrak{K}$ of states, into a direct integral of Hilbert spaces $\mathfrak{K}_\lambda$. This decomposition is such that for each λ the W^* algebra $\mathfrak{N}_\lambda$ induced by $\mathfrak{N}$ in $\mathfrak{K}_\lambda$ is irreducible, i.e., its commutant $\mathfrak{N}_\lambda'$ consists only of multiples of the identity, $\mathfrak{N}_\lambda' = \{\alpha \mathbf{1}\}$. $\mathfrak{N}$ consists of all operators A such that $A\Psi = \{A_\lambda \Psi_\lambda\}$, $\Psi \in \mathfrak{K}$, $\Psi_\lambda \in \mathfrak{K}_\lambda$, where A_λ is σ-measurable and essentially bounded.

In this theory it is assumed that the center $\mathfrak{z}$ of the algebra $\mathfrak{N}$ coincides with its commutant $\mathfrak{N}'$: $\mathfrak{z} = \mathfrak{N} \cap \mathfrak{N}' = \mathfrak{N}'$.[18] Now, in a quantum field theory, a self-adjoint operator which commutes with all observables defines a superselection rule. It is an empirical fact that the spectrum of the self-adjoint operator, corresponding to any of the known superselection rules, is discrete.

On the other hand, it is a mathematical result that a W^* algebra is generated by its unitary operators. A unitary operator $\mathfrak{U}$ which belongs to $\mathfrak{N}'$ and is not a multiplum of the identity is called by Jauch a *super-symmetry*. It is clear that the super-symmetries, together with the identity, form a group—the *super-symmetry group* $\mathfrak{S}$. It is an assumption adopted by the conventional field theory that $\mathfrak{S}$ is a subgroup of the group $\mathfrak{U}$ of multiplicative symmetries.[19] If we accept this assumption as part of the framework of the axiomatic approach to quantum field theory, then, because of Theorem 10, Sec. 1, the decomposition of $\mathfrak{K}$ into a direct integral $\mathfrak{K} = \int^{\oplus} \mathfrak{K}_\lambda \sqrt{d\sigma_\lambda}$ becomes a decomposition in a direct sum of subspaces of $\mathfrak{K}$. If $\mathfrak{S} \equiv \mathfrak{U}$ then $\mathfrak{K}_\lambda \equiv \overline{D_{\varrho(\lambda)}}$ and $\mathfrak{K} = \bigoplus_\lambda \overline{D_{\varrho(\lambda)}} = \bigoplus_\lambda \mathfrak{K}_\lambda$. In general we can say:

Theorem 3. If, in a quantum field theory, the W^* algebra generated by the elements of the supersymmetry group $\mathfrak{S}$ coincides with the center $\mathfrak{z}$ of $\mathfrak{N}$, if $\mathfrak{z} \equiv \mathfrak{N}$, and if $\mathfrak{S}$ is a subgroup of $\mathfrak{U}$, $\mathfrak{S} \subset \mathfrak{U}$, then

$$\mathfrak{K} = \bigoplus_\lambda \mathfrak{K}_\lambda,$$

where each subspace $\mathfrak{K}_\lambda$ is left invariant by all the operators in $\mathfrak{N}$. The algebras $\mathfrak{N}_\lambda$ induced by $\mathfrak{N}$ in $\mathfrak{K}_\lambda$ are irreducible.

ACKNOWLEDGMENT

The author would like to express his gratitude to Professor Wightman for suggesting to him the study of multiplicative symmetries as well as for many very valuable discussions and suggestions. He would also like to thank him for proofreading the first section of the paper.

APPENDIX. WIGHTMAN AXIOMS[1,2]

I. Each physical state is described by a ray $\mathbf{\Psi}$ in a *separable* Hilbert space $\mathfrak{K}$: $\mathbf{\Psi} = \{\alpha\Psi, \alpha \text{ any complex number}\}$ where Ψ is a certain vector from $\mathfrak{K}$. In general, because of superselection rules, the converse does not have to be true: there *can be* rays which do not represent any physical states.

II. A quantum-mechanical physical theory will be called a *field theory with N spinor fields*, each field having spin s_ν, $s_\nu = 0, 1, 2 \cdots (\nu = 1, \cdots N)$,

[17] See J. M. Jauch, Helv. Phys. Acta **33**, 711 (1960); J. M. Jauch and B. Misra, *ibid.* **34**, 699 (1961); J. M. Jauch, "Continuous Geometry and Superselection Rules," CERN 61-14 (1961); A. Galindo, A. Morales, and R. Nuñez-Lagos, J. Math. Phys. **3**, 324 (1962).

[18] This assumption is equivalent to the assumption that $\mathfrak{N}$ contains a maximal Abelian W^* algebra. (See the second and last references of the previous footnote.)

[19] It is evident that in a theory in which $M_{\mu\nu}$ corresponds to observables and which besides possesses improper multiplicative symmetries, $\mathfrak{S}$ does not coincide with $\mathfrak{U}$ (see Theorem 1, Sec. 2).

if there is a set of operator valued functionals

$$\varphi^{(1)}_{-s_1}(f),\quad \varphi^{(1)}_{-s_1+1}(f),\ \cdots\ \varphi^{(1)}_{s_1}(f);\ \cdots\ ;\varphi^{(N)}_{-s_N}(f),$$
$$\varphi^{(N)}_{-s_N+1}(f),\ \cdots\ \varphi^{(N)}_{s_N}(f),$$

$f \in \mathcal{S}_4$ ($\mathcal{S}_r$ is the space of infinitely differentiable functions in r variables, which vanish at infinity, together with *all* their derivatives, faster than any polynomial in their arguments), fulfilling the following conditions:

1. The intersection of the domains of definition of all these $n = \sum_{\nu=1}^{N}(2s_\nu + 1)$ operators for *all* values of their arguments $f \in \mathcal{S}_4$ contains a domain $D \subset \mathfrak{H}$ having the following two properties:

   (a) $\bar{D} = \mathfrak{H}$, i.e., D is everywhere dense in $\mathfrak{H}$.

   (b) $\varphi^{(\nu)}_r(f)D \subset D,\quad f \in \mathcal{S}_4,\quad \nu = 1, \cdots N,$
   $$r = -s_\nu,\quad -s_\nu + 1, \cdots s_\nu.$$

2. $(\Phi, \varphi^{(\nu)}_r(f)\Psi),\quad f \in \mathcal{S}_4,\quad \nu = 1, \cdots N,$
   $$r = -s_\nu, \cdots s_\nu$$
   is a distribution in $\mathcal{S}'_4$, in the sense of Schwartz, whenever $\Phi, \Psi \in D$.

3. When acting on vectors in D, the field operators are linear functionals, i.e.,
   $$\varphi^{(\nu)}_r(af + bg)\Psi = a\varphi^{(\nu)}_r(f)\Psi + b\varphi^{(\nu)}_r(g)\Psi,$$
   $$\nu = 1, \cdots N,\quad r = -s_\nu, \cdots + s_\nu,\quad \Psi \in D,$$
   where a and b are any complex numbers and $f, g \in \mathcal{S}_4$.

III. The field theory is a *relativistic* field theory if unitary (anti-unitary) operators $U(a, \Lambda)$ performing the transition to a relativistically transformed state have the following properties:

1. They form a continuous unitary representation of the inhomogeneous unimodular group $SL(2, C)$ ($\Lambda \to \pm A$ for $\Lambda \in \mathfrak{L}^{\uparrow}_{+}$),
   $$||(U(a', A') - U(a, A))\Psi|| \to 0$$
   $$\text{for}\quad \{a', A'\} \to \{a, A\},\quad \Psi \in \mathfrak{H}.$$

2. $U(a, \Lambda)D \subset D,\quad \{a, \Lambda\} \in \mathcal{P}$.[20]

3. $$U(a, A)\varphi^{(\nu)}_r(f)U(a, A)^{-1} = \sum_{r'=-s_\nu}^{s_\nu} S^{(\nu)}_{rr'}(A)\varphi^{(\nu)}_{r'}(f_{\{a,\Lambda\}}),$$
   $$\nu = 1, \cdots N,\quad f \in \mathcal{S}_4,$$
   $$f_{\{a,\Lambda\}}(x) = f(\Lambda^{-1}(x - a)) \in \mathcal{S}_4,$$
   where $S^{(\nu)}(A)$ is an irreducible representation in a $(2s_\nu + 1)$-dimensional vector space of the homogeneous restricted Lorentz group as well as of the three-dimensional rotation group (double-valued if s_ν is half-integral).

IV. (Spectral conditions) As a consequence of the asumptions of Axiom III, the representation of the translation group can be always written in the form

$$U(a, 1) = \int \exp(ip_\nu a^\nu)\, d\mu(p),\quad \mathfrak{H} = \int^{\oplus} \mathfrak{H}_p \sqrt{d\mu(p)},$$

where $d\mu(p)$ is a projection-valued measure required to fulfill the following conditions:

1. $d\mu(p) = 0$ for p outside the forward light cone, i.e., supp $d\mu(p) \subset \bar{V}_+$.

2. (uniqueness of vacuum state) $\mu(0)\mathfrak{H} = \{\alpha\Psi_0,\ \alpha$ any complex number$\}$ where $\Psi_0 \in \mathfrak{H}$ is a vector, which it is convenient to choose to be of unit norm, i.e., $||\Psi_0|| = 1$; Ψ_0 will be called the vacuum state.

3. $\Psi_0 \in D$ and $U(a, A)\Psi_0 = \Psi_0$.

V. (Cyclicity of the vacuum state) The linear manifold D_0 obtained by applying on Ψ_0 all possible polynomials in the field operators, each field operator being taken at any point $f \in \mathcal{S}_4$, is dense in $\mathfrak{H}$, i.e., $\bar{D}_0 = \mathfrak{H}$.

As a consequence of the nuclear theorem of Schwartz, a unique meaning can be given to the smearing of any product of n field operators ($n = 1, 2, 3, \cdots$) with a function from $\mathcal{S}_{4n}$, when this product is applied on a vector from D. Therefore we can define D_1 as the linear manifold obtained when each term of polynomials in fields applied on Ψ_0 is smeared with an $f \in \mathcal{S}_{4n}$ (in case that the term is the product of n fields). Obviously $D_1 \supset D_0$ and $\bar{D}_1 = \mathfrak{H}$.

VI. (Local commutativity) For any $f, g \in \mathcal{S}_4$ such that $f(x)\cdot g(y) = 0$ if $(x - y)^2 \geq 0$, the relation

$$[\varphi^{(\nu)}_r(f), \varphi^{(\nu')}_{r'}(g)]_{\pm}\Psi = 0,\quad \nu, \nu' = 1, \cdots N,$$
$$r = -s_\nu, \cdots s_\nu,\quad r' = -s_{\nu'}, \cdots s_{\nu'},\quad \Psi \in D$$

is assumed to hold, where the commutator is taken if at least one of the fields is a Bose field (integer spin), and the anticommutator is taken when both fields are Fermi fields (half-integral spin).

VII. (Ruelle's axiom) The "in" and "out" states span the whole Hilbert space $\mathfrak{H}$ of interacting fields, i.e.,

$$\mathfrak{H}^{\text{in}} = \mathfrak{H} = \mathfrak{H}^{\text{out}}.$$

[20] $\mathcal{P}$ denotes the Poincaré group.

Relativistic Particle Dynamics and the S Matrix*

R. FONG† AND J. SUCHER
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 30 July 1963)

Direct-interaction theories are examined from the viewpoint of relativistic scattering theory and the associated concept of "asymptotic covariance." It is pointed out that with any two-particle Hamiltonian which has no bound states there can be associated a variety of representations of the Lie algebra of the inhomogeneous Lorentz group (IHLG), although the S matrix is in general not covariant. It is shown that the requirement of asymptotic covariance ensures both the covariance of the S matrix and the existence of a unique representation of the IHLG to be associated with the relativistic two-particle system. The connection between the Lie algebra, the covariant form of the S matrix, and the uniqueness of $\mathbf{K}$, the generator of pure Lorentz transformations, is thereby clarified. The extension of these considerations to include bound states is made. The form of H given by Bakamjian and Thomas is shown to satisfy asymptotic covariance and, moreover, to be the most general form of interest from the viewpoint of relativistic scattering theory, thereby including as a special case a form of H suggested by Sudarshan. It is also proved that relativistic Hamiltonians of this type do not admit the usual notion of a coupling constant.

1. INTRODUCTION

THE basic physical considerations for a relativistic quantum theory are embodied in the work of Dirac[1] and of Wigner,[2] and have recently been further emphasized by Foldy.[3] These considerations lead to the result that the Hilbert space of a relativistic system is a representation space of the inhomogeneous Lorentz group, the representation being by linear mappings. If we concern ourselves only with the proper inhomogeneous Lorentz group (IHLG), we can take the representation to be unitary. The problem then of finding relativistic theories is, at a minimum, that of finding unitary representations of the IHLG or, equivalently, a set of Hermitian operators that satisfy the well-known commutation relations (Lie bracket relations) for the IHLG and which can then be identified as the infinitesimal generators of a unitary representation of the IHLG.

We are interested here in relativistic systems with interaction. The irreducible representations of the IHLG give only descriptions of elementary systems or particles. Thus, in relativistic particle dynamics, one is interested in the reducible and mathematically equivalent representations, as Dirac[4] has recently emphasized. That is, equivalence from the point of view of representation theory is *not* equivalence from the point of view of physics.

In the case of direct-interaction theories, i.e., theories in which the infinitesimal generators, H, $\mathbf{P}$, $\mathbf{J}$ and $\mathbf{K}$, are functions only of the dynamical variables and which are thus theories in which there are no external fields and the number of particles is conserved, Bakamjian and Thomas[5] and, more recently, Sudarshan[6] have given relativistic theories in the sense that they have displayed nontrivial sets of operators which satisfy these commutation relations. However, these results have not been discussed from the point of view of scattering theory. We shall thus be concerned with the relation between the IHLG and relativistic scattering theory. It is well-known that in relativistic scattering theory the form of the S matrix is such that, considering the simple case of the elastic scattering of two particles,

$$(E_1'E_2')^{\frac{1}{2}}S(\mathbf{q}_1', \mathbf{q}_2'; \mathbf{q}_1, \mathbf{q}_2)(E_1E_2)^{\frac{1}{2}} = \text{invariant}, \qquad (1.1)$$

where E_i, E_i' $(i = 1, 2)$ are the appropriate particle energies.[7,8] We shall thus require that a relativistic theory should not only provide a representation of the IHLG but should also give rise to an S matrix of the form given by Eq. (1.1). Hereafter, an S matrix will be called *covariant* if it satisfies Eq. (1.1).

As will be seen, given any reasonable direct-interaction Hamiltonian H, that does not give rise to a bound state, one can exhibit a set of infinitesimal

* This research was supported in part by the U. S. Air Force and in part by the National Science Foundation.

† Based in part on a dissertation submitted by R. Fong in partial fulfillment of the requirements for a Ph.D. degree at the University of Maryland, July, 1963.

[1] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949).
[2] E. P. Wigner, Nuovo Cimento **3**, 517 (1956).
[3] L. L. Foldy, Phys. Rev. **122**, 275 (1961).
[4] P. A. M. Dirac, Rev. Mod. Phys. **34**, 592 (1962).
[5] B. Bakamjian and L. H. Thomas, Phys. Rev. **92**, 1300 (1953). See also see Ref. 3.
[6] E. C. G. Sudarshan (unpublished). See also E. C. G. Sudarshan, in *Lectures in Theoretical Physics*, Brandeis Summer Institute 1961 (W. A. Benjamin Company, Inc., New York, 1962), Vol. 2.
[7] W. Heisenberg, Z. Phys. **120**, 513 (1943).
[8] C. Møller, Kgl. Danske Videnskab Selskab, Math.-Fys. Medd. **23**, No. 1 (1945).

generators for the IHLG in which H is the infinitesimal operator of time translations. But, to say the least, not all reasonable Hamiltonians—reasonable in the sense that one can apply scattering theory—give rise to a covariant S matrix. *Thus the requirement that one has a representation of the algebra of the IHLG is not sufficient to ensure that the theory yields a covariant S matrix.*

To clarify this situation, a concept of "asymptotic covariance" is introduced. In local relativistic quantum field theory asymptotic covariance is implicit in the basic axioms. However, this is not the case for direct-interaction theories. As will be seen, the explicit requirement of asymptotic covariance, namely that it should be possible to associate with each Lorentz matrix L a linear mapping (L) of Hilbert space onto itself with the property that for the "in" and "out" states $\Psi^{(\pm)}$,

$$\Psi^{(+)}_{\mathbf{q}_1,\mathbf{q}_2} \xrightarrow{(L)} \Psi^{(+)}_{\tilde{\mathbf{q}}_1,\tilde{\mathbf{q}}_2} \quad \text{and} \quad \Psi^{(-)}_{\mathbf{q}_1,\mathbf{q}_2} \xrightarrow{(L)} \Psi^{(-)}_{\tilde{\mathbf{q}}_1,\tilde{\mathbf{q}}_2}, \tag{1.2}$$

where the $\tilde{\mathbf{q}}_i$ are the Lorentz-transformed three-momenta, assures both the uniqueness of the relevant representation of the IHLG and the covariance of the S matrix.

We now give an outline of the following sections. The considerations which lead to the problem of finding a representation of the Lie algebra of the IHLG in relativistic quantum mechanics are briefly reviewed in Sec. 2. In Sec. 3, relativistic center-of-momentum (c.m.) variables are introduced, preliminary to a consideration, in Sec. 4, of two-particle direct-interaction theories. It is shown explicitly that with any two-particle Hamiltonian H which has no bound states there can always be associated a variety of representations of the Lie algebra of the IHLG. The implications of the assumption of asymptotic covariance are studied in Sec. 5, and the connection between the Lie algebra, the covariant form of the S matrix, and the uniqueness of $\mathbf{K}$, the generator of pure Lorentz transformations, is clarified. In Sec. 6, the form of H proposed by Bakamjian and Thomas[5] is considered from the viewpoint of scattering theory. In Sec. 7 the problem of specifying the general class of Hamiltonians which will yield a covariant S matrix is studied. In Sec. 8, the extension of the previous considerations to bound states is made. Section 9 contains a summary of the results and a concluding discussion. Some further aspects of the nature of relativistic "potentials" are considered in the Appendix.

2. THE PROPER INHOMOGENEOUS LORENTZ GROUP

In this section, a brief review is given of the considerations that lead to the problem of finding realizations of the Lie algebra of the IHLG in terms of operators in a Hilbert space.[2]

The principle of relativistic invariance, as formulated by Haag,[9] is incorporated in the following three postulates:

"(a) It should be possible to translate a complete description of a physical system from one coordinate system into every equivalent coordinate system.

"(b) The translation of a dynamically possible description should be again dynamically possible. Expressed in a somewhat more simple language, a succession of events which appears possible to one observer should appear possible also to any other observer.

"(c) The criteria for the dynamical possibility of complete descriptions should be identical for equivalent observers."

Now, with the assumption that all self-adjoint operators are observable, i.e., that any two rays in Hilbert space can be distinguished, postulate (a) implies that the physical system can be described by the same set of rays by all observers and that there exists a well-defined isomorphism between the individual labelings of any two observers. In other words, the same Hilbert space can be used by all observers to describe the physical system and the isomorphism is seen to be an automorphism W of the rays in this Hilbert space. (Foldy[3] has called this Hilbert space the "public" Hilbert space.) Mathematically, if a physical system is seen by two observers A and B to be in the normalized states $|\phi\rangle$ and $|\bar{\phi}\rangle$, respectively, then

$$|\bar{\phi}\rangle = W\,|\phi\rangle.$$

Postulate (b) tells us that transition probabilities are invariant, i.e., $|\langle\phi_1 \mid \phi_2\rangle|^2 = |\langle\bar{\phi}_1 \mid \bar{\phi}_2\rangle|^2$, from which it follows that one need only consider unitary or antiunitary W. Postulate (c) now implies that W can depend only on the relation of the two reference frames of A and B with respect to each other, and not on their relation with respect to some absolute frame of reference in space–time, i.e.,

$$W = W(\mathfrak{L}),$$

where $\mathfrak{L} = \{a, L\}$ is the Lorentz transformation relating the reference frames of A and B. There is then the consistency requirement that for a third observer C, the transition from A to B to C should give the same result as the direct transition from

[9] R. Haag (unpublished). The postulates appear in the given form in Ref. 2.

A to C. This implies that the set of mappings $\{W(\mathfrak{L})\}$ is a representation of the inhomogeneous Lorentz group up to a phase factor. If one is concerned only with the proper inhomogeneous Lorentz group (IHLG), $W(\mathfrak{L})$ may be taken to be unitary and the phase factor may be more or less eliminated.[10]

Thus, the symmetry of the restricted principle of relativity is mathematically embodied in the IHLG. In a relativistic theory one must, therefore, be able to exhibit a unitary representation $W(\mathfrak{L})$ of the IHLG for the system. Equivalently, one must, *at least*, exhibit operators H, $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ that satisfy the well-known commutation relations (C.R.), i.e., the Lie bracket relations, for the IHLG:

$$[P_i, P_j] = 0, \quad [P_i, H] = 0, \quad [J_i, H] = 0, \quad [J_i, J_j] = i\epsilon_{ijk}J_k, \quad [J_i, P_j] = i\epsilon_{ijk}P_k, \tag{2.3a}$$

and

$$[J_i, K_j] = i\epsilon_{ijk}K_k, \quad [H, K_j] = -iP_j, \quad [K_i, K_j] = -i\epsilon_{ijk}J_k, \quad [P_i, K_j] = -i\delta_{ij}H, \tag{2.3b}$$

where we have put $\hbar = c = 1$. One can then identify H, $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ as the infinitesimal generators of time translations, space translations, space rotations, and pure Lorentz transformations, respectively.

3. RELATIVISTIC CENTER-OF-MOMENTUM VARIABLES

Let $\mathbf{p}_i$ denote the three-vector momentum variable associated with a particle of mass m_i $(i = 1, 2)$ in a momentum-space representation. We define four-vectors p_i by

$$p_i = (p_i^\mu) = (p_i^0, \mathbf{p}_i) = (E_i(\mathbf{p}_i), p_i^1, p_i^2, p_i^3),$$

where

$$E_i(\mathbf{p}_i) = (m_i^2 + \mathbf{p}_i^2)^{\frac{1}{2}},$$

and generally denote the result of transforming p_i with a Lorentz matrix $L = (L^\mu{}_\nu)$ by $\bar{p}_i$:

$$\bar{p}_i = Lp_i = (\bar{E}_i, \bar{\mathbf{p}}_i). \tag{3.1a}$$

We denote the three-vector part $\bar{\mathbf{p}}_i$ by

$$\bar{\mathbf{p}}_i = \mathbf{L}(\mathbf{p}_i). \tag{3.1b}$$

For a pure Lorentz transformation with velocity $\mathbf{v}$, we write $L = L(\mathbf{v})$ and have, explicitly,

$$\bar{\mathbf{p}}_i = \mathbf{L}_\mathbf{v}(\mathbf{p}_i) \equiv \mathbf{p}_i + (\gamma - 1)(\mathbf{p}_i \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \gamma E_i \mathbf{v}, \qquad \bar{E}_i = \gamma(E_i - \mathbf{p}_i \cdot \mathbf{v}), \tag{3.2}$$

where $\gamma = (1 - \mathbf{v}^2)^{-\frac{1}{2}}$.

[10] E. P. Wigner, Ann. Math. 40, 149 (1939); lectures in "Group Theoretical Concepts and Methods in Elementary Particle Physics" (Proc. Symp., Istanbul 1962), edited by F. Gürsey (Gordon and Breach, New York, to be published).

It will prove convenient to introduce new variables $\mathbf{P}$ and $\mathbf{k}$ via

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2, \tag{3.3a}$$

$$\mathbf{k} = (\epsilon_2\mathbf{p}_1 - \epsilon_1\mathbf{p}_2)/(\epsilon_1 + \epsilon_2), \tag{3.3b}$$

where

$$\epsilon_i = \tfrac{1}{2}(E_i + w_i),$$

and

$$w_i = w_i(\mathbf{k}) = (m_i^2 + \mathbf{k}^2)^{\frac{1}{2}}. \tag{3.4}$$

$\mathbf{k}$ is simply the three-momentum of particle 1 in the "instantaneous" c.m. system and the form Eq. (3.3b) for $\mathbf{k}$ is obtained, after some manipulation, by choosing $\mathbf{v} = \mathbf{u}$ in Eq. (3.2), where

$$\mathbf{u} = \mathbf{u}(\mathbf{p}_1, \mathbf{p}_2) = (\mathbf{p}_1 + \mathbf{p}_2)/(E_1 + E_2).$$

Thus,

$$\mathbf{k} = \mathbf{L}_\mathbf{u}(\mathbf{p}_1) = -\mathbf{L}_\mathbf{u}(\mathbf{p}_2), \tag{3.5}$$

and

$$(w_1(\mathbf{k}), \mathbf{k}) = L(\mathbf{u})p_1,$$
$$(w_2(\mathbf{k}), -\mathbf{k}) = L(\mathbf{u})p_2.$$

The magnitude of $\mathbf{k}^2$ is determined directly from

$$w_1(\mathbf{k}) + w_2(\mathbf{k}) = (P^2)^{\frac{1}{2}}, \tag{3.6}$$

where

$$P = p_1 + p_2 = (E, \mathbf{P}) = (E_1 + E_2, \mathbf{p}_1 + \mathbf{p}_2).$$

The Jacobian of the *nonlinear* transformation from the variables $\mathbf{p}_1$, $\mathbf{p}_2$ to the variables $\mathbf{P}$, $\mathbf{k}$ may be shown, after some tedious calculation, to have the form

$$J(\mathbf{p}_1, \mathbf{p}_2) = \frac{\partial(\mathbf{p}_1, \mathbf{p}_2)}{\partial(\mathbf{k}, \mathbf{P})} = \frac{E_1E_2}{E_1 + E_2}\frac{w_1 + w_2}{w_1w_2}. \tag{3.7}$$

Let $\mathbf{p}_1^{\mathrm{op}}$, $\mathbf{p}_2^{\mathrm{op}}$ denote the (abstract) momentum operators associated with the particles, acting in a Hilbert space $\mathfrak{H} = \{\Psi\}$ which is spanned by states $|\mathbf{q}_1, \mathbf{q}_2\rangle$ ("product plane waves"), defined by

$$\mathbf{p}_i^{\mathrm{op}}\,|\mathbf{q}_1, \mathbf{q}_2\rangle = \mathbf{q}_i\,|\mathbf{q}_1, \mathbf{q}_2\rangle \qquad (i = 1, 2), \tag{3.8}$$

and normalized so that

$$\langle \mathbf{q}_1, \mathbf{q}_2 \mid \mathbf{q}_1', \mathbf{q}_2'\rangle = \delta(\mathbf{q}_1 - \mathbf{q}_1')\delta(\mathbf{q}_2 - \mathbf{q}_2'), \qquad \int d\mathbf{q}_1\, d\mathbf{q}_2\, |\mathbf{q}_1, \mathbf{q}_2\rangle\langle\mathbf{q}_1, \mathbf{q}_2| = 1. \tag{3.9}$$

The operators $\mathbf{k}^{\mathrm{op}}$ and $\mathbf{P}^{\mathrm{op}}$ are defined through Eqs. (3.3a, b) by replacing $\mathbf{p}_i$ by $\mathbf{p}_i^{\mathrm{op}}$ and $\mathbf{P}$ by $\mathbf{P}^{\mathrm{op}}$, $\mathbf{k}$ by $\mathbf{k}^{\mathrm{op}}$ in these equations. Since

$$[\mathbf{p}_1^{op}, \mathbf{p}_2^{op}] = 0,$$

it follows that

$$[\mathbf{k}^{op}, \mathbf{P}^{op}] = 0,$$

so that $\mathfrak{H}$ can be spanned by simultaneous eigenstates $|\mathbf{l}, \mathbf{Q}\rangle$ of $\mathbf{k}^{op}$ and $\mathbf{P}^{op}$, defined by

$$\begin{aligned} \mathbf{k}^{op} |\mathbf{l}, \mathbf{Q}\rangle &= \mathbf{l} |\mathbf{l}, \mathbf{Q}\rangle, \\ \mathbf{P}^{op} |\mathbf{l}, \mathbf{Q}\rangle &= \mathbf{Q} |\mathbf{l}, \mathbf{Q}\rangle, \end{aligned} \tag{3.10}$$

and normalized so that

$$\begin{aligned} \langle \mathbf{l}, \mathbf{Q} \mid \mathbf{l}', \mathbf{Q}' \rangle &= \delta(\mathbf{l} - \mathbf{l}')\delta(\mathbf{Q} - \mathbf{Q}') \\ \int d\mathbf{l}\, d\mathbf{Q}\, |\mathbf{l}, \mathbf{Q}\rangle\langle \mathbf{l}, \mathbf{Q}| &= 1. \end{aligned} \tag{3.11}$$

Since

$$\begin{aligned} \mathbf{k}^{op} |\mathbf{q}_1, \mathbf{q}_2\rangle &= \mathbf{k}(\mathbf{q}_1, \mathbf{q}_2) |\mathbf{q}_1, \mathbf{q}_2\rangle, \\ \mathbf{P}^{op} |\mathbf{q}_1, \mathbf{q}_2\rangle &= (\mathbf{q}_1 + \mathbf{q}_2) |\mathbf{q}_1, \mathbf{q}_2\rangle, \end{aligned}$$

where $\mathbf{k}(\mathbf{q}_1, \mathbf{q}_2)$ is given by Eq. (3.3b) with $\mathbf{p}_i \to \mathbf{q}_i$, we may take

$$|\mathbf{l}, \mathbf{Q}\rangle = N |\mathbf{q}_1, \mathbf{q}_2\rangle,$$

where

$$\begin{aligned} \mathbf{l} &= \mathbf{k}(\mathbf{q}_1, \mathbf{q}_2), \\ \mathbf{Q} &= \mathbf{q}_1 + \mathbf{q}_2, \end{aligned}$$

and N is determined from Eqs. (3.9) and (3.11). Thus,

$$NN'\delta(\mathbf{q}_1 - \mathbf{q}_1')\delta(\mathbf{q}_2 - \mathbf{q}_2') = \delta(\mathbf{l} - \mathbf{l}')\delta(\mathbf{Q} - \mathbf{Q}'),$$

which yields

$$N^2 = \partial(\mathbf{q}_1, \mathbf{q}_2)/\partial(\mathbf{l}, \mathbf{Q}),$$

so that

$$|\mathbf{l}, \mathbf{Q}\rangle = J^{\frac{1}{2}}(\mathbf{q}_1, \mathbf{q}_2) |\mathbf{q}_1, \mathbf{q}_2\rangle. \tag{3.12}$$

It follows that the wavefunctions $\psi(\mathbf{p}_1, \mathbf{p}_2)$ and $\phi(\mathbf{k}, \mathbf{P})$ of a state Ψ,

$$\begin{aligned} \psi(\mathbf{p}_1, \mathbf{p}_2) &= \langle \mathbf{p}_1, \mathbf{p}_2 \mid \Psi \rangle, \\ \phi(\mathbf{k}, \mathbf{P}) &= \langle \mathbf{k}, \mathbf{P} \mid \Psi \rangle, \end{aligned} \tag{3.13}$$

are related by

$$\psi(\mathbf{p}_1, \mathbf{p}_2) = J^{-\frac{1}{2}}(\mathbf{p}_1, \mathbf{p}_2)\phi(\mathbf{k}(\mathbf{p}_1, \mathbf{p}_2), \mathbf{p}_1 + \mathbf{p}_2). \tag{3.14}$$

Furthermore, if A is an operator in $\mathfrak{H}$, the matrices which represent A in the two different bases are related by

$$\begin{aligned} &\langle \mathbf{p}_1, \mathbf{p}_2| A |\mathbf{p}_1', \mathbf{p}_2'\rangle \\ &= J^{-\frac{1}{2}}(\mathbf{p}_1, \mathbf{p}_2)\langle \mathbf{k}, \mathbf{P}| A |\mathbf{k}', \mathbf{P}'\rangle J^{-\frac{1}{2}}(\mathbf{p}_1', \mathbf{p}_2'). \end{aligned} \tag{3.15}$$

Eqs. (3.14) and (3.15) apply in particular to cases where $\phi = \phi_a$ and $\psi = \psi_a$ are representatives of a state $\Psi = \Psi_a$ which is an eigenfunction of some operator A with eigenvalue a. ϕ_a and ψ_a are then eigenstates of the corresponding matrices

$$\langle \mathbf{k}, \mathbf{P}| A |\mathbf{k}', \mathbf{P}'\rangle \quad \text{and} \quad \langle \mathbf{p}_1, \mathbf{p}_2| A |\mathbf{p}_1', \mathbf{p}_2'\rangle$$

with the same eigenvalue a.

The factor $J^{-\frac{1}{2}}(\mathbf{p}_1, \mathbf{p}_2)$ is *not* simply a normalization constant, since it varies when $\mathbf{p}_1$ and $\mathbf{p}_2$ vary. However, in the particular case where

$$\Psi = |\mathbf{q}_1, \mathbf{q}_2\rangle = J^{-\frac{1}{2}}(\mathbf{q}_1, \mathbf{q}_2) |\mathbf{l}, \mathbf{Q}\rangle,$$

so that

$$\begin{aligned} \psi &\to \delta(\mathbf{p}_1 - \mathbf{q}_1)\delta(\mathbf{p}_2 - \mathbf{q}_2), \\ \phi &\to J^{-\frac{1}{2}}(\mathbf{q}_1, \mathbf{q}_2)\delta(\mathbf{k} - \mathbf{l})\delta(\mathbf{P} - \mathbf{Q}), \end{aligned}$$

the factor $J^{-\frac{1}{2}}(\mathbf{p}_1, \mathbf{p}_2)$ in Eq. (3.14) may be replaced by $J^{-\frac{1}{2}}(\mathbf{q}_1, \mathbf{q}_2)$, which is a constant, and the identity

$$\begin{aligned} &\delta(\mathbf{p}_1 - \mathbf{q}_1)\delta(\mathbf{p}_2 - \mathbf{q}_2) \\ &\quad = J^{-1}(\mathbf{q}_1, \mathbf{q}_2)\delta(\mathbf{k} - \mathbf{l})\delta(\mathbf{P} - \mathbf{Q}) \end{aligned} \tag{3.16}$$

is recovered.

We also note here the well-known identity

$$E_i(\bar{\mathbf{q}}_i)\delta(\bar{\mathbf{q}}_i - \bar{\mathbf{q}}_i') = E_i(\mathbf{q}_i)\delta(\mathbf{q}_i - \mathbf{q}_i'). \tag{3.17}$$

Finally, we define, for later use, operators $\boldsymbol{\varrho}^{op}$ and $\mathbf{R}^{op}$ which are cononically conjugate to $\mathbf{k}^{op}$ and $\mathbf{P}^{op}$,

$$[k_i^{op}, \rho_j^{op}] = -i\delta_{ij}, \quad [P_i^{op}, R_j^{op}] = -i\delta_{ij}. \tag{3.18}$$

Thus, in a representation in which $\mathbf{k}$ and $\mathbf{P}$ are independent variables, we have

$$\boldsymbol{\varrho}^{op} \to i\,\partial/\partial\mathbf{k}, \qquad \mathbf{R}^{op} \to i\,\partial/\partial\mathbf{P}. \tag{3.19}$$

Similarly, if $\mathbf{r}_i^{op}$ is conjugate to $\mathbf{p}_i^{op}$, so that with $\mathbf{p}_1$ and $\mathbf{p}_2$ as independent variables

$$\mathbf{r}_i^{op} \to i\,\partial/\partial\mathbf{p}_i, \tag{3.20}$$

it may be shown that the angular momentum operator, defined by

$$\mathbf{J} = \mathbf{r}_1^{op} \times \mathbf{p}_1^{op} + \mathbf{r}_2^{op} \times \mathbf{p}_2^{op}, \tag{3.21}$$

may be written in the form[5]

$$\mathbf{J} = \mathbf{R}^{op} \times \mathbf{P}^{op} + \mathbf{l}^{op}, \tag{3.22}$$

where

$$\mathbf{l}^{op} = \boldsymbol{\varrho}^{op} \times \mathbf{k}^{op}. \tag{3.23}$$

Thus, the same decomposition of $\mathbf{J}$ holds as in the nonrelativistic case, where the c.m. momentum variables are defined by

$$\begin{aligned} \mathbf{k}_{n.r.} &= (m_2\mathbf{p}_1 - m_1\mathbf{p}_2)/(m_1 + m_2), \\ \mathbf{P} &= \mathbf{p}_1 + \mathbf{p}_2. \end{aligned} \tag{3.24}$$

4. DIRECT-INTERACTION THEORIES AND THE IHLG

For a system of two spinless, noninteracting particles of mass m_1 and m_2, the relevant representation of the IHLG is simply obtained by adding the well-known generators of the IHLG for each of the particles considered separately. Thus, using a subscript "0" to distinguish operators referring to the free system, we take

$$H_0 = \sum_{i=1}^{2} H_i, \quad \mathbf{P}_0 = \sum_{i=1}^{2} \mathbf{p}_i^{op},$$
$$\mathbf{J}_0 = \sum_{i=1}^{2} \mathbf{r}_i^{op} \times \mathbf{p}_i^{op}, \quad \mathbf{K}_0 = \sum_{i=1}^{2} \tfrac{1}{2}(\mathbf{r}_i^{op} H_i + H_i \mathbf{r}_i^{op}), \tag{4.1}$$

where

$$H_i = [m_i^2 + (\mathbf{p}_i^{op})^2]^{\frac{1}{2}}.$$

To introduce interaction one lets $H_0 \to H$, where

$$H = H_0 + V, \tag{4.2}$$

with V a linear Hermitian operator. One must now supplement H with operators $\mathbf{P}$, $\mathbf{J}$, and $\mathbf{K}$ such that the C.R. of the IHLG [Eqs. (2.3a, b)] are satisfied. If, as in the "instant form" of dynamics,[1] one also demands

$$\mathbf{P} = \mathbf{P}_0, \qquad \mathbf{J} = \mathbf{J}_0, \tag{4.3}$$

then the C.R. imply that

$$[V, \mathbf{P}] = 0, \qquad [V, \mathbf{J}] = 0, \tag{4.4}$$

corresponding to conservation of linear and angular momenta for the interacting system. Conversely, if we require that Eqs. (4.3) and (4.4) hold, then the parts of the C.R., Eq. (2.3a), which involve only H, $\mathbf{P}$, and $\mathbf{J}$ will be automatically satisfied, and only those parts [Eq. (2.3b)] which involve $\mathbf{K}$ need further attention.

Now it might be thought that it is in the further attempt to find a $\mathbf{K}$ such that the remaining C.R. [Eq. (2.3b)] are satisfied that the problem of constructing relativistic direct-interaction theories hinges, and that the requirement that such a $\mathbf{K}$ can be found will yield restrictions on V, since some of these C.R. involve H and hence V. We wish to emphasize here that this is not the case, and that a $\mathbf{K}$ can be found for almost any V which has no bound states. In fact, we have

Theorem 1. If H has complete sets of "in" and "out" scattering states, then there exists at least one choice of $\mathbf{K}$ such that H, $\mathbf{P}_0$, $\mathbf{J}_0$, and $\mathbf{K}$ satisfy the C.R. of the IHLG.

Proof: The hypothesis on V means that there exists for each (product) plane-wave eigenstate $|\mathbf{q}_1, \mathbf{q}_2\rangle$ of H_0, an eigenstate $\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}$ of H which behaves asymptotically, in coordinate space, as a plane wave plus relative outgoing or ingoing spherical waves. Although not needed here, we note for later use the formal relation

$$\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} = |\mathbf{q}_1, \mathbf{q}_2\rangle + (E_1 + E_2 - H \pm i\epsilon)^{-1} V |\mathbf{q}_1, \mathbf{q}_2\rangle, \tag{4.5}$$

where $E_i = (m_i^2 + \mathbf{q}_i^2)^{\frac{1}{2}}$ and $\epsilon \to 0_+$. We now define

$$\Omega_\pm = \int d\mathbf{q}_1\, d\mathbf{q}_2\, |\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}\rangle\langle \mathbf{q}_1, \mathbf{q}_2|. \tag{4.6}$$

Since the "in" states $\{\Psi^{(+)}\}$ and "out" states $\{\Psi^{(-)}\}$ separately form complete sets, the operators $\Omega_\pm$ are unitary,

$$\Omega_\pm^\dagger \Omega_\pm = \Omega_\pm \Omega_\pm^\dagger = 1. \tag{4.7}$$

The $\Omega_\pm$ correspond simply to the unitary transformation from the set of plane-wave states to the set of "in" or "out" states, since

$$\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} = \Omega_\pm |\mathbf{q}_1, \mathbf{q}_2\rangle. \tag{4.8}$$

From Eqs. (4.3) and (4.4) it follows that $\mathbf{P}_0$ and $\mathbf{J}_0$ commute with $\Omega_\pm$ so that we may write

$$\mathbf{P}_0 = \Omega_\pm \mathbf{P}_0 \Omega_\pm^{-1}, \qquad \mathbf{J}_0 = \Omega_\pm \mathbf{J}_0 \Omega_\pm^{-1}. \tag{4.9}$$

From Eq. (4.6), it is seen that

$$H\Omega_\pm = \Omega_\pm H_0,$$

so that

$$H = \Omega_\pm H_0 \Omega_\pm^{-1}. \tag{4.10}$$

We now *define* operators $\mathbf{K}_\pm$ by

$$\mathbf{K}_\pm = \Omega_\pm \mathbf{K}_0 \Omega_\pm^{-1}. \tag{4.11}$$

The set of operators $\mathbf{P}_0$, $\mathbf{J}_0$, H, and $\mathbf{K}_+$ are then similarity transforms by the *same* operator Ω_+ of the ten operators $\mathbf{P}_0$, $\mathbf{J}_0$, H_0, and $\mathbf{K}_0$. Hence they satisfy the C.R. of the IHLG. Since Ω_+ is unitary, $\mathbf{K}_+$ is Hermitian, along with $\mathbf{P}_0$, $\mathbf{J}_0$ and H. The same statements hold for the set $\mathbf{P}_0$, $\mathbf{J}_0$, H, and $\mathbf{K}_-$. This completes the proof. (The extension of this theorem to include bound states is made in Sec. 8.)

More generally, if A is a unitary operator which commutes with $\mathbf{P}_0$, $\mathbf{J}_0$, and H_0, we may define

$$\mathbf{K}_{A\pm} = \Omega_\pm A \mathbf{K}_0 (\Omega_\pm A)^{-1}, \tag{4.12}$$

and note that $\mathbf{P}_0$, $\mathbf{J}_0$, H, and $\mathbf{K}_{A\pm}$ satisfy the C.R., since $\mathbf{P}_0$, $\mathbf{J}_0$, and H are also similarity transforms by $\Omega_\pm A$. Thus, an enormous variety of operators $\mathbf{K}$ may be adjoined to $\mathbf{P}_0$, $\mathbf{J}_0$, and almost any reasonable

$H(= H_0 + V)$ to yield a set of 10 generators for a representation of the IHLG.

Both the lack of any restrictions on V and the ambiguity in the choice of $\mathbf{K}$ are connected with the fact that the requirement that one have a representation of the Lie algebra of the IHLG is *not* sufficient to ensure that the theory is relativistic from the viewpoint of scattering theory. As we shall see in the next section, it is this latter requirement that both restricts the possible forms of V and reduces the ambiguity in $\mathbf{K}$, in fact determines $\mathbf{K}$ uniquely and in such a way that, in the absence of bound states,

$$\mathbf{K} = \mathbf{K}_+ = \mathbf{K}_-.$$

5. ASYMPTOTIC COVARIANCE

Consider a two-particle system with Hamiltonian $H = H_0 + V$. We assume that $[\mathbf{P}_0, V] = [\mathbf{J}_0, V] = 0$ and that H possesses complete sets of both "in" and "out" scattering states $\{\Psi^{(+)}_{\mathbf{q}_1,\mathbf{q}_2}\}$, $\{\Psi^{(-)}_{\mathbf{q}_1,\mathbf{q}_2}\}$, where the $\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}$ satisfy Eq. (4.5) and so are normalized according to

$$\langle \Psi^{(\pm)}_{\mathbf{q}_1',\mathbf{q}_2'} \mid \Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} \rangle = \delta(\mathbf{q}_1' - \mathbf{q}_1)\delta(\mathbf{q}_2' - \mathbf{q}_2). \tag{5.1}$$

We now introduce the concept of asymptotic covariance by asking whether, with each Lorentz matrix L, it is possible to associate a linear mapping (L) of the Hilbert space onto itself,

$$\Psi \xrightarrow{(L)} \Psi',$$

in such a way that (L) maps any (improper) eigenstate $\Psi_{\mathbf{q}_1,\mathbf{q}_2}$ of H, which consists asymptotically of a plane wave $|\mathbf{q}_1, \mathbf{q}_2\rangle$ plus scattered waves, into a state which consists asymptotically of a plane wave $|\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2\rangle$ plus similar scattered waves. Here $\bar{\mathbf{q}}_i$ is the Lorentz transform of $\mathbf{q}_i$ [see Eq. (3.1b)],

$$\bar{\mathbf{q}}_i = \mathbf{L}(\mathbf{q}_i) \qquad (i = 1, 2).$$

Thus, in particular, we require that

$$\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} \xrightarrow{(L)} \Psi^{(\pm)}_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2}. \tag{5.2}$$

The physical interpretation of these states in terms of the time development of the corresponding wave packets[11] leads one to expect that, in relativistic theory, such a mapping will exist. If the mapping (L) exists, then one may define an operator $W(L)$, such that

$$W(L)\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} = N_L(\mathbf{q}_1, \mathbf{q}_2)\Psi^{(\pm)}_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2}, \tag{5.3}$$

and choose the factor N_L so that $W(L)$ will be unitary. The normalization of states implied by Eq. (5.1), and the identity of Eq. (3.17) then yields

$$N_L(\mathbf{q}_1, \mathbf{q}_2) = [E_1(\bar{\mathbf{q}}_1)E_2(\bar{\mathbf{q}}_2)]^{\frac{1}{2}}[E_1(\mathbf{q}_1)E_2(\mathbf{q}_2)]^{-\frac{1}{2}}. \tag{5.4}$$

The requirement that for each L a $W(L)$ exists satisfying Eq. (5.3) will be referred to as the requirement of asymptotic covariance. If the $\{W(L)\}$ exist, then it follows readily from Eq. (5.3) that they constitute a unitary representation of the homogeneous Lorentz group (HLG), i.e., they satisfy

$$W(L_1)W(L_2) = W(L_1L_2).$$

The existence of the $\{W(L)\}$ is not simply a matter of definition since the "in" and "out" states are determined, independently, by the Hamiltonian H and a boundary condition. To exhibit the question of the existence of W in the clearest possible way, we note that, with each Lorentz matrix L, we may associate unitary operators $W_\pm(L)$ according to the *definitions*

$$\begin{aligned} W_+(L)\Psi^{(+)}_{\mathbf{q}_1,\mathbf{q}_2} &= N_L(\mathbf{q}_1, \mathbf{q}_2)\Psi^{(+)}_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2}, \\ W_-(L)\Psi^{(-)}_{\mathbf{q}_1,\mathbf{q}_2} &= N_L(\mathbf{q}_1, \mathbf{q}_2)\Psi^{(-)}_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2}. \end{aligned} \tag{5.5}$$

Since the "in" and "out" states are complete sets, Eq. (5.5) determines the operators $W_\pm(L)$ completely. The operator sets $\{W_+(L)\}$ and $\{W_-(L)\}$ then separately provide a unitary representation of the HLG.

On comparing Eq. (5.5), which are definitions, with Eq. (5.3), which is the mathematical formulation of asymptotic covariance, one arrives at

Theorem 2. Asymptotic covariance holds if and only if

$$W_+(L) = W_-(L). \tag{5.6}$$

For if Eq. (5.6) is satisfied, we may define $W(L)$ as

$$W(L) = W_+(L) = W_-(L),$$

and if Eq. (5.6) is not satisfied, there is no $W(L)$ which satisfies Eq. (5.3).

To proceed further, we first define the direct-product representation $\{U(L)\}$ of the HLG associated with the noninteracting two-particle system, with generators $\mathbf{J}_0$, $\mathbf{K}_0$, via

$$U(L)\,|\mathbf{q}_1, \mathbf{q}_2\rangle = N_L(\mathbf{q}_1, \mathbf{q}_2)\,|\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2\rangle. \tag{5.7}$$

From Eqs. (4.8), (5.5), and (5.7), we then get

$$W_\pm(L)\Omega_\pm = \Omega_\pm U(L).$$

Thus, in the assumed absence of bound states,

$$W_\pm(L) = \Omega_\pm U(L)\Omega_\pm^\dagger. \tag{5.8}$$

If L is a pure spatial rotation R, then since

[11] G. C. Wick, Rev. Mod. Phys. 27, 339 (1955).

$[\mathbf{J}_0, \Omega_\pm] = 0$, Eq. (5.8) reduces to

$$W_\pm(R) = U(R),$$

so that Eq. (5.6) is trivially satisfied in this case.

On the other hand, if L is a pure Lorentz transformation, Eq. (5.6) is satisfied if and only if

$$\mathbf{K}_+ = \mathbf{K}_-, \tag{5.9}$$

where the $\mathbf{K}_\pm$ are the generators of the $W_\pm(L)$ and are related to $\mathbf{K}_0$ by Eq. (4.11). But Eq. (5.9) implies, on multiplication on the left by $\Omega_-^\dagger$ and on the right by Ω_+, that

$$[S^{op}, \mathbf{K}_0] = 0, \tag{5.10}$$

where S^{op}, defined by

$$S^{op} = \Omega_-^\dagger \Omega_+, \tag{5.11}$$

is precisely the operator which, when taken between plane-wave states, yields the S matrix,

$$S(\mathbf{q}_1', \mathbf{q}_2'; \mathbf{q}_1, \mathbf{q}_2) = \langle \mathbf{q}_1', \mathbf{q}_2' | S^{op} | \mathbf{q}_1, \mathbf{q}_2 \rangle. \tag{5.12}$$

Eq. (5.10) or, more directly, the corresponding relation for the finite transformation,

$$S^{op} = U^\dagger(L) S^{op} U(L), \tag{5.13}$$

implies that the S matrix is covariant. Conversely, Eq. (1.1) implies that Eq. (5.13) holds, so that $W(L)$ may be defined as $W_\pm(L)$. Then, we have

Theorem 3. The requirements that:

(i) asymptotic covariance hold,
(ii) $\mathbf{K}_+ = \mathbf{K}_-$,
(iii) the S matrix is covariant,

are all equivalent.

We reserve further discussion of the requirement of asymptotic covariance to Sec. 9 and proceed to an examination of a model of direct-interaction theories which satisfies this requirement.

6. THE BAKAMJIAN-THOMAS MODEL

In this section we examine a class of direct-interaction theories from the viewpoint of scattering theory.

A direct interaction theory satisfying the C.R. of the IHLG was proposed some time ago by Bakamjian and Thomas,[5] who stipulated a Hamiltonian H of the form

$$H = [h^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}}, \tag{6.1}$$

where

$$h = h(\mathbf{k}^{op}, \boldsymbol{\varrho}^{op}) \tag{6.2}$$

is a rotationally invariant operator function of $\mathbf{k}^{op}$ and $\boldsymbol{\varrho}^{op}$ only. The associated $\mathbf{K}$ was taken to have the form

$$\mathbf{K} = \tfrac{1}{2}(\mathbf{R}^{op} H + H\mathbf{R}^{op}) - (\mathbf{l}^{op} \times \mathbf{P}^{op})(h + H)^{-1}. \tag{6.3}$$

Here $\boldsymbol{\varrho}^{op}$ and $\mathbf{R}^{op}$ are conjugate to $\mathbf{k}^{op}$ and $\mathbf{P}^{op}$ [see Eqs. (3.18) and (3.19)], and $\mathbf{l}^{op}$ is the internal angular momentum operator [Eq. (3.23)]. The C.R. are then satisfied with $\mathbf{P}^{op} = \mathbf{P}_0$ and $\mathbf{J}^{op} = \mathbf{J}_0$.

If one takes h to have the form

$$h = h_0 + v(\mathbf{k}^{op}, \boldsymbol{\varrho}^{op}), \tag{6.4}$$

where

$$h_0 = [m_1^2 + (\mathbf{k}^{op})^2]^{\frac{1}{2}} + [m_2^2 + (\mathbf{k}^{op})^2]^{\frac{1}{2}}, \tag{6.5}$$

and assumes that v decreases sufficiently rapidly at large distances, the scattering problem is well defined. For then H may be written in the form

$$H = H_0 + V, \tag{6.6}$$

where

$$H_0 = H|_{v=0}, \tag{6.7}$$

and the interaction V, given by

$$V = H - H_0, \tag{6.8}$$

vanishes rapidly at large distances.[12] Moreover, it follows from Eqs. (3.4), (3.6), and (6.5), that H_0, defined by Eq. (6.7), coincides with the H_0 of Eq. (4.1), i.e.,

$$H_0 = [h_0^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}} = H_1 + H_2, \tag{6.9}$$

so that that the interpretation of V as the interaction is justified.

We can now state

Theorem 4. If H has the form (6.1), then the associated S matrix is covariant.

Proof: Let $\phi_{\mathbf{l}}^{(\pm)}(\mathbf{k})$ denote the momentum-space wavefunctions of the "in" and "out" continuum eigenstates of the operator h [defined by Eqs. (6.4), (6.5)], with asymptotic relative momentum $\mathbf{l}$. Thus, $\phi_{\mathbf{l}}^{(\pm)}$ is determined by

$$\phi_{\mathbf{l}}^{(\pm)}(\mathbf{k}) = \delta(\mathbf{k} - \mathbf{l}) + [w(\mathbf{l}) - w(\mathbf{k}) \pm i\epsilon]^{-1} \times \int v(\mathbf{k}, \mathbf{k}') \phi_{\mathbf{l}}^{(\pm)}(\mathbf{k}')\, d\mathbf{k}', \tag{6.10}$$

where

$$w(\mathbf{l}) = (m_1^2 + \mathbf{l}^2)^{\frac{1}{2}} + (m_2^2 + \mathbf{l}^2)^{\frac{1}{2}}, \tag{6.11}$$

[12] For nonlocal V, we mean by "vanishing at large distances" that $\langle \chi | V | \chi \rangle \to 0$ as $|\mathbf{r}_2^c - \mathbf{r}_1^c| \to \infty$ where χ denotes a product of single-particle wave packets centered about $\mathbf{r}_1^c$ and $\mathbf{r}_2^c$.

and

$$v(\mathbf{k}, \mathbf{k}') = \langle \mathbf{k} \,|v(\mathbf{k}^{op}, \varrho^{op})|\, \mathbf{k}' \rangle. \tag{6.12}$$

One then obtains, in the standard manner,

$$\langle \phi_{\mathbf{l}'}^{(-)} \mid \phi_{\mathbf{l}}^{(+)} \rangle = \delta(\mathbf{l}' - \mathbf{l}) - 2\pi i \delta(w' - w) t_{\mathbf{l}',\mathbf{l}}, \tag{6.13}$$

where $w' = w(\mathbf{l}')$ and

$$t_{\mathbf{l}',\mathbf{l}} = \langle \mathbf{l}' \,|v|\, \phi_{\mathbf{l}}^{(+)} \rangle = \langle \mathbf{l}' \,|v + v[w - h + i\epsilon]^{-1} v|\, \mathbf{l} \rangle \,|_{w'=w}. \tag{6.14}$$

It follows from the rotational invariance of v that $t_{\mathbf{l}',\mathbf{l}}$ may be regarded as a function of w and $\cos\theta = \hat{\mathbf{l}}' \cdot \hat{\mathbf{l}}$ only:

$$t_{\mathbf{l}',\mathbf{l}} = t(w, \cos\theta).$$

Now let $\mathbf{p}_1$ and $\mathbf{p}_2$ denote variables related to $\mathbf{k}$ and $\mathbf{P}$ according to Eq. (3.3), and let $\mathbf{q}_1$ and $\mathbf{q}_2$, eigenvalues of $\mathbf{p}_1^{op}$ and $\mathbf{p}_2^{op}$, be similarly related to $\mathbf{l}$ and $\mathbf{Q}$. We define functions $\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}$ of $\mathbf{p}_1$ and $\mathbf{p}_2$ by

$$\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2) = [J(\mathbf{q}_1, \mathbf{q}_2) \times J(\mathbf{p}_1, \mathbf{p}_2)]^{-\frac{1}{2}} \phi_{\mathbf{l}}^{(\pm)}(\mathbf{k}) \delta(\mathbf{P} - \mathbf{Q}), \tag{6.15}$$

and regard these as the momentum space representatives of the (improper) states $\Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}$ in Hilbert space, according to

$$\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2) = \langle \mathbf{p}_1, \mathbf{p}_2 \mid \Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)} \rangle. \tag{6.16}$$

Now the functions $\phi_{\mathbf{l}}^{(\pm)}(\mathbf{k})\delta(\mathbf{P} - \mathbf{Q})$ are eigenfunctions of the matrix $\langle \mathbf{k}, \mathbf{P}|\, H\, |\mathbf{k}', \mathbf{P}' \rangle$ with eigenvalue E,

$$E = [w^2(\mathbf{l}) + \mathbf{Q}^2]^{\frac{1}{2}},$$

since

$$\langle \mathbf{k}, \mathbf{P}\, |H|\, \mathbf{k}', \mathbf{P}' \rangle = \langle \mathbf{k}\, |[h^2(\mathbf{k}^{op}, \varrho^{op}) + \mathbf{P}^2]^{\frac{1}{2}}|\, \mathbf{k}' \rangle \delta(\mathbf{P} - \mathbf{P}'),$$

and Eq. (6.10) implies that $\phi_{\mathbf{l}}^{(\pm)}(\mathbf{k})$ are eigenfunctions of the matrix $\langle \mathbf{k}|\, h(\mathbf{k}^{op}, \varrho^{op})\, |\mathbf{k}' \rangle$ with eigenvalue $w(\mathbf{l})$.

It follows from Eq. (3.15) and the remark thereafter that $\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2)$ is an eigenfunction of the matrix $\langle \mathbf{p}_1, \mathbf{p}_2|\, H\, |\mathbf{p}_1', \mathbf{p}_2' \rangle$ with the same eigenvalue. Thus, also using Eq. (3.6), we get

$$H\Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)} = (E_1(\mathbf{q}_1) + E_2(\mathbf{q}_2))\Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}. \tag{6.17}$$

Substituting Eq. (6.10) into Eq. (6.15) we find, using Eq. (3.16), that

$$\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2) = \delta(\mathbf{p}_1 - \mathbf{q}_1)\delta(\mathbf{p}_2 - \mathbf{q}_2) + \chi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2), \tag{6.18a}$$

where

$$\chi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}(\mathbf{p}_1, \mathbf{p}_2) = [J(\mathbf{q}_1, \mathbf{q}_2)J(\mathbf{p}_1, \mathbf{p}_2)]^{-\frac{1}{2}} \times \delta(\mathbf{P} - \mathbf{Q})[w(\mathbf{l}) - w(\mathbf{k}) \pm i\epsilon]^{-1} v \phi_{\mathbf{l}}^{(\pm)}(\mathbf{k}). \tag{6.18b}$$

Equations (6.17) and (6.18) justify the interpretation of $\Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)}$ as the usual "in" and "out" scattering states. Moreover, it follows from Eqs. (6.10) and (6.15) that the $\Psi^{(\pm)}$ are appropriately normalized,

$$\langle \Psi_{\mathbf{q}_1',\mathbf{q}_2'}^{(\pm)} \mid \Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(\pm)} \rangle = \delta(\mathbf{q}_1' - \mathbf{q}_1)\delta(\mathbf{q}_2' - \mathbf{q}_2).$$

The S matrix is accordingly determined by

$$S = S(\mathbf{q}_1', \mathbf{q}_2'; \mathbf{q}_1, \mathbf{q}_2) = \langle \Psi_{\mathbf{q}_1',\mathbf{q}_2'}^{(-)} \mid \Psi_{\mathbf{q}_1,\mathbf{q}_2}^{(+)} \rangle = \int \psi_{\mathbf{q}_1',\mathbf{q}_2'}^{(-)*}(\mathbf{p}_1, \mathbf{p}_2)\psi_{\mathbf{q}_1,\mathbf{q}_2}^{(+)}(\mathbf{p}_1, \mathbf{p}_2)\, d\mathbf{p}_1\, d\mathbf{p}_2. \tag{6.19}$$

Substitution of Eq. (6.15), followed by change of variables to $\mathbf{k}$, $\mathbf{P}$, yields

$$S = [J(\mathbf{q}_1', \mathbf{q}_2')J(\mathbf{q}_1, \mathbf{q}_2)]^{-\frac{1}{2}} \delta(\mathbf{Q}' - \mathbf{Q}) \langle \phi_{\mathbf{l}'}^{(-)} \mid \phi_{\mathbf{l}}^{(+)} \rangle$$

or, using Eq. (6.13),

$$S = [J(\mathbf{q}_1', \mathbf{q}_2')J(\mathbf{q}_1, \mathbf{q}_2)]^{-\frac{1}{2}} \delta(\mathbf{Q}' - \mathbf{Q}) \times [\delta(\mathbf{l}' - \mathbf{l}) - 2\pi i \delta(w' - w) t_{\mathbf{l}',\mathbf{l}}].$$

Use of Eq. (3.16) (with $\mathbf{p}_i \to \mathbf{q}_i'$) and the identity

$$\delta(\mathbf{Q}' - \mathbf{Q})\delta(w' - w) = [(w' + w)/(E' + E)]\delta(\mathbf{Q}' - \mathbf{Q})\delta(E' - E)$$

then yields

$$S = \delta(\mathbf{q}_1' - \mathbf{q}_1)\delta(\mathbf{q}_2' - \mathbf{q}_2) - 2\pi i \delta(Q' - Q)\Gamma t_{\mathbf{l}',\mathbf{l}},$$

where

$$\Gamma = [J(\mathbf{q}_1', \mathbf{q}_2')J(\mathbf{q}_1, \mathbf{q}_2)]^{-\frac{1}{2}}(w' + w)(E' + E)^{-1}.$$

Γ is to be evaluated on the energy-momentum shell, $Q' = Q$, or $E' = E$, $\mathbf{Q}' = \mathbf{Q}$. Using Eq. (3.7), we get

$$\Gamma = \left(\frac{E_1'E_2'}{E'} \cdot \frac{w'}{w_1'w_2'} \cdot \frac{E_1E_2}{E} \cdot \frac{w}{w_1w_2}\right)^{-\frac{1}{2}} \frac{w' + w}{E' + E}.$$

Since $w' = w \Rightarrow |\mathbf{l}'| = |\mathbf{l}| \Rightarrow w_i' = w_i$, we get

$$\Gamma = (E_1'E_2'E_1E_2)^{-\frac{1}{2}}(w_1w_2),$$

so that

$$(E_1'E_2')^{\frac{1}{2}} S (E_1E_2)^{\frac{1}{2}} = [(E_1')^{\frac{1}{2}}\delta(\mathbf{q}_1' - \mathbf{q}_1)(E_1)^{\frac{1}{2}}] \times [(E_2')^{\frac{1}{2}}\delta(\mathbf{q}_2' - \mathbf{q}_2)(E_2)^{\frac{1}{2}}] - 2\pi i \delta(Q' - Q)F, \tag{6.20}$$

where

$$F = (w_1w_2)t_{\mathbf{l}',\mathbf{l}} = w_1w_2 t(w, \cos\theta) \tag{6.21}$$

is an invariant function of the momenta $\mathbf{q}_i$, $\mathbf{q}_i'$. Since the factors $(E_i')^{\frac{1}{2}}\delta(\mathbf{q}_i' - \mathbf{q}_i)(E_i)^{\frac{1}{2}}$ are also Lorentz invariant functions [Eq. (3.17)], the covariance of S is established.

Now we consider the form for $\mathbf{K}$ [Eq. (6.3)] given by Bakamjian and Thomas, and show that it is the "right" $\mathbf{K}$ in the sense of Sec. 5. For simplicity

we assume that h [Eq. (6.4)] has no bound states. Since the S matrix has already been shown to be covariant, it follows from Sec. 5 that the correct choice for $\mathbf{K}$ is to take [see Eq. (4.11)]

$$\mathbf{K} = \mathbf{K}_+ = \Omega_+ \mathbf{K}_0 \Omega_+^\dagger,$$

where Ω_+ is defined by Eq. (4.6). $\mathbf{K}_0$ is defined by Eq. (4.1). An alternative form for $\mathbf{K}_0$ is[5,13]

$$\mathbf{K}_0 = \tfrac{1}{2}[\mathbf{R}^{op}, H_0]_+ - (\mathbf{l}^{op} \times \mathbf{P}^{op})(h_0 + H_0)^{-1}. \quad (6.22)$$

It may be shown[13] that Ω_+ commutes with $\mathbf{P}^{op}$, $\mathbf{R}^{op}$, and $\mathbf{J}^{op}$, and hence with $\mathbf{l}^{op} = \mathbf{J}^{op} - \mathbf{R}^{op} \times \mathbf{P}^{op}$, so that

$$\mathbf{K}_+ = \Omega_+ \mathbf{K}_0 \Omega_+^\dagger$$
$$= \tfrac{1}{2}[\mathbf{R}^{op}, H]_+ - (\mathbf{l}^{op} \times \mathbf{P}^{op})(h + H)^{-1}, \quad (6.23)$$

where we have used Eq. (4.10) and the relation

$$\Omega_+ h_0 \Omega_+^\dagger = h.$$

The last equation follows from the operator identities

$$h_0 = (H_0^2 - \mathbf{P}^2)^{\frac{1}{2}}, \qquad h = (H^2 - \mathbf{P}^2)^{\frac{1}{2}}.$$

Comparison of Eq. (6.3) and Eq. (6.23) shows that $\mathbf{K} = \mathbf{K}_+$ as asserted.

We note that, in the computation of S, no assumption concerning the existence of bound states was made. Having shown that Eq. (6.1) for H yields a truly covariant theory, we turn to an investigation of the generality of this form.

7. GENERAL FORM OF A RELATIVISTIC HAMILTONIAN

Consider a 2-particle Hamiltonian $H = H_0 + V$, where V commutes with $\mathbf{P}_0$ and $\mathbf{J}_0$, and assume that the scattering problem is well defined. Then scattering states $\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}$ exist which satisfy

$$\mathbf{P}^{op}\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2} = (\mathbf{q}_1 + \mathbf{q}_2)\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}, \quad (7.1)$$

and are eigenfunctions of H with eigenvalue $E_1(\mathbf{q}_1) + E_2(\mathbf{q}_2)$, i.e., they satisfy Eq. (6.17). It follows from Eq. (7.1) that the representatives $\psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}(\mathbf{p}_1, \mathbf{p}_2)$ of $\Psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}$ may be taken to have the form

$$\psi^{(\pm)}_{\mathbf{q}_1,\mathbf{q}_2}(\mathbf{p}_1, \mathbf{p}_2) = \delta(\mathbf{P} - \mathbf{Q})[J(\mathbf{p}_1, \mathbf{p}_2) \times J(\mathbf{q}_1, \mathbf{q}_2)]^{-\frac{1}{2}} U_\pm(\mathbf{k}, \mathbf{l}; \mathbf{Q}), \quad (7.2)$$

without loss of generality. Here $\mathbf{P}$, $\mathbf{k}$, and $\mathbf{Q}$, $\mathbf{l}$ are related to $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{q}_1$, $\mathbf{q}_2$, respectively, via Eq. (3.3), as in Sec. 6, and the functions $U_\pm(\mathbf{k}, \mathbf{l}; \mathbf{Q})$ may be taken as rotationally invariant functions of the variables $\mathbf{k}$, $\mathbf{l}$ and $\mathbf{Q}$.

[13] R. Fong, Ph.D. thesis, University of Maryland, July, 1963. (Available as Tech. Rept. No. 322.)

On inserting Eq. (7.2) into the Schrödinger equation and introducing $\mathbf{k}$ and $\mathbf{P}$ as independent variables one finds, on imposing the usual boundary conditions and the normalization of Eq. (5.1), that the $U_\pm$ must satisfy, on the one hand,

$$U_\pm(\mathbf{k}, \mathbf{l}; \mathbf{Q}) = \delta(\mathbf{k} - \mathbf{l}) + [E(\mathbf{l}, \mathbf{Q}) - E(\mathbf{k}, \mathbf{Q}) \pm i\epsilon]^{-1} \times \int V(\mathbf{k}, \mathbf{k}'; \mathbf{Q}) U_\pm(\mathbf{k}', \mathbf{l}; \mathbf{Q})\, d\mathbf{k}', \quad (7.3)$$

where $V(\mathbf{k}, \mathbf{k}'; \mathbf{Q})$ is defined by

$$\langle \mathbf{k}, \mathbf{P} \,|V|\, \mathbf{k}', \mathbf{P}' \rangle = \delta(\mathbf{P} - \mathbf{P}') V(\mathbf{k}, \mathbf{k}'; \mathbf{P}), \quad (7.4)$$

and $E(\mathbf{l}, \mathbf{Q})$ by

$$E(\mathbf{l}, \mathbf{Q}) = [w^2(\mathbf{l}) + \mathbf{Q}^2]^{\frac{1}{2}}. \quad (7.5)$$

It follows from Eq. (7.3) that the scalar product

$$\langle U_- \mid U_+ \rangle = \int U_-^*(\mathbf{k}, \mathbf{l}'; \mathbf{Q}) U_+(\mathbf{k}, \mathbf{l}; \mathbf{Q})\, d\mathbf{k}$$

is given by

$$\langle U_- \mid U_+ \rangle = \delta(\mathbf{l}' - \mathbf{l}) - 2\pi i \delta[E(\mathbf{l}', \mathbf{Q}) - E(\mathbf{l}, \mathbf{Q})] \times \int V(\mathbf{l}', \mathbf{k}; \mathbf{Q})\, U_+(\mathbf{k}, \mathbf{l}; \mathbf{Q})\, d\mathbf{k}. \quad (7.6)$$

On the other hand, the S matrix is now determined with the help of Eqs. (6.19), (7.2), and (7.6) to be given by

$$S(\mathbf{q}_1', \mathbf{q}_2'; \mathbf{q}_1, \mathbf{q}_2) = \delta(\mathbf{q}_1' - \mathbf{q}_1)\delta(\mathbf{q}_2' - \mathbf{q}_2) - 2\pi i \delta(Q' - Q)[E_1' E_2' E_1 E_2]^{-\frac{1}{2}} \times [w_1 w_2/(w_1 + w_2)]\mathfrak{F}(\mathbf{l}', \mathbf{l}; \mathbf{Q}), \quad (7.7)$$

where $\mathfrak{F}$ is defined for all $\mathbf{l}'$, $\mathbf{l}$, and $\mathbf{Q}$ by

$$\mathfrak{F}(\mathbf{l}', \mathbf{l}; \mathbf{Q}) = -[E(\mathbf{l}', \mathbf{Q}) + E(\mathbf{l}, \mathbf{Q})] \times \int V(\mathbf{l}', \mathbf{l}''; \mathbf{Q}) U_+(\mathbf{l}'', \mathbf{l}; \mathbf{Q})\, d\mathbf{l}''. \quad (7.8)$$

Using Eq. (7.3) we may write $\mathfrak{F}$ in the form

$$\mathfrak{F}(\mathbf{l}', \mathbf{l}; \mathbf{Q}) = -[E(\mathbf{l}', \mathbf{Q}) + E(\mathbf{l}, \mathbf{Q})] \times [E(\mathbf{l}', \mathbf{Q}) - E(\mathbf{l}, \mathbf{Q})] U_+(\mathbf{l}', \mathbf{l}; \mathbf{Q})$$
$$= -[w^2(\mathbf{l}') - w^2(\mathbf{l})] U_+(\mathbf{l}', \mathbf{l}; \mathbf{Q}). \quad (7.9)$$

Since $\mathfrak{F}$ must be invariant on the energy-momentum shell $Q' = Q$, we obtain the result:

Theorem 5. A necessary and sufficient condition that the S matrix is covariant is that, with $\mathbf{l}'$ and $\mathbf{l}$ fixed,

$$\lim_{|\mathbf{l}'| \to |\mathbf{l}|} (\mathbf{l}'^2 - \mathbf{l}^2) U_+(\mathbf{l}', \mathbf{l}; \mathbf{Q}) \quad (7.10)$$

is independent of $\mathbf{Q}$.

The simplest way of satisfying this condition is to have $U_+(\mathbf{k}, 1; \mathbf{Q})$ independent of $\mathbf{Q}$, i.e., $U_+ = U_+(\mathbf{k}, 1)$ only. This is precisely the case for the case for the Bakamjian–Thomas (B–T) form of H, for comparison of Eqs. (6.15) and (7.2) shows that the $U_+(\mathbf{k}, 1; \mathbf{Q})$ coincides with $\phi_1^{(+)}(\mathbf{k})$, and, *a fortiori*, is independent of $\mathbf{Q}$.

Conversely, we may attempt to find the general form of V, or equivalently H, consistent with $U_+ = U_+(\mathbf{k}, 1)$ only.

We note that if the $\phi_1^{(+)}(\mathbf{k})$ are *given* and if v has no bound states, then H may be constructed from H_0 by a transformation U, defined by

$$\langle \mathbf{k}, \mathbf{P} \,|U|\, \mathbf{k}', \mathbf{P}' \rangle = \delta(\mathbf{P} - \mathbf{P}')\phi_{\mathbf{k}'}^{(+)}(\mathbf{k}). \tag{7.11}$$

U is unitary, since the matrix

$$\mathfrak{U}(\mathbf{k}, \mathbf{k}') = \phi_{\mathbf{k}'}^{(+)}(\mathbf{k}) \tag{7.12}$$

is unitary, as a result of the completeness and orthogonality of the $\phi_{\mathbf{k}'}^{(+)}(\mathbf{k})$. It then follows, that

$$H = UH_0U^{-1}. \tag{7.13}$$

More generally, given any unitary matrix $\mathfrak{U}(\mathbf{k}, \mathbf{k}')$, a Hamiltonian H which will have a covariant S matrix may be constructed, via Eq. (7.13) and the generalized definition

$$\langle \mathbf{k}, \mathbf{P} \,|U|\, \mathbf{k}', \mathbf{P}' \rangle = \delta(\mathbf{P} - \mathbf{P}')\mathfrak{U}(\mathbf{k}, \mathbf{k}'),$$

provided that H actually possesses eigenstates which may be interpreted as scattering states. This will be the case only if $\mathfrak{U}(\mathbf{k}, \mathbf{k}')$ has the appropriate singular behavior at $\mathbf{k}' = \mathbf{k}$. The simplest way to ensure such behavior is to generate $\mathfrak{U}$ as the solution of an equation analogous to that satisfied by $\phi_{\mathbf{k}'}^{(+)}(\mathbf{k})$ [Eq. (6.10) with $\mathbf{l} \to \mathbf{k}'$],

$$\mathfrak{U}(\mathbf{k}, \mathbf{k}') = \delta(\mathbf{k} - \mathbf{k}') + [w_c(\mathbf{k}') - w_c(\mathbf{k}) + i\epsilon]^{-1} \times \int v_c(\mathbf{k}, \mathbf{k}'')\mathfrak{U}(\mathbf{k}'', \mathbf{k}')\, d\mathbf{k}'', \tag{7.14}$$

where $w_c(\mathbf{k})$ is a monotonic function of $\mathbf{k}^2$ and $v_c(\mathbf{k}, \mathbf{k}')$ is a suitably well-behaved potential matrix. The resulting H is of the B–T type, since

$$U\mathbf{P}^{op}U^{-1} = \mathbf{P}^{op},$$

so that

$$H = U[h_0^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}}U^{-1} = [h^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}}, \tag{7.15}$$

with

$$h = h(\mathbf{k}^{op}, \boldsymbol{\varrho}^{op}) = Uh_0U^{-1}. \tag{7.16}$$

This method of constructing relativistic Hamiltonians has already been used by Sudarshan.[6] The above considerations show that the class of Hamiltonians obtained in this way is contained in the B–T class. The latter class is actually larger, since, as we shall see, the B–T form admits the existence of bound states, whereas an H of the form of Eq. (7.13) cannot have bound states, since for unitary U, H and H_0 have the same spectrum.[14]

We also see that for any pair $\{w_c(\mathbf{k}), v_c(\mathbf{k}, \mathbf{k}')\}$ which generates a unitary $\mathfrak{U}(\mathbf{k}, \mathbf{k}')$ via Eq. (7.14), the associated pair $\{w(\mathbf{k}), v(\mathbf{k}, \mathbf{k}')\}$, with $v = h - h_0$ is such that the resulting $\phi_{\mathbf{k}'}^{(+)}(\mathbf{k})$ coincides with $\mathfrak{U}(\mathbf{k}, \mathbf{k}')$. Thus, all unitary matrices $\mathfrak{U}$ generated by Eq. (7.14) with w_c and v_c arbitrary may also be generated by Eq. (6.10) with v arbitrary.

We conclude this section with the remark that if a two-body covariant S matrix has a "potential" origin, i.e., is the S matrix of some direct-interaction theory, $\bar{H} = H_0 + \bar{V}$, with H_0 defined by Eq. (4.1), then there always exists a B–T type of Hamiltonian which yields the same S matrix. As proof, we note that if $\bar{V}(\mathbf{k}, \mathbf{k}'; \mathbf{P})$ [see Eq. (7.4)] yields a covariant S matrix, we may *define* a function

$$\bar{v}(\mathbf{k}, \mathbf{k}') = \bar{V}(\mathbf{k}, \mathbf{k}'; \mathbf{0}),$$

and correspondingly an operator $\bar{v}(\mathbf{k}^{op}, \boldsymbol{\varrho}^{op})$, such that

$$\langle \mathbf{k} \,|\bar{v}(\mathbf{k}^{op}, \boldsymbol{\varrho}^{op})|\, \mathbf{k}' \rangle = \bar{v}(\mathbf{k}, \mathbf{k}').$$

Then

$$H = [(h_0 + \bar{v})^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}}$$

is a B–T Hamiltonian whose S matrix coincides with the given S matrix (although, in general, $\bar{H} \neq H$). This follows from Eq. (7.10) and the observation that $\phi_1^{(+)}(\mathbf{k})$ will coincide with $\bar{U}_+(\mathbf{k}, \mathbf{l}; \mathbf{0})$.

Thus, we may assert that the B–T type of Hamiltonian is, in a sense, the most general form of

---

[14] Let H' denote an operator of the form $[h'^2 + (\mathbf{P}^{op})^2]^{1/2}$, with $h' = w' + v'$, $w' = w(\mathbf{k}_{n.r.}{}^{op})$, $v' = v'(\mathbf{k}_{n.r.}, \boldsymbol{\rho}_{n.r.}{}^{op})$, where $\mathbf{k}_{n.r.}{}^{op}$ is defined by Eq. (3.24). Such Hamiltonians have been considered by Foldy.[3] If the operators $\mathbf{p}_1{}^{op}$, $\mathbf{p}_2{}^{op}$ have their usual physical interpretation, the choice $v' = 0$ does not correspond to the absence of interaction since the free Hamiltonian H_0 [see Eq. (4.1)] is not recovered thereby. Furthermore, on expressing the nonrelativistic operator $\mathbf{k}_{n.r.}{}^{op}$ in terms of $\mathbf{k}^{op}$ and $\mathbf{P}^{op}$, it may be seen that for no choice of v' is H' of the B–T form. In general, this form of Hamiltonian can lead to a covariant S matrix only if a reinterpretation[3] of the $\mathbf{p}_i{}^{op}$ is made. In this paper the $\mathbf{p}_i{}^{op}$ always have their original meaning. From the point of view of scattering theory the identification of a simultaneous eigenstate $|\mathbf{q}_1, \mathbf{q}_2\rangle$ of $\mathbf{p}_1{}^{op}$, $\mathbf{p}_2{}^{op}$ as describing noninteracting particles with *kinetic* momenta $\mathbf{q}_1$, $\mathbf{q}_2$ is permissible even in the presence of nonlocal interactions $\bar{V}$, when $\dot{\mathbf{r}}_i{}^{op} = i[H, \mathbf{r}_i{}^{op}] = \mathbf{v}_i{}^{op} + i[V, \mathbf{r}_i{}^{op}] \neq \mathbf{v}_i{}^{op}$ (with $\mathbf{v}_i{}^{op} = \mathbf{p}_i{}^{op}/H_i$). The reason is that if a product wave packet is constructed from the $|\mathbf{q}_1, \mathbf{q}_2\rangle$, then in the limit of infinite separation of the individual packets the expectation values of $\dot{\mathbf{r}}_i{}^{op}$ and $\mathbf{v}_i{}^{op}$ will coincide, provided that V effectively vanishes at large distances. If $\mathbf{r}_i{}^{op}$ is replaced by $\bar{\mathbf{r}}_i{}^{op}$, the Newton–Wigner position operator, and nonlocality of V is defined by $[\bar{\mathbf{r}}_i{}^{op}, V] \neq 0$, the same statement holds.

interest from the viewpoint of covariant scattering theory.[15]

8. BOUND STATES AND DIRECT-INTERACTION THEORIES

In Sec. 5, it was seen that a necessary and sufficient condition for a Hamiltonian H to yield a covariant S matrix is that

$$\Omega_+\mathbf{K}_0\Omega_+^\dagger = \Omega_-\mathbf{K}_0\Omega_-^\dagger,$$

whether or not H has bound states. However, the operator

$$\mathbf{K}_c = \Omega_\pm\mathbf{K}_0\Omega_\pm^\dagger \tag{8.1}$$

may be identified with the $\mathbf{K}$ for the system only in the absence of bound states, for the orthogonality of the continuum and bound states implies that

$$\mathbf{K}_c\Psi_B = 0, \tag{8.2}$$

if Ψ_B is a bound state.

