JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 8, NUMBER 8 AUGUST 1967

Dispersion of Electromagnetic Waves in Molecular Crystals

C. MAVROYANNIS
Division of Pure Chemistry, National Research Council, Ottawa, Canada

(Received 2 December 1966)

The dispersion of electromagnetic waves in molecular crystals has been studied using the second quantization formalism. The excitation spectrum, the Green's functions for the optical exciton and photon field, as well as the corresponding distribution functions, are calculated and discussed. The ground state energy of the crystal is derived in a closed form with the polarization of the medium taken into account explicitly. In the low-density limit the expression for the ground state energy corresponds to the summation of an infinite sequence of terms in a perturbation-theory approach.

1. INTRODUCTION

IN a recent paper[1] the excitation spectrum and the ground state energy of a molecular crystal has been calculated using a microscopic approach. In this treatment a molecular crystal has been considered in the tight binding approximation where an electron and a hole are tightly bound at the same lattice site and only contributions arising from the direct electron–electron interactions up to terms of the order $N^{-\frac{1}{2}}$ in the crystal Hamiltonian have been taken into account, N being the total number of unit cells in the crystal. In the present paper we discuss the same problem in the presence of an electromagnetic field.

The problem is formulated in Sec. 2, where the total Hamiltonian of the system, consisting of the crystal Hamiltonian with direct interactions between the electrons plus the Hamiltonian for the electromagnetic field and the electron–photon interaction, is expressed in the second quantization representation and is used to derive the equation of motion for the two-particle double-time retarded Green's function. The higher-order Green's functions appearing in the equation of motion are decoupled by making use of a procedure which is equivalent to the Hartree–Fock self-consistent field approximation. Then a general equation is developed for the two-particle Green's function which describes under certain conditions the excitation spectrum of Frenkel or Mott excitons in an undeformed lattice.

The equations of motion for the Green's functions of the optical exciton (polariton) and photon field for a molecular crystal are developed in Sec. 3, while the dispersion of electromagnetic waves is discussed in Sec. 4. The expression for the excitation spectrum of optical excitons is identical to that found by Agranovich,[2] who used the Bogolyubov's canonical transformation to diagonalize the crystal Hamiltonian plus the electromagnetic field.

Using the expressions for the optical exciton and photon Green's functions, we derive the corresponding distribution functions which are used in Sec. 4 to average the Hamiltonian of the system. The ground state energy of a molecular crystal is obtained in closed form. The polarization of the medium

[1] C. Mavroyannis, J. Chem. Phys. **42**, 1772 (1965).

[2] V. M. Agranovich, Zh. Eksperim. i Teor. Fiz. **37**, 430 (1959) [English transl.: Soviet Phys.—JETP **10**, 307 (1960)].

resulting from the electromagnetic field of the atoms or molecules in the crystal has been included fully in the expression for the ground state energy. Since no restriction has been made about the strength of the interactions involved, the result is exact in the model under consideration up to terms $N^{-\frac{1}{2}}$ in the total Hamiltonian. In the limiting case where the density of the crystal is small, an expansion of the expression for the ground state energy in powers of the density leads to an infinite sequence of terms corresponding to an expansion in perturbation theory.

Higher-order effects will be the subject of later publications; broadening of the exciton spectrum, scattering and absorption of electromagnetic waves including the phonon field, and the dielectric properties of molecular crystals are discussed.

2. FORMULATION OF THE PROBLEM

The Hamiltonian of a crystal, in which all the molecules are fixed at the lattice sites, is taken in the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm ph} + \mathcal{H}_{\rm int}, \tag{1}$$

where $\mathcal{H}_0$ is the crystal Hamiltonian with direct interactions between the electrons:

$$\mathcal{H}_0 = \sum_{ff'} \langle f|\, L\,|f'\rangle \alpha_f^\dagger \alpha_{f'} + \tfrac{1}{2} \sum_{f,f',f_1,f_1'} \langle f, f_1|\, V\,|f', f_1'\rangle \alpha_f^\dagger \alpha_{f_1}^\dagger \alpha_{f_1'} \alpha_{f'}. \tag{2}$$

The index $f \equiv (s, i, \sigma_1)$, $s = \mathbf{n}\alpha$ ($\mathbf{n}$ being the lattice site), $\alpha\,(= 1, 2 \cdots \sigma)$ enumerates the molecules per unit cell, i designates the electron state, and σ_1 is the spin component of an electron ($\pm\frac{1}{2}$). $\alpha_f^\dagger$ and α_f are the creation and annihilation operators satisfying the Fermi anticommutation relations

$$[\alpha_f, \alpha_{f'}^\dagger]_+ = \delta_{ff'}.$$

$\langle f|\, L\,|f'\rangle$ are the matrix elements corresponding to the additive part of the energy operator, i.e., the kinetic energy and the energy of interaction of an electron with the periodic field of the lattice,

$$\langle f|\, L\,|f'\rangle = \int \psi_f^*(\mathbf{r}_s) \times \left[\frac{-1}{2m}\nabla^2 + \sum_{f} V(\mathbf{r}_s - \mathbf{r}_s')\right] \psi_{f'}(\mathbf{r}_s)\, d\tau,$$

$V(\mathbf{r}_s - \mathbf{r}_s')$ being the potential of an electron at the site s, and the matrix elements,

$$\langle f, f_1|\, V\,|f', f_1'\rangle = \int \psi_f^*(\mathbf{r}_s)\psi_{f_1}^*(\mathbf{r}_{s_1}) \times V(\mathbf{r}_s - r_{s_1})\psi_{f'}(\mathbf{r}_s)\psi_{f_1'}(\mathbf{r}_{s_1})\, d\tau_1\, d\tau_2,$$

$$V(\mathbf{r}_s - \mathbf{r}_{s_1}) = \frac{e^2}{\mathbf{r}^s - \mathbf{r}_{s_1}},$$

correspond to the potential energy of the electron–electron interaction. e^2 denotes the square of the charge divided by the static dielectric constant of the substance and the ψ's are the Wannier functions of the electron states. The system of units with $\hbar = 1$ is used throughout.

$\mathcal{H}_{\rm ph}$ is the Hamiltonian for the transverse radiation field

$$\mathcal{H}_{\rm ph} = \sum_{\mathbf{q},\lambda} cq\beta_{\mathbf{q}\lambda}^\dagger \beta_{\mathbf{q}\lambda}, \tag{3}$$

where $\beta_{\mathbf{q}\lambda}^\dagger$ and $\beta_{\mathbf{q}\lambda}$ are the creation and annihilation operators of a photon with wavevector $\mathbf{q}$ and polarization $\lambda\,(= 1, 2)$, representing the two possible values of polarization perpendicular to the direction of propagation $\mathbf{q}$. The interaction between the electrons and photons in the second quantization representation may be taken as[3]

$$\mathcal{H}_{\rm int} = \mathcal{H}_{\rm int}^{(1)} + \mathcal{H}_{\rm int}^{(2)},$$

where

$$\mathcal{H}_{\rm int}^{(1)} = \sum_{\mathbf{q},\lambda,f_1,f_2} \langle f_1|\, T(\mathbf{q}, \lambda)\,|f_2\rangle \alpha_{f_1}^\dagger \alpha_{f_2} e^{i\mathbf{q}\cdot\mathbf{r}_{s_1}(p)} \tilde{\beta}_{\mathbf{q}\lambda}, \qquad \tilde{\beta}_{\mathbf{q}\lambda} \equiv \beta_{\mathbf{q}\lambda} + \beta_{-\mathbf{q}\lambda}^\dagger, \tag{4}$$

$$\mathcal{H}_{\rm int}^{(2)} = \frac{e^2\pi}{mcV} \sum_{\mathbf{q},\lambda,\mathbf{q}',\lambda'} \frac{(\mathbf{e}_{\mathbf{q}\lambda}\cdot\mathbf{e}_{\mathbf{q}'\lambda'})}{(qq')^{\frac{1}{2}}} [\beta_{\mathbf{q}\lambda}\beta_{\mathbf{q}'\lambda'} f(\mathbf{q} + \mathbf{q}') + \beta_{\mathbf{q}\lambda}\beta_{\mathbf{q}'\lambda'}^\dagger f(\mathbf{q} - \mathbf{q}') + \beta_{\mathbf{q}\lambda}^\dagger\beta_{\mathbf{q}'\lambda'} f(-\mathbf{q} + \mathbf{q}') + \beta_{\mathbf{q}\lambda}^\dagger\beta_{\mathbf{q}'\lambda'}^\dagger f(-\mathbf{q} - \mathbf{q}')], \tag{5a}$$

$$\langle f_1|\, T(\mathbf{q}, \lambda)\,|f_2\rangle = -\sum_p \frac{e_p}{m_p}\left(\frac{2\pi}{Vqc}\right)^{\frac{1}{2}} (\mathbf{j}_{f_1 f_2}^p \cdot \mathbf{e}_{\mathbf{q}\lambda}), \qquad f(\mathbf{q}) = \sum_{f_1 f_2} \langle f_1|\, \sum_p e^{i\mathbf{q}\cdot\mathbf{r}_s^{(p)}}\,|f_2\rangle\, \alpha_{f_1}^\dagger \alpha_{f_2}.$$

$\mathbf{j}_{f_1 f_2}^p$ is the matrix element of the momentum operator of the pth electron of a molecule, $\mathbf{e}_{\mathbf{q}\lambda}\,(= \mathbf{e}_{-\mathbf{q}\lambda})$ is the photon polarization vector, V is the volume of the crystal, and $\mathbf{r}_s^{(p)}$ is the position vector of the pth electron at the lattice site s. In the zero approximation, $\mathcal{H}_{\rm int}^{(2)}$ may be written as

$$\mathcal{H}_{\rm int}^{(2)(0)} = \tfrac{1}{4}\omega_p^2 \sum_{\mathbf{q},\lambda} \frac{1}{cq} \tilde{\beta}_{\mathbf{q}\lambda}^\dagger \tilde{\beta}_{\mathbf{q}\lambda}, \tag{5b}$$

ω_p being the plasma frequency

$$\omega_p^2 = 4\pi e^2 N\sigma S/mV, \tag{5c}$$

where S and N are the total number of electrons in the unit cell and the total number of unit cells in the crystal, respectively. $\mathcal{H}_{\rm int}^{(2)(0)}$ instead of $\mathcal{H}_{\rm int}^{(2)}$ is used in what follows.

To study the exciton spectrum, it is sufficient to consider the two-particle retarded double-time Green's function

$$\langle\langle \alpha_{f_1}^\dagger(t)\alpha_{f_2}(t);\, \alpha_{g_1}^\dagger(t')\alpha_{g_2}(t')\rangle\rangle\,.$$

[3] L. N. Ovander, Fiz. Tverd. Tela 3, 2394 (1961) [English transl.: Soviet Phys.—Solid State 3, 1737 (1962)].

The retarded Green's function for the operators $A(t)$ and $B(t')$ is defined as[4]

$$\langle\langle A(t); B(t')\rangle\rangle = -i\theta(t-t')\langle[A(t), B(t')]_{-\eta}\rangle,$$

where

$$\langle U\rangle = Q^{-1}\,\mathrm{Tr}\,[U\exp(-\beta\mathcal{H})]; \quad Q = \mathrm{Tr}\exp(-\beta\mathcal{H}),$$
$$\beta = (K_B T)^{-1},$$

K_B is Boltzmann's constant, T the absolute temperature, and $\mathcal{H}$ is the total Hamiltonian. The value of η is taken to be either $+1$ or -1 depending upon considerations of convenience, $\theta(t)$ is the step function

$$\theta(t) = 0, \quad t < 0; \quad \theta(t) = 1, \quad t > 0,$$

and the operators $A(t)$ and $B(t')$ are expressed in the Heisenberg representation

$$A(t) = \exp(i\mathcal{H}t)A(0)\exp(-i\mathcal{H}t).$$

The equation of motion for the Fourier transform of the Green's function $\langle\langle A(t); B(t')\rangle\rangle_{(\omega)}$ is given by[5]

$$\omega\langle\langle A(t); B(t')\rangle\rangle_{(\omega)} = \frac{1}{2\pi}\langle[A(t), B(t)]_{-\eta}\rangle + \langle\langle[A(t), \mathcal{H}]_{-}; B(t')\rangle\rangle_{(\omega)}. \quad (6)$$

In what follows we omit the subscript ω for the sake of convenience.

Using the Hamiltonian (1), the equation of motion (6) with $\eta = 1$ for the Fourier transform of the Green's function $\langle\langle\alpha^\dagger_{f_1}\alpha_{f_2}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle$, and the following decoupling procedure for the higher-order Green's functions:

$$\begin{aligned}\langle\langle\alpha^\dagger_{f_1}\alpha^\dagger_{f_2}\alpha_{f_3}\alpha_{f_4}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle &\approx \langle\alpha^\dagger_{f_2}\alpha_{f_3}\rangle\langle\langle\alpha^\dagger_{f_1}\alpha_{f_4}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle - \langle\alpha^\dagger_{f_2}\alpha_{f_4}\rangle\langle\langle\alpha^\dagger_{f_1}\alpha_{f_3}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &+ \langle\alpha^\dagger_{f_1}\alpha_{f_4}\rangle\langle\langle\alpha^\dagger_{f_2}\alpha_{f_3}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle - \langle\alpha^\dagger_{f_1}\alpha_{f_3}\rangle\langle\langle\alpha^\dagger_{f_2}\alpha_{f_4}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \end{aligned} \quad (7)$$

which is equivalent to the Hartree–Fock self-consistent field approximation,[6] we derive the following expression:

$$\begin{aligned}\omega\langle\langle\alpha^\dagger_{f_1}\alpha_{f_2}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle &= \frac{1}{2\pi}(\langle\alpha^\dagger_{f_1}\alpha_{g_2}\rangle\delta_{f_2 g_1} - \langle\alpha_{g_1}\alpha_{f_2}\rangle\delta_{f_1 g_2}) \\ &+ \sum_f F(f_2 f)\langle\langle\alpha^\dagger_{f_1}\alpha_f; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &- \sum_f F(f, f_1)\langle\langle\alpha^\dagger_f\alpha_{f_2}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &+ \sum_{f,f',f''}[\langle f_2, f'|\,V\,|f, f''\rangle - \langle f_2 f'|\,V\,|f'', f\rangle] \\ &\times \langle\alpha^\dagger_{f_1}\alpha_f\rangle\langle\langle\alpha^\dagger_{f'}\alpha_{f''}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &- \sum_{f,f',f''}[\langle f, f'|\,V\,|f_1, f''\rangle - \langle f, f'|\,V\,|f'', f_1\rangle] \\ &\times \langle\alpha^\dagger_f\alpha_{f_2}\rangle\langle\langle\alpha^\dagger_{f'}\alpha_{f''}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &+ \sum_{f,\mathbf{q},\lambda}\langle f_2|\,T(\mathbf{q}\lambda)\,|f\rangle\langle\langle\alpha^\dagger_{f_1}\alpha_f\tilde{\beta}_{\mathbf{q}\lambda}e^{i\mathbf{q}\cdot\mathbf{r}_s^{(p)}}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle \\ &- \sum_{f,\mathbf{q},\lambda}\langle f|\,T(\mathbf{q}\lambda)\,|f_1\rangle\langle\langle\alpha^\dagger_f\alpha_{f_2}\tilde{\beta}_{\mathbf{q}\lambda}e^{i\mathbf{q}\cdot\mathbf{r}_s^{(p)}}; \alpha^\dagger_{g_1}\alpha_{g_2}\rangle\rangle, \end{aligned} \quad (8)$$

where

$$F(f_1 f_2) \equiv \langle f_1|\,L\,|f_2\rangle + \sum_{f'f''}[\langle f_1, f'|\,V\,|f_2, f''\rangle - \langle f_1, f'|\,V\,|f'', f_2\rangle]\langle\alpha^\dagger_{f'}\alpha_{f''}\rangle.$$

The last two terms on the right-hand side of (8) correspond to the electron–photon interaction. Thus the expression (8) describes a two-particle excitation spectrum including retardation effects. We note here that by using the approximation (7), we have neglected pairing of electrons of the superconducting type, $\langle\alpha_{f_1}\alpha_{f_2}\rangle$ and $\langle\alpha^\dagger_{f_1}\alpha^\dagger_{f_2}\rangle$, because they are not of interest for our problem.

If we neglect the last two terms on the right-hand side of (8) and set $f_1 \equiv (s, i, \sigma_1)$, $f_2 \equiv (s_1, j, \sigma_2)$, i and j being the states of the electron and the hole, respectively, then the resulting equation for the Green's function

$$\langle\langle\alpha^\dagger_{si\sigma_1}\alpha_{s_1 j\sigma_2}; \alpha^\dagger_{s_1' j'\sigma_2'}\alpha_{s' i'\sigma_1'}\rangle\rangle$$

describes the exciton spectrum without retardation. In particular, if $s = s_1$, i.e., the electron and the hole are tightly bound at the same lattice site, the resulting equation for $\langle\langle\alpha^\dagger_{si\sigma_1}\alpha_{sj\sigma_2}; \alpha^\dagger_{s'j'\sigma_2'}\alpha_{s'i'\sigma_2}\rangle\rangle$ will correspond to the spectrum of single ($\sigma_1 = \sigma_2$) or triple ($\sigma_1 \neq \sigma_2$) Frenkel exciton. If, on the other hand, $s \neq s_1$, then one obtains the corresponding spectrum of the Mott-type exciton. The spectra of single Frenkel and Mott type excitons have been discussed in this way by Dzyub[7] at finite temperatures. Thus, if we introduce the compound operators $b_{sj} = \alpha^\dagger_{s_0\sigma_1}\alpha_{s_0\sigma_1}$ and $b^\dagger_{sj} = \alpha^\dagger_{sj\sigma_1}\alpha_{s_0\sigma_1}$, and use (8) and its complex conjugate, we may calculate the excitation and average energy corresponding to the Hamiltonian $\mathcal{H}_0$. Therefore, we write the Hamiltonian $\mathcal{H}_0$ in the form

$$\mathcal{H}_0 = \mathcal{H}^{(0)} + \mathcal{H}^{\mathrm{III}}_{\mathrm{int}} + \mathcal{H}^{\mathrm{IV}}_{\mathrm{int}}, \quad (9)$$

where $\mathcal{H}^{(0)}$ represents the bare or mechanical exciton spectrum of a molecular crystal correct to the order $N^{-\frac{1}{2}}$ and is given by

$$\mathcal{H}^{(0)} = \langle\mathcal{H}_0\rangle + \sum_{\mathbf{k},\mu} E_\mu(\mathbf{k})b^\dagger_\mu(\mathbf{k})b_\mu(\mathbf{k}), \quad (10)$$

[4] N. N. Bogolyubov and S. V. Tyablikov, Dokl. Akad. Nauk SSSR **126**, 53 (1959) [English transl.: Soviet Phys.—Doklady **4**, 589 (1959)].

[5] D. N. Zubarev, Usp. Fiz. Nauk **71**, 71 (1960) [English transl.: Soviet Phys.—Usp. **3**, 320 (1960)].

[6] S. V. Tyablikov and V. L. Bonch-Bruevich, Advan. Phys. **11**, 317 (1962).

[7] I. P. Dzyub, Dokl. Acad. Nauk SSSR **130**, 1241 (1960) [English transl.: Soviet Phys.—Doklady **5**, 125 (1959)]; Zh. Eksperim. i Teor. Fiz. **39**, 610 (1960) [English transl.: Soviet Phys.—JETP **12**, 429 (1961)].

where $E_\mu(\mathbf{k})$ is the excitation energy of the μth exciton band with wavevector $\mathbf{k}$. Here μ is a compound index that denotes the exciton band, the corresponding molecular term and the kind of mode, transverse ($\mu_\perp = 1, 2$) or longitudinal ($\mu_\parallel = 3$). In (10) $\mathbf{k}$ is a wavevector in the first Brillouin zone. Effects resulting from configuration mixing of different states may also be included in the expression for $E_\mu(\mathbf{k})$.[1,8] The operators $b_\mu(\mathbf{k})$ and $b_\mu^\dagger(\mathbf{k})$ satisfy the commutation relation

$$[b_\mu(\mathbf{k}), b_{\mu'}^\dagger(\mathbf{k}')]_- = [n_{0\mu}(\mathbf{k}) - n_\mu(\mathbf{k})]\delta_{\mu\mu'}\delta_{\mathbf{k}\mathbf{k}'},$$

where

$$n_0(\mathbf{k}) \equiv \langle\alpha_0^\dagger(\mathbf{k})\alpha_0(\mathbf{k})\rangle \quad \text{and} \quad n_\mu(\mathbf{k}) \equiv \langle\alpha_\mu^\dagger(\mathbf{k})\alpha_\mu(\mathbf{k})\rangle$$

are the occupation numbers of the unperturbed initial and final state, respectively. $\langle\mathcal{H}_0\rangle$ in (10) is the average energy resulting from the direct short- and long-range electron–hole pair interactions including configuration interaction; at zero temperature, it corresponds to the ground state energy arising from the zero-point fluctuations of the excitation field as has been pointed out by Hopfield[9] and Anderson.[10] Expressions for $\langle\mathcal{H}_0\rangle$ have been given in the literature.[1] In (9) $\mathcal{H}_{\text{int}}^{\text{III}}$ and $\mathcal{H}_{\text{int}}^{\text{IV}}$ are higher-order terms proportional to $N^{-\frac{1}{2}}$ and N^{-1}, respectively, and express scattering processes. For example, if the exchange interaction is neglected, then $\mathcal{H}_{\text{int}}^{\text{III}}$ takes the form[3,11]

$$\mathcal{H}_{\text{int}}^{\text{III}} = N^{-\frac{1}{2}} \sum_{\substack{\mathbf{k},\mathbf{q} \\ \mu,\mu',\mu''}} U_{0\mu'',\mu\mu'}(\mathbf{k}-\mathbf{q}) \times b_{\mu''}(\mathbf{k}) b_{\mu'}^\dagger(\mathbf{q}) b_\mu^\dagger(\mathbf{k}-\mathbf{q}) + \text{c.c.}, \quad (11)$$

where

$$U_{0\mu'',\mu\mu'}(\mathbf{k}-\mathbf{q}) = \sum_{s_1 \neq s} \langle s0, s_1\mu''|\, V_{ss_1}\, |s\mu, s_1\mu'\rangle \times u_{\mu''}(\mathbf{k}) u_{\mu'}^\dagger(\mathbf{q}) u_\mu^\dagger(\mathbf{k}-\mathbf{q}) \exp[i(\mathbf{k}-\mathbf{q})(\mathbf{r}_{s_1} - \mathbf{r}_s)],$$

and the u's and v's are the amplitudes of the canonical transformation which diagonalize the unperturbed part of the Hamiltonian, $\mathcal{H}^{(0)}$. Since the v amplitudes are much smaller than the u's,[1,3] they have been omitted in (11). Explicit expressions for the u's and v's have been given elsewhere.[1–3] The expression for $\mathcal{H}_{\text{int}}^{\text{IV}}$ contains four operators but its explicit form is not given here.