The generalization of the previous discussion to include bound states is straightforward if we assume that the eigenvalues $E_\alpha(\mathbf{Q})$ of H corresponding to the bound states $\Psi_{\alpha,\mathbf{Q}}$, with

$$\begin{aligned}\mathbf{P}^{op}\Psi_{\alpha,\mathbf{Q}} &= \mathbf{Q}\Psi_{\alpha,\mathbf{Q}},\\ H\Psi_{\alpha,\mathbf{Q}} &= E_\alpha(\mathbf{Q})\Psi_{\alpha,\mathbf{Q}},\end{aligned} \tag{8.3}$$

depend on $\mathbf{Q}$ in the form

$$E_\alpha(\mathbf{Q}) = (M_\alpha^2 + \mathbf{Q}^2)^{\frac{1}{2}}, \tag{8.4}$$

where M_α is identified as the mass of the state $\Psi_{\alpha,\mathbf{Q}}$. We can then require that, for a fixed α, the states $\Psi_{\alpha,\mathbf{Q}}$ transform under a Lorentz transformation in the same way as the states of a single particle of mass M_α and spin j_α, where j_α is the internal orbital angular momentum determined by

$$\mathbf{J}^2\Psi_{\alpha,0} = j_\alpha(j_\alpha + 1)\Psi_{\alpha,0}.$$

Considering only S states for simplicity, we thus require, for pure Lorentz transformations L,

$$W(L)\Psi_{\alpha,\mathbf{Q}} = N_\alpha(\mathbf{Q})\Psi_{\alpha,\bar{\mathbf{Q}}}, \tag{8.5}$$

where the normalization condition

$$\langle\Psi_{\alpha,\mathbf{Q}'} \mid \Psi_{\alpha,\mathbf{Q}}\rangle = \delta_{\alpha',\alpha}\delta(\mathbf{Q}' - \mathbf{Q}), \tag{8.6}$$

and the requirement that $W(L)$ be unitary determines the factor $N_\alpha(\mathbf{Q})$:

$$N_\alpha(\mathbf{Q}) = [E_\alpha(\bar{\mathbf{Q}})/E_\alpha(\mathbf{Q})]^{\frac{1}{2}}. \tag{8.7}$$

Here $\bar{\mathbf{Q}}$ is determined from Eq. (3.2) with $\mathbf{p} \to \mathbf{Q}$, $E_\mathbf{p} \to E_\alpha(\mathbf{Q})$.

Now let $\mathfrak{B}$ denote the subspace of states spanned by $\{\Psi_{\alpha,\mathbf{Q}}\}$, and let $\mathcal{P}$ denote the projection operator onto $\mathfrak{B}$, so that

$$\mathcal{P} = \sum_\alpha \int d\mathbf{Q}\ |\Psi_{\alpha,\mathbf{Q}}\rangle\langle\Psi_{\alpha,\mathbf{Q}}|. \tag{8.8}$$

Eq. (8.5) will be satisfied if $\mathbf{K}_b$, the reduction of the generator $\mathbf{K}$ to the subspace $\mathfrak{B}$,

$$\mathbf{K}_b = \mathbf{K}\mathcal{P}, \tag{8.9}$$

is defined by

$$\mathbf{K}_b = \tfrac{1}{2}(H_b\mathbf{R} + \mathbf{R}H_b), \tag{8.10}$$

where H_b is the reduction of H to $\mathfrak{B}$,

$$H_b = H\mathcal{P}. \tag{8.11}$$

We now take as the generators for the system the operators $\mathbf{P}_0$, $\mathbf{J}_0$, $\mathbf{K}$, and H, where $\mathbf{K}$ is defined by

$$\mathbf{K} = \mathbf{K}_c + \mathbf{K}_b. \tag{8.12}$$

These operators satisfy the C.R. of the IHLG, whether or not asymptotic covariance is satisfied for the continuum states. For on defining

$$\left\{\begin{aligned}&\mathbf{P}_c = \Omega\mathbf{P}_0\Omega^\dagger, \quad \mathbf{P}_b = \mathbf{P}_0\mathcal{P},\\ &\mathbf{J}_c = \Omega\mathbf{J}_0\Omega^\dagger, \quad \mathbf{J}_b = \mathbf{J}_0\mathcal{P}, \quad H_c = \Omega H_0\Omega^\dagger,\\ &\qquad [\Omega = \Omega_+ \text{ or } \Omega_-],\end{aligned}\right. \tag{8.13}$$

we see that the continuum operators $\mathbf{P}_c$, $\mathbf{J}_c$, $\mathbf{K}_c$, H_c, and the bound-state operators $\mathbf{P}_b$, $\mathbf{J}_b$, $\mathbf{K}_b$ and H_b satisfy the C.R. of the IHLG separately. For the continuum operators, this follows simply from the orthogonality relation $\Omega^\dagger\Omega = 1$(Ω need *not* be unitary). For the bound-state operators it follows from the definition of $\mathbf{K}_b$, [Eq. (8.10)], and the relations[16]

$$\begin{aligned}[\mathcal{P}, \mathbf{R}] = [\mathcal{P}, H] = [\mathcal{P}, \mathbf{P}_0]\\ = [\mathcal{P}, \mathbf{J}_0] = 0; \qquad \mathcal{P}^2 = \mathcal{P}.\end{aligned}$$

The assertion for the full operators $\mathbf{P}_0$, $\mathbf{J}_0$, $\mathbf{K}$, and H now follows from the definition of $\mathbf{K}$, Eq. (8.12), and the identities

$$\begin{aligned}\mathbf{P}_0 = \mathbf{P}_c + \mathbf{P}_b, \qquad \mathbf{J}_0 = \mathbf{J}_c + \mathbf{J}_b,\\ H = H_c + H_b,\end{aligned} \tag{8.14}$$

[15] It should be noted that there exist Hamiltonians which are not of the B–T form but yield a covariant S matrix. For example, if

$$H' = H + \int d\mathbf{Q} \sum_{\alpha,\beta} \Gamma_{\alpha,\beta}(\mathbf{Q})\ |\Psi_{\alpha,\mathbf{Q}}\rangle\langle\Psi_{\alpha,\mathbf{Q}}|,$$

where H is a B–T type of Hamiltonian with bound states $\Psi_{\alpha,\mathbf{Q}}$ and $\Gamma_{\alpha,\beta}(\mathbf{Q})$ is arbitrary, then H' is generally not of the B–T type, although it has the same covariant S matrix as H.

[16] The first relation is a consequence of Eq. (3.18) and the fact that $[\mathbf{P}, h] = 0$. (See Ref. 13.)

which are a simple consequence of the relations

$$[\Omega, \mathbf{P}_0] = [\Omega, \mathbf{J}_0] = 0,$$

and the completeness relation

$$\Omega\Omega^\dagger = 1 - \mathcal{P}. \tag{8.15}$$

The above results show, incidentally, that Theorem 1 may be extended to the case where H has bound states, provided that the bound-state energies have the form of Eq. (8.4).

Consider now the case where H is of the B–T form, so that Eq. (6.1) is satisfied, but Eq. (6.10), with $w(1)$ replaced by M_α, admits normalizable solutions $\phi_\alpha(\mathbf{k})$. Then H possesses bound states $\psi_{\alpha,\mathbf{Q}}$ with

$$\langle \mathbf{k}, \mathbf{P} \mid \Psi_{\alpha,\mathbf{Q}} \rangle = \delta(\mathbf{P} - \mathbf{Q})\phi_\alpha(\mathbf{k}), \tag{8.16}$$

and the eigenvalues have the form of Eq. (8.4). It follows from Eqs. (8.1) and (8.10) that $\mathbf{K}$ is given by

$$\mathbf{K} = \Omega\mathbf{K}_0\Omega^\dagger + (\tfrac{1}{2})[(H\mathcal{P})\mathbf{R} + \mathbf{R}(H\mathcal{P})]. \tag{8.17}$$

On comparing Eqs. (6.3) and (8.17) in the basis $\{\Psi^{(+)}_{\mathbf{q}_1,\mathbf{q}_2}; \Psi_{\alpha,\mathbf{Q}}\}$, in which $\mathbf{P}$ and H are diagonal, the expressions are seen to be identical, so that Eq. (6.3) for $\mathbf{K}$ remains correct when (S-wave) bound states are present.

The extension to bound states with $j_\alpha \neq 0$ is straightforward.

In concluding this section we note that the asymptotic covariance requirement is, in operator form,

$$W(L)\Omega_\pm = \Omega_\pm U(L),$$

so that multiplying by $\Omega_\pm^\dagger$ on the right, we get, on using Eqs. (8.14) and (8.15),

$$W(\mathfrak{L}) = \Omega_\pm U(\mathfrak{L})\Omega_\pm^\dagger + W(\mathfrak{L})\mathcal{P}, \tag{8.18a}$$

or

$$W(\mathfrak{L}) = W_c(\mathfrak{L}) + W_b(\mathfrak{L}), \tag{8.18b}$$

thereby defining $W_c(\mathfrak{L})$ and $W_b(\mathfrak{L})$.[17] This decomposition corresponds to that for the generators [Eqs. (8.12), (8.14)], and explicitly exhibits the reduction of the representation into a continuum part $W_c(\mathfrak{L})$, relevant for asymptotic covariance, and a bound-state part $W_b(\mathfrak{L})$ which, for the case where H has the B–T form, is the direct sum of irreducible representations of mass M_α and spin j_α. The continuum part $W_c(\mathfrak{L})$ may be said to be equivalent in an extended sense to the direct-product representation $U(\mathfrak{L})$, associated with the system of two noninteracting particles. [An extension of the usual definition of equivalence of representations is necessary, because the operators

$$W_c(\mathfrak{L}) = \Omega_\pm U(\mathfrak{L})\Omega_\pm^\dagger$$

represent the IHLG, even though the transformation by Ω_+ is not a unitary or even a similarity transformation. All that is needed is that $\Omega_+^\dagger$ be a left inverse of Ω_+, as is the case. This peculiarity is related to the infinite-dimensional nature of the corresponding representation in terms of matrices. For finite-dimensional matrices A, an equation of the form $A^\dagger A = 1$ implies, of itself, that $AA^\dagger = 1$ so that $A^\dagger = A^{-1}$ and A is unitary.] Eq. (8.18a), when written in the appropriate matrix form, shows that $\{W(\mathfrak{L})\}$ is unitarily equivalent to the direct sum of the direct-product representation $\{U(\mathfrak{L})\}$ and the irreducible representations of mass M_α and spin j_α.

[17] By replacing L by $\mathfrak{L} = \{a, L\}$ in Eq. (8.18), we have returned from the *HLG*, with generators $\mathbf{J}_0$ and $\mathbf{K}$, to the IHLG, with the additional generators $\mathbf{P}_0$ and H.

9. SUMMARY AND DISCUSSION

In part A of this section, the results obtained in Secs. 4 to 8 are summarized. In part B, relativistic potentials are considered from the viewpoint of perturbation theory and the notion of a "coupling constant" is examined in this light. (The details are given in the Appendix.) This is followed by part C which contains a brief comparison of the role of asymptotic covariance in direct-interaction theories and in field theory. Finally, in part D, some remarks are made on the extension of the results to N particles and to production processes.

A. Summary

In Sec. 4, it was seen that if a two-particle Hamiltonian H is invariant under spatial rotations and translations and has no bound states, it is always possible to define a variety of operators $\mathbf{K}$ such that H, $\mathbf{P}_0$, $\mathbf{J}_0$, and $\mathbf{K}$ satisfy the C.R. of the IHLG (Theorem 1). This is a consequence of the existence of unitary transformations which change H_0 to H, but leave $\mathbf{P}_0$ and $\mathbf{J}_0$ unchanged, and which can thus be used to define such $\mathbf{K}$. It follows that satisfying the C.R. is not sufficient to guarantee that the theory is relativistic from the point of view of scattering theory, since for arbitrary H the S matrix is, in general, not covariant.

In Sec. 5, the concept of asymptotic covariance was therefore introduced as an additional requirement for a direct-interaction theory to be relativistic. The operators $\{W(L)\}$ occurring in the definition

of asymptotic covariance were shown, in the absence of bound states, to be unique if they exist at all (Theorem 2). It was then shown that asymptotic covariance holds if and only if the S matrix is covariant and the ambiguity in the definition of $\mathbf{K}$ was clarified (Theorem 3).

In Sec. 6, the form of H introduced by Bakamjian and Thomas[5] was shown to yield a covariant S matrix (Theorem 4) and the expression for $\mathbf{K}$ given by these authors was seen to be "correct" from the viewpoint of asymptotic covariance.

In Sec. 7, the problem of specifying the general form of relativistic 2-particle Hamiltonians was studied and a simple condition, equivalent with the covariance of the S matrix, was obtained on the wavefunctions of the system (Theorem 5). The methods developed were used (i) to provide an alternative proof of the covariance of the S matrix for Hamiltonians of the B–T type, (ii) to show that a class of Hamiltonians considered by Sudarshan[6] is contained in the larger B–T class, and (iii) to prove that any covariant two-body S matrix which has a potential origin is also the S matrix of some potential of the B–T type. The last statement shows that this class of potentials is the most general of interest from the viewpoint of relativistic scattering theory.[15]

In Sec. 8, the discussion was extended to the case where H has bound states. It was shown that if the eigenvalues $E_\alpha(\mathbf{Q})$ of H for the bound states $\Psi_{\alpha,\mathbf{Q}}$ had the form $(M_\alpha^2 + \mathbf{Q}^2)^{\frac{1}{2}}$, then again a variety of $\mathbf{K}$ could be defined so that H, $\mathbf{P}_0$, $\mathbf{J}_0$, and $\mathbf{K}$ satisfy the C.R. of the IHLG. If asymptotic covariance is satisfied and bound states are present, each operator $W(L)$ may be decomposed into two parts; a part $W_c(L)$, acting only on the continuum states and uniquely determined by asymptotic covariance, and a part $W_b(L)$, acting only on the bound states. If $E_\alpha(\mathbf{Q}) = (M_\alpha^2 + \mathbf{Q}^2)^{\frac{1}{2}}$, $W_b(L)$ may be determined uniquely by the requirement that the bound states of mass M_α, spin j_α transform under Lorentz transformations like the states of a particle with the same mass and spin.[18] $\{W(\mathfrak{L})\}$ is then the direct sum of the reducible representation $\{W_c(\mathfrak{L})\}$, equivalent (in an extended sense) to the direct product $\{U(\mathfrak{L})\}$ which describes two non-interacting particles, and $\{W_b(\mathfrak{L})\}$, which is a direct sum of irreducible representations of the IHLG.

[18] Conversely, if the generators H_b, $\mathbf{P}_b$, $\mathbf{J}_b$ and $\mathbf{K}_b$ are required to satisfy the C.R. of the IHLG and Eq. (8.5) is to hold, then $E_\alpha(\mathbf{Q})$ must have the form of Eq. (8.4), since the C.R. imply that $H^2 - \mathbf{P}^2$ commutes with $W_b(\mathfrak{L})$.

B. Further Aspects of Relativistic Potentials

It should be noted that the set $\{V\}$ of potential operators of the B–T type is not a linear set, not being closed under addition. $\{V\}$ is not even closed under multiplication by a constant, so that the usual notion of a variable coupling constant, familiar from nonrelativistic quantum mechanics or quantum field theory, is not applicable. In particular, given a family of direct-interaction Hamiltonians, $H[\lambda] = H_0 + \lambda V$ with V fixed, the S matrix $S[\lambda]$ is in general covariant for, at most, isolated values of λ (including the value $\lambda = 1$) if V is of the B–T type. Another aspect of relativistic potentials V, i.e., which yield a covariant S matrix, is that, in general, it is not possible for the matrix element $\langle \mathbf{p}_1', \mathbf{p}_2' | V | \mathbf{p}_1, \mathbf{p}_2 \rangle$ to depend, except for kinematical factors and a factor $\delta(\mathbf{P}' - \mathbf{P})$, only on the c.m. relative momenta $\mathbf{k}'$ and $\mathbf{k}$. (This is to be contrasted with the case in nonrelativistic quantum mechanics when Galilean invariance is imposed.) Thus, it may be shown that if λV has this form and $S[\lambda]$ is analytic in a neighborhood $\mathfrak{N}$ of $\lambda = 0$, then $S[\lambda]$ cannot in general be covariant for $\lambda \in \mathfrak{N}$. More precise versions of these statements and sketches of their proofs are given in the Appendix.

C. Remarks on Asymptotic Covariance

Asymptotic covariance is usually implicit in other, perhaps more basic, assumptions. In quantum field theory it is unnecessary to stipulate separately equations such as Eqs. (1.1) or (1.2). The reason is that in quantum field theory there exist, by hypothesis, structures [namely the Heisenberg field operators] which are assumed from the outset to carry an irreducible representation of the IHLG, e.g.,

$$U(a, L)A(x)U^{-1}(a, L) = A(Lx + a),$$

for a single scalar field $A(x)$. This, together with the assumption of the asymptotic condition, relating $A(x)$ to the "in" and "out" fields $A_{\text{in}}(x)$ and $A_{\text{out}}(x)$, and the particle interpretation of A_{in} and A_{out}, is sufficient to guarantee, among other things, the covariance of the S matrix. In a Lagrangian field theory, the same result is guaranteed by related, familiar requirements. However, in the case of direct-interaction theories, there is no underlying quantum field available. If it is only required that a representation of the Lie algebra of the IHLG be exhibited, with $\mathbf{P} = \mathbf{P}_0$ and $\mathbf{J} = \mathbf{J}_0$, i.e., the same as in the absence of interaction, the S matrix need not be covariant.

In scattering theory, there is a fundamental concept that an observer can make a comparison

between the interacting systems and free systems. The latter are idealized from the actual systems under consideration and are made to correspond to direct products of irreducible representations of the IHLG. But the IHLG only incorporates the symmetry of the special theory of relativity and says nothing about how observers should make this comparison between the interacting systems and free systems. It therefore seems necessary to introduce explicitly an additional postulate.[19]

D. Extension to N Particles

The extension of Theorem 1 to N particles ($N \geq 3$) is trivial. The construction of N-particle Hamiltonians which satisfy the obvious generalization of the requirement of asymptotic covariance is also straightforward. It is also possible to construct Hamiltonians which describe any number of coupled two-body channels, or which describe production processes, e.g., a system consisting of a coupled two-body and three-body channel, and which satisfy asymptotic covariance.[13] However, the task of ensuring the "separability of the interaction" (see Ref. 3 for a discussion of this concept) is more difficult. For example, the sum of a two-particle Hamiltonian of the B–T type and the free Hamiltonian for a third particle is not expressible in the three-particle form indicated in Ref. 5. We hope to discuss these problems in a future paper.

A complete understanding of the multichannel, N-particle case and a study of the limit $N \to \infty$ would perhaps shed some light on the problem of the explicit construction of a nontrivial relativistic quantum field theory. In conclusion, we believe that this "middle ground" between nonrelativistic quantum mechanics and quantum field theory deserves further exploration, and so share the point of view so well expressed by Foldy.[3]

ACKNOWLEDGMENTS

We would both like to thank Professor P. A. M. Dirac for several stimulating discussions on this subject. One of us (J. S.) would also like to thank Professor R. Haag and Professor E. C. G. Sudarshan for conversations on some of these topics.

APPENDIX: RELATIVISTIC POTENTIALS, COUPLING CONSTANTS, AND PERTURBATION THEORY

Let $H = H[\lambda] = H_0 + \lambda V$ denote a two-particle Hamiltonian, with coupling constant λ, and let $S[\lambda]$ denote the corresponding S matrix. Assume that $S[\lambda]$ is analytic in λ for $|\lambda| < \lambda_m$, so that ordinary perturbation theory is applicable.

One obtains for the S matrix

$$S[\lambda] = S(\mathbf{q}_1', \mathbf{q}_2'; \mathbf{q}_1, \mathbf{q}_2; \lambda)$$
$$= \delta(\mathbf{q}_1' - \mathbf{q}_1)\delta(\mathbf{q}_2' - \mathbf{q}_2) + \sum_{n=1}^{\infty} S^{(n)}\lambda^n, \qquad (1)$$

with

$$S^{(1)} = (-2\pi i)\delta(E' - E)\langle \mathbf{q}_1', \mathbf{q}_2' \,|V|\, \mathbf{q}_1, \mathbf{q}_2\rangle,$$

$$S^{(2)} = (-2\pi i)\delta(E' - E) \int d\mathbf{q}_1'' \, d\mathbf{q}_2''$$
$$\times \frac{\langle \mathbf{q}_1', \mathbf{q}_2' \,|V|\, \mathbf{q}_1'', \mathbf{q}_2''\rangle\langle \mathbf{q}_1'', \mathbf{q}_2'' \,|V|\, \mathbf{q}_1, \mathbf{q}_2\rangle}{(E - E'' + i\epsilon)},$$

where $E'' = E_1(\mathbf{q}_1'') + E_2(\mathbf{q}_2'')$, etc. Now assume that $S[\lambda]$ is covariant for $-\lambda_m < \lambda < \lambda_m$. It follows, from the uniqueness of analytic continuation, that $S[\lambda]$ is covariant for all λ, real or complex, such that $|\lambda| < \lambda_m$. This in turn implies that each coefficient $S^{(n)}$ must be separately covariant, and $S^{(1)}$ in particular.

Use of Eqs. (3.7), (3.15), and (7.4) yields

$$S^{(1)} = (-2\pi i)\delta(Q' - Q)[E_1'E_2'E_1E_2]^{-\frac{1}{2}}F^{(1)}, \qquad (2)$$

where

$$F^{(1)} = [E(1', \mathrm{Q})E(1, \mathrm{Q})]^{\frac{1}{2}}V(1', 1; \mathrm{Q})\{\mu(1')\mu(1)\}^{\frac{1}{2}},$$

with

$$\mu(1) = w_1(1)w_2(1)/w(1),$$

must be invariant on the energy shell. [See Eqs. (6.11) and (7.5) for $w(1)$ and $E(1, \mathrm{Q})$.] The simplest possibility is to set

$$V(1', 1; \mathrm{Q}) = [E(1', \mathrm{Q})E(1, \mathrm{Q})]^{-\frac{1}{2}}\eta(1', 1), \qquad (3a)$$

with η independent of Q, as indicated, but otherwise an arbitrary rotational scalar. Eq. (3a) looks more natural in the $\mathbf{p}_1$, $\mathbf{p}_2$ representation, where it corresponds to the ansatz[20]

$$\langle \mathbf{p}_1', \mathbf{p}_2' \,|V|\, \mathbf{p}_1, \mathbf{p}_2\rangle$$
$$= \delta(\mathbf{P}' - \mathbf{P})(E_1'E_2')^{-\frac{1}{2}}\eta(\mathbf{k}', \mathbf{k})(E_1E_2)^{-\frac{1}{2}}, \qquad (3b)$$

which is the form referred to in Sec. 9: a product

[19] Some further heuristic remarks on asymptotic covariance are made in Ref. 13.

[20] In the language of second quantization, Eq. (3b) corresponds to an interaction Hamiltonian

$$H_1 \propto \int \phi_1^{(-)}(\mathbf{x})\phi_2^{(-)}(\mathbf{x})\phi_1^{(+)}(\mathbf{x})\phi_2^{(+)}(\mathbf{x}) \, d\mathbf{x},$$

if η = const., where $\phi_i^{(+)}$ and $\phi_i^{(-)}$ are the positive and negative frequency parts of the field operator $\phi_i(\mathbf{x})$ associated with the particle of mass m_i. If $\eta \neq$ const., the corresponding H_1 can be similarly expressed as a multiple integral involving also the Fourier transform of η as a form factor.

of kinematical factors and a function dependent only on c.m. momenta. However, this form leads to a noncovariant $S^{(2)}$, as we now show explicitly.

One obtains for $F^{(2)}$, related to $S^{(2)}$ in the same way as $F^{(1)}$ is related to $S^{(1)}$, the result

$$F^{(2)} = \int d1''(\mu(1'')E'')^{-1} \frac{\eta^*(1'', 1')\eta(1'', 1)}{(E - E'' + i\epsilon)},$$

where $E = E(1, \mathbf{Q})$ and $E'' = E(1'', \mathbf{Q})$.

Now $S^{(2)}$ will be covariant if and only if $F^{(2)}$ is independent of $\mathbf{Q}$ on the energy shell, $|1'| = |1|$. On separating $F^{(2)}$ into dispersive and absorptive parts, $F_D^{(2)}$ and $F_A^{(2)}$, by use of the relation

$$(x + i\epsilon)^{-1} = \mathcal{P}(1/x) - i\pi\delta(x),$$

it may be seen that $F_A^{(2)}$ is indeed independent of $\mathbf{Q}$, even off the energy shell, as a consequence of the identity

$$\delta(E - E'') = 2E''\delta(w^2(1) - w^2(1'')).$$

On the other hand, $F_D^{(2)}$ always depends on $\mathbf{Q}$, regardless of the choice of η. To prove this, it is only necessary to consider forward scattering, $1' = 1$. Then

$$F_D^{(2)} \to \mathcal{P}\int d1''(E - E'')^{-1}(E'')^{-1}\mu^{-1}(1'') \times |\eta(1'', 1)|^2. \quad (4)$$

The right-hand side of Eq. (4) may be regarded as a function of $\mathbf{Q}^2$ and 1^2, and its derivative with respect to $\mathbf{Q}^2$ is readily computed to be

$$-\int d1''\mu^{-1}(1'')(2E\,E''^3)^{-1}\,|\eta(1'', 1)|^2,$$

which is never zero if $\eta \not\equiv 0$. This completes the proof.

The method of proof shows, incidentally, that if V has the form of Eq. (3), then $S[\lambda]$ can be covariant for at most a finite number of points in any interval in the interior of the open interval $(-\lambda_m, \lambda_m)$. Further study indicates that in fact only the trivial choice $\lambda = 0$ is possible.

It is also instructive to consider potentials of the B–T type from the viewpoint of perturbation theory. If

$$V = [h^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}} - [h_0^2 + (\mathbf{P}^{op})^2]^{\frac{1}{2}}$$

[see Eqs. (6.1) and (6.9)], and V is replaced by λV, the first-order S matrix is $\lambda S^{(1)}$, with $S^{(1)}$ given by Eq. (2), where $F^{(1)}$ is now given by

$$F^{(1)} = (E'E)^{\frac{1}{2}}\langle 1'\,|V_{\mathbf{Q}}|\,1\rangle[\mu(1')\mu(1)]^{\frac{1}{2}},$$
$$V_{\mathbf{Q}} = [h^2 + \mathbf{Q}^2]^{\frac{1}{2}} - [h_0^2 + \mathbf{Q}^2]^{\frac{1}{2}}.$$

On setting $|1'| = |1|$, a short computation yields

$$dF^{(1)}/d\mathbf{Q}^2 = [\mu^{\frac{1}{2}}(1)]\langle 1'\,|(E - H_{\mathbf{Q}})^2(EH_{\mathbf{Q}})^{-1}|\,1\rangle,$$

where $H_{\mathbf{Q}} = (h^2 + \mathbf{Q}^2)^{\frac{1}{2}}$. On defining

$$|\chi_{1,\mathbf{Q}}\rangle = (E - H_{\mathbf{Q}})\,|1\rangle,$$

and on the insertion of a complete set $\{\phi_{\mathbf{k}'}^{(+)}\}$ of eigenstates of h, one obtains, for $1' = 1$,

$$\frac{dF^{(1)}}{d\mathbf{Q}^2} = [\tfrac{1}{2}\mu(1)]\int d\mathbf{k}'(E(1, Q)E(\mathbf{k}', \mathbf{Q}))^{-1} \times |\langle\chi_{1,\mathbf{Q}}\,|\,\phi_{\mathbf{k}'}^{(+)}\rangle|^2,$$

which is always positive if $v \neq 0$.

Thus $F^{(1)}$ is not an invariant, so that $S^{(1)}$ is not covariant. It follows that (i) $S[\lambda]$ is in general not covariant for $\lambda \neq 1$, and (ii) the partial sums $\sum_{n=1}^{N} S^{(n)}$ are not covariant for any N, although their limit is the covariant $S[1]$.

Note added in proof. It should be noted that although an N-particle system may satisfy asymptotic covariance, the momentum operators $\mathbf{p}_i^{op}$ ($i = 1, 2, \cdots N$) do not transform like Lorentz vectors under the group $\{W(L)\}$ associated with the interacting system. In fact, the equation

$$W^{-1}\mathbf{p}_i^{op}W = \mathbf{L}(\mathbf{p}_i^{op}) \quad (A)$$

can only hold if there is no interaction [so that $W(L) = U(L)$]. For, (A) implies that $\mathbf{p}_i^{op}$ commutes with $W(L)U^{-1}(L)$ and hence with $\mathbf{K} - \mathbf{K}_0$. Since $\mathbf{P} = \mathbf{P}_0 = \sum \mathbf{p}_i^{op}$, it follows that

$$0 = [P_i, K_i - K_{0i}] = -iH + iH_0 = -iV,$$

so that $V = 0$.

This result may be regarded as a quantum analog of a theorem of Currie [D. G. Currie, J. Math. Phys. 4, 1470 (1963)], which states that in classical relativistic Hamiltonian particle mechanics the coordinates of the points which compose the world line of a particle can transform according to the Lorentz transformation law only if there is no interaction.

Field Operators for Bosons with Impenetrable Cores. II. Equations of Motion and General Operator Formalism*

Ko Aizu†
Physics Department, Rikkyo University, Tokyo, Japan
Gianfausto Dell'Antonio†
Istituto di Fisica Teorica, Università di Napoli, Napoli, Italy
AND
Arnold J. F. Siegert‡
Department of Physics, Northwestern University, Evanston, Illinois
(Received 22 October 1963)

Continuing the investigation started in part I, we derive the equations of motion in the Fock representation and a nonrelativistic field theory for bosons interacting as hard spheres.

1. INTRODUCTION

We continue in the present paper the derivation and discussion of an exact nonrelativistic field theory valid for particles with impenetrable cores. This investigation was started by one of us in part I of this paper,[1] and the motivation for this investigation was given there. The field operators, represented as matrices in the Fock representation, were found to satisfy—instead of the standard commutation rules—another set of equations, explicitly stated for spherical cores in Eqs. (A1)–(A5) of I. The operators so defined still admit an interpretation as destruction and creation operators, and the density operator has the usual form. The integral of the density operator over any domain was shown to have the desired property that its eigenvalues are those nonnegative integers which do not exceed the largest number of hard spheres whose centers can be placed simultaneously in the domain.

Section 2 is a direct continuation of part I; the equations of motion for the new operators are derived in the Fock representation. (We consider hard-sphere interaction only; additional integrable interaction potentials would give rise to the conventional interaction term in the equations of motion). We obtain the equation of motion in two forms, which are equivalent as matrix equations in the special representation under certain regularity conditions. The first form is rederived in Sec. 5 and shown to be valid as an operator equation. The second form is an exact generalization of the equation of motion with pseudopotential, in the sense that the differential operator which replaces the potential becomes the pseudopotential asymptotically for small hard-sphere diameter. It should be emphasized, however, that this is not by itself a justification of the use of the pseudopotential, since the field operators also depend on the hard-sphere diameter through the new commutation rules. There are, in fact, some modifications of our equations of motion which are trivially equivalent for field operators with the new commutation rules, but would not be equivalent if the new commutation rules were "approximated" by the standard commutation rules.

It has not yet been possible to show the precise conditions under which the second equation of motion becomes valid also as an *operator* equation in the general formalism. We have shown, instead, in Sec. 3, that its range of validity is sufficient for applications to quantum statistical mechanics.

The Hamiltonian for the system is found to have the same form as the free-particle Hamiltonian written in terms of the new operators. This is of course to be expected, since the hard-sphere interaction is represented, as a constraint on the configuration space, in the new commutation relations. The interaction term in the equations of motion arises accordingly in forming the commutator of the destruction or creation operator with the free-particle Hamiltonian, due to the new commutation rules.

* Program supported by the National Science Foundation and the U. S. Office of Naval Research.
† The major part of this work was carried out while these authors were at the Physics Department, Northwestern University, as research associates under a National Science Foundation grant.
‡ This work was completed while the author was National Science Foundation Senior Post Doctoral Fellow, at the Instituut voor Theoretische Fysica, Universiteit van Amsterdam, Amsterdam, the Netherlands.
[1] A. J. F. Siegert, Phys. Rev. **116**, 1057 (1959) (hereafter referred to as I).

Sections 4 and 5 are intended to show that the scheme obtained in I and Sec. 2 is a legitimate field theory, and various properties of the new algebra are exhibited. It is also shown that the entire formalism can be constructed on a purely operator basis; this can be done by shifting the constraints characteristic of the hard-sphere problem from the states of the system to a set of operators which are used to describe it. The essential advantage in doing so lies in the fact that no unphysical state is ever introduced in the course of the analysis.

An attempt to extend the existing results for the hard-sphere Bose gas by investigating approximations to the general formalism presented here is being carried out by Professor L. Liu, and Professor K. W. Wong at Northwestern University.

2. EQUATIONS OF MOTION IN THE FOCK REPRESENTATION

Let $\mathbf{x}_i$ be a position vector in a three-dimensional cube Ω. Let q_n be the point $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ in the configuration space Ω^n, and let $q_n', q_n'', \cdots$, etc. be the configuration points with component vectors $\mathbf{x}_j', \mathbf{x}_j'', \cdots$, respectively. The volume element $\prod_{j=1}^n d^3x_j$ will be written as dq_n. The restricted configuration space Ω_R^n is defined as the set of all points q_n of Ω^n whose component vectors satisfy the inequalities $|\mathbf{x}_j - \mathbf{x}_k| > a > 0$ for all pairs (j, k).

Let $\Phi(q_n; \nu_n) \equiv \Phi(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_n; \nu_n)$ be symmetric functions in $\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_n$; and let the set of functions $\Phi(q_n; \nu_n)$, numbered by ν_n, be complex orthonormal, and complete in the space of the symmetric functions in Ω_R^n. Let $\Phi(q_n; \nu_n)$ be zero for any q_n not in Ω_R^n and satisfy periodicity conditions, for each of the component vectors, on the boundary of Ω.

The functions $\Phi(q_n; \nu_n)$ can then be expanded in terms of the functions $\Phi(q_m; \nu_m)$, when $m \leq n$, and $q_n \equiv (q_m, \mathbf{x}_{m+1}, \cdots \mathbf{x}_n)$ if q_m is in Ω^m, since the assumed completeness justifies the expansion in Ω_R^m and both functions vanish for q_m not in Ω_R^m.

The field operators are defined in a matrix representation which exhibits them as transformation matrices:

$$\langle \nu_n | \psi^\dagger(\mathbf{x}) | \nu_{N-1} \rangle = N^{-\frac{1}{2}} \int_{\Omega^{N-1}} \Phi^*(q_{N-1}, \mathbf{x}; \nu_N) \times \Phi(q_{N-1}; \nu_{N-1})\, dq_{N-1}. \quad (2.1)$$

(All operator equations in Secs. 2 and 3 of this paper are to be understood as a short notation for equations relating matrix elements.)

This definition is equivalent to that of I, Eq. (3), if

$$\langle q_N'' | \psi^\dagger(\mathbf{x}) | q_{N-1}' \rangle = \sum_{\nu_N, \nu_{N-1}} \Phi(q_N''; \nu_N) \times \langle \nu_N | \psi^\dagger(\mathbf{x}) | \nu_{N-1} \rangle \Phi^*(q_{N-1}'; \nu_{N-1}), \quad (2.2)$$

since

$$\sum_{\nu_N} \Phi(q_N''; \nu_N)\Phi^*(q_{N-1}, \mathbf{x}; \nu_N) = \begin{cases} \delta(q_N''; q_{N-1}, \mathbf{x}) & \text{for } q_N'' \text{ in } \Omega_R^N, \\ 0 & \text{otherwise,} \end{cases} \quad (2.3)$$

where $\delta(q_N''; q_N)$ is the symmetrized δ function defined by Eq. (2.18) below, and, therefore,

$$\langle q_N'' | \psi^\dagger(\mathbf{x}) | q_{N-1}' \rangle = N^{\frac{1}{2}} \int_{\Omega^{N-1}} \delta(q_N''; q_{N-1}, \mathbf{x}) \delta(q_{N-1}; q_{N-1}')\, dq_{N-1} = N^{\frac{1}{2}} \delta(q_N''; q_{N-1}', \mathbf{x}) \text{ for } q_N'' \text{ in } \Omega_R^N, \quad (2.4)$$

and

$$\langle q_N'' | \psi^\dagger(\mathbf{x}) | q_{N-1}' \rangle = 0 \text{ if } q_N'' \text{ is not in } \Omega_R^N, \quad (2.5)$$

in agreement with Eq. (3) of I. The equations derived in I, therefore, remain valid if the definition Eq. (2.1) is used instead of Eq. (3) of I.

In order to derive the equations of motion we choose specially for the functions $\Phi(q_N; \nu_N)$ the solutions of the Schrödinger equation for hard spheres:

$$-\sum_{i=1}^{N-1} \nabla_i^2 \Phi(q_{N-1}; \nu_{N-1}) = E(\nu_{N-1})\Phi(q_{N-1}; \nu_{N-1}) \text{ for } q_{N-1} \text{ in } \Omega_R^{N-1}, \quad (2.6)$$

$$-\left(\sum_{i=1}^{N-1} \nabla_i^2 + \nabla^2\right)\Phi(q_{N-1}, \mathbf{x}; \nu_N) = E(\nu_N)\Phi(q_{N-1}, \mathbf{x}; \nu_N) \text{ for } q_{N-1} \text{ in } \Omega_{R_x}^{N-1}, \quad (2.7)$$

where $\Omega_{R_x}^{N-1}$ is that subset of Ω_R^{N-1} for which $(q_{N-1}, \mathbf{x})$ is in Ω_R^N. Periodicity conditions are imposed as stated above and boundary conditions are $\Phi(q_N; \nu_N) = 0$ for q_N in $\bar{\Omega}_R^N = \Omega^N - \Omega_R^N$ and $\Phi(q_N; \nu_N) \to 0$ when q_N approaches the boundary of $\bar{\Omega}_R^N$.

Rigorous statements of the properties of the Schrödinger functions for hard spheres do not seem to be available in the literature. We have assumed here that the solutions $\Phi(q_N; \nu_N)$ exist and form a complete orthonormal set in the restricted region Ω_R^N, and that the eigenvalues $E(\nu_N)$ are discrete. We have also assumed that these functions have the regularity properties required to obtain, for instance, $\nabla^2\Phi(q_{N-1}, \mathbf{x}; \nu_N)$ in $\Omega_{R_x}^{N-1}$ by differentiating the expansion of $\Phi(q_{N-1}, \mathbf{x}; \nu_N)$ in terms of $\Phi(q_{N-1}; \nu_{N-1})$ term by term.

With both Eqs. (2.6) and (2.7) valid in Ω_{Rx}^{N-1}, we can then multiply Eq. (2.6) by

$$\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)$$

and Eq. (2.7) by

$$\Phi(q_{N-1}; \nu_{N-1}),$$

subtract and integrate over $\Omega_{R_s}^{N-1}$ to obtain

$$\begin{aligned}
&[E(\nu_{N-1}) - E(\nu_N)] \\
&\times \int_{\Omega_{R_s}^{N-1}} \Phi^*(q_{N-1}, \mathbf{x}; \nu_N)\Phi(q_{N-1}; \nu_{N-1})\, dq_{N-1} \\
&= -\sum_{j=1}^{N-1} \int_{\Omega_{R_s}^{N-1}} [\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)\nabla_j^2\Phi(q_{N-1}; \nu_{N-1}) \\
&\quad - \Phi(q_{N-1}; \nu_{N-1})\nabla_j^2\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)]\, dq_{N-1} \\
&\quad + \int_{\Omega_{R_s}^{N-1}} \Phi(q_{N-1}; \nu_{N-1})[\nabla^2\Phi^*(q_{N-1}; \mathbf{x}; \nu_N)]\, dq_{N-1}.
\end{aligned} \tag{2.8}$$

The domain of integration on the left-hand side can be extended to Ω_R^{N-1}, and the integral becomes $N^{-\frac{1}{2}}\langle \nu_N |\psi^\dagger(\mathbf{x})| \nu_{N-1}\rangle$.

The first term on the right-hand side becomes (by symmetry)

$$\begin{aligned}
&-(N-1)\int_{\Omega_{R_s}^{N-1}} dq_{N-2}\, d^3x' \\
&\times [\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N)\nabla'^2\Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1}) \\
&- \Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1})\nabla'^2\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N)] \\
&- (N-1)\int_{\Omega_{R_s}^{N-1}} dq_{N-2} \oint^{\mathbf{x}} d\sigma_{x'} \frac{\mathbf{x} - \mathbf{x}'}{a} \\
&\times [\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N)\nabla'\Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1}) \\
&- \Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1})\nabla'\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N)],
\end{aligned} \tag{2.9}$$

where $\oint^{\mathbf{x}} d\sigma_{x'}$ denotes the integral over the surface of the sphere of radius $a + 0$ with center $\mathbf{x}$. (Since the integrand vanishes on all spheres $|\mathbf{x}_j - \mathbf{x}'| \leq a$, $j = 1, 2, \cdots, N-2$ if $\nabla'\Phi$ and $\nabla'\Phi^*$ is bounded, it is not necessary to restrict the integration over dq_{N-2} further.) Furthermore, $\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N) \to 0$ for $|\mathbf{x}' - \mathbf{x}| \to a$, so that only the second term remains, and one has for the first term on the right-hand side of Eq. (2.8)

$$\begin{aligned}
&(N-1)\int_{\Omega_{R_s}^{N-1}} dq_{N-2} \oint^{\mathbf{x}} d\sigma_{x'} \frac{\mathbf{x} - \mathbf{x}'}{a} \\
&\times \Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1})\nabla'\Phi^*(q_{N-2}, \mathbf{x}', \mathbf{x}; \nu_N) \\
&= (N-1)\int_{\Omega_{R_s}^{N-1}} dq_{N-2} \oint^{\mathbf{x}} d\sigma_{x'} \frac{\mathbf{x} - \mathbf{x}'}{a} \\
&\times \Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1})N^{-\frac{1}{2}}\nabla' \sum_{\nu_{N-1}'} \langle \nu_N |\psi^\dagger(\mathbf{x}')| \nu_{N-1}'\rangle \\
&\times \Phi^*(q_{N-2}, \mathbf{x}; \nu_{N-1}'),
\end{aligned} \tag{2.10}$$

since from Eq. (2.1) follows

$$\begin{aligned}
&\sum_{\nu_{N-1}'} \langle \nu_N |\psi^\dagger(\mathbf{x}')| \nu_{N-1}'\rangle\Phi^*(q_{N-1}; \nu_{N-1}') \\
&\qquad = N^{\frac{1}{2}}\Phi^*(q_{N-1}, \mathbf{x}'; \nu_N),
\end{aligned} \tag{2.11}$$

and with $q_{N-1} \equiv (q_{N-2}, \mathbf{x})$,

$$\begin{aligned}
&\sum_{\nu_{N-1}'} \langle \nu_N |\psi^\dagger(\mathbf{x}')| \nu_{N-1}'\rangle\Phi^*(q_{N-2}, \mathbf{x}; \nu_{N-1}') \\
&\qquad = N^{\frac{1}{2}}\Phi^*(q_{N-2}, \mathbf{x}, \mathbf{x}'; \nu_N).
\end{aligned} \tag{2.12}$$

Since $\Phi^*(q_{N-2}, \mathbf{x}; \nu_{N-1}')$ vanishes, when q_{N-2} is not in $\Omega_{R_s}^{N-2}$, the integral over dq_{N-2} can be extended over Ω^{N-2}, and becomes the mixed density matrix[2]

$$\begin{aligned}
&(N-1)\int_{\Omega^{N-2}} dq_{N-2}\Phi^*(q_{N-2}; \mathbf{x}; \nu_{N-1}')\Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1}) \\
&\qquad = \langle \nu_{N-1}' |\psi^\dagger(\mathbf{x})\psi(\mathbf{x}')| \nu_{N-1}\rangle.
\end{aligned} \tag{2.13}$$

Combining Eqs. (2.9)–(2.13) we thus have for the first term on the right-hand side of Eq. (2.8)

$$\begin{aligned}
&-\sum_{j=1}^{N-1} \int_{\Omega_{R_s}^{N-1}} dq_{N-1}[\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)\nabla_j^2\Phi(q_{N-1}; \nu_{N-1}) \\
&- \Phi(q_{N-1}; \nu_{N-1})\nabla_j^2\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)] = N^{-\frac{1}{2}} \oint^{\mathbf{x}} d\sigma_{x'} \\
&\times \left\langle \nu_N \left| \left[\frac{\mathbf{x} - \mathbf{x}'}{a}\nabla'\psi^\dagger(\mathbf{x}')\right]\psi^\dagger(\mathbf{x})\psi(\mathbf{x}') \right| \nu_{N-1} \right\rangle.
\end{aligned} \tag{2.14}$$

In order to evaluate the last term on the right-hand side of Eq. (2.8), we use Eq. (2.11) to obtain

$$\begin{aligned}
&\int_{\Omega_{R_s}^{N-1}} \Phi(q_{N-1}; \nu_{N-1})\nabla^2\Phi^*(q_{N-1}, \mathbf{x}; \nu_N)\, dq_{N-1} \\
&= N^{-\frac{1}{2}} \sum_{\nu_{N-1}'} \langle \nu_N |\nabla^2\psi^\dagger(\mathbf{x})| \nu_{N-1}'\rangle \\
&\times \int_{\Omega_{R_s}^{N-1}} dq_{N-1}\Phi^*(q_{N-1}; \nu_{N-1}')\Phi(q_{N-1}; \nu_{N-1}).
\end{aligned} \tag{2.15}$$

The integral is identified as the matrix element of the projection operator $P(\mathbf{x}) \equiv 0^{n(s_x)}$ with

$$n(s_x) \equiv \int_{|\mathbf{x}'-\mathbf{x}| \leq a} d^3x'\psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')$$

by changing to the coordinate representation. With

[2] Using Eq. (2.11) and its Hermitian conjugate we have

$$\begin{aligned}
&(N-1)\int_{\Omega^{N-2}} dq_{N-2}\Phi^*(q_{N-2}, \mathbf{x}; \nu_{N-1}')\Phi(q_{N-2}, \mathbf{x}'; \nu_{N-1}) \\
&= \int_{\Omega^{N-2}} dq_{N-2} \sum_{\nu_{N-2}'} \langle \nu_{N-1}'| \psi^\dagger(\mathbf{x}) |\nu_{N-2}'\rangle\Phi^*(q_{N-2}; \nu_{N-2}) \\
&\quad \times \sum_{\nu_{N-2}''} \Phi(q_{N-2}; \nu_{N-2}'')\langle \nu_{N-2}''| \psi(\mathbf{x}') |\nu_{N-1}\rangle \\
&= \sum_{\nu_{N-2}'} \langle \nu_{N-1}'| \psi^\dagger(\mathbf{x}') |\nu_{N-2}'\rangle\langle \nu_{N-2}'| \psi(\mathbf{x}') |\nu_{N-1}\rangle.
\end{aligned}$$

$$\langle \nu'_{N-1} \,|P(\mathbf{x})|\, \nu_{N-1}\rangle$$

$$= \int_{\Omega_{R_x}{}^{N-1}} dq_{N-1}\Phi^*(q_{N-1};\nu'_{N-1})\Phi(q_{N-1};\nu_{N-1}) \quad (2.16)$$

we have

$$\langle q''_{N-1} \,|P(\mathbf{x})|\, q'_{N-1}\rangle$$

$$= \sum_{\nu'_{N-1}}\sum_{\nu_{N-1}} \Phi(q''_{N-1};\nu'_{N-1}) \int_{\Omega_{R_x}{}^{N-1}} dq_{N-1}$$

$$\times \Phi^*(q_{N-1};\nu'_{N-1})\Phi(q_{N-1};\nu_{N-1})\Phi^*(q'_{N-1};\nu_{N-1})$$

$$= \int_{\Omega_{R_x}{}^{N-1}} dq_{N-1}\,\delta(q''_{N-1};\,q_{N-1})\,\delta(q_{N-1};\,q'_{N-1})$$

$$= \begin{cases} \delta(q''_{N-1};\,q'_{N-1}) & \text{if } q'_{N-1} \text{ in } \Omega_{R_x}^{N-1}, \\ 0 & \text{if } q'_{N-1} \text{ not in } \Omega_{R_x}^{N-1}. \end{cases} \quad (2.17)$$

The δ functions used here are symmetrized δ functions, i.e.,

$$\delta(q'_N;\,q_N) \equiv (N!)^{-1}\sum_P \prod_{i=1}^N \delta(\mathbf{x}_i - \mathbf{x}_{P_i}), \quad (2.18)$$

where P_i is the number obtained from i by the permutation P. The three-dimensional δ functions are defined by

$$\int_S \delta(\mathbf{x}' - \mathbf{x})\, d^3x'$$

$$= \begin{cases} 1 \text{ if } \mathbf{x} \text{ in } S \text{ (including the surface, if } S \text{ is a closed region),}^3 \\ 0 \text{ otherwise.} \end{cases}$$

Comparing Eq. (2.17) with the coordinate representation of the density operator [I, Eq. (5)]

$$\langle q''_{N-1} \,|\psi^\dagger(\mathbf{x})\psi(\mathbf{x})|\, q'_{N-1}\rangle = \sum_{k=1}^{N-1} \delta(\mathbf{x} - \mathbf{x}_k)\delta(q''_{N-1};\,q'_{N-1})$$

$$\text{for } q'_{N-1} \text{ in } \Omega_R^{N-1}, \quad (2.20)$$

one sees that in the coordinate representation where both $P(\mathbf{x})$ and

$$n(s_x) = \int_{|\mathbf{x}'-\mathbf{x}|\le a} d^3x'\,\psi^\dagger(\mathbf{x}')\psi(\mathbf{x}') \quad (2.21)$$

are diagonal, $P(\mathbf{x}) = 1$ if $n(s_x) = 0$, and $P(\mathbf{x}) = 0$ if $n(s_x) \neq 0$, so that

$$P(\mathbf{x}) = 0^{n(s_x)}. \quad (2.22)$$

Alternative expressions for $P(\mathbf{x})$ have been derived in I [Eq. (A5) and Eqs. (30), (31)]. Since the eigenvalues of $n(s_x)$ are nonnegative integers and cannot exceed a fixed number m,[4] we can, for instance, write $P(\mathbf{x})$ in the form

$$P(\mathbf{x}) = 1 + \sum_{l=1}^m (-1)^l \binom{n(s_x)}{l}. \quad (2.22')$$

This is a polynomial in $n(s_x)$, and Eq. (2.22') can be used as the definition of $P(\mathbf{x})$ in any representation. We note also that $P(\mathbf{x})$ is the same projection operator which appears in the expression for $\psi(\mathbf{x})\psi^\dagger(\mathbf{x}')$ for $|\mathbf{x} - \mathbf{x}'| \le a$, since the coordinate representation for $\psi(\mathbf{x})\psi^\dagger(\mathbf{x}')$ in this region is given by Eq. (11) of I as

$$\langle q''_{N-1} \,|\psi(\mathbf{x})\psi^\dagger(\mathbf{x}')|\, q'_{N-1}\rangle$$

$$= \begin{cases} \delta(\mathbf{x} - \mathbf{x}')\delta(q''_{N-1};\,q'_{N-1}) \text{ if } q'_{N-1} \text{ in } \Omega_{R_x}^{N-1}, \\ 0 \quad \text{otherwise,} \end{cases} \quad (2.23)$$

so that

$$\psi(\mathbf{x})\psi^\dagger(\mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}')P(\mathbf{x}) \quad (2.24)$$

when $|\mathbf{x} - \mathbf{x}'| \le a$.

Substituting Eqs. (2.1), (2.14), (2.15), and (2.16) into Eq. (2.8), one obtains

$$[E(\nu_{N-1}) - E(\nu_N)]\langle \nu_N \,|\psi^\dagger(\mathbf{x})|\, \nu_{N-1}\rangle$$

$$= \langle \nu_N \,|(\nabla^2\psi^\dagger(\mathbf{x}))P(\mathbf{x})|\, \nu_{N-1}\rangle + \left\langle \nu_N \left| \oint^{\mathbf{x}} d\sigma_{x'} \right.\right.$$

$$\left.\left. \times \left[\frac{\mathbf{x} - \mathbf{x}'}{a}\nabla'\psi^\dagger(\mathbf{x}')\right]\psi^\dagger(\mathbf{x})\psi(\mathbf{x}') \right| \nu_{N-1}\right\rangle. \quad (2.25)$$

Since $\psi^\dagger(\mathbf{x}')$ and $\psi^\dagger(\mathbf{x})$ commute, the last term can be written in the form

$$\left\langle \nu_N \left| \psi^\dagger(\mathbf{x}) \oint^{\mathbf{x}} d\sigma_{x'} \left[\frac{\mathbf{x} - \mathbf{x}'}{a}\nabla'\psi^\dagger(\mathbf{x}')\right]\psi(\mathbf{x}') \right| \nu_{N-1}\right\rangle.$$

We now define $\psi^\dagger(\mathbf{x}, t)$ by

$$\langle \nu_N \,|\psi^\dagger(\mathbf{x}, t)|\, \nu_{N-1}\rangle \equiv \langle \nu_N \,|\psi^\dagger(\mathbf{x})|\, \nu_{N-1}\rangle$$

$$\times \exp\,(i[E(\nu_N) - E(\nu_{N-1})]t). \quad (2.26)$$

Equation (2.25) and its Hermitian conjugate can then be written as equations of motion:

$$i\frac{\partial}{\partial t}\psi^\dagger(\mathbf{x}, t) = (\nabla^2\psi^\dagger(\mathbf{x}, t))\cdot P(\mathbf{x}, t) - \psi^\dagger(\mathbf{x}, t)$$

$$\times \oint^{\mathbf{x}} d\sigma_{x'}\left[\frac{\mathbf{x}' - \mathbf{x}}{a}\nabla'\psi^\dagger(\mathbf{x}', t)\right]\psi(\mathbf{x}', t), \quad (2.27)$$

$$i\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = -P(\mathbf{x}, t)(\nabla^2\psi^\dagger(\mathbf{x}, t))$$

$$+ \oint^{\mathbf{x}} d\sigma_{x'}\psi^\dagger(\mathbf{x}', t)\cdot\frac{\mathbf{x}' - \mathbf{x}}{a}\nabla'\psi(\mathbf{x}', t)\cdot\psi(\mathbf{x}, t), \quad (2.27')$$

where $P(\mathbf{x}, t)$ is defined by

$$P(\mathbf{x}, t) = 0^{n(s_x, t)}, \quad (2.28)$$

[3] This definition is chosen to avoid the value $\frac{1}{2}$ when $\mathbf{x}$ is on the boundary of S.

[4] The actual value of m is irrelevant here.

with

$$n(s_x, t) \equiv \int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x' \psi^\dagger(\mathbf{x}', t)\psi(\mathbf{x}', t). \tag{2.29}$$

The equations of motion can be written in a more convenient form in any representation in which $\nabla(\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) \cdots \psi(\mathbf{x}_n)\psi(\mathbf{x}))$ is bounded for any set of points $\mathbf{x}_1, \cdots \mathbf{x}_n, \mathbf{x}$.

In I we had shown that the projection operator $P(\mathbf{x})$ can be written in the form

$$P(\mathbf{x}) = 1 + \sum_{l=1}^{m} \frac{(-1)^l}{l!} \times \int \cdots \int_{s_x} \prod_{j=1}^{l} \psi^\dagger(\mathbf{x}_j) \prod_{k=1}^{l} \psi(\mathbf{x}_k)\, d^3x_k, \tag{2.30}$$

where s_x is the sphere $|\mathbf{x}_k - \mathbf{x}| \leq a$, and m is the largest eigenvalue of $n(s_x)$. We, therefore, have

$$P(\mathbf{x})\nabla^2\psi(\mathbf{x}) = \nabla^2\psi(\mathbf{x}) + \sum_{l=1}^{m} \frac{(-1)^l}{l!} \times \int \cdots \int_{s_x} \prod_{j=1}^{l} \psi^\dagger(\mathbf{x}_j) \prod_{j=1}^{l} \psi(\mathbf{x}_k)\, d^3x_k \nabla^2\psi(\mathbf{x}). \tag{2.31}$$

To evaluate the first term in the sum we use a test function $f(\mathbf{x})$ and evaluate

$$\iint_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x'\, d^3x f(\mathbf{x})\psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')\nabla^2\psi(\mathbf{x})$$

$$= \int \psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')\, d^3x' \int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x f(\mathbf{x})\nabla^2\psi(\mathbf{x})$$

$$= \int \psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')\, d^3x' \int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x \times \{\nabla\cdot(f(\mathbf{x})\nabla\psi(\mathbf{x})) - \nabla f(\mathbf{x})\cdot\nabla\psi(\mathbf{x})\} \tag{2.32}$$

$$= \int \psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')\, d^3x' \left[\oint^{\mathbf{x}'} d\sigma_x f(\mathbf{x})\cdot\frac{\mathbf{x}'-\mathbf{x}}{a}\nabla\psi(\mathbf{x}) - \int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x\, \nabla f(\mathbf{x})\cdot\nabla\psi(\mathbf{x})\right]. \tag{2.33}$$

Since $\int d^3x' \oint^{\mathbf{x}'} d\sigma_x$ is an integral over the hypersurface $|\mathbf{x}' - \mathbf{x}| = a$ in the six-dimensional $(\mathbf{x}, \mathbf{x}')$ space, it can be replaced by $\int d^3x \oint^{\mathbf{x}} d\sigma_{x'}$ and we have

$$\int f(\mathbf{x})\, d^3x \int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x' \psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')\nabla^2\psi(\mathbf{x})$$

$$= \int f(\mathbf{x})\, d^3x \oint^{\mathbf{x}} d\sigma_{x'}\psi^\dagger(\mathbf{x}')\cdot\frac{\mathbf{x}'-\mathbf{x}}{a}\nabla\psi(\mathbf{x}')\psi(\mathbf{x})$$

$$- \int d^3x\, \nabla f(\mathbf{x})\cdot\int_{|\mathbf{x}'-\mathbf{x}|\leq a} d^3x'\psi^\dagger(\mathbf{x}')\nabla\psi(\mathbf{x}')\psi(\mathbf{x}). \tag{2.34}$$

The second term vanishes since $\nabla\psi(\mathbf{x}')\psi(\mathbf{x})$ vanishes for $|\mathbf{x}' - \mathbf{x}| < a$, and is by assumption finite for $|\mathbf{x}' - \mathbf{x}| = a$.

In order to show that the terms with $l \geq 2$ in Eq. (2.31) vanish, it will be sufficient to show that the term $l = 2$ vanishes, since the argument is the same for all the other terms. Using again a test function $f(\mathbf{x})$, we have

$$\int f(\mathbf{x})\, d^3x \iint_{s_x} d^3x_1\, d^3x_2 \psi^\dagger(\mathbf{x}_1)\psi^\dagger(\mathbf{x}_2)\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\nabla^2\psi(\mathbf{x})$$

$$= \iint d^3x_1\, d^3x_2 \psi^\dagger(\mathbf{x}_1)\psi^\dagger(\mathbf{x}_2)\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) \times \int_{\substack{|\mathbf{x}-\mathbf{x}_1|\leq a \\ |\mathbf{x}-\mathbf{x}_2|\leq a}} d^3x\, f(\mathbf{x})\nabla^2\psi(\mathbf{x})$$

$$= \iint d^3x_1\, d^3x_2 \psi^\dagger(\mathbf{x}_1)\psi^\dagger(\mathbf{x}_2)\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) \times \left\{ \oint^{\mathbf{x}_1}_{|\mathbf{x}-\mathbf{x}_2|\leq a} d\sigma_x f(\mathbf{x})\frac{\mathbf{x}_1-\mathbf{x}}{a}\nabla\psi(\mathbf{x})\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) + \oint^{\mathbf{x}_2}_{|\mathbf{x}-\mathbf{x}_1|\leq a} d\sigma_x f(\mathbf{x})\frac{\mathbf{x}_2-\mathbf{x}}{a}\nabla\psi(\mathbf{x})\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) - \int_{\substack{|\mathbf{x}-\mathbf{x}_1|\leq a \\ |\mathbf{x}-\mathbf{x}_2|\leq a}} d^3x \nabla f(\mathbf{x})\cdot\nabla\psi(\mathbf{x})\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\right\}, \tag{2.35}$$

where $\oint^{\mathbf{x}_i}_{|\mathbf{x}-\mathbf{x}_j|\leq a}$ denotes the surface integral over that part of the sphere $|\mathbf{x} - \mathbf{x}_i| = a$ for which $|\mathbf{x} - \mathbf{x}_j| \leq a$. Since $\nabla\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\psi(\mathbf{x})$ vanishes for $|\mathbf{x} - \mathbf{x}_1| < a$ and for $|\mathbf{x} - \mathbf{x}_2| < a$, the integrands of the two surface integrals vanish in the region of integration, except on the line $|\mathbf{x}-\mathbf{x}_1| = |\mathbf{x}-\mathbf{x}_2| = a$, where the integrands are bounded. The surface integrals, therefore, vanish for $\mathbf{x}_1 \neq \mathbf{x}_2$. For $\mathbf{x}_1 = \mathbf{x}$ they also vanish since $\psi(\mathbf{x}_1)\psi(\mathbf{x}_2) = 0$ for $|\mathbf{x} - \mathbf{x}_2| \leq a$. The last term in Eq. (2.35) vanishes, since $\nabla\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\psi(\mathbf{x})$ vanishes for $|\mathbf{x} - \mathbf{x}_1| < a$ and for $|\mathbf{x} - \mathbf{x}_2| < a$, and is by assumption bounded.

For representations for which $\nabla\psi(\mathbf{x}_1)\cdots\psi(\mathbf{x}_N)\psi(\mathbf{x})$ is bounded, the equations of motion therefore reduce to

$$i\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\nabla^2\psi(\mathbf{x}, t) + \oint^{\mathbf{x}} d\sigma_{x'}\psi^\dagger(\mathbf{x}', t) \times \frac{\mathbf{x}'-\mathbf{x}}{a}(\nabla' - \nabla)\psi(\mathbf{x}', t)\psi(\mathbf{x}, t) \tag{2.36}$$

and its Hermitian conjugate.

We note that

$$\frac{\mathbf{x}'-\mathbf{x}}{a}(\nabla' - \nabla)\psi(\mathbf{x}', t)\psi(\mathbf{x}, t) = \frac{1}{f(\mathbf{x}, \mathbf{x}')}\frac{\mathbf{x}'-\mathbf{x}}{a} \times (\nabla' - \nabla)[f(\mathbf{x}, \mathbf{x}')\psi(\mathbf{x}', t)\psi(\mathbf{x}, t)],$$

for $|\mathbf{x}' - \mathbf{x}| \leq a$ if $(\mathbf{x}' - \mathbf{x})/a(\nabla' - \nabla)f(\mathbf{x}', \mathbf{x})$ is bounded and $f(\mathbf{x}', \mathbf{x}) \not\equiv 0$. Equation (2.36) can, therefore, be written in various forms; we have for instance, with $f(\mathbf{x}', \mathbf{x}) = |\mathbf{x}' - \mathbf{x}|$,

$$i\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = -\nabla^2\psi(\mathbf{x}, t) + \oint^{\mathbf{x}} d\sigma_{x'}\psi^\dagger(\mathbf{x}', t) \times \frac{\mathbf{x}' - \mathbf{x}}{a^2}(\nabla' - \nabla)[|\mathbf{x}' - \mathbf{x}|\,\psi(\mathbf{x}', t)\psi(\mathbf{x}, t)]. \quad (2.37)$$

Equation (2.37) provides the exact formulation of the equation of motion with pseudopotential since

$$[(\mathbf{x}' - \mathbf{x})/a](\nabla' - \nabla) = 2(\partial/\partial r), \quad (2.38)$$

with $r = |\mathbf{x}' - \mathbf{x}|$, and the partial differentiation $\partial/\partial r$ defined with $\mathbf{x}' + \mathbf{x}$ fixed. With this notation, Eq. (2.37) becomes

$$\begin{aligned} i\frac{\partial}{\partial t}\psi(\mathbf{x}, t) &= -\nabla^2\psi(\mathbf{x}, t) + \frac{2}{a}\oint^{\mathbf{x}} d\sigma_{x'}\psi^\dagger(\mathbf{x}', t) \\ &\quad \times \frac{\partial}{\partial r}[r\psi(\mathbf{x}', t)\psi(\mathbf{x}, t)] \\ &= -\nabla^2\psi(\mathbf{x}, t) + \frac{2}{a}\int_\Omega d^3x'\psi^\dagger(\mathbf{x}', t) \\ &\quad \times \delta(r - a - 0)\frac{\partial}{\partial r}[r\psi(\mathbf{x}', t)\psi(\mathbf{x}, t)], \end{aligned} \quad (2.39)$$

where $\delta(r - a - 0)$ is the one-dimensional δ function.[5] The operator

$$(2/a)\delta(r - a - 0)(\partial/\partial r)$$

is, therefore, the exact counterpart of the pseudopotential. The standard form of the pseudopotential

$$8\pi a[\delta(\mathbf{r})(\partial/\partial r)]r$$

is obtained for sufficiently small a, since

$$(4\pi a^2)^{-1}\delta(r - a - 0) \to \delta(\mathbf{r}), \quad (2.40)$$

where $\delta(\mathbf{r})$ is the three-dimensional δ function and $\mathbf{r} \equiv \mathbf{x}' - \mathbf{x}$.

Equation (2.8), from which we derived the equations of motion, can be written in the canonical form

$$i(\partial\psi^\dagger/\partial t) = [\psi^\dagger, H] \quad (2.41)$$

by a rearrangement of the terms, and one finds that

$$H = -\int_\Omega d^3x\psi^\dagger(\mathbf{x}, t)\nabla^2\psi(\mathbf{x}, t) \quad (2.42)$$

as one would expect, since the interaction, considered as a constraint, has been expressed in the new commutation rules.

We can now define a pseudopotential Hamiltonian H_{ps}, so that the equations of motion (Eq. 2.39) are obtained from

$$i(\partial\psi/\partial t) = [\psi, H_{ps}], \quad (2.43)$$

when the new commutation rules are approximated by the standard commutation rules. One shows by straightforward computation that

$$H_{ps} = -\int \psi^\dagger(\mathbf{x}')\nabla'^2\psi(\mathbf{x}')\,d^3x' + \frac{1}{a}\iint d^3x'\,d^3x'' \times \psi^\dagger(\mathbf{x}')\psi^\dagger(\mathbf{x}'')\,\delta(r - a - 0)\frac{\partial}{\partial r}[r\psi(\mathbf{x}')\psi(\mathbf{x}'')]. \quad (2.44)$$

3. APPLICATIONS TO QUANTUM STATISTICAL MECHANICS

The derivation of the commutation rules and equations of motion in I, and in the preceding section, shows their validity only as matrix equations in a particular representation. In Secs. 4 and 5, we will show that the commutation rules and the first form of the equations of motion [Eqs. (2.27) and (2.27′)], understood as operator equations, constitute a formalism which has the properties required of a nonrelativistic field theory. It has not been possible to prove the validity of the second form of the equations of motion [Eq. (2.36)] in this full generality. Since the interest in this second and simpler form of the equations of motion is motivated by possible applications in quantum statistical mechanics, we will show in this section that it yields—together with the modified commutation rules—correct equations for the density matrix.

For this purpose we introduce the operators $\varphi(\mathbf{x}, \beta)$ and their Matsubara conjugates

$$\varphi^*(\mathbf{x}, \beta) \equiv \varphi^\dagger(\mathbf{x}, -\beta), \quad (3.1)$$

where β is a reciprocal temperature variable. We define these operators by the modified commutation rules and equations of motion previously obtained as matrix equations for $\psi(\mathbf{x}, t)$ and $\psi^\dagger(\mathbf{x}, t)$, replacing t by $-i\beta$:

$$[\varphi(\mathbf{x}', \beta), \varphi(\mathbf{x}, \beta)] = [\varphi^*(\mathbf{x}', \beta), \varphi^*(\mathbf{x}, \beta)] = 0, \quad (3.2)$$

$$[\varphi(\mathbf{x}', \beta), \varphi^*(\mathbf{x}, \beta)] = \begin{cases} 0 & \text{for } |\mathbf{x}' - \mathbf{x}| > a, \\ \delta(\mathbf{x}' - \mathbf{x})P(\mathbf{x}) + \varphi^*(\mathbf{x}, \beta)\varphi(\mathbf{x}, \beta) & \\ \qquad \text{for } |\mathbf{x}' - \mathbf{x}| \leq a, & \end{cases} \quad (3.3)$$

where $P(\mathbf{x}, \beta)$ is the projection operator defined as in Eqs. (2.28), (2.29), with $\psi^*(\mathbf{x}, t)\psi(\mathbf{x}, t)$ replaced by $\varphi^*(\mathbf{x}, \beta)\varphi(\mathbf{x}, \beta)$. We assume the equation of motion in the integral form of Eq. (2.36),

[5] The notation $\int \delta(r + a + 0)f(r)dr$ stands for $\lim_{r\to a+0} f(r)$.

$$\varphi(\mathbf{x}, \beta) = \int_\Omega \varphi(\mathbf{x}_1, 0)\, d^3x_1 G_1(\mathbf{x}_1, \mathbf{x}; \beta)$$
$$- \int_0^\beta dt \int_\Omega d^3x_1 \oint^{\mathbf{x}_1} d\sigma' \varphi^*(\mathbf{x}', t) \frac{\mathbf{x}' - \mathbf{x}_1}{a}$$
$$\times (\nabla' - \nabla_1)\varphi(\mathbf{x}', t)\varphi(\mathbf{x}_1, t) G_1(\mathbf{x}_1, \mathbf{x}; \beta - t), \quad (3.4)$$

where $G_1(\mathbf{x}_1, \mathbf{x}; \beta)$ is defined by

$$(\partial/\partial\beta - \nabla^2) G_1(\mathbf{x}_1, \mathbf{x}; \beta) = 0, \quad (3.5)$$
$$G_1(\mathbf{x}_1, \mathbf{x}; 0) = \delta(\mathbf{x}_1 - \mathbf{x}), \quad (3.5')$$

and Born–von Karman boundary conditions on the surface of a cube of volume Ω.

We also assume the existence of a unique vacuum state $|0\rangle$ with

$$\varphi(\mathbf{x}, \beta)\, |0\rangle = 0. \quad (3.6)$$

We will show that

$$\Gamma_N(q_N^0, \beta_0; q_N, \beta)$$
$$\equiv (N!)^{-1} \langle 0| \prod_{l=1}^N \varphi(\mathbf{x}_l, \beta) \prod_{s=1}^N \varphi^*(\mathbf{x}_s^0, \beta_0)\, |0\rangle \quad (3.7)$$

is the symmetrized Green's function $G_N^S(q_N^0, q_N; \beta - \beta_0)$ of the Bloch equation for N hard spheres of diameter a.

For this purpose we derive first a sequence of differentio-integral equations for the functions $G_N^S(q_N^0, q_N; \beta)$ [Eq. (3.12)], then show that the functions $\Gamma_N(q_N^0, 0; q_N, \beta)$ satisfy the same sequence.

The Bloch equation for the Green's function $g(\mathbf{x}^0, \mathbf{x}; \beta)$ for one particle in a nonnegative potential $v(\mathbf{x})$,

$$\partial/\partial\beta\, g(\mathbf{x}^0, \mathbf{x}; \beta) = \nabla^2 g(\mathbf{x}^0, \mathbf{x}; \beta)$$
$$- g(\mathbf{x}^0, \mathbf{x}; \beta) v(\mathbf{x}), \quad (3.8)$$

can be considered as the equation for Brownian motion of a particle in an absorbing medium. If β is interpreted as time and $v(\mathbf{x})d\beta$ as the probability of absorption at $\mathbf{x}$ in the time interval $(\beta, \beta + d\beta)$, then $g(\mathbf{x}^0, \mathbf{x}; \beta)d^3x$ is the probability that the particle, having started at $\mathbf{x}^0$ at time zero, is in the volume element d^3x at $\mathbf{x}$ at the time β.

In the case of the hard-sphere potential

$$v(\mathbf{x}) = \begin{cases} 0 & \text{for } |\mathbf{x}| > a, \\ \infty & \text{for } |\mathbf{x}| \leq a, \end{cases} \quad (3.9)$$

absorption occurs instantaneously with certainty if the particle is at a point $\mathbf{x}$ on the surface of the sphere. The function $g(\mathbf{x}^0, \mathbf{x}, \beta)$ is then the solution of the equation, valid for $|\mathbf{x}^0| \geq a$,

$$g(\mathbf{x}^0, \mathbf{x}; \beta) = g_0(\mathbf{x}^0, \mathbf{x}; \beta) - \int_0^\beta dt \oint^0 d\sigma'$$
$$\times \left[\frac{\mathbf{x}'}{a} \nabla' g(\mathbf{x}^0, \mathbf{x}'; t)\right] g_0(\mathbf{x}', \mathbf{x}; \beta - t), \quad (3.10)$$

where $g_0(\mathbf{x}^0, \mathbf{x}; \beta)$ is the transition probability density in the absence of absorption and $\oint^0 d\sigma'$ is the surface integral over the sphere of radius $a + 0$, with center at the origin. This integral equation can be derived by a simple counting argument. We imagine that the particle, instead of being absorbed, receives a marker, which does not affect its motion. The conditional probability $g(\mathbf{x}^0, \mathbf{x}; \beta)\, d^3x$ of finding an unmarked particle in d^3x at $\mathbf{x}$ at time β is then obtained by subtracting, from the probability $g_0(\mathbf{x}^0, \mathbf{x}; \beta)$ of finding the particle in d^3x at $\mathbf{x}$ at time β with or without marker, the probability that the particle hit the sphere in some surface element $d\sigma_x'$ at $\mathbf{x}$ for the first time in a time interval $(t, t + dt)(0 < t < \beta)$ and then continued its motion (regardless of additional markers) during the remaining time $\beta - t$, so as to be in d^3x at $\mathbf{x}$ at time β. The probability that an unmarked particle which started from $\mathbf{x}^0$ at time zero hits the surface element $d\sigma'$ of the sphere at the point $\mathbf{x}'$ in $(t, t + dt)$ is obtained from the diffusion current density as

$$[(1/a)\mathbf{x}' \cdot \nabla'] g(\mathbf{x}^0, \mathbf{x}'; t)\, d\sigma'\, dt',$$

and the probability that this particle, having started at $\mathbf{x}'$ at time t, is in d^3x at $\mathbf{x}$ at time β is

$$g_0(\mathbf{x}', \mathbf{x}; \beta - t)\, d^3x.$$

We thus obtain

$$g(\mathbf{x}^0, \mathbf{x}; \beta)\, d^3x = g_0(\mathbf{x}^0, \mathbf{x}; \beta)\, d^3x - \int_0^\beta dt \oint^0 d\sigma'$$
$$\times \left[\frac{\mathbf{x}'}{a} \cdot \nabla' g(\mathbf{x}^0, \mathbf{x}'; t)\right] g_0(\mathbf{x}', \mathbf{x}; \beta - t)\, d^3x, \quad (3.10')$$

in agreement with Eq. (3.10).

By a similar argument one derives a differentio-integral equation for the Green's function $G_N(q_N^0, q_N; \beta)$ for N hard spheres of diameter a in terms of $G_{N-1}(q_{N-1}^0, q_{N-1}; \beta) G_1(\mathbf{x}_N^0, \mathbf{x}_N; \beta)$. We interpret $G_N(q_N^0, q_N; \beta)$ as the conditional probability density that N particles in Brownian motion, which started from the configuration q_N^0 at time zero, are in the neighborhood dq_N of the configuration q_N at time β, all without markers. The conditional probability—that particles $1, 2 \cdots N - 1$ are in dq_{N-1} at q_{N-1} at time t, and that the Nth particle hits the surface element $d\sigma_{x_N}$ of the sphere of radius a with center $\mathbf{x}_j$ in the time interval $(t, t + dt)$—is then given by

$$\frac{1}{a} (\mathbf{x}_N - \mathbf{x}_j)(\nabla_N - \nabla_j) G_N(q_N^0, q_N; t)\, d\sigma_{x_N}\, dt\, dq_{N-1}.$$

Since this expression vanishes when $|\mathbf{x}_N - \mathbf{x}_l| < a$ for $l \neq j$, one obtains, by the argument used before,

the equation

$$G_N(q_N^0, q_N; \beta) = G_{N-1}(q_{N-1}^0, q_{N-1}; \beta)G_1(\mathbf{x}_N^0, \mathbf{x}_N; \beta)$$
$$- \int_0^\beta dt \int_{\Omega_R{}^{N-1}} dq'_{N-1} \sum_{i=1}^{N-1} \oint^{x'_i} d\sigma_{x'_N}$$
$$\times \left[\frac{\mathbf{x}'_N - \mathbf{x}'_i}{a}(\nabla'_N - \nabla'_i)G_N(q_N^0, q'_N; t)\right]$$
$$\times G_{N-1}(q'_{N-1}, q_{N-1}; \beta - t)G_1(\mathbf{x}'_N, \mathbf{x}_N; \beta - t) \quad (3.11)$$

for

$$q_N^0 \in \Omega_R^N.$$

This equation can also be derived by a limiting process from a *differentio*-integral equation for the Green's function with finite differentiable potentials (see Appendix I).

The corresponding differentio-integral equation for the Green's function G_N^S symmetric in q_N^0 and q_N can be obtained by application of the symmetrizing operator to G_N as function of q_N^0, and one obtains

$$G_N^S(q_N^0, q_N; \beta)$$
$$= \frac{1}{N}\sum_{i=1}^{N} G_{N-1}^S({}_iq_N^0, q_{N-1}; \beta)G_1(\mathbf{x}_i^0, \mathbf{x}_N; \beta)$$
$$- \int_0^\beta dt \int_{\Omega_R{}^{N-1}} dq'_{N-1} \sum_{i=1}^{N-1} \oint^{x'_i} d\sigma_{x'_N}$$
$$\times \left[\frac{\mathbf{x}'_i - \mathbf{x}'_N}{a}(\nabla'_N - \nabla'_i)G_N^S(q_N^0, q'_N; t)\right]$$
$$\times G_{N-1}(q'_{N-1}, q_{N-1}; \beta - t)G_1(\mathbf{x}_N, \mathbf{x}'_N; \beta - t), \quad (3.12)$$

where

$${}_iq_N^0 \equiv (\mathbf{x}_1^0, \mathbf{x}_2^0, \cdots \mathbf{x}_{i-1}^0, \mathbf{x}_{i+1}^0, \cdots \mathbf{x}_N^0).$$

In the second term one can replace G_{N-1} by G_{N-1}^S without error.

In order to derive the same sequence of differentio-integral equations for the functions $\Gamma_N(q_N^0, \beta_0; q_N, \beta)$, we consider the expression

$$J \equiv \sum_{i=1}^{N-1} \int_0^\beta dt \int d^3x'_i \int_{\Omega^{N-1}} dq'_{N-1} \oint^{x'_i} d\sigma_{x'_N}$$
$$\times \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta) \prod_{k=1}^{N-1} \varphi^*(\mathbf{x}_k, t) |0\rangle$$
$$\times \langle 0| \partial'_{i_N} \prod_{r=1}^{N} \varphi(\mathbf{x}'_r, t)$$
$$\times \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0) |0\rangle G_s(\mathbf{x}'_N, \mathbf{x}_N; \beta - t), \quad (3.13)$$

where

$$\partial'_{i_N} \equiv [(\mathbf{x}'_i - \mathbf{x}'_N)/a](\nabla'_i - \nabla'_N). \quad (3.14)$$

With the assumption of a unique vacuum state, the product of the vacuum expectation values in Eq. (3.14) can be replaced by the vacuum expectation value of the product of the two operators, since their matrix elements between the vacuum state and any nonvacuum state vanish.

We write the resulting expression in the form

$$J = \sum_{i=1}^{N-1} \int_0^\beta dt \int_{\Omega_{N-1}} dq'_{N-1} \oint^{x'_i} d\sigma_{x'_N} \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta)$$
$$\times \prod_{k=1}^{N-1} \varphi^*(\mathbf{x}'_k, t) \prod_{\substack{r=1 \\ r\neq i}}^{N-1} \varphi(\mathbf{x}'_r, t)\partial'_{i_N}\varphi(\mathbf{x}'_i, t)\varphi(x'_N, t) \quad (3.15)$$
$$\times \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0) |0\rangle G_1(\mathbf{x}'_N, \mathbf{x}_N; \beta - t)$$
$$= \sum_{i=1}^{N-1} \int_0^\beta dt \oint^{x'_i} d\sigma_{x'_N} \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta)\varphi^*(\mathbf{x}'_i, t)$$
$$\times \int dq'_{N-2} \prod_{\substack{k=1 \\ k\neq i}}^{N-1} \varphi^*(\mathbf{x}'_k, t) \prod_{\substack{r=1 \\ r\neq i}}^{N-1} \varphi(\mathbf{x}'_r, t)\partial'_{i_N}\varphi(\mathbf{x}'_i, t)$$
$$\times \varphi(\mathbf{x}'_N, t) \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0) |0\rangle G_1(\mathbf{x}'_N, \mathbf{x}_N, \beta - t)$$

Let us now study the operator that appears in J,

$$K_i^{N-2} \equiv \int dq'_{N-2} \prod_{\substack{k=1 \\ k\neq i}}^{N-1} \varphi^*(\mathbf{x}'_k, t) \prod_{\substack{r=1 \\ r\neq i}}^{N-1} \varphi(\mathbf{x}'_r, t), \quad (3.16)$$

where

$$dq'_{N-1} = \prod_{\substack{i=1 \\ i\neq j}}^{N-1} d^3x'_i \,.$$

Relabeling the variables:

$$K_i^{N-2} = \int dq'_{N-2} \prod_{k=1}^{N-2} \varphi^*(\mathbf{x}'_k, t) \prod_{i=1}^{N-2} \varphi(\mathbf{x}'_i, t)$$
$$= \int dq'_{N-3} \prod_{k=1}^{N-3} \varphi^*(\mathbf{x}'_k, t) \int dx'_{N-2}$$
$$\times \varphi^*(\mathbf{x}'_{N-2}, t)\varphi(\mathbf{x}'_{N-2}, t) \prod_{r=1}^{N-3} \varphi(\mathbf{x}'_2, t)$$
$$= \int dq'_{N-3} \prod_{k=1}^{N-3} \varphi^*(\mathbf{x}'_k, t) \prod_{r=1}^{N-3} \varphi(\mathbf{x}'_r, t)$$
$$\times [\mathfrak{N}_i - (N-3)] = K_i^{N-3}[\mathfrak{N}_i - (N-3)], \quad (3.17)$$

where we have introduced the number operator

$$\mathfrak{N}_i \equiv \int dx \varphi^*(\mathbf{x}, t)\varphi(\mathbf{x}, t),$$

and have used its commutation properties with $\varphi(\mathbf{x}'_r, t)$. From (3.17) one has, by induction,

$$K_i^{N-2} = [\mathfrak{N}_i(\mathfrak{N}_i - 1) \cdots][\mathfrak{N}_i - (N-3)]. \quad (3.18)$$

Notice that the interaction we consider does not change the number of particles, so that for all pairs t, t' one has $\mathfrak{N}_t = \mathfrak{N}_{t'} \equiv \mathfrak{N}$. We now make use of (3.18) in Eq. (3.15) and remark that the state vector on which K_t^{N-2} acts is an eigenstate of $\mathfrak{N}_t$ to the eigenvalue $N-2$. It is therefore an eigenstate of K_t^{N-2} to the eigenvalue $(N-2)!$. Summing up, we have now

$$J = (N-2)! \sum_{j=1}^{N-1} \int_0^\beta dt \int_\Omega d^3x_j' \oint^{\mathbf{x}'_j} d\sigma_{x'_N} \times \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta)\varphi^*(\mathbf{x}_j', t) \times \partial'_{jN}\varphi(\mathbf{x}_j', t)\varphi(\mathbf{x}_N', t) \times \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle G(\mathbf{x}_N', \mathbf{x}_N; \beta - t). \tag{3.19}$$

With the assumption that Ω is connected to itself, we have

$$\int_\Omega d^3x_j' \oint^{\mathbf{x}'_j} d\sigma_{x_N'} \cdots = \int_\Omega d^3x_N' \oint^{\mathbf{x}'_N} d\sigma_{x'_j}. \tag{3.20}$$

After making this substitution, we use Eq. (3.4) to obtain from Eq. (3.19)

$$\begin{aligned} J &= (N-2)! \sum_{j=1}^{N-1} \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta) \\ &\times \left\{ \int \varphi(\mathbf{x}', 0)\, d^3x' G_1(\mathbf{x}', \mathbf{x}_N; \beta) \right. \\ &\left. - \varphi(\mathbf{x}_N, \beta) \right\} \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle \\ &= (N-1)! \left\{ \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta) \int_\Omega \varphi(\mathbf{x}', 0)\, d^3x' \right. \\ &\times G_1(\mathbf{x}', \mathbf{x}_N; \beta) \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle \\ &\left. - \langle 0| \prod_{l=1}^{N-1} \varphi(\mathbf{x}_l, \beta) \prod_{s=1}^{N-1} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle \right\}. \end{aligned} \tag{3.21}$$

In order to simplify the first term, we note first that the contributions to the integral come only from the region Ω' where $|\mathbf{x}' - \mathbf{x}_s^0| \leq a$ for at least one of the points $\mathbf{x}_s^0$. When $\mathbf{x}'$ is not in this region, $\varphi(\mathbf{x}', 0)$ commutes with $\prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0)$ and can be pulled to the right to act on the vacuum state.

The region Ω' can be further reduced, since by virtue of Eq. (3.3)

$$\varphi(\mathbf{x}, 0)\varphi^*(\mathbf{x}_k^0, 0) = P(\mathbf{x}_k^0, 0)\delta(\mathbf{x}' - \mathbf{x}_k^0) \quad \text{for} \quad |\mathbf{x}' - \mathbf{x}_k^0| \leq a, \tag{3.22}$$

where $P(\mathbf{x}_k^0, 0)$ is defined by Eq. (2.22) or (2.22'), and Eq. (2.21) [with $\psi^\dagger(\mathbf{x}')\psi(\mathbf{x}')$ replaced by $\varphi^*(\mathbf{x}', 0)\varphi(\mathbf{x}', 0)$]. If the point $\{\mathbf{x}_1^0, \cdots, \mathbf{x}_N^0\}$ is in Ω_R^N, the points are obviously distinct, and we can replace Ω' without error by the union of spheres s_k with centers at $\mathbf{x}_k^0$, and with a radius $\epsilon < a$ chosen small enough so that the spheres do not overlap. We then have

$$\begin{aligned} &\int_\Omega \varphi(\mathbf{x}', 0)\, d^3x' G_1(\mathbf{x}', \mathbf{x}_N; \beta) \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle \\ &= \sum_{k=1}^{N} \int_{s'_k} \varphi(\mathbf{x}', 0)\, d^3x' G_1(\mathbf{x}', \mathbf{x}_N; \beta)\varphi^*(\mathbf{x}_k^0, 0) \\ &\times \prod_{\substack{s=1 \\ s\neq k}}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle = \sum_{k=1}^{N} P(\mathbf{x}_k^0, 0) \\ &\times \prod_{\substack{s=1 \\ s\neq k}}^{N} \varphi^*(\mathbf{x}_s^0, 0)\,|0\rangle G_1(\mathbf{x}_k^0, \mathbf{x}_N; \beta) \end{aligned} \tag{3.23}$$

if $\{\mathbf{x}_1^0, \cdots, \mathbf{x}_N^0\}$ is in Ω_R^N. With this condition, the projection operator $P(\mathbf{x}_k^0, 0)$ can be omitted, since it then commutes with all $\varphi^*(\mathbf{x}_s^0, 0)$ for $K \neq s$, and can be pulled through to act on the vacuum, and be replaced by unity. In order to combine Eqs. (3.14), (3.21), and (3.23), and to compare the result with Eq. (3.12), we use the notation of Eq. (3.7):

$$\Gamma_N(q_N^0, t_0; q_N, t) = \frac{1}{N!} \langle 0| \prod_{r=1}^{N} \varphi(\mathbf{x}_r, t) \prod_{s=1}^{N} \varphi^*(\mathbf{x}_s^0, t_0)\,|0\rangle, \tag{3.24}$$

and obtain from Eqs. (3.14), (3.21), and (3.23)

$$\begin{aligned} J &= N!\,(N-1)! \sum_{j=1}^{N-1} \int_0^\beta dt \int_{\Omega^{N-1}} dq_{N-1}' \oint^{\mathbf{x}'_j} d\sigma_{x'_N} \\ &\times [\partial'_{jN}\Gamma_N(q_N^0, 0; q_N', t)] \\ &\times \Gamma_{N-1}(q_{N-1}', t; q_{N-1}, \beta)G_1(\mathbf{x}_N', \mathbf{x}_N; \beta - t) \\ &= (N-1)! \left\{ \sum_{k=1}^{N} (N-1)!\, \Gamma_{N-1}({}_kq_N^0, 0; q_{N-1}, \beta) \right. \\ &\left. \times G_1(\mathbf{x}_k^0, \mathbf{x}_N; \beta) - N!\, \Gamma_N(q_N^0, 0; q_N, \beta) \right\}, \end{aligned} \tag{3.25}$$

or

$$\begin{aligned} \Gamma_N(q_N^0, 0; q_N, \beta) &= \frac{1}{N} \sum_{k=1}^{N} \Gamma_{N-1}({}_kq_N^0, 0; q_{N-1}, \beta) \\ &\times G_1(\mathbf{x}_N^0, \mathbf{x}_N; \beta) - \int_0^\beta dt \int_{\Omega^{N-1}} dq_{N-1}' \\ &\times \sum_{j=1}^{N-1} \oint^{\mathbf{x}'_j} d\sigma_{x'_N} [\partial'_{jN}\Gamma_N(q_N^0, 0; q_N', t)] \\ &\times \Gamma_{N-1}(q_{N-1}', t; q_{N-1}, \beta)G_1(\mathbf{x}_N', \mathbf{x}_N; \beta - t) \end{aligned} \tag{3.26}$$

for $q_N^0 \in \Omega_R^N$.

From Eq. (3.24) follows also that $\Gamma_N(q_N^0, t_0; q_N, t)$

depends only on $t - t_0$, that

$$\Gamma_N(q_N^0, t_0; q_N, t_0) = \delta(q_N; q_N^0) \quad \text{for} \quad q_N, q_N^0 \in \Omega_R^N, \tag{3.27}$$

and that

$$\Gamma_1(\mathbf{x}^0, t_0; \mathbf{x}, t) = G_1(\mathbf{x}^0, \mathbf{x}; t - t_0). \tag{3.28}$$

In Eq. (3.26) the domain of integration for q'_{N-1} can be reduced to Ω_R^{N-1}, and one sees that the functions $\Gamma_N(q_N^0, t_0; q_N, t)$ satisfy the same recursive system of differentio-integral equations [Eq. (3.12)] as the functions $G_N^S(q_N^0, q_N; t - t_0)$ for $N > 1$. The interpretation of Eq. (3.12) in terms of Brownian motion theory shows that the solution of Eq. (3.12) must be unique, and we can therefore conclude that

$$\Gamma_N(q_N^0, t_0; q_N, t) = G_N^S(q_N^0, q_N; t - t_0) \tag{3.29}$$

for all positive integers N.

4. FORMULATION IN CONFIGURATION SPACE AND USUAL SECOND QUANTIZATION

Let us summarize the situation as it appears in the first-quantization formalism; we want in particular to underline a few results which will be employed in the second quantization.

Let $\Phi(q_N; \nu_N)$ be a Schrödinger wavefunction which describes (in configuration space) a collection of N noninteracting hard spheres of diameter "a", in the state ν_N. We use the notation $q_N \equiv \{\mathbf{x}_1, \cdots \mathbf{x}_N\}$. We shall not specify, for the time being, which statistics are appropriate to our hard spheres, and do not insist therefore on symmetry properties of $\Phi(q_N; \nu_N)$. We look for the set of square-integrable functions which

(a) are twice differentiable in R with respect to each variable, and are, there, solutions of the Schrödinger equation

$$-\left(\sum_{i=1}^{N} \nabla_i^2\right)\Phi(q_N; \nu_N) = E(\nu_N)\Phi(q_N; \nu_N), \tag{4.1}$$

$$\nabla_i^2 \equiv \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial Y_i^2} + \frac{\partial^2}{\partial z_i^2};$$

the domain R is defined by[6]

$$R\colon \{q_N\colon |\mathbf{x}_i - \mathbf{x}_j| > a, \quad \text{all pairs} \quad i, j\};$$

(b) satisfy the auxiliary condition

$$\Phi(q_N; \nu_N) = 0 \tag{4.2}$$

in the complement of R (indicated in what follows with $R^\perp$);

(c) are continuous at the boundary of $R^\perp(\partial R)$ and possess there derivatives in all directions.

We shall assume, without further analysis, that the set $\Phi(q_N; \nu_N)$ provides an orthonormal basis in the space of square-integrable functions satisfying (4.2).

It is now convenient to reduce the given problem to one in which a Schrödinger equation, valid over the entire E^{3N} (the product of the Euclidean three-dimensional spaces of the N variables), takes the place of Eq. (4.1) and of the auxiliary condition (4.2). For this purpose, we introduce the characteristic function C of the domain R

$$C(q_N) \equiv \begin{cases} 1 & \text{if} \quad q_N \in R, \\ 0 & \text{otherwise,} \end{cases} \qquad C^2 = C. \tag{4.3}$$

Let us also remark that $C(q_N) = \prod_{i<k} C(\mathbf{x}_i, \mathbf{x}_k)$. We shall use the same symbols C, R regardless of the number of variables on which they depend.

Notice now that, for those functions $\Phi(q_N)$ for which (4.2) is satisfied, we have

$$\Phi'(q_N) \equiv C(q_N)\Phi(q_N) = \Phi(q_N). \tag{4.4}$$

We shall indicate with R' the manifold of all square-integrable functions which satisfy conditions (b) and (c) as stated above.

Using (4.3) it can be shown that, if the function $\Phi(q_N; \nu_N) \in R'$ satisfies (4.1) in R, it also satisfies in E^{3N} the equation

$$\lim_{\epsilon \to 0} \left\{ -\sum_i \nabla_i^2 + \sum_{i=1}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} [\nabla_i C(\mathbf{x}, \mathbf{x}_k)]_\epsilon C^{-1}(\mathbf{x}_i, \mathbf{x}_k) \right.$$

$$\left. \times C(q_N)\nabla_i \right\} \Phi(q_N; \nu_N) = E(\nu_N)\Phi(q_N; \nu_N), \tag{4.5}$$

where

$$[\nabla_i C(\mathbf{x}_i, \mathbf{x}_k)]_\epsilon \equiv \frac{\mathbf{x}_i - \mathbf{x}_k}{a}\, \delta(|\mathbf{x}_i - \mathbf{x}_k| - a - \epsilon), \qquad \epsilon > 0, \tag{4.6}$$

and

$$C^{-1}(\mathbf{x}_i, \mathbf{x}_k)C(q_N) \equiv \prod_{i \neq j \neq h \neq k} C(\mathbf{x}_j \mathbf{x}_h) \prod_{j \neq i,k} C(\mathbf{x}_i, \mathbf{x}_k) \prod_{h \neq k,i} C(\mathbf{x}_i, \mathbf{x}_h).$$

Conversely, if (4.5) is satisfied everywhere in E^{3N}, then $\Phi(q_N; \nu_N) \in R'$ and is a solution of (4.1) in R.

We shall not give a detailed proof of this here, and limit ourselves to make it plausible with the following heuristic argument.

The second term on the right-hand side of (5) expresses the fact that, going through a point

[6] This domain was called $\Omega_R{}^N$ in the previous paragraphs.

$P \in \partial R$, along a direction $\mathbf{n}$, pointing towards $R^{\perp}$, the function $\Phi(q_N; \nu_N)$ experiences a discontinuity in the derivative precisely equal to its derivative at P in the direction $-\mathbf{n}$. Therefore the derivative at P in the direction $\mathbf{n}$ is zero.

In $R^{\perp}$, only the first term on the left-hand side of (4.5) contributes; the derivative remains therefore zero throughout $R^{\perp}$, i.e., $\Phi(q_N; \nu_N)$ is constant in $R^{\perp}$. But then, in $R^{\perp}$, the left-hand side of (4.5) is zero and we conclude $\Phi(q_N; \nu_N) = 0$ in $R^{\perp}$.[7]

$\Phi(q_N; \nu_N)$ is continuous at the boundary of R; this follows from the fact that we have assumed that the derivatives of $\Phi(q_N; \nu_N)$ exist on ∂R in all directions; Eq. (4.5) would not make sense otherwise.

That $\Phi(q_N; \nu_N)$ satisfies (4.1) at every point of R follows immediately from (4.5), noting that R is an open set.

Equation (4.5) which, as we have remarked, is valid on the entire E^{3N}, will from now on substitute (4.1) and the auxiliary condition (4.2).

At this stage of the game, one would introduce the standard formalism of second quantization and thereby define "creation" and "annihilation" operators on a certain Hilbert space $\mathfrak{H}$; these operators are required to satisfy the usual (canonical) commutation relations. All the results of the first-quantization scheme can then be reproduced, provided one exercises proper care in the choice of the Hamiltonian.

To be sure, one has also to require suitable transformation properties under those groups which left the original Schrödinger theory invariant. The auxiliary condition (4.2) is foreign to these manipulations; it is then introduced at the end as a constraint that a vector in $\mathfrak{H}$ must satisfy in order to correspond to a physically admissible state.

Let us emphasize that, in the standard scheme, the condition (4.2) is irrelevant to the algebra of operators, which is determined by the "free field" commutation relations. If we now construct a Hilbert space $\mathfrak{H}$, operating with these operators on the vacuum, we shall find that some of the vectors in $\mathfrak{H}$ are "unphysical," since they should be associated with states in which two (or more) hard spheres overlap. It may be therefore interesting to give a scheme in which only physical states are used. This formulation may, in fact, prove particularly convenient if one wants to work in momentum space, where the criterion for selecting physical states has a somewhat complicated appearance. It may also turn out to be useful in the investigation of the meaning of commonly accepted approximations.

Before we go on to this formulation, let us recall a few formal features of the usual second-quantization scheme. One introduces creation operators $\psi^{\dagger}(\mathbf{x})$[8] and annihilation operators $\psi_0(\mathbf{x})$, which provide an irreducible representation of the commutation relations

$$[\psi_0(\mathbf{x}), \psi_0^{\dagger}(\mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}) \tag{4.7}$$

on a Hilbert space $\mathfrak{H}$ to which belongs by assumption the "vacuum" state $|0\rangle$, characterized by $\psi_0(\mathbf{x})\,|0\rangle = 0$, all x's.

We shall treat from now on the case of hard spheres satisfying Bose–Einstein statistics; what will be said remains valid, with obvious modifications, for hard spheres satisfying Fermi–Dirac statistics.

Any vector $|\nu\rangle$ in $\mathfrak{H}$ can be written in the form

$$|\nu\rangle = \sum_{N=0}^{\infty} \int \Phi(q_N; \nu_N)\, |q_N\rangle\, d^3q_N, \tag{4.8}$$

where[9]

$$|q_N\rangle \equiv (N!)^{-\frac{1}{2}} \psi_0^{\dagger}(\mathbf{x}_1) \cdots \psi_0^{\dagger}(\mathbf{x}_N)\, |0\rangle.$$

We shall define

$$P_N(q_N) \equiv |q_N\rangle\langle q_N|, \tag{4.9}$$

and introduce the projection operator P:

$$P = \sum_{N=0}^{\infty} \int d^3q_N C(q_N) P(q_N), \tag{4.10}$$

with $C(q_N)$ defined as in (4.3). One can then easily check that condition (4.2) is equivalent to

$$|\nu\rangle \in \mathfrak{H}_P \quad \text{where} \quad \mathfrak{H}_P = P\mathfrak{H}. \tag{4.11}$$

Equation (4.11) is the condition that a vector in $\mathfrak{H}$ has to satisfy in order to be associated with a physically admissible configuration, i.e., to be characterized—in the x representation—by a wave-function

$\Phi(q_N; \nu_N) = \langle q \mid \nu_N \rangle$ which satisfies (4.2). As Hamiltonian, we choose

$$H = \int \nabla\psi_0^{\dagger}(\mathbf{x}) \cdot \nabla\psi_0(\mathbf{x})\, d^3x. \tag{4.12}$$

This choice is justified noting that, from (4.7), (4.8), (4.12), the following equality is obtained:

[7] Note that $E(\nu_N) > 0$ in hard-sphere system without interaction.

[8] The subscript is here used only to differentiate these operators from the $\psi(\mathbf{x})$ used in Sec. 2 and 3, and later in Sec. 5.

[9] We make free use of improper vectors; one can imagine enclosing the system in a box, and letting the volume of the box go to infinity in the end.

$$H\,|\nu_N\rangle = \sum_{N=0}^{\infty} \int d^3q_N \times \lim_{\epsilon\to 0} \{-\sum_i \nabla_i^2 + \sum_i \sum_{k\neq i} [\nabla_i C(\mathbf{x}_i, \mathbf{x}_k)]\cdot \times C^{-1}(\mathbf{x}_i, \mathbf{x}_k) C(q_N)\}\Phi(q_N; \nu_N)\,|q_N\rangle \tag{4.13}$$

if $\Phi(q_N; \nu_N) \in R'$, i.e., $C(q_N)\Phi(q_N; \nu_N) = \Phi(q_N; \nu_N)$, and $\Phi(q_N; \nu_N)$ is continuous on ∂R and has there derivatives in all directions. The Schrödinger equation (4.1) together with the auxiliary condition (4.2) is therefore equivalent to

$$H\,|\nu_N\rangle = E(\nu_N)\,|\nu_N\rangle, \tag{4.14}$$

with H given by (4.12) and $|\nu_N\rangle \in \mathfrak{H}_P$. We also notice that

$$[H, P]\,|\nu_N\rangle = 0 \tag{4.14'}$$

if $|\nu_N\rangle$ is an eigenstate of H belonging to $\mathfrak{H}_P$. We shall assume, as mentioned above, that the eigenstates of H belonging to $\mathfrak{H}_P$, are discrete and that they provide an orthonormal basis in $\mathfrak{H}_P$.

We shall find useful, in what follows, to describe the dynamics of the hard-sphere system using, instead of H, the operator

$$H' = PHP. \tag{4.15}$$

Evidently, if $|\nu\rangle \in \mathfrak{H}_P$ and $H\,|\nu\rangle \in \mathfrak{H}_P$, then $H\,|\nu\rangle = H'\,|\nu\rangle$. Also, from (4.15) it follows that the eigenstates of H which belong to $\mathfrak{H}_P$ are also eigenstates of H'. We have already assumed that the eigenstates of H belonging to $\mathfrak{H}_P$ provide there a complete basis. Therefore the eigenstates of H' are also eigenstates of H belonging to $\mathfrak{H}_P$ and the two Hamiltonians are "equivalent" on $\mathfrak{H}_P$.

In fact, let $|\mu_N\rangle$ be an eigenstate of H':

$$H'\,|\mu_N\rangle = \epsilon(\mu_N)\,|\mu_N\rangle. \tag{4.16}$$

Let $|\nu_N\rangle$ be the eigenstates of H in $\mathfrak{H}_P$ and let $|\mu_N\rangle \neq \alpha\,|\nu_N\rangle$, α a number, for all ν_N's. Since the $|\nu_N\rangle$ form a complete set,

$$|\mu_N\rangle = \sum_{N'} C_{N'}\,|\nu_{N'}\rangle;$$

$$H'\,|\mu_N\rangle = \sum_{N'} C_{N'} E(\nu_{N'})\,|\nu_{N'}\rangle. \tag{4.16'}$$

In fact,

$$\langle\nu_{N'}|\,H'\,|\mu_N\rangle = E(\nu_{N'})\langle\nu_{N'}\,|\,\mu_N\rangle = E(\nu_{N'})C_{N'}.$$

From (4.16) and (4.16') one derives

$$\sum_{N'} [E(\nu_{N'}) - \epsilon(\mu_N)]C_{N'}\,|\nu_{N'}\rangle = 0. \tag{4.17}$$

However, by asumption,

$$E(\nu_N) \neq \epsilon(\mu_N)$$

for all μ's (the case of denumerable degeneracy can be treated in a similar way). Equation (4.17) implies then $C_{N'} = 0$, all N''s, and therefore $|\mu_N\rangle = 0$. Using (4.15) we shall write the equation of motion for the state $|\nu\rangle$ as

$$i(\partial/\partial t)\,|\nu\rangle = H'\,|\nu\rangle$$

provided the right-hand side is defined. To pass to the Heisenberg representation, we set

$$|\nu\rangle_H = e^{-iH't}\,|\nu\rangle; \qquad \psi(\mathbf{x}, t) = e^{-iH't}\psi(\mathbf{x})e^{iH't}.$$

Since $[H', P] = 0$, the condition for a state vector in the Heisenberg representation to represent the "history" of a physical system is

$$P\,|\nu\rangle_H = |\nu\rangle_H.$$

The equation of motion for any operator $\mathcal{O}(\mathbf{x}, t)$ is

$$i(\partial/\partial t)\mathcal{O}(\mathbf{x}, t) = [\mathcal{O}(\mathbf{x}, t), H']. \tag{4.18}$$

5. FORMULATION IN TERMS OF A NEW ALGEBRA

We shall now reformulate the alternative approach to second quantization for a gas of hard spheres, already introduced in I and in Sec. 2, working now consistently within an operator formalism. The advantages of this approach have been mentioned before; the main disadvantage, from the point of view of this section, is the appearance of a "free-field" algebra, far more complicated than the one usually accepted. Changing the commutation relations may seem a somewhat drastic step, but an attempt to deal with the hard-sphere problem using standard commutation relations led [E. Lieb, Proc. Natl. Acad. Sci. U.S. 47, 1000 (1960)] to a formalism in which the Hamiltonian does not exist (as an operator).

Let us introduce in $\mathfrak{H}$ the following operators:

$$\psi(\mathbf{x}) = P\psi_0(\mathbf{x})P, \tag{5.1}$$

$$\psi^\dagger(\mathbf{x}) = P\psi_0^\dagger(\mathbf{x})P, \tag{5.1'}$$

with P defined in (4.9), (4.10). These operators will turn out to have precisely the same matrix elements as the matrix operators introduced in I; this will justify *a posteriori* the use of the same symbol. Using the definition of P one can easily check that the following identities hold:

$$P\psi_0(\mathbf{x})P = \psi_0(\mathbf{x})P, \tag{5.2}$$

$$P\psi_0^\dagger(\mathbf{x})P = P\psi_0^\dagger(\mathbf{x}). \tag{5.2'}$$

Equation (5.2') shows that the $\psi^\dagger(\mathbf{x})$'s form in $\mathfrak{H}_P$ a cyclic representation[10] of some algebra (to be

[10] A set $\{A_i\}$ of operators on a Hilbert space $\mathfrak{K}$ is called cyclic with respect to a vector $\Phi_0 \in \mathfrak{K}$ if the vectors $P(A_i)\Phi_0$ span the entire Hilbert space, when $P(A_i)$ varies over all the polynomials in the A_i's.

exhibited later) with the vacuum as cyclic vector. In other words, the states

$$(N!)^{-\frac{1}{2}}\psi^\dagger(\mathbf{x}_1)\cdots\psi^\dagger(\mathbf{x}_N)\,|0\rangle$$

form a complete basis in $\mathfrak{H}_P$.

The proof goes as follows: we know that any state $|\nu\rangle$ can be approximated arbitrarily by states of the form

$$\int dq_N \Phi_N(\mathbf{x}_1, \cdots \mathbf{x}_N)\psi_0^\dagger(\mathbf{x}_1)\cdots\psi_0^\dagger(\mathbf{x}_N)\,|0\rangle,$$

i.e.,

$$\Big\| |\nu\rangle - \sum_{N=0}^{M}\int dq_N\Phi_N(\mathbf{x}_1, \cdots \mathbf{x}_N) \times \psi_0^\dagger(\mathbf{x}_1)\cdots\psi_0^\dagger(\mathbf{x}_N)\,|0\rangle\Big\| < \epsilon$$

if M is large enough. But, for any vector $|\rho\rangle$ and any projection operator P, $\|P\,|\rho\rangle\| \leq \|\,|\rho\rangle\|$. Therefore,

$$\Big\|P\,|\nu\rangle - \sum_N \int dq_N\Phi_N(\mathbf{x}_1\cdots\mathbf{x}_N) \times P\psi_0^\dagger(\mathbf{x}_1)\cdots\psi_0^\dagger(\mathbf{x}_N)\,|0\rangle\Big\| < \epsilon$$

for M large enough. The interchange of the operator P with the operation $\sum_N \int$ is justified since P is a bounded operator. If $|\nu\rangle \in \mathfrak{H}_P$, $P\,|\nu\rangle = |\nu\rangle$. Using (5.2′) and the identities $P^2 = P$ and $P\,|0\rangle = |0\rangle$, one obtains

$$\Big\| |\nu\rangle - \sum_N \int dq_N\Phi_N(\mathbf{x}_1\cdots\mathbf{x}_N) \times \psi^\dagger(\mathbf{x}_1)\cdots\psi^\dagger(\mathbf{x}_N)\,|0\rangle\Big\| < \epsilon$$

if $|\nu\rangle \in \mathfrak{H}_P$, which proves our assertion.

For various purposes (e.g., to introduce a particle interpretation), it is useful to know that the representation is also irreducible, i.e., that no proper subspace of $\mathfrak{H}_P$ is left invariant by all $\psi(\mathbf{x})$, $\psi^\dagger(\mathbf{x})$. To prove this, suppose $\mathfrak{H}_P'$ were such subspace. Then

$$\psi(\mathbf{x})\mathfrak{H}_P' \subseteq \mathfrak{H}_P', \qquad \psi^\dagger(\mathbf{x})\mathfrak{H}_P' \subseteq \mathfrak{H}_P', \tag{5.3}$$

i.e.,

$$P\psi_0(\mathbf{x})P\mathfrak{H}_P' \subseteq \mathfrak{H}_P', \qquad P\psi_0^\dagger(\mathbf{x})P\mathfrak{H}_P' \subseteq \mathfrak{H}_P'. \tag{5.4}$$

But $P\mathfrak{H}_P' = \mathfrak{H}_P'$ since $\mathfrak{H}_P' \subset \mathfrak{H}_P$. Equation (5.4) reads therefore

$$\psi_0(\mathbf{x})\mathfrak{H}_P' \subseteq \mathfrak{H}_P', \qquad P\psi_0^\dagger(\mathbf{x})\mathfrak{H}_P \subseteq \mathfrak{H}_P. \tag{5.4′}$$

Let $|\chi\rangle \in \mathfrak{H}_P'$; the second relation of (5.4′) implies then

$$\psi_0^\dagger(\mathbf{x})\,|\chi_\beta = \alpha\,|\chi_1\rangle + \beta\,|\chi_2\rangle,$$

where

$$|\chi_1\rangle \in \mathfrak{H}_P', \qquad |\chi_2\rangle \in \bar{\mathfrak{H}}_P',$$

$\bar{\mathfrak{H}}_P' = \mathfrak{H}_P'' - \mathfrak{H}_P'$, and $\mathfrak{H}_P''$ is the set of all vectors which can be obtained from a vector in $\mathfrak{H}_P'$ through a finite number of applications of $\psi_0^\dagger(\mathbf{x})$, and their closure. Notice that (5.4′) implies $\bar{\mathfrak{H}}_P' \subset \mathfrak{H}_P^\perp$. Then $\mathfrak{H}_P' \cup \bar{\mathfrak{H}}_P'$ is a subspace of $\mathfrak{H}$ invariant under all the $\psi_0(\mathbf{x})$'s and $\psi_0^\dagger(\mathbf{x})$'s. Since the representation they provide is, by assumption, irreducible, we conclude $\mathfrak{H}_P' \cup \mathfrak{H}_P$ equals either $\mathfrak{H}$ or $\underline{0}$. In the latter case, $\mathfrak{H}_P' = 0$. In the former, since $\bar{\mathfrak{H}}_P' \subseteq \mathfrak{H}_P^\perp$ one has $\mathfrak{H}_P' = \mathfrak{H}_P$. Q.E.D.

We shall now derive the explicit form of the algebra, a representation of which is provided by the operators $\psi(\mathbf{x})$, $\psi^\dagger(\mathbf{x})$. Using (5.2), (5.2′), and the standard commutation relations for $\psi_0(\mathbf{x})$, $\psi_0^\dagger(\mathbf{x})$, one gets

$$[\psi(\mathbf{x}), \psi(\mathbf{y})] = 0 = [\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{y})], \tag{5.5}$$

$$\psi(\mathbf{x})\psi(\mathbf{y}) = P(\mathbf{x}, \mathbf{y})\psi_0(\mathbf{x})\psi_0(\mathbf{y}), \tag{5.6}$$

$$P(\mathbf{x}, \mathbf{y}) = \sum_N \int C(\mathbf{x}, \mathbf{y}, q_N)P_N(q_N)\,dq_N,$$

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \psi_0(\mathbf{x})P\psi_0^\dagger(\mathbf{y}) - P\psi_0^\dagger(\mathbf{y})\psi_0(\mathbf{x})P$$
$$= P\delta(\mathbf{x} - \mathbf{y}) + P\psi_0(\mathbf{x})(P - 1)\psi_0^\dagger(\mathbf{y})P. \tag{5.7}$$

If we could write the last term in (5.7) as an explicit function of $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$, we would have succeeded in exhibiting the algebra of the $\psi(\mathbf{x})$'s. It is of course rather simple to express P in terms of $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$; the problem of writing in a similar form expressions like $\psi_0(\mathbf{x})(1 - P)\psi_0^\dagger(\mathbf{x})$ is, however, more involved. Some more manipulations are therefore required (and are given below), before the new free-field algebra may be exhibited.

Before we introduce new definitions, however, let us describe the dynamics of a collection of non-interacting hard spheres in terms of $\psi(\mathbf{x})$, $\psi^\dagger(\mathbf{x})$. We have already seen that a Hamiltonian which is equivalent, within the physical states, to (4.1) *and* (4.2) is

$$H' = P\int \nabla\psi_0^\dagger(\mathbf{x})\cdot\nabla\psi_0(\mathbf{x})P\,d^3x. \tag{5.8}$$

With our definition of $\psi(\mathbf{x})$, (5.8) can be written as

$$\int \nabla(P\psi_0^\dagger(\mathbf{x}))\cdot\nabla(\psi_0(\mathbf{x})P)\,d^3x = \int \nabla\psi^\dagger(\mathbf{x})\cdot\nabla\psi(\mathbf{x})\,d^3x \equiv H_{\cdot}. \tag{5.8}$$

Notice that H_s is written completely as a function of the elements of our algebra. In passing from (5.8) to (5.9), use has been made of the identities

$$\nabla(\psi_0(\mathbf{x})P) = (\nabla\psi_0(\mathbf{x}))\cdot P. \tag{5.10}$$

To justify this, consider the expression

$$\int d^3x\langle\rho|\ (\nabla\psi_0(\mathbf{x}))P\ |\chi\rangle\sigma(\mathbf{x}),$$

where $\sigma(\mathbf{x})$ is some "smooth" function of $\mathbf{x}$ and

$$|\rho\rangle,\quad |\chi\rangle \in \mathfrak{H}.$$

We have, if $|\chi\rangle \in \mathfrak{H}_P$,

$$\begin{aligned}
&\int d^3x\ \langle\rho|\ \nabla(\psi_0(\mathbf{x})P)\ |\chi\rangle\sigma(\mathbf{x}) \\
&\quad = -\int d^3x\ \langle\rho|\ \psi_0(\mathbf{x})P\ |\chi\rangle\ \nabla\cdot\sigma(\mathbf{x}) \\
&\quad = -\int d^3x\ \langle\rho|\ \psi_0(\mathbf{x})\ |\chi\rangle\ \nabla\cdot\sigma(\mathbf{x}) \\
&\quad = \int d^3x\ \langle\rho|\ \nabla\cdot\psi_0(\mathbf{x})\ |\chi\rangle\cdot\sigma(\mathbf{x}) \\
&\quad = \int d^3x\ \langle\rho|\ \nabla\psi_0(\mathbf{x})\cdot P\ |\chi\rangle\cdot\sigma(\mathbf{x}).
\end{aligned} \tag{5.11}$$

If $|\chi\rangle \in \mathfrak{H} - \mathfrak{H}_P$,

$$\begin{aligned}
&\langle\rho|\ \nabla(\psi_0(\mathbf{x})P)\ |\chi\rangle = 0 \\
&\qquad = \langle\rho|\ (\nabla\cdot\psi_0(\mathbf{x}))P\ |\chi\rangle.
\end{aligned} \tag{5.12}$$

Q.E.D.

Let us now introduce one more projection operator, so defined:

$$P(\mathbf{x}) = \sum_{N=1}^{\infty}\int dq_N C(\mathbf{x}, q_N)P_N(q_N), \tag{5.13}$$

with P_N, C defined in (4.9), (4.3). The meaning of $P(\mathbf{x})$ is easily found: $P(\mathbf{x})$ selects those states that represent a system in an allowed configuration ($|\mathbf{x}_i - \mathbf{x}_j| > a$, all i's and j's) such that one more hard sphere of diameter a can be added at the point $\mathbf{x}$ (and at a fixed time, that we choose equal to zero) still yielding an allowed configuration.

From (5.13) it follows that

$$P\cdot P(\mathbf{x}) = P(\mathbf{x})\cdot P = P(\mathbf{x}). \tag{5.14}$$

One moment's thought (or a straightforward computation) makes the following identities clear:

$$\psi(\mathbf{x}) \equiv P\psi_0(\mathbf{x})P = P(\mathbf{x})\psi_0(\mathbf{x}), \tag{5.15}$$

$$\psi^\dagger(\mathbf{x}) \equiv P\psi_0^\dagger(\mathbf{x})P = \psi_0^\dagger(\mathbf{x})P(\mathbf{x}), \tag{5.15'}$$

$$\psi_0(\mathbf{x})P(\mathbf{y}) = C(\mathbf{x}, \mathbf{y})P(\mathbf{y})\psi_0(\mathbf{x}),^{11} \tag{5.16}$$

[11] Note that

$$\begin{aligned}
P(\mathbf{x}, \mathbf{y}) &= C(\mathbf{x}, \mathbf{y})P(\mathbf{x})P(\mathbf{y}), \\
P(\mathbf{x})\psi_0^\dagger(\mathbf{y}) &= C(\mathbf{y}, \mathbf{x})\psi_0^\dagger(\mathbf{y})P(\mathbf{x}).
\end{aligned} \tag{5.16'}$$

Using these identities, one then has

$$\begin{aligned}
[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] &= \delta(\mathbf{x} - \mathbf{y})P(\mathbf{x}) \\
&\quad - [1 - C(\mathbf{x}, \mathbf{y})]\psi^\dagger(\mathbf{y})\psi(\mathbf{x}).
\end{aligned} \tag{5.17}$$

Proof:

$$\begin{aligned}
[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] &= P(\mathbf{x})[\psi_0(\mathbf{x}), \psi_0^\dagger(\mathbf{y})]P(\mathbf{y}) \\
&\quad - \psi^\dagger(\mathbf{y})\psi(\mathbf{x}) + P(\mathbf{x})\psi_0^\dagger(\mathbf{y})\psi_0(\mathbf{x})P(\mathbf{y}) \\
&= P(\mathbf{x})\delta(\mathbf{x} - \mathbf{y}) - \psi^\dagger(\mathbf{y})\psi(\mathbf{x}) \\
&\quad + C(\mathbf{x}, \mathbf{y})\psi_0^\dagger(\mathbf{y})P(\mathbf{x})P(\mathbf{y})\psi_0(\mathbf{x}) \\
&= P(\mathbf{x})\delta(\mathbf{x} - \mathbf{y}) - \psi^\dagger(\mathbf{y})\psi(\mathbf{x}) + C(\mathbf{x}, \mathbf{y})\psi^\dagger(\mathbf{y})\psi(\mathbf{x}),
\end{aligned}$$

where use has been made of the identities

$$\begin{aligned}
P(\mathbf{x})P(\mathbf{y}) &= P(\mathbf{y})P(\mathbf{x}), \\
P^2(\mathbf{x}) &= P(\mathbf{x}).
\end{aligned} \tag{5.18}$$

Let us remark that, although a new projection operator has been introduced, we did not change the definition of $\psi(\mathbf{x})$. In particular, the dynamics is still given by the Hamiltonian (5.8).

We shall now express $P(\mathbf{x})$ as a function of $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$. One has, formally,[12]

$$P(\mathbf{x}) = 0^{\int_{|\mathbf{z}-\mathbf{x}|\le a}\psi^\dagger(\mathbf{z})\psi(\mathbf{z})\,d^3z} \tag{5.19}$$

where the symbol 0^A, A being an operator, is defined by the following rule: choose a representation in which A is diagonal. If A has only nonnegative eigenvalues (as is in our case), then 0^A is defined by

$$0^A\ |a\rangle = \begin{cases} |a\rangle & \text{if } a = 0, \\ 0 & \text{if } a > 0; \end{cases} \qquad A\ |a\rangle = a\ |a\rangle. \tag{5.20}$$

Expression (5.17) has thereby been written completely in terms of $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$.

We can now forget its derivation as well as the expression of $\psi(\mathbf{x})$, $\psi^\dagger(\mathbf{x})$ as a function of $\psi_0(\mathbf{x})$ and use (5.17) with (5.5) as the definition of the "free-field algebra" (referred to in what follows as S algebra), appropriate to the hard-sphere problem. A gas of free hard spheres is then described using an irreducible representation of the S algebra; the "vacuum" may be selected as cyclic vector to construct $\mathfrak{H}_P$, the Hilbert space on which the elements of the representation operate. All vectors in $\mathfrak{H}_P$ represent physically permissible states, and no auxiliary condition like (4.2) is introduced.

The evolution in time of the system is exhibited by a unitary transformation $U(t)$, the infinitesimal generator of which is H_s, introduced above in Eq. (5.9).

[12] The formal proof of (5.19) is given in the Appendix.

One has, for every operator A,

$$i(\partial A/\partial t) = [A, H_s]. \tag{5.21}$$

In particular, choosing $A = \psi(\mathbf{x}, t)$ and using in (5.21) the explicit form of H_s and (5.17), one obtains

$$i\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -P(\mathbf{x})\nabla^2\psi(\mathbf{x}, t) + \lim_{\epsilon\to 0+}\int_{|\mathbf{y}-\mathbf{x}|=a+\epsilon} d\sigma_y$$

$$\times\frac{\mathbf{y}-\mathbf{x}}{a}\psi^\dagger(\mathbf{y}, t)\cdot\nabla_y\psi(\mathbf{y}, t)\psi(\mathbf{x}, t). \tag{5.22}$$

For completeness, we want now to give explicitly, as a function of the ψ's, a representation on $\mathfrak{H}_P$ of the group of rotations and translations in three dimensions. We shall restrict ourselves to the latter group, the former being treated in precisely the same fashion. On the Hilbert space $\mathfrak{H}$ of the usual second quantization, a representation is provided by the unitary operator

$$U(\mathbf{a}) = \exp(i\boldsymbol{\Pi}\cdot\mathbf{a}),$$

$$\boldsymbol{\Pi} = i\int\psi_0^\dagger(\mathbf{x})\nabla\psi_0(\mathbf{x})\,d^3x. \tag{5.23}$$

One has

$$\psi_0(\mathbf{x}+\mathbf{a}) = U(\mathbf{a})\psi_0(\mathbf{x})U^\dagger(\mathbf{a}),$$

and also

$$P(\mathbf{x}+\mathbf{a}) = U(\mathbf{a})P(\mathbf{x})U^\dagger(\mathbf{a}),$$

as one can easily see, e.g., from Eq. (II.3) (cf. Appendix II). Therefore,

$$\psi(\mathbf{x}+\mathbf{a}) = U(\mathbf{a})\psi(\mathbf{x})U^\dagger(\mathbf{a}). \tag{5.24}$$

We remark now that

$$[P, U(\mathbf{a})] = 0 \tag{5.25}$$

for all $\mathbf{a}$. In fact,

$$U(\mathbf{a})PU^\dagger(\mathbf{a}) = \sum_{N=0}^{\infty} d^3q_N C(q_N)U(\mathbf{a})P_N(q_N)U^\dagger(\mathbf{a})$$

$$= \sum_{N=0}^{\infty}\int d^3q_N C(q_N+\mathbf{a})P_N(q_N+\mathbf{a}), \tag{5.26}$$

where we have used

$$C(q_N+\mathbf{a}) \equiv C(\mathbf{x}_1+\mathbf{a}, \mathbf{x}_2+\mathbf{a}, \cdots \mathbf{x}_N+\mathbf{a}) = C(q_N)$$

and

$$U(\mathbf{a})P_N(q_N)U^\dagger(\mathbf{a}) = U(\mathbf{a})\,|q_N\rangle\langle q_N|\,U^\dagger(\mathbf{a})$$

$$= |q_N+\mathbf{a}\rangle\langle q_N+\mathbf{a}| = P_N(q_N+\mathbf{a}).$$

From (5.25) it follows

$$[P, \boldsymbol{\Pi}] = 0. \tag{5.27}$$

This could have been proved also directly, by noting that

$$\sum_{i=1}^{N}\frac{\partial}{\partial\mathbf{x}_i}C(\mathbf{x}_1\cdots\mathbf{x}_N) = 0,$$

as C is a function only of the differences $\mathbf{x}_i - \mathbf{x}_j$. Consider now the operator

$$U_P(\mathbf{a}) \equiv PU(\mathbf{a})P. \tag{5.28}$$

It satisfies

$$U_P(\mathbf{a})U_P^\dagger(\mathbf{a}) = P = U_P^\dagger(\mathbf{a})U_P(\mathbf{a}). \tag{5.29}$$

If restricted to $\mathfrak{H}_P$, it is, therefore, unitary. Also,

$$\psi(\mathbf{x}+\mathbf{a}) = U(\mathbf{a})\psi(\mathbf{x})U^\dagger(\mathbf{a}) = U(\mathbf{a})P\psi_0(\mathbf{x})PU^\dagger(\mathbf{a})$$

$$= U_P(\mathbf{a})P\psi_0(\mathbf{x})PU_P^\dagger(\mathbf{a}) = U_P(\mathbf{a})\psi(\mathbf{a})U_P^\dagger(\mathbf{x}). \tag{5.30}$$

Let

$$U_p(\mathbf{a}) = \exp(i\boldsymbol{\Pi}_p\cdot\mathbf{a}), \tag{5.31}$$

where

$$\boldsymbol{\Pi}_P = P\boldsymbol{\Pi}P = iP\int\psi_0^\dagger(\mathbf{x})\nabla\psi_0(\mathbf{x})\,d^3x. \tag{5.32}$$

Using (5.10), (5.2), and (5.2′), one obtains

$$\boldsymbol{\Pi}_P = i\int\psi^\dagger(\mathbf{x})\nabla\psi(\mathbf{x})\,d^3x. \tag{5.33}$$

Equations (5.30), (5.31), and (5.33) exhibit, as a function of the ψ's, a unitary representation on $\mathfrak{H}_P$ of the group of translations in three dimensions. For infinitesimal values of the parameter $\mathbf{a}$, Eqs. (5.30) and (5.31) give

$$i[\partial\psi(\mathbf{x})/\partial\mathbf{x}] = [\boldsymbol{\Pi}_P, \psi(\mathbf{x})]. \tag{5.34}$$

Introducing the Fourier transform of $\psi(\mathbf{x})$,

$$\tilde{\psi}(\mathbf{p}) = \int e^{i\mathbf{p}\cdot\mathbf{x}}\psi(\mathbf{x})\,d^3x,$$

Eq. (5.34) can be rewritten as

$$[\boldsymbol{\Pi}_P, \tilde{\psi}(\mathbf{k})] = -\mathbf{k}\tilde{\psi}(\mathbf{k}),$$
$$[\boldsymbol{\Pi}_P, \tilde{\psi}^\dagger(\mathbf{k})] = \mathbf{k}\tilde{\psi}^\dagger(\mathbf{k}). \tag{5.35}$$

Eq. (5.27) shows that the "total momentum" of a state is still a meaningful quantity, and Eq. (5.35) indicates that, e.g., $\tilde{\psi}^\dagger(\mathbf{k})$ does indeed "increase" the total momentum of a state by the amount $\mathbf{k}$. It should be kept in mind, however, that the momentum of a single particle is not an observable in $\mathfrak{H}_P$. In fact, let $|\mathbf{k}, 1\rangle$ be the state of one particle with momentum $\mathbf{k}$, and let $|\mathbf{p}, N\rangle \in \mathfrak{H}_P^N$ be a state describing N particles with total momentum

equal to $\mathbf{p}$. Then

$$|\mathbf{k}, 1\rangle \otimes |\mathbf{p}, N\rangle \notin \mathfrak{H}_P^{N+1}, \tag{5.36}$$

i.e., the state $|\mathbf{k}, 1\rangle \otimes |\mathbf{p}, N\rangle$ cannot occur as a physical state for a system of $N + 1$ particles.

Therefore, the operator $\psi^\dagger(\mathbf{k})$ is *not* a creation operator for a particle with momentum $\mathbf{k}$.

APPENDIX I

It must be possible to obtain Eq. (3.11) by a limiting process from an equation valid for finite potentials. There is, however, a slight difference between the case of a finite potential, however large, and the case of hard-sphere interaction.

For the case of a finite potential, one has

$$\lim_{\beta\to 0} G_N(q_N^0, q_N; \beta) = \delta(q_N - q_N^0), \tag{I.1}$$

while in the case of impenetrable cores, one has

$$G_N(q_N^0, q_N; \beta) = 0, \tag{I.2}$$

unless q_N^0 and q_N are in Ω_R^N, and, therefore,

$$\lim_{\beta\to 0} G_N(q_N^0, q_N; \beta) = \begin{cases} \delta(q_N - q_N^0) & \text{if } q_N, q_N^0 \in \Omega_R^N, \\ 0 & \text{otherwise.} \end{cases} \tag{I.3}$$

This difference means physically that for any finite potential the entire configuration space is accessible, while the idealization of impenetrable cores excludes part of the configuration space even in the limit of infinite temperature. This requires some precautions in performing the limiting process.

Let $G_N(q_N^0, q_N; \beta)$ and $G_N'(q_N^0, q_N; \beta)$ be Green's functions of the Bloch equation for N particles with interaction potentials $V(q_N)$ and $V'(q_N)$, respectively. (We will omit the subscript N in the following calculation.)

Consider the integral

$$I = \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int G(q_0, q'; t)v(q')e^{-(\beta-t)v(q')} \times G'(q', q; \beta - t)\, dq', \tag{I.4}$$

where α is a real positive number smaller than unity which we will let approach zero at the end; and $v(q) = V(q) - V'(q)$.

This integral can be transformed as follows:

$$\begin{aligned} I &= \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' G(q_0, q'; t)\left[\frac{\partial}{\partial t} e^{-(\beta-t)v(q')}\right] G'(q', q; \beta - t) \\ &= \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int \frac{\partial}{\partial t} [G(q_0, q'; t)e^{-(\beta-t)v(q')}G'(q', q; \beta - t)] \\ &\quad - \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' \Big[\frac{\partial}{\partial t} G(q_0, q', t)e^{-(\beta-t)v(q')}G'(q', q; \beta - t) \\ &\quad + G(q_0, q'; t)e^{-(\beta-t)v(q')} \frac{\partial}{\partial t} G'(q', q; \beta - t)\Big]. \end{aligned} \tag{I.5}$$

With the use of the Bloch equation, the second term can be written as

$$-\int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' e^{-(\beta-t)v(q')}\Big\{\Big[\sum_{i=1}^{N} \nabla_i'^2 - V(q')\Big]G(q_0, q'; t)G'(q', q; \beta - t) - G(q_0, q'; t)\Big[\sum_{i=1}^{N} \nabla_i'^2 - V'(q')\Big]G(q', q; \beta - t)\Big\}.$$

The terms with V and V' add up to I again and we have

$$\begin{aligned} 0 &= \int_{\alpha\beta}^{(1-\alpha)\beta} dt \frac{\partial}{\partial t} \int dq' G(q_0, q'; t)e^{-(\beta-t)v(q')}G'(q', q; \beta - t) - \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' e^{-(\beta-t)v(q')} \\ &\quad \times \Big[\sum_{i=1}^{N} \nabla_i'^2 G(q_0, q'; t)G'(q', q; \beta - t) - G(q_0, q'; t) \sum_{i=1}^{N} \nabla_i'^2 G'(q', q; \beta - t)\Big] \\ &= \int [G(q_0, q'; (1 - \alpha)\beta)e^{-\alpha\beta v(q')}G'(q', q; \alpha\beta) - G(q_0, q'; \alpha\beta)e^{-(1-\alpha)\beta v(q')}G'(q', q; (1 - \alpha)\beta)] \\ &\quad - \int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' e^{-(\beta-t)v(q')} \sum_i \nabla_i'\{[\nabla_i' G(q_0, q'; t)]G'(q', q; \beta - t) - G(q_0, q'; t)\nabla_i' G'(q', q; \beta - t)\}. \end{aligned} \tag{I.6}$$

The last term can be written as

$$-\int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' \sum_j \nabla_j' \{e^{-(\beta-t)v(q')}[(\nabla_j' G)\cdot G' - G\cdot\nabla_j' G']\}$$
$$+\int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' \sum_j \nabla_j' e^{-(\beta-t)v(q')}[(\nabla_j' G)\cdot G' - G\cdot\nabla_j' G'].$$

The first integral vanishes when Born–von Karman conditions are imposed on G, G', and v, and we have

$$0 = \int dq'[G(q_0, q'; (1-\alpha)\beta)e^{-\alpha\beta v(q')}G'(q', q; \alpha\beta) - G(q_0, q'; \alpha\beta)e^{-(1-\alpha)\beta v(q')}G'(q', q; (1-\alpha)\beta)]$$
$$+\int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq' \sum_{j=1}^{N} \nabla_j' e^{-(\beta-t)v(q')}[G'(q', q; \beta-t)\nabla_j' G(q_0, q'; t) - G(q_0, q'; t)\nabla_j' G'(q', q; \beta-t)]. \quad \text{(I.7)}$$

Now we specialize to the case

$$V(q) = \frac{1}{2}\sum_{j,k=1}^{N}{}' u(\mathbf{x}_j - \mathbf{x}_k),$$
$$V'(q) = \frac{1}{2}\sum_{j,k=1}^{N-1}{}' u(\mathbf{x}_j - \mathbf{x}_k),$$

where the prime indicates $j \neq k$, and

$$v(q) = \sum_{j=1}^{N-1} u(\mathbf{x}_j - \mathbf{x}_N),$$

and write to conform with the previous notation

$$G(q_0, q; t) = G_N(q_N^0, q_N; t),$$
$$G'(q_0, q; t) = G_{N-1}(q_{N-1}^0, q_{N-1}; t)G_1(\mathbf{x}_N^0, \mathbf{x}_N; t).$$

If we keep u finite and let α approach zero, we obtain

$$G_N(q_N^0, q_N; \beta) = G_{N-1}(q_{N-1}^0, q_{N-1}; \beta)G_1(\mathbf{x}_N^0, \mathbf{x}_N; \beta)\exp\left[-\beta\sum_{j=1}^{N-1} u(\mathbf{x}_j^0 - \mathbf{x}_N^0)\right]$$
$$-\int_0^\beta dt \int dq_N' \sum_{j=1}^{N} \nabla_j' \exp\left[-(\beta-t)\sum_{j=1}^{N-1} u(\mathbf{x}_j' - \mathbf{x}_N')\right]\{[\nabla_j' G_N(q_N^0, q_N'; t)]$$
$$\times G_{N-1}(q_{N-1}', q_{N-1}; \beta-t)G_1(\mathbf{x}_N', \mathbf{x}_N; \beta-t) - G_N(q_N^0, q_N'; t)$$
$$\times \nabla_j'[G_{N-1}(q_{N-1}', q_{N-1}; \beta-t)]G_1(\mathbf{x}_N', \mathbf{x}_N; \beta-t)\}. \quad \text{(I.8)}$$

In deriving the corresponding equation for the hard-sphere case, we have to start with Eq. (I.7) and perform the limit

$$\begin{aligned} u(\mathbf{x}_j - \mathbf{x}_k) &\to \infty \quad \text{for} \quad |\mathbf{x}_j - \mathbf{x}_k| \leq a, \\ u(\mathbf{x}_j - \mathbf{x}_k) &\to 0 \quad \text{for} \quad |\mathbf{x}_j - \mathbf{x}_k| > a \end{aligned} \quad \text{(I.9)}$$

first, with α and $\beta > 0$ fixed. We then obtain

$$0 = \int dq_N'\Big[G_N(q_N^0, q_N'; (1-\alpha)\beta)\prod_{j=1}^{N-1}\theta(\mathbf{x}_j' - \mathbf{x}_N')G_{N-1}(q_{N-1}', q_N; \alpha\beta)G_1(\mathbf{x}_N', \mathbf{x}_N; \alpha\beta)$$
$$-G_N(q_N^0, q_N'; \alpha\beta)\prod_{j=1}^{N-1}\theta(\mathbf{x}_j' - \mathbf{x}_N')G_{N-1}(q_{N-1}', q_N; (1-\alpha)\beta)G_1(\mathbf{x}_N', \mathbf{x}_N; (1-\alpha)\beta)\Big]$$
$$+\int_{\alpha\beta}^{(1-\alpha)\beta} dt \int dq_N' \sum_{j=1}^{N}\nabla_j' \prod_{l=1}^{N-1}\theta(\mathbf{x}_l' - \mathbf{x}_N')[(\nabla_j' G_N(q_N^0, q_N'; t))G_{N-1}(q_{N-1}', q_{N-1}; \beta-t)$$
$$\times G_1(\mathbf{x}_N', \mathbf{x}_N; \beta-t) - G_N(q_N^0, q_N; t)\nabla_j'(G_{N-1}(q_{N-1}', q_{N-1}; \beta-t))G_1(\mathbf{x}_N', \mathbf{x}_N; (\beta-t))], \quad \text{(I.10)}$$

where

$$\theta(\mathbf{x}_j - \mathbf{x}_N) = \begin{cases} 0 & \text{for} \quad |\mathbf{x}_j - \mathbf{x}_N| \leq a, \\ 1 & \text{for} \quad |\mathbf{x}_j - \mathbf{x}_N| > a. \end{cases} \quad \text{(I.11)}$$

Since the discontinuity of $\theta(\mathbf{x})$ occurs at $|\mathbf{x}| = a + 0$, we have

$$\nabla_i \theta(\mathbf{x}_i - \mathbf{x}_N) = \delta(|\mathbf{x}_i - \mathbf{x}_N| - a - 0)\frac{\mathbf{x}_i - \mathbf{x}_N}{a}, \quad \text{(I.12)}$$

$$\nabla_N \theta(\mathbf{x}_i - \mathbf{x}_N) = \delta(|\mathbf{x}_i - \mathbf{x}_N| - a - 0)\frac{\mathbf{x}_N - \mathbf{x}_i}{a},$$

and

$$\sum_{i=1}^{N} \nabla_i \prod_{l=1}^{N-1} \theta(\mathbf{x}_l - \mathbf{x}_N)\cdot\nabla_i = \sum_{j=1}^{N-1}\left[\prod_{\substack{l=1\\ l\neq j}}^{N-1} \theta(\mathbf{x}_l - \mathbf{x}_N)\right] \times \delta(|\mathbf{x}_j - \mathbf{x}_N| - a - 0)\frac{\mathbf{x}_j - \mathbf{x}_N}{a}(\nabla_j - \nabla_N). \quad \text{(I.13)}$$

Since, in the limit (I.9), the function $G_N(q_N^0, q_N'; t)$ vanishes if $|\mathbf{x}_i' - \mathbf{x}_N'| \leq a$, the factor $\prod_{i=1}^{N-1}\theta(\mathbf{x}_i' - \mathbf{x}_N')$ in the first term of the integrand in the first integral of Eq. (I.10) can be omitted. Under the assumption that, in the limit (I.9), $G_N(q_N^0, q_N; t) \to 0$ if, for any $j \neq N$, $|\mathbf{x}_j - \mathbf{x}_N| \to a$, the second term in the second integral in Eq. (I.10) can be omitted. Also, if $|\mathbf{x}_l - \mathbf{x}_N| \leq a$, for any $l \neq j$, then

$$[(\mathbf{x}_j - \mathbf{x}_N)/a](\nabla_j - \nabla_N)G_N(q_N^0, q_N; t) = 0,$$

so that the factor $\prod_{l=1}^{N-1}\theta(\mathbf{x}_l - \mathbf{x}_N)$ in Eq. (I.10) can be omitted. Taking the limit $\alpha \to 0$, we then obtain

$$G_N(q_N^0, q_N; \beta) = \prod_{j=1}^{N-1}\theta(\mathbf{x}_j^0 - \mathbf{x}_N^0)G_{N-1}(q_{N-1}^0, q_{N-1}; \beta) \times G_1(\mathbf{x}_N^0, \mathbf{x}_N; \beta) - \int_0^\beta dt \int dq_N' \times \sum_{j=1}^{N-1}\delta(|\mathbf{x}_j' - \mathbf{x}_N'| - a - 0) \times \left[\frac{\mathbf{x}_j' - \mathbf{x}_N'}{a}(\nabla_j' - \nabla_N')G_N(q_N^0, q_N'; t)\right] \times G_{N-1}(q_{N-1}', q_{N-1}; \beta - t)G_1(\mathbf{x}_N', \mathbf{x}_N; \beta - t), \quad \text{(I.14)}$$

for $\beta > 0$. For $q_N^0 \in \Omega_R^N$, this agrees with Eq. (3.11).

APPENDIX II

We want to prove here that

$$P(\mathbf{x}) = 0^{\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y \psi^\dagger(\mathbf{y})\psi(\mathbf{y})}.$$

We have

$$P = \sum_{N=0}^{\infty} P_N, \quad \text{(II.1)}$$

with

$$P_N = \left[\prod_{M>N}\left(1 - \frac{N}{M}\right)\right]^{-1} \times \left[\prod_{M>N}\left(1 - \frac{\int \psi_0^\dagger(\mathbf{x})\psi_0(\mathbf{x})\, d^3x}{M}\right)\right] \times \frac{1}{N!}\int C(q_N)\psi_0^\dagger(\mathbf{x}_1)\cdots\psi_0^\dagger(\mathbf{x}_N) \times \psi_0(\mathbf{x}_1)\cdots\psi_0(\mathbf{x}_N)\, d^3q_N, \quad \text{(II.2)}$$

and

$$P(\mathbf{x}) = \sum_{N=0}^{\infty} P_N R_N(\mathbf{x}) = \sum_{N=0}^{\infty}\sum_{M=0}^{\infty} P_N R_M(\mathbf{x}) = PR(\mathbf{x}) = \sum_{N=0}^{\infty} P_N(\mathbf{x}), \quad \text{(II.3)}$$

where

$$R_N(\mathbf{x}) = \left[\prod_{M>N}\left(1 - \frac{N}{M}\right)\right]^{-1} \times \left\{\prod_{M>N}\left[1 - \frac{\int \psi_0^\dagger(\mathbf{x})\psi_0(\mathbf{x})\, d^3x}{M}\right]\right\} \times \frac{1}{N!}\int_{|\mathbf{x}_l - \mathbf{x}_j|>a} d^3q_N \prod_i C(\mathbf{x}, \mathbf{x}_i)\psi_0^\dagger(\mathbf{x}_i)\psi_0(\mathbf{x}_i), \quad \text{(II.4)}$$

$$P_N(\mathbf{x}) = \left[\prod_{M>N}\left(1 - \frac{N}{M}\right)^{-1}\right] \times \left\{\prod_{M>N}\left[1 - \frac{\int \psi_0^\dagger(\mathbf{x})\psi_0(\mathbf{x})\, d^3x}{M}\right]\right\} \times \frac{1}{N!}\int_{|\mathbf{x}_l - \mathbf{x}_j|>a} d^3q_N \prod_i C(\mathbf{x}, \mathbf{x}_i)\psi^\dagger(\mathbf{x}_i)\psi(\mathbf{x}_i). \quad \text{(II.5)}$$

Let us now introduce in R_N' the following decomposition:

$R'_{N;0}$	is (the set of all functions belonging to R_N' and) such that $C(\mathbf{x}, \mathbf{x}_i)f = 0$, all i's;
$R'_{N;k}$	is such that $C(\mathbf{x}, \mathbf{x}_i)f = \delta_{ik}f$;
$R'_{\substack{N;k,j\\ k<j}}$	is such that $C(\mathbf{x}, \mathbf{x}_i)f = (\delta_{ik} + \delta_{ij})f$;
$\vdots$	$\vdots$
$R'_{\substack{N;k_1\cdots k_{N-1}\\ k_1<\cdots<k_{N-1}}}$	is such that $C(\mathbf{x}, \mathbf{x}_i)f = \left(\sum_{m=1}^{N-1}\delta_{ik_m}\right)f$;
$R'_{N;1\cdots N}$	is such that $C(\mathbf{x}, \mathbf{x}_i)f = f$, all i's.

We shall denote by $\{m\}_N$ any sequence of numbers $k_1 \cdots k_m$ taken among the first N cardinals and such that $k_i < k_j$ if $i < j$ $(i, j = 1 \cdots N)$. The sequence that does not contain any element is indicated with 0.

With these notations, R'_N is the direct sum of all $R'_{N;\{m\}_N}$, which are pairwise orthogonal. In fact, if $f(q_N)$ is a function in R'_N, its component in $R'_{N;\{m\}_N}$ is given by

$$P_{N;\{m\}_N} f(q_N) = \prod_{i\in\{m\}_N} C(\mathbf{x}, \mathbf{x}_i) \times \prod_{\substack{i\notin\{m\}_N \\ i\leq N}} [1 - C(\mathbf{x}, \mathbf{x}_i)] f(q_N). \quad \text{(II.6)}$$

$P_{N;\{m\}_N}$, as defined in (II.6), is evidently linear and one has

$$\sum_{\{m\}_N} P_{N;\{m\}_N} = \prod_{i=1}^{N} \{C(\mathbf{x}, \mathbf{x}_i) + [1 - C(\mathbf{x}, \mathbf{x}_i)]\} = 1, \quad \text{(II.7)}$$

$$P_{N;\{m\}_N} \cdot P_{N;\{m'\}_N} = 0 \quad \text{if} \quad \{m\}_N \neq \{m'\}_N \quad \text{(II.8)}$$

(if $\{m\}_N \neq \{m'\}_N$, $P_{N;\{m\}_N} \cdot P_{N;\{m'\}_N}$ contains at least one factor $C(\mathbf{x}, \mathbf{x}_i)\cdot[1 - C(\mathbf{x}, \mathbf{x}_i)] = 0$),

$$P^2_{N;\{m\}_N} = P_{N;\{m\}_N}. \quad \text{(II.8')}$$

We now divide $\mathfrak{H}_{P;N}$[13] into subspaces $\mathfrak{H}^{\{m\}_N}_{P;N}$ according to the rule

$$|\nu\rangle \in \mathfrak{H}^{\{m\}_N}_{P;N} \Leftrightarrow \langle q \mid \nu\rangle \equiv \Phi(q_N; \nu_N) \in R_{N;\{m\}_N}. \quad \text{(II.9)}$$

It follows from (II.8) that $\mathfrak{H}^{\{m\}_N}_{P;N} \cap \mathfrak{H}^{\{m'\}_N}_{P;N} = 0$ if $\{m\}_N \neq \{m'\}_N$, and that

$$\mathfrak{H}_{P;N} = \bigoplus_{\{m\}_N} \mathfrak{H}^{\{m\}_N}_{P;N}. \quad \text{(II.10)}$$

We now notice that both $P_N(\mathbf{x})$ and $0^{\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y\, \psi^\dagger(\mathbf{y})\psi(\mathbf{y})}$ leave each $\mathfrak{H}^{\{m\}}_{P;N}$ invariant or map it to zero; in fact,

$$P_N(\mathbf{x})\,|\nu\rangle \begin{cases} |\nu\rangle & \text{if} \quad |\nu\rangle \in \mathfrak{H}^{\{1\cdots N\}_N}_{P;N}, \\ 0 & \text{if} \quad |\nu\rangle \in \mathfrak{H}^{\{m\}_N}_{P;N}, \quad \{m\}_N \neq \{1 \cdots N\}_N, \end{cases} \quad \text{(II.11)}$$

and

$$\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y\, \psi^\dagger(\mathbf{y})\psi(\mathbf{y})\,|\nu\rangle = [N - P(\{m\}_N)]\,|\nu\rangle \quad \text{if} \quad |\nu\rangle \in \mathfrak{H}^{\{m\}_N}_{P;N}, \quad \text{(II.12)}$$

where $P(\{m\}_N)$ is the number of elements in the sequence $\{m\}$. From (II.12) we have, according to the definition given in the text,

$$0^{\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y\psi^\dagger(\mathbf{y})\psi(\mathbf{y})}\,|\nu\rangle = 0^{[N-P(\{m\}_N)]}\,|\nu\rangle = \begin{cases} |\nu\rangle & \text{if} \quad P(\{m\}_N) = N, \\ 0 & \text{otherwise.} \end{cases} \quad \text{(II.13)}$$

But

$$P(\{m\}_N) = N \quad \text{only if} \quad \{m\}_N = \{1 \cdots N\}_N.$$

Therefore,

$$0^{\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y\psi^\dagger(\mathbf{y})\psi(\mathbf{y})} = P_N(\mathbf{x}) \quad \text{on} \quad \mathfrak{H}_{P;N}. \quad \text{(II.14)}$$

One can also verify that

$$P_N(\mathbf{x})\,|\nu\rangle = 0 \quad \text{if} \quad |\nu\rangle \in \mathfrak{H}_{P,M}, \quad M \neq N. \quad \text{(II.15)}$$

From (II.3), (II.14), (II.15), and (II.10), we then conclude

$$0^{\int_{|\mathbf{y}-\mathbf{x}|\leq a} d^3y\psi^\dagger(\mathbf{y})\psi(\mathbf{y})} = P(\mathbf{x}) \quad \text{on} \quad \mathfrak{H}_P. \quad \text{(II.16)}$$

[13] $\mathfrak{K}_{P;N}$ is composed of those vectors $|\nu\rangle \in \mathfrak{K}_P$ such that $\Phi(q_N; \nu) \equiv \langle q \mid \nu\rangle$ is a function of N variables; one can easily prove that $\mathfrak{K}_{P;N} \cap \mathfrak{K}_{P;N'} = 0'$ for $N \neq N'$, and that $\mathfrak{K}_P = \oplus_N \mathfrak{K}_{P;N}$.

Causality Implies the Lorentz Group

E. C. ZEEMAN
Institut des Hautes Études Scientifiques, Bures-sur-Yvette, Seine et Oise, France
(Received 30 July 1963)

Causality is represented by a partial ordering on Minkowski space, and the group of all automorphisms that preserve this partial ordering is shown to be generated by the inhomogeneous Lorentz group and dilatations.

LET M denote Minkowski space, the real 4-dimensional space–time continuum of special relativity, and let Q denote the characteristic quadratic form on M,

$$Q(x) = x_0^2 - x_1^2 - x_2^2 - x_3^2,$$
$$x = (x_0, x_1, x_2, x_3) \in M.$$

There is a partial ordering on M given by $x < y$ if an event at x can influence an event at y; more precisely, $x<y$ if $y-x$ is a time vector, $Q(y-x)>0$, oriented towards the future, $x_0 < y_0$. Let $f : M \to M$ be a function that is a one-to-one mapping (*we make no assumptions that f is linear or continuous*). We call f a *causal automorphism* if both f and f^{-1} preserve the partial ordering; in other words,

$$x < y \Leftrightarrow fx < fy, \quad \text{all} \quad x, y \in M.$$

The causal automorphisms form a group, which we call the *causality group*.

Let G be the group generated by (i) the orthochronous Lorentz group (linear maps of M that leave Q invariant, and preserve time orientation, but possibly reverse space orientation), (ii) translations of M, and (iii) dilatations of M (multiplication by a scalar).

Theorem. The causality group = G.

Remark 1. The significance of the theorem is that if we interpret the principle of causality mathematically as the set M together with the partial ordering, then the inhomogeneous Lorentz group appears naturally (with dilatations and space reversal) as the symmetry group of M. Therefore the basic invariants of physics, which are the representations of the inhomogeneous Lorentz group, follow naturally from the single principle of causality.

Remark 2. It is easy to see that G is contained in the causality group, since the generators of G preserve the partial ordering. The converse is not obvious at first sight, because there seems no reason why a causal automorphism should be linear or even continuous. In fact, the result depends essentially upon space being more than 1-dimensional. If space were 1-dimensional then the causality group would be much larger than G, and the general causal automorphism would map the space and time axes into curved lines, as is shown by the example below. Thus the typical 2-dimensional picture of Minkowski space to be found in most textbooks is misleading.

Remark 3. The condition for f to be a causal automorphism is a global condition, but is equivalent (by an elementary compactness argument using the transitivity of $<$) to the following local condition: given $x \in M$, then there is a neighborhood N of x such that

$$y < z \Leftrightarrow fy < fz, \quad \text{all} \quad y, z \in N.$$

Intuitively this means we need only think of the principle of causality acting in our laboratories for a few seconds, rather than between distant galaxies forever, and still we are able to deduce the Lorentz group.

Remark 4. There is another relation on M given by $x<\cdot\, y$ if light can go from x to y; more precisely $x <\cdot\, y$ if $y - x$ is a light vector, $Q(y - x) = 0$, oriented towards the future, $x_0 < y_0$. The relation $x <\cdot\, y$ is not a partial ordering because it is not transitive,

$$x <\cdot\, y <\cdot\, z \not\Rightarrow x <\cdot\, z.$$

We shall show in Lemma 1 that, in the definition of causal automorphism, it does not matter whether we use $<$ or $<\cdot$ (or both). Intuitively this means that the Lorentz group can be deduced equally well either from causality between heavy particles, or from causality between photons, or from both. Remark 3 also holds for $<\cdot$, although the argument is slightly more complicated due to the lack of transitivity.

Remark 5. In Remarks 2 and 3, when referring to the "continuity" of f or the "neighborhood" of x, we have implicitly assumed a topology on M, although for the proof of the theorem we assume no topology. It is customary to think of M as having the topology of real 4-dimensional Euclidean space, but there are reasons why this is wrong. In particular:

(i) Euclidean topology is locally homogeneous whereas M is not; every point has its associated light cone separating space vectors from time vectors.

(ii) The group of all homeomorphisms of Euclidean space is vast, and of no physical significance.

In a subsequent paper[1] we suggest alternative topologies for M, which are not homogeneous, and have the property that any homeomorphism maps light cones to light cones. Therefore any homeomorphism preserves or reverses the relation $<\cdot$, and so the group of all homeomorphisms of M with such a topology will be the double cover of G. Consequently, the topology is physically significant because it implies the Lorentz group.

The existence of such topologies on Minkowski space suggest the possibility of similar topologies on the inhomogeneous Lorentz group, finer than the Lie-group topology. Any representation with the Lie-group topology would *a fortiori* be a representation with a finer topology, but not necessarily conversely. This raises the question: are there some new representations of the inhomogeneous Lorentz group?

Example.

Let K denote 2-dimensional Minkowski space with characteristic quadratic form

$$Q(x) = x_0^2 - x_1^2, \qquad x = (x_0, x_1) \in K.$$

Choose new coordinates

$$y_0 = x_0 - x_1, \qquad y_1 = x_0 + x_1.$$

Let $f_0, f_1 : R \to R$ be two arbitrary nonlinear orientation-preserving homeomorphisms of the real line onto itself. Define $f : K \to K$ by

$$f(y_0, y_1) = (f_0 y_0, f_1 y_1).$$

Then f is a causal automorphism, but $f \notin G$ because f is nonlinear. In general, the images of the space and time axes will not be straight lines.

Lemma 1. *Let $f : M \to M$ be a function that is a one-to-one mapping. Then f, f^{-1} preserve the partial ordering $<$ if and only if they preserve the relation $<\cdot$.*

Proof. If $x < y$ implies $f^{-1}x < f^{-1}y$, then $x \not< y$ implies $fx \not< fy$. Therefore if f, f^{-1} preserve $<$, then f preserves $<$ and $\not<$. Now

$$x <\cdot y \Leftrightarrow \begin{cases} x \not< y \\ y < z \Rightarrow x < z. \end{cases}$$

Therefore if f preserves $<$ and $\not<$, then f preserves $<\cdot$. Therefore if f, f^{-1} preserve $<$, then f, f^{-1} preserve $<\cdot$. Conversely,

$$x < y \Leftrightarrow \begin{cases} x \not<\cdot y \\ x <\cdot z <\cdot y, \quad \text{for some} \quad z. \end{cases}$$

Therefore if f preserves $<\cdot$ and $\not<\cdot$, then f preserves $<$, and so if f, f^{-1} preserve $<\cdot$, then f, f^{-1} preserve $<$.

Notation.

If $x \in M$, let C_x denote the *light cone* through x,

$$C_x = \{y; x <\cdot y \quad \text{or} \quad x = y \quad \text{or} \quad y <\cdot x\}.$$

If $x <\cdot y$, we call the line through x and y a *light ray* and denote it by $R_{x,y}$. We deduce

$$R_{x,y} = C_x \cap C_y.$$

Lemma 2. A causal automorphism maps light rays to light rays.

Proof: Let f be a causal automorphism. By Lemma 1, f and f^{-1} preserve $<\cdot$, and so $fC_x = C_{fx}$. Therefore if $x <\cdot y$,

$$fR_{x,y} = f(C_x \cap C_y) = C_{fx} \cap C_{fy} = R_{fx,fy}.$$

Lemma 3. A causal automorphism maps parallel light rays to parallel light rays.

Proof: Let a_1, a_2 be parallel light rays, and let P be the plane through them. There are two cases according to whether or not P is a tangent to all the light cones with vertex in P.

Case (1). Suppose P is not a tangent (this is the usual case). Then P contains two families $\{a\}$, $\{b\}$ of light rays, where $\{a\}$ consists of all lines parallel to a_1 (and a_2), and $\{b\}$ consists of all lines parallel to another direction. If $x \in P$, then the light cone with vertex x meets P in the two light rays through x, one from each family.

Let f be a causal automorphism. The images $\{fa\}$, $\{fb\}$ are families of lines with the property that each fa meets each fb, but no two of any one family meet. There are two possibilities; (i) fa_1 and fa_2 are coplanar, or (ii) they are not. We shall show that (ii) leads to a contradiction. For if fa_1, fa_2 are not coplanar they lie in a 3-dimensional subspace S, say, of M. Then each $fb \subset S$, because

[1] E. C. Zeeman, "The topology of Minkowski space" (to be published).

it meets fa_1 and fa_2, and therefore each $fa \subset S$. Therefore the two families are generators of a nondegenerate quadric surface in S. If this quadric surface is a hyperboloid (meeting S_∞ in a conic, where S_∞ denotes the plane at infinity), then each fa is parallel to some fb (the unique fb through $fa \cap S_\infty$), contradicting that fa meets fb (in a finite point). Alternatively, if the quadric is a paraboloid (meeting S_∞ in two lines), then the directions of $\{fa\}$ are parallel to all the lines in a plane (which meets S_∞ in one of the lines). But all light rays are parallel to the rays in a single light cone (which meets M_∞ in a sphere), and so any plane contains at most two lines parallel to light rays (through the points in M_∞ where the sphere meets the line in $S_\infty \subset M_\infty$), and so again we have a contradiction.

Therefore fa_1, fa_2 must be coplanar, and, since they do not meet, must be parallel.

Case (2). Suppose P is a tangent to all light cones with vertex in P (this is the exceptional case). The argument of Case (1) breaks down because P has the property that it contains only one family of light rays, namely all the lines parallel to a_1 and a_2. The planes through a_1 with this property span a 3-dimensional subspace A_1, say, of M (the tangent prime to the light cones through a_1). Similarly the planes through a_2 with the property span A_2. Choose a_3 parallel to a_1 and a_2, and not in $A_1 \cup A_2$. Then by Case (1), a_1 and a_2 are both parallel to a_3, and hence parallel to each other. The proof of Lemma 3 is complete.