3. GREEN'S FUNCTIONS FOR THE OPTICAL EXCITON AND PHOTON FIELD

The Hamiltonian $\mathcal{H}_{\text{int}}^{(1)}$ is expressed now in the representation where $\mathcal{H}^{(0)}$ is diagonal, i.e.,

$$\mathcal{H}_{\text{int}}^{(1)} = \mathcal{H}_{\text{int}}^{\text{I}} + \mathcal{H}_{\text{int}}^{\text{II}}, \quad (12)$$

where

$$\mathcal{H}_{\text{int}}^{\text{I}} = \frac{i\omega_p}{2} \sum_{\mathbf{k},\lambda,\mu} f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda) \left[\frac{E_\mu(k)}{ck}\right]^{\frac{1}{2}} \tilde{b}_\mu(\mathbf{k}) \tilde{\beta}_\lambda(\mathbf{k}), \quad (13)$$

$$\mathcal{H}_{\text{int}}^{\text{II}} = \frac{i\omega_p}{N^{\frac{1}{2}}} \sum_{\substack{\mathbf{k},\mathbf{q},\lambda \\ \mu,\mu'}} f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}^{\frac{1}{2}}(\mathbf{q},\lambda) \left(\frac{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}}{cq}\right)^{\frac{1}{2}} \times b_\mu^\dagger(\mathbf{k}) b_{\mu'}(\mathbf{k}-\mathbf{q}) \tilde{\beta}_\lambda(\mathbf{q}), \quad (14)$$

$$\tilde{b}_\mu(\mathbf{k}) \equiv b_\mu(-\mathbf{k}) - b_\mu^\dagger(\mathbf{k}),$$

and the coupling constants are given by

$$f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda) = \left(\frac{2m}{e^2\sigma S}\right)^{\frac{1}{2}} \sum_\alpha (\mathbf{e}_{\mathbf{q}\lambda} \cdot \mathbf{P}_{0\mu}^\alpha) E_\mu^{\frac{1}{2}}(\mathbf{k}) \times [u_{\mu\alpha}(\mathbf{k}) + v_{\mu\alpha}(\mathbf{k})] \exp(i\mathbf{k}\cdot\mathbf{r}_{\mathbf{n}\alpha}), \quad (15)$$

$$f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}^{\frac{1}{2}}(\mathbf{q},\lambda) = \left(\frac{2m}{e^2\sigma S}\right)^{\frac{1}{2}} \sum_\alpha (\mathbf{e}_{\mathbf{q}\lambda} \cdot \mathbf{P}_{\mu\mu'}^\alpha) \times E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}^{\frac{1}{2}} u_{\mu\alpha}^\dagger(\mathbf{k}) u_{\mu'\alpha}(\mathbf{k}-\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r}_{\mathbf{n}\alpha}), \quad (16)$$

$$E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu} \equiv E_{\mu'}(\mathbf{k}-\mathbf{q}) - E_\mu(\mathbf{k}).$$

$\mathbf{P}_{0\mu}^\alpha$ and $\mathbf{P}_{\mu\mu'}^\alpha$ are the dipole moment operators of the molecule $\mathbf{n}\alpha$ for the allowed transitions $0 \to \mu$ and $\mu \to \mu'$, respectively. In (14) we have kept only the u's while contributions from terms proportional to the v amplitudes have been neglected. The Hamiltonians $\mathcal{H}_{\text{int}}^{\text{I}}$ and $\mathcal{H}_{\text{int}}^{\text{II}}$ describe dispersion and scattering of the electromagnetic waves, respectively.

We introduce the retarded Green's function

$$\tilde{G}_\mu(\mathbf{k};\omega) \equiv \langle\langle A_\mu(\mathbf{k}); A_\mu^\dagger(\mathbf{k})\rangle\rangle \equiv \left\langle\!\!\left\langle \begin{pmatrix} b_\mu(\mathbf{k}) \\ b_\mu^\dagger(-\mathbf{k}) \end{pmatrix}; [b_\mu^\dagger(\mathbf{k}) b_\mu(-\mathbf{k})] \right\rangle\!\!\right\rangle,$$

and using (6), (9), (10), and (12), we obtain the equation of motion for $\tilde{G}_\mu(\mathbf{k};\omega)$:

$$[\hat{\alpha}\omega - E_\mu(\mathbf{k})]\tilde{G}_\mu(\mathbf{k};\omega) = \frac{I}{2\pi}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) + \langle\langle B_\lambda(\mathbf{k},\mu); A_\mu^\dagger(\mathbf{k})\rangle\rangle + \langle\langle \Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu'); A_\mu^\dagger(\mathbf{k})\rangle\rangle + \langle\langle \hat{V}_{\mathbf{k},\mathbf{q}}^{\alpha\beta}; A_\mu^\dagger(\mathbf{k})\rangle\rangle, \quad (17)$$

where we have made use of the following notation:

$$B_\lambda(\mathbf{k},\mu) \equiv -\frac{i\omega_p}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_\lambda f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda) \times \left[\frac{E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}} \begin{pmatrix} \tilde{\beta}_\lambda(\mathbf{k}) \\ -\tilde{\beta}_\lambda(\mathbf{k}) \end{pmatrix}, \quad (18a)$$

$$\Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu') \equiv (n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \frac{i\omega_p}{N^{\frac{1}{2}}} \sum_{\mathbf{q}\lambda,\mu'} f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}^{\frac{1}{2}}(\mathbf{q},\lambda) \times \left(\frac{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}}{cq}\right)^{\frac{1}{2}} \begin{pmatrix} b_{\mu'}(\mathbf{k}-\mathbf{q})\tilde{\beta}_\lambda(\mathbf{q}) \\ -b_{\mu'}^\dagger(\mathbf{k}-\mathbf{q})\tilde{\beta}_\lambda^\dagger(\mathbf{q}) \end{pmatrix}, \quad (18b)$$

[8] V. M. Agranovich, Fiz. Tverd. Tela **3**, 811 (1961) [English transl.: Soviet Phys.—Solid State **3**, 592 (1961)].

[9] J. J. Hopfield, Phys. Rev. **112**, 1555 (1958).

[10] P. W. Anderson, *Concepts in Solids* (W. A. Benjamin, Inc., New York, 1964), p. 147.

[11] L. N. Ovander, Usp. Fiz. Nauk **86**, 3 (1965) [English transl.: Soviet Phys.—Usp. **8**, 337 (1965)].

$$\hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} \equiv \frac{(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})}{N^{\frac{1}{2}}} \begin{pmatrix} \hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} \\ \hat{U}^{\alpha\beta\dagger}_{-\mathbf{k},-\mathbf{q}} \end{pmatrix},$$
$$\hat{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{18c}$$

$$\hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} = \sum_{\substack{\mathbf{q},\mu',\mu_1 \\ \beta}} \{[U^{\alpha\beta}_{0\mu',\mu_1\mu}(\mathbf{q}) + U^{\alpha\beta}_{0\mu',\mu\mu_1}(\mathbf{k})]b^{\dagger}_{\mu_1}(-\mathbf{q}) + U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k}-\mathbf{q})b_{\mu_1}(q)\}b_{\mu'}(\mathbf{k}-\mathbf{q}). \tag{18d}$$

In (17), the term proportional to N^{-1} has been omitted. Similarly, the equation of motion for the Green's function $\langle\langle B_\lambda(\mathbf{k},\mu); A^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle$ is given by

$$\langle\langle B_\lambda(\mathbf{k},\mu); A^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle = \frac{\omega_p^2}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mu,\lambda} \frac{f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})\hat{\sigma}_3\tilde{G}_\mu(\mathbf{k};\omega)}{(\omega^2 - c^2k^2 - \omega_p^2)} + \langle\langle C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu'); A^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle, \tag{19}$$

where

$$C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu') = \frac{\omega_p^2}{N^{\frac{1}{2}}}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q}\lambda,\mu_1\mu'} f^{\frac{1}{2}}_{0\mu}(\mathbf{k},\lambda) \times \frac{f^{\frac{1}{2}}_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k},\lambda)[E_\mu(\mathbf{k})E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{\frac{1}{2}}}{(\omega^2 - c^2k^2 - \omega_p^2)} \times \begin{pmatrix} b^{\dagger}_{\mu_1}(-\mathbf{q})b_{\mu'}(\mathbf{k}-\mathbf{q}) \\ -b^{\dagger}_{\mu_1}(-\mathbf{q})b_{\mu'}(\mathbf{k}-\mathbf{q}) \end{pmatrix}, \tag{20}$$
$$\hat{\sigma}_3 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Substitution of (19) into (17) leads to

$$\tilde{G}^{(0)-1}_{\mu}(\mathbf{k}\lambda;\omega)\tilde{G}_\mu(\mathbf{k};\omega) = \frac{1}{2\pi}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) + \langle\langle \Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu') + C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu') + \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}}; A^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle \tag{21}$$

with

$$\tilde{G}^{(0)-1}_{\mu}(\mathbf{k}\lambda;\omega) = \hat{\alpha}\omega - E_\mu(\mathbf{k}) - \frac{\omega_p^2}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \times \sum_{\mu,\lambda} \frac{f_{0\mu}(\mathbf{k}\lambda)E_\mu(\mathbf{k})}{(\omega^2 - c^2k^2 - \omega_p^2)}\hat{\sigma}_3. \tag{22}$$

In the same way, we derive the equation of motion for the photon Green's function

$$D_\lambda(\mathbf{k};\omega) \equiv \langle\langle \tilde{\beta}_\lambda(\mathbf{k}); \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle$$

in the form

$$D^{(0)-1}_{\lambda}(\mathbf{k};\omega)D_\lambda(\mathbf{k};\omega) = 1 + \frac{\pi\omega_p^2}{N^{\frac{1}{2}}}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \times \sum_{\mathbf{q},\lambda',\mu} f^{\frac{1}{2}}_{0\mu}(\mathbf{k},\lambda)f^{\frac{1}{2}}_{\mu\mu'}(\mathbf{q}\lambda')\left(\frac{E_\mu(\mathbf{k})E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}}{ckcq}\right)^{\frac{1}{2}} \times \left[\frac{1}{E_\mu(\mathbf{k}) - \omega}\langle\langle b_{\mu'}(\mathbf{k}-\mathbf{q})\tilde{\beta}_{\lambda'}(\mathbf{q}); \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle + \frac{1}{E_\mu(\mathbf{k}) + \omega}\langle\langle b^{\dagger}_{\mu'}(\mathbf{k}-\mathbf{q})\tilde{\beta}^{\dagger}_{\lambda'}(\mathbf{q}); \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle\right] + \frac{2\pi i\omega_p}{N^{\frac{1}{2}}} \sum_{\mathbf{q},\mu_1,\mu'} f^{\frac{1}{2}}_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{k},\lambda)\left(\frac{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}}{ck}\right)^{\frac{1}{2}} \times \langle\langle b^{\dagger}_{\mu_1}(-\mathbf{q})b_{\mu'}(\mathbf{k}-\mathbf{q}); \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle + \frac{2\pi i\omega_p}{N^{\frac{1}{2}}} \sum_{\mu} f^{\frac{1}{2}}_{0\mu}(\mathbf{k},\lambda) \times \left[\frac{E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}}\left[\frac{1}{\omega - E_\mu(\mathbf{k})}\langle\langle \hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}}; \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle + \frac{1}{\omega + E_\mu(\mathbf{k})}\langle\langle \hat{U}^{\alpha\beta\dagger}_{-\mathbf{k},-\mathbf{q}}; \tilde{\beta}^{\dagger}_{\lambda}(\mathbf{k})\rangle\rangle\right], \tag{23}$$

where

$$D^{(0)-1}_{\lambda}(\mathbf{k};\omega) = (\pi/ck)[\omega^2 - c^2k^2 - \omega_p^2 + \alpha_\lambda(\mathbf{k},\omega)E^2_\mu(\mathbf{k})], \tag{24}$$

and $\alpha_\lambda(\mathbf{k},\omega)$ is the frequency-dependent polarizability with polarization λ defined by

$$\alpha_\lambda(\mathbf{k},\omega) = \omega_p^2(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\lambda,\mu} f_{0\mu}(\mathbf{k},\lambda)/[E^2_\mu(\mathbf{k}) - \omega^2]. \tag{25}$$

4. DISPERSION

To study the polariton (dressed exciton) spectrum, we have to neglect the Green's functions that appeared on the right-hand side of (21), i.e.,

$$\tilde{G}^{(0)-1}_{\mu}(\mathbf{k}\lambda;\omega)\tilde{G}^{(0)}_{\mu}(\mathbf{k};\omega) = I(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi. \tag{26}$$

Taking the diagonal and nondiagonal elements of (26), we have, respectively,

$$[c^2k^2 + \omega_p^2 - \omega^2 - \alpha_\lambda(\mathbf{k},\omega)E^2_\mu(\mathbf{k})]G^{(0)}_{\mu}(\mathbf{k};\omega) = (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi \times \left[\frac{c^2k^2 + \omega_p^2 - \omega^2}{\omega - E_\mu(\mathbf{k})} + \tfrac{1}{2}\alpha_\lambda(\mathbf{k},\omega)E_\mu(k)\right], \tag{27}$$

$$[c^2k^2 + \omega_p^2 - \omega^2 - \alpha_\lambda(\mathbf{k},\omega)E^2_\mu(\mathbf{k})]\hat{G}^{(0)}_{\mu}(\mathbf{k};\omega) = (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi \times \tfrac{1}{2}\alpha_\lambda(\mathbf{k},\omega)E_\mu(\mathbf{k}), \tag{28}$$

where

$$G^{(0)}_{\mu}(\mathbf{k};\omega) \equiv \langle\langle b_\mu(\mathbf{k}); b^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle^{(0)}$$

and

$$\hat{G}^{(0)}_{\mu}(\mathbf{k};\omega) \equiv \langle\langle b^{\dagger}_{\mu}(-\mathbf{k}); b^{\dagger}_{\mu}(\mathbf{k})\rangle\rangle^{(0)}$$

are the unperturbed single-particle retarded Green's functions. The longitudinal excitons are not coupled with the radiation field; therefore,

$$G^{(0)}_{\mu\|}(\mathbf{k};\omega) = (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi \frac{1}{\omega - E_{\mu\|}(\mathbf{k})}, \quad \hat{G}^{(0)}_{\mu\|}(\mathbf{k};\omega) = 0. \tag{29}$$

For the transverse excitons, taking into account the

degeneracy with respect to λ, we have from (27) and (28)

$$[(c^2k^2+\omega_p^2-\omega^2)\delta_{ij}-\alpha(\mathbf{k},\omega)\eta_{ij}E_\mu^2(\mathbf{k})]G^{(0)}_{jj,\mu_\perp}(\mathbf{k};\omega)$$
$$=\frac{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})}{2\pi}$$
$$\times\left[\frac{c^2k^2+\omega_p^2-\omega^2}{\omega-E_\mu(\mathbf{k})}\delta_{ij}+\tfrac{1}{2}\alpha(\mathbf{k},\omega)\eta_{ij}E_\mu(\mathbf{k})\right], \quad (30)$$

$$[(c^2k^2+\omega_p^2-\omega^2)\delta_{ij}-\alpha(\mathbf{k},\omega)\eta_{ij}E_\mu^2(\mathbf{k})]\hat{G}^{(0)}_{jj,\mu_\perp}(\mathbf{k};\omega)$$
$$=\frac{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})}{2\pi}\tfrac{1}{2}\alpha(\mathbf{k},\omega)\eta_{ij}E_\mu(\mathbf{k}), \quad (31)$$

with $\eta_{ij}=\delta_{ij}-k_ik_j/k^2$ and $i,j=1,2$. The energy of excitation is obtained from the poles of (30) or (31), i.e., from the zeros of the determinant,

$$\|(c^2k^2+\omega_p^2-\omega^2)\delta_{ij}-\alpha(\mathbf{k},\omega)\eta_{ij}E_\mu^2(\mathbf{k})\|=0, \quad (32a)$$

which leads to the following equation:

$$(c^2k^2+\omega_p^2-\omega^2)^2-(c^2k^2+\omega_p^2-\omega^2)(\alpha_{11}+\alpha_{22})$$
$$\times E_\mu^2(\mathbf{k})+(\alpha_{11}\alpha_{22}-\alpha_{12}^2)E_\mu^4(\mathbf{k})=0, \quad (32b)$$

with the notation $\alpha_{ij}\equiv\alpha_{ij}(\mathbf{k},\omega)\equiv\alpha(\mathbf{k},\omega)\eta_{ij}$ and $\alpha_{ij}\equiv\alpha_{ji}$. Using the sum rule developed in Ref. 2, Eq. (32b) may be written as

$$[c^2k^2-\omega^2\hat{\eta}_+^2(\mathbf{k},\omega)][c^2k^2-\omega^2\hat{\eta}_-^2(\mathbf{k},\omega)]=0, \quad (33)$$

where

$$\hat{\eta}_\pm^2(\mathbf{k},\omega)=1+\tfrac{1}{2}(\alpha_{11}+\alpha_{22})$$
$$\pm\tfrac{1}{2}[(\alpha_{11}-\alpha_{22})^2+4\alpha_{12}^2]^{\frac{1}{2}} \quad (34)$$

is the index of refraction of optical waves in the crystal. Thus, the poles of the Green's function given by the roots of (33) determine the dispersion of the electromagnetic waves in a molecular crystal. The expressions (32b) and (34), apart from the factor $(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})$ that appears in (25) and is of the order of unity, are identical with those derived by Agranovich[2]; we refer to his paper for details where expressions for $\hat{\eta}_\pm^2(\mathbf{k},\omega)$ are given for crystals of definite symmetry.

Using (27), its complex conjugate, and (28), we derive the corresponding spectral functions in the usual way[1,5] and, integrating over ω and taking the limit at zero temperature, we obtain the following expressions for the distribution functions:

$$\langle b_\mu^\dagger(\mathbf{k})b_\mu(\mathbf{k})\rangle_0^{(0)}$$
$$=(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\frac{\omega_p^2}{4}\sum_{\mu,\lambda,\rho}\frac{f_{0\mu}(\mathbf{k},\lambda)}{[E_\mu(\mathbf{k})+\omega_{\rho\lambda}(\mathbf{k})]^2}$$
$$\times\left(\frac{E_\mu(\mathbf{k})}{\omega_{\rho\lambda}(\mathbf{k})}\right)\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(k,\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (35)$$

$$\langle b_\mu(\mathbf{k})b_\mu^\dagger(\mathbf{k})\rangle_0^{(0)}$$
$$=(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\frac{\omega_p^2}{4}\sum_{\mu,\lambda,\rho}\frac{f_{0\mu}(\mathbf{k},\lambda)}{[E_\mu(\mathbf{k})-\omega_{\rho\lambda}(\mathbf{k})]^2}$$
$$\times\left(\frac{E_\mu(\mathbf{k})}{\omega_{\rho\lambda}(\mathbf{k})}\right)\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (36)$$

$$\langle b_\mu(\mathbf{k})b_\mu(-\mathbf{k})\rangle_0^{(0)}$$
$$=-(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\tfrac{1}{4}\sum_\rho\alpha_\lambda[\mathbf{k},\omega_{\rho\lambda}(\mathbf{k})]\left(\frac{E_\mu(\mathbf{k})}{\omega_{\rho\lambda}(\mathbf{k})}\right)$$
$$\times\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (37)$$

where $\omega_{\rho\lambda}{}^2(\mathbf{k})$ is the ρth root of the secular equation

$$\omega^2-c^2k^2-\omega_p^2+\alpha_\lambda(\mathbf{k},\omega)E_\mu^2(\mathbf{k})=0. \quad (38)$$

In the limit where $\omega_{\rho\lambda}(\mathbf{k})=E_\mu(\mathbf{k})$, the expressions (35), (36), and (37) are reduced to those of the bare exciton spectrum, i.e.,

$$\langle b_\mu^\dagger(\mathbf{k})b_\mu(\mathbf{k})\rangle_0^{(0)}=\langle b_\mu(\mathbf{k})b_\mu(-\mathbf{k})\rangle_0^{(0)}=0$$

and

$$\langle b_\mu(\mathbf{k})b_\mu^\dagger(\mathbf{k})\rangle_0^{(0)}=(n_{\mathbf{k}0}-n_{\mathbf{k}\mu}).$$

The factor

$$\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}$$

in the expressions for the distribution functions represents the admixture of the field oscillation to each normal mode.[12]

The photon Green's function in the zero approximation is obtained by neglecting the last three terms on the right-hand side of (23), i.e.,

$$D_\lambda^{(0)}(\mathbf{k};\omega)=(ck/\pi)[\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2]\}^{-1}, \quad (39)$$

which has the same poles as $G_\mu^{(0)}(\mathbf{k};\omega)$. We also have

$$\langle\langle\beta_\lambda(\mathbf{k});\beta_\lambda^\dagger(\mathbf{k})\rangle\rangle^{(0)}=\frac{1}{2\pi}\frac{\omega+ck-\omega^2\alpha_\lambda(\mathbf{k},\omega)/2ck}{\omega^2\hat{\eta}_\lambda(\mathbf{k},\omega)-c^2k^2}, \quad (40)$$

$$\langle\langle\beta_\lambda(-\mathbf{k});\beta_\lambda(\mathbf{k})\rangle\rangle^{(0)}=\frac{1}{2\pi}\frac{\omega^2\alpha_\lambda(\mathbf{k},\omega)/2ck}{\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2}. \quad (41)$$

From (39), (40), and (41), we derive the distribution functions:

$$\langle\beta_\lambda^\dagger(\mathbf{k})\beta_\lambda(\mathbf{k})\rangle_0^{(0)}=\sum_\rho\frac{[ck-\omega_{\rho\lambda}(\mathbf{k})]^2}{4ck\omega_{\rho\lambda}(\mathbf{k})}$$
$$\times\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\lambda\rho}(\mathbf{k})}, \quad (42)$$

[12] U. Fano, Phys. Rev. **103**, 1202 (1956).