Remark. So far, everything we have done applies to the 2-dimensional example above. As yet we have not proved that f maps each light ray linearly, nor have we proved that f maps straight lines other than light rays into straight lines. We prove this in the next lemma, using the fact that the dimension of space is greater than 1.

Lemma 4. A causal automorphism maps each light ray linearly.

Proof: Suppose a, a_1 are parallel light rays, as in Case (1) of Lemma 3. The family $\{b\}$ of parallel light rays meeting a and a_1 determine a linear map $g_1 : a \to a_1$, and if f is a causal automorphism the image family $\{fb\}$ determine a linear map $e_1 : fa \to fa_1$ such that the diagram

$$\begin{array}{ccc} a & \xrightarrow{f} & fa \\ \downarrow{\scriptstyle g_1} & & \downarrow{\scriptstyle e_1} \\ a_1 & \xrightarrow{f} & fa_1 \end{array}$$

is commutative. If a_2 is also parallel to a_1, we can define similar maps g_2, e_2 for the pair a_1, a_2, and maps g_3, e_3 for the pair a_2, a, provided neither of the pairs is exceptional as in Case (2) of Lemma 3. Composing the three diagrams gives a commutative diagram

$$\begin{array}{ccc} a & \xrightarrow{f} & fa \\ \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle e} \\ a & \xrightarrow{f} & fa \end{array}$$

where $g = g_3g_2g_1$ is a translation of a, and $e = e_3e_2e_1$ is a translation of fa (g, e are translations because they are compositions of parallel displacements).

If Minkowski space were 2-dimensional, then any such translations would have to be the identity. But in higher dimensions—and this is where the difference is essential—we claim that any given translation g of a can be obtained in this manner. It suffices to construct an arbitrary translation on one particular light ray, for then the result will be true for all light rays since G is transitive on the set of all light rays.

Let $x = (0, 0, 0, 0)$, $y = (0, -t, 0, t)$, $z = (0, 0, 0, 2t)$ and $x^* = (0, 0, t, t)$. Let a, a_1, a_2 be the light rays through x, y, z, respectively, parallel to the direction $[0, 0, 1, 1]$.
Then

$$x \xrightarrow[g_1]{} y \xrightarrow[g_2]{} z \xrightarrow[g_3]{} x^*.$$

Therefore $gx = x^*$, and the given translation g can be obtained by suitable choice of the parameter t.

Let r, s be coordinates chosen on a, fa such that $f(0) = 0$. Suppose that when g is the translation $r \to r + t$, then e (which is uniquely determined by g) is the translation $s \to s + u$, where $u = u(t)$. Then

$$f(r + t) = fg(r) = ef(r) = f(r) + u(t),$$

for all r, t. Putting $r = 0$, we have

$$f(t) = u(t),$$

and so

$$f(r + t) = f(r) + f(t).$$

Therefore by induction $f(nt) = nf(t)$, for positive and negative integers n. If m is also an integer, then $nf[(m/n)t] = f(mt) = mf(t)$, and so

$$f(rt) = rf(t)$$

for r rational. But the last equation is also true for r real, because f preserves $<\cdot$, and so is order-

preserving on each light ray. Hence f is linear on the light ray, and Lemma 4 is proved.

Lemma 5. A causal automorphism maps parallel equal intervals on light rays to parallel equal intervals.

Proof: Parallel light rays must be mapped with the same linear expansion because the family of parallel light rays meeting them both also remains parallel. (In the exceptional case use a third ray, as in the proof of Lemma 3.) Therefore equal intervals are mapped with the same linear expansion onto equal intervals.

Proof of the Theorem: We are given a causal automorphism $f : M \to M$. We can assume the f keeps the origin fixed, by first composing f with a translation if necessary. Choose four linearly independent vectors v_1, v_2, v_3, v_4 directed along four light rays through the origin: these form a base for the vector-space structure of M, and so an arbitrary vector $x \in M$ can be written

$$x = \sum x_i v_i, \qquad x_i \text{ scalar.}$$

Let $g : M \to M$ be the linear map given by

$$gx = x_i(fv_i).$$

We shall show that f is linear by proving that $f = g$. For each i, $1 \leq i \leq 4$, let M_i denote the i-dimensional vector subspace spanned by v_j, $1 \leq j \leq i$. We shall show that $f = g$ on M_i by induction on i. The induction starts with $i = 1$ by Lemma 4, and finishes with $i = 4$. Assume the induction for $i - 1$. Given $x \in M_i$, write $x = y + x_i v_i$, where $y \in M_{i-1}$. Then the interval from y to x is parallel and equal in length to the vector $x_i v_i$. By Lemma 5 the interval from fy to fx is parallel and equal in length to $f(x_i v_i)$. Therefore

$$\begin{aligned} fx &= fy + f(x_i v_i) \\ &= gy + g(x_i v_i), \text{ by induction and by Lemma 4,} \\ &= gx, \text{ because } g \text{ is linear.} \end{aligned}$$

This completes the inductive step, and the proof that f is linear.

Since f preserves $<\cdot$, the light cone, $Q(x) = 0$, through the origin is kept fixed. Therefore, multiplying f by a scalar if necessary, we deduce that f leaves Q invariant. In other words, in modulo multiplication by a translation and a dilatation, f is a time-orientation-preserving element of the Lorentz group. Therefore $f \in G$, and the proof of the theorem is complete.

ACKNOWLEDGMENT

The author is indebted to the physicists of the Institut des Hautes Études Scientifiques, and in particular to F. Lurçat, for their patience in explaining to him which mathematics in modern physics is significant.

New Methods for Reduction of Group Representations Using an Extension of Schur's Lemma*

JOHN GABRIEL
Argonne National Laboratory, Argonne, Illinois
(Received 1 November 1963)

The generators of the group or group algebra are used in an analog of the Lie–Cartan method, which can be applied to finite or infinite groups. This gives a mean for reduction of an arbitrary group representation, using only the matrix representatives of the generators. It is a set of algorithms using pivotal condensation and can easily be coded as a digital computer program. Connections with Lie–Cartan theory are suggested, the reduction of the symmetric group discussed, and methods for the reduction of representations of the $n \times n$ unitary, orthogonal, and proper orthogonal groups suggested.

I. INTRODUCTION

THE process derived is the formalization of a procedure which was deduced semi-intuitively from an analysis of the place of group theory in quantum mechanics.

The method used for proof is the most pedestrian of a number of alternative techniques which all seem likely to work. It has been chosen because it defines the numerical procedure without requiring the user to have a very sophisticated understanding of group representation theory and group algebras. None of the more concise alternatives does this nearly so well, and although the proof given is more tedious than it needs to be because it is nonspecialized, this nonspecialization makes it easier to program for a computer or to write as input for a "Brooker–Irons" type of compiler[1,2] such as COGENT[3] which is being written for use at Argonne. Since the proof and method in the form given do not suggest the way that they were first obtained, it is worthwhile to outline some of the considerations which led to the discovery.

Because the quantum mechanical states which span an irreducible representation of the symmetry group for the Hamiltonian have the same energy, one way to reduce a representation would be to construct a Hamiltonian known to have no accidental degeneracy and transform its matrix representative to diagonal form.

The difficulty in this suggestion is the need to guarantee no accidental degeneracy. This can only be done in particular cases and is difficult to generalize. The solution is to consider not a particular Hamiltonian, but the set of all operators which commute with every group element. The set of all possible Hamiltonians is a subset of these and if all possible Hamiltonians are degenerate then *accidental* degeneracy is not an appropriate term. Since if two operators A and B commute with every group element, so do AB and BA; the set of all operators which commute with the group is an algebra.[4] It is called the commutator algebra of the group. Lemma 2 establishes that the reduction of a representation of the commutator algebra into its irreducible components also reduces the group representation.

If it were necessary to construct the matrix representative of every group element and check that it commuted with a matrix A to prove that A represented some element of the commutator algebra, then this approach would be no better than the usual one using the operators $e_{ij}^{\alpha} = (n_\alpha/g) \sum_g U_{ji}^{\alpha}(P^{-1})P$, since it would implicitly require representatives of every group element. This necessity in the usual method is one reason for its lack of popularity in applications of the Pauli Principle to SCF calculations for example.

But it is almost obvious that it is necessary and sufficient that A commute with the group generators[5] for it to commute with the group, and Lemma 1

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

[1] E. T. Irons, Working Paper No. 47, Commun. Research Division, Institute for Defense Analyses, Princeton, New Jersey; also Commun. Assoc. Computing Mach. 4, 51 (1961).

[2] R. A. Brooker and D. Morris, Computer J. 3, 168 (1960); 3, 220 (1961); J. Assoc. Computing Mach. 9, 1 (1962).

[3] J. Reynolds, Appl. Math. Div., Argonne National Laboratory, Argonne, Illinois (private communication).

[4] Readers who would like further information on algebras are referred to a concise treatment in H. Weyl, *Group Theory and Quantum Mechanics* (Methuen and Company, Ltd., London, 1931; Dover Publications, Inc., New York), pp. 165–170 and 302–309.

[5] W. Lederman, *Introduction to the Theory of Finite Groups* (Oliver and Boyd, London, England, 1953; Interscience Publishers, Inc., New York, 1953), Sec. 14, Eq. (2.37).

provides a formal proof of this. It is very easy to construct the commutator algebra $\mathfrak{A}$ of the cyclic group generated by a single element P. All that is necessary is for $\mathfrak{A}$ not to mix eigenvectors of P belonging to different eigenvalues. With P in diagonal form, Lemma 2 proves this. The result implies that S, the space to be reduced, can be reduced into irreducible invariant subspaces S_λ labelled by the eigenvalues λ of P, and that these afford irreducible representations of $\mathfrak{A}$ through submatrices A_λ of matrices A, which leave the spaces S_λ invariant. Suppose now that another group element Q is introduced, by being given as a matrix in the basis with P diagonal. If the order[6] of the matrix of Q is known, its eigenvalues are also known, and a matrix H can be found by pivotal condensation[7,8] which makes $H^{-1}QH$ diagonal. This allows S to be reduced into $S = \sum_\mu S_\mu$ with respect to (wrt) the commutator algebra $\mathfrak{B}$ of Q also, and H gives a (not necessarily unique) relation between the two bases—one which reduces $\mathfrak{A}$, and the other which reduces $\mathfrak{B}$. The elements common to $\mathfrak{A}$ and $\mathfrak{B}$ form a subalgebra $\mathfrak{C}$ which leaves both S_μ and S_λ invariant. Lemma 3, which is a generalization of Schur's Lemma, indicates how H might be used to reduce the S_μ and S_λ wrt $\mathfrak{C}$; Lemma 4 specifies the process of reduction, first in a theoretical way, and then by a detailed algorithm for the process. Having reduced the commutator algebra of P and Q, the next step is to add R, a third element to the group, repeat the analysis, and continue doing this until all the group generators have been included.

The process works because of a number of interrelated facts, but the following remarks underlie all of them.

Reduction of a previously irreducible space when a group or algebra is restricted by deletion of some of its elements is a procedure that is fairly easy to analyze, because spaces that were irreducible remain invariant and afford (implicitly) reducible representations of the new group.

Adding a new group element does not leave invariant spaces invariant, and therefore the analysis of the new representation has to be started by finding the way that the old irreducible subspaces of the group are coupled by the new element.

The natural way to look for these is to remark that vectors which are not coupled by the group are coupled by its commutator algebra, and inclusion of another group element deletes elements of the commutator algebra. This fact is used implicitly in a method for reduction of the symmetric group[9-11] whose origins are about thirty years old,[12,13] and also in Weyl's work[4,14] on group structure.

There is one other aspect of the problem which shows the importance of Lemma 3. The use of projection operators depends on the orthogonality relations which in turn depend on Schur's Lemma. When two representations are equivalent, there exists a nonsingular matrix H such that $HU(S)H^{-1} = U'(S)$, where U and U' are equivalent representations of S. This implies $HU(S) = U'(S)H$, and taking Kronecker products, $[U'(S^{-1})] \times [HU(S)] = [U'(S^{-1})] \times [U(S)H]$. Summing over the group and using the orthogonality relations gives a set of simultaneous linear equations for the elements of H. This is the reality behind the projection-operator formalism, and it shows that Schur's Lemma must be used in the reduction of an arbitrary representation, and is why use of Lemma 3, which is another way of expressing the above result, is a vital part of the reduction.

The essence of this work is to recognize the above implications, and prove six lemmas which are suggested by them. All six are straightforward except for Lemma 5, and the last two thirds of Lemma 3. Readers who are not familiar with numerical methods for matrix analysis are advised to omit these portions of the paper on a first reading since the rest is independent of them, except insofar as they define the numerical method and prove the validity of results which the reader may be ready to take for granted until he comprehends the main line of argument.

After this preamble it is appropriate to state and prove the Lemmas.

II. FORMAL PROOF

Lemma 1

If $P_1, P_2, \cdots P_f$ are the matrix representatives of the f generators of a group G in a space S, then

[6] See Ref. 5, Sec. 8, Definition 5; the possible latent roots are powers of the hth roots of unity.

[7] A. C. Aitken, *Determinants and Matrices*, (Oliver and Boyd, London, England; Interscience Publishers, Inc., New York, 1951), Sec. 29.

[8] H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, (Blackie & Son, London, 1932; Dover Publications, Inc., New York), Chap. III; Chap. V, Sec. 4, p. 49, Theorem II; Chap. VI.

[9] J. R. Gabriel, Proc. Camb. Phil. Soc. **57**, 330 (1961).

[10] E. M. Corson, *Perturbation Methods in Quantum Mechanics* (Blackie & Son, Ltd., London, 1951).

[11] T. Yamanouchi, Proc. Phys. Math. Soc. **Japan 17**, 274 (1935).

[12] R. Serber, Phys. Rev. **45**, 461 (1934).

[13] J. H. Van Vleck, Phys. Rev. **45**, 405 (1934).

[14] H. Weyl, *The Classical Groups* (Princeton, University Press, Princeton, New Jersey, 1939; Oxford University Press, Oxford, England, 1946).

a matrix A commutes with every matrix representing an element of G in S if and only if it commutes with each of the f matrices $P_1, P_2, \cdots P_f$.

Proof: The condition is necessary since the P_i, $i = 1 \cdots f$ are group elements.

It is sufficient because if $AP = PA$, $P^{-1}A = AP^{-1}$ and A commutes with P^{-1}. Therefore A commutes with every product of P's and their reciprocals, and the products are the elements of G.

If γ is the group algebra of G, the condition is also necessary and sufficient for A to commute with every element of γ, since an element of γ is a polynomial in elements of G.

If the P_i are generators of γ, but not of G, A still commutes with every element of γ since elements of γ are quotients of polynomials in the P_i by nonsingular polynomials in the P_i.

Corollary. If $\mathfrak{A}_1, \mathfrak{A}_2, \cdots \mathfrak{A}_f$ are the commutator algebras of $P_1, P_2, \cdots P_f$, the commutator algebra of γ contains just those elements common to $\mathfrak{A}_1, \mathfrak{A}_2, \cdots \mathfrak{A}_f$.

Lemma 2

Let $\mathfrak{A}$ be the set of all matrices A which commute with all the matrix representatives of a group G in a space S. Suppose S is divided into subspaces $S(p, r)$, r fixed, $p = 1, 2 \cdots p_{\max}(r)$ which afford the same representation of $\mathfrak{A}$ i.e., *each of the submatrices* $A(p, r)$, $p = 1, 2, \cdots p_{\max}(r)$ which represent $\mathfrak{A}$ in $S(p, r)$ are the same. Spaces with different r values afford inequivalent representations.

Then if the matrices P_i of G are partitioned according to the spaces $S(p, r)$, the submatrix $P_i(pr, qs) = 0$ if $r \neq s$, and is a multiple of the unit submatrix $\lambda_{pq}(r, P_i)1$ if $r = s$.

Proof: Since all the nonzero submatrices of A lie on the diagonal, $AP_i = P_iA$ implies

$$A(p, r)P_i(p, r; q, s) = P_i(p, r; q, s)A(q, s),$$

and Schur's Lemma shows

$$P_i(p, r; q, s) = 0 \quad \text{if} \quad r \neq s$$
$$= \lambda_{pq}1 \quad \text{if} \quad r = s,$$

since $A(p, r)$ is irreducible.

Corollary. Suppose the base vectors of $S(p, r)$ are called $x(p, r, m)$ where two vectors with the same m and r but different p values belong to the same row of the matrix $A(r)$ which represents $\mathfrak{A}$ in each of the spaces $S(p, r) = p = 1, 2, \cdots p_{\max}(r)$.

The lemma shows that the vectors $x(p, r, m)$, r, m fixed $p = 1 \cdots p_{\max}(r)$ span a space $\sum (m, r)$ which affords a representation Γ_r of G. If $d(r)$ is the dimension of $S(p, r)$, $\sum (m, r)$ occurs $d(r)$ times.

Because $\mathfrak{A}$ includes *all* the matrices which commute with the representation of G, and no matrix of $\mathfrak{A}$ couples $S(p, r)$ to $S(p' r)$, the spaces $\sum (m, r)$ each afford an irreducible representation of G, since if some $\lambda_{pp'}(r, P_i)$ were zero for all i, then matrices A would exist coupling $S(p, r)$ and $S(p'r)$.

(This last result has to be worded a little differently if S is not fully reducible, but the modification is obvious.)

Notation. Lemma 2 can be put most concisely as follows. Define $e(p, r; q, s)$ as having ones down the diagonal of the $(p, r; q, s)$ submatrix and zeros elsewhere; then $e(p, r; q, s)\, e(q, s; t, u) = e(p, r; t, u)$. Define $\epsilon(p, r) = e(p, r; p, r)$; ϵ has eigenvalues 0 or 1 since $\epsilon^2(p, r) = \epsilon(p, r)$, and those eigenvectors belonging to the eigenvalue 1 span $S(p, r)$. Since the $S(p, r)$ span S, $\sum_{p,r} \epsilon(p, r) = 1$, the unit matrix, and if P_i is an element of the group $\epsilon(p, r)P_i\epsilon(q, s) = \delta_{rs}\lambda_{pq}(P_i, r)e(p, r; q, s)$. The $\epsilon(p, r)$ define the reduction of S and A belongs to the commutator algebra of the group if $\epsilon(p, r)A\epsilon(qs) = \delta_{pq}\delta_{rs}A(pr)$. These ϵ's are called primitive idempotents in the reduction of S.[4]

Lemma 3

If $\mathfrak{A}$ and $\mathfrak{B}$ are two sets of matrices and a matrix H exists such that $\mathfrak{A}H = H\mathfrak{B}$, i.e., given an element A of $\mathfrak{A}$, an element B of $\mathfrak{B}$ can always be found such that $AH = HB$ and vice versa. Then:

1. If H is neither zero, nor square and nonsingular, $\mathfrak{A}$ and $\mathfrak{B}$ are reducible but not necessarily fully reducible.

2. An explicit algorithm exists for reduction of $\mathfrak{A}$ and $\mathfrak{B}$ by pivotal condensation of H.

3. If A and B are fully reducible, then A is reduced into two submatrices A_{11} and A_{22}, and B into B_{11} and B_{22}. If H is of rank r, A_{11} and B_{11} are equivalent and both $r \times r$.

4. The congruence transformations derived from pivotal condensation of H, to reduce A and B through $A_U = UAU^{-1}$ and $B_X = XBX^{-1}$ also reduce the basis space.

Proof: If H is as described and of rank r, rearrange the base vectors for $\mathfrak{A}$ and $\mathfrak{B}$ to put a nonsingular $r \times r$ submatrix in the leading position of H.

Partition H accordingly,

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}.$$

H_{11} is square and nonsingular. Suppose H has n rows and m columns; because it is of rank r, only r columns are linearly independent. Since the basis

has been rearranged to put a nonsingular $r \times r$ submatrix in the leading position, the first r columns are linearly independent, and each of the last $m - r$ is a linear combination of the first r. Consider the post multiplication

$$\begin{bmatrix} H_{11} & H_{21} \\ H_{12} & H_{22} \end{bmatrix} \begin{bmatrix} 1 & F \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{21} + H_{11}F \\ H_{12} & H_{22} + H_{12} \end{bmatrix}.$$

This adds a possibly different linear combination (determined by F) of the first r columns to each of the last $m - r$. Choosing $F = -H_{11}^{-1}H_{21}$, which can be done since H_{11} is nonsingular, makes $H_{21} + H_{11}F = 0$. It also makes $H_{22} + H_{12}F = 0$ because adding a linear combination of the first r columns to one of the last $m - r$ which annihilates the top r elements must also annihilate the rest, because all the elements of this column can be annihilated by adding the correct linear combination of the first r, and since the rows of H_{11} are linearly independent, no other F matrix except $F = -H_{11}^{-1}H_{21}$ would annihilate the top r elements of this column.

The change of basis induced in the carrier space of B by the reciprocal of this matrix does not reduce the space because the two subspaces which result overlap; i.e., an arbitrary vector of the carrier space does not uniquely determine two vectors, one in each space whose sum is the arbitrary vector.[15] This can be seen from the example on the cubic group which is given after Lemma 4, where the change of basis in the $+1$ eigenspace of C_{4z} from $[z^2, (x^2+y^2)/\sqrt{2}]$ to $[(x^2+y^2+z^2)/\sqrt{2}, (x^2+y^2)/\sqrt{2}]$ patently does not give the right spherical harmonics, although it condenses the necessary column of H.

The proper transformation is to $[(x^2+y^2+z^2)/\sqrt{2}, (x^2 + y^2 - 2z^2)/\sqrt{2}]$ or, with subsequent normalization, to $[(x^2 + y^2 + z^2)/\sqrt{3}, (x^2 + y^2 - 2z^2)/\sqrt{6}]$.

The generalization of this is to use

$$X = \begin{bmatrix} 1 & F \\ -F^\dagger & 1 \end{bmatrix},$$

where $F^\dagger$ is the Hermitian conjugate of F, for post multiplication.

The post-multiplied H is

$$HX = \begin{bmatrix} H_{11} - H_{21}F^\dagger & 0 \\ H_{12} - H_{22}F^\dagger & 0 \end{bmatrix}.$$

Since $H_{21} + H_{11}F = 0$, and $H_{22} + H_{12}F = 0$,

$$HX = \begin{bmatrix} H_{11}(1 + FF^\dagger) & 0 \\ H_{12}(1 + F^\dagger F) & 0 \end{bmatrix}.$$

[15] See Ref. 4, p. 3, line 33, *et seq.*; or Ref. 14, p. 9, line 27, *et seq.*; also Ref. 14, 2nd ed., p. 295, Theorem 10.4A.

$FF^\dagger$ is square, nonnegative-definite,[8] and therefore $(1 + FF^\dagger)$ is square, nonsingular, positive-definite. Similarly, $(1 + F^\dagger F)$ is square nonsingular, positive-definite.

Therefore $(1 + FF^\dagger)^{\frac{1}{2}}$ and $(1 + F^\dagger F)^{\frac{1}{2}}$, and their reciprocals exist and are real if F is real.

Defining Y by

$$Y = X \begin{bmatrix} (1 + FF^\dagger)^{-\frac{1}{2}} & 0 \\ 0 & (1 + F^\dagger F)^{-\frac{1}{2}} \end{bmatrix},$$

partitioned matrix multiplication shows that $Y^\dagger Y = 1$, i.e., that Y is unitary.

Since this result is to be used for numerical computation, the algorithm for calculating Y from F which is read from the original H should be outlined. The only difficulty is in the calculation of $(1 + FF^\dagger)^{\frac{1}{2}}$. It is not necessary to reduce $FF^\dagger$ to classical canonical form to do this, since all the post multiplication of X to give Y does is to perform a Schmidt[16] orthogonalization on the m columns of X without mixing the first r with the last $m - r$. The matrix Y is the X^{-1} of statement 4 of this lemma. Exactly similar arguments lead to the matrix U and a transformed H consisting of a nonsingular $r \times r$ matrix H'_{11} in the leading position and zeros elsewhere.

The matrix H'_{11} must be nonsingular because U and X are nonsingular and so the rank of UHX^{-1} is r (by the Binet–Cauchy theorem).[17] Thus

$$UHX^{-1} = \begin{bmatrix} H'_{11} & 0 \\ 0 & 0 \end{bmatrix}.$$

Partitioning $A_u = UAU^{-1}$ and $B_x = XBX^{-1}$ conformably to this,

$$A_u = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad B_x = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

$AH = HB$ requires

$$A_{11}H'_{11} = H'_{11}B_{11},$$
$$A_{21}H'_{11} = 0,$$
$$0 = H'_{11}B_{12},$$

and, since H'_{11} is nonsingular, $A_{21} = 0$, $B_{12} = 0$, and $B_{11} = H_{11}^{-1} A_{11} H_{11}$.

Lemma 4

If $\mathfrak{A}$ is the commutator algebra of a matrix P which operates in a space S, and $\mathfrak{B}$ is the commutator

[16] E. Bodewig, *Matrix Calculus* (North-Holland Publishing Company, Amsterdam, 1956), pp. 22–26. See also Ref. 8, Chap. VII, Sec. 10, p. 95; and Ref. 14, 2nd ed., p. 295, Theorem 10.4A.

[17] See Ref. 7, Chap. V. Secs. 39–41.

algebra of Q which also operates in S, then the representation of $\mathfrak{C}$, the commutator algebra of P and Q, can be reduced if a matrix H, which maps the irreducible subspaces S_α of S wrt $\mathfrak{A}$ onto the irreducible subspaces S_β of S wrt $\mathfrak{B}$, is known.

Proof: Suppose that C commutes with P and Q; then it also commutes with every matrix X of the pencil $x = (P + xQ)/(1 + x)$. Lemma 2 shows that degenerate eigenvectors of X which do not depend on x must span invariant subspaces wrt C.

The space S can be divided into subspaces $S_{\lambda\mu}$ each of which is spanned by simultaneous eigenvectors of X belonging to the eigenvalues λ of P and μ of Q, and a remainder $S(x)$ which is spanned by eigenvectors of P or eigenvectors of Q, but every eigenvector of X in $S(x)$ depends on x.

No two eigenvectors of X in $S(x)$ belong to the same eigenvalue of X for all x, for if so they span an invariant subspace of $S(x)$ in which X is represented by a multiple of the unit matrix for all x and the special values $x = 0$ and $x = \infty$ show that they are simultaneous eigenvectors of P and Q.

Because of this $C(x)$, the representative of C in $S(x)$ is a diagonal matrix.

The only other question is: Is $C(x)$ an arbitrary diagonal matrix, or do P and Q force degeneracies?

The special case where Q is a diagonal matrix D_2 and $P = HD_1H^{-1}$ where D_1 is another diagonal matrix, may be used to discuss this. The space $\sum_{\lambda\mu} S_{\lambda\mu}$ is an invariant space wrt H and H^{-1} because it is a simultaneous eigenspace, and, because both H and H^{-1} leave $\sum_{\lambda\mu} S_{\lambda\mu}$ invariant, H leaves $S(x)$ invariant. Since $S(x)$ is invariant wrt H, there is a submatrix $H(x)$ representing H in $S(x)$. Because $H(x)$ only mixes base vectors of $S(x)$ with other base vectors of $S(x)$, $C'(x) = H^{-1}C(x)H$ is also diagonal for the same reasons as $C(x)$ is. Since $C'(x)$ and $C(x)$ are equivalent, they have the same eigenvalues, i.e., the diagonal elements of $C'(x)$ are a rearrangement of those of $C(x)$.

This means that the rows and columns of $H(x)$, and the base vectors of $S(x)$ and $H(x)S(x)$ can be rearranged so that $C(x)$ and $C'(x)$ appear with multiples of the unit submatrix down their diagonals, and if H is conformably partitioned, there is just one nonzero submatrix in each row of submatrices, which is also nonsingular, and similarly for the columns.

Lemma 5

Given two algebras and a space which affords representations of both of them:

If a matrix is known which transforms a basis that displays one algebra in terms of its irreducible representations into an analogous basis for the other, then an explicit algorithm exists for reduction of the subalgebra of elements common to both algebras.

Proof: $\mathfrak{A}$ and $\mathfrak{B}$ are two algebras having a common subalgebra $\mathfrak{C}$ containing all the elements common to $\mathfrak{A}$ and $\mathfrak{B}$. When it is necessary to distinguish between $\mathfrak{C}$ regarded as a subalgebra of $\mathfrak{A}$, and $\mathfrak{C}$ as a subalgebra of $\mathfrak{B}$, the names $\mathfrak{C}_A$ and $\mathfrak{C}_B$ will be used.

$\mathfrak{A}$ and $\mathfrak{B}$ are represented in a vector space S by sets of matrices A and B. S can be reduced into irreducible subspaces S_α of dimension n_α spanned by base vectors $x_{\alpha i}$ wrt $\mathfrak{A}$. Analogous definitions of subspaces S_β, dimensions n_β, and base vectors $x_{\beta j}$ are made wrt $\mathfrak{B}$. When it is necessary to distinguish between S reduced wrt $\mathfrak{A}$ according to $S = \sum_\alpha S_\alpha$ and the analogous reduction wrt $\mathfrak{B}$, the names S_A and S_B will be used.

Let S_B be mapped on S_A by a nonsingular matrix H, i.e., the $x_{\alpha i}$ and the $x_{\beta j}$ are connected by

$$x_{\alpha i} = \sum_\beta \sum_{j=1}^{n_\beta} H_{\alpha i,\beta j} x_{\beta j}.$$

The submatrix $H_{\alpha i,\beta j}; i = 1 \cdots n_\alpha, j = 1 \cdots n_\beta$ is called $H_{\alpha\beta}$. The reciprocal matrix of H maps S_B onto S_A. The appropriate submatrices are called $H^{-1}_{\beta\alpha}$. These should not be confused with the reciprocals of the $H_{\alpha\beta}$, when these are square and nonsingular which are called $(H_{\alpha\beta})^{-1}$. Call the matrices of $\mathfrak{A}$ which belong to $\mathfrak{C}$, A_C. Define B_C analogously. Then $B_C = H^{-1}A_CH$. The spaces S_α are invariant but not necessarily irreducible wrt $\mathfrak{C}$. Call the part of A_C in S_α $A_{C\alpha}$. Define $B_{C\beta}$ analogously. Then

$$A_{C\alpha}H_{\alpha\beta} = H_{\alpha\beta}B_{C\beta}.$$

Either $H_{\alpha\beta}$ is square and nonsingular, zero, or Lemma 3 applies, and both S_α and S_β are reducible. In many cases it is known that reducibility implies complete reducibility. When this is true, the process of Lemma 3 reduces S_α into two invariant subspaces according to $S_\alpha = S_{\alpha'} + S_{\alpha''}$, and S_β similarly according to $S_\beta = S_{\beta'} + S_{\beta''}$.

In the not fully reducible case the invariance of $S_{\alpha''}$ and $S_{\beta''}$ must be tested by examining $H^{-1}_{\beta\alpha}$. If S_α is found to be reducible but not completely reducible, all of the argument following this point can be repeated except that H and H^{-1} must be considered together, and reducible but not fully reducible spaces marked especially. This will not be done here since it will cause unnecessary confusion, and the reader who needs this case (usually in

connection with the Lorentz group) should be able to construct the necessary argument himself.

To recapitulate, pivotal condensation of the submatrix $H_{\alpha\beta}$, in the fully reducible case leads to the construction of two matrices U_α and X_β, according to Lemma 3. The reduction of S_α by U_α into $S_{\alpha'} + S_{\alpha''}$, and S_β by X_β into $S_{\beta'} + S_{\beta''}$, $S_{\alpha'}$ and $S_{\beta'}$ afford equivalent representations of $\mathfrak{C}$ in S_A and S_B.

If all the submatrices in the α block of rows of H are premultiplied by U_α, and the β columns are similarly post-multiplied by X_β^{-1}, the reduction of S_α and S_β is included in the new H matrix. The other nonzero submatrices in the new α rows and β columns can then be divided further to show the coupling of $S_{\alpha'}$ and $S_{\alpha''}$ to other spaces in S_B, and $S_{\beta'}$ and $S_{\beta''}$ to other spaces in S_A. If any of the corresponding submatrices are nonzero and not square and nonsingular, the subspaces can be reduced further until all the submatrices in the old α block of rows and β block of columns are either zero or nonsingular, and the subspaces descended from S_α in S_A and S_β in S_B cannot be further reduced by condensation of the α rows or β columns of H. Subspaces of S_α and S_β coupled by nonsingular submatrices afford equivalent representations of $\mathfrak{C}$ in S_A and S_B.

When this has been done, the subspaces S_α of S_A and S_β of S_B can be relabeled by their parent α or β and another name corresponding to the $'$, $''$, etc.

The transformed matrix H which takes account of the above reduction has nonsingular or zero submatrices in the block of rows whose parent is S_α, and similarly for S_β. Consider the rows of the new H which belong to a single subspace whose parent is S_α, and call it S_σ.

Using the notation of Lemma 4, there is no guarantee that the basis vectors have been arranged to put the matrices $C(x)$ and $C'(x)$ in the same form, or that the $S_{\lambda\mu}$ have been separated at this stage; there may be more than one nonzero, nonsingular matrix in the σ block of rows; one of them may occur at the intersection with a column τ whose parent is not S_β, and which may therefore contain a nonzero singular or rectangular submatrix $H_{\lambda\tau}$. S_λ which belongs to S_A, and S_τ which belongs to S_B may both be reduced by condensation of $H_{\lambda\tau}$. The nonsingular, square submatrix $H_{\sigma\tau}$ forces a corresponding reduction of S_σ in S_A since S_σ in S_A and S_τ in S_B afford equivalent representations of $\mathfrak{C}$. Moreover the space $S_{\lambda'}$, which is coupled to $S_{\tau'}$ by the nonzero submatrix $H_{\lambda'\tau'}$ of the condensed $H_{\lambda\tau}$, is equivalent to the $S_{\sigma'}$ whose separation from $S_{\sigma''}$ is forced by the nonsingular submatrix $H_{\sigma\tau}$ once S_τ and S_λ are reduced.

This procedure can be continued until S_A and S_B have been reduced into invariant subspaces wrt $\mathfrak{C}$, and the corresponding submatrices of H are all either zero or square and nonsingular.

Lemma 4 shows that where more than one nonzero submatrix is found in a set of rows, the subspaces S_β can be rearranged so as to put all these together, and a similar rearrangement of the $S\alpha$'s can be done, so that the submatrices are assembled into one larger submatrix, and that when this has been done as far as possible each larger submatrix is square, nonsingular, and defines an invariant subspace wrt $\mathfrak{C}$. This procedure is applied to the first two generators P_1 and P_2 and leads to the reduction of the space wrt the commutator algebra of the pencil $P_1 + xP_2$. When this has been done the eigenvectors of P_3 must be found, and also the new matrix H relating the commutator algebra of an arbitrary member of the pencil (which is analogous to P in Lemma 4) to the commutator algebra of P_3 which is analogous to Q. The process of Lemma 4 is repeated to give the reduction of the commutator algebra of P_1, P_2 and P_3.

This can be repeated and will always derive the commutator algebra of the pencil $X_K = \sum_{i=1}^{k} x_i P_i$ from the eigenvectors of P_K and the commutator algebra of X_{K-1}.

This proves the theorem.

Example for Lemmas 4 and 5

At this stage an example will explain the process of Lemma's 4 and 5 better than many words. The symmetrized Kronecker square of the cubic group in 3-space will be reduced. The base vectors for the reducible representation are x^2, y^2, z^2, xy, yz, zx. The two fourfold rotations one about $0Z$, and the other about $0Y$, are generators for the cubic group and the classification according to eigenvalues of C_{4v} is displayed above the upper matrix in Table I, while the eigenvalues and eigenvectors of C_{4z} are shown along the left-hand edge. The matrix is H, and is partitioned into the 2×2 matrices $H_{+,+}$, $H_{+,-}$, $H_{-,+}$, and H_{--}; the 2×1 matrices $H_{+,-i}$, $H_{-,i}$, $H_{-,-i}$; the 1×2 matrices $H_{+i,+i}$, H_{+i-i}, $H_{-i,+i}$ and $H_{-i,+i}$.

The only submatrices which are not already condensed as far as possible, are H_{+-} and H_{-+} which are both of rank one. According to Lemma 3,

$$U_+ = \begin{bmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & 1 \end{bmatrix}$$

TABLE I. Reduction of the symmetrized Kronecker square of the cubic group.

A.	$+1$	$+1$	-1	-1	$+i$	$-i$
	y^2	$(x^2+z^2)/\sqrt{2}$	xz	$(x^2-z^2)/\sqrt{2}$	$y\frac{(x+iz)}{\sqrt{2}}$	$y\frac{(x-iz)}{\sqrt{2}}$
$+1$ z^2	0	$1/\sqrt{2}$	0	$-1/\sqrt{2}$	0	0
$+1$ $(x^2+y^2)/\sqrt{2}$	$1/\sqrt{2}$	1/2	0	1/2	0	0
-1 xy	0	0	0	0	$1/\sqrt{2}$	$1/\sqrt{2}$
-1 $(x^2-y^2)/\sqrt{2}$	$-1/\sqrt{2}$	1/2	0	1/2	0	0
$+i$ $z(x+iy)/\sqrt{2}$	0	0	$1/\sqrt{2}$	0	1/2	$-1/2$
$-i$ $z(x-iy)/\sqrt{2}$	0	0	$\sqrt{2}1/$	0	$-1/2$	1/2

B.	$+1$	$+1'$	-1	$-1'$	$+i$	$-i$
	$\frac{x^2+y^2+z^2}{\sqrt{3}}$	$\frac{x^2-2y^2+z^2}{\sqrt{6}}$	xz	$\frac{(x^2-z^2)}{\sqrt{2}}$	$\frac{y(x+iz)}{\sqrt{2}}$	$\frac{y(x-iz)}{\sqrt{2}}$
$+1$ $\frac{x^2+y^2+z^2}{\sqrt{3}}$	1	0	0	0	0	0
$+1'$ $\frac{x^2+y^2-2z^2}{\sqrt{6}}$	0	$-1/2$	0	$\sqrt{3/2}$	0	0
-1 xy	0	0	0	0	$1/\sqrt{2}$	$1/\sqrt{2}$
$-1'$ $(x^2-y^2)/\sqrt{2}$	0	$\sqrt{3/2}$	0	1/2	0	0
$+i$ $z(x+iy)/\sqrt{2}$	0	0	$1/\sqrt{2}$	0	1/2	$-1/2$
$-i$ $z(x-iy)/\sqrt{2}$	0	0	$1/\sqrt{2}$	0	$-1/2$	1/2

C.	$+1$	$+1'$	-1	-1	$+i$	$-i$
	$\frac{x^2+y^2+z^2}{\sqrt{3}}$	$\frac{x^2-2y^2+z^2}{\sqrt{6}}$	$\frac{x^2-z^2}{\sqrt{2}}$	xz	$\frac{y(x+iz)}{\sqrt{2}}$	$\frac{y(x-iz)}{\sqrt{2}}$
$\frac{x^2+y^2+z^2}{\sqrt{3}}$	1	0	0	0	0	0
$\frac{x^2+y^2-2z^2}{\sqrt{6}}$	0	$-1/2$	$\sqrt{2}/2$	0	0	0
$\frac{x^2-z^2}{\sqrt{2}}$	0	$\sqrt{3}/2$	1/2	0	0	0
xy	0	0	0	0	$1/\sqrt{2}$	$1/\sqrt{2}$
$z\frac{(x+iy)}{\sqrt{2}}$	0	0	0	$1/\sqrt{2}$	1/2	$-1/2$
$z\frac{(x-iy)}{\sqrt{2}}$	0	0	0	$1/\sqrt{2}$	$-1/2$	1/2

will do what is necessary in $S(z+)$, the eigenvector space of C_{4z} belonging to $+1$. However it is convenient to normalize the rows and columns to unity to give

$$U_+ = \begin{bmatrix} \sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{2}{3}} & \sqrt{\frac{1}{3}} \end{bmatrix}.$$

An analogous transformation of $S(y, +)$ performed simultaneously sets up the eigenvectors shown in the first H partitioning (see proof of Lemma 3), lower matrix, and the corresponding transformations of H give the matrix illustrated.

The condensation subdivides $S(z+)$ into two subspaces, $S(z-)$ is already divided before the calculation started. The vectors of $S(z+)$ are numbered 1 to 6; 1 is not coupled to any other vector. Two and four are coupled by nonsingular submatrices in column 2 or column 4, i.e., they afford two equivalent representations of $\mathfrak{C}$ or span an irreducible representation of the cubic group. Spaces 2, 5, and 6 span another irreducible representation. The final rearrangement B is shown in Table I, Part C.

One more lemma remains to be proved. It is concerned with the relation between generators in the same class.

If the generators P and Q are in the same class, then a group element R can be chosen such that $Q = R^{-1}PR$; then obviously P and R or Q and R are at least as good a pair of generators as P and Q, and may be better because R may generate group elements which P or Q do not.

Lemma 6

If P, Q, and R are as described above, and P and Q commute, then R does not further reduce any simultaneous eigenspaces of P and Q, but it does couple each simultaneous eigenspace of P and Q belonging to the pair of eigenvalues (λ, μ) to a corresponding eigenspace belonging to (μ, λ) by a square, nonsingular submatrix. All other submatrices of R in the same block of rows or columns are zero.

Proof: If A commutes with P and R, then $AR^{-1}PR = R^{-1}PRA$, and A commutes with Q also.

$$\begin{aligned} Q = A^{-1}QA &= A^{-1}R^{-1}PRA \\ &= A^{-1}R^{-1}A\ A^{-1}PRA\ A^{-1}RA \\ &= (A^{-1}RA)^{-1}P(A^{-1}RA). \end{aligned}$$

This does not imply $A^{-1}RA = R$, as is shown by Turnbull and Aitken[18] where the solution of $Q = X^{-1}PX$ is discussed. However, R maps eigenvectors of P onto eigenvectors of Q belonging to the same eigenvalue, for: let x_λ be an eigenvector of P belonging to the eigenvalue λ. Then $QRx_\lambda = RPx_\lambda = \lambda Rx_\lambda$.

It follows that if the S_α are eigenspaces of P, and the S_β are eigenspaces of Q, the submatrices $H_{\alpha\beta}$ are submatrices of R, and they are either zero or nonsingular according as $\lambda_\alpha \neq \lambda_\beta$ or $\lambda_\alpha = \lambda_\beta$.

Because P and Q commute as well as being nonsingular, an argument similar to that of Lemma 4 shows that a value of x can be found which makes the eigenvalues $(\lambda+\mu x)/(1+x)$ of $X = (P+xQ)/(1+x)$ distinct for all distinct pairs (λ, μ). This separates the $S_{\lambda\mu}$ as far as possible, and all that R does is to force equivalent representations of $\mathfrak{a}$ in $S_{\lambda\mu}$ and $S_{\mu\lambda}$.

In this case, P and Q do not generate all the group elements which would be generated by P and R, but they are sufficient to reduce the commutator algebra of P, Q, and R.

III. APPLICATIONS

So far it has been shown how to reduce a group representation if the group generators are known. The practical application of the method thus depends on knowledge of the group generators, and a wise choice from the several alternative sets, each of which generates a given group, can make great simplifications in the calculations.

In the case of infinite groups such as $GL(n)$, $Su(n)$, $O(n)$, and $O + (n)$, it is sufficient to reduce the smallest subgroup whose linear closure spans the group algebra, e.g., in the case of $O + (3)$, the cubic group generated by $C_4(x)$ and $C_4(y)$ is asequate, and for $O + (n)$, the generalization of this to the group generated by the $n - 1$ fourfold plane rotations $C_4(i)$ $i = 1, 2, \cdots n - 1$ defined by: $C_4(i)$ is $x_i \to x_{i+1}$, $x_{i+1} \to -x_i$ all other $x_j \to x_j$.

Alternatively, the Lie generators may be used. Some of these may be redundant by Lemma 6, since $I_1I_2 - I_2I_1$ is the infinitesimal analog of $P_1^{-1}P_2P_1$.

With this much preamble some examples can be outlined.

The Permutation or Symmetric Group

A set of generators are the disjoint interchanges (12), (23), (34), $\cdots$ $(n - 1, n)$, $(n, 1)$.[9] These form two sets, each set commuting among themselves: σ_1 containing (12), (34), etc. and σ_2 (23), (45), etc. Since the eigenvalues of an interchange are ± 1, the elements of the algebra,

[18] See Ref. 8, Chap. X, problems I, II, and III.

$$S_1 = (12) + 2(34) + 4(56) + \cdots,$$

$$S_2 = (23) + 2(45) + 4(67) + \cdots,$$

have eigenvectors which are simultaneous eigenvectors of σ_1 and σ_2, respectively. They correspond to the P and Q of Lemma 6. The cyclic permutation $(1, 2, 3 \cdots n)$ corresponds to the matrix H of Lemma's 3, 4, and 5, and so S_2 need not be used explicitly.

The Proper Orthogonal Group

It is easily verified that if $C_4(i)$ is defined as in the start of this section; then

$$X_i(\alpha) = \tfrac{1}{2}[1 + C_4(i)^2] + \tfrac{1}{2}[1 - C_4(i)^2]\cos\alpha + [C_4(i) - \tfrac{1}{2}\{1 + C_4(i)^2\}]\sin\alpha$$

is a plane rotation[19] by α in the plane defined by x_i and x_{i+1} and leaves all other coordinates unchanged.

It is also well known that an arbitrary $n \times n$ orthogonal matrix can be factored into a product of plane rotations[16] (generalization of Euler angles), and it follows at once that the n matrices $C_4(i)$ generate a finite group (the generalized Cubic Group) whose linear closure in a field is the algebra of $O + (n)$ in that field. In the case of $O + (n)$, we are usually concerned with real matrices, the field is the field of real numbers, and the enveloping algebra is the real linear closure of the group, i.e., the set of all linear combinations with real coefficients.

Since $C_4(i)$ is of order 4, its possible eigenvalues are ± 1 and $\pm \mathbf{j}$, and the imaginary ones seem incompatible with the requirement of real coefficients.

This difficulty may be avoided as follows: the imaginary eigenvectors are $x_i + jx_{i+1}$ and $x_i - jx_{i+1}$. If only real transformations to diagonal form are permitted, these eigenvectors are a conjugate complex pair which span a two-dimensional space belonging to the eigenvalue -1, of $\{C_4(i)\}^2$. The argument used on the symmetric group may now be repeated using the n matrices $\{C_4(i)\}^2$. In this case all the n matrices commute and only one linear combination,

$$S = C_4(1)^2 + 2C_4(2)^2 + \cdots 2^{n-1}C_4(n)^2,$$

is needed. The reduction of O_3 is illustrated in Table II.

The Improper Orthogonal Group

The single improper element σ defined by $x_1 \to -x_1$, all other x_1 unchanged, must be added to the generalized cubic group discussed above.

Su(n) and *GL(n)*

These two groups have the same enveloping algebra, which is the linear closure of the generalized cubic group defined for $O + (3)$ over the complex field. Here complex eigenvectors are permitted and the proper generators are the matrices $C_4(i)$. These may be divided into two commuting sets, σ_1 containing $C_4(1)$, $C_4(3)$, etc., and σ_2 containing $C_4(2)$, $C_4(4)$, etc. The calculation is the same as for the symmetric group and the transformation corresponding to $(1, 2, 3 \cdots n)$ is

$$x_i \to x_{i+1} i = 1 \cdots n - 1, \qquad x_n \to x_1.$$

Generators for Groups in General

The methods given so far are adequate for theoretical physics because the generators are known for all the groups which are of interest. However, it is interesting to speculate on applications to pure mathematics and the study of group structure. Ideally one would like to find a nonredundant set of generators for an arbitrary group, using the minimum of knowledge of its structure. The author has not been completely successful in doing this, but the following method certainly uses very little explicit knowledge of the group, although the generators obtained are redundant. However they are not more in number than the elements in the smallest class for simple groups and probably even less redundant for more complicated groups.

Simple Groups

These are groups having no invariant subgroup other than themselves and the unit element.[20] Therefore a complete set of conjugate elements, (i.e., a class) are a set of generators, since a class generates an invariant subgroup. Lemma 6 shows that, unless all the members of the class commute (which is a trivial case since then they can all be simultaneously transformed to diagonal form), then some members of the class are redundant; for if P and Q are in the class, so is $R = P^{-1}QP$, and is redundant.

Suppose that **C** is a class of a simple group **G**, and that in **C** there is a set $\mathfrak{d}_0$ of elements which commute with each other, but do not all commute with any other element of **C**.

[19] F. D. Murnaghan, *The Unitary and Rotation Groups* (Spartan Books, Washington, D. C., 1962) Chap. 2.

[20] R. D. Carmichael, *Groups of Finite Order* (Dover Publications, Inc., New York, 1956), Chap. II, Sec. IV, p. 48. See also, W. Burnside, *Groups of Finite Order*, 2nd ed. (Cambridge University Press, Cambridge, England, 1911; Dover Publications, Inc., New York, 1955), Chaps. IV and V.

TABLE II. Reduction of symmetrized Kronecker square of O_3.

A.		+1	+1	−1	−1	$(i, -i)$	$(i, -i)$
		y^2	$(x^2+z^2)/\sqrt{2}$	$(x^2-z^2)/\sqrt{2}$	xy	xy	yz
+1	z^2	0	$1/\sqrt{2}$	0	$-1/\sqrt{2}$	0	0
+1	$(x^2+y^2)/\sqrt{2}$	$1/\sqrt{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
−1	$(x^2-y^2)/\sqrt{2}$	$-1\sqrt{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
−1	xy	0	0	0	0	1	0
$(i, -i)$	yz	0	0	0	0	0	1
$(i, -i)$	zx	0	0	0	1	0	0

B.		+1	+1′	−1	−1	$(i, -i)$	$(i, -i)$
		$\frac{x^2+y^2+z^2}{\sqrt{3}}$	$\frac{x^2-2z^2+y^2}{\sqrt{6}}$	$(x^2-z^2)/\sqrt{2}$	xz	xy	yz
+1	$\frac{x^2+y^2+z^2}{\sqrt{3}}$	1	0	0	0	0	0
+1′	$\frac{x^2-2y^2+z^2}{\sqrt{6}}$	0	$-\frac{1}{2}$	$+\sqrt{3}/2$	0	0	0
−1	$\frac{x^2-y^2}{\sqrt{2}}$	0	$\sqrt{3}/2$	$-1/2$	0	0	0
−1′	xy	0	0	0	0	1	0
$(i, -i)$	yz	0	0	0	0	0	1
$(i, -i)$	zx	0	0	0	1	0	0

Let γ_0 be the group generated by δ_0. The group γ_0 is not invariant wrt conjugation with every element of **C**, for if it were, then γ_0 would be an invariant Abelian subgroup of the group generated by **C**, i.e., of **G**, and since **G** is simple, this cannot occur except trivially.

Therefore, there exists some element of R_1 of **C** for which $R_1^{-1}\gamma_0 R_1 = \gamma_1 \neq \gamma_0$, i.e., there exists γ_1, a subgroup conjugate to γ_0, generated by $\delta_1 = R_1^{-1}\delta_0 R_1$. If δ_0 and δ_1 do not exhaust C, then there exists R_2 in **C** such that either

$$R_2^{-1}\gamma_0 R_2 = \gamma_2 \neq \gamma_1 \quad \text{or} \quad \gamma_0,$$

$$R_2^{-1}\gamma_1 R_2 = \gamma_2 \neq \gamma_1 \quad \text{or} \quad \gamma_0;$$

or conjugation by element of **C** maps γ_0 and γ_1 onto themselves, or each other.

This process may be continued until **C** has been exhausted, or a set of conjugate subgroups have been obtained which are mapped onto each other or themselves by elements of C. Lemma 6 shows that the commuting sets δ_0, δ_1, δ_2 $\cdots$ etc. are sufficient to reduce any representation.

Extension to General Groups

Suppose now that **G** is not simple, but has a maximal invariant subgroup **F**,[21] their orders being g and f, respectively, and assume that the representa-

[21] See Ref. 20, Carmichael, Chap. III; Burnside, Chap. V.

tion of **F** has already been reduced. The expansion of **G** in cosets[22] wrt **F** is

$$G = \mathbf{F} + S_1\mathbf{F} + S_2\mathbf{F} + \cdots S_t\mathbf{F}.$$

Let **C** be a class of **F**. Consider the linear combination

$$H = \sum \alpha_{\mathbf{C}} \sum_{\mathbf{C}} P,$$

where the α's are numbers. This is a linear combination of class sums with arbitrary coefficients $\alpha_{\mathbf{C}}$, and commutes with every element of **F**. Partition **H** and each of the S_i into submatrices conformably to the reduction of **F**. Using the notation of Lemma 2 $\mathbf{H}(p, r; q, s) = \delta_{pq}\delta_{rs}(f\alpha_c/\chi_{\mathbf{C}})\mathbf{1}$ where $\chi_{\mathbf{C}}^r$ is the character of the class **C** in the rth irreducible representation.

Because **F** is an invariant subgroup of **G**, $S_i(p, r; q, s)\mathbf{H}(q, s; t, u) = \mathbf{H}(p, r; v, w)S_i(v, w; t, u)$ which implies $S_i(p, r; q, s) = 0$ unless $r = s$, i.e., the S_i only couple multiple occurrences of the same irreducible representation of **F**. Assemble all the $S_i(p, r; q, r)$ into one square submatrix, and call it $S_i(r)$. Then the way the $S_i(r)$ have been constructed makes them faithfully represent the quotient group **G/F**. Since **F** is maximal, **G/F** is simple, and the method given for simple groups applies to the reduction of the S_i.

Any finite group has a composition series of quotient groups which are simple and so the method applies to all finite groups. For matrix representations of infinite groups, the finite subgroup which spans the group algebra must be used.

The results of Sec. IV can be extended to give a more detailed discussion of group structure by considering the analysis of the structure of simple groups more carefully in the ways suggested by the Refs. 1 and 2. This is the finite-group analog of the Cartan theory of continuous groups.

[22] See Ref. 5, Chap. II, Sec. 12, Theorem 3; also Ref. 20, Carmichael, Chap. II, Sec. 10, p. 44; and Burnside, Chap. III, Sec. 22, p. 26.

ACKNOWLEDGMENTS

This paper really presents an alternative to the usual point of view on reduction of group representations. It has been developed slowly over a period of several years, and owes much to comments by mathematicians and theoretical physicists at The University of Otago, and the Department of Scientific and Industrial Research, Applied Mathematics Laboratory, in New Zealand; The Atomic Energy Research Establishment, Harwell, England; and Argonne National Laboratory. The author gratefully acknowledges many discussions with staff and students at these establishments.

In addition the author is particularly grateful to M. Hamermesh of Argonne National Laboratory for encouragement to complete this work, and to M. J. D. Powell of A.E.R.E. for precise and constructive criticism of an early draft of the paper.

Representations for the Nonrelativistic Coulomb Green's Function*

KOICHI MANO
Wentworth Institute, Boston, Massachusetts
(Received 24 October 1963)

Derivation of the representations for the nonrelativistic Coulomb Green's function is discussed. It is shown that the recently published representations, both in integral and closed forms, can be obtained from an expression for the difference of the diverging and converging wave solutions of the Green's function which is given in terms of the wavefunctions summed by Gordon.

INTRODUCTION

REPRESENTATIONS for the nonrelativistic Coulomb Green's function have recently been obtained both in integral[1] and closed forms.[2] In the present note we wish to derive these expressions from a slightly different viewpoint, having the expression[3]

$$G(\mathbf{r}, \mathbf{r}'; k) \equiv G^{(+)}(\mathbf{r}, \mathbf{r}'; k) - G^{(-)}(\mathbf{r}, \mathbf{r}'; k)$$
$$= -\pi i k \int d\omega_k \psi^{(S)}(\mathbf{r}, \mathbf{k}) \psi^{(S)}(\mathbf{r}', \mathbf{k})^* \qquad (1)$$

as the starting point. In Eq. (1), $G^{(\pm)}$ represent the diverging and converging wave solutions of the Green's function and $\psi^{(S)}$ stands for either $\psi^{(+)}$ or $\psi^{(-)}$ with

$$\psi^{(\pm)}(\mathbf{r}, \mathbf{k}) = (2\pi)^{-\frac{3}{2}} e^{-\pi/2ka} \Gamma(1 \pm n)$$
$$\times F(\mp n, 1; \pm i(kr \mp \mathbf{k}\cdot\mathbf{r})), \qquad (2)$$

where F denotes the confluent hypergeometric function ${}_1F_1$ and n is written for i/ka.

It is seen therefore that $G^{(+)}$ or $G^{(-)}$ is characterized as the diverging or converging wave part of the last expression of Eq. (1) when the substitution from Eq. (2) is made into it.

I.

In this section we show the derivation of a representation which is the contour integral version of that obtained in Ref. 1.

Before going any further we mention that the procedure in Ref. 1 consists of first constructing in the usual manner the partial-wave Green's function from the solutions of the radial equation, and then summing it over all the angular momentum states to arrive at the final result. Clearly, the present work proceeds essentially in the reversed order, as can be seen from the fact that the expression for the ψ's given by Eq. (2) represent the wavefunctions already summed over all states.

Out of several alternatives for the integral on the right-hand side of Eq. (1), we consider

$$\int d\omega_k \psi^{(+)}(\mathbf{r}_<, \mathbf{k}) \psi^{(+)}(\mathbf{r}_>, \mathbf{k})^*, \qquad (3)$$

where $\mathbf{r}_>$ and $\mathbf{r}_<$ have the usual meaning.

In order to rewrite Eq. (3) we first give contour integral representations for the function $F(a, c; x)$ when c is a positive integer:

$$F(a, c; x) = \frac{C(a, c; P_\pm)}{B(a, c-a)}$$
$$\times \int_{P_\pm}^{(0+,1+)} du e^{xu} u^{a-1} (1-u)^{c-a-1}. \qquad (4)$$

In Eq. (4), B stands for the Beta function and P_+ (or P_-) represents the starting point of the integration taken on the upper (or lower) portion of the path which encloses 0 and 1. The values of the C's, $C(a, c; P_+) = -C(c-a, c; P_-) = (e^{2\pi i a} - 1)^{-1}$, are fixed by deforming the path in the usual way in which we assumed that $\arg u = \arg(1-u) \to 0$ at the starting point when it approaches the real axis. We may add that, although one form of Eq. (4) can be obtained from the other, it turns out in some connections to be helpful to have both forms available.

Next, we note the following representation for the decomposition of an F function with c equal to an integer,[4]

* Work performed under contract with the Air Force Cambridge Research Laboratories.

[1] E. H. Wichmann and C. H. Woo, J. Math. Phys. **2**, 178 (1961). This paper is referred to as WW hereafter.

[2] L. Hostler and R. H. Pratt, Phys. Rev. Letters, **10**, 469 (1963).

[3] K. Mano, J. Math. Phys. **4**, 522 (1963), Eqs. (39) and (43+). The notations used here are the same as those in this paper except for the trivial change that the general form of the Coulomb field is considered here with $a = \hbar^2/\mu ZZ'e^2$ for arbitrary Z and Z'.

[4] This can be obtained from the decomposition formula for the F function. See, for example, Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 1, p. 259, Eq. (7).

$$F(a, c; x) = \frac{1}{B(a, c-a)} \times \left[\frac{e^{-\pi i(c-a-1)}}{1-e^{-2\pi i a}} \int_{\infty}^{(1+)} du e^{xu} u^{a-1}(u-1)^{c-a-1} + \frac{e^{+\pi i(a-1)}}{e^{+2\pi i a}-1} \int_{-\infty}^{(0+)} du e^{xu}(-u)^{a-1}(1-u)^{c-a-1} \right], \quad (5)$$

where arg u = arg $(u-1) = 0$ and arg $(-u)$ = arg $(1-u) = 0$ for the first and second integral, respectively. It is to be remembered that the former has to do with the diverging wave part and the latter with the converging one when $x = +i(kr_> - \mathbf{k}\cdot\mathbf{r}_>)$.

Use of the Kummer transformation and the substitution of Eqs. (4) and (5) into F of the first and second ψ functions, respectively, of Eq. (3) give rise to an expression for G in which an integral

$$\Omega \equiv \int d\omega_k e^{i\mathbf{k}\cdot(\mathbf{r}_< u' - \mathbf{r}_> u)} \quad (6)$$

is included. This integral can be evaluated as

$$\Omega = 4\pi \frac{\sin k\,|\mathbf{r}_< u' - \mathbf{r}_> u|}{k\,|\mathbf{r}_< u' - \mathbf{r}_> u|} = -\frac{(4\pi)^2}{2ik} G_0(\mathbf{r}_< u', \mathbf{r}_> u; k) \quad (7)$$

by the use of the Gegenbauer's formulas.[5] In Eq. (7),

$$|\mathbf{r}_< u' - \mathbf{r}_> u| = [(r_< u')^2 + (r_> u)^2 - 2(\mathbf{r}_< u')\cdot(\mathbf{r}_> u)]^{\frac{1}{2}},$$

and $G_0 = G_0^{(+)} - G_0^{(-)}$, with $G_0^{(\pm)}$ standing for the Green's function for the free particle.

In this way we arrive at the contour integral representation for $G^{(+)}$ given by

$$G^{(+)}(\mathbf{r}, \mathbf{r}'; k) = \frac{-n^2}{|\Gamma(1+n)|^2} \left[\frac{1}{e^{2\pi i n}-1} \times \int_{P_+}^{(0+,1+)} du' e^{ikr_<(1-u')} u'^n (1-u')^{-n-1} \right] \times \left[\frac{1}{e^{2\pi i n}-1} \int_{\infty}^{(1+)} du e^{ikr_>(u-1)} u^{-n}(u-1)^{n-1} \right] \times G_0^{(+)}(\mathbf{r}_< u', \mathbf{r}_> u; k), \quad (8)$$

where the $\int_{P_+}^{(0+,1+)}$ form is employed for the F function in the first ψ of Eq. (3). The corresponding expression for $G^{(-)}$ is self-explanatory. It is the contour integral representation mentioned earlier. In the limit $n \to 0$ we see that the residue calculation of the right-hand side of Eq. (8) produces $G_0^{(+)}$ correctly.

[5] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, London, England, 1944), 2nd ed., p. 368, Eq. (1); p. 363, Eq. (3).

That Eq. (8) for $G^{(+)}$ is equivalent to Eq. (31) of WW can be seen by observing the following. If we assume that Re $n > 0$, i.e., that k possesses an infinitesimal positive imaginary part, we may replace the contour integral in the second set of brackets of Eq. (8) by an ordinary integral. If we further suppose that the similar argument applies also to the integral in the first set of brackets, then we have for Eq. (8)

$$G^{(+)} = \frac{n^2}{4\pi\,|\Gamma(1+n)|^2} \times \int_0^1 du' \int_1^{\infty} du\, u'^n (1-u')^{-n-1} u^{-n}(u-1)^{n-1} \times |\mathbf{r}_< u' - \mathbf{r}_> u|^{-1} \exp ik[r_<(1-u') + r_>(u-1) + |\mathbf{r}_< u' - \mathbf{r}_> u|]. \quad (9)$$

We notice that this is identical with the expression one would obtain from Eq. (31) of WW by performing integrations by parts on s and t as our n corresponds to $-i\epsilon$ there. As is evident, however, the condition $1 >$ Re $(-n) > 0$ required for the rewriting of the first integral in Eq. (8) is incompatible with the similar condition for the second one. This is the very reason why it is necessary to introduce the partial derivative with respect to t in the integrand of WW Eq. (31), in order to modify the condition to Re $(1-n) > 0$ which can then be made compatible with Re $n > 0$. Incidentally, it should be noted that the partial derivative with respect to s in WW Eq. (31) is not required except for reasons of symmetry.

We might add that the following representation for an F function with c equal to an integer:

$$F(a, c; x) = \frac{C(a, \frac{1}{2}c; P_\pm)}{B(a, \frac{1}{2}c - a)} \times \int_{P_\pm}^{(0+,1+)} du\, u^{a-1}(1-u)^{\frac{1}{2}c-a-1} F(\tfrac{1}{2}c, c; xu),$$

which may be given in terms of the Bessel function, leads to the expression

$$L_l(r, k) = -\frac{\Gamma(1+n)}{2\pi i} e^{-i\eta(l,k)} e^{-\pi/2ka \pm \pi/ka} \times \int_{P_\pm}^{(0+,1+)} du\, u^n (1-u)^{-n-1} e^{ikr(1-u)} j_l(kru),$$

which corresponds to Eq. (2) of WW.

The decomposition of the F function in [3]

$$L_l(r, k) = N_l(r, k) F(l+1-n, 2l+2; +2ikr)$$

gives rise to the expression for $H_l^{(+)}$, the diverging wave part of L_l,

$$H_l^{(+)}(r, k) = -\frac{\Gamma(1-n)}{2\pi} e^{+i\eta(l,k)} e^{+3\pi/2k\alpha} \times \int_{\infty}^{(1+)} du\, u^{-n}(u-1)^{n-1} e^{ikr(u-1)} h_l^{(1)}(kru).$$

This equation corresponds to Eq. (5) of WW.

II.

To obtain the closed form given in Ref. 2, we may take

$$\int d\omega_k \psi^{(+)}(\mathbf{r}_>, \mathbf{k}) \psi^{(+)}(\mathbf{r}_<, \mathbf{k})^*$$

for the integral in Eq. (1). With the use of the identity

$$F(a, c; x) = e^{x/2} x^{-\frac{1}{2}c} M_{\frac{1}{2}c-a, \frac{1}{2}(c-1)}(x),$$

where $M_{\kappa,\mu}(x)$ is the Whittaker function, and the expression for the product of two Whittaker functions of different arguments,[6] we obtain

$$G = -\frac{ik}{2\pi^2} e^{-\pi/k\alpha} K_n(\mathbf{r}_>, \mathbf{r}_<) \tag{10}$$

with

$$K_n(\mathbf{r}_>, \mathbf{r}_<) = e^{ik(r_>-r_<)} \int d\omega_k \times \int_0^{\infty} d\rho\, \rho^3 e^{-\rho^2} \int_0^{\pi/2} d\varphi (\cos\varphi)^{1+2n} (\sin\varphi)^{1-2n} \times J_0(2\rho[+ikr_>(1-\cos\theta_>)]^{\frac{1}{2}} \cos\varphi) \times J_0(2\rho[-ikr_<(1-\cos\theta_<)]^{\frac{1}{2}} \sin\varphi), \tag{11}$$

where $\theta_>$ denotes the angle between $\mathbf{k}$ and $\mathbf{r}_>$ and similarly for $\theta_<$. If both of the J_0's in Eq. (11) are represented by a Neumann series[7]

$$(\tfrac{1}{2}pz)^{\mu-\nu} J_\nu(pz) = p^\mu \sum_{l=0}^{\infty} \frac{\Gamma(\mu+l)}{l!\,\Gamma(\nu+1)} \times {}_2F_1(\mu+l, -l; \nu+1; p^2)(\mu+2l) J_{\mu+2l}(z), \tag{12}$$

with $\mu = 1$, $\nu = 0$, and $p^2 = \frac{1}{2}(1-\cos\theta)$, the integration over the angles of the $\mathbf{k}$ space can be performed immediately when the ${}_2F_1$ function in Eq. (12) is recognized as the Legendre polynomial. We may then proceed to carry out the integration over ρ by the help of the Weber's second exponential integral.[8] The result is given by

$$K_n = 2\pi \int_0^{\pi/2} d\varphi (\cos\varphi)^{1+2n} (\sin\varphi)^{1-2n} e^{-ik(r_>+r_<)\cos 2\varphi} \times \left\{ \frac{1}{k(r_> r_<)^{\frac{1}{2}} \sin 2\varphi} \sum_l (2l+1) \times P_l(\cos\Theta) I_{2l+1}(2k[r_> r_<]^{\frac{1}{2}} \sin 2\varphi) \right\}, \tag{13}$$

where Θ is the angle between $\mathbf{r}_>$ and $\mathbf{r}_<$. Rewriting I_{2l+1} in terms of J_{2l+1} by

$$J_{2l+1}(iz) = I_{2l+1}(z) e^{i\pi\frac{1}{2}(2l+1)} \qquad (-\pi < \arg z \le \tfrac{1}{2}\pi),$$

and using Eq. (12) once again with $\mu = 1$, $\nu = 0$, and $p^2 = \frac{1}{2}(1+\cos\Theta)$, we see that the bracketed expression in Eq. (13) is equal to $J_0(+ik[\sigma^2-\rho^2]^{\frac{1}{2}} \sin 2\varphi)$, where $\sigma = r_> + r_<$ and $\rho = |\mathbf{r}_> - \mathbf{r}_<|$. With a trivial change of the variable, Eq. (11) becomes[9]

$$K_n = \frac{\pi}{2} \int_0^{\pi} d\varphi\, e^{+ik\sigma\cos\varphi} J_0(-ik[\sigma^2-\rho^2]^{\frac{1}{2}} \sin\varphi) \times (\cot\tfrac{1}{2}\varphi)^{-2n} \sin\varphi, \tag{14}$$

in which use is made of the relation $J_0(-z) = J_0(z)$. From the identity

$$zJ_0(z) = zJ_1'(z) + J_1(z) = (d/dz)[zJ_1(z)],$$

we find

$$+ik\rho \sin\varphi J_0(-ik[\sigma^2-\rho^2]^{\frac{1}{2}} \sin\varphi) = (\partial/\partial\rho)\{[\sigma^2-\rho^2]^{\frac{1}{2}} J_1(-ik[\sigma^2-\rho^2]^{\frac{1}{2}} \sin\varphi)\},$$

so that Eq. (14) can be written

$$K_n = \frac{\pi}{2} \frac{1}{ik\rho} \frac{\partial}{\partial\rho} \left\{ [\sigma^2-\rho^2]^{\frac{1}{2}} \int_0^{\pi} d\varphi\, e^{+ik\sigma\cos\varphi} \times J_1(-ik[\sigma^2-\rho^2]^{\frac{1}{2}} \sin\varphi)(\cot\tfrac{1}{2}\varphi)^{-2n} \right\}. \tag{15}$$

It is immediately seen that the bracketed expression in Eq. (15) is essentially equal to the product of two Whittaker functions of arguments $-ikx$ and $-iky$ if we write $x = \sigma + \rho$ and $y = \sigma - \rho$.[10] Therefore, G now reads

$$G(\mathbf{r}, \mathbf{r}'; k) = \frac{|\Gamma(1+n)|^2}{4\pi\,|\mathbf{r}-\mathbf{r}'|} e^{-\pi/k\alpha} \times \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) \mathfrak{M}_{-n,\frac{1}{2}}(-ikx) \mathfrak{M}_{-n,\frac{1}{2}}(-iky), \tag{16}$$

where

$$\mathfrak{M}_{\kappa,\mu}(x) = M_{\kappa,\mu}(x)/\Gamma(1+2\mu).$$

[6] H. Buchholz, *Die Konfluente Hypergeometrische Funktion*, Part II of Ergebnisse der Angewandten Mathematik (Springer-Verlag, Berlin, 1953), p. 83, Eq. (2), with care for an obvious misprint in the power of ρ.

[7] See Ref. 5, p. 140 (Sec. 5.21), Eq. (3).

[8] See Ref. 5, p. 395.

[9] This can also be obtained from Ref. 3, Eq. (5b) by a procedure analogous to that given above.

[10] See Ref. 6, p. 85, Eq. (3α).

The decomposition of $\mathfrak{M}_{-n,\frac{1}{2}}(-ikx)$ into two W functions leads to the forms of $G^{(+)}$ and $G^{(-)}$ which are equivalent to those in Ref. 2 since our $-n$ corresponds to their $i\nu$.

Finally, we may remark that the expression for G given by Eq. (16), or that given by the combination of Eqs. (10) and (14), reduce to G_0 in the limit $n \to 0$. To see this we only need to notice the relation $\mathfrak{M}_{0,\frac{1}{2}\mu}(-ikz) \propto (\pi z)^{\frac{1}{2}} J_{\frac{1}{2}\mu}(k\frac{1}{2}z) \propto \sin k\frac{1}{2}z$ for Eq. (16), or apply the second finite Sonine integral[11] to Eq. (14).

[11] See Ref. 5, p. 376.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Theory of Vibrational Structure in Optical Spectra of Impurities in Solids. II. Multiplets

D. E. McCumber
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received 11 October 1963)

The results of a previous paper describing vibrational structure in the spectra of singlet-to-singlet optical transitions of impurities in solids are extended to a general class of multiplet-to-multiplet transitions. The transitions considered are those between degenerate or nondegenerate multiplets which are not interconnected by the interaction coupling the impurities to the lattice phonons. The phonon field couples impurity states within each multiplet but does not couple states in different multiplets. The results apply to systems which display strong Jahn–Teller distortions as well as to those which do not. As a specific application of our results we consider the singlet-to-doublet transitions of simple three-level systems having C_3 and C_{3v} symmetry.

1. INTRODUCTION

In a previous paper[1] we derived expressions relevant to vibrational structure in the spectra of singlet-to-singlet optical transitions of impurities in solids. The vibrational structure we considered arose from phonon–impurity interactions which shifted the impurity energy levels but which did not mix the different impurity states. In the present paper we extend our previous results to transitions between degenerate or nondegenerate multiplets. We restrict ourselves to phonon–impurity interactions which couple the impurity states within each multiplet but which do not couple states from different multiplets. In the special case for which the multiplets are singlets, our present results reduce to those of I.

Briefly, the vibrational structure we discuss arises whenever the characteristics of the lattice in the neighborhood of the impurity are different for the ground-state and excited-state impurity multiplets.[2–9] For singlet impurity states only those properties of the lattice which have the full point-group symmetry of the impurity site can be sensitive to the state occupation; however, for degenerate impurity multiplets, lattice properties with more complicated symmetries (such as those responsible for Jahn–Teller effects) can be sensitive to the impurity state. In spite of this rather important difference between the singlet and multiplet states of an impurity in a lattice, their optical spectra are qualitatively very similar. For each set of degenerate states forming a basis for an irreducible representation of the impurity-site symmetry group, the spectrum will display a single sharp no-phonon line accompanied by adjacent vibrational structure.

[1] D. E. McCumber, J. Math. Phys. **5**, 221 (1964). In what follows we refer to this paper as I. We indicate particular equations in the paper by the prefix I; for example, we indicate Eq. (1.9b) of I by (I1.9b). In our usage the terms *singlet* and *multiplet* do not refer exclusively to spin multiplicity, but indicate the total number of relevant initial or final electronic states.

[2] M. Lax, J. Chem. Phys. **20**, 1752 (1952).
[3] S. I. Pekar, Uspekhi Fiz. Nauk SSSR **50**, 197 (1953).
[4] M. Lax and E. Burstein, Phys. Rev. **100**, 592 (1955).
[5] H. Gummel and M. Lax, Ann. Phys. **2**, 28 (1957).
[6] D. L. Dexter, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1958), Vol. 6, p. 353.
[7] J. J. Markham, Rev. Mod. Phys. **31**, 956 (1959).
[8] B. S. Gourary and A. A. Maradudin, J. Phys. Chem. Solids **13**, 88 (1960).
[9] E. O. Kane, Phys. Rev. **119**, 40 (1960); W. E. Lamb, Jr., Phys. Rev. **55**, 190 (1939); E. D. Trifonov, Dokl. Akad. Nauk (USSR) **147**, 826 (1962) [English transl.: Soviet Phys.—Doklady **7**, 1105 (1963)].

The decomposition of $\mathfrak{M}_{-n,\frac{1}{2}}(-ikx)$ into two W functions leads to the forms of $G^{(+)}$ and $G^{(-)}$ which are equivalent to those in Ref. 2 since our $-n$ corresponds to their $i\nu$.

Finally, we may remark that the expression for G given by Eq. (16), or that given by the combination of Eqs. (10) and (14), reduce to G_0 in the limit $n \to 0$. To see this we only need to notice the relation $\mathfrak{M}_{0,\frac{1}{2}\mu}(-ikz) \propto (\pi z)^{\frac{1}{2}} J_{\frac{1}{2}\mu}(k\frac{1}{2}z) \propto \sin k\frac{1}{2}z$ for Eq. (16), or apply the second finite Sonine integral[11] to Eq. (14).

[11] See Ref. 5, p. 376.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Theory of Vibrational Structure in Optical Spectra of Impurities in Solids. II. Multiplets

D. E. McCumber
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received 11 October 1963)

The results of a previous paper describing vibrational structure in the spectra of singlet-to-singlet optical transitions of impurities in solids are extended to a general class of multiplet-to-multiplet transitions. The transitions considered are those between degenerate or nondegenerate multiplets which are not interconnected by the interaction coupling the impurities to the lattice phonons. The phonon field couples impurity states within each multiplet but does not couple states in different multiplets. The results apply to systems which display strong Jahn–Teller distortions as well as to those which do not. As a specific application of our results we consider the singlet-to-doublet transitions of simple three-level systems having C_3 and C_{3v} symmetry.

1. INTRODUCTION

In a previous paper[1] we derived expressions relevant to vibrational structure in the spectra of singlet-to-singlet optical transitions of impurities in solids. The vibrational structure we considered arose from phonon–impurity interactions which shifted the impurity energy levels but which did not mix the different impurity states. In the present paper we extend our previous results to transitions between degenerate or nondegenerate multiplets. We restrict ourselves to phonon–impurity interactions which couple the impurity states within each multiplet but which do not couple states from different multiplets. In the special case for which the multiplets are singlets, our present results reduce to those of I.

Briefly, the vibrational structure we discuss arises whenever the characteristics of the lattice in the neighborhood of the impurity are different for the ground-state and excited-state impurity multiplets.[2–9] For singlet impurity states only those properties of the lattice which have the full point-group symmetry of the impurity site can be sensitive to the state occupation; however, for degenerate impurity multiplets, lattice properties with more complicated symmetries (such as those responsible for Jahn–Teller effects) can be sensitive to the impurity state. In spite of this rather important difference between the singlet and multiplet states of an impurity in a lattice, their optical spectra are qualitatively very similar. For each set of degenerate states forming a basis for an irreducible representation of the impurity-site symmetry group, the spectrum will display a single sharp no-phonon line accompanied by adjacent vibrational structure.

[1] D. E. McCumber, J. Math. Phys. **5**, 221 (1964). In what follows we refer to this paper as I. We indicate particular equations in the paper by the prefix I; for example, we indicate Eq. (1.9b) of I by (I1.9b). In our usage the terms *singlet* and *multiplet* do not refer exclusively to spin multiplicity, but indicate the total number of relevant initial or final electronic states.

[2] M. Lax, J. Chem. Phys. **20**, 1752 (1952).
[3] S. I. Pekar, Uspekhi Fiz. Nauk SSSR **50**, 197 (1953).
[4] M. Lax and E. Burstein, Phys. Rev. **100**, 592 (1955).
[5] H. Gummel and M. Lax, Ann. Phys. **2**, 28 (1957).
[6] D. L. Dexter, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1958), Vol. 6, p. 353.
[7] J. J. Markham, Rev. Mod. Phys. **31**, 956 (1959).
[8] B. S. Gourary and A. A. Maradudin, J. Phys. Chem. Solids **13**, 88 (1960).
[9] E. O. Kane, Phys. Rev. **119**, 40 (1960); W. E. Lamb, Jr., Phys. Rev. **55**, 190 (1939); E. D. Trifonov, Dokl. Akad. Nauk (USSR) **147**, 826 (1962) [English transl.: Soviet Phys.—Doklady **7**, 1105 (1963)].

[Additional vibrational structure deriving from the same no-phonon line can lie near distant no-phonon lines when we are concerned with nondegenerate multiples. (An example is briefly discussed in Sec. 8.)] The extent of the vibrational structure adjacent to the no-phonon line depends upon the strength of the phonon–impurity interaction—more precisely, upon the difference of the ground-state and excited-state interactions—but, for moderate interactions, typically extends to several times the *restrahlen* frequencies.

Our basic results have the form of Baker–Hausdorff,[10] generalized cumulant (or semi-invariant),[11] or "linked-cluster"[12] expansions formally very similar to the singlet-to-singlet expansion of Eq. (I1.9b).[1] The essential difference between the present expansions and those in I stems from the fact that the general phonon interaction operator $A(t)$ in the expansions of I has here been replaced by a phonon operator *matrix* $\mathbf{A}(t)$ whose matrix components couple the different states of the impurity multiplets.

In I we remarked that the linked-cluster expansion (I1.9b) terminates after two terms when the phonon Hamiltonian is harmonic and the phonon coupling operator $A(t)$ is linear in the phonon annihilation–creation operators $\{a_q, a_q^\dagger\}$.[13] This particular result does not apply to the more general expansions of this paper. The matrix character of the coupling $\mathbf{A}(t)$ generally precludes the cancellations necessary for the expansions to terminate. In spite of this fact we believe that the Baker–Hausdorff or generalized cumulant expansions developed here provide useful representations of the optical spectra of impurities in solids.