$$\langle\beta_\lambda(\mathbf{k})\beta_\lambda^\dagger(\mathbf{k})\rangle_0^{(0)} = \sum_\rho \frac{[ck+\omega_{\rho\lambda}(\mathbf{k})]^2}{4ck\omega_{\rho\lambda}(\mathbf{k})} \times \left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (43)$$

$$\langle\beta_\lambda(\mathbf{k})\beta_\lambda(-\mathbf{k})\rangle_0^{(0)} = \frac{1}{4ck}\sum_\rho \omega_{\rho\lambda}(\mathbf{k})\alpha_\lambda[\mathbf{k},\omega_{\rho\lambda}(\mathbf{k})] \times \left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (44)$$

$$\langle\tilde{\beta}_\lambda^\dagger(\mathbf{k})\tilde{\beta}_\lambda(\mathbf{k})\rangle_0^{(0)} = ck\sum_\rho \frac{1}{\omega_{\rho\lambda}(k)}\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (45)$$

$$\langle[\beta_\lambda(\mathbf{k}),\beta_\lambda^\dagger(\mathbf{k})]_-\rangle_0^{(0)} = \sum_\rho \left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}. \quad (46)$$

In the absence of dispersion, i.e., when $\hat{\eta}_\lambda^2(\mathbf{k},\omega)=1$, then $\omega_{\rho\lambda}(\mathbf{k})=ck$,

$$\langle\beta_\lambda^\dagger(\mathbf{k})\beta_\lambda(\mathbf{k})\rangle_0^{(0)} = \langle\beta_\lambda(\mathbf{k})\beta_\lambda(-\mathbf{k})\rangle_0^{(0)} = 0,$$
$$\langle\beta_\lambda(\mathbf{k})\beta_\lambda^\dagger(\mathbf{k})\rangle_0^{(0)} = \langle\tilde{\beta}_\lambda^\dagger(\mathbf{k})\tilde{\beta}_\lambda(\mathbf{k})\rangle_0^{(0)} = 1.$$

It is interesting to note here that if one makes use of Bogolyubov's canonical transformation to diagonalize the Hamiltonian $\mathcal{H}^0+\mathcal{H}_{\rm int}^{\rm I}+\mathcal{H}_{\rm int}^{(2)(0)}+\mathcal{H}_{\rm ph}$, as was done in Refs. 2 and 3, it is easy to show that the amplitudes $u_{\mu\rho}(\mathbf{k})$, $v_{\mu\rho}(\mathbf{k})$, $u_{\rho\lambda}(\mathbf{k})$, and $v_{\rho\lambda}(\mathbf{k})$ of the canonical transformation are related to the distribution functions at zero temperature by

$$\langle b_\mu(\mathbf{k})b_\mu^\dagger(\mathbf{k})\rangle_0^{(0)} = |u_{\mu\rho}(\mathbf{k})|^2,$$
$$\langle b_\mu^\dagger(\mathbf{k})b_\mu(\mathbf{k})\rangle_0^{(0)} = |v_{\mu\rho}(\mathbf{k})|^2, \quad (47a)$$

$$\langle\beta_\lambda(\mathbf{k})\beta_\lambda^\dagger(\mathbf{k})\rangle_0^{(0)} = |u_{\rho\lambda}(\mathbf{k})|^2,$$
$$\langle\beta_\lambda^\dagger(\mathbf{k})\beta_\lambda(\mathbf{k})\rangle_0^{(0)} = |v_{\rho\lambda}(\mathbf{k})|^2, \quad (47b)$$

$$\langle b_\mu(\mathbf{k})b_\mu(-\mathbf{k})\rangle_0^{(0)} = |u_{\mu\rho}(\mathbf{k})v_{\mu\rho}(\mathbf{k})|,$$
$$\langle\beta_\lambda(\mathbf{k})\beta_\lambda(-\mathbf{k})\rangle_0^{(0)} = |u_{\rho\lambda}(\mathbf{k})v_{\rho\lambda}(\mathbf{k})|. \quad (47c)$$

5. GROUND-STATE ENERGY

The ground-state energy of a molecular crystal, correct to the order $N^{-\frac{1}{2}}$, i.e., neglecting contributions from $\mathcal{H}_{\rm int}^{\rm II}$, $\mathcal{H}_{\rm int}^{\rm III}$, and $\mathcal{H}_{\rm int}^{\rm IV}$, is given by

$$\langle\mathcal{H}_1\rangle_0^{(0)} = \langle\mathcal{H}_0\rangle_0 + \sum_{k,\mu} E_\mu(\mathbf{k})\langle b_\mu^\dagger(\mathbf{k})b_\mu(\mathbf{k})\rangle_0^{(0)} + \sum_{k,\lambda} ck\langle\beta_\lambda^\dagger(\mathbf{k})\beta_\lambda(\mathbf{k})\rangle_0^{(0)} + \frac{i\omega_p}{2}\sum_{\mathbf{k},\lambda,\mu} f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda)\left[\frac{E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}} \times \langle[b_\mu(-\mathbf{k})-b_\mu^\dagger(\mathbf{k})]\tilde{\beta}_\lambda^\dagger(\mathbf{k})\rangle_0^{(0)} + \omega_p^2/4\sum_{k,\lambda}\frac{1}{ck}\langle\tilde{\beta}_\lambda^\dagger(\mathbf{k})\tilde{\beta}_\lambda(\mathbf{k})\rangle_0^{(0)}. \quad (48)$$

Using the expression for the Green's function

$$\frac{i\omega_p}{2}\sum_{\mathbf{k},\lambda,\mu} f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda)\left[\frac{E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}}\langle\langle b_\mu^\dagger(\mathbf{k})-b_\mu(-\mathbf{k});\tilde{\beta}_\lambda^\dagger(\mathbf{k})\rangle\rangle^{(0)} = \frac{1}{2\pi}\sum_{\mathbf{k}}\frac{\alpha_\lambda(\mathbf{k},\omega)E_\mu^2(\mathbf{k})}{\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2}, \quad (49)$$

we derive the corresponding expression for the distribution function

$$\frac{i\omega_p}{2}\sum_{\mathbf{k},\lambda,\mu} f_{0\mu}^{\frac{1}{2}}(\mathbf{k},\lambda)\left[\frac{E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}}\langle[b_\mu(-\mathbf{k})-b_\mu^\dagger(\mathbf{k})]\tilde{\beta}_\lambda^\dagger(\mathbf{k})\rangle_0^{(0)} = -\tfrac{1}{2}\sum_{\mathbf{k},\rho}\alpha_\lambda[\mathbf{k},\omega_{\rho\lambda}(\mathbf{k})]\frac{E_\mu^2(\mathbf{k})}{\omega_{\rho\lambda}(\mathbf{k})}\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})} \quad (50)$$

Substituting (35), (42), and (50), into (48) and rearranging, we find

$$\langle\mathcal{H}_1\rangle_0^{(0)} = \langle\mathcal{H}_0\rangle_0 + \tfrac{1}{2}\sum_{\mathbf{k},\rho}[\omega_{\rho\lambda}(\mathbf{k})-E_\mu(\mathbf{k})] + \tfrac{1}{2}\sum_{\mathbf{k},\rho}[E_\mu(\mathbf{k})-ck]\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}, \quad (51)$$

with $\omega_{\rho\lambda}(\mathbf{k})$ given by the roots of the secular equation (38). The formula (51) gives the ground-state energy of a molecular crystal in a closed form exact to the order $N^{-\frac{1}{2}}$. The polarization of the medium resulting from the radiation field of atoms or molecules in the crystal has been fully taken into account in (51). Both the second and third term on the right-hand side of (51) describe the fluctuations of the polarization field in the medium; in the absence of dispersion, i.e., when $\hat{\eta}_\lambda(\mathbf{k},\omega)=1$, $ck=\omega_{\rho\lambda}(\mathbf{k})$ and both terms disappear. In the static case, when $\omega_{\rho\lambda}(\mathbf{k})=0$, (51) becomes

$$\langle\mathcal{H}_1\rangle_0^{(0)} = \langle\mathcal{H}_0\rangle_0 + \tfrac{1}{2}\sum_{k,\mu}E_\mu(\mathbf{k})\left(\frac{1}{\hat{\eta}_\lambda^2(\mathbf{k},0)}-1\right), \quad (52)$$

where the second term on the right-hand side of (52) gives the local field correction to the energy of excitation resulting from the static polarization field. In (52), $\hat{\eta}_\lambda^2(\mathbf{k},0)$ is now the static dielectric constant of the medium. $\langle\mathcal{H}_0\rangle_0$ is given by[1]

$$\langle\mathcal{H}_0\rangle_0 = \tfrac{1}{2}\sum_{\mathbf{k},\mu}[-\Delta_\mu+E_\mu(\mathbf{k})] \quad (53)$$

and Δ_μ is the excitation energy of an isolated molecule. Substitution of (53) into (52) yields

$$\langle\mathcal{H}_1\rangle_0^{(0)} = \tfrac{1}{2}\sum_{\mathbf{k},\mu}\left[-\Delta_\mu+\frac{E_\mu(\mathbf{k})}{\hat{\eta}_\lambda^2(\mathbf{k},0)}\right]. \quad (54)$$

Thus, the effect of the static local polarization field is to reduce the energy of excitation by a factor $1/\hat{\eta}_\lambda^2(\mathbf{k},0)$.

If the density of a crystal is small, $N/V \ll 1$, then one may expand all terms in (51) in power series in the density and the result is an infinite sequence of terms that corresponds to an expansion in perturbation theory for the ground-state energy including retardation. When sums are taken over $\mathbf{k}$, the first nonvanishing retarded term turns out to be proportional to e^4 and R^{-7} for intermolecular distances $R \gg c/\Delta_\mu$, a result which is in agreement with one found in the literature.[13,14] This is, of course, a small contribution to the binding energy of the crystal. But, in the opposite limit of large densities or large polarizabilities, one has to compute (51) for a crystal to find its binding energy. In this case, it is not obvious at first sight to what extent the last two terms of (51) contribute to the binding energy of the crystal unless an actual computation is performed. In deriving (51), we have made no assumption about the strength of the interaction and thus our formalism is more useful for the case where a perturbation expansion is not applicable. The expression (51) will be used in the future to compute binding energies of molecular crystals.

In our calculation, we have neglected higher-order effects which are expressed by the Green's functions that appear in the right-hand side of (21) and (23) for the exciton and photon field, respectively. They describe scattering processes and we shall deal with them in later publications.

ACKNOWLEDGMENT

The author is indebted to Dr. N. Matsudaira for reading the manuscript.

[13] H. B. G. Casimir and D. Polder, Phys. Rev. 73, 360 (1948).
[14] C. Mavroyannis and M. J. Stephen, Mol. Phys. 5, 629 (1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Scattering of Electromagnetic Waves in Molecular Crystals

C. MAVROYANNIS
Division of Pure Chemistry, National Research Council, Ottawa, Canada

(Received 2 December 1966)

The scattering of electromagnetic waves in molecular crystals has been investigated by considering the polarization operator of the system that results from the Dyson equation in the first approximation. The Green's function for the photon field is then calculated and an expression for the frequency- and wavevector-dependent index of refraction of optical waves is developed including nonlinear effects. The excitation spectrum has been studied by evaluating the spectral intensity for the photon field. Expressions are derived for the electronic contribution to the spectral width and energy shift for physical processes: (i) resonance Raman scattering and (ii) resonances occurring when either two excitons are created or one exciton is created and the other is absorbed by a single incident photon. A theory for the dielectric permeability of the crystal is developed and a relation between the dielectric permeability, the polarization operator, and the photon Green's function is established. Contributions to the binding energy of the crystal resulting from the dispersion and scattering of the polarization waves at finite temperatures have been calculated and expressed in terms of the excitation energies and the index of refraction of the medium.

I. INTRODUCTION

AS is well known, the interaction of electromagnetic waves with crystalline matter can be treated either phenomenologically or microscopically. In the first, the characteristic properties of the medium remain undetermined but they are included in the dielectric permeability that appears as a parameter in the theory and may be derived from the experiment. The mathematical simplicity and the usefulness in explaining experimental data are the main points that make this method of great value.

The microscopic theory, on the other hand, consists of diagonalizing the total Hamiltonian of the system, crystal plus electromagnetic field. Its mathematical formulation is much more complicated than that of the phenomenological method but, in the final result, all the properties of the medium are revealed and properly explained. The advantage of this method is that it gives full physical insight into what happens in the crystal when the electromagnetic field acts upon it and answers the question for the importance of the interactions involved. The

If the density of a crystal is small, $N/V \ll 1$, then one may expand all terms in (51) in power series in the density and the result is an infinite sequence of terms that corresponds to an expansion in perturbation theory for the ground-state energy including retardation. When sums are taken over $\mathbf{k}$, the first nonvanishing retarded term turns out to be proportional to e^4 and R^{-7} for intermolecular distances $R \gg c/\Delta_\mu$, a result which is in agreement with one found in the literature.[13,14] This is, of course, a small contribution to the binding energy of the crystal. But, in the opposite limit of large densities or large polarizabilities, one has to compute (51) for a crystal to find its binding energy. In this case, it is not obvious at first sight to what extent the last two terms of (51) contribute to the binding energy of the crystal unless an actual computation is performed. In deriving (51), we have made no assumption about the strength of the interaction and thus our formalism is more useful for the case where a perturbation expansion is not applicable. The expression (51) will be used in the future to compute binding energies of molecular crystals.

In our calculation, we have neglected higher-order effects which are expressed by the Green's functions that appear in the right-hand side of (21) and (23) for the exciton and photon field, respectively. They describe scattering processes and we shall deal with them in later publications.

ACKNOWLEDGMENT

The author is indebted to Dr. N. Matsudaira for reading the manuscript.

[13] H. B. G. Casimir and D. Polder, Phys. Rev. 73, 360 (1948).
[14] C. Mavroyannis and M. J. Stephen, Mol. Phys. 5, 629 (1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Scattering of Electromagnetic Waves in Molecular Crystals

C. MAVROYANNIS
Division of Pure Chemistry, National Research Council, Ottawa, Canada

(Received 2 December 1966)

The scattering of electromagnetic waves in molecular crystals has been investigated by considering the polarization operator of the system that results from the Dyson equation in the first approximation. The Green's function for the photon field is then calculated and an expression for the frequency- and wavevector-dependent index of refraction of optical waves is developed including nonlinear effects. The excitation spectrum has been studied by evaluating the spectral intensity for the photon field. Expressions are derived for the electronic contribution to the spectral width and energy shift for physical processes: (i) resonance Raman scattering and (ii) resonances occurring when either two excitons are created or one exciton is created and the other is absorbed by a single incident photon. A theory for the dielectric permeability of the crystal is developed and a relation between the dielectric permeability, the polarization operator, and the photon Green's function is established. Contributions to the binding energy of the crystal resulting from the dispersion and scattering of the polarization waves at finite temperatures have been calculated and expressed in terms of the excitation energies and the index of refraction of the medium.

I. INTRODUCTION

AS is well known, the interaction of electromagnetic waves with crystalline matter can be treated either phenomenologically or microscopically. In the first, the characteristic properties of the medium remain undetermined but they are included in the dielectric permeability that appears as a parameter in the theory and may be derived from the experiment. The mathematical simplicity and the usefulness in explaining experimental data are the main points that make this method of great value.

The microscopic theory, on the other hand, consists of diagonalizing the total Hamiltonian of the system, crystal plus electromagnetic field. Its mathematical formulation is much more complicated than that of the phenomenological method but, in the final result, all the properties of the medium are revealed and properly explained. The advantage of this method is that it gives full physical insight into what happens in the crystal when the electromagnetic field acts upon it and answers the question for the importance of the interactions involved. The

conventional perturbation theory, which is applicable for the interaction of the electromagnetic field with a rarefied gas, is not valid in the case of interaction with a condensed medium. This is so because, although for certain frequencies the effect of the electromagnetic field on the medium may be regarded as a small perturbation, the reciprocal effect of the condensed medium on the electromagnetic field cannot be regarded as weak and accounts for the deviations that occurred from its behavior in vacuum. Thus, for an accurate description of the scattering of electromagnetic waves in crystals, the scattering amplitudes must be expressed in terms of quantities that have been calculated exactly in the zero approximation.

In an other paper,[1] which we refer to here as I, the crystal Hamiltonian for a molecular crystal plus the Hamiltonian for the electromagnetic field has been diagonalized exactly in the zero approximation and expressions for the excitation spectrum, index of refraction, and distribution functions for the exciton and photon field, respectively, as well as for the ground state energy of the crystal have been developed and discussed. The calculation is based on the tight-binding approximation, where the photons of the electromagnetic field are acting upon electron–hole pairs that are tightly bound at the lattice sites. Such a model is appropriate for describing the excitation spectrum of a molecular crystal.

The present study, which is a continuation of I, is concerned with the scattering of electromagnetic waves in molecular crystals. The theory is developed in Sec. II, where the Dyson equation is derived for the exciton and photon field, respectively, and the corresponding Green's functions have been expressed in terms of the polarization operators. Formulas for the exciton Green's functions are obtained by considering the polarization operator in the first approximation that corresponds to the physical process where an exciton decays into two excitons. The polarization operator for the photon field has been expressed in terms of the exciton Green's functions. The results obtained in the zero approximation are in agreement with those found in I.

In Sec. III the theory is applied to the physical process, where a dressed exciton $(\mathbf{k}, \mu)$, with wavevector $\mathbf{k}$ in the μth excitation band, decays into two excitons: a bare exciton $(\mathbf{k} - \mathbf{q}, \mu')$ and a dressed one $(\mathbf{q}, \mu_1)$ with wavevectors $\mathbf{k} - \mathbf{q}$ and $\mathbf{q}$ in the μ'th and μ_1th excitation bands, respectively. All formulas derived in the zero approximation are now renormalized to take account of the interactions. Expressions are derived for the renormalized energy of excitation and the exciton and photon Green's functions, respectively. The expression for the perturbed index of refraction consists of three terms: the usual linear polarizability, the Raman polarizability tensor, and a cooperative polarizability arising from the correlation of the radiation field with the intermolecular interactions in the crystal and proportional to the sixth power of the electronic charge. In these expressions, quantities referring to the excitons $(\mathbf{k}, \mu)$ and $(\mathbf{q}, \mu_1)$, $(\mathbf{k} - \mathbf{q}, \mu')$ are correct in the first and zero approximation, respectively, with the dispersion for the dressed exciton $(\mathbf{q}, \mu_1)$ taken into account explicitly.

The spectral intensity for the photon field is evaluated (by considering the imaginary part of the photon Green's function in the first approximation) and the photon excitation spectrum is discussed. Formulas are developed for the energy shift and spectral width corresponding to the process of resonance Raman scattering. The resonance processes, occurring when either two excitons are created or one exciton is created and the other is absorbed by a single photon, are studied and expressions for their energy shifts and spectral widths are established. It is found that the spectral widths for the processes in question depend on the polarizabilities and the spontaneous emission probabilities for the transitions under consideration.

A theory of the dielectric permeability for a molecular crystal has been developed in Sec. IV, where a general relation between the dielectric permeability, the polarization operator of the system, and the photon Green's function is established. It is shown that, if only the linear term in the polarization operator is retained, the relation becomes identical with that derived by Dzyaloshinskii *et al.*[2] by means of the diagram technique. There is also a term proportional to the square of the polarization operator which, in the final result, accounts for the difference between the index of refraction and the dielectric permeability of the crystal. The transverse and longitudinal dielectric permeabilities of the crystal have been calculated in the zero and first approximation, respectively. It is shown that the expression for the imaginary part of the transverse dielectric permeability in the first approximation describes, under the same conditions, the same excitation spectrum as that given by the spectral intensity for the photon field.

[1] C. Mavroyannis, J. Math. Phys. **8**, 1515 (1967).

[2] I. Ye. Dzyaloshinskii, E. M. Lifshitz, and L. P. Pitaevskii, Advan. Phys. **10**, 165 (1965); A. A. Abrikosov, L. P. Gorkov, and I. Ye. Dzyaloshinskii, *Quantum Field Theoretical Methods in Statistical Physics* (Pergamon Press, Ltd., London 1965), 2nd ed., p. 260.

Finally, in Sec. V, the total Hamiltonian of the crystal is averaged in the first approximation. Contributions to the binding energy of the crystal arising from the dispersion and scattering of the polarization waves at finite temperatures are calculated and discussed. They have been expressed in terms of the index of refraction and the excitation energies of both the bare and dressed exciton $(\mathbf{k}, \mu)$.

The effects on the physical processes of the exciton–phonon interaction and the scattering by impurities discussed here will be the subject of a later publication, where the temperature dependence of the dielectric permeability is investigated as well.

II. THEORY

The total Hamiltonian for a molecular crystal, where the molecules are rigidly attached to the lattice sites, has been given in I in the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm ph} + \mathcal{H}_{\rm int}^{\rm I} + \mathcal{H}_{\rm int}^{\rm II} + \mathcal{H}_{\rm int}^{(2)(0)}, \quad (1)$$

where

$$\mathcal{H}_0 = \langle \mathcal{H}_0 \rangle + \sum_{\mathbf{k},\mu} E_\mu(\mathbf{k}) b_\mu^\dagger(\mathbf{k}) b_\mu(\mathbf{k}) + \mathcal{H}_{\rm int}^{\rm III} + \mathcal{H}_{\rm int}^{\rm IV}, \quad (2)$$

$$\mathcal{H}_{\rm ph} = \sum_{\mathbf{k},\lambda} ck \beta_\lambda^\dagger(\mathbf{k}(\beta_\lambda(\mathbf{k}), \quad (3)$$

$$\mathcal{H}_{\rm int}^{(2)(0)} = \frac{\omega_p^2}{4} \sum_{\mathbf{k},\lambda} \frac{1}{ck} \tilde{\beta}_\lambda^\dagger(\mathbf{k}) \tilde{\beta}_\lambda(\mathbf{k}), \quad (4)$$

$$\mathcal{H}_{\rm int}^{\rm I} = \frac{i\omega_p}{2} \sum_{\mathbf{k},\lambda,\mu} \left[\frac{f_{0\mu}(\mathbf{k}, \lambda) E_\mu(\mathbf{k})}{ck} \right]^{\frac{1}{2}} \tilde{b}_\mu(\mathbf{k}) \tilde{\beta}_\lambda(\mathbf{k}), \quad (5)$$

$$\mathcal{H}_{\rm int}^{\rm II} = \frac{i\omega_p}{N^{\frac{1}{2}}} \sum_{\substack{\mathbf{k},\mathbf{q},\lambda \\ \mu,\mu'}} [f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q}, \lambda) E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu/cq}]^{\frac{1}{2}} \times b_\mu^\dagger(\mathbf{k}) b_{\mu'}(\mathbf{k} - \mathbf{q}) \tilde{\beta}_\lambda(\mathbf{q}) \quad (6)$$

$$\mathcal{H}_{\rm int}^{\rm III} = 1/N^{\frac{1}{2}} \sum_{\substack{\mathbf{k},\mathbf{q},\beta \\ \mu,\mu',\mu_1}} U_{0\mu,\mu'\mu_1}^{\alpha\beta}(\mathbf{k} - \mathbf{q}) b_\mu(\mathbf{k}) b_{\mu_1}^\dagger(\mathbf{q}) \times b_{\mu'}^\dagger(\mathbf{k} - \mathbf{q}) + \text{c.c.}, \quad (7)$$

$$\mathcal{H}_{\rm int}^{\rm IV} = \frac{2}{N} \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}' \\ \mu,\mu',\mu_1,\mu_1',\beta}} U_{\mu\mu',\mu_1\mu_1'}^{\alpha\beta}(\mathbf{k} - \mathbf{k}') b_\mu^\dagger(\mathbf{k}) \times b_{\mu_1}(\mathbf{k}') b_{\mu'}^\dagger(\mathbf{q}) b_{\mu_1'}(\mathbf{q}') \delta_{\mathbf{k}+\mathbf{q},\mathbf{k}'+\mathbf{q}'}, \quad (8)$$

$$\tilde{b}_\mu(\mathbf{k}) \equiv b_\mu(-\mathbf{k}) - b_\mu^\dagger(\mathbf{k})$$

and

$$\tilde{\beta}_\lambda(\mathbf{k}) \equiv \beta_\lambda(\mathbf{k}) + \beta_\lambda^\dagger(-\mathbf{k}), \quad \hbar \equiv 1.$$

$E_\mu(\mathbf{k})$ is the energy of excitation of the bare exciton $(\mathbf{k}, \mu)$, $\mathbf{k}$ is a wavevector in the first Brillouin zone, and μ is a compound index that indicates the exciton band, the corresponding molecular term and the kind of mode, transverse $(\mu_\perp)$ or longitudinal $(\mu_\parallel)$. The quantities $b_\mu^\dagger(\mathbf{k})$, $b_\mu(\mathbf{k})$ and $\beta_\lambda^\dagger(\mathbf{k})$, $\beta_\lambda(\mathbf{k})$ are the creation and annihilation operators for the exciton and photon field, respectively. The photon operators satisfy Bose statistics while the exciton operators satisfy the commutation relation

$$[b_\mu(\mathbf{k}), b_{\mu'}^\dagger(\mathbf{k}')]_- = (n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \delta_{\mathbf{k}\mathbf{k}'} \delta_{\mu\mu'},$$

where $n_{\mathbf{k}0} \equiv \langle \alpha_{\mathbf{k}0}^\dagger \alpha_{\mathbf{k}0} \rangle$ and $n_{\mathbf{k}\mu} \equiv \langle \alpha_{\mathbf{k}\mu}^\dagger \alpha_{\mathbf{k}\mu} \rangle$ are the occupation numbers for the hole (valence band) and for the electron in the $(\mathbf{k}, \mu)$ excitation band, respectively. The quantity $\langle \mathcal{H}_0 \rangle$ is the average energy due to direct electron–hole pair interactions correct to the order $N^{-\frac{1}{2}}$, where N is the total number of unit cells in the crystal; its expression is given elsewhere.[3] In (3)–(8) the coupling constants are given by

$$f_{0\mu}^{\frac{1}{2}}(\mathbf{k}, \lambda) = \left(\frac{2m}{e^2 \sigma S} \right)^{\frac{1}{2}} \sum_\alpha (\mathbf{e}_{\mathbf{k}\lambda} \cdot \mathbf{P}_{0\mu}^\alpha) E_\mu^{\frac{1}{2}}(\mathbf{k}) \times [u_{\mu\alpha}(\mathbf{k}) + v_{\mu\alpha}(\mathbf{k})] \exp(i\mathbf{k} \cdot \mathbf{r}_{\mathbf{n}\alpha}), \quad (9)$$

$$f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}^{\frac{1}{2}}(\mathbf{q}, \lambda) = \left(\frac{2m}{e^2 \sigma S} \right)^{\frac{1}{2}} \sum_\alpha (\mathbf{e}_{\mathbf{q}\lambda} \cdot \mathbf{P}_{\mu\mu'}^\alpha) E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}^{\frac{1}{2}} \times u_{\mu\alpha}^*(\mathbf{k}) u_{\mu'\alpha}(\mathbf{k} - \mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_{\mathbf{n}\alpha}), \quad (10)$$

$$U_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q}) = \sum_{s_1 \neq s} \langle s0, s_1\mu | V_{ss_1} | s\mu', s_1\mu_1 \rangle \times u_\mu(\mathbf{k}) u_{\mu'}^*(\mathbf{q}) u_{\mu'}(\mathbf{k} - \mathbf{q}) \times \exp[i(\mathbf{k} - \mathbf{q}) \cdot (\mathbf{r}_s - \mathbf{r}_{s_1})], \quad (11)$$

$$E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu} \equiv E_{\mu'}(\mathbf{k} - \mathbf{q}) - E_\mu(\mathbf{k}).$$

The u's and v's are the amplitudes of the canonical transformation that diagonalize the unperturbed part of the Hamiltonian. In (6)–(8) we have retained only the u's while the v's, being much smaller than the u's, have been neglected. The lattice sites where the molecules α and β are located are $s \equiv \mathbf{n}\alpha$ and $s_1 \equiv \mathbf{m}\beta$; S and σ indicate the number of electrons and molecules in the unit cell, respectively. In the formulas (1)–(11) we have used the same notation as in I, where details are given. We have included here the expression (8), which was omitted in I, where the factor of 2 arises from the transformation of electron operators to the corresponding exciton operators; i.e., if $\alpha_{sf}^\dagger$ and $\alpha_{sf'}$ are the electron operators for the states f and $f'(f \neq f')$, respectively, then $\alpha_{sf}^\dagger \alpha_{sf'} = 2b_{sf}^\dagger b_{sf'}$ with $b_{sf'} = \alpha_{s0}^\dagger \alpha_{sf'}$. In (1) we have retained only $\mathcal{H}_{\rm int}^{(2)(0)}$ and neglected the higher-order terms which give rise to photon–photon scattering processes because these terms are important only in the x-ray region of frequencies.[4]

[3] C. Mavroyannis, J. Chem. Phys. **42**, 1772 (1965).
[4] L. N. Ovander, Fiz. Tverd. Tela **3**, 2394 (1961) [English transl.: Soviet Phys.—Solid State **3**, 1737 (1962)].

The equation of motion for the retarded Green's function

$$\tilde{G}_\mu(\mathbf{k};\omega) \equiv \langle\langle A_\mu(\mathbf{k}); A_\mu^\dagger(\mathbf{k})\rangle\rangle \equiv \left\langle\!\!\left\langle \begin{pmatrix} b_\mu(\mathbf{k}) \\ b_\mu^\dagger(-\mathbf{k}) \end{pmatrix}; (b_\mu^\dagger(\mathbf{k}) b_\mu(-\mathbf{k})) \right\rangle\!\!\right\rangle$$

has been derived in I in the form

$$\begin{aligned}[\hat{\alpha}\omega - E_\mu(\mathbf{k}) - 2/N(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})\textstyle\sum_{\mathbf{q},\mu',\beta} U^{\alpha\beta}_{\mu\mu',\mu'\mu}(\mathbf{q}) \\ \times (n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'})]\tilde{G}_\mu(\mathbf{k};\omega) \\ = (I/2\pi)(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) + \langle\langle B(\mathbf{k},\mu); A_\mu^\dagger(\mathbf{k})\rangle\rangle \\ + \langle\langle \Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu'); A_\mu^\dagger(\mathbf{k})\rangle\rangle + \langle\langle \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}}; A_\mu^\dagger(\mathbf{k})\rangle\rangle \\ + \langle\langle W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}; A_\mu^\dagger(\mathbf{k})\rangle\rangle, \end{aligned} \tag{12}$$

where

$$B(\mathbf{k},\mu) \equiv -\frac{i\omega_p}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \times \sum_\lambda \left[\frac{f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})}{ck}\right]^{\frac{1}{2}} \begin{pmatrix} \tilde{\beta}_\lambda(\mathbf{k}) \\ -\tilde{\beta}_\lambda(\mathbf{k}) \end{pmatrix}, \tag{13}$$

$$\begin{aligned}\Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu') &\equiv \frac{i\omega_p}{N^{\frac{1}{2}}}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\lambda,\mu'} [f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q},\lambda)E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu/c\mathbf{q}}]^{\frac{1}{2}} \\ &\times \begin{pmatrix} b_{\mu'}(\mathbf{k}-\mathbf{q})\tilde{\beta}_\lambda(\mathbf{q}) \\ -b_{\mu'}^\dagger(\mathbf{k}-\mathbf{q})\tilde{\beta}_\lambda^\dagger(\mathbf{q}) \end{pmatrix}, \end{aligned} \tag{14a}$$

$$\hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} \equiv \frac{(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})}{N^{\frac{1}{2}}} \begin{pmatrix} \hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} \\ \hat{U}^{\alpha\beta\dagger}_{-\mathbf{k},-\mathbf{q}} \end{pmatrix}, \tag{14b}$$

$$\begin{aligned}\hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} &\equiv \sum_{\mathbf{q},\mu',\mu_1,\beta} \{[U^{\alpha\beta}_{0\mu',\mu_1\mu}(-\mathbf{q}) + U^{\alpha\beta}_{0\mu',\mu\mu_1}(\mathbf{k})]b_{\mu_1}^\dagger(-\mathbf{q}) \\ &+ U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k}-\mathbf{q})b_{\mu_1}(\mathbf{q})\}b_{\mu'}(\mathbf{k}-\mathbf{q}), \end{aligned} \tag{14c}$$

$$W_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'} \equiv \frac{2}{N}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \begin{pmatrix} \hat{W}^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'} \\ \hat{W}^{\alpha\beta\dagger}_{-\mathbf{k},-\mathbf{k}',\mathbf{q}-,-\mathbf{q}'} \end{pmatrix}, \tag{14d}$$

$$\begin{aligned}\hat{W}^{\alpha\beta}_{\mathbf{k},\mathbf{k},'\mathbf{q},\mathbf{q}'} &\equiv \sum_{\substack{\mathbf{k}',\mathbf{q},\mathbf{q}',\beta \\ \mu',\mu_1,\mu_1'}} [U^{\alpha\beta}_{\mu\mu',\mu_1\mu_1'}(\mathbf{k}-\mathbf{k}') \\ &+ U_{\mu'\mu,\mu_1\mu_1'}(\mathbf{q}-\mathbf{k}')] \\ &\times b_{\mu'}^\dagger(\mathbf{q})b_{\mu_1}(\mathbf{k}')b_{\mu_1'}(\mathbf{q}')\delta_{\mathbf{q}+\mathbf{k},\mathbf{q}'+\mathbf{k}'}, \end{aligned} \tag{14e}$$

$$\hat{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In (12), we have included the term resulting from $\mathcal{H}_{\text{int}}^{\text{IV}}$. The equation of motion for the Green's function $\langle\langle B(\mathbf{k},\mu); A_\mu^\dagger(\mathbf{k})\rangle\rangle$ is given by

$$\begin{aligned}\langle\langle B(\mathbf{k},\mu); A_\mu^\dagger(\mathbf{k})\rangle\rangle &= \frac{\omega_p^2}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mu,\lambda} \frac{f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})}{(\omega^2 - c^2k^2 - \omega_p^2)} \hat{\sigma}_3 \tilde{G}_\mu(k;\omega) \\ &+ \langle\langle C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu'); A_\mu^\dagger(\mathbf{k})\rangle\rangle, \end{aligned} \tag{15}$$

where

$$\begin{aligned}C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu') &\equiv (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})\frac{\omega_p^2}{N^{\frac{1}{2}}} \\ &\times \sum_{\mathbf{q},\lambda,\mu_1,\mu'} [f_{0\mu}(\mathbf{k},\lambda)f_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k},\lambda)E_\mu(\mathbf{k})E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{\frac{1}{2}} \\ &\times (\omega^2 - c^2k^2 - \omega_p^2)^{-1} \begin{pmatrix} b_{\mu_1}^\dagger(-\mathbf{q})b_{\mu'}(\mathbf{k}-\mathbf{q}) \\ -b_{\mu_1}^\dagger(-\mathbf{q})b_{\mu'}(\mathbf{k}-\mathbf{q}) \end{pmatrix}, \\ \hat{\sigma}_3 &= \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \end{aligned} \tag{16}$$

Substitution of (15) into (12) yields the following equation:

$$\begin{aligned}\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega)\tilde{G}_\mu(\mathbf{k};\omega) &= \frac{I}{2\pi}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) + \langle\langle \Gamma_{\mathbf{q},\lambda}(\mathbf{k}-\mathbf{q},\mu') + C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu') \\ &+ \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} + W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}; A_\mu^\dagger(\mathbf{k})\rangle\rangle, \end{aligned} \tag{17}$$

where

$$\begin{aligned}\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) &= \hat{\alpha}\omega - E_\mu(\mathbf{k}) - \frac{\omega_p^2}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mu,\lambda} \frac{f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})}{(\omega^2 - c^2k^2 - \omega_p^2)} \hat{\sigma}_3 \\ &- \frac{2}{N}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\mu',\beta} U^{\alpha\beta}_{\mu\mu',\mu'\mu}(\mathbf{q})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'}) \end{aligned} \tag{18}$$

is the unperturbed Green's function for the exciton spectrum. The last term on the right-hand side of (18) describes the direct interaction between the bare excitons $(\mathbf{k},\mu)$ and $(\mathbf{k}-\mathbf{q},\mu')$. To proceed further, we have to consider the equations of motion for the Green's functions that appear on the right-hand side of (17), which are found to be

$$\begin{aligned}\langle\langle \Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu'); A_\mu^\dagger(\mathbf{k})\rangle\rangle\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) &= \langle\langle \Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu'); \Gamma_{\mathbf{q}'\lambda'}^\dagger(\mathbf{k}-\mathbf{q}',\mu'') \\ &+ C^\dagger(\mathbf{q}'\mu_1', \mathbf{k}-\mathbf{q}',\mu'') + \hat{V}^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'} + W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'}\rangle\rangle, \end{aligned} \tag{19}$$

$$\begin{aligned}\langle\langle C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu'); A_\mu^\dagger(\mathbf{k})\rangle\rangle\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) &= \langle\langle C(\mathbf{q}\mu_1, \mathbf{k}-\mathbf{q}\mu'); C^\dagger(\mathbf{q}'\mu_1', \mathbf{k}-\mathbf{q}'\mu'') \\ &+ \Gamma_{\mathbf{q}'\lambda'}^\dagger(\mathbf{k}-\mathbf{q}',\mu'') + \hat{V}^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'} + W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'}\rangle\rangle, \end{aligned} \tag{20}$$

$$\begin{aligned}\langle\langle V^{\alpha\beta}_{\mathbf{k},\mathbf{q}}; A_\mu^\dagger(\mathbf{k})\rangle\rangle\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) &= \langle\langle \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}}; \hat{V}^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'} + \Gamma_{\mathbf{q}'\lambda'}^\dagger(\mathbf{k}-\mathbf{q}'\mu'') \\ &+ C^\dagger(\mathbf{q}'\mu', \mathbf{k}-\mathbf{q}'\mu'') + W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'}\rangle\rangle, \end{aligned} \tag{21}$$

$$\begin{aligned}\langle\langle W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}; A_\mu^\dagger(\mathbf{k})\rangle\rangle\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) &= \frac{1}{2\pi}\langle [W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}, A_\mu^\dagger(\mathbf{k})]_-\rangle \\ &+ \langle\langle W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}; W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'} + \Gamma_{\mathbf{q}'\lambda'}^\dagger(\mathbf{k}-\mathbf{q}',\mu'') \\ &+ C^\dagger(\mathbf{q}'\mu_1', \mathbf{k}-\mathbf{q}'\mu'') + V^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'}\rangle\rangle. \end{aligned} \tag{22}$$

In deriving (19)–(22), we have made the assumption that the average values of the single operators $\langle b_\mu^\dagger(\mathbf{k})\rangle$, $\langle b_\mu(\mathbf{k})\rangle$, $\langle \beta_\lambda^\dagger(\mathbf{k})\rangle$, and $\langle \beta_\lambda(\mathbf{k})\rangle$ are equal to zero.

Substituting expressions (19)–(22) into (17), we

derive Dyson's equation

$$\begin{aligned}\tilde{G}_\mu(\mathbf{k};\omega) &= \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega) + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)P_1(\mathbf{k};\omega)\tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\\ &= \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega) + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\tilde{P}(\mathbf{k};\omega)\tilde{G}_\mu(\mathbf{k};\omega),\end{aligned} \tag{23}$$

where

$$\begin{aligned}P_1(\mathbf{k};\omega) = \left(\frac{2\pi}{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})}\right)^2\Big\{&\langle\langle\Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu')\\ &+ C(\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu') + \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} + W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'};\\ &\quad \Gamma^\dagger_{\mathbf{q}'\lambda'}(\mathbf{k}-\mathbf{q}',\mu'') + C^\dagger(\mathbf{q}'\mu',\mathbf{k}-\mathbf{q}'\mu'')\\ &+ \hat{V}^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'} + W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'}\rangle\rangle\\ &+ \frac{1}{(2\pi)}\langle[W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'}, A^\dagger_\mu(\mathbf{k})]_-\rangle\Big\},\end{aligned} \tag{24}$$

and the polarization operator, $\tilde{P}(\mathbf{k};\omega)$, is given by

$$\tilde{P}(\mathbf{k};\omega) = P_1(\mathbf{k};\omega)[1 + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)P_1(\mathbf{k};\omega)]^{-1}. \tag{25}$$

Then (23) may be written as

$$[\tilde{G}_\mu^{(0)-1}(\mathbf{k},\lambda;\omega) - \tilde{P}(\mathbf{k};\omega)]\tilde{G}_\mu(\mathbf{k};\omega) = (n_{\mathbf{k}0}-n_{\mathbf{k}\mu})I/2\pi. \tag{26}$$

In general, the polarization operator is a complex quantity and may be taken as

$$\tilde{P}(\mathbf{k};\omega) = \operatorname{Re}\tilde{P}(\mathbf{k};\omega) + i\operatorname{Im}\tilde{P}(\mathbf{k};\omega), \tag{27}$$

where

$$\operatorname{Re}\tilde{P}(\mathbf{k};\omega) = \frac{\operatorname{Re}P_1(\mathbf{k};\omega) + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\{[\operatorname{Re}P_1(\mathbf{k};\omega)]^2 + [\operatorname{Im}P_1(\mathbf{k};\omega)]^2\}}{[1+\tilde{G}_\mu^{(0)}(\mathbf{k};\lambda;\omega)\operatorname{Re}P_1(\mathbf{k};\omega)]^2 + [\tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\operatorname{Im}P_1(\mathbf{k};\omega)]^2}, \tag{28}$$

$$\operatorname{Im}\tilde{P}(\mathbf{k};\omega) = \frac{\operatorname{Im}P_1(\mathbf{k};\omega)}{[1+\tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\operatorname{Re}P_1(\mathbf{k};\omega)]^2 + [\tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\operatorname{Im}P_1(\mathbf{k};\omega)]^2}. \tag{29}$$

We may now distinguish two types of spectra: If $\omega^2 = \omega_r^2(\mathbf{k})$ are the roots of the equation

$$I + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)\operatorname{Re}P_1(\mathbf{k};\omega) = 0, \tag{30}$$

then, in the neighborhood of these frequencies $\omega^2 \sim \omega_r^2(\mathbf{k})$, the function $\operatorname{Im}\tilde{P}(\mathbf{k};\omega)$ is a Lorentzian line if $\operatorname{Re}P_1(\mathbf{k};\omega)$ and $\operatorname{Im}P_1(\mathbf{k};\omega)$ vary slowly with ω. At $\omega^2 = \omega_r^2(\mathbf{k})$, (28) and (29) are reduced to

$$\operatorname{Re}\tilde{P}[\mathbf{k};\omega_r(\mathbf{k})] = \tilde{G}_\mu^{(0)-1}[\mathbf{k},\lambda;\omega_r(\mathbf{k})], \tag{31}$$

$$\operatorname{Im}\tilde{P}^{-1}[\mathbf{k};\omega_r(\mathbf{k})] = \tilde{G}_\mu^{(0)}[\mathbf{k},\lambda;\omega_r(\mathbf{k})]\operatorname{Im}P_1[\mathbf{k};\omega_r(k)]\tilde{G}_\mu^{(0)}[\mathbf{k},\lambda;\omega_r(\mathbf{k})], \tag{32}$$

respectively. Substitution of (31) and (32) into (26) yields

$$\operatorname{Re}\tilde{G}_\mu[k;\omega_r(\mathbf{k})] = 0, \tag{33}$$

while

$$\operatorname{Im}\tilde{G}_\mu[\mathbf{k};\omega_r(\mathbf{k})] = \tilde{G}_\mu^{(0)}[\mathbf{k},\lambda;\omega_r(\mathbf{k})]\operatorname{Im}P_1[\mathbf{k};\omega_r(\mathbf{k})]\tilde{G}_\mu^{(0)}[\mathbf{k},\lambda;\omega_r(k)]. \tag{34}$$

Then, if $\operatorname{Im}P_1[\mathbf{k};\omega_r(\mathbf{k})] \ll \omega_r^2(\mathbf{k})$, the system resonates at frequencies $\omega^2 = \omega_r^2(\mathbf{k})$. The spectrum of (34) corresponds to the strong coupling case and could be attributed to the spectrum of localized or trapped excitons. In the limit when $\operatorname{Im}P_1(\mathbf{k};\omega)\to 0$, then the expression (29) becomes a delta function, i.e.,

$$\operatorname{Im}\tilde{P}(\mathbf{k};\omega) \approx \pi\delta[1 + \tilde{G}_\mu^{(0)}(\mathbf{k},\lambda;\omega)P_1(\mathbf{k};\omega)]. \tag{35}$$

The second type of spectrum that is of interest to us is that of free exciton which occurs at frequencies ω, where

$$1 + \tilde{G}_\mu^{(0)}(\mathbf{k};\lambda;\omega)P_1(\mathbf{k};\omega) \neq 0. \tag{36}$$

When (36) holds, we expand the denominator in (25) in a power series of the coupling constant and, retaining the first nonvanishing term, we have

$$\tilde{P}(\mathbf{k};\omega) \approx P_1^{(1)}(\mathbf{k};\omega), \tag{37}$$

where

$$\begin{aligned}P_1^{(1)}(\mathbf{k};\omega) = \left[\frac{2\pi}{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})}\right]\Big\{&\langle\langle\Gamma_{\mathbf{q}\lambda}(\mathbf{k}-\mathbf{q},\mu')\\ &+ C(\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu') + \hat{V}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} + W^{\alpha\beta}_{\mathbf{k},\mathbf{k}',\mathbf{q},\mathbf{q}'};\\ &\quad \Gamma^\dagger_{\mathbf{q}'\lambda'}(\mathbf{k}-\mathbf{q}',\mu'') + C^\dagger(\mathbf{q}'\mu_1',\mathbf{k}-\mathbf{q}'\mu'')\\ &+ \hat{V}^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{q}'} + W^{\alpha\beta\dagger}_{\mathbf{k},\mathbf{k}'',\mathbf{q}'',\mathbf{q}_1'}\rangle\rangle^{(0)}\\ &+ \frac{1}{2\pi}\langle[W^{\alpha\beta}_{\mathbf{k},\mathbf{k}'\mathbf{q},\mathbf{q}'}, A^\dagger_\mu(\mathbf{k})]_-\rangle^{(0)}\Big\}.\end{aligned} \tag{38}$$