As we noted in Sec. 3 of I, simple expansions of spectral correlation functions, in powers of the phonon–impurity coupling, relate to moment analyses of the spectra. [Kubo calls time-dependent expansions of the type (I1.9a) generalized moment expansions.[11]] Where spectral details are important, moment methods are generally unsuitable. One approach, which avoids the most serious pitfalls of the simple perturbative expansion but which is different from the cumulant-expansion method we use below, is based upon an analysis of the frequency-plane singularites of the Laplace or Fourier transforms of the spectral correlation functions.[14,15] This approach is quite useful for determining the widths and positions of the no-phonon lines,[16] but it does not *conveniently* indicate the strengths of those lines nor the characteristics of the vibrational structure, especially when the phonon–impurity interaction is strong. Moreover, it inevitably involves a power-series expansion (of a characteristic-frequency function[17]) which is intrinsically no better than the Baker–Hausdorff expansion.[18] Since the Baker–Hausdorff expansions we have developed display the spectral properties in a convenient form even when the phonon–impurity interactions are strong, those expansions should prove useful for the analysis of many real physical systems. As is the case for all expansion methods, however, the utility of the Baker–Hausdorff or cumulant expansion method can only be verified by a comparison of its predictions with experimental results.

In Sec. 2 we introduce detailed notation. We also discuss the fundamental reciprocity between emission and absorption spectra for systems whose multiplet populations are initially in thermal equilibrium. In Sec. 3 we briefly discuss the role played in our formalism by the subgroup of lattice symmetry operations which leave the impurity site invariant. In the following Sec. 4 we derive expressions describing the absorption spectrum appropriate to the excitation of a degenerate multiplet from a singlet. These singlet-to-multiplet results are generalized in the next section where we consider transitions between degenerate multiplets and discuss the role played by Jahn–Teller effects. In Sec. 6, expressions of greater generality are derived which describe transitions between arbitrary multiplets. As a simple application of our results, we consider in Sec. 7 the singlet-to-doublet transitions of simple three-level impurity systems having C_3 and C_{3v} impurity-site symmetry. In the final section we summarize our results and mention some areas which require further study.

---

[10] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962); R. Englman and P. Levi, J. Math. Phys. **4**, 105 (1963).

[11] R. Kubo, J. Phys. Soc. Japan **17**, 110 (1962).

[12] D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic Press Inc., New York, 1961), pp. 35–44.

[13] R. C. O'Rourke, Phys. Rev. **91**, 265 (1953); S. Kiode, Z. Naturforsch. **15a**, 123 (1960).

[14] D. E. McCumber, Bull. Am. Phys. Soc. **8**, 256 (1963).

[15] D. E. McCumber, Phys. Rev. **130**, 2271 (1963).

[16] D. E. McCumber and M. D. Sturge, J. Appl. Phys. **34**, 1682 (1963).

[17] This expansion is analogous to the mass-operator expansions of elementary-particle field theories. Compare also, Ref. 12, pp. 65–69.

[18] For the special harmonic-Hamiltonian linear-coupling case for which the expansion (I1.9b) terminates, the frequency-function expansion is rather unsatisfactory. Considerable care is required to avoid a spurious width for the no-phonon line. This fact provided much of the motivation for our utilizing the cumulant or Baker–Hausdorff techniques reported here and in I.

2. EMISSION AND ABSORPTION SPECTRA

In this paper we restrict ourselves to a single localized impurity system imbedded in an insulating crystal. The lattice vibrations of the crystal constitute the only mechanism by which significant amounts of energy are transported to and from the impurity site. We consider an impurity system which in the absence of phonon–impurity interactions has two sets of energy eigenstates, each set forming what we call a "multiplet" (Fig. 1). We characterize the states of each multiplet by the multiplet index $M = 0$ or 1 and by auxiliary indices μ or ν necessary if the M-multiplet dimension $l_M > 1$. For definiteness we refer to the $M = 0$ and $M = 1$ multiplets as the ground-state and excited-state multiplets, respectively.

In a rigid (static) lattice we characterize the ground-state levels by their energies $\mathcal{E}_{0\mu}$, $\mu = 1$ to l_0, and in a second-quantized notation by the annihilation–creation operators $\{\psi_{0\mu}, \psi^\dagger_{0\mu}\}$. The corresponding quantities for the excited-state levels are the energies $\mathcal{E}_{1\nu}$, $\nu = 1$ to l_1, and the operators $\{\psi_{1\nu}, \psi^\dagger_{1\nu}\}$. The annihilation–creation operators satisfy the familiar Fermi–Dirac anticommutation relations

$$\begin{aligned}[\psi_{m\sigma}, \psi_{m'\sigma'}]_+ &= [\psi^\dagger_{m\sigma}, \psi^\dagger_{m'\sigma'}]_+ = 0,\\ [\psi_{m\sigma}, \psi^\dagger_{M'\sigma'}]_+ &= \delta(M, M')\delta(\sigma, \sigma').\end{aligned} \tag{2.1}$$

It follows from these expressions that the operator $\psi^\dagger_{M\sigma}\psi_{M\sigma}$ with eigenvalues 0 and 1 measures the occupation of the sublevel σ of the multiplet $M = 0, 1$.

The Hamiltonians we consider in this paper have the general structure

$$\begin{aligned}H = &\sum_\mu \mathcal{E}_{0\mu}\psi^\dagger_{0\mu}\psi_{0\mu} + \sum_\nu \mathcal{E}_{1\nu}\psi^\dagger_{1\nu}\psi_{1\nu} + H_p\\ &+ \sum_{\mu\mu'} \psi^\dagger_{0\mu}A_{0(\mu\mu')}\psi_{0\mu'} + \sum_{\nu\nu'} \psi^\dagger_{1\nu}A_{1(\nu\nu')}\psi_{1\nu'},\end{aligned} \tag{2.2}$$

where H_p describes the lattice vibrations (phonons) of the host lattice, and where the last two terms describe the phonon–impurity interactions. In those last terms, $A_{0(\mu\mu')}$ and $A_{1(\nu\nu')}$ are phonon operators which commute with the impurity operators $\{\psi_{M\sigma}, \psi^\dagger_{M\sigma}\}$ at equal times. *By assumption,* the phonon interactions do not couple the $M = 0$ and $M = 1$ multiplets. It follows that the total occupation

$$M = \sum_\nu \psi^\dagger_{1\nu}\psi_{1\nu} \tag{2.3}$$

of the excited-state multiplet is a constant of the motion.[19]

Let O_{01} and O_{10} be any two operators which connect the M eigenvectors of the phonon–impurity system as indicated by the following matrix elements:

$$\begin{aligned}&\langle M, \cdots| O_{01} |M, \cdots\rangle\\ &\qquad = \langle M, \cdots| O_{10} |M, \cdots\rangle = 0,\\ &\langle 1, \cdots| O_{01} |0, \cdots\rangle\\ &\qquad = \langle 0, \cdots| O_{10} |1, \cdots\rangle = 0,\\ &\langle 0, \cdots| O_{01} |1, \cdots\rangle \neq 0\\ &\qquad \text{and} \quad \langle 1, \cdots| O_{10} |0, \cdots\rangle \neq 0.\end{aligned} \tag{2.4}$$

We introduce normalized thermal expectation values $\langle M\beta| \cdots |M\beta\rangle$ such that, if O is an arbitrary operator,

$$\begin{aligned}\langle M\beta| O |M\beta\rangle = &\operatorname{tr}_M [O \exp(-\beta H/\hbar)]/\operatorname{tr}_M\\ &\times [\exp(-\beta H/\hbar)],\end{aligned} \tag{2.5}$$

where tr_M designates the trace over all states of the Hamiltonian (2.2) for which M has its specified eigenvalue. Here β is determined by the temperature $T = \hbar/k\beta$ which characterizes the lattice excitation and the ocupation of the impurity states within any multiplet M.

If the phonon–impurity system is initially described by a temperature $T = \hbar/k\beta$ and the multiplet occupation parameter M, then the absorption and emission spectra appropriate to the operators $\{O_{01}, O_{10}\}$ are, respectively,

$$\begin{aligned}\bar{f}_{01}(t - t') &\equiv \langle 0\beta| O_{01}(t)O_{10}(t') |0\beta\rangle\\ &\equiv \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')}\bar{S}_{01}(\omega),\end{aligned} \tag{2.7a}$$

$$\begin{aligned}\bar{f}_{10}(t - t') &\equiv \langle 1\beta| O_{10}(t')O_{01}(t) |0\beta\rangle\\ &\equiv \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')}\bar{S}_{10}(\omega).\end{aligned} \tag{2.7b}$$

Here, and throughout this paper, Heisenberg time dependence is understood: $O(t) = \exp(iHt/\hbar)O \exp(-iHt/\hbar)$. It is an immediate consequence of

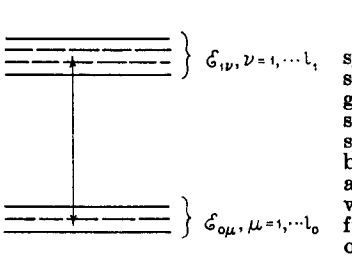

FIG. 1. Simple impurity system with two optically separated multiplets. The ground-state and excited-state multiplets are respectively characterized by the eigenvalues $M = 0$ and 1. Individual states within the multiplets are further identified by secondary indices μ and ν.

[19] Our dual use of the symbol M is deliberate. If the ground-state multiplet ($M = 0$) is occupied, then $M = 0$ in (2.3); if the excited-state multiplet ($M = 1$) is occupied, then $M = 1$ in (2.3).

the cyclic property of the trace in the expectation values (2.5) that[20,21]

$$\bar{S}_{01}(\omega) = \bar{S}_{10}(\omega) \exp[\beta(\omega - \mu)], \tag{2.8}$$

where

$$\exp(\beta\mu) = \mathrm{tr}_0[\exp(-\beta H/\hbar)]/\mathrm{tr}_1[\exp(-\beta H/\hbar)]. \tag{2.9}$$

The parameter μ is a type of "chemical potential." Equation (2.8) is in essence a detailed-balance relation. A special case of this important result connecting thermal-equilibrium absorption and emission spectra was previously indicated in Sec. 2 of I.

Note that Eq. (2.8) refers to absorption and emission spectra at the *same frequency* ω and that it makes no reference whatsoever to the fact that the absorption and emission spectra are sometimes nearly mirror images about a no-phonon line. The result (2.8) is considerably more general than the mirror-image property which depends upon rather specialized circumstances.

Equation (2.8) predicts that for broadband spectra the dominant components in emission lie at lower frequencies (Stokes' shift) than the dominant components in absorption, but it gives no quantitative information about their relative strengths. Fowler and Dexter[22] discuss reasons why the important components of the emission and absorption spectra may differ in shape and strength as well as in location.

In this paper we shall be concerned with the following special absorption and emission spectra appropriate to systems initially at a temperature $T = \hbar/k\beta$:

$$f_{01}(t - t')_{(\mu\nu;\mu'\nu')} \equiv \langle 0\beta|\,(\psi^{\dagger}_{0\mu}\psi_{1\nu})(t)(\psi^{\dagger}_{1\nu'}\psi_{0\mu'})(t')\,|0\beta\rangle; \tag{2.10a}$$

$$f_{10}(t - t')_{(\mu\nu;\mu'\nu')} \equiv \langle 1\beta|\,(\psi^{\dagger}_{1\nu'}\psi_{0\mu'})(t')(\psi^{\dagger}_{0\mu}\psi_{1\nu})(t)\,|1\beta\rangle. \tag{2.10b}$$

We shall indicate in detail in a later paper the specific connection between these functions (or their Fourier transforms) and experimentally observed spectra. Briefly, the observed spectra will be described by the functions (2.10) suitably weighted by matrix-element coefficients which reflect the structure in terms of the impurity operators $\{\psi_{M\sigma}, \psi^{\dagger}_{M\sigma}\}$ of the moment operators which govern the coupling of laboratory fields to the phonon–impurity system[23]

Since the functions (2.10) are special cases of the functions (2.7), their Fourier transforms are connected by the relation (2.8). With no loss of generality we may therefore confine our discussion to the absorption functions (2.10a). The properties of the emission functions (2.10b) may be easily inferred from Eq. (2.8).

In the following sections we distinguish several general cases. We first consider in Sec. 4 the case for which the ($M = 0$) ground state is a singlet and the ($M = 1$) excited state is a degenerate multiplet. In Sec. 5 we consider the more complicated case in which both states are degenerate multiplets. In Sec. 6 we discuss the extension of our results to nondegenerate multiplets. Section 3 is devoted to some group-theroetical preliminaries.

3. IMPURITY-SITE SYMMETRY

Let $G_{\mathfrak{b}}$ be the group (of order $g_{\mathfrak{b}} \geq 1$) of crystal symmetry operations which leave the impurity site $\mathfrak{b}$ invariant. This group is a point group and a subgroup of the full crystal symmetry group. Its operations leave the impurity, the lattice-phonon, and the interaction parts of the Hamiltonian (2.2) separately invariant.

Let $\bar{G}_{\mathfrak{b}}$ be the group (of order $\bar{g}_{\mathfrak{b}} \geq g_{\mathfrak{b}}$) of operations which leave the *impurity* part (but not necessarily the lattice-phonon or the interaction parts) of the Hamiltonian invariant. This point group is not necessarily a subgroup of the full crystal symmetry group, but it does contain $G_{\mathfrak{b}}$ as a subgroup. We assume that the two sets of operators $\{\psi_{M\sigma}\}$ and $\{\psi^{\dagger}_{M\sigma}\}$ appropriate to the multiplet M are separately basis functions for (reducible) complex-conjugate unitary representations of $\bar{G}_{\mathfrak{b}}$. Ruling out "accidental" degeneracies, we call the multiplet M degenerate if $\{\psi_{M\sigma}\}$ and $\{\psi^{\dagger}_{M\sigma}\}$ are basis functions for complex-conjugate ℓ_M-dimensional irreducible unitary representations (reps) of *both* $\bar{G}_{\mathfrak{b}}$ and $G_{\mathfrak{b}}$.[24] The multiplet is nondegenerate if the $G_{\mathfrak{b}}$ reps are reducible.

[20] P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).

[21] L. Landau, Zh. Eksperim. i Teor. Fiz. **34**, 262 (1958) [English transl.: Soviet Phys.—JETP **7**, 182 (1959)].

[22] W. B. Fowler and D. L. Dexter, Phys. Rev. **128**, 2154 (1962).

[23] The functions we compute are very closely related to the normalized spectral functions $S_{mk}{}^{\alpha,\sigma}(E)$ in Eqs. (2.1) and (2.2) of Dexter, Ref. 6. As in those equations, the coefficients relating the impurity-lattice spectral functions to observed electromagnetic fluorescence, stimulated-emission, and absorption spectra will be frequency dependent. Because of this frequency dependence the observed spectra will be slightly different from those we compute; these differences will be more significant as the relative width of the spectral region increases.

[24] Following M. A. Melvin [Rev. Mod. Phys. **28**, 18 (1956)], we shall for brevity speak of a unitary irreducible representation as a rep.

Consider first the case in which the multiplet M is degenerate and the operators $\{\psi_{M\sigma}\}$, $\{\psi^\dagger_{M\sigma}\}$ are respectively basis functions for the reps Γ_M, Γ^*_M of $G_\mathfrak{b}$. The interaction term of (2.2) appropriate to the multiplet M has the form

$$H_{IM} = \sum_{\sigma\sigma'} \psi^\dagger_{M\sigma} A_{M(\sigma\sigma')} \psi_{M\sigma'}. \tag{3.1}$$

The $A_{M(\sigma\sigma')}$ are operator functions of the phonon or lattice coordinates. We write

$$A_{M(\sigma\sigma')} = \sum_L \sum_{\lambda\in L} p(M, \sigma\sigma'; L\lambda) A^{L\lambda}, \tag{3.2}$$

where the numerical coefficients $p(M, \sigma\sigma'; L\lambda)$ are matrix elements and where the new phonon operators $A^{L\lambda}$ transform with respect to operations of the group $G_\mathfrak{b}$ as the basis functions λ of the rep Γ_L. The invariance of the interaction (3.1) with respect to the operations of $G_\mathfrak{b}$ ensures that for all $\mathbf{g}$ in $G_\mathfrak{b}$

$$H_{IM} = \sum_{\substack{\bar\sigma\bar\sigma' \\ \sigma\sigma'}} \sum_L \sum_{\lambda\bar\lambda} p(M, \sigma\sigma'; L\lambda) [\Gamma^*_M(\mathbf{g})_{\sigma\bar\sigma} \psi^\dagger_{M\bar\sigma}] \times [A^{L\bar\lambda} \Gamma_L(\mathbf{g})_{\bar\lambda\lambda}][\psi_{M\bar\sigma'} \Gamma_M(\mathbf{g})_{\bar\sigma'\sigma'}], \tag{3.3}$$

or, if the rep Γ_L does not appear in the direct product $\Gamma^*_M \otimes \Gamma_M$ more than once,

$$p(M, \sigma\sigma'; L\lambda) = \sum_{\bar\sigma\bar\sigma'} \sum_{\bar\lambda} \Gamma^*_M(\mathbf{g})_{\bar\sigma\sigma} \Gamma_L(\mathbf{g})_{\lambda\bar\lambda} \Gamma_M(\mathbf{g})_{\sigma'\bar\sigma'} \times p(M, \sigma\sigma'; L\bar\lambda) = V\begin{bmatrix} M^* & M & L \\ \sigma^* & \sigma' & \lambda \end{bmatrix} P(M \,||L||\, M). \tag{3.4}$$

The double-barred matrix elements $P(M \,||L||\, M)$ are independent of the indices $(\sigma, \sigma', \lambda)$ and the V coefficients are universal functions. The V coefficients have been tabulated for important point groups with real reps by Griffith.[24] An important special case of (3.4) obtains when Γ_L is the identity representation of $G_\mathfrak{b}$. For that case,

$$p(M, \sigma\sigma'; 1) = \delta(\sigma, \sigma') l_M^{-\frac{1}{2}} P(M \,||1||\, M). \tag{3.5}$$

If the multiplet M is nondegenerate, we may select sets of linear combinations of the $\{\psi_{M\sigma}\}$ which are basis functions for one or more reps Γ_m of $G_\mathfrak{b}$. Taking these new operators and their Hermitian adjoints as the ψ's and $\psi^\dagger$'s of our theory we rewrite (3.1) somewhat more explicitly:

$$H_{IM} = \sum_{m,m'\in M} \sum_{\sigma\in m} \sum_{\sigma'\in m'} \psi_{m\sigma} A_{(m\sigma, m'\sigma')} \psi_{m'\sigma'}. \tag{3.6}$$

Using the same arguments as before, we conclude that

$$A_{(m\sigma, m'\sigma')} = \sum_L \sum_{\lambda\in L} p_M(m\sigma, m'\sigma'; L\lambda) A^{L\lambda}, \tag{3.7}$$

where, if the rep Γ_L appears in the direct product $\Gamma^*_m \otimes \Gamma_{m'}$ no more than once,

$$p_M(m\sigma, m'\sigma'; L\lambda) = V\begin{bmatrix} m^* & m' & L \\ \sigma^* & \sigma' & \lambda \end{bmatrix} P_M(m \,||L||\, m'). \tag{3.8}$$

For the special case in which Γ_L is the identity representation,

$$p_M(m\sigma, m'\sigma'; L\lambda) = \delta(m, m')\delta(\sigma, \sigma') l_m^{-\frac{1}{2}} P_M(m \,||L||\, m'). \tag{3.9}$$

Here the function $\delta(m, m')$ only ensures that the m and m' reps be identical. It does not preclude coupling between two different sets of multiplet-M states which might transform according to the *same* rep Γ_m of $G_\mathfrak{b}$. To distinguish the various possibilities in the latter case, we would introduce supplemental state indices which would appear in the double-barred matrix elements but not in the V coefficients.

In the following sections we shall be concerned with expectation values of the type defined in Eq. (2.5). Since the Hamiltonian (2.2) is invariant with respect to the operations in $G_\mathfrak{b}$, the expectation values $\langle M\beta| \cdots |M\beta\rangle$ are likewise invariant. This implies that, if as in Eq. (3.2) we separate any operator O into components $O^{L\lambda}$ which transform as basis functions λ of the rep Γ_L, then $\langle M\beta|O^{L\lambda}|M\beta\rangle = 0$ if Γ_L is not the identity rep of $G_\mathfrak{b}$. As a special case, it follows that for a degenerate multiplet M

$$\langle \bar M\beta|\, \psi_{M\sigma}(t)\psi^\dagger_{M\sigma'}(t')\, |\bar M\beta\rangle = \delta(\sigma, \sigma')\langle \bar M\beta|\, \psi_{M\sigma''}(t)\psi^\dagger_{M\sigma''}(t')\, |\bar M\beta\rangle, \tag{3.10}$$

independent of σ''. For nondegenerate multiplets the correlation functions (3.10) could contain nondiagonal elements only for indices belonging to complex-conjugate basis functions of complex-conjugate reps of $G_\mathfrak{b}$, a result analogous to that in Eq. (3.9). As a second special case it follows from the $G_\mathfrak{b}$ invariance of the expectation value (2.5) and from the orthonormality of the V coefficients[25] that, if the ground- and excited-state multiplets are each degenerate, then

$$\langle \bar M\beta|\, (\psi^\dagger_{0\mu}\psi_{1\nu})(t)(\psi^\dagger_{1\nu'}\psi_{0\mu'})(t')\, |\bar M\beta\rangle = \delta(\mu, \mu')\delta(\nu, \nu') \times \langle \bar M\beta|\, (\psi^\dagger_{0\mu''}\psi_{1\nu''})(t)(\psi^\dagger_{1\nu''}\psi_{0\mu''})(t')\, |\bar M\beta\rangle, \tag{3.11a}$$

$$\langle \bar M\beta|\, (\psi^\dagger_{1\nu'}\psi_{0\mu'})(t')(\psi^\dagger_{0\mu}\psi_{1\nu})(t)\, |\bar M\beta\rangle = \delta(\mu, \mu')\delta(\nu, \nu') \times \langle \bar M\beta|\, (\psi^\dagger_{1\nu''}\psi_{0\mu''})(t')(\psi^\dagger_{0\mu}\psi_{1\nu})\, |\bar M\beta\rangle, \tag{3.11b}$$

independent of μ'' and ν''.

[25] J. S. Griffith, *The Irreducible Tensor Method for Molecular Symmetry Groups* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962). Cf. especially Chap. 2 and Appendix A.

Briefly, the expressions (3.10) are independent of σ'' and the expressions (3.11) independent of μ'' and ν'' because in the expectation value (2.5) we average over all initial states and in (3.10) and (3.11) we sum over all final states. This implies that in functions of the type considered in Eq. (3.11) we may require that the operator products $\psi^\dagger_{0\mu}(t)\psi_{0\mu'}(t')$ and $\psi_{1\nu}(t)\psi^\dagger_{1\nu'}(t')$ *separately* transform as basis functions for the identity representation of G_b. If either multiplet is nondegenerate, nondiagonal elements are admitted only for indices belonging to the complex-conjugate basis functions of complex-conjugate reps of G_b.

The preceding results are not invalidated by Jahn–Teller distortions.[26–28] Such distortions do *not* lift the degeneracy of G_b-degenerate multiplets. The reason for this fact, which might appear at first glance to be in contradiction to familiar properties of Jahn–Teller distortions, is that for each static Jahn–Teller distortion there are other equivalent static Jahn–Teller distortions connected by operations of G_b.[29] In the absence of symmetry-destroying external fields, all of these distorted configurations are equally likely and contribute symmetrically to initial-state averages and final-state sums. This is no longer true with symmetry-destroying external fields, but for such cases one should replace the zero-field symmetry group G_b by the new real-field symmetry group for which our general remarks still obtain.

4. SINGLET-TO-MULTIPLET TRANSITIONS

In this section we develop expressions describing the excitation of a degenerate multiplet from a singlet. By using Eq. (2.8), one can easily apply these results to the reciprocal process for which the initial state is the degenerate multiplet and the final state the singlet.

When the $M = 0$ state is a singlet, the (μ, μ') indices in the Hamiltonian (2.2) and in the correlation functions (2.10) are superfluous and may be omitted. Noting further that the *degenerate* excited-state multiplet energies $\mathcal{E}_{1\nu}$ are independent of ν, we rewrite our basic Hamiltonian (2.2) in the form

$$H = \mathcal{E}_0\psi^\dagger_0\psi_0 + \mathcal{E}_1 \sum_\nu \psi^\dagger_{1\nu}\psi_{1\nu} + H_p + \psi^\dagger_0 A_0\psi_0 + \sum_{\nu\nu'} \psi^\dagger_{1\nu}A_{1(\nu\nu')}\psi_{1\nu'}. \quad (4.1)$$

Here $\mathcal{E}_0$, $\mathcal{E}_1$ are the noninteracting impurity electronic energies, H_p is the lattice-phonon Hamiltonian, and A_0, $A_{1(\nu\nu')}$ are phonon operators for the phonon–impurity interactions. The reality of H requires that $A^\dagger_0 = A_0$ and $A^\dagger_{1(\nu\nu')} = A_{1(\nu'\nu)}$.

With the indices (μ, μ') suppressed, the correlation function (2.10a) describing the singlet-to-multiplet transitions is

$$f_{01}(t - t')_{(\nu\nu')} = \langle 0\beta|\ (\psi^\dagger_0\psi_{1\nu})(t)(\psi^\dagger_{1\nu'}\psi_0)(t')\ |0\beta\rangle. \quad (4.2)$$

The only states of the phonon–impurity system which enter this function are those for which the total electronic occupation

$$\psi^\dagger_0\psi_0 + \sum_\nu \psi^\dagger_{1\nu}\psi_{1\nu} = 1. \quad (4.3)$$

If we solve Eq. (4.3) for $\psi^\dagger_0\psi_0$, we can replace (4.1) by the effective Hamiltonian

$$H_{eff} = \hbar\omega_{10} \sum_\nu \psi^\dagger_{1\nu}\psi_{1\nu} + H_{p0} + \sum_{\nu\nu'} \psi^\dagger_{1\nu}A_{10(\nu\nu')}\psi_{1\nu'}, \quad (4.4)$$

where $\hbar\omega_{10} = \mathcal{E}_1 - \mathcal{E}_0$, $H_{p0} = H_p + A_0$, and

$$A_{10(\nu\nu')} = A_{1(\nu\nu')} - A_0\delta(\nu, \nu'). \quad (4.5)$$

For this new Hamiltonian the two-state correlation function (4.2) is equal to the simpler one-state function

$$f_{01}(t - t')_{(\nu\nu')} = \langle 0\beta|\ \psi_{1\nu}(t)\psi^\dagger_{1\nu'}(t')\ |0\beta\rangle. \quad (4.6)$$

As in Eq. (2.5), the expectation value in (4.6) involves the trace over all states for which the excited-state occupation (2.3) is $M = 0$. Equations (4.4) and (4.6) are, respectively, very similar to Eqs. (I1.6) and (I1.7) basic to the analysis in I.

An important feature of the singlet-to-single and singlet-to-multiplet cases is the fact that we can reduce the two-state spectral functions (2.10) or (4.2) to the more tractable one-state functions (4.6). In the multiplet-to-multiplet case to be treated below, no such simplification is possible.

One feature of the effective Hamiltonian (4.4) is particularly noteworthy. This is the fact that in the phonon–impurity interaction the separate coupling operators A_0 and $A_{1(\nu\nu')}$ of (4.1) have been replaced by the single difference operator (4.5). Physically this is not unexpected because the transition spectrum reflects the *relative* energies of the phonon–impurity system in the two impurity configurations, not their *absolute* energies. Absolute energies are normally reflected in one-state functions of the type (4.6). The two-state functions (4.2),

[26] H. A. Jahn and E. Teller, Proc. Roy. Soc. (London) **A161**, 220 (1937).
[27] H. A. Jahn, Proc. Roy. Soc. (London) **164A**, 117 (1938).
[28] U. Öpik and M. H. L. Pryce, Proc. Roy. Soc. (London) **A238**, 425 (1957).
[29] A. Abragam and M. H. L. Pryce, Proc. Phys. Soc. (London) **A63**, 409 (1950).

which reflect the relative energies of the impurity configurations, contain important correlations which are absent in products of one-state functions (4.6).[14,30] For the Hamiltonain (4.1) the one-state functions would of course involve the separate operators A_0 and $A_{1(\nu\nu')}$. The functions (4.6) have relevance in the present situation only because the effective Hamiltonian (4.4) replaces the Hamiltonian (4.1).

It is notationally convenient to represent the set of $(\ell_1)^2$ operators (4.5) as a ℓ_1-dimensional square matrix $\mathbf{A}_{10}$ with $(\nu\nu')$ components $A_{10(\nu\nu')}$ and to interpret the absorption correlation functions (4.6) as the $(\nu\nu')$ components of a matrix function $\mathbf{F}_{01}(t-t')$. If matrix algebra [in the $(\nu\nu')$ indices] is understood, we may then easily establish by procedures entirely analogous to those used in Sec. 5 of I that

$$\mathbf{F}_{01}(t) = \mathbf{F}(t)^0\mathbf{G}_0(t), \tag{4.7}$$

where

$$\mathbf{F}(t)^0 = \mathbf{1}e^{-i\omega_{10}t}, \tag{4.8}$$

and for $M = 0$ or 1

$$\begin{aligned}\mathbf{G}_M(t) &= \mathbf{1} - \frac{i}{\hbar}\int_0^t dt_1\,\langle M\beta|\,\mathbf{A}_{10}(t_1)\,|M\beta\rangle \\ &\quad + \left(\frac{i}{\hbar}\right)^2\int_0^t dt_1\int_0^{t_1} dt_2 \\ &\quad \times \langle M\beta|\,\mathbf{A}_{10}(t_1)\mathbf{A}_{10}(t_2)\,|M\beta\rangle - \cdots . \qquad (4.9a)\end{aligned}$$

$$\begin{aligned}&= \exp\left\{-\frac{i}{\hbar}\int_0^t dt_1\,\langle M\beta|\,\mathbf{A}_{10}(t_1)\,|M\beta\rangle + \left(\frac{i}{\hbar}\right)^2\right. \\ &\times\left[\int_0^t dt_1\int_0^{t_1} dt_2\,\langle M\beta|\,\mathbf{A}_{10}(t_1)\mathbf{A}_{10}(t_2)\,|M\beta\rangle\right. \\ &- \tfrac{1}{2}\int_0^t dt_1\,dt_2\,\langle M\beta|\,\mathbf{A}_{10}(t_1)\,|M\beta\rangle \\ &\left.\left.\times\,\langle M\beta|\,\mathbf{A}_{10}(t_2)\,|M\beta\rangle\right] - \cdots\right\}. \qquad (4.9b)\end{aligned}$$

As we have already noted in Sec. 2, the time dependence of the operators in (4.9) and in the equations which follow is that appropriate to the Heisenberg picture: $\mathbf{A}_{10}(t) = \exp(iH_{eff}t/\hbar)\mathbf{A}_{10}\exp(-iH_{eff}t/\hbar)$.

The preceding expressions are formally identical to the singlet expressions (I1.7) through (I1.9), except that various components are now matrices with respect to the impurity states whereas previously they were scalars.

The physical interpretation of our present results parallels that in I. The parallel is particularly close for degenerate multiplets since, as we have noted in Eq. (3.10), the matrix functions in Eq. (4.7) are multiples of the unit matrix. In fact, because the expectation value (2.5) is invariant with respect to the operations of the site symmetry group G_b, *each* expectation value in the expansions (4.9) is separately a multiple of the unit matrix.[31] The absorption spectrum predicted from the cumulant expansion (4.9b) displays a single sharp no-phonon line accompanied by vibrational structure. These features, which have been discussed in detail for the singlet system in Sec. 2 of I, are mentioned briefly in the singlet-to-multiplet example treated below in Sec. 7.

As we have discussed in Sec. 3 of I, the purely perturbative expansion (4.9a) is useful for the computation of spectral moments. Such moments are useful when the no-phonon line is weak and the vibrational structure has a broad smooth envelope.[2,4,5]

A notable difference between the results of this section and those for the singlet system of I is related to the matrix character of the present coupling operators $\mathbf{A}_{10}(t)$. When the phonon Hamiltonian H_p in (4.4) is harmonic—that is, bilinear in the phonon annihilation–creation operators $\{a_q, a_q^\dagger\}$—and the coupling operators $\mathbf{A}_{10(\nu\nu')}$ are linear in the $\{a_p, a_q^\dagger\}$, the Baker–Hausdorff or generalized cumulant expansion (I1.9b) for the singlet system rigorously terminates after its second term, whereas the corresponding multiplet expansion (4.9b) does not generally terminate. Factorizations of the type used in Sec. 4 of I still apply to the *phonon* parts of the $\mathbf{A}_{10}(t)$ operators; however, those factorizations will not "disentangle" the higher-order terms of the expansions (4.9) unless the *matrix* parts of the $\mathbf{A}_{10}(t)$ commute. A typical "tangled" component which appears in the factored fourth-order terms of Eqs. (4.9) is

$$\begin{aligned}&\sum_{\nu'\nu''\nu'''}\langle A_{10(\nu\nu')}(t)A_{10(\nu''\nu''')}(t')\rangle \\ &\quad\times\langle A_{10(\nu'\nu'')}(t')A_{10(\nu'''\nu''')}(t'')\rangle. \qquad (4.10a)\end{aligned}$$

A cancellation of the type discussed in Sec. 4 of I could be arranged if we could replace this component by

$$\begin{aligned}&\sum_{\nu'\nu''\nu'''}\langle A_{10(\nu\nu'')}(t)A_{10(\nu'''\nu')}(t')\rangle \\ &\quad\times\langle A_{10(\nu'\nu''')}(t')A_{10(\nu''\nu'')}(t'')\rangle. \qquad (4.10b)\end{aligned}$$

[30] R. Brout, Phys. Rev. **107**, 664 (1957).

[31] This does *not* imply that off-diagonal components of the operator matrices $\mathbf{A}_{10}(t)$ do not contribute. Rather, it is the statement that whatever off-diagonal elements do contribute in a particular expectation value must be accompanied by "compensating" elements in the same expectation value. This feature is apparent in the example of Sec. 7.

This is generally possible only if the matrix $\mathbf{A}_{10}(t)$ is diagonal.

5. MULTIPLET-TO-MULTIPLET TRANSITIONS

In this section we develop expressions describing the absorption spectrum appropriate to transitions from one degenerate multiplet to another. For this case the Hamiltonian (2.2) takes the form

$$H = \mathcal{E}_0 \sum_{\mu} \psi^{\dagger}_{0\mu}\psi_{0\mu} + \mathcal{E}_1 \sum_{\nu} \psi^{\dagger}_{1\nu}\psi_{1\nu} + H_p + \sum_{\mu\mu'} \psi^{\dagger}_{0\mu}A_{0(\mu\mu')}\psi_{0\mu'} + \sum_{\nu\nu'} \psi^{\dagger}_{1\nu}A_{1(\nu\nu')}\psi_{1\nu'}, \tag{5.1}$$

where $A^{\dagger}_{0(\mu\mu')} = A_{0(\mu'\mu)}$ and $A^{\dagger}_{1(\nu\nu')} = A_{1(\nu'\nu)}$. The correlation functions appropriate to the absorption spectrum are those in Eq. (2.10a). They involve states of the phonon–impurity system for which the total electronic occupation

$$\sum_{\mu} \psi^{\dagger}_{0\mu}\psi_{0\mu} + \sum_{\nu} \psi^{\dagger}_{1\nu}\psi_{1\nu} = 1. \tag{5.2}$$

This equation, which is the analog of the very useful Eq. (4.3), cannot be used to significantly simplify the Hamiltonian (5.1) when both ground and excited states are degenerate multiplets. We cannot here replace the two-state excitation functions (2.10a) by simpler one-state functions of the type (4.6). Multiplet-to-multiplet transitions are inherently more difficult to treat than are transitions involving singlets.

In order to treat the multiplet-to-multiplet case, a compact notation is essential. Paralleling the treatment of Sec. 4 we use $(\ell_0\ell_1)$-dimensional square matrices. The new phonon–impurity operator coupling matrix $\mathbf{A}_{10}$ for this case has the $(\mu\nu;\ \mu'\nu')$ components

$$[\mathbf{A}_{10}]_{(\mu\nu;\mu'\nu')} = \delta(\mu, \mu')A_{1(\nu\nu')} - \delta(\nu, \nu')(A^{T}_{0})_{(\mu\mu')} = \delta(\mu, \mu')A_{1(\nu\nu')} - \delta(\nu, \nu')A_{0(\mu'\mu)}. \tag{5.3}$$

It reduces to the expected form (4.5) when the $M = 0$ level is a singlet. We also define matrix functions $\mathbf{F}_{01}(t - t')$ and $\mathbf{F}_{10}(t - t')$ whose $(\mu\nu;\ \mu'\nu')$ components are the functions (2.10a) and (2.10b), respectively. Proceeding as before, we may easily establish that for this case the analog of Eq. (4.7) is

$$\mathbf{F}_{01}(t) = l_0^{-1}\mathbf{F}(t)^0\mathbf{G}_0(t). \tag{5.4}$$

Here

$$\mathbf{F}(t)^0 = \mathbf{1}e^{-i\omega_{10}t}, \qquad \omega_{10} \equiv (\mathcal{E}_1 - \mathcal{E}_0)/\hbar. \tag{5.5}$$

and

$$\mathbf{G}_M(t) = \langle M\beta|\ \mathbf{\Psi}_M\ |M\beta\rangle - \frac{i}{\hbar}\int_0^t dt_1\ \langle M\beta|\ \mathbf{A}_{10}(t_1)\mathbf{\Psi}_M\ |M\beta\rangle + \left(\frac{i}{\hbar}\right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \langle M\beta|\ \mathbf{A}_{10}(t_1) \times \mathbf{A}_{10}(t_2)\mathbf{\Psi}_M\ |M\beta\rangle - \cdots \tag{5.6a}$$

$$= \exp\left\{-\frac{i}{\hbar}\int_0^t dt_1\ \langle M\beta|\ \mathbf{A}_{10}(t_1)\mathbf{\Psi}_M\ |M\beta\rangle + \left(\frac{i}{\hbar}\right)^2\left[\int_0^t dt_1 \int_0^{t_1} dt_2\ \langle M\beta|\ \mathbf{A}_{10}(t_1)\mathbf{A}_{10}(t_2)\mathbf{\Psi}_M\ |M\beta\rangle - \tfrac{1}{2}\int_0^t dt_1\, dt_2\ \langle M\beta|\ \mathbf{A}_{10}(t_1)\mathbf{\Psi}_M\ |M\beta\rangle \times \langle M\beta|\ \mathbf{A}_{10}(t_2)\mathbf{\Psi}_M\ |M\beta\rangle\right] - \cdots\right\}. \tag{5.6b}$$

The integer ℓ_M is the degeneracy of the multiplet M. In (5.6) $\mathbf{\Psi}_M$ is the operator matrix whose $(\mu\nu;\ \mu'\nu')$ components are

$$[\mathbf{\Psi}_0]_{(\mu\nu;\mu'\nu')} = \delta(\nu, \nu')l_0\psi^{\dagger}_{0\mu}\psi_{0\mu'}, \qquad [\mathbf{\Psi}_1]_{(\mu\nu;\mu'\nu')} = \delta(\mu, \mu')l_1\psi^{\dagger}_{1\nu}\psi_{1\nu'}. \tag{5.7}$$

In going from (5.6a) to (5.6b), we used the fact implied by Eq. (3.10) that $\langle M\beta|\ \mathbf{\Psi}_M\ |M\beta\rangle = 1$. If the state M is a singlet, we everywhere set $\mathbf{\Psi}_M = \mathbf{1}$ in Eqs. (5.6) and obtain the results of Sec. 4.

Equations (5.4)–(5.6) have a formal structure very similar to that in Eqs. (4.10). However, the expressions of the present section are significantly more difficult to evaluate than are those previously derived. The expectation values in Eqs. (5.6) contain a $\mathbf{\Psi}_M$ weighting not previously present. It is *not* generally sufficient to use the approximation

$$\langle M\beta|\ \mathbf{A}_{10}(t_1) \cdots \mathbf{A}_{10}(t_n)\mathbf{\Psi}_M\ |M\beta\rangle \approx \langle M\beta|\ \mathbf{A}_{10}(t_1) \cdots \mathbf{A}_{10}(t_n)\ |M\beta\rangle\langle M\beta|\ \mathbf{\Psi}_M\ |M\beta\rangle. \tag{5.8}$$

Briefly, and somewhat too simply, the factorization (5.8) is inadaquate because it precludes from the initial-state configuration all static distortions of the impurity lattice environment which do not have the full G_b symmetry. While symmetric distortions are indeed the only admissible distortions if the initial state is a singlet, symmetry-destroying Jahn–Teller distortions can and often do occur when the initial impurity state belongs to a degenerate multiplet in the absence of phonon–impurity interactions.[26–28] Somewhat more generally, even if static Jahn–Teller distortions do not occur, the left-hand side of Eq. (5.8) will contain correlations between the dynamic lattice configuration and the electron motion not present in the factored right-hand side.[32–35]

[32] H. C. Longuet-Higgins, U. Öpik, M. H. L. Pryce, and R. A. Sack, Proc. Roy. Soc. (London) **A244**, 1 (1958).
[33] A. D. Liehr and W. Moffitt, J. Chem. Phys. **25**, 1074 (1956).
[34] W. Moffitt and A. D. Liehr, Phys. Rev. **106**, 1195 (1957).
[35] W. Moffitt and W. Thorson, Phys. Rev. **108**, 1251 (1957).

A related complication present in the calculation of the elements of Eqs. (5.6) obtains because the time dependence of the operator matrices $\mathbf{A}_{10}(t)$ is determined by the fully coupled Hamiltonian (2.2), not by the noninteracting phonon Hamiltonian. Since the phonon interactions mix the different electronic components of the initial multiplet, the time dependence of $\mathbf{A}_{10}(t)$ will indirectly involve the dynamics of the phonon-coupled initial electronic states. Actually, this complication is not entirely new to Eqs. (5.6); it is also present in Sec. 4 in a slightly simpler form. In going from the Hamiltonian (4.1) to the effective Hamiltonian (4.4), we were required to replace the phonon Hamiltonian H_p by $H_{p0} = H_p + A_0$. The operator A_0 embodies the modifications induced in the phonon spectrum as a consequence of the impurity being present in its ground state ($M = 0$). These modifications might include, for example, the generation of local phonon modes.[36,37] The system in the present section is slightly more complicated than that of Sec. 4 because the impurity system has more ground-state degrees of freedom.

Although the results of this section are mathematically more complicated than those presented in Sec. 4, or in I, their physical interpretation parallels that of the cases discussed previously. Again, group theory ensures through Eqs. (3.11) that the matrix functions in Eqs. (5.4)–(5.6) are multiples of the unit matrix. As in Eq. (4.9), each expectation value in the expansion (5.6) is separately a multiple of the unit matrix. Like the singlet-to-singlet spectra of I and the singlet-to-multiplet spectra of Sec. 4, the multiplet-to-multiplet spectra display *single* sharp no-phonon lines accompanied by vibrational structure. In this sense all these spectra are qualitatively alike. An experimental consequence is that the presence or absence of a static or dynamic Jahn–Teller effect will not be reflected in a splitting of the no-phonon line.[29,38] Jahn–Teller effects will be apparent in the no-phonon line only if the full G_b site symmetry is destroyed by static electric or magnetic fields or by selective electromagnetic pumping (which destroys the initial-state thermal distribution).[39,40] In some cases Jahn–Teller effects will also manifest themselves in qualitative differences between the vibrational structure in emission and that in absorption.[32]

[36] E. W. Montroll and R. B. Potts, Phys. Rev. **100**, 525 (1955).
[37] P. Mazur, E. W. Montroll, and R. B. Potts, J. Washington Acad. Sci. **46**, 1 (1956).
[38] The lines observed in magnetic-resonance spectra are analogous to the no-phonon lines of optical spectra. Cf. D. E. McCumber, Phys. Rev. **133**, A163 (1964).
[39] B. Bleaney, K. D. Bowers, and R. S. Trenam, Proc. Roy. Soc. (London) **A228**, 157 (1955).
[40] S. Geschwind and J. P. Remeika, J. Appl. Phys. **33**, 370 (1962).

6. NONDEGENERATE MULTIPLETS

In the preceding two sections we considered transitions involving degenerate multiplets. It is not difficult to extend our formal results to the case of nondegenerate multiplets. We here treat the general case in which both the initial and final states are such multiplets. Just as the results of Sec. 5 reduce in a straightforward manner to those of Sec. 4 when one of the multiplets is a singlet, both of these examples follow as special cases of the system considered below.

As before, our basic Hamiltonian is of the type (2.2) in which the phonons do not couple the ground-state ($M = 0$) and the excited-state ($M = 1$) multiplets. We introduce the operator matrix $\mathbf{A}_{10}(t)$ by the definition (5.3) of the preceding section. Likewise we introduce the matrix functions $\mathbf{F}_{01}(t-t')$ and $\mathbf{F}_{10}(t-t')$ whose $(\mu\nu;\mu'\nu')$ components are the functions (2.10). We define the $(\mu\nu;\mu'\nu')$ components of the diagonal matrix function $\mathbf{F}(t)^0$ to be

$$[\mathbf{F}(t)^0]_{(\mu\nu;\mu'\nu')} = \delta(\mu,\mu')\delta(\nu,\nu') \times \exp[-i(\mathcal{E}_{1\nu} - \mathcal{E}_{0\mu})t/\hbar]. \quad (6.1)$$

This definition is equivalent to (5.6) when the multiplets are degenerate. We next introduce a new operator matrix $\tilde{\mathbf{A}}_{10}(t)$ by the matrix equation

$$\tilde{\mathbf{A}}_{10}(t) = \mathbf{F}(-t)^0\mathbf{A}_{10}(t)\mathbf{F}(t)^0. \quad (6.2)$$

For degenerate multiplets, $\tilde{\mathbf{A}}_{10}(t) = \mathbf{A}_{10}(t)$.

If the integer ℓ_M is the number of states in the multiplet M and if the matrix operator $\mathbf{\Psi}_M$ is defined as in Eqs. (5.7), we define a new matrix operator $\bar{\mathbf{\Psi}}_M$ such that

$$\bar{\mathbf{\Psi}}_M = \mathbf{\Psi}_M(\langle M\beta|\ \mathbf{\Psi}_M\ |M\beta\rangle)^{-1}. \quad (6.3)$$

This clearly has the property that $\langle M\beta|\ \bar{\mathbf{\Psi}}_M\ |M\beta\rangle = 1$. We next define as the analog of the function $\mathbf{G}_M(t)$ of Eqs. (5.6) the matrix function

$$\bar{\mathbf{G}}_M(t) = \Big\{1 - \frac{i}{\hbar}\int_0^t dt_1\,\langle M\beta|\ \tilde{\mathbf{A}}_{10}(t_1)\bar{\mathbf{\Psi}}_M\ |M\beta\rangle + \Big(\frac{i}{\hbar}\Big)^2\int_0^t dt_1\int_0^{t_1} dt_2\,\langle M\beta|\ \tilde{\mathbf{A}}_{10}(t_1) \times \tilde{\mathbf{A}}_{10}(t_2)\bar{\mathbf{\Psi}}_M\ |M\beta\rangle + \cdots\Big\}\langle M\beta|\ \mathbf{\Psi}_M\ |M\beta\rangle \quad (6.4a)$$

$$= \Big(\exp\Big\{-\frac{i}{\hbar}\int_0^t dt_1\,\langle M\beta|\ \tilde{\mathbf{A}}_{10}(t_1)\bar{\mathbf{\Psi}}_M\ |M\beta\rangle$$

$$+\left(\frac{i}{\hbar}\right)^2\Big[\int_0^t dt_1\int_0^{t_1} dt_2\,\langle M\beta|\,\bar{\mathbf{A}}_{10}(t_1)\bar{\mathbf{A}}_{10}(t_2)\bar{\mathbf{\Psi}}_M\,|M\beta\rangle$$

$$-\tfrac{1}{2}\int_0^t dt_1\,dt_2\,\langle M\beta|\,\bar{\mathbf{A}}_{10}(t_1)\bar{\mathbf{\Psi}}_M\,|M\beta\rangle$$

$$\times\langle M\beta|\,\bar{\mathbf{A}}_{10}(t_2)\bar{\mathbf{\Psi}}_M\,|M\beta\rangle\Big]-\cdots\Big\}\Big)$$

$$\times\langle M\beta|\,\mathbf{\Psi}_M\,|M\beta\rangle. \tag{6.4b}$$

For degenerate multiplets, $\bar{\mathbf{\Psi}}_M = \mathbf{\Psi}_M$ and $\bar{\mathbf{G}}_M(t) = \mathbf{G}_M(t)$.

Using these definitions, we may easily establish by the techniques used previously that

$$\mathbf{F}_{01}(t) = l_0^{-1}\mathbf{F}(t)^0\bar{\mathbf{G}}_0(t). \tag{6.5}$$

This equation is the required generalization of Eq. (5.4).

Using the arguments of Sec. 3, we may verify that, if the operators $\{\psi_{M\sigma}\}$ of the multiplet M can be grouped into d_M linearly independent sets of degenerate basis functions determining reps of G_b, then each of the functions (6.5) will typically display $d = d_0d_1 \geq 1$ distinct spectra. Each spectrum will have one sharp no-phonon line plus its associated vibrational structure. If for each multiplet the G_b reps are all distinct, the functions (6.5) and each expectation value in Eqs. (6.4) will be diagonal with $d = d_0d_1$ distinct diagonal elements. If a particular G_b rep is repeated within a given multiplet, the functions (6.5) may contain off-diagonal elements connecting the corresponding basis functions in the repeated rep. These off-diagonal elements describe a phonon-induced mixing of the corresponding basis states. If in this case the functions are diagonalized by an appropriate unitary transformation, the same transformation will generally not simultaneously diagonalize the separate expectation values in the expansions (6.4).

7. A SINGLE-TO-DOUBLET EXAMPLE

As a simple application of the preceding results we consider the singlet-to-doublet transitions of a simple three-level impurity system. We consider two cases, that in which the site symmetry group G_b is C_3 (reps A and $C_\pm$) and that in which the symmetry group is C_{3v} (reps A_1, A_2, and E). Many of our results can immediately be extended to systems of higher symmetry, when the actual symmetry group contains C_3 or C_{3v} as a subgroup. The cubic case $G_b = O_h$ is a notable example.

The energy-level diagram of the impurity is indicated for the C_3 and C_{3v} cases in Fig. 2. For C_{3v} symmetry the excited state is a degenerate

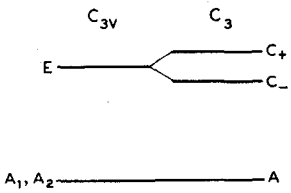

FIG. 2. Energy-level diagrams of a simple impurity system imbedded in crystals with site symmetries C_3 and C_{3v}.

doublet whose characteristic operators $\{\psi_\pm\}$ and $\{\psi_\mp^\dagger\}$ are bases for the two-dimensional rep E of C_{3v}. In a site of C_3 symmetry the doublet is split into two states whose operators are, respectively, bases for the one-dimensional reps C_+ of C_3. We assume that the singlet ground state belongs to either the A_1 or A_2 reps of C_{3v} and to the identity rep A of C_3.

Since we are concerned with singlet-to-doublet transitions, we can simplify our mathematical expressions by using an effective Hamiltonian similar to (4.4) and one-state spectral functions similar to (4.6). For the C_3 system we use the effective Hamiltonian

$$H_{\text{eff}} = \hbar\omega_+\psi_+^\dagger\psi_+ + \hbar\omega_-\psi_-^\dagger\psi_- + H_{p0} + \begin{bmatrix}\psi_+\\ \psi_-\end{bmatrix}^\dagger \mathbf{A}\begin{bmatrix}\psi_+\\ \psi_-\end{bmatrix}, \tag{7.1}$$

where $\mathbf{A}$ is the 2×2 phonon-operator matrix

$$\mathbf{A} = \begin{bmatrix} A_0 & iA_- \\ -iA_+ & \bar{A}_0\end{bmatrix}. \tag{7.2}$$

The phonon operators $A_0 = A_0^\dagger$ and $\bar{A}_0 = \bar{A}_0^\dagger$ are invariant with respect to the operations of C_3; the operators $A_\pm$ are related by Hermitian conjugation ($A_+^\dagger = A_-$) and respectively transform like basis functions of the reps $C_\pm$ of C_3. For the C_{3v} system $\omega_+ = \omega_-$; the operators $A_0 = \bar{A}_0$ are invariant with respect to the operations of C_{3v}; and the operators $A_\pm$ form a basis for the rep E of C_{3v}.

The one-state correlation functions relevant to the absorption spectrum are contained in the 2×2 matrix function

$$\mathbf{F}_{01}(t-t') = \begin{bmatrix}\langle 0\beta|\,\psi_+(t)\psi_+^\dagger(t')\,|0\beta\rangle & \langle 0\beta|\,\psi_+(t)\psi_-^\dagger(t')\,|0\beta\rangle\\ \langle 0\beta|\,\psi_+(t)\psi_+^\dagger(t')\,|0\beta\rangle & \langle 0\beta|\,\psi_-(t)\psi_-^\dagger(t')\,|0\beta\rangle\end{bmatrix}. \tag{7.3}$$

Simplifying the expressions of Sec. 6 to singlet-to-multiplet systems, we find that with the Hamiltonian (7.1)

$$\mathbf{F}_{01}(t) = \mathbf{F}(t)^0\bar{\mathbf{G}}_0(t), \tag{7.4}$$

where here

$$\mathbf{F}(t)^0 = \begin{bmatrix} \exp(-i\omega_+ t) & 0 \\ 0 & \exp(-i\omega_- t) \end{bmatrix} \tag{7.5}$$

and

$$\bar{\mathbf{G}}_0(t) = \exp\left\{-\frac{i}{\hbar}\int_0^t dt_1 \langle 0\beta|\, \bar{\mathbf{A}}(t_1)\, |0\beta\rangle + \left(\frac{i}{\hbar}\right)^2\left[\int_0^t dt_1 \int_0^{t_1} dt_2 \langle 0\beta|\, \bar{\mathbf{A}}(t_1)\bar{\mathbf{A}}(t_2)\, |0\beta\rangle - \tfrac{1}{2}\int_0^t dt_1\, dt_2 \langle 0\beta|\, \bar{\mathbf{A}}(t_1)\, |0\beta\rangle\langle 0\beta|\, \bar{\mathbf{A}}(t_2)\, |0\beta\rangle\right] - \cdots\right\}, \tag{7.6}$$

with ($\Delta \equiv \omega_+ - \omega_-$)

$$\bar{\mathbf{A}}(t) = \mathbf{F}(-t)^0 \mathbf{A}(t)\mathbf{F}(t)^0 = \begin{bmatrix} A_0(t) & iA_-(t)e^{i\Delta t} \\ -iA_+(t)e^{-i\Delta t} & \bar{A}_0(t) \end{bmatrix}. \tag{7.7}$$

For C_3, symmetry these expressions correctly reduce to expressions of the type previously derived in Sec. 4.

Since the two excited states of the impurity system belong to different reps of C_3, it follows from the symmetry arguments outlined in Secs. 3 and 6 that with C_3 symmetry the matrix function $\mathbf{F}_{01}(t)$ is diagonal and has the two independent nonzero components

$$f_{01}^{\pm}(t - t') = \langle 0\beta|\, \psi_{\pm}(t)\psi_{\pm}^{\dagger}(t')\, |0\beta\rangle. \tag{7.8}$$

From Eqs. (7.4)–(7.7) we find that

$$f_{01}^{+}(t) = e^{-i\omega_+ t}\exp\left\{-\frac{it}{\hbar}\langle 0\beta|\, A_0\, |0\beta\rangle + \left(\frac{i}{\hbar}\right)^2\int_0^t dt_1 \int_0^{t_1} dt_2\, (\langle 0\beta|\, [A_0(t_1)A_0(t_2) + A_-(t_1)e^{i\Delta(t_1-t_2)}A_+(t_2)]\, |0\beta\rangle - [\langle 0\beta|\, A_0\, |0\beta\rangle]^2) - \cdots\right\} \tag{7.9a}$$

and

$$f_{01}^{-}(t) = e^{-i\omega_- t}\exp\left\{-\frac{it}{\hbar}\langle 0\beta|\, \bar{A}_0\, |0\beta\rangle + \left(\frac{i}{\hbar}\right)^2\int_0^t dt_1 \int_0^{t_1} dt_2\, (\langle 0\beta|\, [\bar{A}_0(t_1)\bar{A}_0(t_2) + A_+(t_1)e^{-i\Delta(t_1-t_2)}A_-(t_2)]\, |0\beta\rangle - [\langle 0\beta|\, \bar{A}_0\, |0\beta\rangle]^2) - \cdots\right\}. \tag{7.9b}$$

For C_3, symmetry $\mathbf{F}_{01}(t)$ is a multiple of the unit matrix and $f_{01}^{+}(t) = f_{01}^{-}(t)$.

To facilitate the physical interpretation of these results, it is convenient to consider a specific phonon system. We choose a simple harmonic Hamiltonian

$$H_{p0} = \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}}(a_{\mathbf{q}}^{\dagger}a_{\mathbf{q}} + \tfrac{1}{2}). \tag{7.10}$$

Here $\{a_{\mathbf{q}}, a_{\mathbf{q}}^{\dagger}\}$ are phonon annihilation–creation operators having the familiar commutation relations

$$[a_{\mathbf{q}}, a_{\mathbf{q}'}] = [a_{\mathbf{q}}^{\dagger}, a_{\mathbf{q}'}^{\dagger}] = 0, \quad [a_{\mathbf{q}}, a_{\mathbf{q}'}^{\dagger}] = \delta(\mathbf{q}, \mathbf{q}'). \tag{7.11}$$

The $\mathbf{q}$ summation in (7.10) extends over a complete set of normal lattice modes. We next assume that the phonon-impurity coupling operators have simple linear forms:

$$A_0 = \sum_{\mathbf{q}} [c_{\mathbf{q}}^0 a_{\mathbf{q}}^{\dagger} + (c_{\mathbf{q}}^0)^* a_{\mathbf{q}}], \tag{7.12a}$$

$$\bar{A}_0 = \sum_{\mathbf{q}} [\bar{c}_{\mathbf{q}}^0 a_{\mathbf{q}}^{\dagger} + (\bar{c}_{\mathbf{q}}^0)^* a_{\mathbf{q}}], \tag{7.12b}$$

$$A_{\pm} = \sum_{\mathbf{q}} [c_{\mathbf{q}}^{\pm} a_{\mathbf{q}}^{\dagger} + (c_{\mathbf{q}}^{\mp})^* a_{\mathbf{q}}]. \tag{7.12c}$$

For all values of the numerical coupling coefficients $\{c_{\mathbf{q}}^0, \bar{c}_{\mathbf{p}}^0, c_{\mathbf{q}}^{\pm}\}$ these operators satisfy the required relations $A_0^{\dagger} = A_0$, $\bar{A}_0^{\dagger} = \bar{A}_0$, and $A_+^{\dagger} = A_-$.

Using properties of the Hamiltonian (7.10),[41] we find that only even-order terms of the expansion (7.9) are different from zero. Typically,

$$f_{01}^{+}(t) = e^{-i\omega_+ t}\exp\left\{\left(\frac{i}{\hbar}\right)^2\int_0^t dt_1 \int_0^{t_1} dt_2 \times \langle 0\beta|\, [A_0(t_1)A_0(t_2) + A_-(t_1)e^{i\Delta(t_1-t_2)}A_+(t_2)]\, |0\beta\rangle + \cdots\right\} \tag{7.13a}$$

$$= e^{-i\omega_+ t}\exp\left\{\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}[n(\omega)\rho_0(\omega) + n(\omega - \Delta) \times \rho_+(\omega - \Delta)](e^{i\omega t} - 1 - i\omega t)/(\hbar\omega)^2 + \cdots\right\}. \tag{7.13b}$$

Paralleling the treatment of Sec. 2 in I, we have introduced in the last equation the spectral functions

$$\rho_0(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}[A_0(t), A_0] = 2\pi\sum_{\mathbf{q}} |c_{\mathbf{q}}^0|^2\, [\delta(\omega - \omega_{\mathbf{q}}) - \delta(\omega + \omega_{\mathbf{q}})], \tag{7.14}$$

$$\rho_{\pm}(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}[A_{\pm}(t), A_{\mp}] = 2\pi\sum_{\mathbf{q}} [|c_{\mathbf{q}}^{\mp}|^2\, \delta(\omega - \omega_{\mathbf{q}}) - |c_{\mathbf{q}}^{\pm}|^2\, \delta(\omega + \omega_{\mathbf{q}})].$$

[41] Cf. Eq. (I4.3).

These functions reflect the density of phonon states weighted by the phonon-impurity coupling coefficients. They have the properties

$$\rho_0(\omega) = -\rho_0(-\omega) = \rho_0(\omega)^*, \qquad \rho_\pm(\omega) = -\rho_\mp(-\omega) = \rho_\pm(\omega)^*. \tag{7.15}$$

In Eq. (7.13) we have also used the thermal weighting function

$$n(\omega) = [\exp(\beta\omega) - 1]^{-1}. \tag{7.16}$$

For C_3, symmetry $\rho_+(\omega) = \rho_-(\omega)$, and if $\bar{\rho}_0(\omega)$ is the $\bar{A}_0$ spectral function analogous to $\rho_0(\omega)$, $\bar{\rho}_0(\omega) = \rho_0(\omega)$.

If we approximate $f_{01}^+(t)$ by the components explicitly indicated in Eq. (7.13b), it is useful and instructive to rewrite those components in the form

$$f_{01}^+(t) = e^{-\gamma_{1+}} e^{-i\bar{\omega}_+ t} e^{-\frac{1}{2}\Gamma_+|t|} \exp\left\{\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{i\omega t}}{\hbar^2(\omega + i\epsilon t)^2} \times [n(\omega)\rho_0(\omega) + n(\omega - \Delta)\rho_+(\omega - \Delta)]\right\}, \tag{7.17a}$$

where for $\epsilon = 0^+$

$$\gamma_{1+} = \int_0^\infty \frac{d\omega}{2\pi} \left\{\frac{[1 + 2n(\omega)]\rho_0(\omega)}{(\hbar\omega)^2} + \frac{1}{\hbar^2}\left(\frac{n(\omega)\rho_+(\omega)}{[\omega + \Delta + i\epsilon t]^2} + \frac{[1 + n(\omega)]\rho_-(\omega)}{(\omega - \Delta - i\epsilon t]^2}\right)\right\}, \tag{7.17b}$$

$$\bar{\omega}_+ = \omega_+ - \frac{1}{\hbar^2}\int_0^\infty \frac{d\omega}{2\pi}\left\{\frac{\rho_0(\omega)}{\omega} + \rho_-(\omega)\frac{P}{\omega - \Delta} + n(\omega)\left[\rho_-(\omega)\frac{P}{\omega - \Delta} - \rho_+(\omega)\frac{P}{\omega + \Delta}\right]\right\}, \tag{7.17c}$$

and

$$\Gamma_+ = (1/\hbar^2)[1 + n(\Delta)]\rho_-(\Delta). \tag{7.17d}$$

The corresponding expressions for $f_{01}^-(t)$ follow by replacing $\rho_0(\omega)$ by $\bar{\rho}_0(\omega)$, $\rho_\pm(\omega)$ by $\rho_\mp(\omega)$, and Δ by $(-\Delta)$. In these equations we have assumed that, as is usually the case in three-dimensional systems, the linear-coupling spectral functions (7.14) vanish at least as fast as ω^3 as $\omega \to 0$, so that $\rho(\omega)n(\omega)/\omega^2$ is regular at $\omega = 0$. Temperature ($T = \hbar/k\beta$) enters Eqs. (7.17) only through the function $n(\omega)$.

The results (7.17) are very similar to those obtained in Sec. 2 of I. By expanding the last exponential in (7.17a), one obtains from the first term in the expansion a representation of the no-phonon line while the other terms describe the vibrational structure. When $\rho_+(\omega) = \rho_-(\omega)$ and $\Delta \equiv \omega_+ - \omega_- = 0$, the difference between these results and those in I rests in the replacement of the effective density $\rho_0(\omega)$ of totally symmetric phonons by the sum $[\rho_0(\omega) + \rho_\pm(\omega)]$ of totally symmetric and $C_\pm$ (or E) phonon densities. For this case the energy shifts $\bar{\omega}_\pm - \omega_\pm$ are temperature independent and the no-phonon lines have zero width ($\Gamma_\pm = 0$). When $\rho_+(\omega) \neq \rho_-(\omega)$ or $\Delta \neq 0$, the results (7.17) differ somewhat from the results of I. For example, if $\rho_+(\omega) = \rho_-(\omega)$ but $\Delta \neq 0$, Eqs. (7.17) reflect a direct phonon coupling between the excited states.[42] This direct coupling produces a well-known temperature dependence in the position of the no-phonon lines; it also gives those lines a finite width.[16,42] A sensitive measure of the direct phonon coupling is the temperature dependence of the no-phonon-line frequency difference

$$\bar{\omega}_+ - \bar{\omega}_- = \Delta - \frac{1}{\hbar^2}\int_0^\infty \frac{d\omega}{2\pi}[1 + 2n(\omega)] \times \left[\rho_-(\omega)\frac{P}{\omega - \Delta} - \rho_+(\omega)\frac{P}{\omega + \Delta}\right]. \tag{7.18}$$

Another important characteristic of the direct phonon coupling is the line-width ratio

$$\Gamma_+/\Gamma_- = [1 + n(\Delta)]/n(\Delta) = \exp(\beta\Delta). \tag{7.19}$$

Its behavior reflects the low-temperature ($\beta \to \infty$) limits $\Gamma_\pm \to \rho_\mp(|\Delta|)/\hbar^2$, $\Gamma_\mp \to 0$, respectively appropriate to $\Delta \gtrless 0$.

In C_3 symmetry, where $\rho_+(\omega)$ may be different from $\rho_-(\omega)$, the frequency difference (7.18) is not necessarily zero when $\Delta = 0$ nor are the expressions (7.17c) and (7.18) temperature independent. These features reflect the fact that the phonons responsible for exciting the state (+) from the state (−) are generally different from those exciting (−) from (+). In C_{3v} symmetry $\rho_+(\omega) = \rho_-(\omega)$ and these distinctions no longer exist.

While it is true that, if allowance is made for the frequency difference Δ and for the difference between the spectral functions $\rho_+(\omega)$ and $\rho_-(\omega)$, the multiplet spectral functions are very similar to the singlet functions computed in I, there are fundamental differences which are not apparent in Eqs. (7.17). These differences only appear in terms of higher order than those indicated in Eqs. (7.13) and (7.17). They derive from the fact noted in earlier sections that the phonon–impurity coupling is governed in the multiplet case by nondiagonal *matrices* $\mathbf{A}(t)$. Because these matrices do not commute for different time arguments—compare Eqs. (4.10)—the expansion (7.6) does not terminate after

[42] J. H. Van Vleck, Phys. Rev. 57, 426 (1940).

two terms even with the harmonic Hamiltonian (7.10) and the linear coupling (7.12), whereas the corresponding singlet-to-singlet expansion would terminate for this special situation.

Let us investigate the lowest-order correction terms to (7.17) for the mathematically simpler case of C_{3v} symmetry for which $f_{01}(t) = f_{01}^{\pm}(t)$, $\Delta = 0$, $\rho_0(\omega) = \bar{\rho}_0(\omega)$, and $\rho_+(\omega) = \rho_-(\omega)$. Proceeding as in our derivation of Eqs. (7.13) and (7.17), we find after considerable manipulation that to terms of fourth order in the phonon–impurity interactions ($\epsilon = 0^+$)

$$f_{01}(t) = e^{-\gamma_t} e^{-i\bar{\omega}t} e^{-\frac{1}{2}\Gamma|t|} \times \exp\left\{\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{n(\omega)e^{i\omega t}}{(\hbar\omega)^2} [\rho_0(\omega) + \rho_{\pm}(\omega)] - \int_{-\infty}^{\infty} \frac{d\omega\, d\omega'}{(2\pi)^2} \frac{n(\omega)n(\omega')}{(\hbar^2\omega\,\omega')^2} \rho_{\pm}(\omega)\rho_{\pm}(\omega') \times \left[\frac{\omega' e^{i\omega t} - \omega e^{i\omega' t}}{\omega' - \omega} + \frac{\omega\,\omega'}{(\omega + \omega' + i\epsilon t)^2} e^{i(\omega+\omega')t}\right]\right\}, \tag{7.20a}$$

where

$$\gamma_t = \int_0^{\infty} \frac{d\omega}{2\pi} \frac{[1 + 2n(\omega)]}{(\hbar\omega)^2} [\rho_0(\omega) + \rho_{\pm}(\omega)] - \int_{-\infty}^{\infty} \frac{d\omega\, d\omega'}{(2\pi)^2} \frac{n(\omega)n(\omega')}{(\hbar^2\omega\,\omega')^2} \frac{\rho_{\pm}(\omega)\rho_{\pm}(\omega')}{(\omega + \omega' + i\epsilon t)^2} \times [\omega\,\omega' + (\omega + \omega')^2], \tag{7.20b}$$

$$\bar{\omega} = \omega_{\pm} - \int_0^{\infty} \frac{d\omega}{2\pi} \frac{1}{\hbar^2\omega} [\rho_0(\omega) + \rho_{\pm}(\omega)] + 2\int_0^{\infty} \frac{d\omega\, d\omega'}{(2\pi)^2} \frac{\rho_{\pm}(\omega)\rho_{\pm}(\omega')}{\hbar^4\omega} \times [1 + 2n(\omega)] \frac{P}{(\omega')^2 - \omega^2}, \tag{7.20c}$$

and

$$\Gamma = 2\int_0^{\infty} \frac{d\omega}{2\pi} \left(\frac{\rho_{\pm}(\omega)}{\hbar^2\omega}\right)^2 n(\omega)[1 + n(\omega)]. \tag{7.20d}$$

The two higher-order terms of (7.20a) not present in Eqs. (7.17) affect the observed spectrum rather differently. The last term [involving $(\omega + \omega')$] contributes a "two-phonon" component to the vibrational structure of (7.20a), while the other term [involving the $(\omega - \omega')$ denominator] modifies the "one-phonon" vibrational structure. Although both of these terms are reflected in the no-phonon strength parameter γ_t, only the "two-phonon" component is reflected in the no-phonon frequency $\bar{\omega}$ and the line-width Γ. Note in particular that even with $\rho_+(\omega) = \rho_-(\omega)$ and $\Delta = 0$ the two-phonon corrections generate a finite width Γ. Equation (7.20d) for that width has a structure we have elsewhere associated with Raman scattering by the impurity.[16]

It is not difficult to understand the physical reason why, even for the simple Hamiltonian (7.10) and the linear coupling (7.12), the spectral functions for multiplet transitions require higher-order corrections of the type appearing in Eqs. (7.20). Such corrections are *not* required only if operators of the type (7.12) couple to electronic operators which are constants of the motion. The totally symmetric phonons do not contribute corrections in Eqs. (7.20) because they couple to the constant of the motion $M = (\psi_+^{\dagger}\psi_+ + \psi_-^{\dagger}\psi_-)$; a similar statement applies to the singlet transitions discussed in I. If phonons couple to operators which are not constants of the motion—the $C_{\pm}(C_3)$ and the $E(C_{3v})$ phonons of this section are examples—then the phonon–impurity coupling depends upon the instantaneous state of the impurity as well as upon the state of the lattice. Since the present state of the impurity clearly depends upon the past phonon interactions, the phonon–impurity interaction involves correlations between several different phonons. These correlations are not relevant when the coupling proceeds through an electronic operator which is a constant of the motion. Briefly, the "two-phonon" corrections in Eqs. (7.20) reflect the simplest phonon correlations transmitted by the impurity dynamics. The higher-order "one-phonon" components in Eqs. (7.20) are corrections to the original one-phonon term induced by the two-phonon correlations[43]

8. DISCUSSION

In the preceding sections we have derived Baker–Hausdorff[10] or generalized cumulant[11] expansions describing the spectra of a rather general class of optical transitions of an impurity imbedded in a crystal lattice and in interaction with the phonons of that lattice. The results derived in I and in Secs. 4 and 5 of this paper are all special cases of the general multiplet-to-multiplet expansions discussed in Sec. 6. Those expressions relate to transitions between arbitrary impurity multiplets not connected by the phonon–impurity interaction.

[43] V. Ambegaokar, J. M. Conway, and G. Baym (to be published) consider related phenomena for the scattering of neutrons by crystals. In their case lattice anharmonicities responsible for phonon lifetimes correlate one-phonon and two-phonon processes. As a consequence the neutron scattering cross sections contain small correlation corrections which must be taken into account if the experimental data are to be accurately interpreted in terms of phonon spectra. Cf. also B. V. Thompson, Phys. Rev. **131**, 1420 (1963).

In addition to phonon–impurity systems with Hamiltonians of the form (2.2) our results apply to systems whose optical properties are accurately described by an *effective* Hamiltonian of that type. Briefly, this extended class of systems includes all those systems for which the multiplet-connecting phonon interactions involve states whose energy separations are large relative to phonon energies. They are necessarily systems for which the nonradiative transition probability between multiplets is small and the phonon relaxation time short relative to the spontaneous-emission radiative transition probability, since it is only for such systems that one can meaningfully introduce quasistationary initial ensembles of the type (2.5) which have definite total multiplet populations and a fixed lattice temperature. We shall discuss these points in detail in a later paper.

Also to be discussed later is the detailed relation of the basic spectral correlation functions (2.10) to experimentally observable spectra. This connection involves specific assumptions about the coupling of the phonon–impurity system to laboratory fields and reflects the appropriate coupling matrix elements and selection rules. Connections to the electromagnetic field, for example, would involve the matrix-element coefficients with which one would construct the various multipole operators from the impurity fields $\{\psi_{M\sigma}, \psi^{\dagger}_{M\sigma}\}$ and from the phonon fields $\{a_{\mathbf{q}}, a^{\dagger}_{\mathbf{q}}\}$. We can anticipate from the electromagnetic example that the external coupling to the phonon–impurity system will generally proceed through the impurity fields and through the phonon fields. Both couplings eventually lead to expressions involving correlation functions of the type (2.10); however, phonon assistance is reflected in the spectra in different ways. For example, phonon-assisted transitions—these are transitions whose total integrated intensity in the spectral region of interest increases with the strength of the phonon–impurity coupling—can exhibit a no-phonon line if the external coupling proceeds through the phonon field but cannot exhibit such a line if the external coupling proceeds through the impurity fields.

The simple system of Sec. 7 illustrates this last point. We assume that the frequency separation $\Delta = \omega_+ - \omega_-$ in the system of Sec. 7 is large and that the spectral region under investigation is the neighborhood of $\omega = \omega_-$. We also assume that the laboratory field couples only to the impurity excited state (+) and that matrix elements coupling that field to the state (−) vanish. The observed spectrum will be proportional to the Fourier transform of the correlation function $f^{+}_{01}(t)$ defined in Eq. (7.8). As is clear from Eqs. (7.9), (7.13), and (7.17), this function displays a sharp no-phonon line near $\omega = \omega_+$ but not near $\omega = \omega_-$, the region of experimental interest. If we expand the last exponential in Eq. (7.17a), we find that

$$f^{+}_{01}(t) = e^{-\gamma_1 t} e^{-i\bar{\omega}_+ t} e^{-\frac{1}{2}\Gamma_+ |t|} \left\{ 1 + \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \times \left(\frac{e^{i\omega t}\rho_0(\omega)}{[\hbar\omega]^2} + \frac{e^{i(\omega+\Delta)t}\rho_+(\omega)}{[\hbar(\omega+\Delta)]^2} \right) n(\omega) + \cdots \right\}. \tag{8.1}$$

The second term involves single-phonon frequencies added to or subtracted from $(\bar{\omega}_+ - \Delta) \approx \omega_-$. The spectrum described by $f^{+}_{01}(t)$ will involve components near ω_- as well as near ω_+; however, the no-phonon line will lie exclusively at $\omega = \bar{\omega}_+ \approx \omega_+$. Vibrational structure in the phonon neighborhood of $\omega = (\bar{\omega}_+ - n\Delta)$ for integral $n \geq 2$ is precluded from the higher-order terms of the expansion (8.1) by correction terms to Eqs. (7.17) of the type considered for C_{3v} symmetry at the end of Sec. 7. This is not surprising, because such frequencies do not appear in any term of the perturbative expansion derived from Eq. (6.4a).

ACKNOWLEDGMENTS

I am indebted to C. G. B. Garrett, D. A. Kleinman, M. Lax, and M. D. Sturge for constructive suggestions and stimulating conversations.

Exact Conditions for the Preservation of a Canonical Distribution in Markovian Relaxation Processes

H. C. ANDERSEN* AND I. OPPENHEIM†
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts
AND
KURT E. SHULER AND GEORGE H. WEISS‡
National Bureau of Standards, Washington, D. C.
(Received 9 November 1963)

Necessary and sufficient conditions have been determined for the exact preservation of a canonical distribution characterized by a time-dependent temperature (canonical invariance) in Markovian relaxation processes governed by a master equation. These conditions, while physically realizable, are quite restrictive so that canonical invariance is the exception rather than the rule. For processes with a continuous energy variable, canonical invariance requires that the integral master equation is exactly equivalent to a Fokker–Planck equation with linear transition moments of a special form. For processes with a discrete energy variable, canonical invariance requires, in addition to a special form of the level degeneracy, equal spacing of the energy levels and transitions between nearest-neighbor levels only. Physically, these conditions imply that canonical invariance is maintained only for weak interactions of a special type between the relaxing subsystem and the reservoir. It is also shown that canonical invariance is a sufficient condition for the exponential relaxation of the mean energy. A number of systems (hard-sphere Rayleigh gas, Brownian motion, harmonic oscillators, nuclear spins) are discussed in the framework of the above theory. Conditions for *approximate* canonical invariance valid up to a certain order in the energy are also developed and then applied to nuclear spins in a magnetic field.

I. INTRODUCTION

THERE exists a wide class of nonequilibrium systems whose relaxation can be described, at least to a very good approximation, by a *master equation* characteristic of a Markovian stochastic process. One class of such systems, with which this paper is concerned, is that of a dilute subsystem (with number density n_s) dispersed in a heat bath (with number density n_h) with $n_s \ll n_h$. It is then assumed that the subsystem, which has been prepared in an initial nonequilibrium distribution, relaxes to the (time-invariant) equilibrium distribution characteristic of the heat bath solely through interactions with the heat bath, the inequality $n_s \ll n_h$ being taken sufficiently strong that interactions between the subsystem particles can be neglected compared to the subsystem–heat-bath interactions.

In previous studies on the relaxation of such subsystem–heat-bath ensembles, it has been shown, exactly and analytically, that for *some* such ensembles an initial canonical distribution of the subsystem (i.e, a Maxwell–Boltzmann distribution in energy or velocity) will relax to the final equilibrium canonical distribution corresponding to that of the heat bath via a continuous (in time) sequence of canonical distributions with a time-dependent "temperature". We shall refer to this property as *canonical invariance*. Examples of such exact canonical invariance are: the vibrational relaxation of a set of harmonic oscillators subject to Landau–Teller transition probabilities in contact with a heat bath,[1,2] the (translational) relaxation of a hard-sphere Rayleigh gas,[3] and the (translational) relaxation of a Lorentz gas with a Maxwellian (r^{-5}) force law.[4] For many other systems, however, it has been shown that an initial canonical distribution is not preserved during the relaxation process.[5] Some examples of these are: the relaxation of a subsystem of harmonic oscillators with exponentially varying transition probabilities,[6] the relaxation of a subsystem of anharmonic oscillators,[7] the relaxa-

* Summer Student, National Bureau of Standards.
† Consultant, National Bureau of Standards.
‡ Consultant, National Bureau of Standards. Present Address: Rockefeller Institute, New York, N. Y.

[1] R. J. Rubin and K. E. Shuler, J. Chem. Phys. **25**, 59 (1956).
[2] E. W. Montroll and K. E. Shuler, J. Chem. Phys. **26**, 454 (1957).
[3] K. Andersen and K. E. Shuler, J. Chem. Phys. **40**, 633 (1964).
[4] O. I. Osipov, Bull. Moscow Univ. Ser. III **1**, 13 (1961) and Ref. 3.
[5] We are discussing here exact analytical results and are not concerned, at this time, with approximate theoretical treatments or with the approximate fitting of experimental data.
[6] R. J. Rubin and K. E. Shuler, J. Chem. Phys. **25**, 68 (1956).
[7] N. W. Bazley, E. W. Montroll, R. J. Rubin, and K. E. Shuler, J. Chem. Phys. **28**, 700 (1958).

tion of a subsystem of rigid rotators,[8] the relaxation of a hard-sphere Lorentz gas,[3] and the relaxation of a subsystem of spins in contact with a lattice.[9]

The only variables entering into the relaxation equation (the master equation) which can effect the preservation or nonpreservation of the initial distribution during relaxation are the transition probabilities per unit time between the quantum states (or the energy states), the degeneracies of the energy states, and, in discrete quantum systems, the spacing of the energy levels and their total number, i.e., finite or infinite. It is of interest to determine the necessary and sufficient conditions on these variables for canonical invariance for Markovian relaxation processes. This program is carried through in the subsequent sections of this paper.