The superscript (0) means that the Green's functions must be evaluated in the zero approximation, i.e., disregarding the interaction terms in the total Hamiltonian but including dispersion. Taking the diagonal, $G_\mu^{(1)}(\mathbf{k};\omega) \equiv \langle\langle b_\mu(\mathbf{k}); b_\mu^\dagger(\mathbf{k})\rangle\rangle^{(1)}$, and nondiagonal, $\hat{G}_\mu^{(1)}(\mathbf{k};\omega) \equiv \langle\langle b_\mu^\dagger(-\mathbf{k}); b_\mu^\dagger(\mathbf{k})\rangle\rangle^{(1)}$, matrix elements of (26), we find

$$\begin{aligned}G_\mu^{(1)}(\mathbf{k};\omega) = &\frac{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})}{2\pi}[\omega + E_\mu(\mathbf{k}) + P_{\mathbf{k}\mu}(-\omega)]\\ &\times\left\{\left[\omega - \frac{P_{\mathbf{k}\mu}(\omega)-P_{\mathbf{k}\mu}(-\omega)}{2}\right]^2\right.\\ &\left.- \xi_{\mathbf{k}\mu}^{(1)}(\omega)\left[E_\mu(\mathbf{k}) + \frac{P_{\mathbf{k}\mu}(\omega)+P_{\mathbf{k}\mu}(-\omega)}{2} - \hat{P}'_{\mathbf{k}\mu}(\omega)\right]\right\}^{-1},\end{aligned} \tag{39}$$

$$\hat{G}_{\mu}^{(1)}(\mathbf{k};\omega) = -\frac{(n_{k0}-n_{k\mu})}{2\pi}\hat{P}'_{k\mu}(\omega) \times \left\{\left[\omega - \frac{P_{k\mu}(\omega)-P_{k\mu}(-\omega)}{2}\right]^2 - \mathcal{E}_{k\mu}^{(1)}(\omega)\left[E_{\mu}(\mathbf{k}) + \frac{P_{k\mu}(\omega)+P_{k\mu}(-\omega)}{2} - \hat{P}'_{k\mu}(\omega)\right]\right\}^{-1}, \tag{40}$$

where

$$\mathcal{E}_{k\mu}^{(1)}(\omega) = E_{\mu}(\mathbf{k}) + \frac{P_{k\mu}(\omega)+P_{k\mu}(-\omega)}{2} + \hat{P}'_{k\mu}(\omega), \tag{41}$$

$$2\hat{P}'_{k\mu}(\omega) = \hat{P}_{k\mu}(\omega) + \hat{P}_{k\mu}(-\omega), \tag{42}$$

$$P_{k\mu}(\pm\omega) = \frac{\omega_p^2}{2}(n_{k0}-n_{k\mu})\sum_{\mu,\lambda}\frac{f_{0\mu}(\mathbf{k},\lambda)E_{\mu}(\mathbf{k})}{(\omega^2-c^2k^2-\omega_p^2)} + \frac{2}{N}(n_{k0}-n_{k\mu})\sum_{\mathbf{q},\mu',\beta}U_{\mu\mu',\mu'\mu}^{\alpha\beta}(\mathbf{q}) \times (n_{\mathbf{k}-\mathbf{q}0}-n_{\mathbf{k}-\mathbf{q}\mu'}) + P_{k\mu}^{(1)}(\pm\omega), \tag{43}$$

$$\hat{P}_{k\mu}(\pm\omega) = -\frac{\omega_p^2}{2}(n_{k0}-n_{k\mu})\sum_{\mu,\lambda}\frac{f_{0\mu}(\mathbf{k},\lambda)E_{\mu}(\mathbf{k})}{(\omega^2-c^2k^2-\omega_p^2)} + \frac{2}{N}(n_{k0}-n_{k\mu})\sum_{\mathbf{q},\mu',\beta}U_{\mu\mu',\mu'\mu}^{\alpha\beta}(\mathbf{q}) \times (n_{\mathbf{k}-\mathbf{q}0}-n_{\mathbf{k}-\mathbf{q}\mu'}) + \hat{P}_{k\mu}^{(1)}(\pm\omega). \tag{44}$$

The quantities $P_{k\mu}^{(1)}(\pm\omega)$ and $\hat{P}_{k\mu}^{(1)}(\pm\omega)$ are the diagonal and nondiagonal elements of (38), respectively. Using (39) and (40), we derive an expression for the Green's function

$$g_{\mu}^{(1)}(\mathbf{k};\omega) \equiv G_{\mu}^{(1)}(\mathbf{k};\omega) + G_{\mu}^{(1)}(\mathbf{k};-\omega) - \hat{G}_{\mu}^{(1)}(\mathbf{k};\omega) - \hat{G}_{\mu}^{(1)}(\mathbf{k};-\omega)$$

in the form

$$g_{\mu}^{(1)}(\mathbf{k};\omega) = \frac{(n_{k0}-n_{k\mu})}{\pi}\mathcal{E}_{k\mu}^{(1)}(\omega)\left\{\left[\omega - \frac{P_{k\mu}(\omega)-P_{k\mu}(-\omega)}{2}\right]^2 - \mathcal{E}_{k\mu}^{(1)}(\omega)\left[E_{\mu}(\mathbf{k}) + \frac{P_{k\mu}(\omega)+P_{k\mu}(-\omega)}{2} - \hat{P}'_{k\mu}(\omega)\right]\right\}^{-1}. \tag{45}$$

The Green's functions (39), (40), and (45) have the same poles that give the energies of excitation for the perturbed energy spectrum.

The equation of motion for the photon Green's function $D_{k\lambda}(\omega) \equiv \langle\langle\tilde{\beta}_{\lambda}(\mathbf{k});\tilde{\beta}_{\lambda}^{\dagger}(\mathbf{k})\rangle\rangle$ is easily derived by means of the Hamiltonian (1) as

$$D_{k\lambda}(\omega) = D_{k\lambda}^{(00)}(\omega) + D_{k\lambda}^{(00)}(\omega)\prod_{k\lambda}(\omega)D_{k\lambda}^{(00)}(\omega) = D_{k\lambda}^{(00)}(\omega) + D_{k\lambda}^{(00)}(\omega)\hat{\prod}_{k\lambda}(\omega)D_{k\lambda}(\omega), \tag{46}$$

where $D_{k\lambda}^{(00)}(\omega)$ is the unperturbed Green's function in the absence of dispersion given by

$$D_{k\lambda}^{(00)}(\omega) = (ck/\pi)(\omega^2-c^2k^2-\omega_p^2)^{-1}, \tag{47}$$

and the function $\prod_{k\lambda}(\omega)$ is equal to

$$\prod_{k\lambda}(\omega) = \left(\frac{\pi^2}{ck}\right)\left\{\omega_p^2\sum_{\mu}f_{0\mu}(\mathbf{k},\lambda)E_{\mu}(\mathbf{k})g_{\mu}(\mathbf{k};\omega) + \frac{2\omega_p^2}{N^{\frac{1}{2}}} \times \sum_{\mathbf{q},\mu',\mu_1,\mu}[f_{0\mu}(\mathbf{k},\lambda)f_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k},\lambda)E_{\mu}(\mathbf{k})E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{\frac{1}{2}} \times [\langle\langle b_{\mu}(\mathbf{k}) - b_{\mu}^{\dagger}(-\mathbf{k});\, b_{\mu'}^{\dagger}(\mathbf{k}-\mathbf{q})b_{\mu_1}(-\mathbf{q})\rangle\rangle + \langle\langle b_{\mu'}(\mathbf{k}-\mathbf{q})b_{\mu_1}^{\dagger}(-\mathbf{q});\, b_{\mu}^{\dagger}(\mathbf{k}) - b_{\mu}(-\mathbf{k})\rangle\rangle] + \frac{4\omega_p^2}{N}\sum_{\mathbf{q},\mu_1,\mu'}f_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k},\lambda)E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1} \times \langle\langle b_{\mu'}(\mathbf{k}-\mathbf{q})b_{\mu_1}^{\dagger}(-\mathbf{q});\, b_{\mu'}^{\dagger}(\mathbf{k}-\mathbf{q})b_{\mu_1}(-\mathbf{q})\rangle\rangle\right\}. \tag{48}$$

The polarization operator $\hat{\prod}_{k\lambda}(\omega)$ and the function $\prod_{k\lambda}(\omega)$ are related by

$$\hat{\prod}_{k\lambda}(\omega) = \prod_{k\lambda}(\omega)[1 + D_{k\lambda}^{(00)}(\omega)\prod_{k\lambda}(\omega)]^{-1} \tag{49}$$

or

$$\prod_{k\lambda}(\omega) = [1 - \hat{\prod}_{k\lambda}(\omega)D_{k\lambda}^{(00)}(\omega)]^{-1}\hat{\prod}_{k\lambda}(\omega) = \hat{\prod}_{k\lambda}(\omega)[1 + D_{k\lambda}(\omega)\hat{\prod}_{k\lambda}(\omega)]. \tag{50}$$

In the zero approximation [i.e., retaining only the first terms in the expressions (43) and (44)] we have, from (45),

$$g_{\mu}^{(0)}(\mathbf{k};\omega) = \frac{(n_{k0}-n_{k\mu})}{\pi} \times \frac{(\omega^2-c^2k^2-\omega_p^2)E_{\mu}(\mathbf{k})}{[\omega^2-E_{\mu}^2(\mathbf{k})][\omega^2\hat{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2]}, \tag{51}$$

where $\hat{\eta}_{\lambda}^2(\mathbf{k},\omega)$ is the square of the unperturbed index of refraction defined by

$$\hat{\eta}_{\lambda}^2(\mathbf{k},\omega) = 1 + \alpha_{\lambda}(\mathbf{k},\omega), \tag{52a}$$

and

$$\alpha_{\lambda}(\mathbf{k},\omega) = \omega_p^2(n_{k0}-n_{k\mu})\sum_{\mu,\lambda}\frac{f_{0\mu}(\mathbf{k},\lambda)}{[E_{\mu}^2(\mathbf{k})-\omega^2]} \tag{52b}$$

is the frequency- and wavevector-dependent polarizability with polarization λ. In the absence of dispersion, i.e., when $\alpha_{\lambda}(\mathbf{k},\omega) \approx 0$, Eq. (51) is reduced to

$$g_{\mu,b}^{(0)}(\mathbf{k};\omega) = [(n_{k0}-n_{k\mu})/\pi]E_{\mu}(\mathbf{k})[\omega^2-E_{\mu}^2(\mathbf{k})]^{-1}, \tag{52c}$$

which describes the bare exciton spectrum. Substituting (51) into (48) and retaining only the first term on the right-hand side of (48), we obtain

$$\prod_{k\lambda}^{(0)}(\omega) = -(\pi/ck)\omega^2[\hat{\eta}_{\lambda}^2(\mathbf{k},\omega) - \hat{\eta}_{\infty}^2] \times (\omega^2-c^2k^2-\omega_p^2)[\omega^2\hat{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2]^{-1}, \tag{53}$$

with

$$\hat{\eta}_\infty^2 = 1 - (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})\omega_p^2/\omega^2.$$

When Eq. (53) is inserted into (49), the expression for the polarization operator in the zero approximation turns out to be

$$\prod_{\mathbf{k}\lambda}^{(0)}(\omega) = -\left(\frac{\pi}{ck}\right)\omega^2[\hat{\eta}_\lambda^2(\mathbf{k}, \omega) - \hat{\eta}_\infty^2]. \quad (54)$$

This relation shows that the polarization operator is entirely determined by the properties of the medium. Substituting (54) into (46), we have

$$D_{\mathbf{k}\lambda}^{(0)}(\omega) = (ck/\pi)[\omega^2\hat{\eta}_\lambda^2(\mathbf{k}, \omega) - c^2k^2]^{-1}. \quad (55)$$

The Green's function derived here for the exciton and photon field, respectively, in the zero approximation are in agreement with those obtained in I, where the unperturbed excitation spectrum has been discussed. A combination of (51), (52c), (54), and (55) yields

$$\frac{g_\mu^{(0)}(\mathbf{k}; \omega)}{g_{\mu,b}^{(0)}(\mathbf{k}; \omega)} = 1 + \hat{\prod}_{\mathbf{k}\lambda}^{(0)}(\omega)D_{\mathbf{k}\lambda}^{(0)}(\omega), \quad (56)$$

which indicates that the ratio of the exciton Green's functions with and without dispersion taken into account depends on the polarization operator of the medium and the photon Green's function.

III. SCATTERING

We here evaluate the Green's functions for the exciton and photon field given by (45) and (46), respectively, for the process where the dressed exciton $(\mathbf{k}, \mu)$ decays into a bare exciton $(\mathbf{k} - \mathbf{q}, \mu')$ and a dressed one $(\mathbf{q}, \lambda')$. For the bare exciton $(\mathbf{k} - \mathbf{q}, \mu')$, the oscillator strength $f_{0\mu'}(\mathbf{k} - \mathbf{q})$ for the transition $0 \to \mu'$ is not significant and is taken equal to zero. Then, using the fact that in the complex ω plane the relations $P_{\mathbf{k}\mu}^{(1)}(-\omega) = P_{\mathbf{k}\mu}^{(1)}(\omega)$ and $\hat{P}_{\mathbf{k}\mu}^{(1)}(-\omega) = \hat{P}_{\mathbf{k}\mu}^{(1)}(\omega)$ hold, we evaluate the Green's functions that appear on the right-hand side of (38) in the zero approximation by means of the Hamiltonian (1). Then, substituting the result into (45), we obtain

$$g_\mu^{(1)}(\mathbf{k}; \omega) = \frac{(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})}{\pi} \times \frac{(\omega^2 - c^2k^2 - \omega_p^2)\mathcal{E}_{\mathbf{k}\mu}^{(1)}(\omega)}{[\omega^2 - \mathcal{E}_{\mathbf{k}\mu}^2(\omega)][\omega^2\hat{\eta}_\lambda^2(\mathbf{k}, \omega) - c^2k^2]}. \quad (57)$$

In the absence of dispersion for the exciton $(\mathbf{k}, \mu)$, expression (57) becomes

$$g_{\mu,b}^{(1)}(k; \omega) = [(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/\pi]\mathcal{E}_{\mathbf{k}\mu}^{(1)}(\omega)[\omega^2 - \mathcal{E}_{\mathbf{k}\mu}^2(\omega)]^{-1}, \quad (58)$$

which corresponds to the bare exciton spectrum. In expressions (57) and (58) we have taken

$$\mathcal{E}_{\mathbf{k}\mu}^2(\omega) \equiv \mathcal{E}_{\mathbf{k}\mu}^{(1)}(\omega)\mathcal{E}_{\mathbf{k}\mu}^{(2)}(\omega),$$

where $\mathcal{E}_{\mathbf{k}\mu}(\omega)$ is now the perturbed energy of excitation of the bare exciton $(\mathbf{k}, \mu)$, and $\mathcal{E}_{\mathbf{k}\mu}^{(1)}(\omega)$, $\mathcal{E}_{\mathbf{k}\mu}^{(2)}(\omega)$ are given by

$$\begin{aligned}\mathcal{E}_{\mathbf{k}\mu}^{(1)}(\omega) = {} & E_\mu(\mathbf{k}) + 4/N(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\mu',\beta} U_{\mu\mu',\mu'\mu}^{\alpha\beta}(\mathbf{q})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'}) \\ & + \frac{(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})}{N} \sum_{\mathbf{q},\mu',\mu_1,\beta} |U_{0\mu,\mu'\mu_1}^{\alpha\beta}(\mathbf{k} - \mathbf{q}) + U_{0\mu,\mu_1\mu'}^{\alpha\beta}(\mathbf{q})|^2 (\omega - E_{\mathbf{k}-\mathbf{q}\mu'} - E_{\mathbf{q}\mu_1})^{-1} \\ & \times (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'}) + \frac{\omega_p^2}{N}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\mu',\mu_1,\lambda',\beta} \Bigg| [f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q}, \lambda')E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}]^{\frac{1}{2}} \\ & + \frac{1}{2}\left[\frac{U_{0\mu',\mu_1\mu}^{\alpha\beta}(-\mathbf{q}) + 2U_{0\mu',\mu\mu_1}^{\alpha\beta}(\mathbf{k}) + U_{0\mu_1,\mu'\mu}^{\alpha\beta}(\mathbf{k} - \mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1} - \omega} + \frac{U_{0\mu,\mu'\mu_1}^{\alpha\beta}(\mathbf{k} - \mathbf{q}) + U_{0\mu,\mu_1\mu'}^{\alpha\beta}(\mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu'} + E_{\mathbf{q}\mu_1} - \omega}\right] \\ & \times [f_{0\mu_1}(\mathbf{q}, \lambda')E_{\mathbf{q}\mu_1}]^{\frac{1}{2}}(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})\Bigg|^2 \tilde{D}_{\mathbf{q}\lambda'}^{(0)}(\mathbf{k} - \mathbf{q}, \mu'; \omega) \end{aligned} \quad (59)$$

and

$$\begin{aligned}\mathcal{E}_{\mathbf{k}\mu}^{(2)}(\omega) = {} & E_\mu(\mathbf{k}) + \frac{(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})}{N} \sum_{\mathbf{q},\mu',\mu_1,\beta} |U_{0\mu,\mu'\mu_1}^{\alpha\beta}(\mathbf{k} - \mathbf{q}) + U_{0\mu,\mu_1\mu'}^{\alpha\beta}(\mathbf{q})|^2 \\ & \times (\omega - E_{\mathbf{k}-\mathbf{q}\mu'} - E_{\mathbf{q}\mu_1})^{-1}(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'}) + (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})\frac{\omega_p^2}{N} \sum_{\mathbf{q},\mu'\mu_1,\lambda',\beta} \Bigg| [f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q}, \lambda')E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}]^{\frac{1}{2}} \\ & + \frac{1}{2}\left[\frac{U_{0\mu',\mu_1\mu}^{\alpha\beta}(-\mathbf{q}) - U_{0\mu_1,\mu'\mu}^{\alpha\beta}(\mathbf{k} - \mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1} - \omega} + \frac{U_{0\mu,\mu'\mu_1}^{\alpha\beta}(\mathbf{k} - \mathbf{q}) + U_{0\mu,\mu_1\mu'}^{\alpha\beta}(\mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu'} + E_{\mathbf{q}\mu_1} - \omega}\right] \\ & \times [f_{0\mu_1}(\mathbf{q}, \lambda')E_{\mathbf{q}\mu_1}]^{\frac{1}{2}}(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})\Bigg|^2 \tilde{D}_{\mathbf{q}\lambda'}^{(0)}(\mathbf{k} - \mathbf{q}, \mu'; \omega), \end{aligned} \quad (60)$$

where

$$\tilde{D}_{\mathbf{q}\lambda'}^{(0)}(\mathbf{k} - \mathbf{q}, \mu'; \omega) = (2\pi/cq)\langle\langle b_{\mu'}(\mathbf{k} - \mathbf{q})\tilde{\beta}_{\lambda'}(\mathbf{q}); b_{\mu'}^\dagger(\mathbf{k} - \mathbf{q})\tilde{\beta}_{\lambda'}^\dagger(\mathbf{q})\rangle\rangle^{(0)}. \quad (61)$$

In deriving (59) and (60), we have neglected terms proportional to the distribution function $\langle b^\dagger_{\mu_1}(\mathbf{q}) b_{\mu_1}(\mathbf{q})\rangle$ because their contributions are negligibly small. The Green's function $\tilde{D}^{(0)}_{\mathbf{q}}(\mathbf{k}-\mathbf{q}, \mu'; \omega)$ is easily evaluated by using the unperturbed Hamiltonian; we find

$$\tilde{D}^{(0)}_{\mathbf{q}\lambda'}(\mathbf{k}-\mathbf{q}, \mu'; \omega) = \frac{(n_{k-q0} - n_{k-q\mu'})}{(cq)} \times [(\omega - E_{\mathbf{k}-\mathbf{q}\mu'})^2 \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega - E_{\mathbf{k}-\mathbf{q}\mu'}) - c^2q^2]^{-1} \times \Big((\omega - E_{\mathbf{k}-\mathbf{q}\mu'}) \langle \tilde{\beta}^\dagger_{\lambda'}(\mathbf{q}) \tilde{\beta}_{\lambda'}(\mathbf{q})\rangle^{(0)} + cq \langle [\beta_{\lambda'}(\mathbf{q}) - \beta^\dagger_{\lambda'}(-q)] \tilde{\beta}_{\lambda'}(\mathbf{q})\rangle^{(0)} + i\omega_p \sum_{\mu_1} \frac{[f_{0\mu_1}(\mathbf{q}, \lambda') E_{q\mu_1} cq]^{\frac{1}{2}}}{(\omega - E_{\mathbf{k}-\mathbf{q}\mu'})^2 - E^2_{\mathbf{q}\mu_1}} \times \{ (\omega - E_{\mathbf{k}-\mathbf{q}\mu'}) \langle [b_{\mu_1}(\mathbf{q}) - b^\dagger_{\mu_1}(-\mathbf{q})] \tilde{\beta}^\dagger_{\lambda'}(\mathbf{q})\rangle^{(0)} + E_{\mathbf{q}\mu_1} \langle [b_{\mu_1}(\mathbf{q}) + b^\dagger_{\mu_1}(-\mathbf{q})] \tilde{\beta}_{\lambda'}(\mathbf{q})\rangle^{(0)} \} \Big). \quad (62)$$