Sections II and III are devoted primarily to the formal mathematical development. We begin our treatment in Sec. II by deriving, from the master equation in energy space, the necessary and sufficient conditions for canonical invariance for systems with continuous energy-level spectra. We then apply these results to three specific examples: Brownian Motion (Ornstein–Uhlenbeck processes), the Rayleigh gas, and classical harmonic oscillators. In Sec. III we establish the connection between the quantum-state master equation and the master equation in energy space, and then derive the necessary and sufficient conditions for systems with discrete energy levels, i.e., quantum systems, for both finite- and infinite-level systems. The quantal results are then applied to the specific examples of the relaxation of (Landau–Teller) harmonic oscillators and to spin-lattice relaxation systems. In Sec. IV we derive the conditions for the *approximate* preservation of a canonical distribution for a finite discrete energy-level system with a level spacing small compared to kT. This case is of interest in connection with certain spin-relaxation problems. In Sec. V we then consider the physical implications of our analysis, and the relation of our results to some previous work on relaxation processes.

Our findings may be summarized broadly as follows. The conditions for exact canonical invariance, while physically realizable, are quite restrictive. Canonical invariance in a relaxation process is thus the "exception" rather than the rule. Whenever canonical invariance obtains, the mean energy of the subsystem undergoes a simple exponential relaxation of the type discussed previously by Shuler, Weiss, and Andersen.[10] Canonical invariance is, however, only a sufficient and not a necessary condition for such an exponential relaxation of the mean energy. The analysis presented here permits the ready determination of canonical invariance from the form of the relevant parameters (transition-probability kernel, degeneracy, level spacing, etc.), and obviates the need for an explicit solution of the relaxation equation. Left open in the present study is the important question as to approximate canonical invariance, i.e., the extent of the deviation from canonical invariance, under a weakening of the conditions for exact invariance. We plan to consider this problem in a subsequent paper.

II. SYSTEMS WITH A CONTINUOUS ENERGY SPECTRUM

Necessary and Sufficient Conditions

We first consider subsystems, such as a classical gas, in which the individual particles can have any energy greater than zero. Let us define $P(\epsilon, t)d\epsilon$ as the probability that a subsystem particle has an energy between ϵ and $\epsilon + d\epsilon$ at time t. We assume that $P(\epsilon, t)$ satisfies the master equation

$$\frac{\partial P(\epsilon, t)}{\partial t} = \int_0^\infty [B(\epsilon \mid \epsilon')P(\epsilon', t) - B(\epsilon' \mid \epsilon)P(\epsilon, t)]\, d\epsilon'. \tag{2.1}$$

The transition probabilities per unit time, for transitions from state ϵ' to ϵ, $B(\epsilon \mid \epsilon')$, are taken to be independent of time, i.e., we consider a stationary process. We also use detailed balancing,

$$B(\epsilon \mid \epsilon')P(\epsilon', \infty) = B(\epsilon' \mid \epsilon)P(\epsilon, \infty). \tag{2.2}$$

As $t \to \infty$, the function $P(\epsilon, t)$ will approach the Maxwell–Boltzmann distribution

$$P(\epsilon, \infty) = g(\epsilon)e^{-\beta(\infty)\epsilon}/Q[\beta(\infty)], \tag{2.3}$$

where $\beta(\infty) \equiv [kT(\infty)]^{-1}$ with $T(\infty)$ being the heat-bath temperature, the partition function $Q[\beta] = \int_0^\infty g(\epsilon)e^{-\beta\epsilon}\, d\epsilon$, and $g(\epsilon)$ is the degeneracy, i.e., the density of states with energy ϵ.

We wish to investigate the conditions under which a canonical distribution is preserved during the relaxation process, i.e., the conditions under which Eq. (2.1) has a solution of the form

[8] R. Herman and K. E. Shuler, J. Chem. Phys. 29, 366 (1958).

[9] Except for the trivial case of spin ½, which involves only two energy levels. The relaxation of spin systems will be discussed in more detail in the body of the paper.

[10] K. E. Shuler, G. H. Weiss, and K. Andersen, J. Math. Phys. 3, 550 (1962).

$$P(\epsilon, t) = g(\epsilon)e^{-\beta(t)\epsilon}/Q[\beta(t)]. \tag{2.4}$$

Such a solution would enable us to define a time-dependent temperature $T(t) = 1/k\beta(t)$ for all times $0 \leq t \leq \infty$ for the relaxation process. To find the necessary conditions we assume that a solution of the form (2.4) exists. Then

$$\frac{\partial P(\epsilon, t)}{\partial t} = -\left\{\dot{\beta}(t)\epsilon + \frac{\dot{Q}[\beta(t)]}{Q[\beta(t)]}\right\}P(\epsilon, t), \tag{2.5}$$

where the dot denotes differentiation with respect to t. Since

$$\frac{\dot{Q}[\beta(t)]}{Q[\beta(t)]} = -\frac{\dot{\beta}(t)\int_0^\infty \epsilon g(\epsilon)e^{-\beta(t)\epsilon}\,d\epsilon}{\int_0^\infty g(\epsilon)e^{-\beta(t)\epsilon}\,d\epsilon} = -\dot{\beta}(t)\bar{\epsilon}(t), \tag{2.6}$$

where $\bar{\epsilon}(t)$ is the average energy of the particles at time t, Eq. (2.5) reduces to

$$\partial P(\epsilon, t)/\partial t = \dot{\beta}(t)[\bar{\epsilon}(t) - \epsilon]P(\epsilon, t). \tag{2.7}$$

The detailed balance condition (2.2) can be used to rewrite Eq. (2.1) as

$$\frac{\partial P(\epsilon, t)}{\partial t} = P(\epsilon, t)\int_0^\infty B(\epsilon' \mid \epsilon) \times [\exp\{-[\beta(t) - \beta(\infty)](\epsilon' - \epsilon)\} - 1]\,d\epsilon'. \tag{2.8}$$

With the definition

$$\alpha(t) \equiv \beta(t) - \beta(\infty), \tag{2.9}$$

the master equation (2.1) can finally be written as

$$\dot{\alpha}(t)[\bar{\epsilon}(t) - \epsilon] = \int_0^\infty B(\epsilon' \mid \epsilon) \times \{\exp[-\alpha(t)(\epsilon' - \epsilon)] - 1\}\,d\epsilon', \tag{2.10}$$

where we have made use of Eqs. (2.7) and (2.8) above. Equation (2.10) will serve as our starting point for the determination of the conditions for canonical invariance.

We assume that the right side of Eq. (2.10) can be expressed as

$$\int_0^\infty B(\epsilon' \mid \epsilon)\{\exp[-\alpha(t)(\epsilon' - \epsilon)] - 1\}\,d\epsilon' = \sum_{m=1}^{\infty}\frac{(-1)^m}{m!}b_m(\epsilon)\alpha^m, \tag{2.11}$$

where

$$b_m(\epsilon) = \int_0^\infty B(\epsilon' \mid \epsilon)(\epsilon' - \epsilon)^m\,d\epsilon'. \tag{2.12}$$

We now determine the coefficients $b_m(\epsilon)$ by deriving an equation for $\bar{\epsilon}$ as a function of time and then relating $\bar{\epsilon}(t)$ to $\alpha(t)$ for a particular class of degeneracies $g(\epsilon)$. This will yield an explicit expression in terms of $\alpha(t)$ for the left-hand side (lhs) of Eq. (2.10). The use of a Taylor expansion of the lhs of Eq. (2.10) in powers of α will then permit us to determine the $b_m(\epsilon)$.

Combining (2.10) and (2.11) we obtain

$$\dot{\alpha}(t)[\bar{\epsilon}(t) - \epsilon] = \sum_{m=1}^{\infty}\frac{(-1)^m}{m!}b_m(\epsilon)\alpha^m. \tag{2.13}$$

We first note that, for Eq. (2.13) to hold for all ϵ, the $b_m(\epsilon)$ must be of the form

$$b_m(\epsilon) = b_m^0 + \epsilon b_m^1; \qquad m = 1, 2, \cdots, \tag{2.14}$$

where b_m^0 and b_m^1 are constants independent of ϵ. This follows from the fact that the lhs of Eq. (2.13) is linear in ϵ. We can now find the equation satisfied by $\bar{\epsilon}(t)$. From Eq. (2.12) it follows that, for $m = 1$,

$$\int_0^\infty B(\epsilon' \mid \epsilon)(\epsilon' - \epsilon)\,d\epsilon' = b_1^0 + \epsilon b_1^1. \tag{2.15}$$

It has been shown in Ref. (10) that a transition moment $b_1(\epsilon)$ of the form displayed in Eq. (2.15) leads directly to the relation

$$d\bar{\epsilon}(t)/dt = b_1^0 + b_1^1\bar{\epsilon}(t) \tag{2.16}$$

for the exponential relaxation of the mean energy.

To find the relationship between $\alpha(t)$ and $\bar{\epsilon}(t)$ we must introduce an ansatz about the density of states $g(\epsilon)$. We shall use the form[11]

$$g(\epsilon) = \epsilon^n e^{p\epsilon}, \tag{2.17}$$

where $n > -1$ (but not necessarily integer) to ensure the normalizability of the canonical distribution. The density of states of a classical free particle is of this form with $p = 0$ and $n = -\frac{1}{2}$, 0, or $\frac{1}{2}$, depending upon whether the motion takes place in one, two, or three dimensions.

For the density of states given by (2.17), the mean energy $\bar{\epsilon}(t)$ is

$$\bar{\epsilon}(t) = \int_0^\infty \epsilon g(\epsilon)e^{-\beta(t)\epsilon}\,d\epsilon \Big/ \int_0^\infty g(\epsilon)e^{-\beta(t)\epsilon}\,d\epsilon = \frac{n+1}{\alpha(t) + \beta(\infty) - p}. \tag{2.18}$$

An expression for $\alpha(t)$ can be obtained by differentiating (2.18) and using Eq. (2.16),

$$\dot{\alpha}(t) = -[\alpha(t) + \beta(\infty) - p] \times \left\{\frac{b_1^0[\alpha(t) + \beta(\infty) - p]}{n+1} + b_1^1\right\}. \tag{2.19}$$

[11] The form of $g(\epsilon)$ in Eq. (2.17) can be shown to be the continuum analog of the discrete degeneracy g_l whose necessary and sufficient form for canonical invariance can be determined unequivocally (see Sec. III and Appendix III).

When $\alpha = 0$, $\beta(t) = \beta(\infty)$ and the ensemble will be in its equilibrium distribution with $\dot{\alpha}(t) \equiv \dot{\beta}(t) = 0$. From this it follows that

$$b_1^1 = -b_1^0[\beta(\infty) - p]/(n + 1), \qquad (2.20)$$

and

$$\dot{\alpha}(t) = [-b_1^0\alpha(t)/(n + 1)][\alpha(t) + \beta(\infty) - p]. \qquad (2.21)$$

Combining these results with Eq. (2.18), we finally obtain

$$\dot{\alpha}(t)\bar{\epsilon}(t) = -b_1^0\alpha(t) \qquad (2.22)$$

as the desired relation between $\dot{\alpha}$ and $\bar{\epsilon}$. We can now use Eqs. (2.20)–(2.22) to reexpress the lhs of Eq. (2.10) and equate it to the expansion (2.11). This yields

$$b_1^0\left\{\left[\frac{(\beta(\infty) - p)\epsilon}{n + 1} - 1\right]\alpha + \frac{\epsilon}{n + 1}\alpha^2\right\} = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} b_m(\epsilon)\alpha^m. \qquad (2.23)$$

Comparing coefficients of α we then find

$$\begin{aligned} b_1(\epsilon) &= b_1^0\{1 - \epsilon[\beta(\infty) - p]/(n + 1)\} = b_1^0(1 - A\epsilon), \\ b_2(\epsilon) &= 2b_1^0\epsilon/(n + 1) = b_1^0(A'\epsilon), \\ b_m(\epsilon) &= 0; \qquad m \geq 3. \end{aligned} \qquad (2.24)$$

Clearly these relations could not be realized if $B(\epsilon' \mid \epsilon)$ were a function in the ordinary sense, since $B(\epsilon' \mid \epsilon) \geq 0$ and $(\epsilon' - \epsilon)^m > 0$ for $\epsilon' \neq \epsilon$ for even m. Hence $B(\epsilon' \mid \epsilon)$ must be expressed in terms of distributions, i.e.,

$$B(\epsilon' \mid \epsilon) = b_0(\epsilon)\delta(\epsilon' - \epsilon) - b_1(\epsilon)\delta^{(1)}(\epsilon' - \epsilon) + \tfrac{1}{2}b_2(\epsilon)\delta^{(2)}(\epsilon' - \epsilon), \qquad (2.25)$$

where $\delta^{(m)}(\epsilon)$ is the mth derivative of the Dirac delta function. When the above expression is substituted into the master equation (2.1) and the resulting equation integrated by parts, one obtains

$$\frac{\partial P(\epsilon, t)}{\partial t} = -\frac{\partial}{\partial \epsilon}[b_1(\epsilon)P(\epsilon, t)] + \tfrac{1}{2}\frac{\partial^2}{\partial \epsilon^2}[b_2(\epsilon)P(\epsilon, t)], \qquad (2.26)$$

with $b_1(\epsilon)$ and $b_2(\epsilon)$ given by Eq. (2.24). We have thus obtained the interesting result that the transition probability $B(\epsilon' \mid \epsilon)$ which gives rise to canonical invariance is of the form which leads to the exact equivalence between the integral master equation (2.1) and the Fokker–Planck equation (2.26). Put in other words, it is only for relaxation processes described by the Fokker–Planck equation (2.26) with the $b_n(\epsilon)$ given by (2.24) that canonical invariance obtains for the distribution function. It can readily be verified by direct substitution that the canonical distribution (2.4) is a solution of the relaxation equation (2.26).

To summarize then, we have shown here that the necessary and sufficient conditions for canonical invariance for a Markovian relaxation process with a continuous energy variable are:

(a) *The transition probability $B(\epsilon' \mid \epsilon)$ is a sum of Dirac delta functions and their derivatives as shown in Eq. (2.25).*

(b) *The moments $b_m(\epsilon)$ are given by Eq. (2.24).*

Condition (a) implies that the master equation is exactly equivalent to the Fokker–Planck equation (2.26). The above results pertain to degeneracies $g(\epsilon)$ as given by Eq. (2.17).

Relaxation of the Temperature and Mean Energy

An explicit expression for the relaxation of the temperature $T(t) \equiv 1/k\beta(t)$ can readily be obtained from Eqs. (2.20) and (2.21). The integration of (2.21) leads to

$$\left[\frac{T(t) - T(\infty)}{T(0) - T(\infty)}\right]\left[\frac{T(0) - T_p}{T(t) - T_p}\right] = e^{b_1^1 t}, \qquad (2.27)$$

with $T_p \equiv 1/kp$. When $p = 0$, so that $g(\epsilon) = \epsilon^n$, one obtains the simple exponential temperature relaxation

$$[T(t) - T(\infty)]/[T(0) - T(\infty)] = e^{b_1^1 t}. \qquad (2.28)$$

It can readily be verified from Eq. (2.28) that $b_1^1 < 0$.

The differential equation for the relaxation of the mean energy has already been given [Eq. (2.16)]. Its solution is

$$[\bar{\epsilon}(t) - \bar{\epsilon}(\infty)]/[\bar{\epsilon}(0) - \bar{\epsilon}(\infty)] = e^{b_1^1 t}. \qquad (2.29)$$

This result is independent of the value of p in Eq. (2.17) for the density of states. Since $\bar{\epsilon}$ is proportional to kT for $g(\epsilon) = \epsilon^n$ with a canonical $P(\epsilon, t)$, the exponential temperature relaxation in (2.28) is equivalent to the exponential energy relaxation of Eq. (2.29).

We have thus demonstrated that canonical invariance is a sufficient condition for the exponential relaxation of the mean energy.[12]

[12] Canonical invariance is, however, not a necessary condition. This can readily be seen from the fact that it is necessary to specify *both* $b_1(\epsilon)$ and $b_2(\epsilon)$ for canonical invariance [Eq. (2.24)], whereas only $b_1(\epsilon)$ needs to be specified explicitly for the exponential relaxation of the mean energy [see Ref. (10)].

The Conditional Probability

The Fokker–Planck equation (2.26) can be used to obtain the conditional probability $W(\epsilon, t; \epsilon', 0)$ which corresponds to the transition probabilities $B(\epsilon' \mid \epsilon)$ given in Eq. (2.25). The conditional probability $W(\epsilon, t; \epsilon', 0)$, which is defined so that $W(\epsilon, t; \epsilon', 0)d\epsilon$ is the probability that a molecule has energy between ϵ and $\epsilon + d\epsilon$ at time t given that it has energy ϵ' at time $t = 0$, is the solution of Eq. (2.26) for the initial condition

$$P(\epsilon, 0) = \delta(\epsilon' - \epsilon). \tag{2.30}$$

To simplify the following expressions, we transform to dimensionless units of energy x and time τ, by letting

$$x = \epsilon[\beta(\infty) - p], \qquad \tau = -b_1^1 t. \tag{2.31}$$

Equation (2.26) then becomes

$$\frac{\partial P(x, \tau)}{\partial \tau} = P(x, \tau) + (x - n + 1)\frac{\partial P(x, \tau)}{\partial x} + x\frac{\partial^2 P(x, \tau)}{\partial x^2}. \tag{2.32}$$

Before evaluating $W(\epsilon, t; \epsilon', 0)$ we shall first develop the general solution of (2.32). With the transformation

$$P(x, \tau) = x^n e^{-x}\psi(x, \tau), \tag{2.33}$$

Eq. (2.32) becomes

$$\frac{\partial \psi(x, \tau)}{\partial \tau} = (n + 1 - x)\frac{\partial \psi(x, \tau)}{\partial x} + x\frac{\partial^2 \psi(x, \tau)}{\partial x^2}. \tag{2.34}$$

The rhs of Eq. (2.34) defines the linear operator

$$\mathfrak{D} \equiv (n + 1 - x)(\partial/\partial x) + x(\partial^2/\partial x^2),$$

so that (2.34) can be written compactly as

$$\partial \psi(x, \tau)/\partial \tau = \mathfrak{D}\psi(x, \tau). \tag{2.35}$$

The eigenfunctions of the operator $\mathfrak{D}$ are the generalized Laguerre polynomials L_m^n, i.e.,

$$\mathfrak{D}L_m^n(x) = -mL_m^n(x). \tag{2.36}$$

Since the L_m^n form a complete set of functions, we can write

$$\psi(x, \tau) = \sum_{m=0}^{\infty} c_m L_m^n(x)\Omega_m(\tau). \tag{2.37}$$

It can easily be shown that

$$\Omega_m(\tau) = e^{-m\tau}. \tag{2.38}$$

The general eigenfunction solution of (2.32) is thus

$$P(x, \tau) = x^n e^{-x} \sum_{m=0}^{\infty} c_m L_m^n(x) e^{-m\tau}. \tag{2.39}$$

The Laguerre polynomials have the orthogonality relation

$$\int_0^\infty x^n e^{-x} L_m^n(x) L_r^n(x)\, dx = \frac{\delta_{mr}\Gamma(m + n + 1)}{m!}, \tag{2.40}$$

so that

$$c_m = \frac{m!}{\Gamma(n + m + 1)} \int_0^\infty L_m^n(x) P(x, 0)\, dx. \tag{2.41}$$

For the initial condition $P(x, 0) = \delta(x - x_0)$, we obtain

$$c_m = m!\, L_m^n(x_0)/\Gamma(n + m + 1), \tag{2.42}$$

which yields

$$W(x, \tau; x_0, 0) = x^n e^{-x} \sum_{m=0}^{\infty} \frac{m!}{\Gamma(n + m + 1)} \times L_m^n(x_0) L_m^n(x) e^{-m\tau} = \left(\frac{x}{x_0}\right)^{\frac{1}{2}n} (1 - e^{-\tau})^{-1} e^{\frac{1}{2}n\tau} \times \exp\left[\frac{(x + x_0 e^{-\tau})}{1 - e^{-\tau}}\right] I_n\left[\frac{2(x x_0 e^{-\tau})^{\frac{1}{2}}}{1 - e^{-\tau}}\right], \tag{2.43}$$

where I_n is the nth-order modified Bessel function of the first kind.[13] This is the unique form of the conditional probability which preserves the canonical distribution. For $n = \frac{1}{2}$, Eq. (2.43) becomes[3]

$$W(x, \tau; x_0, 0) = \frac{e^{\frac{1}{2}\tau}}{2[\pi x_0(1 - e^{-\tau})]^{\frac{1}{2}}} \left\{\exp\left[-\frac{[x^{\frac{1}{2}} - (x_0 e^{-\tau})^{\frac{1}{2}}]^2}{1 - e^{-\tau}}\right] - \exp\left[-\frac{[x^{\frac{1}{2}} + (x_0 e^{-\tau})^{\frac{1}{2}}]^2}{1 - e^{-\tau}}\right]\right\}, \tag{2.44}$$

where the conditional probability can be seen to be the difference between two Gaussian terms. It is of course well known that Gaussian conditional probabilities (in velocity space) lead exactly to Fokker–Planck equations. The development presented here shows that while a Gaussian conditional probability is sufficient for the equivalence of the master equation and the Fokker–Planck equation, it is by no means a necessary form of $W(x, \tau; x_0, 0)$.

[13] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, p. 189.

Examples

We shall now apply the above results to some specific examples.

A. Harmonic Oscillator

Rubin and Shuler[1] have derived a partial differential equation to describe the collisional relaxation of an ensemble of harmonic oscillators in a heat bath for the case that $h\nu/kT(\infty) \ll 1$. (The frequency of the oscillator is ν and h is Planck's constant.) Physically, this corresponds to replacing the discrete quantum oscillator by a quasiclassical one whose transitions are still governed by the quantal transition probabilities. Their Fokker–Planck equation is

$$\frac{1}{k_{10}}\frac{\partial u(y,\,t)}{\partial t} = \theta u + (\theta y + 1)\frac{\partial u}{\theta y} + y\frac{\partial^2 u}{\partial y^2}, \qquad (2.45)$$

where $u(y,\,t)$ is defined so that for integral y it is equal to the probability that an oscillator is in the yth quantum state at time t, θ is $h\nu/kT(\infty)$, and the unit of energy is chosen such that $h\nu = 1$. We can compare their result with the Fokker–Planck equation (2.26) derived here. When the indicated differentiation is performed, Eq. (2.26) becomes

$$\frac{\partial P(\epsilon,\,t)}{\partial t} = \frac{b_1^0}{n+1}\left\{\epsilon\frac{\partial^2 P}{\partial \epsilon^2} + [\epsilon[\beta(\infty) - p] + 1 - n]\frac{\partial P}{\partial \epsilon} + [\beta(\infty) - p]P\right\}. \qquad (2.46)$$

With $n = 0$, $p = 0$ for the harmonic oscillator [$g(\epsilon) = 1$], it can readily be seen that Eqs. (2.45) and (2.46) are identical except for notation. The canonical invariance and exponential relaxation of the mean energy found by Rubin and Shuler for their quasiclassical oscillators is thus seen to be consistent with the general formulation developed here.

B. Rayleigh Gas

Andersen and Shuler[3] derived the following Fokker–Planck equation for the relaxation of a hard-sphere Rayleigh gas:

$$\frac{\partial P(x,\,\tau)}{\partial \tau} = k_R\frac{\partial}{\partial x}\left\{(x - \tfrac{3}{2})P(x,\,\tau) + \frac{\partial}{\partial x}[xP(x,\,\tau)]\right\}. \qquad (2.47)$$

Here x is the reduced dimensionless energy, $x = \epsilon/kT_2$, where T_2 is the temperature of the heat bath and ϵ is the kinetic energy of the subsystem particles. We can transform our Eq. (2.26) to one analogous to (2.47) by transforming to the reduced energy variable $x = \epsilon[\beta(\infty) - p]$. Equation (2.26) then becomes

$$\frac{\partial P(x,\,\tau)}{\partial \tau} = \frac{b_1^0[\beta(\infty) - p]}{n+1}\frac{\partial}{\partial x} \times \left\{(x - n - 1)P(x,\,\tau) + \frac{\partial}{\partial x}[xP(x,\,\tau)]\right\}. \qquad (2.48)$$

For the three-dimensional Rayleigh gas, $n = \frac{1}{2}$ and $p = 0$, and (2.48) is thus identical with (2.47) with $2b_1^0/3kT_2 = k_R$. The translation relaxation of a hard-sphere Rayleigh gas in energy space is thus shown to be, as already demonstrated explicitly in Ref. 3, another specific example of canonical invariance.

C. Ornstein–Uhlenbeck Process

Bowen and Meijer[14] have shown that for a one-dimensional Ornstein–Uhlenbeck relaxation process (i.e., Brownian motion in the absence of external forces), a velocity distribution which is initially Gaussian will remain Gaussian in form. This is true for two and three dimensions as well. The Fokker–Planck equation for the *velocity* probability density function for three-dimensional Brownian motion is[15]

$$\partial W(\mathbf{u},\,t)/\partial t = \gamma\nabla_{\mathbf{u}}(W\mathbf{u}) + [\gamma/\beta(\infty)m]\nabla_{\mathbf{u}}^2 W, \qquad (2.49)$$

where $\mathbf{u}$ is the velocity of the particle, m is its mass, and γ is the friction coefficient (Chandrasekhar's β). When this equation is converted to an equation for the *energy* probability density function, the result is of the form of Eq. (2.26) with

$$b_1(\epsilon) = [3\gamma/\beta(\infty)][1 - \tfrac{2}{3}\epsilon\beta(\infty)], \qquad b_2(\epsilon) = [3\gamma/\beta(\infty)](\tfrac{4}{3}\epsilon). \qquad (2.50)$$

This is in agreement with Eqs. (2.24), since $p = 0$ and $n = \frac{1}{2}$ for three-dimensional Brownian motion. The "Gaussian invariance" of the Ornstein–Uhlenbeck process is thus shown to be another specific example of canonical invariance.

III. SYSTEMS WITH A DISCRETE ENERGY SPECTRUM

Transition from the Quantum State to the Energy-Level Master Equation

We now consider systems which have a discrete energy spectrum such that

[14] J. J. Bowen and P. H. E. Meijer, Physica **26**, 485 (1960).
[15] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

$$0 = \epsilon_0 < \epsilon_1 < \epsilon_2 < \cdots ,$$

where every energy-level ϵ_i can be reached from every level ϵ_k in a finite number of steps, i.e., the energy-level system is irreducible in the sense of Markov chains. Each level has a degeneracy, g_i, and there may be a finite or infinite number of levels. In the finite case let there be $N + 1$ levels, so that the highest energy is ϵ_N.

We use two quantum numbers, i and σ_i, to designate a quantum state. The first denotes the energy level and the second denotes the quantum state within that level. The value of σ_i is an integer between 1 and g_i. We define $P_i^{\sigma_i}(t)$ to be the probability that a system is in state i, σ_i at time t, and assume that it satisfies a master equation:

$$\frac{dP_i^{\sigma_i}(t)}{dt} = \sum_{j=0}^{N} \sum_{\sigma_j=1}^{g_j} [B_{ij}^{\sigma_i\sigma_j} P_j^{\sigma_j} - B_{ji}^{\sigma_j\sigma_i} P_i^{\sigma_i}],$$

$$i = 0, 1, \cdots N; \quad \sigma_i = 1, 2, \cdots g_i. \tag{3.1}$$

The transition probabilities $B_{ij}^{\sigma_i\sigma_j}$, which denote the probability per unit time for a transition from state j, σ_j to state i, σ_i, are again taken to be independent of time. We define $B_{ii}^{\sigma_i\sigma_i} = 0$, for $i = 0, 1, \cdots N$, and $\sigma_i = 1, 2, \cdots g_i$. The equilibrium distribution is given by

$$P_i^{\sigma_i}(\infty) = e^{-\beta(\infty)\epsilon_i}/Q[\beta(\infty)], \tag{3.2}$$

where $Q[\beta] = \sum_{i=0}^{N} \sum_{\sigma_i=1}^{g_i} e^{-\beta\epsilon_i} = \sum_{i=0}^{N} g_i e^{-\beta\epsilon_i}$. We require that the degeneracies g_i be such that $Q[\beta]$ remains finite for the case $N = \infty$. Detailed balancing holds in the form

$$B_{ij}^{\sigma_i\sigma_j} P_j^{\sigma_j}(\infty) = B_{ji}^{\sigma_j\sigma_i} P_i^{\sigma_i}(\infty). \tag{3.3}$$

We now wish to examine the conditions under which the *quantum-state master equation* preserves the form of the canonical ensemble, i.e., the conditions under which Eq. (3.1) has a solution of the form

$$P_i^{\sigma_i}(t) = e^{-\beta(t)\epsilon_i}/Q[\beta(t)]. \tag{3.4}$$

In Appendix I it is shown that this is possible if and only if:

(a) the corresponding energy-level master equation

$$\frac{dP_i(t)}{dt} = \sum_{j=0}^{N} [B_{ij}P_j - B_{ji}P_i] \tag{3.5}$$

has a canonically invariant solution of the form

$$P_i(t) = g_i e^{-\beta(t)\epsilon_i}/Q[\beta(t)], \tag{3.6}$$

and (b)

$$\sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j} \text{ is independent of } \sigma_i. \tag{3.7}$$

In Eq. (3.5), the quantity $P_i(t)$ is defined as the probability that a system is in the energy level ϵ_i at time t, and B_{ij} is the transition probability per unit time for a transition from energy level ϵ_j to ϵ_i. These "energy level" variables are related to the corresponding "quantum state" variables by

$$P_i(t) = \sum_{\sigma_i=1}^{g_i} P_i^{\sigma_i}(t), \tag{3.8}$$

$$B_{ij} = \frac{1}{g_j} \sum_{\sigma_i=1}^{g_i} \sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j}, \tag{3.9}$$

with $B_{ii} = 0$. As is shown in Appendix I, the B_{ij} satisfy detailed balancing with

$$B_{ij}P_j(\infty) = B_{ji}P_i(\infty). \tag{3.10}$$

The condition (3.7) on the quantum-state transition probabilities $B_{ij}^{\sigma_i\sigma_j}$ has the following physical interpretation. If the population of a state is dependent only upon its energy (as is the case in a Boltzmann distribution) then the rate at which transitions are made from level j to the states in level i must be the same for all states in level i. This condition ensures that, if at one time $P_i^{\sigma_i}$ is independent of σ_i, then this independence holds for all subsequent times.

Necessary and Sufficient Conditions

We now wish to find the necessary and sufficient conditions for Eq. (3.5) to have a solution of the form of (3.6). That is, we wish to determine the conditions imposed, if any, on the degeneracies g_i, the transition probabilities per unit time B_{ij}, and the level spacing $\epsilon_i - \epsilon_j$ for $P_i(t)$ to be canonically invariant. As in the previous section [see Eq. (2.10)] we find that the existence of such a solution, together with the detailed balance condition, implies that

$$\alpha(t)[\bar{\epsilon}(t) - \epsilon_i] = \sum_{j=0}^{N} B_{ji}[e^{-\alpha(t)(\epsilon_j - \epsilon_i)} - 1], \tag{3.11}$$

where $\alpha(t) = \beta(t) - \beta(\infty)$. We define the function $B(\epsilon; i)$ of the continuous variable ϵ and the discrete variable i by

$$B(\epsilon; i) = \sum_{j=0}^{N} B_{ji}\delta(\epsilon - \epsilon_j). \tag{3.12}$$

The Laplace transform $\mathfrak{B}(\alpha; i)$ of this function is

$$\mathfrak{B}(\alpha; i) \equiv \mathfrak{L}\{B(\epsilon; i)\}$$

$$\equiv \int_0^\infty e^{-\alpha\epsilon} B(\epsilon; i)\, d\epsilon = \sum_{j=0}^{N} B_{ji} e^{-\alpha\epsilon_j}. \tag{3.13}$$

Equation (3.11) can now be written as

$$\dot{\alpha}[\bar{\epsilon} - \epsilon_i] = e^{\alpha\epsilon_i}\mathfrak{B}(\alpha; i) - \mathfrak{B}(0; i). \tag{3.14}$$

From this we find

$$\mathfrak{B}(\alpha; i) = e^{-\alpha\epsilon_i}[\mathfrak{B}(0; i) + \dot{\alpha}\bar{\epsilon} - \epsilon_i\dot{\alpha}] \tag{3.15}$$

and

$$B(\epsilon; i) = \mathfrak{L}^{-1}\{e^{-\alpha\epsilon_i}[\mathfrak{B}(0; i) + \dot{\alpha}\bar{\epsilon} - \epsilon_i\dot{\alpha}]\}$$
$$= \mathfrak{B}(0; i)\delta(\epsilon - \epsilon_i) + \mathfrak{L}^{-1}\{e^{-\alpha\epsilon_i}[\dot{\alpha}\bar{\epsilon} - \epsilon_i\dot{\alpha}]\}, \tag{3.16}$$

where $\mathfrak{L}^{-1}$ denotes the inverse Laplace-transform operator. By successively setting $i = 0$ and $i = 1$ we can rewrite the last term of (3.16) to obtain (see Appendix II)

$$B(\epsilon; i) = \left[\mathfrak{B}(0; i) - \frac{\epsilon_i}{\epsilon_1}\mathfrak{B}(0; 1) + \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\mathfrak{B}(0; 0)\right]\delta(\epsilon - \epsilon_i)$$
$$- \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\sum_{j=0}^{N} B_{j0}\delta(\epsilon - \epsilon_i - \epsilon_j) + \frac{\epsilon_i}{\epsilon_1}\sum_{j=0}^{N} B_{j1}\delta(\epsilon - \epsilon_i - \epsilon_j + \epsilon_1). \tag{3.17}$$

A comparison of Eq. (3.17) with Eq. (3.12) then yields

$$B_{ki} = \frac{\epsilon_i}{\epsilon_1}\sum_{j=0}^{N} B_{j1}\Delta(\epsilon_i + \epsilon_j - \epsilon_1, \epsilon_k) - \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\sum_{j=0}^{N} B_{j0}\Delta(\epsilon_i + \epsilon_j, \epsilon_k), \tag{3.18}$$

where $\Delta(a, b)$ is the Kronecker delta

$$\Delta(a, b) = 1, \quad a = b,$$
$$= 0, \quad a \neq b.$$

We shall now show that, if the system has more than two energy levels, Eq. (3.18) implies that $\epsilon_i = i\epsilon_1$ for $i = 0, 1, 2, \cdots$. For let us consider the expression for B_{ki} for $i \geq 2$ and $i > k$, where B_{ki} is to be greater than zero. Then the second summation can never contribute since $\epsilon_i + \epsilon_j > \epsilon_k$ for all j. The first sum contributes a term only for values of j such that

$$\epsilon_i - \epsilon_1 + \epsilon_j = \epsilon_k. \tag{3.19}$$

This can only occur (for $i \geq 2$) for $j = 0$. For $j = 1$, $B_{ki} = 0$ since $B_{11} = 0$; for $j \geq 2$, Eq. (3.19) cannot hold with $i > k$. Hence, Eq. (3.19) reduces to

$$\epsilon_i - \epsilon_1 = \epsilon_k. \tag{3.20}$$

The argument for $i < k \geq 2$ follows from detailed balance. Since we are concerned here only with sets of levels which are irreducible (in the sense of Markov chains), it follows that

$$\epsilon_i = i\epsilon_1, \tag{3.21}$$

as asserted. The energy levels are thus uniformly spaced.

Next we will show that only a small number of the B_{j0} and B_{j1} can differ from zero, and that they can be expressed in terms of B_{01}. Consider first the expression for B_{1i} for $i \geq 2$. According to Eqs. (3.18) and (3.21) and the argument of the last paragraph we must have

$$B_{12} = 2B_{01},$$
$$B_{1i} = 0, \quad i > 2. \tag{3.22}$$

From detailed balance we conclude that

$$B_{i1} = 0, \quad i > 2,$$
$$B_{21} = (2g_2/g_1)B_{01}e^{-\beta(\infty)\epsilon_1}. \tag{3.23}$$

Furthermore, it is easily verified that $B_{0i} = 0$ for $i \geq 2$ since

$$B_{0i} = i\sum_j B_{j1}\Delta(i - 1 + j, 0)$$
$$- (i - 1)\sum_j B_{j0}\Delta(i + j, 0), \tag{3.24}$$

and both Kronecker deltas are equal to zero when $i > 1$. The transition probability B_{10} can be written in terms of B_{01} as

$$B_{10} = (g_1/g_0)B_{01}e^{-\beta(\infty)\epsilon_1}. \tag{3.25}$$

Collecting now the results of Eqs. (3.18), (3.21), (3.23), and (3.25), we can finally write the general expression

$$B_{ki} = B_{01}\left\{\frac{\epsilon_i}{\epsilon_1}\Delta(i - 1, k) + e^{-\beta(\infty)\epsilon_i}\right.$$
$$\left.\times\left[2\frac{\epsilon_i}{\epsilon_1}\frac{g_2}{g_1} - \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\frac{g_1}{g_0}\right]\Delta(i + 1, k)\right\}, \tag{3.26}$$

which is valid both for $i > k$ and $i < k$. It is immediately evident that B_{ki} is nonzero only when $k = i \pm 1$. We have thus obtained the interesting result that, for canonical invariance, transitions can occur *only* between nearest-neighbor levels.

The above results are in accord with the results in Sec. II for continuous energy-level systems where

we have shown that canonical invariance obtains only when the kinetic equation is a Fokker–Planck equation as displayed in (2.26). The kinetic equation for discrete energy-level systems with equally spaced levels and with transitions only between adjacent levels is equivalent, in the limit of vanishing level spacing, to a Fokker–Planck equation.

We shall now derive the necessary and sufficient conditions on the degeneracy g_k for canonical invariance. We rewrite Eq. (3.26) as

$$B_{ki} = B_{01}[i\Delta(i-1,k) + e^{-\beta(\infty)\epsilon_1}(ai+b)\Delta(i+1,k)], \quad (3.27)$$

where we have used the result $\epsilon_i = i\epsilon_1$ of Eq. (3.21) and where we have set

$$a = \frac{2g_2}{g_1} - \frac{g_1}{g_0} \quad \text{and} \quad b = \frac{g_1}{g_0}. \quad (3.28)$$

From detailed balancing we find

$$\frac{g_{k+1}}{g_k} = \frac{B_{k+1,k}e^{-\beta(\infty)\epsilon_1}}{B_{k,k+1}} = \frac{a(k+f)}{k+1}, \quad (3.29)$$

where $f = b/a$. From Eq. (3.29) it then follows that

$$g_k = \binom{k+f-1}{k} a^k g_0. \quad (3.30)$$

Equation (3.29) is the general condition on the degeneracies for canonical invariance. It is valid for both an infinite-level system and a finite-level system. If the number of levels is infinite, we can readily show that $a \geq 1$. It is easily seen with the aid of Eq. (3.29) that a cannot be zero or negative. Also, if $0 \leq a < 1$, then Eq. (3.29) leads to $\lim_{k\to\infty} g_{k+1}/g_k = a < 1$; hence from (3.30), it follows that $\lim_{k\to\infty} g_k = 0$, which is impossible since all the g_k are positive integers. Therefore $a \geq 1$. In this case f must be positive, and Eq. (3.30) is the general form of the degeneracy for canonical invariance of an *infinite*-level system. The requirement that $Q(\beta)$ be finite when $N = \infty$ implies that

$$1 < ae^{-\epsilon_1\beta(t)}, \quad (3.31)$$

and therefore that $\beta(t) > 0$.

A further condition applies when N, the index of the highest energy level is finite. Since transitions still take place only between nearest neighbors, it is sufficient to consider an infinite set of levels and require that there be no transitions between levels N and $N+1$, i.e., that $B_{N+1,N} = 0$. According to Eq. (3.27), this requires that $aN + b = 0$ or

$$f = -N. \quad (3.32)$$

This result is also necessary. While this value of f cannot be substituted directly into Eq. (3.30) since we would then have gamma functions of negative integer arguments, we can substitute it into Eq. (3.29) and again use recurrence relations. In this way we obtain the recurrence formula

$$\frac{g_{k+1}}{g_k} = \frac{b(1-k/n)}{k+1}, \quad (3.33)$$

and the form of g_k for finite N is then found to be

$$g_k = \binom{N}{k}\left(\frac{g_1}{Ng_0}\right)^k g_0. \quad (3.34)$$

It can now readily be verified by direct substitution that the canonical distribution (3.6), with the g_i, ϵ_i, and B_{ij} as given above, is a solution of the relaxation equation (3.5), so that the above necessary conditions are also sufficient.

To summarize then, we have shown that the necessary and sufficient conditions for canonical invariance for a Markovian relaxation process with a discrete energy variable are:

(a) *The sum* $\sum_{\sigma_j} B_{ij}^{\sigma_i\sigma_j}$ *is independent of* σ_i [*Eq. (3.7)*]. This means that, if the population of a state is dependent only on its energy (as it is in the Boltzmann distribution), then the rate at which transitions are made from level j to the states in level i must be the same for all states in level i.

(b) *The energy levels are uniformly spaced with* $\epsilon_i = i\epsilon_1$ [*Eq. (3.21)*].

(c) *Transitions take place only between nearest-neighbor levels* [*Eq. (3.26)*].

(d) *The transition probability per unit time* B_{ki} *is a sum of Kronecker deltas as shown in Eq. (3.27).*

(e) *The degeneracy* g_k *is as given in Eq. (3.30) for* $N = \infty$ *and in Eq. (3.34) for* N *finite.*

Relaxation of the Temperature and Mean Energy

We now investigate the relaxation of the mean energy $\bar{\epsilon}(t)$ and the temperature $T(t) \equiv [k\beta(t)]^{-1}$ for the canonically invariant distribution function $P_i(t)$ as given by Eq. (3.6). Substitution of Eq. (3.27) into Eq. (3.11), together with the use of the definition

$$\bar{\epsilon}(t) = \sum_k g_k\epsilon_k e^{-\beta(t)\epsilon_k} \Big/ \sum_k g_k e^{-\beta(t)\epsilon_k} \quad (3.35)$$

and the relation $\epsilon_k = k\epsilon_1$ [Eq. (3.21)] with g_k as given by Eq. (3.30), leads to the differential equation

$$\dot{\beta}(t) = -(B_{01}/\epsilon_1) \times \{[e^{\beta(t)\epsilon_1} - a][e^{-\beta(\infty)\epsilon_1} - e^{-\beta(t)\epsilon_1}]\} \quad (3.36)$$

for the time-dependent temperature function $\beta(t)$. The solution of this equation is

$$\beta(t) = \frac{1}{\epsilon_1} \ln\left[\frac{e^{\beta(\infty)\epsilon_1} + aDe^{-b_{01}t}}{1 + De^{-b_{01}t}}\right], \quad (3.37)$$

where

$$D = [e^{\beta(0)\epsilon_1} - e^{\beta(\infty)\epsilon_1}]/[a - e^{\beta(0)\epsilon_1}] \quad (3.38)$$

and

$$b_{01} = B_{01}[1 - ae^{-\beta(\infty)\epsilon_1}]. \quad (3.39)$$

The inequality in Eq. (3.31) assures that $b_{01} > 0$. This is also consistent with Eq. (3.37) as $t \to \infty$. Equation (3.37) is the desired result for the relaxation of the "temperature" $\beta(t) \equiv [kT(t)]^{-1}$ for a canonically invariant distribution function. It will be noted that, in the discrete energy-level system (quantum case), the temperature does not undergo a simple exponential relaxation.

We now wish to investigate the relaxation of the mean energy $\bar{\epsilon}(t)$. It can readily be verified from Eq. (3.27) that

$$\sum_{k=0}^{\infty} B_{ki}(\epsilon_k - \epsilon_i) = -b_{01}\epsilon_i + b\epsilon_1 B_{01} e^{-\beta(\infty)\epsilon_1} = -b_{01}\epsilon_i + d, \quad (3.40)$$

with d being a constant independent of ϵ_i. Equation (3.40) is the discrete analogue of Eq. (2.15) with the quantity $\sum_k B_{ki}(\epsilon_k - \epsilon_i)$ being the discrete analog of the transition moment $b_1(\epsilon)$. As is shown in some detail in Ref. 10, Eq. (3.40) leads to the relaxation equation

$$d\bar{\epsilon}(t)/dt = -b_{01}\bar{\epsilon} + d \quad (3.41)$$

for the mean energy $\bar{\epsilon}(t)$, with the solution

$$[\bar{\epsilon}(t) - \bar{\epsilon}(\infty)]/[\bar{\epsilon}(0) - \bar{\epsilon}(\infty)] = e^{-b_{01}t}. \quad (3.42)$$

We have thus shown that canonical invariance implies the exponential relaxation of the mean energy.

It should be noted that, for a finite number of levels, nothing we have done restricts $\beta(t) \equiv [kT(t)]^{-1}$ to positive values. All the above necessary and sufficient conditions for canonical invariance and the results obtained for the relaxation of the temperature and mean energy will hold and are still valid for subsystems with an initial canonical distribution with *negative* temperatures. For an ensemble consisting of a subsystem with an initial negative temperature $T_s(0) < 0$ and a heat bath with a positive temperature $T_h > 0$, the temperature relaxation of a canonically invariant subsystem distribution would proceed from $T_s(0) < 0$ to $T_s(\infty) = T_h$ via an "infinite temperature" $T_s(t) = \infty$ at some time t.

Finally we return to mention an apparent discrepancy between the analyses in the continuous and the discrete case. We have shown, in the discrete case, that the equations themselves determine the allowable form of the degeneracies [cf. Eq. (3.29) *et seq.*]. On the other hand, we chose, apparently somewhat arbitrarily, a particular form for $g(\epsilon)$ in the continuous case which conveniently allowed us to evaluate several integrals. It can be shown however (Appendix III), by starting from the results for a discrete set of levels and passing properly to the limit where the spacing approaches zero, that the resulting form of the degeneracy $g(\epsilon)$ is the one which we used in Eq. (2.17).

Examples

There are two physically important systems which satisfy all the conditions for the exact preservation of the form of the canonical ensemble. The first is a nuclear spin of $\frac{1}{2}$ (or a group of identical, noninteracting spins of $\frac{1}{2}$) in a magnetic field interacting with a lattice (heat bath). This, however, is a somewhat trivial example since a two-level system can always be characterized by a temperature.

The other example is that of a system of harmonic oscillators in weak interaction with a heat bath. Montroll and Shuler[2] derived an equation for the collisional relaxation of an ensemble of harmonic oscillators of which a small fraction is excited to an initial vibrational nonequilibrium distribution, while a large excess of unexcited oscillators serves as a heat bath. This equation is

$$\frac{dx_n(t)}{dt} = \frac{k_{10}}{1 - e^{-\theta}}\{ne^{-\theta}x_{n-1} - [n + (n+1)e^{-\theta}]x_n + (n+1)x_{n+1}\} \quad (3.43)$$

$$n = 0, 1, \cdots,$$

where $x_n(t)$ is the fraction of excited oscillators in level n at time t, and $\theta = h\nu/kT(\infty)$. They found that an initial Boltzmann distribution relaxes to the final Boltzmann distribution at the heat-bath temperature via a continuous sequence of Boltzmann distributions. The harmonic-oscillator energy levels are nondegenerate, so that our Eq. (3.27) reduces to

$$B_{ki} = B_{01}[i\Delta(i-1, k) + e^{-\beta(\infty)\epsilon_1}(i+1)\Delta(i+1, k)]. \quad (3.44)$$

This is the form of the transition probability used by Montroll and Shuler based on the Landau–Teller prescription; its use in the master equation (3.5) yields

$$\frac{dP_i(t)}{dt} = B_{01}\{ie^{-\beta(\infty)\epsilon_1}P_{i-1}(t) - [i + (i+1)e^{-\beta(\infty)\epsilon_1}] \times P_i(t) + (i+1)P_{i+1}(t)\}, \quad i = 0, 1, \cdots, \tag{3.45}$$

which is identical with Eq. (3.43).

It should be pointed out that spin systems in general do not obey the conditions derived here for canonical invariance. While the levels will be equally spaced in a magnetic field, and while transitions may be only between adjacent levels, the degeneracies are not of the required form (3.34). The relaxation of spin systems (except for the trivial case of spin $\frac{1}{2}$) can therefore not be described by a "spin temperature" in a rigorous sense. However, as will be discussed in the next section, there can be under certain physical conditions an approximate preservation of the canonical distribution which would permit one to ascribe an approximate "spin temperature" to the relaxation of such systems.

It should be noted that the analysis of this section clearly shows that the relaxation of a system of quantal rotators cannot be canonically invariant since the energy levels are not uniformly spaced ($\epsilon_i \neq i\epsilon_1$) for such systems. This result is in agreement with the calculations of Herman and Shuler[8] on the relaxation of a system of rigid rotators.

IV. APPROXIMATE CANONICAL INVARIANCE

Up to this point we have only considered the conditions under which the form of the canonical ensemble is preserved *exactly*. These conditions are very restrictive and rule out many systems of physical interest. For example, a system of nuclear spins with $I = 1$ in a magnetic field cannot exhibit exact canonical invariance since the degeneracies do not satisfy Eq. (3.34).

It may, however, be a useful procedure under certain conditions to consider the *approximate* preservation of the canonical distribution during a Markovian relaxation process. In this section we shall study the special case in which there are a finite number of energy levels with a spacing $\Delta\epsilon$ small compared with kT, i.e., $\Delta\epsilon/kT \ll 1$. If the form of the canonical ensemble were preserved exactly, then the order of magnitude of both sides of the master equation (3.1) would be $B\beta\epsilon$, where B is a typical value of the $B_{ij}^{\sigma_i\sigma_j}$, β is the larger of $\beta(0)$, and $\beta(\infty)$, and ϵ is approximately $\epsilon_N - \epsilon_0$. We shall investigate the conditions under which canonical invariance is maintained when quantities of the order of $B\beta^2\epsilon^2$ are neglected. Only an outline of the proofs will be given since the methods are similar to those already described in the previous sections.

The quantal time-dependent canonical distribution function is

$$P_i^{\sigma_i}(t) = \frac{e^{-\beta(t)\epsilon_i}}{\sum_{j=0}^{N} g_j e^{-\beta(t)\epsilon_j}} = \frac{1 - \beta(t)\epsilon_i + O(\beta^2\epsilon^2)}{R - \beta(t)S + O(\beta^2\epsilon^2)}$$

$$= \frac{1}{R}\left[1 + \beta(t)\left(\frac{S}{R} - \epsilon_i\right)\right] + O(\beta^2\epsilon^2), \quad i = 0, 1, \cdots N, \tag{4.1}$$

where $R = \sum_{i=0}^{N} g_i$ and $S = \sum_{i=0}^{N} g_i\epsilon_i$. We now assume that these $P_i^{\sigma_i}$ will satisfy Eq. (3.1) when quantities of order $B\beta^2\epsilon^2$ are neglected. By the same methods used in Appendix I for the exact invariance it can be shown that this is the case if and only if:

(a) the equations

$$\frac{dP_i(t)}{dt} = \sum_{j=0}^{N}[B_{ij}P_j - B_{ji}P_i], \quad i = 0, 1, \cdots N \tag{4.2}$$

possess a solution of the form

$$P_i(t) = \frac{g_i}{R}\left[1 + \beta(t)\left(\frac{S}{R} - \epsilon_i\right)\right] + O(\beta^2\epsilon^2), \tag{4.3}$$

and (b)

$$\sum_{i=0}^{N}\left\{g_i\left[B_{ij} - \frac{g_i}{g_j}\sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j}\right] - g_i\left[B_{ji} - \sum_{\sigma_j=1}^{g_j} B_{ji}^{\sigma_j\sigma_i}\right]\right\} + O(B\beta^2\epsilon^2) = 0, \tag{4.4}$$

$$\sum_{i=0}^{N}\left\{g_i\left(\frac{S}{R} - \epsilon_j\right)\left[B_{ij} - \frac{g_i}{g_j}\sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j}\right] - g_i\left(\frac{S}{R} - \epsilon_i\right)\left[B_{ji} - \sum_{\sigma_j=1}^{g_j} B_{ji}^{\sigma_j\sigma_i}\right]\right\} + O(B\beta\epsilon^2) = 0, \quad \sigma_i = 1, 2, \cdots g_i. \tag{4.5}$$

The B_{ij} are defined in Eq. (3.9). As might be expected, these necessary and sufficient conditions are less stringent than those derived for the *exact* preservation of the form of the canonical ensemble. A simple example for which condition (b) is satisfied is that of nondegenerate energy levels where all the g_k are unity.

As in the exact case [see Eq. (3.11)], we find that condition (a) implies that

$$\dot{\beta}(t)\left(\frac{S}{R}-\epsilon_i\right) = -[\beta(t)-\beta(\infty)] \times \sum_{j=0}^{N} B_{ji}(\epsilon_j-\epsilon_i) + O(B\beta^2\epsilon^2), \quad (4.6)$$

which can be rearranged to give

$$\frac{\dot{\beta}(t)}{\beta(t)-\beta(\infty)} = \frac{-\sum_{j=0}^{N} B_{ji}(\epsilon_j-\epsilon_i)}{S/R-\epsilon_i} + O(B\beta\epsilon). \quad (4.7)$$

Equation (4.7) can be fulfilled only if

(i) $\dfrac{\sum_{j=0}^{N} B_{ji}(\epsilon_j-\epsilon_i)}{S/R-\epsilon_i}$ is independent of i for $\epsilon_i \neq S/R$,

when quantities of order $B\beta\epsilon$ are neglected, or

(ii) $\sum_{j=0}^{N} B_{ji}(\epsilon_j-\epsilon_i)$ is of $O(B\beta\epsilon^2)$ when $\epsilon_i = S/R$.

If we now set the lhs of Eq. (4.7) equal to a constant b (independent of i) we obtain

$$[\beta(t)-\beta(\infty)]/[\beta(0)-\beta(\infty)] = e^{-bt} \quad (4.8)$$

for the relaxation of the "temperature" $\beta(t) \equiv 1/kT(t)$ with $1/b$ equal to the "relaxation time".

To summarize then, in order for a quantum-state master equation to preserve the form of the canonical distribution to within first order in $\beta\epsilon$, it is necessary and sufficient that

(a) $\dfrac{\sum_{j=0}^{N} B_{ji}(\epsilon_j-\epsilon_i)}{\frac{S}{R}-\epsilon_i}$

be independent of i to within quantities of order B. (4.9)

(b) $$\sum_{i=0}^{N}\left\{g_i\left[B_{ii}-\frac{g_i}{g_j}\sum_{\sigma_j=1}^{g_j}B_{ij}^{\sigma_i\sigma_j}\right] - g_i\left[B_{ii}-\sum_{\sigma_j=1}^{g_j}B_{ji}^{\sigma_j\sigma_i}\right]\right\} + O(B\beta^2\epsilon^2) = 0. \quad (4.10)$$

(c) $$\sum_{i=0}^{N}\left\{g_i\left(\frac{S}{R}-\epsilon_j\right)\left[B_{ii}-\frac{g_i}{g_j}\sum_{\sigma_j=1}^{g_j}B_{ij}^{\sigma_i\sigma_j}\right] - g_i\left(\frac{S}{R}-\epsilon_i\right)\left[B_{ii}-\sum_{\sigma_j=1}^{g_j}B_{ji}^{\sigma_j\sigma_i}\right]\right\} + O(B\beta\epsilon^2) = 0. \quad (4.11)$$

There are $(N+1)^2$ quantities in the B_{ij} matrix. All the diagonal elements, B_{ii}, are taken to be zero, so there can be as many as $N(N+1)$ nonzero B_{ij}'s. The detailed balance conditions impose $N[\frac{1}{2}(N+1)]$ independent linear relationships among the B_{ij}, and the condition that b be independent of i imposes at most N additional linear relationships. The number of matrix elements B_{ij} (with $i \neq j$) minus the number of relationships is at least $N[\frac{1}{2}(N-1)]$. Therefore, it is always possible in the case considered here to construct a set of B_{ij} (which may or may not represent a valid description of a physical system) which lead to canonical invariance to within order $\beta\epsilon$, independent of the spacing of the energy levels (as long as $\Delta\epsilon/kT \ll 1$) and independent of the form of the degeneracies. This conclusion follows directly from the conditions (4.9)–(4.11) above for $N \geq 2$. When $N = 1$, there are only two energy levels in the system and the canonical distribution can always be preserved exactly.

Example

The ortho-hydrogen molecule in a magnetic field forms an example of a system which exhibits the approximate preservation of the canonical distribution of the type discussed above.[16] There are three nondegenerate nuclear states with energies 0, ϵ_1, and $2\epsilon_1$. The condition that b be independent of i leads to the following two conditions:

$$B_{10}+2B_{20} = 2B_{02}+B_{12}+O(B\beta\epsilon), \qquad -B_{01}+B_{21} = O(B\beta\epsilon). \quad (4.12)$$

The detailed balance conditions are

$$B_{01} = B_{10}+O(B\beta\epsilon), \quad B_{02} = B_{20}+O(B\beta\epsilon), \quad B_{12} = B_{21}+O(B\beta\epsilon). \quad (4.13)$$

The solution of Eqs. (4.12) and (4.13) is

$$B_{12} = B_{21}+O(B\beta\epsilon) = B_{01}+O(B\beta\epsilon) = B_{10}+O(B\beta\epsilon), \quad B_{02} = B_{20}+O(B\beta\epsilon). \quad (4.14)$$

The relaxation time b is then given by

$$b = B_{10}+2B_{20}. \quad (4.15)$$

Needler and Opechowski[16] found that the transition probabilities applicable to their model satisfied these equations.

V. DISCUSSION

The conditions for the preservation of an exact canonical distribution in Markovian relaxation processes are extremely severe. They involve restrictions

[16] G. T. Needler and W. Opechowski, Can. J. Phys. 39, 870 (1961).

upon the interaction of the system of interest with the reservoir and upon the energy-level spectrum of the system itself. In this section, we discuss briefly the physical import of some of the imposed conditions, the relationship of the process described here to Brownian motion in an arbitrary field of force, and the connection between our results and those of previous investigators.

The forms of the transition probabilities given by Eq. (2.25) in the continuum case and by Eq. (3.27) in the discrete case, imply that each interaction between the system and the reservoir must be weak. However, while weak interactions are a necessary condition for canonical invariance, they are not a sufficient condition. This is evident since the forms of the coefficients $b_1(\epsilon)$ and $b_2(\epsilon)$ must be specified by Eq. (2.24) in the continuum case, and the forms of the transition probabilities $B_{i,i+1}$ and $B_{i,i-1}$ must be specified by Eq. (3.27) in the discrete case. The deeper *physical* significance, if any, of these particular forms of $b_m(\epsilon)$ and $B_{i,j}$ are not clear to us at this time.

The forms of the $b_m(\epsilon)$ and $B_{i,j}$ for canonical invariance also yield, as is to be expected, the correct expressions for the energy fluctuation of a canonical distribution at the temperature $T(t)$. In the continuum case, for instance, it can readily be shown that the mean value of the jth power of the energy is related to the jth power of the mean energy by

$$\overline{\epsilon^j}(t) = [(j+n)!/n!\,(n+1)^j][\overline{\epsilon(t)}]^j, \tag{5.1}$$

when $g(\epsilon)$ is given by (2.17) with $p = 0$. From Eq. (5.1) one can readily obtain the following expression for the fluctuation (mean-square deviation) of the energy:

$$\overline{\epsilon^2}(t) - [\bar{\epsilon}(t)]^2 = (n+1)^{-1}[\bar{\epsilon}(t)]^2. \tag{5.2}$$

The fluctuation of the energy for a canonical ensemble is usually written as

$$\overline{\epsilon^2} - (\bar{\epsilon})^2 = kT^2C_v, \tag{5.3}$$

where C_v is the heat capacity at constant volume. For $\bar{\epsilon}(t) = (n+1)kT(t)$, as will be the case for $g(\epsilon) = \epsilon^n$ and with $C_v(t) = d\bar{\epsilon}(t)/dT(t)$, Eq. (5.2) is equivalent to (5.3). Thus, the form of the $b_m(\epsilon)$ assures that the fluctuation of the energy at all time t are those of a canonical ensemble at the temperature $T(t)$. An analogous result can be obtained for the discrete energy-level spectrum. The forms of the b's and B's are necessary and sufficient for this canonical fluctuation.

We shall now investigate whether Brownian motion in an arbitrary field of force preserves the canonical form of the probability density $P(\mathbf{R}, \mathbf{u}, t)$. The quantity $P(\mathbf{R}, \mathbf{u}, t)\, d\mathbf{R}\, d\mathbf{u}$ is the probability that the Brownian particle has a postion between $\mathbf{R}$ and $\mathbf{R} + d\mathbf{R}$ and a velocity between $\mathbf{u}$ and $\mathbf{u} + d\mathbf{u}$ at time t. The Fokker–Planck equation for Brownian motion in a field of force $\mathbf{F}(\mathbf{R})$ can be derived from the Langevin equation and is given by

$$\frac{\partial P}{\partial t} + \mathbf{u}\cdot\nabla_R P + \frac{\mathbf{F}(\mathbf{R})}{m} - \nabla_u P = \gamma\nabla_u\cdot(P\mathbf{u}) + \frac{\gamma}{m\beta(\infty)}\nabla_u^2 P, \tag{5.4}$$

where m is the mass of the Brownian particle, γ is the friction coefficient, and $\beta(\infty) = [kT(\infty)]^{-1}$ where $T(\infty)$ is the temperature of the heat bath and the final equilibrium temperature of the system. Substitution of the canonical form

$$P(\mathbf{R}, \mathbf{u}, t) = e^{-\beta(t)H(\mathbf{r},\mathbf{u})} \Big/ \int e^{-\beta(t)H(\mathbf{r},\mathbf{u})}\, d\mathbf{R}\, d\mathbf{u}, \tag{5.5}$$

where the Hamiltonian $H(\mathbf{R}, \mathbf{u})$ is given by

$$H(\mathbf{R}, \mathbf{u}) = \tfrac{1}{2}mu^2 + V(\mathbf{R}), \tag{5.6}$$

and the potential $V(\mathbf{R})$ is related to the force $\mathbf{F}(\mathbf{R})$ by

$$\mathbf{F}(\mathbf{R}) = -\nabla_R V(\mathbf{R}), \tag{5.7}$$

into (5.4) yields

$$\dot{\beta}(t)[\bar{\epsilon}(t) - H(\mathbf{R}, \mathbf{u})] = \gamma[\beta(t)mu^2 - 3][\beta(t)/\beta(\infty) - 1], \tag{5.8}$$

with a mean energy $\bar{\epsilon}(t)$ given by

$$\bar{\epsilon}(t) = \int H(\mathbf{R}, \mathbf{u})P(\mathbf{R}, \mathbf{u}, t)\, d\mathbf{R}\, d\mathbf{u}. \tag{5.9}$$

When $H(\mathbf{R}, \mathbf{u})$ is a function of $\mathbf{R}$, Eq. (5.8) can be satisfied only when $\dot{\beta}(t) = 0$ and $\beta(t) = \beta(\infty)$. Thus, a canonical distribution cannot be preserved for all times t for Brownian motion *in a field of force.*

Other investigators have studied solutions of equations similar to Eq. (3.5) whose forms remain invariant in time. Mathews, Shapiro, and Falkoff[17] have shown that the joint probability distribution $P(n_1, \cdots n_i, \cdots ; t)$ which is the probability that there are n_1 particles in state 1, n_2 particles in state 2, $\cdots$, at time t will preserve the multinomial form

$$P(n_1, \cdots n_i, \cdots ; t) = (N!/\prod_i n_i!) \prod_i [P_i(t)]^{n_i} \tag{5.10}$$

[17] P. M. Mathews, I. I. Shapiro, and D. L. Falkoff, Phys. Rev. **120**, 1 (1960).

during a Markovian relaxation process for an initial multinomial distribution $P(n_1, \cdots n_i, \cdots ; 0)$. In Eq. (5.10), $P_i(t) \equiv \bar{n}_i(t)/N$ is the probability that a particle is in state i at time t and $N = \sum n_i$ is the total number of particles in the system. The multinomial invariance of the joint probability distribution $P(n_1, \cdots n_i, \cdots ; t)$ as shown in (5.10) requires none of the restrictive assumptions which had to be made in the body of this paper to assure the canonical invariance of the distribution function $P_i(t)$. This is not too surprising since the requirement of a specific time-invariant form for the distribution function $P_i(t)$ itself is much more stringent than the more general multinomial form of the joint probability distribution. Sher and Primakoff[18] have considered spin systems in which there are interactions between the particles (or degrees of freedom) of the subsystem as well as between the subsystem and the reservoir. This more general question of possible exact or approximate canonical invariance, when there are both subsystem–subsystem interactions and subsystem–reservoir interactions, is an important extension of the present study. We plan to investigate this problem in subsequent publications.

APPENDIX I. CONDITIONS ON THE $B_{ij}^{\sigma_i\sigma_i}$

We assume that Eq. (3.1) has a solution of the form of Eq. (3.4). If we define

$$P_i(t) \equiv \sum_{\sigma_i=1}^{g_i} P_i^{\sigma_i} = g_i P_i^{\sigma_i} = g_i e^{-\beta(t)\epsilon_i}/Q[\beta(t)], \qquad i = 0, 1, \cdots N, \tag{I1}$$

we have from Eq. (3.1)

$$\frac{dP_i}{dt} = \sum_{\sigma_i=1}^{g_i} \sum_{i=0}^{N} \sum_{\sigma_j=1}^{g_j} [B_{ij}^{\sigma_i\sigma_j} P_j^{\sigma_j} - B_{ji}^{\sigma_j\sigma_i} P_i^{\sigma_i}] = \sum_{i=0}^{N} [B_{ij}P_j - B_{ji}P_i], \quad i = 0, 1, \cdots N, \tag{I2}$$

where the B_{ij} are defined by Eq. (3.9), i.e., $B_{ij} = 1/g_j \sum_{\sigma_i} \sum_{\sigma_j} B_{ij}^{\sigma_i\sigma_j}$. [We are at liberty to choose $B_{ii} = 0$, $i = 0, 1, \cdots N$, since their values do not enter into Eq. (I.2)]. We can obtain another equation for dP_i/dt from Eq. (3.1),

$$\frac{dP_i}{dt} = g_i \frac{dP_i^{\sigma_i}}{dt} = \sum_{j=0}^{N} \left[P_i \left(\sum_{\sigma_j=1}^{g_j} \frac{g_i}{g_j} B_{ij}^{\sigma_i\sigma_j} \right) - P_i \left(\sum_{\sigma_j=1}^{g_j} B_{ji}^{\sigma_j\sigma_i} \right) \right], \qquad \sigma_i = 1, 2, \cdots g_i. \tag{I3}$$

[18] A. Sher and H. Primakoff, Phys. Rev. 119, 178 (1960); 130, 1267 (1963).

Comparing Eqs. (I2) and (I3) and using Eq. (I1) we find that

$$0 = \sum_{j=0}^{N} \left\{ g_i e^{-\beta\epsilon_j} \left[B_{ij} - \frac{g_i}{g_j} \sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j} \right] - g_i e^{-\beta\epsilon_i} \left[B_{ji} - \sum_{\sigma_j=1}^{g_j} B_{ji}^{\sigma_j\sigma_i} \right] \right\}. \tag{I4}$$

Equation (I4) must hold for all β between $\beta(0)$ and $\beta(\infty)$. This is possible only if the coefficient of each $e^{-\beta\epsilon_j}$ is identically zero. This condition yields

$$B_{ij} - \frac{g_i}{g_j} \sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j} = 0. \tag{I5}$$

This in turn implies that the sum $\sum_{\sigma_j=1}^{g_j} B_{ij}^{\sigma_i\sigma_j}$ is independent of σ_i.

By reversing the argument, it is possible to show that conditions (I5) and (3.6) are sufficient, as well as necessary, to ensure that the quantum-state master equation preserves the canonical ensemble.

If we use Eqs. (3.3) and (I1) we find

$$\begin{aligned} B_{ij}P_j(\infty) &= \left(\frac{1}{g_j} \sum_{\sigma_j=1}^{g_j} \sum_{\sigma_i=1}^{g_i} B_{ij}^{\sigma_i\sigma_j} \right) [g_j P_j^{\sigma_j\prime}(\infty)] \\ &= \left[\frac{1}{g_j} \sum_{\sigma_j=1}^{g_j} \sum_{\sigma_i=1}^{g_i} B_{ij}^{\sigma_i\sigma_j} \frac{P_i^{\sigma_i}(\infty)}{P_j^{\sigma_j}(\infty)} \right] [g_j P_j^{\sigma_j\prime}(\infty)] \\ &= \left(\frac{1}{g_i} \sum_{\sigma_j=1}^{g_j} \sum_{\sigma_i=1}^{g_i} B_{ji}^{\sigma_j\sigma_i} \right) [g_i P_i^{\sigma_i\prime}(\infty)] \\ &= B_{ji}P_i(\infty). \end{aligned} \tag{I6}$$

for the detailed balance relation for the B_{ij}.

APPENDIX II. EVALUATION OF $\mathfrak{L}^{-1}\{e^{-\alpha\epsilon_i}[\dot{\alpha}\bar{\epsilon} - \epsilon_i\dot{\alpha}]\}$

We can solve for $\mathfrak{L}^{-1}\{\dot{\alpha}\bar{\epsilon}\}$ by setting $i = 0$ in Eq. (3.16). This yields

$$\mathfrak{L}^{-1}\{\dot{\alpha}\bar{\epsilon}\} = B(\epsilon; 0) - \mathfrak{B}(0; 0)\delta(\epsilon). \tag{II1}$$

Using a familiar result from the theory of Laplace transforms we note that

$$\mathfrak{L}^{-1}\{e^{-\alpha\epsilon_i}\dot{\alpha}\bar{\epsilon}\} = B(\epsilon - \epsilon_i; 0) - \mathfrak{B}(0; 0)\delta(\epsilon - \epsilon_i). \tag{II2}$$

This result can now be substituted into Eq. (3.16) to yield

$$B(\epsilon; i) = [\mathfrak{B}(0; i) - \mathfrak{B}(0; 0)]\delta(\epsilon - \epsilon_i) + B(\epsilon - \epsilon_i; 0) - \epsilon_i \mathfrak{L}^{-1}\{e^{-\alpha\epsilon_i}\dot{\alpha}\}. \tag{II3}$$

An expression for the last term in Eq. (II3) can be obtained by setting $i = 1$. This leads to the final result

$$B(\epsilon; i) = \left[\mathfrak{B}(0; i) - \frac{\epsilon_i}{\epsilon_1}\mathfrak{B}(0; 1) + \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\mathfrak{B}(0; 0)\right]\delta(\epsilon - \epsilon_i) - \left(\frac{\epsilon}{\epsilon_1} - 1\right)B(\epsilon - \epsilon_i; 0) + \frac{\epsilon_i}{\epsilon_1}B(\epsilon - \epsilon_i + \epsilon_1; 1)$$

$$= \left[\mathfrak{B}(0; i) - \frac{\epsilon_i}{\epsilon_1}\mathfrak{B}(0; 1) + \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\mathfrak{B}(0; 0)\right]\delta(\epsilon - \epsilon_i)$$

$$- \left(\frac{\epsilon_i}{\epsilon_1} - 1\right)\sum_{j=0}^{N} B_{j0}\delta(\epsilon - \epsilon_i - \epsilon_j) + \frac{\epsilon_i}{\epsilon_1}\sum_{j=0}^{N} B_{j1}\delta(\epsilon - \epsilon_i + \epsilon_1 - \epsilon_j). \quad \text{(II4)}$$

APPENDIX III. PASSAGE FROM g_i TO $g(\epsilon)$

Let us assume that ϵ_1 is so small that $\beta\epsilon_1 \ll 1$, where β is the larger of $\beta(0)$ and $\beta(\infty)$. In this case, the discrete spectrum can be approximately represented by a continuous one. Let us define a density of states, $g(\epsilon)$, which is a continuous function of ϵ. In order for $g(\epsilon)$ to represent the discrete spectrum, it must be true that

$$g(i\epsilon_1) \approx g_i/\epsilon_1 \quad \text{(III1)}$$

if g_i is a slowly varying function of i. With the aid of Eq. (3.29) we then find that

$$\frac{g[(i+1)\epsilon_1]}{g(i\epsilon_1)} = \frac{a(i+f)}{i+1}, \quad \text{(III2)}$$

or

$$\frac{g(\epsilon+\epsilon_1)}{g(\epsilon)} = \frac{a(\epsilon/\epsilon_1 + f)}{\epsilon/\epsilon_1 + 1}, \quad \text{(III3)}$$

with $f=b/a$ [see Eq. (3.29)]. Since $g(\epsilon+\epsilon_1) \approx g(\epsilon)+\epsilon_1[dg(\epsilon)/d\epsilon] + O(\epsilon_1^2)$, one can write, using (III3),

$$\frac{dg(\epsilon)}{d\epsilon} = g(\epsilon)\left[\frac{(\epsilon/\epsilon_1)(a-1) + af - 1}{\epsilon + \epsilon_1}\right]. \quad \text{(III4)}$$

For $\epsilon \gg \epsilon_1$, the solution of this equation is

$$g(\epsilon) = \epsilon^{b-1}e^{[(a-1)/\epsilon_1]\epsilon}. \quad \text{(III5)}$$

If we now let $p = (a-1)/\epsilon_1$ and $n = b-1$, we obtain Eq. (2.17).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Elastic, Electromagnetic, and Other Waves in a Random Medium*

FRANK C. KARAL, JR. AND JOSEPH B. KELLER
Courant Institute of Mathematical Sciences, New York University, New York, N. Y.

Propagation of any type of wave in a random medium is analyzed on the assumption that the medium differs slightly from a homogeneous medium. An equation satisfied by the average wave is deduced which is correct through terms of order ϵ^2, where ϵ measures the deviation of the medium from homogeneity. From this equation, the propagation constant of the medium is determined. The general formulation applies to any type of linear differential or integral equation with random coefficients. It is applied to time-harmonic waves satisfying the reduced wave equation, to the equations of elasticity and to Maxwell's equations. The propagation constant for the average or coherent wave is complex even for a nondissipative medium, because the coherent wave is continually scattered by the inhomogeneities and converted into the incoherent wave. The propagation velocity of the average wave is also diminished by the inhomogeneities. This propagation constant depends upon certain trigonometric integrals of the auto- and cross-correlation functions of the coefficients in the original equations, i.e., of the various coefficients characterizing the medium. To illustrate the results, media with particular random variations are considered and the propagation constants are determined for them.

1. INTRODUCTION

We wish to examine wave propagation in a medium which differs slightly from a homogeneous medium. We represent this medium by a random medium, i.e., a family of media each member of which differs slightly from the homogeneous medium. By employing perturbation theory we determine the wave in each member of the family up to second order in ϵ, where ϵ measures the deviation of the medium from homogeneity. Then we compute the average or mean wave up to second order in ϵ, and find an equation which it satisfies. From this equation we determine the effective propagation constant for the mean wave, when it is a plane wave. This effective propagation constant is the main goal of our analysis. It reveals an attenuation of the wave as well as an alteration in phase velocity. In addition we obtain some results about the amplitude of the mean plane wave.

We first present our analysis in general terms and then apply it to special kinds of waves. These include scalar, elastic, and electromagnetic waves. Where possible we compare our results with previous ones. The general considerations have been discussed before[1] and applied to scalar waves,[1,2] but even in that case we present some new results. However our main applications are to elastic waves in a medium in which the Lamé constants and the density are all random and to electromagnetic waves in media in which the dielectric constant, the permeability, and the conductivity are all random.

Let us first point out that a random medium is a mathematical model of a complex medium. It consists of a family of media together with a probability distribution over the members of the family. The random medium represents the particular complex medium well if each member of the family is similar to it, or if the mean of the random medium is close to the given medium and if the variance is small.

In order to represent wave propagation in the given complex medium, we consider propagation in the random medium. This means that we consider a random wave. This is just a family of waves, one in each medium of the family, together with the probability of the corresponding medium. The mean of the random wave, and its other statistics, are assumed to provide information about the wave in the given complex medium.

2. GENERAL CONSIDERATIONS

Let u_0 denote a wave in a homogeneous medium characterized by a linear operator L. Then u_0 is a solution of the equation

$$Lu_0 = 0. \tag{1}$$

Let α designate a different medium characterized by the operator $L - \epsilon L_1(\alpha) - \epsilon^2 L_2(\alpha) + O(\epsilon^3)$.

* The research reported in this paper was supported by the American Petroleum Institute (Project 61) and the American Chemical Society (Petroleum Research Fund Grant in Aid 436-A).

[1] J. B. Keller, "Stochastic equations and wave propagation in random media," in *Proceedings of the 17th Symposium on Applied Mathematics* (American Mathematical Society, New York, 1964).

[2] J. B. Keller, "Wave propagation in random media," in *Proceedings of the 13th Symposium on Applied Mathematics* (American Mathematical Society, New York, 1960).

Here ϵ is a measure of the departure of the medium from homogeneity while $L_1(\alpha)$ and $L_2(\alpha)$ are perturbing operators representing the effects of inhomogeneity. A wave $u(\alpha)$ in this medium satisfies

$$[L - \epsilon L_1(\alpha) - \epsilon^2 L_2(\alpha) + O(\epsilon^3)]u(\alpha) = 0. \quad (2)$$

We assume that α varies over a certain set A and that $p(\alpha)$ is the corresponding probability distribution in A. Then we designate expectation values with respect to $p(\alpha)$ by angular brackets

$$\langle f \rangle = \int_A f(\alpha)p(\alpha)\, d\alpha. \quad (3)$$

Our objective is to find an equation for $\langle u \rangle$.

Let us rewrite (2) as the new equation

$$u = u_0 + \epsilon L^{-1}(L_1 + \epsilon L_2)u + O(\epsilon^3). \quad (4)$$

To see that (4) implies (2) we merely apply L to both sides of (4). Now iteration of (4) yields the solution u as a series in ϵ, which we write as

$$u = u_0 + \epsilon L^{-1}L_1u_0 + \epsilon^2(L^{-1}L_1L^{-1}L_1 + L^{-1}L_2)u_0 + O(\epsilon^3). \quad (5)$$

Upon taking the expectation value of (5) we obtain

$$\langle u \rangle = u_0 + \epsilon L^{-1}\langle L_1 \rangle u_0 + \epsilon^2 L^{-1}(\langle L_1L^{-1}L_1 \rangle + \langle L_2 \rangle)u_0 + O(\epsilon^3). \quad (6)$$

From (6) we find that

$$\begin{aligned} u_0 &= \langle u \rangle - \epsilon L^{-1}\langle L_1 \rangle u_0 + O(\epsilon^2) \\ &= \langle u \rangle - \epsilon L^{-1}\langle L_1 \rangle \langle u \rangle + O(\epsilon^2). \end{aligned} \quad (7)$$

Substitution of (7) into (6) yields the following equation for $\langle u \rangle$:

$$\langle u \rangle = u_0 + \epsilon L^{-1}\langle L_1 \rangle \langle u \rangle + \epsilon^2 L^{-1}[\langle L_1L^{-1}L_1 \rangle - \langle L_1 \rangle L^{-1}\langle L_1 \rangle + \langle L_2 \rangle]\langle u \rangle + O(\epsilon^3). \quad (8)$$

Application of L to both sides of (8) yields a somewhat simpler equation for $\langle u \rangle$ which is

$$L\langle u \rangle - \epsilon\langle L_1 \rangle \langle u \rangle - \epsilon^2[\langle L_1L^{-1}L_1 \rangle - \langle L_1 \rangle L^{-1}\langle L_1 \rangle + \langle L_2 \rangle]\langle u \rangle = O(\epsilon^3). \quad (9)$$

When $\langle L_1 \rangle = 0$, which is a common case, (9) simplifies to

$$(L - \epsilon^2\langle L_1L^{-1}L_1 \rangle - \epsilon^2\langle L_2 \rangle)\langle u \rangle = O(\epsilon^3). \quad (10)$$

Equation (8), (9), or (10) becomes an explicit equation for $\langle u \rangle$ when the $O(\epsilon^3)$ term is omitted. Then any one of them becomes an equation of the type we want.

In many applications a wave is represented by an n component vector function of the position variable x. Then $u = u(x)$ is an n component vector and L, L_1, and L_2 are nth-order matrices of operators, each component of which may be a differential operator. In such cases the operator L^{-1} which appears in (10) may be represented as an integral operator, the kernel of which is the Green's matrix (also called the Green's tensor or dyadic) $G(x, x')$. G is an nth-order matrix defined by the equation

$$LG(x, x') = I\delta(x - x'). \quad (11)$$

Here I is the nth-order identity matrix and δ is the Dirac delta function. In terms of G, L^{-1} becomes

$$L^{-1}f = \int G(x, x')f(x')\, dx'. \quad (12)$$

Then the term in (10) which involves L^{-1} becomes

$$\langle L_1L^{-1}L_1 \rangle \langle u \rangle = \left\langle L_1(x) \int G(x, x')L_1(x')\langle u(x') \rangle\, dx' \right\rangle. \quad (13)$$

The notation $L_1(x)$ indicates that L_1 operates on functions of x and, if L_1 is not constant, is to be evaluated at x. When (13) is used in (10) and $O(\epsilon^3)$ is omitted, (10) becomes

$$L(x)\langle u(x) \rangle - \epsilon^2\left\langle L_1(x) \int G(x, x')L_1(x')\langle u(x') \rangle\, dx' \right\rangle - \epsilon^2\langle L_2(x) \rangle \langle u(x) \rangle = 0. \quad (14)$$

This is our main equation for $\langle u(x) \rangle$ when $\langle L_1 \rangle = 0$. If $\langle L_1 \rangle \neq 0$, we have instead from (9) and (13), up to $O(\epsilon^3)$,

$$\begin{aligned} &L(x)\langle u(x) \rangle - \epsilon\langle L_1(x) \rangle \langle u(x) \rangle \\ &\quad + \epsilon^2\Big[\langle L_1(x) \rangle \int G(x, x')\langle L_1(x') \rangle \langle u(x') \rangle\, dx' \\ &\quad - \left\langle L_1(x) \int G(x, x')L_1(x')\langle u(x') \rangle\, dx' \right\rangle \\ &\quad + \langle L_2(x) \rangle \langle u(x) \rangle \Big] = 0. \end{aligned} \quad (15)$$

3. SCALAR WAVES

We shall now apply the preceding considerations to a scalar wave $u(\mathbf{x})$ which satisfies the reduced wave equation in a slightly nonuniform medium,

$$\Delta u + k_0^2[1 + \epsilon\mu(\mathbf{x})]^2 u = 0. \quad (16)$$

Here $L = \Delta + k_0^2$, $L_1 = -2k_0^2\mu(\mathbf{x})$ and $L_2 = -k_0^2\mu^2(\mathbf{x})$. Let us assume that $\langle \mu(\mathbf{x}) \rangle = 0$ so that (14) applies. The Green's function of L is a scalar which we

make unique by requiring it to satisfy the radiation condition. Then G is given by

$$G(\mathbf{x}, \mathbf{x}') = -e^{ik_0|\mathbf{x}-\mathbf{x}'|}/4\pi\,|\mathbf{x} - \mathbf{x}'|. \tag{17}$$

Now (14) becomes

$$(\Delta + k_0^2)\langle u\rangle + \frac{\epsilon^2 k_0^4}{\pi}\left\langle \mu(\mathbf{x}) \int \frac{e^{ik_0|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|}\,\mu(\mathbf{x}')\right\rangle \times \langle u(\mathbf{x}')\rangle\, d\mathbf{x}' + \epsilon^2 k_0^2\langle\mu^2(\mathbf{x})\rangle\langle u(\mathbf{x})\rangle = 0. \tag{18}$$

In order to simplify (3) it is convenient to introduce the correlation coefficient of the inhomogeneity $\mu(\mathbf{x})$, defined by

$$N(\mathbf{x}, \mathbf{x}') = \langle\mu(\mathbf{x})\mu(\mathbf{x}')\rangle/\langle\mu^2(\mathbf{x})\rangle. \tag{19}$$

We shall now assume that $\langle\mu^2(\mathbf{x})\rangle$ is a constant and that N is a function only of the distance $|\mathbf{x} - \mathbf{x}'|$. Then (18) becomes

$$(\Delta + k_0^2 + \epsilon^2 k_0^2\langle\mu^2\rangle)\langle u(\mathbf{x})\rangle + \frac{\epsilon^2 k_0^4\langle\mu^2\rangle}{\pi}\int \frac{e^{ik_0 r}}{r} N(r)\langle u(\mathbf{x} + \mathbf{r})\rangle\, d\mathbf{r} = 0. \tag{20}$$

Let us now seek a plane-wave solution of (20), with amplitude A and propagation vector $\mathbf{k}$, of the form

$$\langle u(\mathbf{x})\rangle = Ae^{i\mathbf{k}\cdot\mathbf{x}}. \tag{21}$$

We must first perform the integration in (20) over the surface of a sphere of radius r centered at $\mathbf{x}$. To do so we make use of a mean-value theorem, which applies to any solution φ of

$$(\Delta + k^2)\varphi = 0. \tag{22}$$

It is

$$\frac{1}{4\pi r^2}\int \varphi(\mathbf{x} + \mathbf{r})\, ds = \frac{\sin kr}{kr}\,\varphi(\mathbf{x}). \tag{23}$$

Since the wave (21) satisfies (22), the theorem (23) applies. Upon using it in (20), we obtain

$$\left(\Delta + k_0^2 + \epsilon^2 k_0^2\langle\mu^2\rangle + \epsilon^2 4k_0^4\langle\mu^2\rangle k^{-1} \times \int_0^\infty e^{ik_0 r}\sin krN(r)\, dr\right)\langle u(\mathbf{x})\rangle = 0. \tag{24}$$

We observe that this equation is of the form (22) with k^2 given by

$$k^2 = k_0^2 + \epsilon^2 k_0^2\langle\mu^2\rangle + \epsilon^2 4k_0^4\langle\mu^2\rangle k^{-1} \times \int_0^\infty e^{ik_0 r}\sin krN(r)\, dr. \tag{25}$$

This is an equation for k. Upon solving it in powers of ϵ we obtain

$$\frac{k^2}{k_0^2} = 1 + \epsilon^2\langle\mu^2\rangle - 2i\epsilon^2 k_0\langle\mu^2\rangle \times \int_0^\infty (e^{2ik_0 r} - 1)N(r)\, dr + O(\epsilon^3). \tag{26}$$

It is customary to set $k/k_0 = n^*$, and to call n^* the refractive index of the medium. Thus the right side of (26) gives $(n^*)^2$ for the "effective" medium in which $\langle u\rangle$ propagates. The imaginary part of k is the attenuation coefficient for a wave in the medium. From (26) it is

$$\alpha = \operatorname{Im} k = \epsilon^2 k_0^2\langle\mu^2\rangle\int_0^\infty (1 - \cos 2k_0 r)N(r)\, dr. \tag{27}$$

We finally note that every solution of (20) satisfies (24), and not merely plane-wave solutions. This follows from the representation of any solution as a superposition of plane waves.

The results (24)–(27) have been derived before[1-3] in essentially the same way. In (24), k_0 appeared in place of k, but by (25) this is consistent since terms of order ϵ^3 have been omitted from (24). The expression (27) for α has also been derived previously on the basis of considerations of the energy scattered out of a beam.[4] That derivation shows that $\alpha > 0$, which has also been shown directly from (27) by Meecham.[3] (See also Ref. 1.)

Let us now apply (26) to a medium with the simple correlation function

$$N(r) = e^{-r/a}. \tag{28}$$

Here $a > 0$ is called the "correlation length" of the random inhomogeneity. From (26) and (28) we obtain, upon neglecting $O(\epsilon^3)$,

$$k^2/k_0^2 = 1 + \epsilon^2\langle\mu^2\rangle[1 - (2ik_0 a)^2/(1 - 2ik_0 a)]. \tag{29}$$

This is an explicit expression for $(k/k_0)^2$. It becomes infinite when $k_0 a$ is infinite, which shows that then (25) must be solved more accurately. Upon using (28) in (25) and omitting $O(\epsilon^3)$, (25) becomes

$$k^2 = k_0^2 + \epsilon^2 k_0^2\langle\mu^2\rangle - 2i\epsilon^2 k_0^4\langle\mu^2\rangle k^{-1} \times [(a^{-1} - ik_0 - ik)^{-1} - (a^{-1} - ik_0 + ik)^{-1}]. \tag{30}$$

This is an equation which can be solved for k. When $k_0 a$ is finite, it has the solution (29) while when $k_0 a$ is sufficiently large ($k_0 a \gg 1/\epsilon\langle\mu^2\rangle^{\frac{1}{2}}$) the solution is

$$k = k_0(1 + \epsilon\langle\mu^2\rangle^{\frac{1}{2}} + i/2k_0 a) + O(\epsilon^2). \tag{31}$$

[3] W. C. Meecham, "On radiation in a randomly inhomogeneous medium," Space Technology Laboratories Report, Los Angeles California; September 1961.

[4] L. A. Chernov, *Wave Propagation in a Random Medium* (McGraw-Hill Book Company, Inc., New York, 1960).

Thus the attenaution coefficient is proportional to $1/2a$ but independent of ϵ when the correlation length is very long compared to the wavelength, while Re $(k - k_0)$ is proportional to ϵ. On the other hand, both are proportional to ϵ^2 when the correlation length is comparable to the wavelength or smaller. The solution of (30) for $k_0 a$ finite describes the transition from one form to the other. The result (31) for Re k applies when $N(r) \equiv 1$, and does not depend upon the special form (28) as we see from (25) and (30). Thus the behavior of the real part of the propagation constant described above is typical for any correlation function.

4. ELASTIC WAVES

Let us consider the displacement vector $\mathbf{u}(\mathbf{x})$ in a time-harmonic elastic wave motion. According to the linear theory of elasticity, $\mathbf{u}$ satisfies the equation

$$(\lambda + \mu)\nabla(\nabla\cdot\mathbf{u}) + \mu\nabla^2\mathbf{u} + \nabla\lambda(\nabla\cdot\mathbf{u}) + \nabla\mu\times(\nabla\times\mathbf{u}) + 2(\nabla\mu\cdot\nabla)\mathbf{u} + \omega^2\rho\mathbf{u} = 0. \quad (32)$$

Here ρ is the density of the medium, λ and μ are its Lamé constants, and ω is the angular frequency of the motion. We assume that $\rho(\mathbf{x})$, $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ differ slightly from the constants ρ_0, λ_0 and μ_0 so we write

$$\rho(\mathbf{x}) = \rho_0 + \epsilon\rho_1(\mathbf{x}), \qquad \lambda(\mathbf{x}) = \lambda_0 + \epsilon\lambda_1(\mathbf{x}), \qquad \mu(\mathbf{x}) = \mu_0 + \epsilon\mu_1(\mathbf{x}). \quad (33)$$

We assume that

$$\langle\rho_1(\mathbf{x})\rangle = \langle\lambda_1(\mathbf{x})\rangle = \langle\mu_1(\mathbf{x})\rangle = 0. \quad (34)$$

When (33) is used in (32), (32) becomes of the form (2) with $L_2 = 0$ and L and L_1 given by

$$L\mathbf{u} = (\lambda_0 + \mu_0)\nabla(\nabla\cdot\mathbf{u}) + \mu_0\nabla^2\mathbf{u} + \omega^2\rho_0\mathbf{u} \quad (35)$$

$$L_1\mathbf{u} = (\lambda_1 + \mu_1)\nabla(\nabla\cdot\mathbf{u}) + \mu_1\nabla^2\mathbf{u} + \nabla\lambda_1(\nabla\cdot\mathbf{u}) + \nabla\mu_1\times(\nabla\times\mathbf{u}) + 2(\nabla\mu_1\cdot\nabla)\mathbf{u} + \omega^2\rho_1\mathbf{u}. \quad (36)$$

It follows from (34) that $\langle L_1\rangle = 0$. Thus (14) applies. To utilize it we need the Green's tensor of L, defined by (35). This tensor depends only upon $\mathbf{r} = \mathbf{x} - \mathbf{x}'$ and is given by

$$\mathbf{G}(\mathbf{x},\mathbf{x}') = +\frac{1}{\omega^2\rho_0}\left[\delta(\mathbf{r}) - \frac{e^{ik_c r}}{4\pi r^3}(1 - ik_c r) + \frac{e^{ik_s r}}{4\pi r^3}(1 - ik_s r - k_s^2 r^2)\right]\mathbf{I} + \frac{1}{4\pi\omega^2\rho_0 r^3}[e^{ik_c r}(3 - 3ik_c r - k_c^2 r^2) - e^{ik_s r}(3 - 3ik_s r - k_s^2 r^2)]\hat{\mathbf{r}}\hat{\mathbf{r}}. \quad (37)$$

Here $\hat{\mathbf{r}}$ is a unit vector in the direction of $\mathbf{r}$, $\mathbf{I}$ is the unit dyadic, and the compressional and shear propagation constants k_c and k_s are given by

$$k_c = \omega[\lambda_0 + 2\mu_0)/\rho_0]^{\frac{1}{2}}, \qquad k_s = \omega(\mu_0/\rho_0)^{\frac{1}{2}}. \quad (38)$$

It will prove convenient to denote the coefficients of $\mathbf{I}$ and $\hat{\mathbf{r}}\hat{\mathbf{r}}$ in (37) by $G_1(r)$ and $G_2(r)$, respectively, and to write $\mathbf{G}$ in the form

$$\mathbf{G}(\mathbf{x},\mathbf{x}') = G_1(r)\mathbf{I} + G_2(r)\hat{\mathbf{r}}\hat{\mathbf{r}}. \quad (39)$$

We shall now assume that $\langle\mathbf{u}(\mathbf{x})\rangle$ is a plane wave given by

$$\langle\mathbf{u}(\mathbf{x})\rangle = \mathbf{A}e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (40)$$

To determine the propagation vector $\mathbf{k}$ and the amplitude vector $\mathbf{A}$ we shall insert (40) into (14), making use of (35)–(37). In doing this we are first led to calculate $L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle$, which is

$$L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle = [\{\omega^2\rho_1(\mathbf{x}') - \mu_1(\mathbf{x}')k^2 + i(\nabla'\mu_1(\mathbf{x}')\cdot\mathbf{k})\}\mathbf{A} + \{-\lambda_1(\mathbf{x}')(\mathbf{k}\cdot\mathbf{A}) - \mu_1(\mathbf{x}')(\mathbf{k}\cdot\mathbf{A}) + i(\nabla'\mu_1(\mathbf{x}')\cdot\mathbf{A})\}\mathbf{k} + i\nabla'\lambda_1(\mathbf{x}')(\mathbf{k}\cdot\mathbf{A})]e^{i\mathbf{k}\cdot\mathbf{x}'}. \quad (41)$$

In (41), ∇' denotes the gradient with respect to $\mathbf{x}'$. Next we calculate $L_1(\mathbf{x})\mathbf{G}(\mathbf{x},\mathbf{x}')$, making use of the form (39) for $\mathbf{G}$. The result is, after some simplification,

$$\begin{aligned} L_1(\mathbf{x})\mathbf{G}(\mathbf{x},\mathbf{x}') = {} & \Big[\{\lambda_1(\mathbf{x}) + \mu_1(\mathbf{x})\}\Big\{\frac{1}{r}\frac{\partial G_1}{\partial r} + \frac{1}{r}\frac{\partial G_2}{\partial r}\Big\} + \mu_1(\mathbf{x})(\nabla^2 G_1) + \Big\{\frac{\partial G_1}{\partial r} + \frac{G_2}{r}\Big\} \\ & \times(\nabla\mu_1(\mathbf{x})\cdot\hat{\mathbf{r}}) + \omega^2\rho_1(\mathbf{x})G_1\Big]\mathbf{I} \\ & + \Big[\{\lambda_1(\mathbf{x}) + \mu_1(\mathbf{x})\}\Big\{-\frac{1}{r}\frac{\partial G_1}{\partial r} + \frac{\partial^2 G_1}{\partial r^2} - \frac{1}{r}\frac{\partial G_2}{\partial r} + \frac{\partial^2 G_2}{\partial r^2}\Big\} + \mu_1(\mathbf{x})(\nabla^2 G_2) \\ & + 2(\nabla\mu_1(\mathbf{x})\cdot\hat{\mathbf{r}})\frac{\partial G_2}{\partial r} + \omega^2\rho_1(\mathbf{x})G_2\Big]\hat{\mathbf{r}}\hat{\mathbf{r}} \\ & + \Big[\frac{\partial G_1}{\partial r} + \frac{\partial G_2}{\partial r}\Big]\nabla\lambda_1(\mathbf{x})\hat{\mathbf{r}} \\ & + \Big[\frac{\partial G_1}{\partial r} - \frac{G_2}{r}\Big]\nabla\mu_1(\mathbf{x})\hat{\mathbf{r}}. \end{aligned} \quad (42)$$

We now multiply (42) by (40) to form

$$L_1(\mathbf{x})\mathbf{G}(\mathbf{x},\mathbf{x}')L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle.$$

The product is a very long expression so we shall not write it out. Then we compute the mean value of the product to obtain $\langle L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle\rangle$. In doing so we find that nine correlation functions of the form $\langle\lambda_1(\mathbf{x})\lambda_1(\mathbf{x}')\rangle$, $\langle\lambda_1(\mathbf{x})\mu_1(\mathbf{x}')\rangle$, etc., occur. We assume that each of these correlations depends only upon the distance $r = |\mathbf{x} - \mathbf{x}'|$ between $\mathbf{x}$ and $\mathbf{x}'$, which means that the medium is statistically homogeneous and isotropic (at least insofar as second-order moments are concerned). Therefore we may write $\langle\lambda_1(\mathbf{x})\lambda_1(\mathbf{x}')\rangle = R_{\lambda\lambda}(r)$, $\langle\lambda_1(\mathbf{x})\mu_1(\mathbf{x}')\rangle = R_{\lambda\mu}(r)$, etc., which defines the correlation functions $R_{\lambda\lambda}$, $R_{\lambda\mu}$, etc. After some rearrangement of terms and some simplification we obtain

$$\langle L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle\rangle e^{-i\mathbf{k}\cdot\mathbf{x}'} = E_1\mathbf{A} + E_2(\hat{\mathbf{r}}\cdot\mathbf{k})\mathbf{A} + B_1(\mathbf{k}\cdot\mathbf{A})\hat{\mathbf{r}} + B_2(\hat{\mathbf{r}}\cdot\mathbf{A})\hat{\mathbf{r}} + B_3(\mathbf{k}\cdot\mathbf{A})(\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}} + B_4(\hat{\mathbf{r}}\cdot\mathbf{A})(\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}} + C_1(\mathbf{k}\cdot\mathbf{A})\mathbf{k} + C_2(\hat{\mathbf{r}}\cdot\mathbf{A})\mathbf{k}. \quad (43)$$

In (12) the eight quantities E_i, B_i, and C_i are scalars depending upon k, r, ω, and the correlation functions $R_{\lambda\lambda}$, $R_{\lambda\mu}$, etc., but independent of $\hat{\mathbf{r}}$. They are defined by the following equations, in which primes denote derivatives with respect to r.

$$E_1 = \left\{\frac{G_1'}{r} + \frac{G_2'}{r}\right\}[\omega^2(R_{\lambda\rho} + R_{\mu\rho}) - k^2(R_{\lambda\mu} + R_{\mu\mu})] + (\nabla^2G_1)[\omega^2R_{\mu\rho} - k^2R_{\mu\mu}] + \left\{G_1' + \frac{G_2}{r}\right\} \times [\omega^2R_{\mu\rho}' - k^2R_{\mu\mu}'] + \omega^2G_1[\omega^2R_{\rho\rho} - k^2R_{\rho\mu}], \quad (44)$$

$$E_2 = \left\{\frac{G_1'}{r} + \frac{G_2'}{r}\right\}[-iR_{\lambda\mu}' - iR_{\mu\mu}'] + (\nabla^2G_1)[-iR_{\mu\mu}'] + \left\{G_1' + \frac{G_2}{r}\right\}[-iR_{\mu\mu}''] - i\omega^2G_1R_{\rho\mu}', \quad (45)$$

$$B_1 = i\left\{\frac{G_1'}{r} + \frac{G_2'}{r}\right\}[-R_{\lambda\lambda}' - R_{\mu\lambda}'] - i(\nabla^2G_1)R_{\mu\lambda}' - i\left\{G_1' + \frac{G_2}{r}\right\}R_{\mu\lambda}' - i\omega^2G_1R_{\rho\lambda}' + i\left\{-\frac{G_1'}{r} + G_1'' - \frac{G_2'}{r} + G_2''\right\}[-R_{\lambda\lambda}' - R_{\mu\lambda}'] - i(\nabla^2G_2)R_{\mu\lambda}' - i2G_2'R_{\mu\lambda}'' - i\omega^2G_2R_{\rho\lambda}' - i\{G_1' + G_2'\}R_{\lambda\lambda}'' - i\left\{G_1' - \frac{G_2}{r}\right\}R_{\mu\lambda}'', \quad (46)$$

$$B_2 = \left\{-\frac{G_1'}{r} + G_1'' - \frac{G_2'}{r} + G_2''\right\}[\omega^2(R_{\lambda\rho} + R_{\mu\rho}) - k^2(R_{\lambda\mu} + R_{\mu\mu})] + (\nabla^2G_2)[\omega^2R_{\mu\rho} - k^2R_{\mu\mu}] + 2G_2'[\omega^2R_{\mu\rho}' - k^2R_{\mu\mu}'] + \omega^2G_2[\omega^2R_{\rho\rho} - k^2R_{\rho\mu}] + \{G_1' + G_2'\}[\omega^2R_{\lambda\rho}' - k^2R_{\lambda\mu}'] + \left\{G_1' - \frac{G_2}{r}\right\}[\omega^2R_{\mu\rho}' - k^2R_{\mu\mu}'], \quad (47)$$

$$B_3 = \left\{-\frac{G_1'}{r} + G_1'' - \frac{G_2'}{r} + G_2''\right\} \times [-(R_{\lambda\lambda} + 2R_{\mu\lambda} + R_{\mu\mu})] + (\nabla^2G_2)[-R_{\mu\lambda} - R_{\mu\mu}] + 2G_2'[-R_{\mu\lambda}' - R_{\mu\mu}'] + \omega^2G_2[-R_{\rho\lambda} - R_{\rho\mu}] + \{G_1' + G_2'\}[-R_{\lambda\lambda}' - R_{\lambda\mu}'] + \left\{G_1' - \frac{G_2}{r}\right\}[-R_{\mu\lambda}' - R_{\mu\mu}'], \quad (48)$$

$$B_4 = +\left\{-\frac{G_1'}{r} + G_1'' - \frac{G_2'}{r} + G_2''\right\}[-iR_{\lambda\mu}' - iR_{\mu\mu}'] + (\nabla^2G_2)[-iR_{\mu\mu}'] + 2G_2'[-iR_{\mu\mu}''] + \omega^2G_2[-iR_{\rho\mu}'] + \{G_1' + G_2'\}[-iR_{\lambda\mu}''] + \left\{G_1' - \frac{G_2}{r}\right\}[-iR_{\mu\mu}''] + \left\{-\frac{G_1'}{r} + G_1'' - \frac{G_2'}{r} + G_2''\right\}[-iR_{\lambda\mu}' - iR_{\mu\mu}'] + (\nabla^2G_2)[-iR_{\mu\mu}'] + 2G_2'[-iR_{\mu\mu}''] + \omega^2G_2[-iR_{\rho\mu}'] + \{G_1' + G_2'\}[-iR_{\lambda\mu}''] + \left\{G_1' - \frac{G_2}{r}\right\}[-iR_{\mu\mu}''], \quad (49)$$

$$C_1 = \left\{\frac{G_1'}{r} + \frac{G_2'}{r}\right\}[-(R_{\lambda\lambda} + 2R_{\mu\lambda} + R_{\mu\mu})] + (\nabla^2G_1)[-R_{\mu\lambda} - R_{\mu\mu}] + \left\{G_1' + \frac{G_2}{r}\right\} \times [-R_{\mu\lambda}' - R_{\mu\mu}'] + \omega^2G_1[-R_{\rho\lambda} - R_{\rho\mu}], \quad (50)$$

$$C_2 = \left\{\frac{G_1'}{r} + \frac{G_2'}{r}\right\}[-iR_{\lambda\mu}' - iR_{\mu\mu}'] + (\nabla^2G_1)[-iR_{\mu\mu}'] + \left\{G_1' + \frac{G_2}{r}\right\}[-iR_{\mu\mu}''] + \omega^2G_1[-iR_{\rho\mu}']. \quad (51)$$

To compute $\langle L_1(\mathbf{x}) \int \mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle\, d\mathbf{x}'\rangle$ we multiply (43) by $e^{i\mathbf{k}\cdot\mathbf{x}'}$ and integrate with respect to $\mathbf{x}'$. In evaluating the integral it is convenient to set $\mathbf{x}' = \mathbf{x} - \mathbf{r}$ and integrate with respect to $\mathbf{r}$. Then we set $d\mathbf{r} = dr\, dS$, where dS denotes the area element on a sphere of radius r centered at x and obtain

$$\left\langle L_1(\mathbf{x}) \int \mathbf{G}(\mathbf{x}, \mathbf{x}')\langle\mathbf{u}(\mathbf{x}')\rangle\, d\mathbf{x}'\right\rangle = \left[+\mathbf{A}\int E_1 \int e^{-i\mathbf{k}\cdot\mathbf{r}}\, dS\, dr + \mathbf{A}\int E_2 \int (\hat{\mathbf{r}}\cdot\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}\, dS\, dr\right.$$

$$+ (\mathbf{k}\cdot\mathbf{A}) \int B_1 \int \hat{\mathbf{r}} e^{-ik\cdot r}\, dS\, dr$$

$$+ \int B_2 \int (\hat{\mathbf{r}}\cdot\mathbf{A})\hat{\mathbf{r}} e^{-ik\cdot r}\, dS\, dr$$

$$+ (\mathbf{k}\cdot\mathbf{A}) \int B_3 \int (\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}} e^{-ik\cdot r}\, dS\, dr$$

$$+ \int B_4 \int (\hat{\mathbf{r}}\cdot\mathbf{A})(\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}} e^{-ik\cdot r}\, dS\, dr$$

$$+ \mathbf{k}(\mathbf{k}\cdot\mathbf{A}) \int C_1 \int e^{-ik\cdot r}\, dS\, dr$$

$$+ \mathbf{k} \int C_2 \int (\hat{\mathbf{r}}\cdot\mathbf{A}) e^{-ik\cdot r}\, dS\, dr \Big] e^{ik\cdot x}. \tag{52}$$

The integration with respect to dS can be done explicitly in terms of the basic integral

$$\int e^{-ik\cdot r}\, dS = f(r). \tag{53}$$

Here $f(r)$ is defined by

$$f(r) = (4\pi r/k) \sin kr. \tag{54}$$

The integrals are evaluated in Appendix 1. The result of integrating (52) is

$$\left\langle L_1(\mathbf{x}) \int \mathbf{G}(\mathbf{x}, \mathbf{x}') L_1(\mathbf{x}') \langle \mathbf{u}(\mathbf{x}') \rangle\, d\mathbf{x}' \right\rangle$$

$$= \Big[+\mathbf{A} \int E_1 f\, dr + ik\mathbf{A} \int E_2 \frac{1}{r} \frac{\partial f}{\partial k}\, dr$$

$$+ ik\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \int B_1 \frac{1}{r} \frac{\partial f}{\partial k}\, dr$$

$$- \int B_2 \frac{1}{r^2} \left[\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) + \mathbf{A} \frac{1}{k} \frac{\partial f}{\partial k} \right] dr$$

$$- k^2 \hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \int B_3 \frac{1}{r^2} \frac{\partial^2 f}{\partial k^2}\, dr$$

$$- ik \int B_4 \frac{1}{r^3} \left[\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \frac{\partial}{\partial k} \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) \right.$$

$$\left. + \mathbf{A} \frac{1}{k} \frac{\partial}{\partial k} \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) \right] dr + k^2 \hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \int C_1 f\, dr$$

$$+ ik\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A}) \int C_2 \frac{1}{r} \frac{\partial f}{\partial k}\, dr \Big] e^{ik\cdot x}. \tag{55}$$

In (24), $\hat{\mathbf{k}}$ denotes a unit vector in the direction of $\mathbf{k}$.

Equation (24) can be rewritten simply in the form

$$\left\langle L_1(\mathbf{x}) \int \mathbf{G}(\mathbf{x}, \mathbf{x}') L_1(\mathbf{x}') \langle \mathbf{u}(\mathbf{x}') \rangle\, d\mathbf{x}' \right\rangle$$

$$= [D_1\mathbf{A} + D_2\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A})] e^{ik\cdot x}. \tag{56}$$

Here D_1 and D_2 are defined by the integrals

$$D_1 = \int E_1 f\, dr + ik \int E_2 \frac{1}{r} \frac{\partial f}{\partial k}\, dr - \frac{1}{k} \int B_2 \frac{1}{r^2} \frac{\partial f}{\partial k}\, dr$$

$$- i \int B_4 \frac{1}{r^3} \frac{\partial}{\partial k} \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) dr, \tag{57}$$

$$D_2 = ik \int B_1 \frac{1}{r} \frac{\partial f}{\partial k}\, dr - \int B_2 \frac{1}{r^2} \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) dr$$

$$- k^2 \int B_3 \frac{1}{r^2} \frac{\partial^2 f}{\partial k^2}\, dr$$

$$- ik \int B_4 \frac{1}{r^3} \frac{\partial}{\partial k} \left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k} \frac{\partial f}{\partial k} \right) dr$$

$$+ k^2 \int C_1 f\, dr + ik \int C_2 \frac{1}{r} \frac{\partial f}{\partial k}\, dr. \tag{58}$$

Now from (35)

$$L(\mathbf{x})\langle \mathbf{u}(\mathbf{x}) \rangle = [-(\lambda_0 + \mu_0)k^2(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}}$$

$$- (\mu_0 k^2 - \omega^2 \rho_0)\mathbf{A}] e^{ik\cdot x}. \tag{59}$$

Finally with the insertion of (56) and (59), (14) becomes

$$[\{\mu_0 k^2 - \omega^2 \rho_0 + \epsilon^2 D_1\}\mathbf{A} + \{(\lambda_0 + \mu_0)k^2$$

$$+ \epsilon^2 D_2\}(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}}] e^{ik\cdot x} = 0. \tag{60}$$

To analyze (60) it is convenient to denote by $A_\parallel$ and $A_\perp$ the components of $\mathbf{A}$ respectively parallel and perpendicular to the propagation vector $\mathbf{k}$, to which $\hat{\mathbf{k}}$ is parallel. Then (60) may be written as the two equations

$$A_\parallel [(\lambda_0 + 2\mu_0)k^2 - \omega^2 \rho_0$$

$$+ \epsilon^2 (D_1 + D_2)] e^{ik\cdot x} = 0, \tag{61}$$

$$A_\perp [\mu_0 k^2 - \omega^2 \rho_0 + \epsilon^2 D_1] e^{ik\cdot x} = 0. \tag{62}$$

From these equations we see that there are two cases, in which either $A_\parallel \neq 0$ or $A_\perp \neq 0$ while the other is zero. When $A_\parallel \neq 0$ we say that the average wave is a longitudinal motion. Then from (61) it follows that the propagation constant k is determined by the equation

$$(\lambda_0 + 2\mu_0)k^2 - \omega^2 \rho_0 + \epsilon^2 (D_1 + D_2) = 0. \tag{63}$$

When $\epsilon^2 = 0$, the solution of the equation is $k = k_c$. Therefore to obtain a solution correct to order ϵ^2 we may merely replace k by k_c in D and D_2 and find

$$k^2 = k_c^2 - [\epsilon^2/(\lambda_0 + 2\mu_0)][D_1(k_c) + D_2(k_c)]. \tag{64}$$

In the other case, when $A_\perp \neq 0$, we say that the average motion is transverse. From (62) we see that k then satisfies the equation

$$\mu_0 k^2 - \omega^2 \rho_0 + \epsilon^2 D_1 = 0. \tag{65}$$

When $\epsilon^2=0$, the solution of (65) is $k=(\omega^2\rho_0/\mu_0)^{\frac{1}{2}}=k_s$. The solution up to order ϵ^2 is

$$k^2 = k_s^2 - (\epsilon^2/\mu_0)D_1(k_s). \tag{66}$$

In both cases the imaginary part of k is the attenuation coefficient, which we see is of order ϵ^2 (but see the next section).

In concluding this section, we observe that (64) and (66) are explicit expressions for k^2 in the longitudinal and transverse cases, respectively. To compute them the expressions for D_1 and D_2, given by (67) and (68), must be evaluated. These expressions involve various integrals of the correlation functions of the random inhomogeneity. In the next section these results will be examined and simplified in several special cases, and an example will be given.

5. SPECIAL ELASTIC MEDIA AND AN EXAMPLE

Let us first consider the simplifications of the above results which occur when all the correlation functions except $R_{\lambda\lambda}$ are zero. This will be the case, for example, if ρ and μ are constant while only λ is variable. Then it follows from (44)–(51) that $E_1 = E_2 = B_2 = B_4 = C_2 = 0$ and

$$B_1 = -i(G_1'' + G_2'')R_{\lambda\lambda}' - i(G_1' + G_2')R_{\lambda\lambda}'', \tag{67}$$

$$B_3 = (G_1'/r - G_1'' + G_2'/r - G_2'')R_{\lambda\lambda} - (G_1' + G_2')R_{\lambda\lambda}', \tag{68}$$

$$C_1 = -r^{-1}(G_1' + G_2')R_{\lambda\lambda}. \tag{69}$$

From (57) we then find $D_1 = 0$ and from (58)

$$D_2 = ik\int B_1\frac{1}{r}\frac{\partial f}{\partial k}\,dr - k^2\int B_3\frac{1}{r^2}\frac{\partial^2 f}{\partial k^2}\,dr + k^2\int C_1 f\,dr. \tag{70}$$

In the longitudinal case $A_1 \neq 0$, k^2 still satisfies (63) and is given approximately by (64), with $D_1 = 0$. However, in the transverse case $A_\perp \neq 0$, (65) yields $k = k_s$ up to order ϵ^2. Thus to order ϵ^2, random inhomogeneities in λ do not affect the mean transverse wave but do affect the mean longitudinal wave. This might have been expected since λ is essentially a coefficient of compressional elasticity and the longitudinal wave is a compressional wave while the transverse wave is a shear wave.

Next we shall consider the case in which all correlation functions except $R_{\rho\rho}$ are zero. This will be the case, for example, if the elastic coefficients are constant while the density ρ is random. It follows from (44)–(51) that $E_1 = \omega^4 G_1 R_{\rho\rho}$, $B_2 = \omega^4 G_2 R_{\rho\rho}$ and $E_2 = B_1 = B_3 = B_4 = C_1 = C_2 = 0$. Then from (57) and (58)

$$D_1 = \omega^4\int G_1 R_{\rho\rho} f\,dr - \omega^4\int\frac{1}{r^2}G_2 R_{\rho\rho}\frac{1}{k}\frac{\partial f}{\partial k}\,dr, \tag{71}$$

$$D_2 = -\omega^4\int\frac{1}{r^2}G_2 R_{\rho\rho}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right)dr. \tag{72}$$

The progagation constant k^2 still satisfies (63) in the longitudinal case and (65) in the transverse case, and it is given to order ϵ^2 by (64) and (66) in the respective cases, with D_1 and D_2 given by (71) and (72).

To obtain explicit results, let us choose for $R_{\rho\rho}$ the simple correlation function

$$R_{\rho\rho}(r) = \langle\rho_1^2\rangle e^{-r/a}. \tag{73}$$

Here $\langle\rho_1^2\rangle$ is the mean square value of ρ_1 and a is the correlation length of the inhomogeneity. We now insert (73) into (71) and (72), together with the expressions in (37) for G_1 and G_2 and the definition (54) of f. Upon using the mean-value theorems (see Appendix I) to evaluate the resulting integrals, we obtain

$$D_1 = \omega^4\int G_1 R_{\rho\rho} f\,dr - \omega^4\int\frac{1}{r^2}G_2 R_{\rho\rho}\frac{1}{k}\frac{\partial f}{\partial k}\,dr, \tag{74}$$

$$D_2 = -\omega^4\int\frac{1}{r^2}G_2 R_{\rho\rho}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right)dr. \tag{75}$$

When the integrals in (74) and (75) are evaluated (see Appendix II) and used in (4.32), we obtain the longitudinal case

$$\frac{k^2}{k_c^2} = 1 + \epsilon^2\frac{\langle\rho_1^2\rangle}{\rho_0^2}\left[-\frac{a^{-1}}{k}\left(1 + \frac{a^{-2}}{k^2} + \frac{k_c^2}{k^2}\right)\right.$$
$$\times\cot^{-1}\frac{(a^{-1} - ik_c)}{k} + \frac{a^{-1}}{k}\left(1 + \frac{a^{-2}}{k^2} + \frac{k_s^2}{k^2}\right)$$
$$\times\cot^{-1}\frac{(a^{-1} - ik_s)}{k} + \frac{k_c^2}{k^2 + (a^{-1} - ik_c)^2}$$
$$\left. + i\frac{a^{-1}}{k}\left(\frac{k_c}{k} - \frac{k_s}{k}\right)\right]. \tag{76}$$

Similarly, in the transverse case, (65) yields

$$\frac{k^2}{k_s^2} = 1 + \epsilon^2\frac{\langle\rho_1^2\rangle}{2\rho_0^2}$$
$$\times\left[-\frac{a^{-1}}{k}\left(1 + \frac{a^{-2}}{k^2} + \frac{k_s^2}{k^2}\right)\cot^{-1}\frac{(a^{-1} - ik_s)}{k}\right.$$
$$+ \frac{a^{-1}}{k}\left(1 + \frac{a^{-2}}{k^2} + \frac{k_c^2}{k^2}\right)\cot^{-1}\frac{(a^{-1} - ik_c)}{k}$$
$$\left. + \frac{2k_s^2}{k^2 + (a^{-1} - ik_s)^2} + i\frac{a^{-1}}{k}\left(\frac{k_s}{k} - \frac{k_c}{k}\right)\right]. \tag{77}$$

Equations (76) and (77) can now be solved for k. When $k_c a$ and $k_s a$ are finite, the solutions for the longitudinal and transverse propagation constants up to terms in ϵ^2 are

$$\frac{k^2}{k_c^2} = 1 + \epsilon^2 \frac{\langle \rho_1^2 \rangle}{\rho_0^2} \left[-\frac{1}{k_c a}\left(2 + \frac{1}{k_c^2 a^2}\right) \cot^{-1}\left(\frac{1 - ik_c a}{k_c a}\right) + \frac{1}{k_c a}\left(1 + \frac{v_c^2}{v_s^2} + \frac{1}{k_c^2 a^2}\right) \cot^{-1}\left(\frac{1 - iv_c v_s^{-1} k_c a}{k_c a}\right) + \frac{k_c^2 a^2}{k_c^2 a^2 + (1 - ik_c a)^2} - i \frac{1}{k_c a}\left(\frac{v_c}{v_s} - 1\right)\right] \tag{78}$$

and

$$\frac{k^2}{k_s^2} = 1 + \epsilon^2 \frac{\langle \rho_1^2 \rangle}{2\rho_0^2} \left[-\frac{1}{k_s a}\left(2 + \frac{1}{k_s^2 a^2}\right) \cot^{-1}\left(\frac{1 - ik_s a}{k_s a}\right) + \frac{1}{k_s a}\left(1 + \frac{v_s^2}{v_c^2} + \frac{1}{k_s^2 a^2}\right) \cot^{-1}\left(\frac{1 - iv_s v_c^{-1} k_s a}{k_s a}\right) + \frac{2k_s^2 a^2}{k_s^2 a^2 + (1 - ik_s a)^2} + i \frac{1}{k_s a}\left(1 - \frac{v_s}{v_c}\right)\right]. \tag{79}$$

The above results are explicit expressions for $(k/k_c)^2$ and $(k/k_s)^2$. They both become infinite when $k_c a$ and $k_s a$ become infinite, respectively, which shows that (76) and (77) must be solved more accurately. When $k_c a$ is sufficiently large ($k_c a \gg \rho_0/\epsilon\langle \rho_1^2 \rangle^{\frac{1}{2}}$) the longitudinal propagation constant is

$$k = k_c[1 + \tfrac{1}{2}\epsilon\langle \rho_1^2 \rangle^{\frac{1}{2}}/\rho_0 + i/2k_c a] + O(\epsilon^2). \tag{80}$$

Thus the longitudinal attenuation coefficient is proportional to $1/2a$ but independent of ϵ when the correlation length is very long compared to the wavelength, while Re $(k - k_c)$ is proportional to ϵ. On the other hand, both Im (k) and Re $(k - k_c)$ are proportional to ϵ^2 when the correlation length is comparable to the wavelength or smaller. The solution of (76) for $k_c a$ finite describes the transition from one form to the other.

When $k_s a$ is sufficiently large ($k_s a \gg \rho_0/\epsilon\langle \rho_1^2 \rangle^{\frac{1}{2}}$), the transverse propagation constant is

$$k = k_s[1 + \tfrac{1}{2}\epsilon\langle \rho_1^2 \rangle^{\frac{1}{2}}/\rho_0 + i/2k_s a] + O(\epsilon^2). \tag{81}$$

The behavior of the transverse propagation constant is the same as that described for the longitudinal propagation constant.

6. ELECTROMAGNETIC WAVES

Let us now apply our general theory to electromagnetic wave propagation in a medium in which the dielectric constant ϵ', the magnetic permeability μ, and the electrical conductivity σ are all random functions of position. We shall consider time-harmonic waves of angular frequency ω and therefore it is convenient to introduce the effective dielectric constant $\epsilon = \epsilon' + i\omega^{-1}\sigma$. Then it follows from Maxwell's equations that the electric field vector $\mathbf{E}$ satisfies the equation

$$\nabla \times \nabla \times \mathbf{E} - \omega^2 \mu \epsilon \mathbf{E} - \mu^{-1} \nabla \mu \times \nabla \times \mathbf{E} = 0. \tag{82}$$

As usual, the corresponding equation for the magnetic field $\mathbf{H}$ is obtained from (82) by interchanging $\mathbf{E}$ with $\mathbf{H}$ and ϵ with $-\mu$. Therefore we needn't consider the equation for $\mathbf{H}$ since we can obtain the results for $\mathbf{H}$ from those for $\mathbf{E}$ by making the indicated interchanges.

We now assume that ϵ', μ, and σ are of the form

$$\epsilon' = \epsilon_0'[1 + \eta\epsilon_1'(\mathbf{x})], \tag{83}$$

$$\mu = \mu_0[1 + \eta\mu_1(\mathbf{x})], \tag{84}$$

$$\sigma = \sigma_0[1 + \eta\sigma_1(\mathbf{x})]. \tag{85}$$

Then ϵ may be written in a similar form which defines ϵ_0 and ϵ_1,

$$\epsilon = \epsilon_0[1 + \eta\epsilon_1(\mathbf{x})] = (\epsilon_0' + i\omega^{-1}\sigma_0) \times \left[1 + \eta \frac{\epsilon_0'\epsilon_1' + i\omega^{-1}\sigma_0\sigma_1}{\epsilon_0' + i\omega^{-1}\sigma_0}\right]. \tag{86}$$

Here we have denoted the small parameter by η to avoid confusion with the effective dielectric constant. The quantities ϵ_0', μ_0, and σ_0 are constants. Upon inserting (83)–(85) into (82) and setting $k_0^2 = \omega^2\mu_0\epsilon_0$, we obtain

$$\nabla \times \nabla \times \mathbf{E} - k_0^2\mathbf{E} = \eta[k_0^2(\mu_1 + \epsilon_1)\mathbf{E} + \nabla\mu_1 \times \nabla \times \mathbf{E}] + \eta^2[k_0^2\mu_1\epsilon_1\mathbf{E} - \mu_1\nabla\mu_1 \times \nabla \times \mathbf{E}] + O(\eta^3). \tag{87}$$

This equation is of the form (2) with L, L_1, and L_2 given by

$$L = \nabla \times \nabla \times - k_0^2, \tag{88}$$

$$L_1 = k_0^2(\mu_1 + \epsilon_1) + \nabla\mu_1 \times \nabla \times , \tag{89}$$

$$L_2 = k_0^2\mu_1\epsilon_1 - \mu_1\nabla\mu_1 \times \nabla \times . \tag{90}$$

We shall also assume that $\langle \mu_1 \rangle = \langle \epsilon_1 \rangle = 0$, which implies that $\langle L_1 \rangle = 0$. Then the mean field $\langle \mathbf{E} \rangle$ satisfies (14).

In (14) there occurs the Green's tensor $\mathbf{G}(\mathbf{x}, \mathbf{x}')$ associated with the operator L defined by (88). This tensor is given by

$$\mathbf{G}(\mathbf{x}, \mathbf{x}') = (\mathbf{I} - k_0^{-2}\nabla\nabla')e^{ik_0|\mathbf{x}-\mathbf{x}'|}/(|\mathbf{x} - \mathbf{x}'|). \tag{91}$$

As before $\mathbf{I}$ is the unit dyadic and ∇' is the gradient with respect to $\mathbf{x}'$. Upon carrying out the differentiations in (91) we can write $\mathbf{G}$ in the form

$$\mathbf{G}(\mathbf{x}, \mathbf{x}') = G_1(r)\mathbf{I} + G_2(r)\hat{\mathbf{r}}\hat{\mathbf{r}}. \tag{92}$$

Here G_1 and G_2 are defined by

$$G_1(r) = (-1 + ik_0r + k_0^2r^2)e^{ik_0r}/4\pi k_0^2r^3, \quad (93)$$

$$G_2(r) = (3 - 3ik_0r - k_0^2r^2)e^{ik_0r}/4\pi k_0^2r^3. \quad (94)$$

To determine the propagation constant k for a plane wave in the random medium, we shall seek a plane-wave solution of (14) of the form

$$\langle \mathbf{E}(\mathbf{x}) \rangle = \mathbf{A}e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (95)$$

We first compute $L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle$ and obtain

$$L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle = \{k_0^2[\mu_1(\mathbf{x}') + \epsilon_1(\mathbf{x}')]\mathbf{A} - i[\nabla'\mu_1(\mathbf{x}')\cdot\mathbf{k}]\mathbf{A} + i[\nabla'\mu_1(\mathbf{x}')\cdot\mathbf{A}]\mathbf{k}\}e^{i\mathbf{k}\cdot\mathbf{x}'}. \quad (96)$$

Secondly we compute $L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')$ using (92) for $\mathbf{G}$. After some simplification we obtain

$$L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}') = k_0^2\{\mu_1(\mathbf{x}) + \epsilon_1(\mathbf{x})\}G_1\mathbf{I} - \frac{\partial G_1}{\partial r}[\nabla\mu_1(\mathbf{x})\cdot\hat{\mathbf{r}}]\mathbf{I} + \frac{G_2}{r}[\nabla\mu_1(\mathbf{x})\cdot\hat{\mathbf{r}}]\mathbf{I} + k_0^2\{\mu_1(\mathbf{x}) + \epsilon_1(\mathbf{x})\}G_2\hat{\mathbf{r}}\hat{\mathbf{r}} + \frac{\partial G_1}{\partial r}\nabla\mu_1(\mathbf{x})\hat{\mathbf{r}} - \frac{G_2}{r}\nabla\mu_1(\mathbf{x})\hat{\mathbf{r}}. \quad (97)$$

The third step is to multiply (96) by (97) to form $L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle$. This leads to a rather long product which we shall not write out. The fourth step is to take the expectation value of this product which yields $\langle L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle \rangle$. In evaluating this expectation we find that three correlation functions occur. We shall assume that each depends only upon $r = |\mathbf{x} - \mathbf{x}'|$ and write for them $R_{\mu\mu}(r) = \langle \mu_1(\mathbf{x})\mu_1(\mathbf{x}') \rangle$, $R_{\mu\epsilon}(r) = \langle \mu_1(\mathbf{x})\epsilon_1(\mathbf{x}') \rangle$, and $R_{\epsilon\epsilon}(r) = \langle \epsilon_1(\mathbf{x})\epsilon_1(\mathbf{x}') \rangle$. Then after some simplification we can write the result in the form

$$\langle L_1(\mathbf{x})\mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle \rangle = [F\mathbf{A} + C(\hat{\mathbf{r}}\cdot\mathbf{k})\mathbf{A} - C(\hat{\mathbf{r}}\cdot\mathbf{A})\mathbf{k} + B(\hat{\mathbf{r}}\cdot\mathbf{A})\hat{\mathbf{r}}]e^{i\mathbf{k}\cdot\mathbf{x}'}. \quad (98)$$

The scalars F, B, and C introduced in (98) do not depend upon $\hat{\mathbf{r}}$ and are defined by

$$F(r) = k_0^4G_1[R_{\mu\mu} + 2R_{\mu\epsilon} + R_{\epsilon\epsilon}] - k_0^2(G_1' - r^{-1}G_2)(R_{\mu\mu}' + R_{\mu\epsilon}'), \quad (99)$$

$$B(r) = k_0^4G_2[R_{\mu\mu} + 2R_{\mu\epsilon} + R_{\epsilon\epsilon}] + k_0^2(G_1' - r^{-1}G_2)(R_{\mu\mu}' + R_{\mu\epsilon}'), \quad (100)$$

$$C(r) = ik_0^2G_1(R_{\mu\mu}' + R_{\mu\epsilon}') - i(G_1' - r^{-1}G_2)R_{\mu\mu}''. \quad (101)$$

In (99)–(101), primes denote differentiation with respect to r.

Finally we integrate (98) with respect to $\mathbf{x}'$. As before it is convenient to set $\mathbf{x}' = \mathbf{x} - \mathbf{r}$ and to integrate with respect to $\mathbf{r}$. In doing so we write $d\mathbf{r} = dr\,dS$ where dS denotes the area element on a sphere of radius r centered at $\mathbf{x}$. The integrations with respect to dS can be done in terms of the integral (53) and expressed in terms of $f(r)$ defined by (54) (see Appendix I). The result may be written as

$$\left\langle L_1(\mathbf{x}) \int \mathbf{G}(\mathbf{x}, \mathbf{x}')L_1(\mathbf{x}')\langle \mathbf{E}(\mathbf{x}') \rangle\, d\mathbf{x}' \right\rangle = [D\mathbf{A} + M(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}}]e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (102)$$

where $\hat{\mathbf{k}}$ is a unit vector in the direction of $\mathbf{k}$, while D and M are constants defined by

$$D = +\int_0^\infty \left[Ff + ikCr^{-1}\frac{\partial f}{\partial k} - Br^{-2}k^{-1}\frac{\partial F}{\partial k}\right]dr, \quad (103)$$

$$M = -\int_0^\infty \left[ikCr^{-1}\frac{\partial f}{\partial k} + Br^{-2}\left(\frac{\partial^2 f}{\partial k^2} - k^{-1}\frac{\partial f}{\partial k}\right)\right]dr. \quad (104)$$

To write (14) we must still calculate $L(\mathbf{x})\langle \mathbf{E}(\mathbf{x}) \rangle$ and $\langle L_2(\mathbf{x}) \rangle\langle \mathbf{E}(\mathbf{x}) \rangle$. A straightforward calculation yields

$$L(\mathbf{x})\langle \mathbf{E}(\mathbf{x}) \rangle = [(k^2 - k_0^2)\mathbf{A} - k^2(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}}]e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (105)$$

To compute $\langle L_2(\mathbf{x}) \rangle$ we find from (90) that

$$\langle L_2(\mathbf{x}) \rangle = k_0^2\langle \mu_1\epsilon_1 \rangle - \langle \mu_1\nabla\mu_1 \rangle \times \nabla \times. \quad (106)$$

But $\langle \mu_1\nabla\mu_1 \rangle = \frac{1}{2}\langle \nabla\mu_1^2 \rangle = \frac{1}{2}\nabla\langle \mu_1^2 \rangle = \frac{1}{2}\nabla R_{\mu\mu}(0) = 0$ and $\langle \mu_1\epsilon_1 \rangle = R_{\mu\epsilon}(0)$, so

$$\langle L_2(\mathbf{x}) \rangle = k_0^2R_{\mu\epsilon}(0). \quad (107)$$

Upon assembling the results (102), (105), and (106) we can write (14) in the form

$$[(k^2 - k_0^2)\mathbf{A} - k^2(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}} - \eta^2D\mathbf{A} - \eta^2M(\hat{\mathbf{k}}\cdot\mathbf{A})\hat{\mathbf{k}} - \eta^2k_0^2R_{\mu\epsilon}(0)\mathbf{A}]e^{i\mathbf{k}\cdot\mathbf{x}} = 0. \quad (108)$$

If we denote by $A_\parallel$ and $A_\perp$ the components of $\mathbf{A}$ respectively parallel and perpendicular to $\mathbf{k}$ or $\hat{\mathbf{k}}$, we can separate (108) into the two equations

$$A_\parallel\{k_0^2 + \eta^2[k_0^2R_{\mu\epsilon}(0) + D + M]\}e^{i\mathbf{k}\cdot\mathbf{x}} = 0, \quad (109)$$

$$A_\perp\{k^2 - k_0^2 - \eta^2[k_0^2R_{\mu\epsilon}(0) + D]\}e^{i\mathbf{k}\cdot\mathbf{x}} = 0. \quad (110)$$

These equations indicate that there may be solutions in which either $A_\parallel \neq 0$ or $A_\perp \neq 0$ while the other is zero. In the first case k must satisfy the equation

$$k_0^2 + \eta^2[k_0^2R_{\mu\epsilon}(0) + D + M] = 0. \quad (111)$$

In (111) k occurs in D and M. When $\eta = 0$, (111)

implies that $k_0 = 0$ which means that $\omega = 0$. Thus without random inhomogeneities, only static plane waves can be longitudinal. When $\eta \neq 0$, (111) may have solutions for nonzero frequencies, as we shall see when we consider an example in the next section.

In the second case, $A_\perp \neq 0$, k must satisfy the equation

$$k^2 - k_0^2 - \eta^2[k_0^2 R_{\mu\epsilon}(0) + D] = 0. \tag{112}$$

When $\eta = 0$ then $k = k_0$, while when η is small we can solve (112) to order η^2 by setting $k = k_0$ in $D(k)$. Thus

$$k^2 = k_0^2 + \eta^2[k_0^2 R_{\mu\epsilon}(0) + D(k_0)]. \tag{113}$$

To the same order we may then write

$$k = k_0 + \tfrac{1}{2}\eta^2[k_0 R_{\mu\epsilon}(0) + k_0^{-1} D(k_0)]. \tag{114}$$

When $\sigma = 0$ then k_0 is real and from (114) the attenuation coefficient Im k is given by

$$\text{Im } k = (\eta^2/2k_0)\ \text{Im } D(k_0). \tag{115}$$

Equations (111), (112), and their consequences are the main results of this section.

When $\sigma \neq 0$ the correlation functions $R_{\epsilon\epsilon}$ and $R_{\mu\epsilon}$ may be written in terms of the correlation functions involving ϵ_1', σ_1, and μ_1. There are nine such functions which we shall denote by $R_{\epsilon'\epsilon'}$, $R_{\epsilon'\sigma}$, etc. Upon using the definition of ϵ_1 in (86) we have

$$R_{\mu\epsilon} = \langle \mu_1(\mathbf{x})\epsilon_1(\mathbf{x}')\rangle = \frac{\epsilon_0'}{\epsilon_0' + i\omega^{-1}\sigma_0} R_{\mu\epsilon'} + \frac{i\omega^{-1}\sigma_0}{\epsilon_0' + i\omega^{-1}\sigma_0} R_{\mu\sigma}, \tag{116}$$

$$R_{\epsilon\epsilon} = \langle \epsilon_1(\mathbf{x})\epsilon_1(\mathbf{x}')\rangle = \left(\frac{\epsilon_0'}{\epsilon_0' + i\omega^{-1}\sigma_0}\right)^2 R_{\epsilon'\epsilon'} + \frac{2i\omega^{-1}\epsilon_0'\sigma_0}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2} R_{\epsilon'\sigma} - \frac{\omega^{-2}\sigma_0^2}{(\epsilon_0' + i\omega'\sigma_0)^2} R_{\sigma\sigma}. \tag{117}$$

7. SPECIAL ELECTROMAGNETIC MEDIA AND AN EXAMPLE

The preceding results can be simplified somewhat in various special cases. First of all let us suppose that all correlation functions involving ϵ' and σ are zero. This is the case, in particular, if only μ is random. Then $R_{\epsilon\epsilon} = R_{\mu\epsilon} = 0$ and (99)–(101) for F, B, and C simplify a little, while D and M are still given by (103) and (104).

A more useful case is that in which all correlation functions involving μ and σ vanish. Then $R_{\epsilon\epsilon} = [\epsilon_0'/(\epsilon_0' + i\omega^{-1}\sigma_0)]R_{\epsilon'\epsilon'}$, while all other correlation functions are zero. From (99)–(101), it follows that $C = 0$, $F = k_0^4 G_1 R_{\epsilon\epsilon}$, and $B = k_0^4 G_2 R_{\epsilon\epsilon}$. Then (103) and (104) become, when (54) is used for $f(r)$,

$$D = +\frac{4\pi k_0^4(\epsilon_0')^2}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2}\int_0^\infty R_{\epsilon'\epsilon'}(r)\Big[G_1(r)^{-1}kr \times \sin kr - G_2(r)k^{-1}r^{-1}\frac{\partial}{\partial k}\left(\frac{1}{k}\sin kr\right)\Big]dr, \tag{118}$$

$$M = -\frac{4\pi k_0^4(\epsilon_0')^2}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2}\int_0^\infty G_2(r)R_{\epsilon'\epsilon'}(r)r^{-1} \times \left[\frac{\partial^2}{\partial k^2}\left(\frac{1}{k}\sin kr\right) - k^{-1}\frac{\partial}{\partial k}\left(\frac{1}{k}\sin kr\right)\right]dr. \tag{119}$$

The propagation constant k for longitudinal waves is determined by (111) with the above values of D and M, while for transverse waves it is determined by (112) in which only D occurs.

Let us now consider the special case in which

$$R_{\epsilon'\epsilon'}(r) = \langle(\epsilon_1')^2\rangle e^{-a^{-1}r}. \tag{120}$$

Then upon using the results of Appendix II, we obtain for the equation for the propagation constant of transverse waves the result

$$\frac{k^2}{k_0^2} = 1 + \frac{\eta^2}{2}\frac{(\epsilon_0')^2\langle(\epsilon_1')^2\rangle}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2} \times \left[\left\{\frac{2}{3} + \frac{(1 + ik_0a)}{(ka)^2} + \frac{2(k_0a)^2}{(ka)^2 + (1 - ik_0a)^2}\right\} - \frac{1}{ka}\left\{1 + \frac{1 + (k_0a)^2}{(ka)^2}\right\}\cot^{-1}\frac{1 - ik_0a}{ka}\right]. \tag{121}$$

When k_0a is finite the solution of (121) up to terms in η^2 is given by (112). It is

$$\frac{k^2}{k_0^2} = 1 + \frac{\eta^2}{2}\frac{(\epsilon_0')^2\langle(\epsilon_1')^2\rangle}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2} \times \left[\left\{\frac{2}{3} + \frac{(1 + ik_0a)}{(k_0a)^2} + \frac{2(k_0a)^2}{(k_0a)^2 + (1 - ik_0a)^2}\right\} - \frac{1}{k_0a}\left\{1 + \frac{1 + (k_0a)^2}{(k_0a)^2}\right\}\cot^{-1}\frac{1 - ik_0a}{k_0a}\right]. \tag{122}$$

The above result is an explicit expression for k^2/k_0^2, where k is the transverse propagation constant. Equation (122) becomes infinite when k_0a becomes infinite, which shows that (121) must be solved more accurately. When k_0a is sufficiently large $[k_0a \gg 1/\eta\langle(\epsilon_1')^2\rangle^{\frac{1}{2}}]$, we find from (121) that the transverse propagation constant is

$$k = k_0[1 + \tfrac{1}{2}\eta\langle(\epsilon_1')^2\rangle^{\frac{1}{2}} + i/2k_0a] + O(\eta^2). \tag{123}$$

Thus the transverse attenuation coefficient is proportionate to $1/2a$ but independent of η when the correlation length is very long compared to

the wavelength, while Re $(k - k_0)$ is proportional to η. On the other hand, both Im (k) and Re $(k - k_0)$ are proportional to η^2 when the correlation length is comparable to the wavelength or smaller. The solution of (121) for $k_0 a$ finite describes the transition from one form to the other.

The equation for the propagation constant of longitudinal waves is given by

$$1 = \eta^2 \frac{(\epsilon_0')^2((\epsilon_1')^2)}{(\epsilon_0' + i\omega^{-1}\sigma_0)^2}\left[\left\{\frac{2}{3} + \frac{(1 + ik_0a)}{k^2a^2}\right\} - \left\{\frac{1}{ka} + \frac{(1 + k_0^2a^2)}{k^3a^3}\right\}\cot^{-1}\frac{(1 - ik_0a)}{ka}\right]. \quad (124)$$

This equation does not have solutions for which $|ka|$ is either large or small. We have not investigated it for intermediate values.

APPENDIX I. MEAN-VALUE THEOREMS

The basic mean-value theorem which we shall use

$$\int e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = \frac{4\pi r}{k}\sin kr \equiv f(r, k), \quad (1)$$

where $\mathbf{k}$ is the propagation vector, $\mathbf{r}$ is the radius vector, and the surface integration ds is performed on a sphere of radius r. Let $\mathbf{k} = k\hat{\mathbf{k}}$ and $\mathbf{r} = r\hat{\mathbf{r}}$, where $\hat{\mathbf{k}}$ and $\hat{\mathbf{r}}$ are unit propagation and radial vectors, respectively. It can be shown that

$$\nabla_{\mathbf{k}} = \hat{\mathbf{k}}\frac{\partial}{\partial \mathbf{k}}, \quad (2)$$

$$\nabla_{\mathbf{k}}\hat{\mathbf{k}} = \frac{1}{k}(\mathbf{I} - \hat{\mathbf{k}}\hat{\mathbf{k}}), \quad (3)$$

where $\nabla_{\mathbf{k}}$ denotes the gradient with respect to $\mathbf{k}$ and $\mathbf{I}$ is the unit dyadic. By differentiating (1) and using (2) and (3), the following new mean-value theorems can be established easily:

$$\int \hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = \frac{i\hat{\mathbf{k}}}{r}\frac{\partial f}{\partial k}, \quad (4)$$

$$\int \hat{\mathbf{r}}\hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{1}{r^2}\left[\hat{\mathbf{k}}\hat{\mathbf{k}}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \frac{1}{k}\frac{\partial f}{\partial k}\mathbf{I}\right], \quad (5)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = \frac{ik}{r}\frac{\partial f}{\partial k}, \quad (6)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{A})e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = \frac{i}{r}(\hat{\mathbf{k}}\cdot\mathbf{A})\frac{\partial f}{\partial k}, \quad (7)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{A})\hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{1}{r^2}\left[\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A})\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \frac{1}{k}\frac{\partial f}{\partial k}\mathbf{A}\right], \quad (8)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{k\hat{\mathbf{k}}}{r^2}\frac{\partial^2 f}{\partial k^2}, \quad (9)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{A})(\hat{\mathbf{r}}\cdot\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{k(\hat{\mathbf{k}}\cdot\mathbf{A})}{r^2}\frac{\partial^2 f}{\partial k^2}, \quad (10)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{A})\hat{\mathbf{r}}\hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{i}{kr^3}\left[\hat{\mathbf{k}}\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A})k\frac{\partial}{\partial k}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) - 2\hat{\mathbf{k}}\hat{\mathbf{k}}(\mathbf{A}\cdot\hat{\mathbf{k}})\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \mathbf{I}(\hat{\mathbf{k}}\cdot\mathbf{A})\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \mathbf{A}\hat{\mathbf{k}}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \hat{\mathbf{k}}\mathbf{A}\frac{\partial}{\partial k}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right)\right], \quad (11)$$

$$\int (\hat{\mathbf{r}}\cdot\mathbf{A})(\hat{\mathbf{r}}\cdot\mathbf{k})\hat{\mathbf{r}}e^{-i\mathbf{k}\cdot\mathbf{r}}\, ds = -\frac{ik}{r^3}\left[\hat{\mathbf{k}}(\hat{\mathbf{k}}\cdot\mathbf{A})\frac{\partial}{\partial k}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right) + \mathbf{A}\frac{1}{k}\frac{\partial}{\partial k}\left(\frac{\partial^2 f}{\partial k^2} - \frac{1}{k}\frac{\partial f}{\partial k}\right)\right]. \quad (12)$$

APPENDIX II. USEFUL INTEGRAL RELATIONS

In working out the examples, the following results have been found useful:

$$\int_0^\infty e^{-\alpha x}\sin kx\, dx = \frac{k}{k^2 + \alpha^2},$$

$$\int_0^\infty e^{-\alpha x}\frac{\sin kx}{x}\, dx = \cot^{-1}\frac{\alpha}{k},$$

$$\int_0^\infty (e^{-\alpha x} - e^{-\beta x})\frac{\sin kx}{x^2}\, dx = k\left[\frac{\beta}{k}\cot^{-1}\frac{\beta}{k} - \frac{\alpha}{k}\cot^{-1}\frac{\alpha}{k} + \tfrac{1}{2}\log\frac{k^2 + \beta^2}{k^2 + \alpha^2}\right],$$

$$\int_0^\infty e^{-\alpha x}\frac{1}{x}\left(\frac{\sin kx}{kx} - \cos kx\right)dx = 1 - \frac{\alpha}{k}\cot^{-1}\frac{\alpha}{k},$$

$$\int_0^\infty e^{-\alpha x}\frac{1}{x^2}\left(\frac{\sin kx}{kx} - \cos kx\right)dx = -\frac{k}{2}\left[\frac{\alpha}{k} - \left(1 + \frac{\alpha^2}{k^2}\right)\cot^{-1}\frac{\alpha}{k}\right],$$

$$\int_0^\infty (e^{-\alpha x} - e^{-\beta x})\frac{1}{x^3}\left(\frac{\sin kx}{kx} - \cos kx\right)dx = \frac{k^2}{2}\left[-\frac{1}{3k^2}(\beta^2 - \alpha^2) + \frac{\beta}{k}\cot^{-1}\frac{\beta}{k} - \frac{\alpha}{k}\cot^{-1}\frac{\beta}{k} + \frac{1}{3}\left(\frac{\beta}{k}\right)^3\cot^{-1}\frac{\beta}{k} - \frac{1}{3}\left(\frac{\alpha}{k}\right)^3\cot^{-1}\frac{\alpha}{k} + \frac{1}{3}\log\frac{k^2 + \beta^2}{k^2 + \alpha^2}\right],$$

$$\int_0^\infty e^{-\alpha x}\frac{1}{x^3}\left\{-kx\sin kx + 3\left(\frac{\sin kx}{kx} - \cos kx\right)\right\}dx = \frac{k^2}{2}\left[\frac{1}{6} + \left(\frac{\alpha}{k}\right)^2 - \frac{\alpha}{k}\cot^{-1}\frac{\alpha}{k} - \left(\frac{\alpha}{k}\right)^3\cot^{-1}\frac{\alpha}{k}\right].$$

A Theorem on the Conductivity of a Composite Medium*

JOSEPH B. KELLER
Courant Institute of Mathematical Sciences, New York University, New York, N. Y.

A composite medium consisting of a rectangular lattice of identical parallel cylinders of arbitrary cross section is considered. The cylinders have conductivity σ_2 and are imbedded in a medium of conductivity σ_1. Simple properties of the conductivity tensor of the composite medium are deduced from the theory of harmonic functions.

LET us consider a composite medium consisting of a rectangular lattice of identical parallel cylinders of any cross section, having electrical conductivity σ_2 imbedded in a medium of conductivity σ_1. Let the x and y axes lie along axes of the lattice and let $2X$ and $2Y$ be the lattice spacings in the x and y directions. When a static electric field of average field strength E_x is applied to the medium parallel to the x axis, the resulting current density will be a periodic function of x and y. If this current density is averaged with respect to y, the average current density j_x must be in the x direction and must be independent of x by conservation of charge. In terms of j_x and E_x we define the effective conductivity $\Sigma_x(\sigma_1, \sigma_2)$ in the x direction to be $\Sigma_x(\sigma_1, \sigma_2) = j_x/E_x$. The first argument of Σ_x denotes the conductivity of the medium surrounding the cylinders and the second argument denotes that of the medium constituting the cylinders. In a similar way we define Σ_y. The object of this note is to prove the following theorem:

Theorem. Let a medium contain a rectangular lattice of identical parallel cylinders, each of which is symmetric in the x and y axes, which are the lattice axes. Then the effective conductivities Σ_x and Σ_y of the composite medium in the x and y directions are related by

$$\sigma_1/\Sigma_x(\sigma_1, \sigma_2) = \Sigma_y(\sigma_2, \sigma_1)/\sigma_2. \tag{1}$$

The first argument of Σ_x or Σ_y denotes the conductivity of the medium surrounding the cylinders and the second argument that of the medium constituting the cylinders. The theorem also applies to thermal and other conductivities when the corresponding potentials are harmonic functions.

An important corollary of this theorem follows when the lattice is square and when each cylinder is symmetric in the line $x = y$. In this case the x and y directions are equivalent so $\Sigma_x(\sigma_1, \sigma_2) = \Sigma_y(\sigma_1, \sigma_2)$. Then we may omit the subscript and write $\Sigma = \Sigma_x = \Sigma_y$. Therefore the theorem yields the corollary

Corollary 1. When the lattice is square and each cylinder is symmetric in the line $x = y$,

$$\sigma_1/\Sigma(\sigma_1, \sigma_2) = \Sigma(\sigma_2, \sigma_1)/\sigma_2. \tag{2}$$

Another corollary of the theorem results from the fact that Σ_x and Σ_y have the dimensions of conductivity, so they are homogeneous of degree one in σ_1 and σ_2. Therefore $\Sigma_x(\sigma_1, \sigma_2)/\sigma_1 = \Sigma_x(1, \sigma_2/\sigma_1)$ and $\Sigma_y(\sigma_2, \sigma_1)/\sigma_2 = \Sigma_y(1, \sigma_1/\sigma_2)$. Now the theorem yields the further corollary

Corollary 2.

$$1/\Sigma_x(1, \sigma_2/\sigma_1) = \Sigma_y(1, \sigma_1/\sigma_2). \tag{3}$$

When the hypothesis of Corollary 1 is satisfied, Corollary 2 yields

$$1/\Sigma(1, \sigma_2/\sigma_1) = \Sigma(1, \sigma_1/\sigma_2). \tag{4}$$

The special cases of (3) and (4) in which $\sigma_2/\sigma_1 = \infty$ and $\sigma_1/\sigma_2 = 0$ were proved previously.[1] It is interesting to note that the approximate expression for Σ derived by Rayleigh[2] for a square lattice of circular cylinders satisfies (4).

The results (2) and (4) also apply to the average conductivity of a statistically homogeneous isotropic random distribution of cylinders of one medium in another medium. This can be proved by appropriately adapting the following proof.

To prove the theorem we first consider the harmonic function $\varphi(x, y)$ which is the potential corresponding to the applied field of average strength unity in the x direction. By symmetry, the lines $x = 0$ and $x = X$ are equipotential lines while $y = 0$ and $y = Y$ are field lines. If we let

* The research reported in this paper was sponsored by the National Science Foundation under Grant No. NSF G 19671.

[1] J. B. Keller, J. Appl. Phys. **34**, 991 (1963).
[2] Lord Rayleigh, Phil. Mag. Ser. 5 **34**, 481 (1892).

$\varphi(0, y) = 0$ then $\varphi(X, y) = X$ while $\varphi_y(x, 0) = \varphi_y(x, Y) = 0$. On the interface S between the two media, continuity of potential and of the normal component of current yield

$$\varphi^+ = \varphi^- \quad \text{on} \quad S, \tag{5}$$

$$\sigma_1 \, \partial\varphi^+/\partial n = \sigma_2 \, \partial\varphi^-/\partial n \quad \text{on} \quad S. \tag{6}$$

Here φ^+ and φ^- denote the values of φ outside S and inside S, respectively, while $\partial/\partial n$ denotes the derivative along the outward normal to S and σ_1 is the conductivity outside S. The average current across the line $x = X$ is

$$j_x = Y^{-1} \int_0^Y \sigma_1 \varphi_x^+(X, y) \, dy$$

$$= \sigma_1 Y^{-1} \int_0^Y \psi_y(X, y) \, dy = \sigma_1 Y^{-1} \psi_0. \tag{7}$$

Here ψ is the harmonic function conjugate to φ which has the value zero on the field line $y = 0$ and the unknown value ψ_0 on the field line $y = Y$.

From (5) it follows that $\partial\psi^+/\partial n = \partial\psi^-/\partial n$ on S and from (6) that $\sigma_1 \partial\psi^+/\partial s = \sigma_2 \partial\psi^-/\partial s$ on S where $\partial/\partial s$ denotes differentiation along S. Upon integrating the last relation and noting that $\psi^+ = \psi^-$ at $y = 0$, it follows that $\sigma_1\psi^+ = \sigma_2\psi^-$. We also note that $\psi_x(0, y) = \psi_x(X, y) = 0$. Therefore we define Φ by the relations

$$\Phi^+ = Y\psi^+/\psi_0, \tag{8}$$

$$\Phi^- = \sigma_2 Y\psi^-/\sigma_1\psi_0. \tag{9}$$

Then Φ is a harmonic function satisfying the conditions $\Phi_x(0, y) = \Phi_x(X, t) = 0$; $\Phi(x, 0) = 0$, $\Phi(x, Y) = Y$, $\Phi^+ = \Phi^-$ on S, and $\sigma_2 \partial\Phi^+/\partial n = \sigma_1 \partial\Phi^-/\partial n$ on S. Thus Φ is the potential corresponding to an applied field of average strength unity in the y direction when the conductivity outside S is σ_2 and that inside S is σ_1. The average current in the y direction is then

$$j_y = X^{-1} \int_0^X \sigma_2 \Phi_y(x, Y) \, dx = \sigma_2 X^{-1} Y \psi_0^{-1}$$

$$\times \int_0^X \psi_y(x, Y) \, dx = \sigma_2 X^{-1} Y \psi_0^{-1}$$

$$\times \int_0^X \varphi_x(x, Y) \, dx = \sigma_2 Y \psi_0^{-1}. \tag{10}$$

Since $E_x = E_y = 1$, it follows that $\Sigma_x = j_x$ and $\Sigma_y = j_y$. Then from (7) and (10) the result (1) of the theorem follows.

Scattering of Electromagnetic Waves by a Composite Cylinder*

MOHAMMED AZIZUL ISLAM
Advanced Communications Section, AVCO Corporation, Wilmington, Massachusetts
(Received 15 October 1963)

The scattering of an obliquely incident electromagnetic wave by a composite cylinder has been obtained using fundamental electromagnetic principles. The general result has been reduced to simpler forms for certain special cases.

1. INTRODUCTION

THE recent scientific interest in space exploration and guided missiles has presented many new and radically different communication and instrumentation problems. For example, one is interested to know the scattering cross section of a re-entry vehicle which is covered with a plasma sheath, or the kind of dielectric coating one should have in order that the scattering cross section may be appreciably reduced.

The problem of scattering of electromagnetic waves by obstacles has received the attention of various authors.[1-9] Many of them, however, treated the case of normally incident electromagnetic waves only. A consistent approach to the problem of scattering by cylindrical structures has been described by Wait.[10]

The problem to be considered in this paper is that of scattering of an obliquely incident plane electromagnetic wave to a composite dielectric cylinder. The approach is similar to that of Stratton[11] and Wait,[10] who treated the problem of a single dielectric cylinder.

2. STATEMENT OF THE PROBLEM AND FORMULATION

The geometry of the composite dielectric structure is as shown in Fig. 1. A plane electromagnetic wave is incident on a composite infinitely long cylinder making an angle θ with the axis of the cylinder.

Let $\mathbf{E}^i$ be the incident electric field, $\hat{e}$ the unit vector along the direction of $\mathbf{E}^i$, $\mathbf{k}_0$ the propagation vector, θ the angle of incidence; $\hat{e}$ will be assumed to be parallel to the plane $\phi = 0$, i.e., in the x–z plane. The time variation is assumed to be according to the factor $e^{i\omega t}$. Since the $\mathbf{H}$ field is assumed to be perpendicular to the axis of cylinder, the wave will be called TM wave. When the $\mathbf{E}$ field is perpendicular to the axis, the wave will be termed a TE wave.

The equation of a plane wave is

$$\mathbf{E}^i = \hat{e}E_0 e^{-i\mathbf{k}_0\cdot\mathbf{R}}.$$

In circular cylindrical coordinate system,

$$x = r\cos\phi, \qquad y = r\sin\phi, \qquad z = z.$$

Therefore,

$$\mathbf{E}^i = \hat{e}E_0 \exp\left[-jk_0 z\cos\theta + j\lambda_0 r\cos\phi\right], \tag{1}$$

where $\lambda_0 = k_0\sin\theta$.

Using the expansion[12] of $e^{iz\sin\theta}$,

$$e^{iz\sin\theta} = \sum_{n=-\infty}^{\infty} e^{in\theta}J_n(z), \tag{2}$$

one writes

$$e^{i\lambda_0 r\cos\phi} = e^{i\lambda_0 r\sin[\frac{1}{2}\pi-\phi]} = \sum_{n=-\infty}^{\infty} (j)^n e^{-in\phi}J_n(\lambda_0 r). \tag{3}$$

Substituting (3) in (1), one obtains

$$\mathbf{E}^i = \hat{e}E_0 \sum_{n=-\infty}^{\infty} (j)^n J_n(\lambda_0 r)F_n, \tag{4}$$

where $F_n = \exp\left[-jk_0 z\cos\theta - jn\phi\right]$.

Therefore, the z component of $\mathbf{E}^i$ will be

$$E_z^i = \mathbf{E}^i\cdot\hat{z} = E_0\sin\theta \sum_{n=-\infty}^{\infty} (j)^n J_n(\lambda_0 r)F_n. \tag{5}$$

Since the cylinder is of infinite length, there is no

* This research was supported by Air Force Ballistic Systems Division under contract AF04(694)239, and partially by the National Aeronautics and Space Administration, under Grant No. 355 to Northeastern University, Boston, Massachusetts.

[1] R. D. Kodis, J. Res. Natl. Bur. Std. **D65**, 19 (1961).
[2] C. Yeh and Z. A. Kaprielian, Can. J. Phys. **41**, 143 (1963).
[3] J. H. Van Vleck, F. Bloch, and M. Hamermesh, J. Appl. Phys. **18**, 274 (1947).
[4] S. H. Dike and D. D. King, Proc. IRE **40**, 853 (1952).
[5] E. S. Cassedy and J. Fainberg, IRE Trans. Antennas Propagation **8**, 1 (1960).
[6] C. T. Tai, J. Appl. Phys. **23**, 909 (1952).
[7] J. E. Storer and J. Sevick, J. Appl. Phys. **25**, 369 (1954).
[8] P. Ya. Ufimtsev, Rad. Eng. Electron. Phys. (USSR) **2**, 241 (1962).
[9] R. W. P. King and T. T. Wu, *The Scattering and Diffraction of Waves* (Harvard University Press, Cambridge, Massachusetts, 1959).
[10] J. R. Wait, *Electromagnetic Radiation from Cylindrical Structures* (Pergamon Press, Inc., New York, 1959).
[11] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), p. 524.
[12] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Part 1, Chap. 5, p. 620.

discontinuity along the z direction. As such, the z variation of all field components must be the same as that of the incident field, i.e., according to the factor $e^{-jk_0 z\cos\theta}$. The z component of the scattered field will then be

$$E_z^s = \sum_{n=-\infty}^{\infty} a_n^s H_n^{(2)}(\lambda_0 r) F_n. \tag{6}$$

a_n^s is a coefficient to be determined from suitable boundary condition, and $H_n^{(2)}(\lambda_0 r)$ is the Hankel function of the second kind, defined as $H_n^{(2)}(\lambda_0 r) = J_n(\lambda_0 r) - jN_n(\lambda_0 r)$, where $J_n(\lambda_0 r)$ and $N_n(\lambda_0 r)$ are two independent solutions of the Bessel equation. This function is chosen to ensure proper behavior of the wave at infinity. Henceforth, the superscript (2) from $H_n^{(2)}(\lambda_0 r)$ will be omitted.

Because of the assumed polarization, the z component of the $\mathbf{H}^i$ field is zero. The z component of the scattered field, however, will not in general be zero, except in the case of normally incident wave or of a perfectly conducting cylinder.[13] Thus, even though, the incident wave is purely TM, the resultant field have to be constructed as a superposition of a set TE and TM waves.