The distribution functions in (62) are calculated at zero temperature by making use of the corresponding Green's functions in the zero approximation as was done in I with the result that

$$\langle \tilde{\beta}^\dagger_{\lambda'}(\mathbf{q}) \tilde{\beta}_{\lambda'}(\mathbf{q})\rangle^{(0)}_0 = \sum_\rho \left[\frac{cq}{\omega_{\rho\lambda'}(\mathbf{q})}\right] \left[\frac{d\omega^2}{d\omega^2} \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega)\right]^{-1}_{\omega=\omega_{\rho\lambda'}(\mathbf{q})}, \quad (63a)$$

$$\langle [\beta_{\lambda'}(\mathbf{q}) - \beta^\dagger_{\lambda'}(-\mathbf{q})] \tilde{\beta}^\dagger_{\lambda'}(\mathbf{q})\rangle^{(0)}_0 = \sum_\rho \left[\frac{d\omega^2}{d\omega^2} \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega)\right]^{-1}_{\omega=\omega_{\rho\lambda'}(\mathbf{q})}, \quad (63b)$$

$$i\omega_p \sum_{\mu_1} [f_{0\mu_1}(\mathbf{q}, \lambda') E_{\mathbf{q}\mu_1} \mid cq]^{\frac{1}{2}} \langle [b_{\mu_1}(\mathbf{q}) - b^\dagger_{\mu_1}(-\mathbf{q})] \tilde{\beta}^\dagger_{\lambda'}, (\mathbf{q})\rangle^{(0)}_0 = -\sum_\rho \alpha_{\lambda'}[\mathbf{q}, \omega_{\rho\lambda'}(\mathbf{q})] \frac{E^2_{\mathbf{q}\mu_1}}{\omega_{\rho\lambda'}(\mathbf{q})} \left[\frac{d\omega^2}{d\omega^2} \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega)\right]^{-1}_{\omega=\omega_{\rho\lambda'}(\mathbf{q})}, \quad (63c)$$

$$i\omega_p \sum_{\mu_1} [f_{0\mu_1}(\mathbf{q}; \lambda') E_{\mathbf{q}\mu_1} \mid cq]^{\frac{1}{2}} E_{\mathbf{q}\mu_1} \times \langle [b_{\mu_1}(\mathbf{q}) + b^\dagger_{\mu_1}(-\mathbf{q})] \tilde{\beta}_{\lambda'}(\mathbf{q})\rangle^{(0)}_0 = \sum_\rho \alpha_{\lambda'}[\mathbf{q}, \omega_{\rho\lambda'}(\mathbf{q})] E^2_{\mathbf{q}\mu_1} \left[\frac{d\omega^2}{d\omega^2} \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega)\right]^{-1}_{\omega=\omega_{\rho\lambda'}(\mathbf{q})}, \quad (63d)$$

where $\omega^2_{\rho\lambda}(\mathbf{q})$ is the ρth root of the secular equation

$$\omega^2 \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega) - c^2q^2 = 0. \quad (64)$$

Substitution of (63a)–(63d) into (62) yields

$$\tilde{D}^{(0)}_{\mathbf{q}\lambda'}(\mathbf{k}-\mathbf{q}, \mu'; \omega) = (n_{k-q0} - n_{k-q\mu'})[(\omega - E_{\mathbf{k}-\mathbf{q}\mu'})^2 \hat{\eta}^2_{\lambda'}(\mathbf{q}, \omega - E_{\mathbf{k}-\mathbf{q}\mu'}) - c^2q^2]^{-1} \sum_\rho \{\omega - E_{\mathbf{k}-\mathbf{q}\mu'} + \omega_{\rho\lambda'}(q) - \frac{\alpha_{\lambda'}(\mathbf{q}, \omega_{\rho\lambda'}(\mathbf{q})) E^2_{\mathbf{q}\mu_1}}{(\omega - E_{\mathbf{k}-\mathbf{q}\mu'})^2 - E^2_{\mathbf{q}\mu_1}} [\omega - E_{k-\mathbf{q}\mu'} - \omega_{\rho\lambda'}(\mathbf{q})]\} 1/\omega_{\rho\lambda'}(\mathbf{q}) \left[\frac{d\omega^2}{d\omega^2} \eta^2_{\lambda'}(\mathbf{q}, \omega)\right]^{-1}_{\omega=\omega_{\rho\gamma'}(\mathbf{q})}. \quad (65)$$

Thus, the dispersion of the photon $(\mathbf{q}, \lambda')$ is fully included in expression (65). Let us now examine the physical meaning of each term on the right-hand side of (59). The first term $E_\mu(\mathbf{k})$, which is the largest of all, is the energy of excitation of the bare exciton $(\mathbf{k}, \mu)$ including the electron–hole pair interactions in the zero approximation. The second term indicates the instantaneous interaction between two bare excitons $(\mathbf{k}, \mu)$ and $(\mathbf{k}-\mathbf{q}, \mu')$. The third term corresponds to the instantaneous interaction leading to the creation or absorption of two bare excitons $(\mathbf{q}, \mu_1)$ and $(\mathbf{k}-\mathbf{q}, \mu')$ simultaneously; it includes both coherent and incoherent processes. The fourth term consists of three contributions: the first is the radiative interaction between two excitons $(\mathbf{k}, \mu)$ and $(\mathbf{k}-\mathbf{q}, \mu')$ through the exchange of the photon $(\mathbf{q}, \lambda')$, while the remaining two are of higher order in the electronic charge and describe the cooperative effect arising from the correlation of the intermolecular interactions with the radiation field where either the exciton $(\mathbf{q}, \mu_1)$ is created and the other $(\mathbf{k}-\mathbf{q}, \mu')$ is absorbed or both excitons are created or absorbed. It is clear that under normal circumstances the second, third and fourth terms on the right-hand side of (59) give a small correction to the real energy $E_\mu(\mathbf{k})$ but they may be of importance in the case of resonance. The imaginary part of (59) is given by

$$\mathrm{Im}\, \delta^{(1)}_{\mathbf{k}\mu}(\omega) = \pi \frac{(n_{k0} - n_{k\mu})}{N} \sum_{\mathbf{q}, \mu', \mu_1 \beta} |U^{\alpha\beta}_{0\mu, \mu'\mu_1}(\mathbf{k}-\mathbf{q}) + U^{\alpha\beta}_{0\mu, \mu_1\mu'}(\mathbf{q})|^2 \times (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'}) \delta(\omega - E_{\mathbf{k}-\mathbf{q}\mu'} - E_{\mathbf{q}\mu_1}) + \frac{\omega_p^2}{N} (n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \mathrm{Im} \sum_{\mathbf{q}, \mu_1, \mu', \lambda', \beta} \Big| [f_{\mathbf{k}\mu, \mathbf{k}-\mathbf{q}\mu'}(\mathbf{q}\lambda') E_{\mathbf{k}-\mathbf{q}\mu', \mathbf{k}\mu}]^{\frac{1}{2}} \cdot \frac{1}{2} \left[\frac{U^{\alpha\beta}_{0\mu', \mu_1\mu}(-\mathbf{q}) + 2U^{\alpha\beta}_{0\mu', \mu\mu_1}(\mathbf{k}) + U^{\alpha\beta}_{0\mu_1, \mu'\mu}(\mathbf{k}-\mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu', \mathbf{q}\mu_1} - \omega} + \frac{U^{\alpha\beta}_{0\mu, \mu'\mu_1}(\mathbf{k}-\mathbf{q}) + U^{\alpha\beta}_{0\mu, \mu_1\mu'}(\mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu'} + E_{\mathbf{q}\mu_1} - \omega} \right] \times [f_{0\mu_1}(\mathbf{q}, \lambda') E_{\mathbf{q}\mu_1}]^{\frac{1}{2}} (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1}) \Big|^2 \tilde{D}^{(0)}_{\mathbf{q}\lambda'}(\mathbf{k}-\mathbf{q}, \mu'; \omega). \quad (66)$$

The first term on the right-hand side of (66) is different from zero when $\omega = E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'}$. The first contribution from the second term is the radiative one which, when the sum over $\mathbf{q}$ is taken in the limit of $\operatorname{Im} \tilde{\eta}^2_\lambda(\mathbf{q}, \omega - E_{\mathbf{k}-\mathbf{q}\mu'}) \to 0$, is equal to the spontaneous emission probability (in energy units) for the transition $\mu \to \mu'$, in the dipole approximation, while the remaining ones give small contributions when $\omega = E_{\mathbf{k}-\mathbf{q}\mu'} \pm E_{\mathbf{q}\mu_1}$. Of course, the expression for $\operatorname{Im} \mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)$ must be compared with the corresponding terms arising from the coupling of intermolecular forces with the phonon field, but at very low temperatures both effects could be of equal importance. A similar discussion holds for the terms appearing in the expression for $\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)$. Thus, $\mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)$ and $\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)$ are the renormalized energies of excitation correct to the first approximation.

In expression (57), $\tilde{\eta}^2_\lambda(\mathbf{k}, \omega)$ is the square of the perturbed index of refraction defined by

$$\tilde{\eta}^2_\lambda(\mathbf{k}, \omega) \approx 1 + \alpha^{(1)}_\lambda(\mathbf{k}, \omega) - \alpha^{(2)}_{\mathbf{k}\lambda}(\mathbf{q}, \lambda', \omega) + \beta_{\mathbf{k}\lambda}(\omega), \tag{67}$$

where $\alpha^{(1)}_\lambda(\mathbf{k}, \omega)$ is the frequency and wavevector dependent polarizability expressed in terms of the renormalized exciton energies, i.e.,

$$\alpha^{(1)}_\lambda(\mathbf{k}, \omega) = (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})\omega_p^2 \sum_{\mu,\lambda} f_{0\mu}(\mathbf{k}, \lambda)/(\mathcal{E}^2_{\mathbf{k}\mu}(\omega) - \omega^2). \tag{68}$$

The quantities $\alpha^{(2)}_{\mathbf{k}\lambda}(q, \lambda', \omega)$ and $\beta_{\mathbf{k}\lambda}(\omega)$ are the nonlinear polarizabilities given by the following expressions:

$$\begin{aligned}\alpha^{(2)}_{\mathbf{k}\lambda}(\mathbf{q}, \lambda', \omega) = \frac{\omega_p^4}{N} \sum_{\mathbf{q},\mu_1,\mu',\mu_1,\beta} \left[\frac{E_\mu(\mathbf{k})}{\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)}\right] \Big| &(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})[f^\alpha_{0\mu}(\mathbf{k}, \lambda) f^\beta_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q}, \lambda') E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}/E_\mu(\mathbf{k})]^{\frac{1}{2}} [\mathcal{E}_{\mathbf{k}\mu}(\omega) - (\omega)]^{-1} \\ &+ (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})[f^\beta_{0\mu_1}(\mathbf{q}, \lambda') f^\alpha_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k}, \lambda) E_{\mathbf{q}\mu_1}/E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{\frac{1}{2}} (\omega - E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1})^{-1} \\ &+ \tfrac{1}{2}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu})(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1}) \frac{f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda)}{[\mathcal{E}_{\mathbf{k}\mu}(\omega) - \omega]} \left[\frac{U^{\alpha\beta}_{0\mu',\mu_1\mu}(-\mathbf{q}) - U^{\alpha\beta}_{0\mu_1,\mu'\mu}(\mathbf{k} - \mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1} - \omega}\right. \\ &+ \left.\frac{U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q}) + U_{0\mu,\mu_1\mu'}(\mathbf{q})}{E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'} - \omega}\right]\left[\frac{f_{0\mu_1}(\mathbf{q}, \lambda') E_{\mathbf{q}\mu_1}}{E_\mu(\mathbf{k})}\right]^{\frac{1}{2}} \Big|^2 \tilde{D}^{(0)}_\lambda(\mathbf{k} - \mathbf{q}, \mu'; \omega)\end{aligned} \tag{69}$$

and

$$\begin{aligned}\beta_{\mathbf{k}\lambda}(\omega) = \frac{\omega_p^2}{2N}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\lambda,\mu_1,\mu_1,\mu',\beta} &\left| f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda) \frac{[U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q}) - U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q})]}{\mathcal{E}_{\mathbf{k}\mu}(\omega) - \omega}\right. \\ &+ \left.\frac{[U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q}) + U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q})]}{E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'} + \omega} f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda)\right|^2 \frac{(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'})}{\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)(E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'})} \\ + \frac{\omega_p^2}{4N}(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \sum_{\mathbf{q},\lambda,\mu,\mu',\mu_1,\beta} &\left| f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda) \frac{[U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q}) + U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q})]}{\mathcal{E}_{\mathbf{k}\mu}(\omega) - \omega}\right. \\ &- \left.\frac{[U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q}) + U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q})]}{E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'} \pm \omega} f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda)\right|^2 \frac{(n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'})}{\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)(E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'})}.\end{aligned} \tag{70}$$

In deriving (67) we have made use of the f-sum rule in the limit $\mathbf{k} \to 0$ in order to eliminate terms proportional to ω^{-2}. The first two terms in (69) give the Raman polarizability tensor while the third term is a correlation effect of higher order in the electronic charge. $\beta_{\mathbf{k}\lambda}(\omega)$ is a cooperative polarizability, proportional to the sixth power in the electronic charge, arising from the correlation of intermolecular interactions with the radiation field which results in producing double exciton states. The transition probability that leads to double excitation of a pair of atoms by a single photon has been considered by Dexter,[5] who employed the first-order corrections in perturbation theory induced by the electronic interactions to the zero-order wavefunctions. The same method has been used by Jortner and Rice[6] to discuss the cooperative exciton states in molecular crystals. In expressions (69) and (70), the excitation energies for the exciton $(\mathbf{k}, \mu)$ and for the excitons $(\mathbf{q}, \mu_1)$ and $(\mathbf{k} - \mathbf{q}, \mu')$ are correct in the first and zero approximation, respectively. The dispersion of the photon $(\mathbf{q}, \lambda')$ in (69) is expressed by the function $\tilde{D}^{(0)}_{\mathbf{q}\lambda'}(\mathbf{k} - \mathbf{q}, \mu'; \omega)$. Since $\alpha^{(2)}_{\mathbf{k}\lambda}(\mathbf{q}, \lambda', \omega)$ and $\beta_{\mathbf{k}\lambda}(\omega)$ are both functions of ω, they are applicable to such phenomena as that of resonances occurring in the region of frequencies of either the incoming or the scattered radiation.

We now calculate the photon Green's function, $D_{\mathbf{k}\lambda}(\omega)$, in the first approximation. Substituting the expression for $g^{(1)}_\mu(\mathbf{k}; \omega)$ given by (57) into (48) and

[5] D. L. Dexter, Phys. Rev. **126**, 1962 (1962).

[6] J. Jortner and S. A. Rice, J. Chem. Phys. **44**, 3364 (1966).

evaluating the remaining Green's functions that appear on the right-hand side of (48) by means of (19) in the same approximation as it was done for $g_{\mu}^{(1)}(\mathbf{k};\omega)$, we find the quantity for $\prod_{\mathbf{k}\lambda}^{(1)}(\omega)$:

$$\prod_{\mathbf{k}\lambda}^{(1)}(\omega)=\left(\frac{\pi}{ck}\right)\left[1-\omega^2\frac{[\tilde{\eta}_{\lambda}^2(k,\omega)-\hat{\eta}_{\infty}^2]}{[\omega^2\tilde{\eta}_{\lambda}^2(k,\omega)-c^2k^2]}\right]\times\{-\omega^2[\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-\hat{\eta}_{\infty}^2]\}. \quad (71)$$

Substitution of (71) into (49) gives the expression for the polarization operator in the form

$$\tilde{\prod}_{\mathbf{k}\lambda}^{(1)}(\omega)=-\left(\frac{\pi}{ck}\right)\omega^2[\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-\hat{\eta}_{\infty}^2], \quad (72)$$

and, from (46), we derive the following expression for the photon Green's function:

$$D_{\mathbf{k}\lambda}^{(1)}(\omega)=(ck/\pi)[\omega^2\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2]^{-1}. \quad (73)$$

Comparison of (72) and (73) with (54) and (55) shows that the polarization operator and the photon Green's function in the first approximation are just the corresponding unperturbed ones renormalized because of the interactions. Similarly, combining (72), (73), (57), and (50), we obtain the relation

$$\frac{g_{\mu}^{(1)}(\mathbf{k};\omega)}{g_{\mu,0}^{(1)}(k;\omega)}=1+\tilde{\prod}_{\mathbf{k}\lambda}^{(1)}(\omega)D_{\mathbf{k}\lambda}^{(1)}(\omega), \quad (74)$$

which corresponds to the expression (56) for the same relation in the zero approximation.

Knowing the Green's function for the photon field, we can calculate the corresponding spectral intensity by the relation[7]

$$J_{\mathbf{k}\lambda}^{(1)}(\omega)=2\,\mathrm{Im}\,D_{\mathbf{k}\lambda}^{(1)}(\omega)(1-e^{\beta\omega})^{-1}=\left(\frac{2ck}{\pi}\right)\frac{\omega^2\,\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)(e^{\beta\omega}-1)^{-1}}{[\omega^2\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2]^2+[\omega^2\,\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)]^2}, \quad (75)$$

where we have taken

$$\tilde{\eta}_{\lambda}^2(\mathbf{k}\omega)=\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)+i\,\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega).$$

Here $\beta=(k_BT)^{-1}$, where K_B is Boltzmann's constant and T the absolute temperature. Using the relation

$$\omega^2\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2=\sum_{\rho}[\omega^2-\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})]\left[\frac{d\omega^2}{d\omega^2}\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\right]_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}, \quad (76)$$

where $\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$ is the ρth root of the secular equation

$$\omega^2\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)-c^2k^2=0, \quad (77)$$

we rewrite (75) as

$$J_{\mathbf{k}\lambda}^{(1)}(\omega)=\left(\frac{2ck}{\pi}\right)\sum_{\rho}\left[\frac{d\omega^2}{d\omega^2}\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\right]_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}^{-1}\times\frac{\gamma_{k\lambda}(\omega)(e^{\beta\omega}-1)^{-1}}{[\omega^2-\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})]^2+\gamma_{\mathbf{k}\lambda}^2(\omega)}, \quad (78)$$

where $\gamma_{\mathbf{k}\lambda}(\omega)$ is equal to

$$\gamma_{\mathbf{k}\lambda}(\omega)=\omega^2\,\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\times\sum_{\rho}\left[\frac{d\omega^2}{d\omega^2}\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\right]_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}^{-1}. \quad (79)$$

As $\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)$ goes to zero, the spectral intensity $J_{\mathbf{k}\lambda}^{(1)}(\omega)$ tends to a delta-shape distribution, i.e.,

$$J_{\mathbf{k}\lambda}^{(1)}(\omega)=2ck\sum_{\rho}\left[\frac{d\omega^2}{d\omega^2}\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\right]_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}^{-1}\times\delta[\omega^2-\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})](e^{\beta\omega}-1)^{-1},\quad \text{for } \mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\to 0. \quad (80)$$

In the limiting case when $\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)$ may be considered to be very small but finite, $\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\ll 1$, the function (78) has a steep maximum at some value $\omega^2\approx\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$, provided that $d\,\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)/d\omega^2\ll 1$. In such a case the center of an absorption line described by the function (78) will appear Lorentzian, while the wings may not. If we take into account that at $\omega^2\sim\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$ the function $\mathrm{Im}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)$ varies slowly with ω^2 and we take $\gamma_{\mathbf{k}\lambda}(\omega)\approx\gamma_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$, then (78) becomes

$$J_{\mathbf{k}\lambda}^{(1)}(\omega)\approx\left(\frac{2ck}{\pi}\right)\sum_{\rho}\left[\frac{d\omega^2}{d\omega^2}\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2(\mathbf{k},\omega)\right]_{\omega=\tilde{\omega}_{\lambda}(\mathbf{k})}^{-1}\times\frac{\gamma_{k\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})](e^{\beta\omega}-1)^{-1}}{[\omega^2-\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})]^2+\gamma_{\mathbf{k}\lambda}^2[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]}. \quad (81)$$

The function (75) or (78) describes the behavior of the photon excitation spectrum in the whole range of frequencies ω, while (81) is restricted to those in the neighborhood of $\omega^2\approx\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$. The function (81) is a Lorentzian line with maximum at $\omega^2\approx\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$; the square of the energy shift is equal to

$$\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})\{\mathrm{Re}\,\tilde{\eta}_{\lambda}^2[\mathbf{k},\tilde{\omega}_{\rho\lambda}(\mathbf{k})]-1\}$$

and the spectral width is of order of magnitude of $\tilde{\gamma}_{\mathbf{k}}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]=|\gamma_{\mathbf{k}}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]/\tilde{\omega}_{\rho\lambda}(\mathbf{k})|$ in energy units. The energy of excitation $\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$ is determined by the equation

$$\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})\,\mathrm{Re}\,\tilde{\eta}_{\lambda}^2[\mathbf{k},\tilde{\omega}_{\rho\lambda}(\mathbf{k})]-c^2k^2=0, \quad (82)$$

provided that $\mathrm{Im}\,\tilde{\eta}_{\lambda}^2[\mathbf{k},\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\ll 1$.

Let us consider the case of resonance Raman scattering that is observed when the frequency of the

[7] V. L. Bonch-Bruevich and S. V. Tyablikov, *Green's Function Method in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962), p. 24.

incident radiation falls in an absorption band.[8] We describe the process where the frequency of the incident radiation ck is in the neighborhood of the absorption band $(\mathbf{k}, \mu)$, while the frequency of the scattered radiation is in the transparent region of the crystal. In such a process, the energies $\mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)$ and $\mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega)$, given by the expressions (59) and (60), respectively, are real. Then taking the imaginary part of (67), we find that the expression for $\operatorname{Im} \tilde{\eta}^2_\lambda(\mathbf{k}, \omega)$ is

$$\operatorname{Im} \tilde{\eta}^2_\lambda(\mathbf{k}, \omega) = \operatorname{Im} \left(\frac{\omega_p^2}{\omega^2}\right) \sum_{\mu,\lambda} \frac{f_{0\mu}(\mathbf{k}, \lambda) E_\mu(\mathbf{k}) \mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)}{[\mathcal{E}^2_{\mathbf{k}\mu}(\omega) - \omega^2]}. \quad (83)$$

In deriving (83), we have retained only the largest term. Substituting (83) into (79), we get

$$\gamma_{\mathbf{k}\lambda}(\omega) = \pi\omega_p^2 \sum_{\mu,\lambda,\rho} f_{0\mu}(\mathbf{k}, \lambda) E_\mu(\mathbf{k}) \mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega) \times \delta[\mathcal{E}^2_{\mathbf{k}\mu}(\omega) - \omega^2] \left[\frac{d\omega^2}{d\omega^2} \operatorname{Re} \tilde{\eta}^2_\lambda(\mathbf{k}, \omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}. \quad (84)$$

The spectral width of the absorption band for resonance Raman scattering in energy units is of the order of

$$\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] = \pi \left[\frac{\omega_p^2}{\tilde{\omega}_{\rho\lambda}(\mathbf{k})}\right] \sum_{\mu,\lambda,\rho} f_{0\mu}(\mathbf{k}, \lambda) E_\mu(\mathbf{k}) \mathcal{E}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] \times \delta\{\mathcal{E}^2_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] - \tilde{\omega}^2_{\rho\lambda}(\mathbf{k})\} \times \left[\frac{d\omega^2}{d\omega^2} \operatorname{Re} \tilde{\eta}^2(\mathbf{k}, \omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}, \quad (85)$$

with $\mathcal{E}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ given by expression (59) with $\omega = \tilde{\omega}_{\rho\lambda}(\mathbf{k})$. Expression (85) gives the electronic contribution to the spectral width, and at very low temperatures it could be comparable with the contribution arising from the coupling of intermolecular interactions with the phonon field. The energy of excitation $\tilde{\omega}^2_{\rho\lambda}(\mathbf{k})$ is determined by Eq. (82), with $\operatorname{Re} \tilde{\eta}^2_\lambda[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ obtained by taking the real part of expression (67). In the case of exact resonance, terms having a principal value may be taken equal to zero. Then $\operatorname{Re} \tilde{\eta}^2_\lambda[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ turns out to be

$$\operatorname{Re} \tilde{\eta}^2_\lambda[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})] = 1 - \frac{\omega_p^4}{N} \sum_{\mathbf{q},\mu_1,\lambda,\lambda'} \frac{E_\mu(\mathbf{k})}{\mathcal{E}^{(2)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]} \Big| (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1}) [f_{0\mu_1}(\mathbf{q}, \lambda') f_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k}, \lambda) E_{\mathbf{q}\mu_1}/E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{\frac{1}{2}} \times [\omega_{\rho\lambda}(\mathbf{k}) - E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}]^{-1} \Big|^2 \tilde{D}^{(0)}_{\mathbf{q}\lambda'}[\mathbf{k} - \mathbf{q}\mu'; \tilde{\omega}_{\rho\lambda}(\mathbf{k})] + \omega_p^2/N(n_{\mathbf{k}0} - n_{\mathbf{k}\mu}) \times \sum_{\mathbf{q},\mu_1,\mu,\mu',\beta} |f^{\frac{1}{2}}_{0\mu}(\mathbf{k}, \lambda)[U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k} - \mathbf{q}) + U^{\alpha\beta}_{0\mu;\mu_1\mu'}(\mathbf{q})]/[E_{\mathbf{k}-\mathbf{q}\mu'} + E_{\mathbf{q}\mu'} - \tilde{\omega}_{\rho\lambda}(\mathbf{k})]|^2 \times (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'})/\mathcal{E}^{(2)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})](E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'}). \quad (86)$$

The second and third term on the right-hand side of (86) will cause a small shift to the energy of the incoming radiation. For an approximate evaluation of (86), one may replace $\tilde{\omega}^2_{\rho\lambda}(\mathbf{k})$ by c^2k^2.

We now consider the resonance that occurs when $\omega = E_{\mathbf{q}\mu_1} + E_{\mathbf{k}-\mathbf{q}\mu'}$, corresponding to the process where two excitons $(\mathbf{q}\mu_1)$ and $(\mathbf{k} - \mathbf{q}, \mu')$ are created simultaneously by a single photon. Such a process may be also regarded as the decay of a dressed exciton into two excitons.[9] Taking the imaginary part of (67) for the process in question and retaining only the largest terms and using (79), we have

$$\gamma_{\mathbf{k}\lambda}(\omega) = \operatorname{Im} \Big\{ \alpha^{(1)}_\lambda(\mathbf{k}, \omega) E_\mu(\mathbf{k}) \mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega) + \omega^2\omega_p^2/N \times \sum_{\mathbf{q},\lambda',\mu_1} (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1}) f_{0\mu_1}(\mathbf{q}, \lambda')(E_{\mathbf{q}\mu_1}/E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}) \times \alpha_{\mathbf{k}\lambda}(\mathbf{q}\mu_1, \mathbf{k} - \mathbf{q}\mu', \omega) \tilde{D}^{(0)}_{\mathbf{q}\lambda'}(\mathbf{k} - \mathbf{q}, \mu'; \omega) \Big\} \times \sum_\rho \left[\frac{d\omega^2}{d\omega^2} \operatorname{Re} \tilde{\eta}^2_\lambda(\mathbf{k}, \omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}, \quad (87)$$

where

$$\alpha_{\mathbf{k}\lambda}(\mathbf{q}\mu_1, \mathbf{k} - \mathbf{q}\mu', \omega) = \omega_p^2 \sum_{\mu',\lambda} f_{\mathbf{q}\mu_1,\mathbf{k}-\mathbf{q}\mu'}(-\mathbf{k}, \lambda)/(E^2_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1} - \omega^2) \quad (88)$$

is the frequency and wavevector dependent polarizability corresponding to the transition $(\mathbf{q}, \mu_1) \to (\mathbf{k} - \mathbf{q}, \mu')$. For the process under consideration, $\gamma_{\mathbf{k}\lambda}(\omega)$, given by (87), is also a shape function. The calculation is facilitated if we take only the real part of $\mathcal{E}^2_{\mathbf{k}\mu}(\omega)$ in the denominator of the first term on the right-hand side of (87) while in the expression for the function $\tilde{D}^{(0)}_\lambda(\mathbf{k} - \mathbf{q}, \mu'; \omega)$ we take the limit when

$$\operatorname{Im} \tilde{\eta}^2_{\lambda'}(\mathbf{q}, \omega - E_{\mathbf{k}-\mathbf{q}\mu'}) \to 0.$$

Using these approximations, we obtain from (87) the following expression for the spectral width:

$$\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] \approx \Big\{ \alpha^{(1)}_\lambda[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})][E_\mu(\mathbf{k})/\tilde{\omega}_{\rho\lambda}(\mathbf{k})] \times \operatorname{Im} \mathcal{E}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] + \tilde{\omega}_{\rho\lambda}(\mathbf{k})\omega_p^2/N \sum_{\mathbf{q},\mu_1,\lambda'} (n_{\mathbf{q}0} - n_{\mathbf{q}\mu_1}) \times f_{0\mu_1}(\mathbf{q}, \lambda')(E_{\mathbf{q}\mu_1}/E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}) \alpha_{\mathbf{k}\lambda}[\mathbf{q}\mu_1, \mathbf{k} - \mathbf{q}\mu', \tilde{\omega}_{\rho\lambda}(\mathbf{k})] \times \operatorname{Im} \tilde{D}^{(0)}_{\mathbf{q}\lambda'}[\mathbf{k} - \mathbf{q}, \mu'; \tilde{\omega}_{\rho\lambda}(\mathbf{k})] \Big\} \times \sum_\rho \left[\frac{d\omega^2}{d\omega^2} \operatorname{Re} \tilde{\eta}^2_\lambda(k, \omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}, \quad (89)$$

[8] L. N. Ovander, Fiz. Tverd. Tela 4, 1471 (1962) [English transl.: Soviet Phys.—Solid State 4, 1081 (1962)].

[9] L. N. Ovander, Fiz. Tverd. Tela 4, 294 (1962) [English transl.: Soviet Phys.—Solid State 4, 212 (1962)].

where

$$\begin{aligned}&\operatorname{Im}\tilde{D}^{(0)}_{\mathbf{q}\lambda'}[\mathbf{k}-\mathbf{q},\mu';\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\\&\approx\pi(n_{\mathbf{k}-\mathbf{q}0}-n_{\mathbf{k}-\mathbf{q}\mu'})\sum_{\rho'}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}\mu'}+\omega_{\rho'\lambda'}(\mathbf{q})]\\&\times\omega^{-1}_{\rho'\lambda'}(\mathbf{q})\left[\frac{d\omega^2}{d\omega^2}\hat{\eta}^2_{\lambda'}(\mathbf{q},\omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho'\lambda'}(\mathbf{q})}\\&\times\delta\{[\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}\mu'}]^2\hat{\eta}^2_{\lambda'}[\mathbf{q},\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}\mu'}]-c^2q^2\}\end{aligned}\tag{90}$$

in the limit when $\operatorname{Im}\hat{\eta}^2_{\lambda'}[\mathbf{q},\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}\mu'}]\to 0$. In the derivation of (90), only the first term in the expression (65) has been retained. The energy of excitation $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$ is determined by the equation

$$\begin{aligned}&\tilde{\omega}^2_{\rho\lambda}(\mathbf{k})\Bigg\{1+\alpha^{(1)}_{\lambda}[\mathbf{k},\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\Bigg[1+1/N\\&\times\sum_{\mathbf{q},\mu_1,\mu',\beta}\frac{|U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k}-\mathbf{q})+U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q})|^2}{E_{\mu}(\mathbf{k})(E_{\mathbf{q}\mu_1}+E_{\mathbf{k}-\mathbf{q}\mu'})}\\&\times(n_{\mathbf{q}0}-n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}\mu_1}-n_{\mathbf{k}-\mathbf{q}\mu'})\Bigg]\Bigg\}-c^2k^2=0,\end{aligned}\tag{91}$$

which is derived by taking the real part of (67), equating to zero all terms having a principal value, and employing (82). The second and the third term on the left-hand side of (91) will cause an energy shift which, apart from the very small third term, is entirely determined by the frequency and wavevector dependent polarizability. From (91) we see that, apart from very small corrections, the energy of excitation $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$ is equal to the unperturbed energy of excitation of the dressed exciton $(\mathbf{k},\mu)$ derived in I. In (89), the expression for $\operatorname{Im}\tilde{\varepsilon}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ is given by

$$\begin{aligned}&\operatorname{Im}\tilde{\varepsilon}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\\&\approx\omega_p^2/N(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\sum_{\mathbf{q},\mu_1,\mu',\lambda',\beta}\Bigg|[f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q},\lambda')E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}]^{\frac{1}{2}}\\&+\frac{1}{2}\left[\frac{U^{\alpha\beta}_{0\mu',\mu_1\mu}(-\mathbf{q})+2U^{\alpha\beta}_{0\mu',\mu\mu_1}(\mathbf{k})+U^{\alpha\beta}_{0\mu_1,\mu'\mu}(\mathbf{k}-\mathbf{q})}{E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{q}\mu_1}-\tilde{\omega}_{\rho\lambda}(\mathbf{k})}\right]\\&\times[f_{0\mu_1}(\mathbf{q},\lambda')E\mathbf{q}_{\mu_1}]^{\frac{1}{2}}(n_{\mathbf{q}0}-n_{\mathbf{q}\mu_1})\Bigg|^2\\&\times\operatorname{Im}\tilde{D}^{(0)}_{\mathbf{q},\lambda'}[\mathbf{k}-\mathbf{q},\mu';\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\\&+\pi(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\sum_{\mathbf{q},\mu_1,\mu',\beta}|U^{\alpha\beta}_{0\mu,\mu'\mu_1}(\mathbf{k}-\mathbf{q})+U^{\alpha\beta}_{0\mu,\mu_1\mu'}(\mathbf{q})|^2\\&\times(n_{\mathbf{q}0}-n_{\mathbf{q}\mu_1})(n_{\mathbf{k}-\mathbf{q}0}-n_{\mathbf{k}-\mathbf{q}\mu'})\\&\times\delta[\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}\mu'}-E_{\mathbf{q}\mu_1}].\end{aligned}\tag{92}$$

To evaluate (92), the sum over $\mathbf{q}$ is replaced by an integral then, after averaging over angles of $\mathbf{q}$, summing over polarizations and integrating over q, we find in the dipole approximation

$$\begin{aligned}&\operatorname{Im}\tilde{\varepsilon}^{(1)}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\\&\approx 4/3c^3(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})(n_{\mathbf{k}-\mathbf{q}'0}-n_{\mathbf{k}-\mathbf{q}'\mu'})\sum_{\mu',\rho}|\mathbf{P}_{\mu\mu'}|^2\\&\times E^2_{\mathbf{k}-\mathbf{q}'\mu',\mathbf{k}\mu}|\omega_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}|\,|u_{\mathbf{k}\mu}|^2\,|u_{\mathbf{k}-\mathbf{q}'\mu'}|^2\\&\times\{\operatorname{Re}\hat{\eta}_{\lambda'}[\mathbf{q}',\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}]+\hat{\eta}_{\lambda'}[\mathbf{q},\omega_{\rho'\lambda'}(\mathbf{q}')]\}\\&\times\left[\frac{d\omega^2}{d\omega^2}\hat{\eta}^2_{\lambda}(\mathbf{q}',\omega)\right]^{-1}_{\omega=\omega_{\rho'\lambda'}(\mathbf{q}')}+O(e^4),\end{aligned}\tag{93}$$

where $(c\mathbf{q}')^2$ is given by the equations

$$\begin{aligned}&[\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}]^2\operatorname{Re}\hat{\eta}^2_{\lambda}[\mathbf{q}',\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}]\\&\qquad=(cq')^2=\omega^2_{\rho'\lambda'}(\mathbf{q}')\hat{\eta}^2_{\lambda'}[\mathbf{q}',\omega_{\rho'\lambda'}(\mathbf{q})].\end{aligned}\tag{94}$$

In deriving (93), we have retained only the first term on the right-hand side of (92), which is proportional to the square of the electronic charge e^2. The remaining terms are of higher order in e arising from intermolecular interactions, are easily evaluated, and may be important for crystals having large transition dipole moments. Expression (93) is the spontaneous emission probability in energy units for the transition $\mu\to\mu'$ expressed in terms of energies which are correct in the zero approximation; effects arising from the dispersion of the photon $(\mathbf{q},\lambda')$ are included as well. In the same way, we can evaluate the second term on the right-hand side of (89). Then, taking into account (93), we derive an expression for the spectral width:

$$\begin{aligned}&\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\\&\approx 4/3c^3\Bigg\{(n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\alpha^{(1)}_{\lambda}[\mathbf{k},\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\left[\frac{E_{\mu}(\mathbf{k})}{\tilde{\omega}_{\rho\lambda}(\mathbf{k})}\right]\sum_{\mu'}|\mathbf{P}_{\mu\mu'}|^2\\&\times E^2_{\mathbf{k}-\mathbf{q}'\mu',\mathbf{k}\mu}|u_{\mathbf{k}\mu}|^2|u_{\mathbf{k}-\mathbf{q}'\mu'}|^2+\tilde{\omega}_{\rho\lambda}(\mathbf{k})\sum_{\mu_1}|\mathbf{P}_{0\mu_1}|^2\\&\times(E^2_{\mathbf{q}'\mu_1}/E_{\mathbf{k}-\mathbf{q}'\mu',\mathbf{k}\mu})(n_{\mathbf{q}'0}-n_{\mathbf{q}'\mu_1})|u_{\mathbf{q}'\mu_1}|^2\\&\times\alpha_{\mathbf{k}\lambda}[\mathbf{q}'\mu_1,\mathbf{k}-\mathbf{q}'\mu',\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\Bigg\}\sum_{\rho,\rho'}|\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}|\\&\times\{\operatorname{Re}\hat{\eta}_{\lambda'}[\mathbf{q}',\tilde{\omega}_{\rho\lambda}(\mathbf{k})-E_{\mathbf{k}-\mathbf{q}'\mu'}]+\hat{\eta}_{\lambda'}[\mathbf{q}';\omega_{\rho'\lambda'}(\mathbf{q}')]\}\\&\times\left[\frac{d\omega^2}{d\omega^2}\hat{\eta}^2_{\lambda'}(\mathbf{q}',\omega)\right]^{-1}_{\omega=\omega_{\rho'\lambda'}(\mathbf{q}')}\\&\times\left[\frac{d\omega^2}{d\omega^2}\operatorname{Re}\hat{\eta}^2_{\lambda}(\mathbf{k},\omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}+O(e^6).\end{aligned}\tag{95a}$$

Expression (95a) for $\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ gives the spectral width at the maximum energy $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$, and may be regarded as corresponding to the process where a dressed exciton $(\mathbf{k},\mu)$ with energy $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$ creates two excitons, a dressed $(\mathbf{q}',\mu_1)$ and a bare $(\mathbf{k}-\mathbf{q}',\mu')$, with energies $\omega_{\rho'\lambda'}(\mathbf{q}')$ and $E_{\mathbf{k}-\mathbf{q}'\mu'}$, respectively. It is easily seen from the whole calculation that the expression for $\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ holds also for the reverse

process, where two excitons recombine into one. From (95a) we can see that the spectral width $\tilde{\gamma}_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ depends on the values of the polarizabilities and spontaneous emission probabilities of the states in question. The higher-order terms in the electronic charge that have been disregarded in (95a) result from the correlation of the intermolecular interactions with the radiation field. Their contribution to the spectral width is much smaller than that of the radiative terms considered in (95a), and they should be taken into consideration only when their magnitude is significant—as could be the case for crystals having large transition dipole moments.

In the same fashion, one can study the resonance occurring when $\omega = \pm(E_{\mathbf{q}\mu_1} - E_{\mathbf{k}-\mathbf{q}\mu'})$, which corresponds to the process where the exciton $(\mathbf{q}, \mu_1)$ is created while the exciton $(\mathbf{k} - \mathbf{q}, \mu')$ is absorbed (and vice versa) by a single photon. The spectral width for such a process is given by expression (95a) if we replace the polarizability $\alpha_{\mathbf{k}\lambda}[\mathbf{q}'\mu_1, \mathbf{k} - \mathbf{q}'\mu', \tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ in the second term on the right-hand side of (95a) by its principal value. The energy of excitation $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$ for the process in question is determined by the equation

$$\tilde{\omega}^2_{\rho\lambda}(\mathbf{k})\{1 + \alpha^{(1)}_{\lambda}[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})] + \beta_{\mathbf{k}\lambda}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\} - c^2k^2 = 0, \tag{95b}$$

if integrals having a principal value are disregarded. Thus, we see that, in this case, the first term on the right-hand side of (95a) gives practically the spectral width at the maximum energy $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$, given by (95b). We should note here that, in the derivation of (86), (91), and (95b), terms having a principal value have been disregarded; this is justifiable only in the case of exact resonance, but for energies near resonance these terms should be also included in the corresponding expressions.

In the limiting case when $\operatorname{Im} \tilde{\eta}^2_{\lambda}(\mathbf{k}, \omega) \to 0$, the spectral intensity $J^{(1)}_{\mathbf{k}\lambda}(\omega)$ becomes a delta function given by (80), and the energy of excitation is determined by equation (77), with $\operatorname{Re} \tilde{\eta}^2_{\lambda}(\mathbf{k}, \omega)$ obtained from the expression (67), which is now real. For such a process, $\alpha^{(2)}_{\mathbf{k}\lambda}(\mathbf{q}, \lambda', \omega)$ and $\beta_{\mathbf{k}\lambda}(\omega)$ may be interpreted as the probability amplitudes for Raman scattering and double excitation, respectively, and they may be utilized for the evaluation of the intensities of the scattered radiation for the processes in question.

IV. DIELECTRIC PERMEABILITY OF MOLECULAR CRYSTALS

We perform here a calculation for the dielectric permeability of a molecular crystal. As is well known, the passage of electromagnetic waves through dielectric media is characterized by a complex dielectric permeability tensor

$$\epsilon_{ij}(\mathbf{k}, \omega) = \operatorname{Re} \epsilon_{ij}(\mathbf{k}, \omega) + i \operatorname{Im} \epsilon_{ij}(\mathbf{k}, \omega),$$

which is related to the complex electrical conductivity tensor,

$$\sigma_{ij}(\mathbf{k}, \omega),$$

by

$$\epsilon_{ij}(\mathbf{k}, \omega) = \delta_{ij} + (4\pi i/\omega)\sigma_{ij}(\mathbf{k}, \omega). \tag{96}$$

Thus, in order to find the expression for $\epsilon_{ij}(\mathbf{k}, \omega)$, we need to calculate the complex conductivity tensor.

The electrical conductivity tensor is given by[10]

$$\sigma_{ij}(\mathbf{k}, \omega) = i\int_{-\infty}^{+\infty} dt\theta(t)e^{-i\mathbf{k}\cdot\mathbf{r}_{n\alpha}} \times \langle[j^i_{m\beta}(x_0), P^j_{n\alpha}(x'_0)]_-\rangle e^{i\omega t}, \tag{97}$$

where $j^i_{m\beta}(x_0)$ and $P^j_{n\alpha}(x'_0)$ are the i and j components of the current density and dipole moment operators of the molecules $(\mathbf{m}\beta)$ and $(\mathbf{n}\alpha)$ at time x_0 and x'_0, respectively; $\mathbf{k}$ is the wavevector of the external field and $\theta(t)$ is the usual step function. Integrating (97) by parts and using the relation

$$(d/dt')P_{n\alpha}(t') = \sum_{n\alpha} j_{n\alpha}(t'),$$

we have

$$\sigma_{ij}(\mathbf{k}, \omega) = -\frac{1}{\omega}\langle[j^i_{m\beta}(t), P^j_{n\alpha}(t)]_-\rangle e^{-i\mathbf{k}\cdot\mathbf{r}_{n\alpha}} + \frac{2\pi i}{\omega}\sum_{n\alpha} e^{-i\mathbf{k}\cdot\mathbf{r}_{n\alpha}}\langle\langle j^i_{m\beta}(t); j^j_{n\alpha}(t')\rangle\rangle_{(\omega)}, \tag{98}$$

where the quantity $\langle\langle j^i_{m\beta}(t); j^j_{n\alpha}(t')\rangle\rangle_{(\omega)}$ is the Fourier transform of the "current–current" retarded double time Green's function defined[10] by (see, for instance, I)

$$\langle\langle j^i_{m\beta}(t); j^j_{n\alpha}(t')\rangle\rangle = -i\theta(t - t')\langle[j^i_{m\beta}(t), j^j_{n\alpha}(t')]\rangle.$$

If we now express in (98) the current and momentum operators in the second quantization representation, take the term of the conductivity tensor that corresponds to the interband transitions, and, by means of (2), express the operators in the representation where $\mathcal{H}_0$ is diagonal, then substitution of the result into (96) yields the following relation:

$$\omega^2[\epsilon_{ij}(\mathbf{k}, \omega) - \hat{\eta}^2_{\infty}\delta_{ij}] = -\left(\frac{ck}{\pi}\right)\prod_{\mathbf{k},ij}(\omega)\delta_{ij}, \tag{99}$$

where the expression for $\prod_{\mathbf{k},ij}(\omega)$ is given by (48) with the exception that, instead of the directions of polarization $\lambda(\lambda = \mu_{\perp} = 1, 2)$, we have the components $i, j(i, j = 1, 2, 3)$. The relation (99) corresponds to the normal waves, and, since the normal waves in any

[10] D. N. Zubarev, Usp. Fiz. Nauk 71, 71 (1960) [English transl.: Soviet Phys.—Usp. 3, 320 (1960)]; see also Ref. 7, p. 107.

direction are either transverse or longitudinal, the expression for $\prod_{\mathbf{k},ij}(\omega)$ includes both kinds of waves. Using (50), we may rewrite (99) in the form

$$\omega^2[\epsilon_{ij}(\mathbf{k},\omega)-\hat{\eta}_\infty^2\delta_{ij}] = -\left(\frac{ck}{\pi}\right)\hat{\prod}_{\mathbf{k},ij}(\omega)[1+\hat{\prod}_{\mathbf{k},ij}(\omega)D_{\mathbf{k},ij}(\omega)]\delta_{ij}. \quad (100)$$

This is an exact formula for the model under consideration, relating the dielectric permeability tensor to the polarization operator and the Green's function for the photon field. If we retain only the linear term in the polarization operator on the right-hand side of (100) and replace ω by $i|\omega_n|$, the resulting expression for $\epsilon(\mathbf{k}, i|\omega_n|)$ is identical (apart from ω_p^2 which is also included in the polarization operator and cancelled in the final result) to that derived by Dzyaloshinskii *et al.*[2] using the diagram technique. Formula (100) is applicable for any isotropic crystal (anisotropic effects could also be included) for which the tight binding picture is justifiable, where the photons of the electromagnetic field act on electron–hole pairs that are tightly bound at the lattice sites.