Therefore (the summation in the sequel runs from $n = -\infty$ to $+\infty$),

$$\begin{aligned} H_z^i &= 0, \\ H_z^s &= \sum b_n^s H_n(\lambda_0 r) F_n. \end{aligned} \tag{7}$$

It is now necessary to relate the other field components with E_z^i, E_z^s, H_z^i, and H_z^s. To do this, one starts with Maxwell's equations and obtains the following relationships:

$$\begin{aligned} E_\phi &= \frac{1}{k^2 - \alpha^2}\left[-j\frac{\alpha}{r}\frac{\partial E_z}{\partial \phi} + j\omega\mu\frac{\partial H_z}{\partial r}\right], \\ H_\phi &= \frac{1}{k^2 - \alpha^2}\left[-j\frac{\alpha}{r}\frac{\partial H_z}{\partial \phi} - j\omega\epsilon\frac{\partial E_z}{\partial r}\right]. \end{aligned} \tag{8}$$

In Eqs. (8), k, μ, ϵ are the properties of the medium and $\partial/\partial z = -j\alpha$. For the present case, the z dependence is as $e^{-jk_0 z\cos\theta}$. Then, $\alpha = k_0\cos\theta$. It is now easy to write the ϕ component of the incident and the scattered field.

Thus, from (5), (6), (7), and (8),

$$E_\phi^i = \sum \frac{-\alpha n E_0 \sin\theta (j)^n J_n(\lambda_0 r) F_n}{r\lambda_0^2}, \tag{9}$$

$$E_\phi^s = \sum \left[\frac{-\alpha n a_n^s}{r\lambda_0^2} H_n(\lambda_0 r) + \frac{j\omega\mu_0 b_n^s}{\lambda_0} H_n'(\lambda_0 r)\right] F_n, \tag{10}$$

[13] J. R. Wait, Can. J. Phys. 33, 189 (1955).

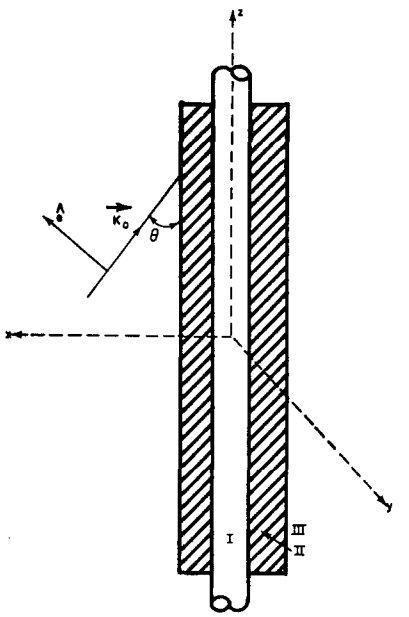

FIG. 1. Composite cylinder and the coordinate system. Region I: $0 \le r \le a$; μ_1, ϵ_1, σ_1, R_1. Region II: $a \le r \le b$; μ_2, ϵ_2, σ_2, R_2. Region III: $b \le r$; μ_0, ϵ_0, R_0.

$$H_\phi^i = \sum \frac{-jk_0^2 E_0 \sin\theta}{\omega\mu_0\lambda_0} (j)^n J_n'(\lambda_0 r) F_n, \tag{11}$$

$$H_\phi^s = \sum \left[\frac{-\alpha n}{r\lambda_0^2} b_n^s H_n(\lambda_0 r) - \frac{jk_0^2}{\omega\mu_0\lambda_0} a_n^s H_n'(\lambda_0 r)\right] F_n. \tag{12}$$

The prime sign with J_n and H_n indicates differentiation, and, as indicated previously, the summation runs over n from $-\infty$ to ∞. Thus (5), (6), (7), and (9)–(12) are the field representations in Region III. Since the radial components are not required for matching boundary conditions, these have not been set up. It is now easy to set up the field components in Regions II and I following the same principles and observing that in Region II both J_n's and H_n's are permissible solutions and in Region I, only the J_n's are permissible solutions. These are

Region II

$$E_z^{(2)} = \sum [b_n J_n(\lambda_2 r) + c_n H_n(\lambda_2 r)] F_n, \tag{13}$$

$$H_z^{(2)} = \sum [a_n J_n(\lambda_2 r) + e_n H_n(\lambda_2 r)] F_n, \tag{14}$$

$$E_\phi^{(2)} = \sum \left\{-\frac{\alpha n}{r\lambda_2^2}[b_n J_n(\lambda_2 r) + c_n H_n(\lambda_2 r)] + \frac{j\omega\mu_2}{\lambda_2}[a_n J_n'(\lambda_2 r) + e_n H_n'(\lambda_2 r)]\right\} F_n, \tag{15}$$

$$H_\phi^{(2)} = \sum \left\{-\frac{jk_2^2}{\omega\mu_2\lambda_2}[b_n J_n'(\lambda_2 r) + c_n H_n'(\lambda_2 r)] - \frac{\alpha n}{r\lambda_2^2}[a_n J_n(\lambda_2 r) + e_n H_n(\lambda_2 r)]\right\} F_n. \tag{16}$$

Region I

$$E_z^{(1)} = \sum d_n J_n(\lambda_1 r) F_n, \tag{17}$$

$$H_z^{(1)} = \sum f_n J_n(\lambda_1 r) F_n, \tag{18}$$

$$E_\phi^{(1)} = \sum \left[-\frac{\alpha n}{r\lambda_1^2} d_n J_n(\lambda_1 r) + \frac{j\omega\mu_1}{\lambda_1} f_n J_n'(\lambda_1 r) \right] F_n, \tag{19}$$

$$H_\phi^{(2)} = \sum \left[-\frac{jk_1^2}{\omega\mu_1\lambda_1} d_n J_n'(\lambda_1 r) - \frac{\alpha n}{r\lambda_1^2} f_n J_n(\lambda_1 r) \right] F_n. \tag{20}$$

All summations run over n from $-\infty$ to ∞, and the superscripts (1) and (2) indicate the regions in which they are valid.

As indicated previously,

$$\alpha = k_0 \cos\theta, \qquad \lambda_0 = k_0 \sin\theta,$$

$$\lambda_1 = (k_1^2 - k_0^2 \cos^2\theta)^{\frac{1}{2}}, \qquad \lambda_2 = (k_2^2 - k_0^2 \cos^2\theta)^{\frac{1}{2}},$$

$$F_n = \exp[-jn\phi - jk_0 z \cos\theta].$$

a_n, b_n, etc. are suitable coefficients to be determined from the boundary conditions. These are

at $r = b$

$$\begin{aligned} H_z^i + H_z^s &= H_z^{(2)}, \qquad H_\phi^i + H_\phi^s = H_\phi^{(2)}; \\ E_z^i + E_z^s &= E_z^{(2)}, \qquad E_\phi^i + E_\phi^s = E_\phi^{(2)}; \end{aligned} \tag{21}$$

at $r = a$

$$\begin{aligned} H_z^{(2)} &= H_z^{(1)}, \qquad H_\phi^{(2)} = H_\phi^{(1)}; \\ E_z^{(2)} &= E_z^{(1)}, \qquad E_\phi^{(2)} = E_\phi^{(1)}. \end{aligned} \tag{22}$$

3. DETERMINATION OF THE COEFFICIENTS

Since at $r = a$, $E_z^{(1)} = E_z^{(2)}$, one must have

$$\sum d_n J_n(\lambda_1 a) \exp[-jn\phi - jk_0 z \cos\theta] = \sum [b_n J_n(\lambda_2 a) + c_n H_n(\lambda_2 a)] \times \exp[-jn\phi - jk_0 z \cos\theta].$$

Multiplying both sides of above by $e^{jp\phi}$ and integrating with ϕ between the limits 0 to 2π, the above reduces to

$$d_n J_n(\lambda_1 a) = b_n J_n(\lambda_2 a) + c_n H_n(\lambda_2 a). \tag{23}$$

In a similar manner, the boundary condition (21) and (22) applied to the proper field expressions yield the following:

$$f_n J_n(\lambda_1 a) = a_n J_n(\lambda_2 a) + e_n H_n(\lambda_2 a), \tag{24}$$

$$\frac{\alpha n}{a\lambda_1^2} d_n J_n(\lambda_1 a) - \frac{j\omega\mu_1}{\lambda_1} f_n J_n'(\lambda_1 a) = \frac{\alpha n}{a\lambda_2^2} [b_n J_n(\lambda_2 a) + c_n H_n(\lambda_2 a)] + \frac{j\omega\mu_2}{\lambda_2} [a_n J_n'(\lambda_2 a) + e_n H_n'(\lambda_2 a)], \tag{25}$$

$$-\frac{jk_1^2}{\omega\mu_1\lambda_1} d_n J_n'(\lambda_1 a) - \frac{\alpha n}{a\lambda_1^2} f_n J_n(\lambda_1 a) = -\frac{jk_2^2}{\omega\mu_2\lambda_2} [b_n J_n'(\lambda_2 a) + c_n H_n'(\lambda_2 a)] - \frac{\alpha n}{a\lambda_2^2} [a_n J_n(\lambda_2 a) + e_n H_n(\lambda_2 a)], \tag{26}$$

$$E_0 \sin\theta (j)^n J_n(\lambda_0 b) + a_n^s H_n(\lambda_0 b) = b_n J_n(\lambda_2 b) + c_n H_n(\lambda_2 b), \tag{27}$$

$$b_n^s H_n(\lambda_0 b) = a_n J_n(\lambda_2 b) + e_n H_n(\lambda_2 b), \tag{28}$$

$$\frac{\alpha n}{b\lambda_0^2} [E_0 \sin\theta (j)^n J_n(\lambda_0 b) + a_n^s H_n(\lambda_0 b)] - \frac{j\omega\mu_0}{\lambda_0} b_n^s H_n'(\lambda_0 b) = \frac{\alpha n}{b\lambda_2^2} [b_n J_n(\lambda_2 b) + c_n H_n(\lambda_2 b)] - \frac{j\omega\mu_2}{\lambda_2} [a_n J_n'(\lambda_2 b) + e_n H_n'(\lambda_2 b)], \tag{29}$$

$$\frac{-jk_0^2}{\omega\mu_0\lambda_0} [E_0 \sin\theta (j)^n J_n'(\lambda_0 b) + a_n^s H_n'(\lambda_0 b)] - \frac{\alpha n}{b\lambda_0^2} b_n^s H_n(\lambda_0 b) = -\frac{jk_2^2}{\omega\mu_2\lambda_2} [b_n J_n'(\lambda_2 b) + c_n H_n'(\lambda_2 b)] - \frac{\alpha n}{b\lambda_2^2} [a_n J_n(\lambda_2 b) + e_n H_n(\lambda_2 b)]. \tag{30}$$

There are thus eight equations [(23)–(30)], and eight unknowns (a_n, b_n, c_n, d_n, e_n, f_n, a_n^s, and b_n^s).

By manipulating the equations and using the value of the Wronskian,

$$J_n(\lambda_0 b) H_n'(\lambda_0 b) - H_n(\lambda_0 b) J_n'(\lambda_0 b) = -(2j/\pi\lambda_0 b),$$

the above eight equations may be reduced to the following set:

$$\begin{aligned} A_n b_n + B_n c_n &= D_n a_n + E_n e_n, \\ H_n b_n + G_n c_n &= A_n a_n + B_n e_n, \\ L_n b_n + M_n c_n &= P_n a_n + Q_n e_n, \\ R_n b_n + S_n c_n &= L_n a_n + M_n e_n - T_n, \end{aligned} \tag{31}$$

where a_n, b_n, c_n and e_n are the unknowns, and the quantities A_n, B_n, etc. are defined as follows:

$$
\begin{aligned}
A_n &= \frac{\alpha n}{a} J_n(\lambda_2 a)\left[\frac{1}{\lambda_1^2} - \frac{1}{\lambda_2^2}\right], \\
B_n &= \frac{\alpha n}{a} H_n(\lambda_2 a)\left[\frac{1}{\lambda_1^2} - \frac{1}{\lambda_2^2}\right], \\
D_n &= -\frac{j\omega\mu_2}{\lambda_2} J_n'(\lambda_2 a) + \frac{j\omega\mu_1}{\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} J_n(\lambda_2 a), \\
E_n &= -\frac{j\omega\mu_2}{\lambda_2} H_n'(\lambda_2 a) + \frac{j\omega\mu_1}{\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} H_n(\lambda_2 a), \\
H_n &= -\frac{jk_1^2}{\omega\mu_1\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} J_n(\lambda_2 a) + \frac{jk_2^2}{\omega\mu_2\lambda_2} J_n'(\lambda_2 a), \\
G_n &= -\frac{jk_1^2}{\omega\mu_1\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} H_n(\lambda_2 a) + \frac{jk_2^2}{\omega\mu_2\lambda_2} H_n'(\lambda_2 a), \\
L_n &= \frac{\alpha n}{b} J_n(\lambda_2 b)\left[\frac{1}{\lambda_0^2} - \frac{1}{\lambda_2^2}\right], \\
M_n &= \frac{\alpha n}{b} H_n(\lambda_2 b)\left[\frac{1}{\lambda_0^2} - \frac{1}{\lambda_2^2}\right], \\
P_n &= -\frac{j\omega\mu_2}{\lambda_2} J_n'(\lambda_2 b) + \frac{j\omega\mu_0}{\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} J_n(\lambda_2 b), \\
Q_n &= -\frac{j\omega\mu_2}{\lambda_2} H_n'(\lambda_2 b) + \frac{j\omega\mu_0}{\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} H_n(\lambda_2 b), \\
R_n &= -\frac{jk_0^2}{\omega\mu_0\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} J_n(\lambda_2 b) + \frac{jk_2^2}{\omega\mu_2\lambda_2} J_n'(\lambda_2 b), \\
S_n &= -\frac{jk_0^2}{\omega\mu_0\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} H_n(\lambda_2 b) + \frac{jk_2^2}{\omega\mu_2\lambda_2} H_n'(\lambda_2 b), \\
T_n &= \frac{2k_0^2(j)^n}{\pi\omega\mu_0 b}\frac{1}{\lambda_0^2}\frac{E_0 \sin\theta}{H_n(\lambda_0 b)}.
\end{aligned}
\tag{32}
$$

The solutions for the unknowns may be now formally written down.

$$
\begin{aligned}
b_n &= \begin{vmatrix} 0 & B_n & -D_n & -E_n \\ 0 & G_n & -A_n & -B_n \\ 0 & M_n & -P_n & -Q_n \\ -T_n & S_n & -L_n & -M_n \end{vmatrix} \Delta^{-1}, \\
c_n &= \begin{vmatrix} A_n & 0 & -D_n & -E_n \\ H_n & 0 & -A_n & -B_n \\ L_n & 0 & -P_n & -Q_n \\ R_n & -T_n & -L_n & -M_n \end{vmatrix} \Delta^{-1}, \\
a_n &= \begin{vmatrix} A_n & B_n & 0 & -E_n \\ H_n & G_n & 0 & -B_n \\ L_n & M_n & 0 & -Q_n \\ R_n & S_n & -T_n & -M_n \end{vmatrix} \Delta^{-1},
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
e_n &= \begin{vmatrix} A_n & B_n & -D_n & 0 \\ H_n & G_n & -A_n & 0 \\ L_n & M_n & -P_n & 0 \\ R_n & S_n & -L_n & -T_n \end{vmatrix} \Delta^{-1}, \\
\Delta &= \begin{vmatrix} A_n & B_n & -D_n & -E_n \\ H_n & G_n & -A_n & -B_n \\ L_n & M_n & -P_n & -Q_n \\ R_n & S_n & -L_n & -M_n \end{vmatrix},
\end{aligned}
$$

and $|\ |$ represents the determinant. Once a_n, b_n, c_n, and e_n are known, the other unknowns may easily be found. These known coefficients may then be substituted in (5), (6), (7), and (9)–(20) to obtain the different field components. This then is the complete formal solution of the general problem. The result for the case of the magnetic vector in the $\phi = 0$ plane and the electric vector along the y direction may be obtained using the analogy between the magnetic and electric quantities, namely, by substituting in the above results H for E, $-E$ for H, and μ and ϵ interchanged throughout.

4. SPECIAL CASES

It is interesting to see how these complicated results behave in certain very special cases.

Case 1

If one lets $\mu_2 = \mu_0$, $\epsilon_2 = \epsilon_0$, the above problem should reduce to the problem of a single dielectric cylinder. With the above substitution, (32) becomes

$$
\begin{aligned}
A_n &= \frac{\alpha n}{a} J_n(\lambda_0 a)\left[\frac{1}{\lambda_1^2} - \frac{1}{\lambda_0^2}\right], \\
B_n &= \frac{\alpha n}{a} H_n(\lambda_0 a)\left[\frac{1}{\lambda_1^2} - \frac{1}{\lambda_0^2}\right], \\
D_n &= -\frac{j\omega\mu_0}{\lambda_0} J_n'(\lambda_0 a) + \frac{j\omega\mu_1}{\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} J_n(\lambda_0 a), \\
E_n &= -\frac{j\omega\mu_0}{\lambda_0} H_n'(\lambda_0 a) + \frac{j\omega\mu_1}{\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} H_n(\lambda_0 a), \\
F_n &= -\frac{jk_1^2}{\omega\mu_1\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} J_n(\lambda_0 a) + \frac{jk_0^2}{\omega\mu_0\lambda_0} J_n'(\lambda_0 a), \\
G_n &= -\frac{jk_1^2}{\omega\mu_1\lambda_1}\frac{J_n'(\lambda_1 a)}{J_n(\lambda_1 a)} H_n(\lambda_0 a) + \frac{jk_0^2}{\omega\mu_0\lambda_0} H_n'(\lambda_0 a), \\
L_n &= M_n = Q_n = S_n = 0, \\
P_n &= 2\omega\mu_0/\pi\lambda_0^2 b H_n(\lambda_0 b),
\end{aligned}
\tag{34}
$$

$$R_n = -2k_0^2/\pi\omega\mu_0\lambda_0^2 b H_n(\lambda_0 b),$$

$$T_n = \frac{2k_0^2(j)^n}{\pi\omega\mu_0\lambda_0^2 b}\frac{E_0 \sin\theta}{H_n(\lambda_0 b)}.$$

Therefore (31) reduces to

$$\begin{aligned} A_n b_n + B_n c_n - D_n a_n - E_n e_n &= 0, \\ H_n b_n + G_n c_n - A_n a_n - B_n e_n &= 0, \\ P_n a_n = 0, \qquad R_n b_n &= -T_n. \end{aligned} \tag{35}$$

These may be solved in the usual way. The results are

$$a_n = 0,$$

$$b_n = (j)^n E_0 \sin\theta,$$

$$b_n^s = e_n = \frac{k_0}{\omega\mu_0} E_0 \sin\theta (j)^n \times \left\{\frac{2}{\pi v^2}\left(\frac{1}{u^2} - \frac{1}{v^2}\right)\frac{n\cos\theta}{[H_n(v)]^2 D}\right\},$$

$$f_n = b_n^s[H_n(v)/J_n(u)], \tag{36}$$

$$a_n^s = c_n = (j)^n E_0 \sin\theta \times \left\{-\frac{J_n(v)}{H_n(v)} - 2j\frac{\left[\frac{H_n'(v)}{vH_n(v)} - \frac{K}{u}\frac{J_n'(u)}{J_n(u)}\right]}{\pi v^2[H_n(v)]^2 D}\right\},$$

$$d_n = [E_0 \sin\theta (j)^n J_n(v) + a_n^s H_n(v)][1/J_n(u)],$$

where

$$D = \left[\frac{H_n'(v)}{vH_n(v)} - \frac{K}{u}\frac{J_n'(u)}{J_n(u)}\right] \times \left[\frac{H_n'(v)}{vH_n(v)} - \frac{N^2}{uK}\frac{J_n'(u)}{J_n(u)}\right] - \left[\frac{1}{u^2} - \frac{1}{v^2}\right]^2 n^2 \cos^2\theta,$$

$$N = k_1^2/k_0^2, \qquad K = \mu_1/\mu_0,$$

$$u = \lambda_1 a, \qquad v = \lambda_0 a.$$

These results are exactly what one expected and agree with Wait's[13] results for a single dielectric cylinder. The corresponding results for normal incidence may easily be obtained by letting $\theta = 90°$ in (36). The back-scattering cross section for this infinite dielectric cylinder may be defined as

$$\sigma_B \text{ per unit length} = \lim_{r\to\infty} 2\pi r \left|\frac{E^s}{E^i}\right|^2 = \frac{P_{\text{omni}}^s}{S^i}, \tag{37}$$

where P_{omni}^s is the total power reradiated per unit length of an ideal omnidirectional scatterer that maintains the same field E^s at a radial distance r for all values of ϕ as that maintained by the actual scattering cylinder specifically in the direction toward the source ($\phi = 0$). Thus, for the infinite cylinder under consideration,

$$\sigma_B \text{ per unit length} = \lim_{\substack{r\to\infty \\ \theta=\frac{1}{2}\pi \\ \phi=0}} 2\pi r \times \left|\frac{\sum_{n=-\infty}^{\infty} a_n^s H_n(\lambda_0 r)\exp[-jn\phi - jk_0 z\cos\theta]}{E_0 \sin\theta \exp[-jk_0 z\cos\theta + jk_0 r\sin\theta\cos\phi]}\right|^2$$

$$= 2\pi r \left|\frac{\sum_{n=-\infty}^{\infty} a_n^s\left(\frac{2}{\pi k_0 r}\right)^{\frac{1}{2}} \exp\{-jk_0 r + j[\frac{1}{4}\pi + \frac{1}{2}(n\pi)]\}}{E_0 \exp[jk_0 r]}\right|^2,$$

where the asymptotic expression for $H_n(k_0 r)$ has been used. Since only the absolute value is of interest, the above reduces to

$$\sigma_B \text{ per unit length} = \frac{4}{E_0 k_0}\left|\sum_{n=-\infty}^{\infty} a_n^s\right|^2 = \frac{4}{E_0 k_0}\left|\sum_{n=0}^{\infty}\epsilon_n a_n^s\right|^2 = \frac{4}{k_0}\left|\sum_{n=0}^{\infty}\epsilon_n (j)^n W_n\right|^2, \tag{38}$$

where

$$\epsilon_n = 2 \quad \text{for} \quad n \neq 0$$
$$= 1 \quad \text{for} \quad n = 0,$$

and

$$W_n = a_n^s/[E_0(j)^n].$$

With $\theta = \frac{1}{2}\pi$ and $\mu_1 = \mu_0$ (assumed), from (36),

$$W_n = \frac{a_n^s}{E_0(j)^n} = -\frac{J_n(v)}{H_n(v)} - 2j\left\{\frac{1}{\pi v^2[H_n(v)]^2\left[\frac{H_n'(v)}{vH_n(v)} - \frac{N^2}{uK}\frac{J_n'(u)}{J_n(u)}\right]}\right\}$$

$$= \frac{-k_1 J_n'(k_1 a)J_n(k_0 a) + k_0 J_n(k_1 a)J_n'(k_0 a)}{k_1 J_n'(k_1 a)H_n(k_0 a) - k_0 J_n(k_1 a)H_n'(k_0 a)}. \tag{39}$$

This result agrees with King and Wu's result (c.f. Ref. 9, p. 69).

Case 2: Highly Conducting Cylinder with a Dielectric Coaxial Cylinder Outside

This is the most important case. Here in Region I, one assumes $\sigma_1/\omega\epsilon_1 \gg 1$. Therefore,

$$k_1^2 = -\mu_1(\partial/\partial t)\left(\epsilon_1\frac{\partial}{\partial t} + \sigma_1\right) \simeq -j\omega\mu_1\sigma_1,$$

$$k_1^2/k_0^2 \simeq -j\omega\mu_1\sigma_1/\omega^2\mu_0\epsilon_0 \to \text{very large},$$

$$\lambda_1^2 = k_1^2 - k_0^2\cos^2\theta \simeq k_1^2 \simeq -j\omega\mu_1\sigma_1.$$

With these restrictions, (32) reduces to the following:

$$\begin{aligned}
A_n &\simeq -(\alpha n/a)[J_n(\lambda_2 a)/\lambda_2^2],\\
B_n &\simeq -(\alpha n/a)[H_n(\lambda_2 a)/\lambda_2^2],\\
D_n &\simeq -(j\omega\mu_2/\lambda_2)J_n'(\lambda_2 a),\\
E_n &\simeq -(j\omega\mu_2/\lambda_2)H_n'(\lambda_2 a),\\
H_n &\simeq -(jk_1^2/\omega\mu_1\lambda_1)J_n(\lambda_2 a)[J_n'(\lambda_1 a)/J_n(\lambda_1 a)],\\
G_n &\simeq -(jk_1^2/\omega\mu_1\lambda_1)H_n(\lambda_2 a)[J_n'(\lambda_1 a)/J_n(\lambda_1 a)].
\end{aligned} \tag{40}$$

L_n, M_n, P_n, Q_n, R_n, S_n, and T_n are the same as in (32).

One may substitute (40) into (31) and solve for the unknowns. From (40) it is apparent that

$$\begin{aligned}
\frac{G_n}{H_n} &\simeq \frac{H_n(\lambda_2 a)}{J_n(\lambda_2 a)} \simeq \frac{B_n}{A_n},\\
A_n/H_n &\simeq B_n/H_n \simeq 0,\\
c_n &= -[J_n(\lambda_2 a)/H_n(\lambda_2 a)]b_n,\\
e_n &\simeq -[J_n'(\lambda_2 a)/H_n'(\lambda_2 a)]a_n.
\end{aligned}$$

Before solving for the unknowns, some comments regarding the boundary conditions at the surface of a perfect conductor are in order. At the surface of a perfect conductor, the tangential components of the electric field and the normal component of the **B** field must be zero. The normal component of **D** is zero inside the perfect conductor, but is not zero, in general, just outside the conductor, because of the presence of the surface charge density. Likewise, the tangential magnetic field inside a perfect conductor is zero, but is not zero just outside it. This discontinuity gives rise to the surface current density. Thus, the boundary conditions required to solve a boundary-value problem involving a perfect conductor surface are either

$$E_\phi = 0, \qquad E_z = 0 \quad \text{at the surface,} \tag{41}$$

or

$$\partial/\partial r(rH_\phi) = 0, \quad \partial/\partial r(H_z) = 0 \quad \text{at the surface.}$$

One has a choice now to continue with the solution of (31) with (40) or reformulate the problem with the new setup. If the formulation is correct, both must yield the same result. If instead of solving (31) with (40), one prefers to work directly with the perfect conductor and the outer dielectric cylinder, the procedure will be to start with (5)–(16) and apply boundary conditions (21) and the condition that at $r = a$,

$$E_z^{(2)} = E_\phi^{(2)} = 0, \tag{42}$$

This will yield the following set of relationships:

$$\begin{aligned}
&E_0 \sin\theta j^n J_0(\lambda_0 b) + a_n^s H_n(\lambda_0 b)\\
&\qquad = b_n J_n(\lambda_2 b) + c_n H_n(\lambda_2 b),\\
&b_n^s H_n(\lambda_0 b) = a_n J_n(\lambda_2 b) + e_n H_n(\lambda_2 b),\\
&-\frac{\alpha n}{b\lambda_0^2}[E_0 \sin\theta j^n J_n(\lambda_0 b) + a_n^s H_n(\lambda_0 b)] + \frac{j\omega\mu_0}{\lambda_0} b_n^s H_n'(\lambda_0 b)\\
&\qquad = -\frac{\alpha n}{b\lambda_2^2}[b_n J_n(\lambda_2 b) + c_n H_n(\lambda_2 b)]\\
&\qquad\quad + \frac{j\omega\mu_2}{\lambda_2}[a_n J_n'(\lambda_2 b) + e_n H_n'(\lambda_2 b)],\\
&-\frac{jk_0^2}{\omega\mu_0\lambda_0}[E_0 \sin\theta j^n J_n'(\lambda_0 b) + a_n^s H_n'(\lambda_0 b)] - \frac{\alpha n}{b\lambda_0^2} b_n^s H_n(\lambda_0 b)\\
&\qquad = -\frac{jk_2^2}{\omega\mu_2\lambda_2}[b_n J_n'(\lambda_2 b) + c_n H_n'(\lambda_2 b)]\\
&\qquad\quad - \frac{\alpha n}{b\lambda_2^2}[a_n J_n(\lambda_2 b) + e_n H_n(\lambda_2 b)],\\
&b_n J_n(\lambda_2 a) + c_n H_n(\lambda_2 a) = 0,\\
&a_n J_n'(\lambda_2 a) + e_n H_n'(\lambda_2 a) = 0.
\end{aligned} \tag{43}$$

Here again one has a set of six equations and six unknowns. As pointed out earlier, solving for the unknowns in (43) will be the same as in (31) with (40). By either or both methods, one obtains the following results:

$$\begin{aligned}
&b_n\left[\frac{1}{\lambda_0^2} - \frac{1}{\lambda_2^2}\right]\frac{\alpha n}{b} X\\
&\qquad = a_n\left[\frac{j\omega\mu_0}{\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} Y - \frac{j\omega\mu_2}{\lambda_2} W\right],
\end{aligned} \tag{44}$$

and

$$\begin{aligned}
&a_n\left[\frac{1}{\lambda_0^2} - \frac{1}{\lambda_2^2}\right]\frac{\alpha n}{b} Y = b_n\left[\frac{jk_2^2}{\omega\mu_2\lambda_2} V\right.\\
&\qquad \left. - \frac{jk_0^2}{\omega\mu_0\lambda_0}\frac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} X\right] + \frac{2k_0^2 j^n}{\pi\omega\mu_0\lambda_0^2 b}\frac{E_0 \sin\theta}{H_n(\lambda_0 b)}.
\end{aligned} \tag{45}$$

From (44) and (45), one obtains

$$a_n = \frac{\dfrac{2k_0^2 j^n E_0 \sin\theta}{\pi\omega\mu_0\lambda_0^2 b H_n(\lambda_0 b)}}{\left[\dfrac{1}{\lambda_0^2} - \dfrac{1}{\lambda_2^2}\right]\dfrac{\alpha n}{b} Y + \left[\dfrac{\kappa_0^2}{\omega\mu_0\lambda_0}\dfrac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} X - \dfrac{k_2^2}{\omega\mu_2\lambda_2} V\right]\left[\dfrac{\dfrac{\omega\mu_2}{\lambda_2} W - \dfrac{\omega\mu_0}{\lambda_0}\dfrac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)} Y}{\left[\dfrac{1}{\lambda_0^2} - \dfrac{1}{\lambda_2^2}\right]\dfrac{\alpha n}{b} X}\right]}. \tag{46}$$

In (44), (45), and (46) above, the functions X, Y, W, and V are defined as follows:

$$X = \frac{J_n(\lambda_2 b)H_n(\lambda_2 a) - J_n(\lambda_2 a)H_n(\lambda_2 b)}{H_n(\lambda_2 a)},$$

$$Y = \frac{J_n(\lambda_2 b)H_n'(\lambda_2 a) - J_n'(\lambda_2 a)H_n(\lambda_2 b)}{H_n'(\lambda_2 a)}, \tag{47}$$

$$W = \frac{J_n'(\lambda_2 b)H_n'(\lambda_2 a) - J_n'(\lambda_2 a)H_n'(\lambda_2 b)}{H_n'(\lambda_2 a)},$$

$$V = \frac{J_n'(\lambda_2 b)H_n(\lambda_2 a) - J_n(\lambda_2 a)H_n'(\lambda_2 b)}{H_n(\lambda_2 a)}.$$

Since a_n is now known from (46), b_n is obtained using (44) and (46), and hence c_n and e_n also are obtained from the last two relationships of (43). From the second equation of (43),

$$b_n^s = a_n[J_n(\lambda_2 b)/H_n(\lambda_0 b)] + e_n[H_n(\lambda_2 b)/H_n(\lambda_0 b)].$$

a_n and e_n both being known, b_n^s can be found. From the first equation of (43),

$$a_n^s = [1/H_n(\lambda_0 b)][b_n X - E_0 \sin\theta j^n J_n(\lambda_0 b)].$$

Thus

$$a_n^s = -\left\{\frac{E_0 \sin\theta j^n J_n(\lambda_0 b)}{H_n(\lambda_0 b)} + \frac{\dfrac{2jk_0^2 j^n E_0 \sin\theta}{\pi\omega\mu_0\lambda_0^2 b[H_n(\lambda_0 b)]^2}\left[\dfrac{\omega\mu_2}{\lambda_2}W - \dfrac{\omega\mu_0}{\lambda_0}\dfrac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)}Y\right]}{\left[\dfrac{1}{\lambda_0^2} - \dfrac{1}{\lambda_2^2}\right]^2\left(\dfrac{\alpha n}{b}\right)^2 Y + \left[\dfrac{k_0^2 H_n'(\lambda_0 b)}{\omega\mu_0\lambda_0 H_n(\lambda_0 b)} - \dfrac{k_2^2}{\omega\mu_2\lambda_2}\dfrac{V}{X}\right]\left[\dfrac{\omega\mu_2}{\lambda_2}W - \dfrac{\omega\mu_0}{\lambda_0}\dfrac{H_n'(\lambda_0 b)}{H_n(\lambda_0 b)}Y\right]}\right\}$$

$$\times\left[J_n(\lambda_0 b) + \frac{\dfrac{2j}{\pi b\omega\mu_0 H_n(\lambda_0 b)}}{\dfrac{k_0^2 H_n'(\lambda_0 b)}{\omega\mu_0\lambda_0 H_n(\lambda_0 b)} - \dfrac{k_2^2}{\omega\mu_2\lambda_2}\dfrac{V}{X}}\right].$$

Thus one obtains the complete solution of the problem. It is interesting to see how a_n^s behaves for the case of normal incidence. For this case,

$$\alpha = k_0 \cos\theta = 0,$$

$$\lambda_0 = k_0 \sin\theta = k_0,$$

$$\lambda_2 = [k_2^2 - k_0^2\cos^2\theta]^{\frac{1}{2}} = k_2.$$

Also assume $\mu_2 = \mu_0$.
Then

$$a_n^s = -\frac{E_0 j^n}{H_n(\lambda_0 b)}$$

Note:

$$\frac{2j}{\pi b} = -\lambda_0[J_n(\lambda_0 b)H_n'(\lambda_0 b) - J_n'(\lambda_0 b)H_n(\lambda_0 b)];$$

also

$$\frac{V}{X} = \frac{J_n'(\lambda_2 b)H_n(\lambda_2 a) - J_n(\lambda_2 a)H_n'(\lambda_2 b)}{J_n(\lambda_2 b)H_n(\lambda_2 a) - J_n(\lambda_2 a)H_n(\lambda_2 b)} = \frac{J_n(\lambda_2 a)N_n'(\lambda_2 b) - J_n'(\lambda_2 b)N_n(\lambda_2 a)}{J_n(\lambda_2 a)N_n(\lambda_2 b) - J_n(\lambda_2 b)N_n(\lambda_2 a)}.$$

Therefore,

$$a_n^s = -\frac{E_0 j^n}{H_n(\lambda_0 b)}\left[J_n(\lambda_0 b) + \frac{2j/\pi b[J_n(\lambda_2 b)N_n(\lambda_2 a) - J_n(\lambda_2 a)N_n(\lambda_2 b)]}{\{k_0 H_n'(\lambda_0 b)[J_n(\lambda_2 b)N_n(\lambda_2 a) - J_n(\lambda_2 a)N_n(\lambda_2 b)] - k_2 H_n(\lambda_0 b)[J_n'(\lambda_2 b)N_n(\lambda_2 a) - J_n(\lambda_2 a)N_n'(\lambda_2 b)]\}}\right]$$

$$= E_0 j^n\left\{\frac{k_2 J_n(k_0 b)[J_n(k_2 a)N_n'(k_2 b) - J_n'(k_2 b)N_n(k_2 a)] - k_0 J_n'(k_0 b)[J_n(k_2 a)N_n(k_2 b) - J_n(k_2 b)N_n(k_2 a)]}{-k_2 H_n(k_0 b)[J_n(k_2 a)N_n'(k_2 b) - J_n'(k_2 b)N_n(k_2 a)] + k_0 H_n'(k_0 b)[J_n(k_2 a)N_n(k_2 b) - J_n(k_2 b)N_n(k_2 a)]}\right\}. \tag{48}$$

Obviously, the coefficient of the back-scattering cross section $W_n = a_n^s/E_0 j^n$ = the quantity in { } above. This result agrees with Ref. 9, p. 69, Eq. (18.3).

The solution for the problem of a single conducting cylinder may simply be obtained by letting $a = b$, $\lambda_2 = \lambda_0$, $\mu_2 = \mu_0$ in the result for the above case.

For this case, from (47),

$$X = W = 0,$$
$$Y = -[2j/\pi\lambda_0 a H_n'(\lambda_0 a)], \tag{49}$$
$$V = 2j/\pi\lambda_0 a H_n(\lambda_0 a).$$

Then from (46) and (43),

$$a_n = e_n = 0. \tag{50}$$

From (45) and (49),

$$b_n = j^n E_0 \sin\theta. \tag{51}$$

Therefore,

$$c_n = -[J_n(\lambda_0 a)/H_n(\lambda_0 a)]j^n E_0 \sin\theta, \tag{52}$$

$$a_n^s = -[J_n(\lambda_0 a)/H_n(\lambda_0 a)]j^n E_0 \sin\theta = c_n. \tag{53}$$

From (43) and (50),

$$b_n^s = 0. \tag{54}$$

Looking back into the field components in the various regions, it is observed that for the case of a single conducting cylinder there is no z component of the H field. One may thus make an important observation that, for the case of a single conducting cylinder, if the incident wave is TE, only TE modes are excited. Similarly, it can be shown easily that, for this case, if the incident wave is TM, only TM modes are excited. However, for a dielectric cylinder with obliquely incident wave, a combination of TE and TM modes will be required to satisfy the boundary conditions even though the incident wave may be just TE or TM.

The results of various simplified cases may thus be obtained from the general formulation.

The author wishes to acknowledge his indebtedness to Professor Sheldon S. Sandler of Northeastern University, Boston, Massachusetts for his encouragement and guidance, and to C. R. Mullin of AVCO/RAD, Wilmington, Massachusetts for various illuminating discussions on this subject.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Application of Operational Methods to the Analysis of Uniform Plasmas

SAUL SILVEN*
McDonnell Aircraft Corporation, St. Louis, Missouri
(Received 16 July 1963)

Various aspects of plasma theory are derived by the application of operational methods to the equation of electron motion. Discussed herein are the complex tensor dielectric constant, the tensor permittivity and conductivity, plasma transient behavior, and plasma oscillations.

INTRODUCTION

THE plasma model discussed in this paper is considered to be composed of equal densities of free electrons and heavy positive ions distributed uniformly throughout the volume under discussion, and neutral particles whose population density is very large with respect to that of the charges. The density of particles is supposed to be large enough to present the appearance of a macroscopically continuous medium, electrically neutral at all points. The collision frequencies are considered to be independent of electron velocity.[1]

The dispersive effects of thermal motions and pressure gradients are neglected. These effects can be analyzed as indicated by Chandrasekhar[2] and Jackson.[3]

* Present address: Lockheed Missiles and Space Company, Sunnyvale, California.

[1] I. P. Shkarofsky, M. P. Bachynski, and T. W. Johnston, *Electromagnetic Effects of Re-entry* (Pergamon Press, Inc., New York, 1961), p. 24.

[2] S. Chandrasekhar, *Plasma Physics* (University of Chicago Press, Chicago, 1960), p. 143.

[3] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New York, 1962), pp. 335–340.

THE EQUATION OF MOTION

Consider an electron–ion pair originally at rest at point 1 (Fig. 1). If the effect of an applied force field is sensitive to charge, the electron–ion pair separates. If the ion mass is much greater than the electron mass, the electron moves to point 2 while the ion is considered to remain approximately at point 1. The symbol $\mathbf{r}$ thus represents electron displacement.

The forces on an electron may be due to the presence of an electric field (alternating or electrostatic), collisions with neutral particles (the average effect of which is a frictional force), and a static magnetic field. The electron velocities are assumed to be much less, numerically, than the speed of light, so that the effect of an alternating magnetic field is negligible in comparison to that of the associated electric field.

Newton's law of motion becomes

$$m\ddot{\mathbf{r}} = -e\mathbf{E} - m\nu\dot{\mathbf{r}} + e\mathbf{B} \times \dot{\mathbf{r}}, \tag{1}$$

where m is the mass of electron, $-e$ the charge on electron, $\mathbf{E}$ the total electric field, ν the collision

From (43) and (50),

$$b_n^s = 0. \tag{54}$$

Looking back into the field components in the various regions, it is observed that for the case of a single conducting cylinder there is no z component of the H field. One may thus make an important observation that, for the case of a single conducting cylinder, if the incident wave is TE, only TE modes are excited. Similarly, it can be shown easily that, for this case, if the incident wave is TM, only TM modes are excited. However, for a dielectric cylinder with obliquely incident wave, a combination of TE and TM modes will be required to satisfy the boundary conditions even though the incident wave may be just TE or TM.

The results of various simplified cases may thus be obtained from the general formulation.

The author wishes to acknowledge his indebtedness to Professor Sheldon S. Sandler of Northeastern University, Boston, Massachusetts for his encouragement and guidance, and to C. R. Mullin of AVCO/RAD, Wilmington, Massachusetts for various illuminating discussions on this subject.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 4 APRIL 1964

Application of Operational Methods to the Analysis of Uniform Plasmas

SAUL SILVEN*
McDonnell Aircraft Corporation, St. Louis, Missouri
(Received 16 July 1963)

Various aspects of plasma theory are derived by the application of operational methods to the equation of electron motion. Discussed herein are the complex tensor dielectric constant, the tensor permittivity and conductivity, plasma transient behavior, and plasma oscillations.

INTRODUCTION

THE plasma model discussed in this paper is considered to be composed of equal densities of free electrons and heavy positive ions distributed uniformly throughout the volume under discussion, and neutral particles whose population density is very large with respect to that of the charges. The density of particles is supposed to be large enough to present the appearance of a macroscopically continuous medium, electrically neutral at all points. The collision frequencies are considered to be independent of electron velocity.[1]

The dispersive effects of thermal motions and pressure gradients are neglected. These effects can be analyzed as indicated by Chandrasekhar[2] and Jackson.[3]

* Present address: Lockheed Missiles and Space Company, Sunnyvale, California.

[1] I. P. Shkarofsky, M. P. Bachynski, and T. W. Johnston, *Electromagnetic Effects of Re-entry* (Pergamon Press, Inc., New York, 1961), p. 24.

[2] S. Chandrasekhar, *Plasma Physics* (University of Chicago Press, Chicago, 1960), p. 143.

[3] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New York, 1962), pp. 335–340.

THE EQUATION OF MOTION

Consider an electron–ion pair originally at rest at point 1 (Fig. 1). If the effect of an applied force field is sensitive to charge, the electron–ion pair separates. If the ion mass is much greater than the electron mass, the electron moves to point 2 while the ion is considered to remain approximately at point 1. The symbol $\mathbf{r}$ thus represents electron displacement.

The forces on an electron may be due to the presence of an electric field (alternating or electrostatic), collisions with neutral particles (the average effect of which is a frictional force), and a static magnetic field. The electron velocities are assumed to be much less, numerically, than the speed of light, so that the effect of an alternating magnetic field is negligible in comparison to that of the associated electric field.

Newton's law of motion becomes

$$m\ddot{\mathbf{r}} = -e\mathbf{E} - m\nu\dot{\mathbf{r}} + e\mathbf{B} \times \dot{\mathbf{r}}, \tag{1}$$

where m is the mass of electron, $-e$ the charge on electron, $\mathbf{E}$ the total electric field, ν the collision

frequency, **B** the static magnetic field, and **r** the electron displacement vector. Equation (1) is normalized by dividing by the mass:

$$\ddot{\mathbf{r}} - \omega_c \mathbf{i} \times \dot{\mathbf{r}} + \nu \dot{\mathbf{r}} = -\gamma \mathbf{E}, \tag{2}$$

where γ is the electronic charge-to-mass ratio, $\omega_c = \gamma |\mathbf{B}|$, angular cyclotron rate, and **i** is a unit vector in direction of **B**. The electric field **E** is made up of two parts; the electric displacement **D**, and the induced volume polarization **P**. The relationship is

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \tag{3}$$

where ϵ_0 is the permittivity of free space. The volume polarization is equal to the induced dipole moment per unit volume; therefore

$$\mathbf{P} = -N e \mathbf{r}, \tag{4}$$

where N represents the electron density. Substituting for **E** in Eq. (1),

$$\ddot{\mathbf{r}} - \omega_c \mathbf{i} \times \dot{\mathbf{r}} + \nu \dot{\mathbf{r}} + \omega_p^2 \mathbf{r} = -\gamma \epsilon_0^{-1} \mathbf{D}, \tag{5}$$

where $\omega_p^2 = Ne\epsilon_0^{-1}$. The quantity ω_p will be shown later in this paper to be the resonant angular rate of a collisionless plasma, in the absence of a static magnetic field.

Either Eq. (2) or Eq. (5) may be used to analyze steady-state behavior, i.e., if the general nature of the electron motion is known. For the analysis of transient behavior, Eq. (5) must be used.

TENSOR DIELECTRIC CONSTANT

It is desired to determine the dielectric tensor **K** of the plasma, such that

$$\mathbf{D} = \epsilon_0 \mathbf{K} \cdot \mathbf{E}. \tag{6}$$

From Eqs. (3) and (4) it is seen that **P**, and therefore **r**, must be determined. If the time-derivative operator is represented by the tensor **S**, Eq. (2) can be written

$$[\mathbf{S}^2 - \omega_c \mathbf{i} \times \mathbf{S} + \nu \mathbf{S}] \cdot \mathbf{r}(\mathbf{S}) = -\gamma \mathbf{E}(\mathbf{S}), \tag{7}$$

as **r** and **E** are now functions of **S**. Combining Eq. (7) with Eqs. (3), (4), and (6) shows that[4]

$$\mathbf{K}(\mathbf{S}) = \mathbf{I} + \omega_p^2 [\mathbf{S}^2 - \omega_c \mathbf{i} \times \mathbf{S} + \nu \mathbf{S}]^{-1}, \tag{8}$$

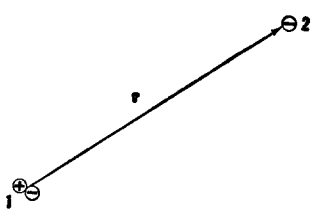

FIG. 1. Displacement of electron.

[4] An inverse $\mathbf{T}^{-1}$ of a tensor **T** is defined by $\mathbf{T}^{-1} \cdot \mathbf{T} = \mathbf{I}$.

where **I** is the idemtensor. If the fields are sinusoidal (elliptically polarized), the operator **S** can be written as $j\omega\mathbf{I}$, with ω representing electrical angular rate and $j = (-1)^{\frac{1}{2}}$. Equation (8) becomes

$$\mathbf{K}(\omega) = \mathbf{I} - \omega_p^2 [(\omega^2 - j\nu\omega)\mathbf{I} + j\omega_c \omega \mathbf{i} \times]^{-1}. \tag{9}$$

Upon introduction of the following standard substitutions[5]

$$\omega_p^2 = X\omega^2, \qquad \omega_c = Y\omega, \qquad \nu = Z\omega, \tag{10}$$

Eq. (9) becomes

$$\mathbf{K}(\omega) = \mathbf{I} - X[(1 - jZ)\mathbf{I} + jY\mathbf{i} \times]^{-1}. \tag{11}$$

It is desirable to express $\mathbf{K}(\omega)$ in terms of the eigenvectors of $j\mathbf{i} \times$. The relationship is found to be[6]

$$j\mathbf{i} \times = \mathbf{i}_1\mathbf{i}_1 - \mathbf{i}_2\mathbf{i}_2, \tag{12}$$

where $\mathbf{i}_1$ and $\mathbf{i}_2$ are unit circularly polarized vectors rotating in the right-handed and left-handed sense,[7] respectively, when viewed in a direction opposite to **i**. Since **i**, $\mathbf{i}_1$, and $\mathbf{i}_2$ form an orthogonal set, the idemtensor can be written as

$$\mathbf{I} = \mathbf{ii} + \mathbf{i}_1\mathbf{i}_1 + \mathbf{i}_2\mathbf{i}_2. \tag{13}$$

Upon taking the indicated inverse, Eq. (11) becomes

$$\mathbf{K}(\omega) = \left[1 - \frac{X}{1 - jZ}\right]\mathbf{ii} + \left[1 - \frac{X}{1 - Y - jZ}\right]\mathbf{i}_1\mathbf{i}_1 + \left[1 - \frac{X}{1 + Y - jZ}\right]\mathbf{i}_2\mathbf{i}_2. \tag{14}$$

The permittivity ε and conductivity σ tensors are given by the real and imaginary parts of **K**;

$$\begin{aligned}\varepsilon(\omega) &= \epsilon_0 R(\mathbf{K}) \\ &= \epsilon_0\left[1 - \frac{X}{1 + Z^2}\right]\mathbf{ii} \\ &\quad + \epsilon_0\left[1 - \frac{X(1 - Y)}{(1 - Y)^2 + Z^2}\right]\mathbf{i}_1\mathbf{i}_1 \\ &\quad + \epsilon_0\left[1 - \frac{X(1 + Y)}{(1 + Y)^2 + Z^2}\right]\mathbf{i}_2\mathbf{i}_2,\end{aligned} \tag{15}$$

$$\begin{aligned}\sigma(\omega) &= \omega\epsilon_0 I(\mathbf{K}) \\ &= \omega\epsilon_0\left[\frac{XZ}{1 + Z^2}\right]\mathbf{ii} + \omega\epsilon_0\left[\frac{XZ}{(1 - Y)^2 + Z^2}\right]\mathbf{i}_1\mathbf{i}_1 \\ &\quad + \omega\epsilon_0\left[\frac{XZ}{(1 + Y)^2 + Z^2}\right]\mathbf{i}_2\mathbf{i}_2.\end{aligned} \tag{16}$$

[5] K. G. Budden, *Radio Waves in the Ionosphere* (Cambridge University Press, Cambridge, England, 1961), 1st ed., pp. 24, 26, 27.

[6] See Appendix I.

[7] Right-handed, as used here, means counterclockwise as viewed in a direction antiparallel to the magnetic field. Left-handed means clockwise.

TRANSIENT BEHAVIOR

In order to analyze the motions within a plasma in the absence of an externally applied field, Eq. (5) must be used with $\mathbf{D} = \mathbf{0}$. Any oscillations involve an interchange of energy between the kinetic energy of the electron motion and the potential energy of the plasma configuration. The equation of motion is

$$\ddot{\mathbf{r}} - \omega_c \mathbf{i} \times \dot{\mathbf{r}} + \nu \dot{\mathbf{r}} + \omega_p^2 \mathbf{r} = \mathbf{0}. \tag{17}$$

For a nontrivial solution to exist, the operator $\mathbf{S}$ must be found so as to satisfy

$$\mathbf{S}^2 - \omega_c \mathbf{i} \times \mathbf{S} + \nu \mathbf{S} + \omega_p^2 \mathbf{I} = \mathbf{0}. \tag{18}$$

The solutions of Eq. (18) are given by

$$2\mathbf{S} = (\nu \pm a)(\mathbf{I} - \mathbf{ii}) + (\omega_c \pm b)\mathbf{i} \times + (-\nu \pm c)\mathbf{ii}, \tag{19}$$

where

$$a\sqrt{2} = \{(\nu^2 - 4\omega_p^2 - \omega_c^2) + [(\nu^2 - 4\omega_p^2 - \omega_c^2)^2 + 4\nu^2\omega_c^2]^{\frac{1}{2}}\}^{\frac{1}{2}}, \tag{20}$$

$$b\sqrt{2} = \{-(\nu^2 - 4\omega_p^2 - \omega_c^2) + [(\nu^2 - 4\omega_p^2 - \omega_c^2)^2 + 4\nu^2\omega_c^2]^{\frac{1}{2}}\}^{\frac{1}{2}}, \tag{21}$$

$$c = (\nu^2 - 4\omega_p^2)^{\frac{1}{2}}, \tag{22}$$

and the upper or lower sign on $\pm a$ requires the upper or lower sign on $\pm b$, respectively.

In order to analyze motion in the time domain, it is necessary to form the operator $\exp(\mathbf{S}t)$, defined by

$$\exp(\mathbf{S}t) = \mathbf{I} + \sum_n (n!)^{-1}(\mathbf{S}t)^n, \qquad n = 1, 2, 3, \cdots. \tag{23}$$

The operator indicated in Eq. (19) transforms to the time domain by

$$\exp(\mathbf{S}t) = (\mathbf{I} - \mathbf{ii}) \exp[\tfrac{1}{2}(-\nu \pm a)t] \times \cos[\tfrac{1}{2}(\omega_c \pm b)t] + (\mathbf{i} \times) \exp[\tfrac{1}{2}(-\nu \pm a)t] \times \sin[\tfrac{1}{2}(\omega_c \pm b)t] + (\mathbf{ii}) \exp[\tfrac{1}{2}(-\nu \pm c)t]. \tag{24}$$

The first two terms on the right-hand side of Eq. (24) show the form of the transient behavior in planes perpendicular to the magnetic field. A study of Eqs. (20) and (21) shows that the quantities a and b are real, with $a < \nu$, so that these motions are a combination of exponentially damped elliptic orbits. The third term on the right-hand side of Eq. (24) shows the form of the transient behavior parallel to the magnetic field. Equation (22) shows that this transient behavior is either a purely exponential decay or an exponentially damped oscillation, depending on whether $\nu > 2\omega_p$ or $\nu < 2\omega_p$. It is noted that the transient behavior parallel to the magnetic field is independent of ω_c. This is due to the fact that a magnetic field has no effect on a charge moving parallel to it.

OSCILLATIONS

It is obvious that sustained self-oscillations are possible only if $\nu = 0$, so that none of the energy being interchanged is absorbed by the neutral particles during collisions. In this case, Eqs. (20), (21), and (22) become

$$a = 0, \tag{25}$$

$$b = (\omega_c^2 + 4\omega_p^2)^{\frac{1}{2}}, \tag{26}$$

$$c = 2j\omega_p. \tag{27}$$

Equation (24) becomes

$$\exp(\mathbf{S}t) = (\mathbf{I} - \mathbf{ii}) \cos\{[\tfrac{1}{2}(\omega_c^2 + 4\omega_p^2)^{\frac{1}{2}} \pm \omega_c]t\} \pm (\mathbf{i} \times) \sin\{[\tfrac{1}{2}(\omega_c^2 + 4\omega_p^2)^{\frac{1}{2}} \pm \omega_c]t\} + (\mathbf{ii}) \cos(\omega_p t). \tag{28}$$

The upper sign indicates a right-handed circular polarization at the higher frequency and the lower sign indicates a left-handed circular polarization at the lower frequency.[7] The third term in Eq. (28) indicates oscillations parallel to the magnetic field at the plasma resonant frequency.

ACKNOWLEDGMENT

The author wishes to express thanks to R. H. Hamilton of McDonnell Aircraft Coproration for his discussions with the author concerning vector and dyadic operations over a field of complex numbers.

APPENDIX I. DETERMINATION OF THE EIGENVECTORS OF $j\mathbf{i}\times$

The eigenvectors $\mathbf{A}_i$ of $j\mathbf{i} \times$ are defined by

$$j\mathbf{i} \times \mathbf{A}_i = k_i \mathbf{A}_i, \tag{29}$$

where the k_i are scalar eigenvalues associated with the $\mathbf{A}_i$. If the eigenvectors are assumed to have the form

$$\mathbf{A}_i = \mathbf{A}_{i1} + j\mathbf{A}_{i2}, \tag{30}$$

with $\mathbf{A}_{i1}$ and $\mathbf{A}_{i2}$ real, it is seen that

$$\mathbf{i} \times \mathbf{A}_{i1} = k_i \mathbf{A}_{i2}, \qquad -\mathbf{i} \times \mathbf{A}_{i2} = k_i \mathbf{A}_{i1}. \tag{31}$$

Solving Eqs. (31) leads to

$$(1 - k_i^2)\mathbf{A}_{i1} = \mathbf{i}(\mathbf{i} \cdot \mathbf{A}_{i1}). \tag{32}$$

The eigenvectors in the plane normal to $\mathbf{i}$ have the eigenvalues $k_1 = +1$ and $k_2 = -1$. Equations (31) show that $\mathbf{A}_{i1}$ and $\mathbf{A}_{i2}$ are equal in magnitude and, since the fields are sinusoidal, the operator $j = (-1)^{\frac{1}{2}}$ indicates that $\mathbf{A}_{i2}$ leads $\mathbf{A}_{i1}$ in phase by one-fourth of a cycle. The eigenvalue $k_1 = +1$ is thus associated with a right-handed (with respect to $\mathbf{i}$) circularly polarized eigenvector ($\mathbf{A}_{12} = +\mathbf{i} \times \mathbf{A}_{11}$), and the eigenvalue $k_2 = -1$ is associated with a left-handed circularly polarized eigenvector[7] ($\mathbf{A}_{22} = -\mathbf{i} \times \mathbf{A}_{21}$). An eigenvector $\mathbf{A}_3$, linearly polarized parallel to i, requires that $k_3 = 0$.

If $\mathbf{j}$ and $\mathbf{k}$ are two unit vectors forming an orthogonal rectangular set with $\mathbf{i}$, then the unit eigenvectors can be represented by $\mathbf{i}_1$ and $\mathbf{i}_2$, where

$$\sqrt{2}\,\mathbf{i}_1 = \mathbf{j} + j\mathbf{k}, \qquad \sqrt{2}\,\mathbf{i}_2 = \mathbf{j} - j\mathbf{k}. \tag{33}$$

Forming the dyadic $\mathbf{i}_1\mathbf{i}_1 - \mathbf{i}_2\mathbf{i}_2$ shows that[8]

$$\mathbf{i}_1\mathbf{i}_1 - \mathbf{i}_2\mathbf{i}_2 = j(\mathbf{kj} - \mathbf{jk}); \tag{34}$$

since, in rectangular coordinates,

$$\mathbf{i} \times = \mathbf{kj} - \mathbf{jk}, \tag{35}$$

it follows that

$$j\mathbf{i} \times = \mathbf{i}_1\mathbf{i}_1 - \mathbf{i}_2\mathbf{i}_2. \tag{12}$$

APPENDIX II. DECOMPOSITION OF AN ELLIPTICALLY POLARIZED FIELD

If Eqs. (14), (15), and (16) are to be useful, it must be shown that any elliptically polarized field can be composed to a linear sum of the eigenvectors. Any elliptically polarized field can be represented by the expression

$$\mathbf{E}(t) = \mathbf{E}_1 \cos \omega t + \mathbf{E}_2 \sin \omega t, \tag{36}$$

where $\mathbf{E}_1$ and $\mathbf{E}_2$ are constant vectors. If there is a preferred direction in space (e.g., that of a static magnetic field), the component of $\mathbf{E}(t)$ linearly polarized in this direction is given by

$$\begin{aligned}\mathbf{E}_A(t) &= (\mathbf{ii})\cdot\mathbf{E}(t) \\ &= \mathbf{i}(\mathbf{i}\cdot\mathbf{E}_1) \cos \omega t + \mathbf{i}(\mathbf{i}\cdot\mathbf{E}_2) \sin \omega t,\end{aligned} \tag{37}$$

where $\mathbf{i}$ is the unit vector in the preferred direction. The remainder of the field is given by

$$\mathbf{E}_B(t) = (\mathbf{I} - \mathbf{ii})\cdot\mathbf{E}(t), \tag{38}$$

which is perpendicular to $\mathbf{i}$. $\mathbf{E}_B(t)$ can now be decomposed into right and left-handed circularly polarized components.

Since both components are perpendicular to $\mathbf{i}$, the right-handed component can be expressed by[7]

$$\mathbf{E}_R(t) = \mathbf{F} \cos \omega t + \mathbf{i} \times \mathbf{F} \sin \omega t, \tag{39}$$

and the left-handed component by

$$\mathbf{E}_L(t) = \mathbf{G} \cos \omega t - \mathbf{i} \times \mathbf{G} \sin \omega t. \tag{40}$$

The constant vectors $\mathbf{F}$ and $\mathbf{G}$ can be determined by setting

$$\mathbf{E}_R(t) + \mathbf{E}_L(t) = \mathbf{E}_B(t), \tag{41}$$

and equating the coefficients of $\cos \omega t$ and $\sin \omega t$ separately. The results are

$$\begin{aligned}\mathbf{F} &= \tfrac{1}{2}[(\mathbf{I} - \mathbf{ii})\cdot\mathbf{E}_1 - \mathbf{i} \times \mathbf{E}_2], \\ \mathbf{G} &= \tfrac{1}{2}[(\mathbf{I} - \mathbf{ii})\cdot\mathbf{E}_1 + \mathbf{i} \times \mathbf{E}_2].\end{aligned} \tag{42}$$

[8] In the formation of the dyadic on the left in terms of $\mathbf{j}$ and $\mathbf{k}$, it is important to note that the dyad $\mathbf{j}(j\mathbf{k}) = -j(\mathbf{jk})$, as shown in a study of vector operations in a complex domain.

Expansions of Integrals of Bessel Functions of Large Order*

C. F. CURTISS
University of Wisconsin Theoretical Chemistry Institute, Madison, Wisconsin
(Received 21 June 1963)

Series expansions in powers of $1/\nu$ of integrals over x of the product of the Bessel function $J_\nu(\nu x/y)$, $[J_\nu(\nu x/y)]^2$, or $J_\nu(\nu x/y)\,J_{\nu+1}(\nu x/y)$ with an arbitrary function of a restricted class are developed.

ASYMTOTIC expressions for Bessel functions of large order and large argument have been considered extensively.[1] In the following we develop series expansions of certain integrals involving such Bessel functions. For this purpose let us define a set of functions

$$F_m(y,\nu) = \int_0^\infty x^m e^{-x} J_\nu(\nu x/y)\,dx, \tag{1}$$

where $J_\nu(z)$ is a Bessel function. The integral may be rewritten simply in the form of a Hankel transform,

$$F_m(\nu/z,\nu) = z^{-\frac{1}{2}} \int_0^\infty x^{m-\frac{1}{2}} e^{-x} J_\nu(xz)(xz)^{\frac{1}{2}}\,dx, \tag{2}$$

included in the tabulation of Erdélyi.[2] From this tabulation, it is found that (if $\nu > -m-1$ and $m \geq 0$)

$$F_m(y,\nu) = (-1)^m \nu^{-1} y^{m+1} G^{(m)}(y,\nu), \tag{3}$$

where $G^{(m)}(y,\nu)$ is the mth derivative with respect to y of the function

$$G(y,\nu) = [(y/\nu)^2+1]^{-\frac{1}{2}}\{[(y/\nu)^2+1]^{\frac{1}{2}} - (y/\nu)\}^\nu. \tag{4}$$

The function $G(y,\nu)$ may be expanded in powers of $1/\nu$. In a straightforward manner it is found that for $0 < y < \nu$

$$G(y,\nu) = \sum_s \frac{(-1)^s \nu!\, y^s}{s!\,(\nu-s)!\,\nu^s}\left[1+\frac{y^2}{\nu^2}\right]^{\frac{1}{2}(\nu-s-1)}$$

$$= \sum_{s,t} \frac{(-1)^s \nu!\,[\frac{1}{2}(\nu-s-1)]!\, y^{s+2t}}{s!\,t!\,(\nu-s)!\,[\frac{1}{2}(\nu-s-1)-t]!\,\nu^{s+2t}}$$

$$= \sum_{s,t} \frac{(-1)^s y^{s+2t}}{s!\,t!\,(2\nu)^t}\left(1-\frac{1}{\nu}\right)\left(1-\frac{2}{\nu}\right)\cdots$$

$$\times\left(1-\frac{s-1}{\nu}\right)\left(1-\frac{s+1}{\nu}\right)$$

$$\times\left(1-\frac{s+3}{\nu}\right)\cdots\left(1-\frac{s+2t-1}{\nu}\right). \tag{5}$$

Thus the function may be written in the form

$$G(y,\nu) = \sum_{n=0}^\infty \nu^{-n} G_n(y). \tag{6}$$

In particular,

$$G_0(y) = e^{-y}, \qquad G_1(y) = 0,$$
$$G_2(y) = -\tfrac{1}{6}(3y^2 - y^3)e^{-y}. \tag{7}$$

It is clear that in general, $G_n(y)$ is of the form

$$G_n(y) = \sum_{s,u} \frac{(-1)^s y^{s+2n-2u}}{s!\,(n-u)!\,2^{n-u}}\, g_{nu}(s), \tag{8}$$

where $g_{nu}(s)$ is a polynomial in s of degree $2u$, and may be written in the form

$$g_{nu}(s) = \sum_{v=0}^{2u} g_{nuv} s(s-1)\cdots(s-2u+v+1), \tag{9}$$

where the g_{nuv} are numerical coefficients. From the last two equations it follows that

$$G_n(y) = e^{-y} \sum_u \sum_{v=0}^{2u} g_{nuv} \frac{(-1)^v y^{2n-v}}{(n-u)!\,2^{n-u}}. \tag{10}$$

That is,

$$G_n(y) = e^{-y} \sum_{i=n}^{2n} A'_{ni} y^i, \tag{11}$$

where the A'_{ni} are numerical constants. In particular, from Eqs. (7), it is seen that

$$A'_{00} = 1, \qquad A'_{11} = A'_{12} = 0,$$
$$A'_{22} = -\tfrac{1}{2}, \quad A'_{23} = \tfrac{1}{6}, \quad A'_{24} = 0. \tag{12}$$

* This research was carried out at the University of Wisconsin Theoretical Chemistry Institute in part under Grant NsG-275-62(4180) with the National Aeronautics and Space Administration, and in part under a grant from the National Science Foundation. This research was also supported in part by a grant from the University Research Committee with funds supplied by the Wisconsin Alumni Research Foundation.

[1] A. Erdélyi, J. Math. Phys. 1, 16 (1960).

[2] *Tables of Integral Transforms*, edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1954), Vol. 2.

From the definition of the function $F_m(y, \nu)$, Eq. (1), and the usual Bessel differential equation, it follows that

$$\left[\frac{\partial^2}{\partial y^2} + \frac{1}{y}\frac{\partial}{\partial y} - \frac{\nu^2}{y^2}\right]F_m(y, \nu) = -\frac{\nu^2}{y^4}F_{m+2}(y, \nu). \quad (13)$$

Then from the form of the function given by Eq. (3), it is seen that

$$(y^2 + \nu^2)G^{(m+2)}(y, \nu) + (2m + 3)yG^{(m+1)}(y, \nu) + [(m + 1)^2 - \nu^2]G^{(m)}(y, \nu) = 0, \quad (14)$$

and from the series expansion [Eq. (6)] one finds that

$$G_{n+2}^{(m+2)}(y) - G_{n+2}^{(m)}(y) = -y^2G_n^{(m+2)}(y) - (2m + 3)yG_n^{(m+1)}(y) - (m + 1)^2G_n^{(m)}(y). \quad (15)$$

In the special case of $m = 0$, this is a set of coupled second-order differential equations,

$$G_{n+2}^{(2)}(y) - G_{n+2}(y) = -y^2G_n^{(2)}(y) - 3yG_n^{(1)}(y) - G_n(y). \quad (16)$$

It follows from induction that if this equation is satisfied, the equations for all values of m are satisfied.

Let us define a set of numerical coefficients, A_{ni}, by the recursion relation

$$j(j - 1)A_{nj} = 2(j - 1)A_{n,j-1} - A_{n-2,j-4} + (2j - 3)A_{n-2,j-3} - (j - 1)^2A_{n-2,j-2}, \quad (17)$$

along with the initial conditions

$$A_{00} = 1; \quad A_{0j} = 0, \quad j \neq 0; \quad A_{1j} = 0; \quad A_{nj} = 0, \quad j < n. \quad (18)$$

From these relations, it follows that

$$A_{22} = -\tfrac{1}{2}; \quad A_{23} = \tfrac{1}{6}; \quad A_{2j} = 0, \quad j \neq 2, 3; \quad A_{3j} = 0; \quad (19)$$

$$A_{44} = \tfrac{3}{8}; \quad A_{45} = -\tfrac{19}{120}; \quad A_{46} = \tfrac{1}{72}; \quad A_{4j} = 0, \quad j \neq 4, 5, 6.$$

Further, let us define the set of functions

$$\mathcal{G}_n(y) = C_{n1}e^{-\nu} + C_{n2}e^{\nu} + e^{-\nu}\sum_j A_{nj}y^j, \quad (20)$$

where the C_{n1} and C_{n2} are arbitrary coefficients.

Let us now consider the set of differential equations [Eqs. (16)]. The functions $G_0(y)$ and $G_1(y)$ are given by Eqs. (7). The equation with $n = 0$ may be considered as a second-order differential equation, which partially describes the function $G_2(y)$. This function is, also, known to be of the form described by Eq. (11). It may readily be shown that the function $\mathcal{G}_2(y)$ defined above is a solution of this equation. Furthermore, since the equation is a second-order differential equation, this expression, with arbitrary values of C_{n1} and C_{n2}, represents the most general solution of the equation. The only solution of Eq. (16) with $n = 0$, which is consistent with Eq. (11), is obtained by setting $C_{21} = C_{22} = 0$. This function is thus $G_2(y)$ [as found earlier, Eq. (7)].

Having obtained the function $G_2(y)$, we may now consider Eq. (16) with $n = 2$. In a similar manner, it is found that the only solution of the equation consistent with (11) is $\mathcal{G}_4(y)$ with $C_{41} = C_{42} = 0$, and thus this function is $G_4(y)$. Continuing the process one finds that, for even n,

$$G_n(y) = e^{-\nu}\sum_i A_{ni}y^i. \quad (21)$$

In a similar manner, starting with the $n = 1$ equation and the known function $G_1(y)$, one finds that the last equation is also valid for odd n. Thus the A'_{ni} introduced by Eq. (11) are equal to the A_{ni} defined by Eqs. (17) and (18).

In summary one finds from Eqs. (3), (6), and (21), that the function $F_m(y, \nu)$ defined by Eq. (1) is

$$F_m(y, \nu) = (-1)^m y^{m+1}\sum_{n,i}\nu^{-n-1}A_{ni}\frac{\partial^m}{\partial y^m}(y^ie^{-\nu}). \quad (22)$$

From elementary properties of derivatives it is readily shown that

$$\frac{\partial^m}{\partial y^m}(y^ie^{-\nu}) = e^{-\nu}\sum_k \frac{(-1)^{m-k}m!\, j!\, y^{i-k}}{k!\,(m - k)!\,(j - k)!}. \quad (23)$$

Thus

$$\frac{\partial^m}{\partial y^m}(y^ie^{-\nu}) = (-1)^{m-i}y^{i-m}\frac{\partial^i}{\partial y^i}(y^me^{-\nu}), \quad (24)$$

and the expression for $F_m(y, \nu)$ may be written in the form

$$F_m(y, \nu) = \sum_{n,i}(-1)^i\nu^{-n-1}A_{ni}y^{i+1}\frac{\partial^i}{\partial y^i}(y^me^{-\nu}). \quad (25)$$

Let $f(x)$ be an arbitrary function of x defined on the range $0 \leq x \leq \infty$ of a class of functions which may be expanded in uniformly convergent series of Laguerre functions. That is, we restrict the consideration to functions $f(x)$, which may be written as powers series in x multiplied by $\exp(-x)$. It follows from the last equation that the integral of the product of such a function with $J_\nu(\nu x/y)$ may be expanded in negative powers of ν in the form

$$\int_0^\infty f(x)J_\nu(\nu x/y)\,dx = \sum_{n,i}(-1)^i\nu^{-n-1}A_{ni}y^{i+1}f^{(i)}(y), \quad (26)$$

where $f^{(i)}(y)$ is the jth derivative of $f(y)$.

A similar expansion of an integral involving the square of the Bessel function may also be obtained. Let us consider a transform of an arbitrary function $f(x)$,

$$\int_0^\infty f(x)[J_\nu(\nu x/y)]^2\,dx. \tag{27}$$

A generalization of Neumann's integral[3] may be written in the form

$$[J_\nu(x)]^2 = (2/\pi)\int_0^{\frac{1}{2}\pi} J_{2\nu}(2x\cos\theta)\,d\theta \tag{28}$$

or

$$[J_\nu(\nu x/y)]^2 = (2/\pi)\int_0^x (x^2 - z^2)^{-\frac{1}{2}} J_{2\nu}(2\nu z/y)\,dz. \tag{29}$$

If the last expression is used in the integral of Eq. (27), and the order of the integrations interchanged, it is found that

$$\int_0^\infty f(x)[J_\nu(\nu x/y)]^2\,dx = \int_0^\infty H(z)J_{2\nu}(2\nu z/y)\,dz, \tag{30}$$

where

$$H(z) = (2/\pi)\int_z^\infty f(x)(x^2 - z^2)^{-\frac{1}{2}}\,dx \tag{31}$$

$$= (2/\pi)\int_0^\infty f(z\cosh x)\,dx.$$

We now restrict the consideration to functions $f(x)$ such that the function $H(z)$ defined above is in the class of functions to which Eq. (26) applies. It then follows from Eqs. (26), (27), and (30) that

$$\int_0^\infty f(x)[J_\nu(\nu x/y)]^2\,dx = \sum_{n,i}(-1)^i(2\nu)^{-n-1}y^{i+1}A_{ni}H^{(i)}(y). \tag{32}$$

Expressions for other integrals involving products of Bessel functions of orders which differ by an integer may readily be obtained from the last result. We consider in particular the case in which the orders differ by unity. By differentiation and use of a standard recursion relation, one finds that

$$J_\nu(\nu x/y)J_{\nu+1}(\nu x/y) = (y/x)[J_\nu(\nu x/y)]^2 - (y/2\nu)(\partial/\partial x)[J_\nu(\nu x/y)]^2. \tag{33}$$

If

$$\lim_{x\to 0} f(x)[J_\nu(\nu x/y)]^2 = 0, \tag{34}$$

and

$$\lim_{x\to\infty} f(x)[J_\nu(\nu x/y)]^2 = 0, \tag{35}$$

it follows by integration by parts that

$$\int_0^\infty f(x)J_\nu(\nu x/y)J_{\nu+1}(\nu x/y)\,dx = y\int_0^\infty [(1/x)f(x) + (2\nu)^{-1}f^{(1)}(x)][J_\nu(\nu x/y)]^2\,dx. \tag{36}$$

If the function $f(x)$ is such that Eq. (32) is valid, we then find that

$$\int_0^\infty f(x)J_\nu(\nu x/y)J_{\nu+1}(\nu x/y)\,dx = \sum_{n,i}(-1)^i(2\nu)^{-n-1}y^{i+2}A_{ni}H_1^{(i)}(y), \tag{37}$$

$$H_1(z) = (2/\pi)\int_z^\infty [(1/x)f(x) + (2\nu)^{-1}f^{(1)}(x)] \times (x^2 - z^2)^{-\frac{1}{2}}\,dx, \tag{38}$$

$$= (2/\pi)\int_0^\infty [(z\cosh x)^{-1}f(z\cosh x) + (2\nu)^{-1}f^{(1)}(z\cosh x)]\,dx. \tag{39}$$

For some purposes it is interesting and convenient to describe these integrations in terms of expansions of the Bessel functions and of the squares and products of the Bessel functions. Let $\delta(x)$ be the usual "delta function", a function which is zero everywhere except at the origin and which is normalized to unity, and let $\delta^{(i)}(x)$ represent the ith derivative of this function. From Eq. (26), it follows that

$$J_\nu(\nu x/y) = \sum_{n,i}\nu^{-n-1}A_{ni}y^{i+1}\delta^{(i)}(x - y), \tag{40}$$

in the sense that the integral on x over the interval $0 \le x \le \infty$ of the product of this series and any function of the class described earlier is correct.

Let us define the function $g(x)$ by

$$\begin{aligned} g(x) &= x^{-\frac{1}{2}}, \qquad x > 0, \\ g(x) &= 0, \qquad x \le 0. \end{aligned} \tag{41}$$

It then follows from Eq. (32) that the square of the Bessel function may be represented by the series

$$[J_\nu(\nu x/y)]^2 = (2/\pi)\sum_{n,i}(-1)^i(2\nu)^{-n-1} \times A_{ni}y^{i+1}(\partial/\partial y)^i g(x^2 - y^2), \tag{42}$$

in the same sense as the previous series. In terms of the function $g(x)$, $H_1(z)$ defined by Eq. (39) may be rewritten

[3] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, Cambridge, England, 1952), p. 150.

$$H_1(z) = (2/\pi)\int_0^\infty f(x)[(1/x)g(x^2 - z^2) - (2\nu)^{-1}(\partial/\partial x)g(x^2 - z^2)]\,dx \quad (43)$$

by integration by parts. Thus it is found from Eq. (38) that

$$J_\nu(\nu x/y)J_{\nu+1}(\nu x/y) = (2/\pi)\sum_{n,i}(-1)^i(2\nu)^{-n-1}y^{i+2}A_{ni} \times (\partial/\partial y)^i[(1/x)g(x^2 - y^2) - (2\nu)^{-1}(\partial/\partial x)g(x^2 - y^2)]. \quad (44)$$

It is easily shown that

$$(\partial/\partial y)^i g(x^2 - y^2) = \sum_{k,l} B_{kl}^{(i)} y^k g^{(l)}(x^2 - y^2), \quad (45)$$

where $g^{(l)}(x)$ is the lth derivative of $g(x)$ and the $B_{kl}^{(i)}$ are determined by the recursion relation

$$B_{kl}^{(i)} = (k+1)B_{k+1,l}^{(i-1)} - 2B_{k-1,i-1}^{(i-1)} \quad (46)$$

and the initial conditions

$$B_{00}^{(0)} = 1; \qquad B_{kl}^{(0)} = 0, \qquad k \neq 0 \text{ or } l \neq 0. \quad (47)$$

Thus it is found from the last expression for the square of the Bessel function [Eq. (42)] that

$$[J_\nu(\nu x/y)]^2 = (2/\pi)\sum_{n,k,l}(2\nu)^{-n-1}D_{nlk}y^{k+1}g^{(l)}(x^2 - y^2), \quad (48)$$

and from Eq. (44) that

$$J_\nu(\nu x/y)J_{\nu+1}(\nu x/y) = (2/\pi)x^{-1}\sum_{n,k,l}(2\nu)^{-n-1} \times [D_{nlk} - 2x^2 D_{n-1,l-1,k}]y^{k+2}g^{(l)}(x^2 - y^2), \quad (49)$$

where

$$D_{nlk} = \sum_i (-1)^i A_{ni}B_{k-i,l}^{(i)}. \quad (50)$$

The nonzero D_{nlk} for $n \leq 4$ are

$$\begin{aligned} &D_{000} = 1, \\ &D_{212} = 1, \quad D_{224} = -4, \quad D_{236} = \tfrac{4}{3}, \\ &D_{424} = \tfrac{8}{2}, \quad D_{436} = -\tfrac{116}{3}, \quad D_{448} = \tfrac{124}{3}, \\ &D_{4,5,10} = -\tfrac{176}{15}, \quad D_{4,6,12} = \tfrac{8}{9}. \end{aligned} \quad (51)$$

The last expressions for the square and product of Bessel functions may be simplified by considering the special case in which y is unity. Thus it follows from Eqs. (48), (49), and (50) that[4]

$$[J_\nu(\nu x)]^2 = (2/\pi)\sum_{n,l}\nu^{-n}C_{nl}g^{(l)}(x^2 - 1), \quad (52)$$

and

$$J_\nu(\nu x)J_{\nu+1}(\nu x) = (2/\pi)x^{-1} \times \sum_{n,l}\nu^{-n}[C_{nl} - x^2 C_{n-l,l-1}]g^{(l)}(x^2 - 1), \quad (53)$$

where

$$C_{nl} = 2^{-n}\sum_k D_{n-1,l,k} = 2^{-n}\sum_{i,k}(-1)^i A_{n-1,i}B_{kl}^{(i)}. \quad (54)$$

The nonzero C_{nl} for $n \leq 5$ are

$$\begin{aligned} &C_{10} = \tfrac{1}{2}, \\ &C_{31} = \tfrac{1}{8}, \quad C_{32} = -\tfrac{1}{2}, \quad C_{33} = \tfrac{1}{6}, \\ &C_{52} = \tfrac{9}{64}, \quad C_{53} = -\tfrac{29}{24}, \quad C_{54} = \tfrac{31}{24}, \\ &C_{55} = -\tfrac{11}{30}, \quad C_{56} = \tfrac{1}{36}. \end{aligned} \quad (55)$$

Expressions for integrals involving products of Bessel functions of orders which differ by an integer greater than unity may be obtained from these results and the usual recursion relations of the Bessel functions.

[4] From arguments based on scattering theory, E. F. Gurnee and J. L. Magee [J. Chem. Phys. **26**, 1237 (1957)] have previously obtained an expression differing negligibly from the first term of this series as an approximation to the square of the Bessel function.