In the first approximation, the expression for $\prod_{\mathbf{k},ij}(\omega)\delta_{ij}$ is given by (71), i.e.,

$$\prod^{(1)}_{\mathbf{k},ij}(\omega)\delta_{ij} = \left(\frac{\pi}{ck}\right)\left\{1-\frac{\omega^2[\tilde{\eta}_\lambda^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]}{[\omega^2\tilde{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2]}\right\} \times\{-\omega^2[\tilde{\eta}_i^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]\}. \quad (101)$$

If scattering effects are disregarded and only dispersion is considered, (101) becomes

$$\prod^{(0)}_{\mathbf{k},ij}(\omega)\delta_{ij} = \left(\frac{\pi}{ck}\right)\left\{1-\frac{\omega^2[\hat{\eta}_\lambda^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]}{[\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2]}\right\} \times\{-\omega^2[\hat{\eta}_i^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]\}. \quad (102)$$

In Formulas (101) and (102), $i=j=\lambda=1,2$ denote the transverse components perpendicular to the direction of propagation, while $i=j=3$ is the longitudinal one. Substituting (101) or (102) into (99) and separating the transverse and longitudinal components, we have, respectively, the following. Transverse:

$$\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega) = \tilde{\eta}_\lambda^2(\mathbf{k},\omega) - \frac{\omega^2[\tilde{\eta}_\lambda^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]^2}{[\omega^2\tilde{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2]}, \quad (103)$$

or

$$\epsilon^{(0)}_{\lambda,\perp}(\mathbf{k},\omega) = \hat{\eta}_\lambda^2(\mathbf{k},\omega) - \frac{\omega^2[\hat{\eta}_\lambda^2(\mathbf{k},\omega)-\hat{\eta}_\infty^2]^2}{[\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2]}; \quad (104)$$

Longitudinal:

$$\epsilon^{(1)}_{\|}(\mathbf{k},\omega) = \eta^{(1)2}_{\|}(\mathbf{k},\omega) = 1+\alpha^{(1)}_{\|}(\mathbf{k},\omega), \quad (105)$$

or

$$\epsilon^{(0)}_{\|}(\mathbf{k},\omega) = \hat{\eta}^2_{\|}(\mathbf{k},\omega) = 1+\alpha_{\|}(\mathbf{k},\omega). \quad (106)$$

In (103) and (104), the notation $\epsilon^{(1)}_{\lambda\lambda,\perp}(\mathbf{k},\omega)\equiv\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega)$ has been introduced. The perturbed and unperturbed longitudinal polarizabilities are defined by

$$\alpha^{(1)}_{\|}(\mathbf{k},\omega) = (n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\omega_p^2\sum_{\mu_\|} f_{0\mu}(\mathbf{k})/[\mathcal{E}^2_{\mathbf{k}\mu_\|}(\omega)-\omega^2], \quad (107)$$

and

$$\alpha_{\|}(\mathbf{k},\omega) = (n_{\mathbf{k}0}-n_{\mathbf{k}\mu})\omega_p^2\sum_{\mu_\|} f_{0\mu}(\mathbf{k})/[E^2_{\mu_\|}(\mathbf{k})-\omega^2], \quad (108)$$

respectively. In (107),

$$\mathcal{E}^2_{\mathbf{k}\mu_\|}(\omega)\equiv\mathcal{E}^{(1)}_{\mathbf{k}\mu_\|}(\omega)\mathcal{E}^{(2)}_{\mathbf{k}\mu_\|}(\omega),$$

where $\mathcal{E}_{\mathbf{k}\mu_\|}(\omega)$ is the perturbed energy of the longitudinal exciton $(\mathbf{k},\mu_\|)$ and $\mathcal{E}^{(1)}_{\mathbf{k}\mu_\|}(\omega)$, $\mathcal{E}^{(2)}_{\mathbf{k}\mu_\|}(\omega)$ are given by the longitudinal parts of expressions (59) and (60), respectively. Expression (106) is in agreement with that found by Agranovich and Ginzburg.[11] The frequencies of the longitudinal waves are obtained by equating to zero the right-hand side of (105) or (106).

Using the results of I, one can easily prove that, for real $\hat{\eta}_\lambda^2(\mathbf{k},\omega)$, the following relation holds:

$$(\pi/ck)D^{(0)-1}_{\mathbf{k}\lambda}(\omega)/\omega^2 = \hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2/\omega^2 = \left\{1+\sum_\rho\frac{\omega^2_{p\lambda}(\mathbf{k})}{\omega^2-\omega^2_{\rho\lambda}(\mathbf{k})}\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}\right\}^{-1} \quad (109)$$

or

$$\hat{\eta}_\lambda^2(\mathbf{k},\omega) = \frac{c^2k^2}{\omega^2} + \left\{1+\sum_\rho\frac{\omega^2_{\rho\lambda}(\mathbf{k})}{\omega^2-\omega^2_{\rho\lambda}(\mathbf{k})}\left(\frac{d\omega^2}{d\omega^2}\hat{\eta}_\lambda^2(\mathbf{k},\omega)\right)^{-1}_{\omega=\omega_{\rho\lambda}(\mathbf{k})}\right\}^{-1}, \quad (110)$$

where $\omega^2_{\rho\lambda}(\mathbf{k})$ is the ρth root of the secular equation

$$\omega^2\hat{\eta}_\lambda^2(\mathbf{k},\omega)-c^2k^2=0. \quad (111)$$

It can easily be shown that the formula (110) is identical to that derived by Agranovich and Konobeev[12] for the unperturbed part of the dielectric permeability. Expression (110) is also valid when scattering is taken into account, provided that $\hat{\eta}_\lambda^2(\mathbf{k},\omega)$ in (110) and (111) is replaced by the real part of $\tilde{\eta}_\lambda^2(\mathbf{k},\omega)$ and $\omega_{\rho\lambda}(\mathbf{k})$ by $\tilde{\omega}_{\rho\lambda}(\mathbf{k})$. The last term on the right-hand side of (103) or (104) is a second-order correction, arising from the term which is proportional to the square of the polarization operator and accounting for the

[11] V. M. Agranovich and V. L. Ginzburg, Usp. Fiz. Nauk 77, 663 (1962) [English transl.: Soviet Phys.—Usp. 5, 675 (1963)].

[12] V. M. Agranovich and Yu. V. Konobeev, Fiz. Tverd. Tela 5, 2544 (1963) [English transl.: Soviet Phys.—Solid State 5, 1858 (1964)].

difference between the square of the index of refraction and the dielectric permeability. We write the expression for $\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega)$ as

$$\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega) = \operatorname{Re}\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega) + i\operatorname{Im}\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega), \quad (112)$$

where $\operatorname{Re}\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega)$ is the real part of the expression (103) while the imaginary part of $\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega)$ is given by

$$\begin{aligned}\operatorname{Im}\epsilon^{(1)}_{\lambda,\perp}(\mathbf{k},\omega) &= (\omega^2 - c^2k^2 - \omega_p^2)^2/\omega^2(\pi/ck)\operatorname{Im} D^{(1)}_{\mathbf{k}\lambda}(\omega)\\ &= \frac{(\omega^2 - c^2k^2 - \omega_p^2)^2}{\omega^2}\\ &\quad\times\frac{\omega^2\operatorname{Im}\tilde{\eta}^2_\lambda(\mathbf{k},\omega)}{[\omega^2\operatorname{Re}\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]^2 + [\omega^2\operatorname{Im}\tilde{\eta}^2_\lambda(\mathbf{k},\omega)]^2}.\end{aligned} \quad (113)$$

In the limiting case when $\operatorname{Im}\tilde{\eta}^2_\lambda(\mathbf{k},\omega) \ll 1$ and $d\operatorname{Im}\tilde{\eta}^2_\lambda(\mathbf{k},\omega)/d\omega^2 \ll 1$, the spectrum described by (113) is the same under the same conditions as that of the spectral intensity for the photon field $J^{(1)}_{\mathbf{k}\lambda}(\omega)$, determined by either (75) or (78). The most useful application of formula (100) is the case when the exciton–phonon interaction is included in the expression for the polarization operator and photon Green's function. This will be the subject of a later publication, where the temperature dependence of the dielectric permeability will be discussed.

V. AVERAGE ENERGY OF THE CRYSTALS

The average energy of the crystal in the first approximation, where scattering effects are included, is obtained by averaging the total Hamiltonian, i.e.,

$$\begin{aligned}\langle\mathcal{H}\rangle^{(1)} &= \langle\mathcal{H}_0\rangle + \sum_{\mu,\mathbf{k}} E_\mu(\mathbf{k})\langle b^\dagger_\mu(\mathbf{k})b_\mu(\mathbf{k})\rangle^{(1)}\\ &\quad + \langle\mathcal{H}_{\mathrm{ph}} + \mathcal{H}^{(2)(0)}_{\mathrm{int}} + \mathcal{H}^{\mathrm{I}}_{\mathrm{int}} + \mathcal{H}^{\mathrm{II}}_{\mathrm{int}} + \mathcal{H}^{\mathrm{III}}_{\mathrm{int}} + \mathcal{H}^{\mathrm{IV}}_{\mathrm{int}}\rangle^{(1)},\end{aligned} \quad (114)$$

where

$$\langle\mathcal{H}_{\mathrm{ph}}\rangle^{(1)} = \sum_{\mathbf{k},\lambda} ck\langle\beta^\dagger_\lambda(\mathbf{k})\beta_\lambda(\mathbf{k})\rangle^{(1)}, \quad (115)$$

$$\langle\mathcal{H}^{(2)(0)}_{\mathrm{int}}\rangle^{(1)} = \omega_p^2/4\sum_{\mathbf{k},\lambda} 1/ck\langle\tilde{\beta}^\dagger_\lambda(\mathbf{k})\tilde{\beta}_\lambda(\mathbf{k})\rangle^{(1)}, \quad (116)$$

$$\begin{aligned}\langle\mathcal{H}^{\mathrm{I}}_{\mathrm{int}}\rangle^{(1)} &= \tfrac{1}{2}i\omega_p\sum_{k,\lambda,\mu}[f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})/ck]^{\frac{1}{2}}\\ &\quad\times\langle\tilde{b}_\mu(\mathbf{k})\tilde{\beta}_\lambda(\mathbf{k})\rangle^{(1)},\end{aligned} \quad (117)$$

$$\begin{aligned}\langle\mathcal{H}^{\mathrm{II}}_{\mathrm{int}}\rangle^{(1)} &= i\omega_p/N^{\frac{1}{2}}\sum_{\mathbf{k},\mathbf{q},\lambda}[f_{\mathbf{k}\mu,\mathbf{k}-\mathbf{q}\mu'}(\mathbf{q},\lambda)E_{\mathbf{k}-\mathbf{q}\mu',\mathbf{k}\mu}/cq]^{\frac{1}{2}}\\ &\quad\times\langle b^\dagger_\mu(\mathbf{k})b_{\mu'}(\mathbf{k}-\mathbf{q})\tilde{\beta}_\lambda(\mathbf{q})\\ &\quad + b^\dagger_{\mu'}(\mathbf{k}-\mathbf{q})b_\mu(\mathbf{k})\tilde{\beta}^\dagger_\lambda(\mathbf{q})\rangle^{(1)},\end{aligned} \quad (118)$$

$$\langle\mathcal{H}^{\mathrm{III}}_{\mathrm{int}}\rangle^{(1)} = 1/N^{\frac{1}{2}}\sum_{\mathbf{k},\mu}\langle b^\dagger_\mu(\mathbf{k})\hat{U}^{\alpha\beta}_{\mathbf{k},\mathbf{q}} + \hat{U}^{\alpha\beta\dagger}_{-\mathbf{k},-\mathbf{q}}b_\mu(\mathbf{k})\rangle^{(1)}, \quad (119)$$

$$\begin{aligned}\langle\mathcal{H}^{\mathrm{IV}}_{\mathrm{int}}\rangle^{(1)} &= 2/N\sum_{\mathbf{q},\mu',\beta}U^{\alpha\beta}_{\mu\mu',\mu'\mu}(\mathbf{q})(n_{\mathbf{k}-\mathbf{q}0} - n_{\mathbf{k}-\mathbf{q}\mu'})\\ &\quad\times\langle b^\dagger_\mu(\mathbf{k})b_\mu(\mathbf{k}) + b_\mu(\mathbf{k})b^\dagger_\mu(\mathbf{k})\\ &\quad + b^\dagger_\mu(\mathbf{k})b^\dagger_\mu(-\mathbf{k}) + b_\mu(-\mathbf{k})b_\mu(\mathbf{k})\rangle^{(1)}.\end{aligned} \quad (120)$$

The single-particle Green's functions $G^{(1)}_\mu(\mathbf{k};\omega)$ and $\hat{G}^{(1)}_\mu(\mathbf{k};\omega)$ are obtained from (39) and (40), respectively, as

$$\begin{aligned}G^{(1)}_\mu(\mathbf{k};\omega) &= (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi\\ &\quad\times\left\{\frac{(\omega^2 - c^2k^2 - \omega_p^2)[\omega + \mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)]^2}{[\omega^2 - \mathcal{E}^2_{\mathbf{k}\mu}(\omega)][\omega^2\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]} - 1\right\}\frac{1}{2\mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)},\end{aligned} \quad (121)$$

$$\begin{aligned}\hat{G}^{(1)}_\mu(\mathbf{k};\omega) &= (n_{\mathbf{k}0} - n_{\mathbf{k}\mu})/2\pi\\ &\quad\times\left\{\frac{(\omega^2 - c^2k^2 - \omega_p^2)[\omega^2 - \mathcal{E}^{(1)2}_{\mathbf{k}\mu}(\omega)]}{[\omega^2 - \mathcal{E}^2_{\mathbf{k}\mu}(\omega)][\omega^2\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]} - 1\right\}\frac{1}{2\mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega)},\end{aligned} \quad (122)$$

while the Green's function $\langle\langle\tilde{b}_\mu(\mathbf{k});\tilde{\beta}_\lambda(\mathbf{k})\rangle\rangle^{(1)}$ is easily evaluated by means of (15) and (20):

$$\begin{aligned}&i\omega_p/2\sum_{\mathbf{k},\mu,\lambda}[f_{0\mu}(\mathbf{k},\lambda)E_\mu(\mathbf{k})/ck]^{\frac{1}{2}}\langle\langle\tilde{b}_\mu(\mathbf{k});\tilde{\beta}_\lambda(\mathbf{k})\rangle\rangle^{(1)}\\ &\qquad = \tfrac{1}{2}\pi\sum_{\mathbf{k},\lambda}\frac{\omega^2[\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - 1] + \omega_p^2}{[\omega^2\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]}.\end{aligned} \quad (123)$$

The photon Green's functions $D^{(1)}_{\mathbf{k}\lambda}(\omega)$ and

$$\langle\langle\beta_\lambda(\mathbf{k});\beta^\dagger_\lambda(\mathbf{k})\rangle\rangle^{(1)}$$

are given by expressions (73) and

$$\begin{aligned}&\langle\langle\beta_\lambda(\mathbf{k});\beta^\dagger_\lambda(\mathbf{k})\rangle\rangle^{(1)}\\ &\qquad = \tfrac{1}{2}\pi\frac{\omega + ck - \omega^2[\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - 1]/2ck}{[\omega^2\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]},\end{aligned} \quad (124)$$

respectively.

To evaluate the distribution functions that appear in the expression (114)–(120) we need to know the corresponding expressions for the Green's functions in the first approximation. Thus, using (121)–(124) and the expressions for (118) and (119) that were obtained by calculating (19) and (21), respectively, we find the expression for the average energy:

$$\begin{aligned}\langle\mathcal{H}\rangle^{(1)} &= \langle\mathcal{H}_0\rangle + 1/\pi\sum_{\mathbf{k},\mu}\int_{-\infty}^{+\infty}d\omega\operatorname{Im}\frac{(1 - e^{\beta\omega})^{-1}}{[\omega^2\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - c^2k^2]}\Big(\omega[\omega + ck + (\omega^2 - c^2k^2 - \omega_p^2)/(\omega - \mathcal{E}_{\mathbf{k}\mu}(\omega))]\\ &\quad + \frac{(\omega^2 - c^2k^2 - \omega_p^2)}{\omega + \mathcal{E}_{\mathbf{k}\mu}(\omega)}\{E_\mu(\mathbf{k})[\omega - \tfrac{1}{2}(\mathcal{E}^{(1)}_{\mathbf{k}\mu}(\omega) + \mathcal{E}^{(2)}_{\mathbf{k}\mu}(\omega))]/[\omega - \mathcal{E}_{\mathbf{k}\mu}(\omega)] - \mathcal{E}_{\mathbf{k}\mu}(\omega)\}\\ &\quad - \tfrac{1}{2}\{\omega^2[\tilde{\eta}^2_\lambda(\mathbf{k},\omega) - 1] + \omega_p^2\}\{[1 - E_\mu(\mathbf{k})/\mathcal{E}_{\mathbf{k}\mu}(\omega)]\}\Big).\end{aligned} \quad (125)$$

To carry out the integration over ω in (125), we assume that ω is far from any of the absorption frequencies of the crystal, which means that $\tilde{\eta}_\lambda^2(\mathbf{k}, \omega)$ is real. Employing the relation

$$2\,\mathrm{Im}\,\frac{1}{[\omega^2\tilde{\eta}_\lambda^2(k,\omega) - c^2k^2]} = \sum_\rho \left[\frac{d\omega^2}{d\omega^2}\tilde{\eta}_\lambda^2(\mathbf{k},\omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}\left[\frac{\pi}{\tilde{\omega}_{\rho\lambda}(\mathbf{k})}\right] \times \{\delta[\omega - \tilde{\omega}_{\rho\lambda}(\mathbf{k})] - \delta[\omega + \tilde{\omega}_{\rho\lambda}(\mathbf{k})]\}, \tag{126}$$

where $\tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$ is the ρth root of the secular equation (77), we integrate over ω with the result

$$\langle\mathcal{H}\rangle^{(1)} = \langle\mathcal{H}_0\rangle + \tfrac{1}{2}\sum_{\mathbf{k},\rho}\Big(\{1 + \Omega_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})\coth\tfrac{1}{2}\beta\tilde{\omega}_{\rho\lambda}(\mathbf{k}) - \mathcal{E}_{\mathbf{k}\mu}] + (\mathcal{E}_{\mathbf{k}\mu} - ck)\Big)\left[\frac{d\omega^2}{d\omega^2}\tilde{\eta}_\lambda^2(\mathbf{k},\omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}$$
$$+ \tfrac{1}{2}\sum_{\mathbf{k},\rho}\Omega_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]\Bigg([\mathcal{E}_{\mathbf{k}\mu} - E_\mu(\mathbf{k})]\left[1 + \frac{\tilde{\omega}_{\rho\lambda}(\mathbf{k})}{2\mathcal{E}_{\mathbf{k}\mu}}\coth\tfrac{1}{2}\beta\tilde{\omega}_{\rho\lambda}(\mathbf{k})\right] - \{E_\mu(\mathbf{k})(\mathcal{E}_{\mathbf{k}\mu}^{(1)} + \mathcal{E}_{\mathbf{k}\mu}^{(2)}) - \mathcal{E}_{\mathbf{k}\mu}[E_\mu(\mathbf{k}) + \mathcal{E}_{\mathbf{k}\mu}]\}$$
$$\times \coth\tfrac{1}{2}\beta\tilde{\omega}_{\rho\lambda}(\mathbf{k})/2\tilde{\omega}_{\rho\lambda}(\mathbf{k})\Bigg)\left[\frac{d\omega^2}{d\omega^2}\tilde{\eta}_\lambda^2(\mathbf{k},\omega)\right]^{-1}_{\omega=\tilde{\omega}_{\rho\lambda}(\mathbf{k})}. \tag{127}$$

The function $\Omega_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ is defined as

$$\Omega_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})] = -(ck/\pi)\,\tilde{\Pi}_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]/\mathcal{E}_{\mathbf{k}\mu}^2 - \tilde{\omega}_{\rho\lambda}^2(\mathbf{k}). \tag{128}$$

The quantity $\tilde{\Pi}_{\mathbf{k}\lambda}^{(1)}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ is the polarization operator given by expression (77) with $\omega^2 = \tilde{\omega}_{\rho\lambda}^2(\mathbf{k})$; the notation $\mathcal{E}_{\mathbf{k}\mu} \equiv \mathcal{E}_{\mathbf{k}\mu}[\tilde{\omega}_{\rho\lambda}(\mathbf{k})]$ has been introduced.

In order to identify the terms appearing on the right-hand side of (127), we refer to the zero approximation which occurs when we disregard scattering, i.e., when $\mathcal{E}_{\mathbf{k}\mu} = \mathcal{E}_{\mathbf{k}\mu}^{(1)} = \mathcal{E}_{\mathbf{k}\mu}^{(2)} \to E_\mu(\mathbf{k})$ and

$$\tilde{\eta}_\lambda^2[\mathbf{k}, \tilde{\omega}_{\rho\lambda}(\mathbf{k})] \to \tilde{\eta}_\lambda^2[\mathbf{k}, \omega_{\rho\lambda}(\mathbf{k})].$$

In this case the last term on the right-hand side of (127) goes to zero. Using the fact that

$$1 + \Omega_{\mathbf{k}\lambda}^{(0)}[\omega_{\rho\lambda}(\mathbf{k})] = \left[\frac{d\omega^2}{d\omega^2}\tilde{\eta}_\lambda^2(\mathbf{k},\omega)\right]_{\omega=\omega_{\rho\lambda}(\mathbf{k})},$$

we obtain

$$\langle\mathcal{H}\rangle^{(0)} = \langle\mathcal{H}_0\rangle + \tfrac{1}{2}\sum_{\mathbf{k},\rho}\Big(\omega_{\rho\lambda}(\mathbf{k})\coth\tfrac{1}{2}\beta\omega_{\rho\lambda}(\mathbf{k}) - E_\mu(\mathbf{k}) + [E_\mu(\mathbf{k}) - ck]\{1 + \Omega_{\mathbf{k}\lambda}^{(0)}[\omega_{\rho\lambda}(\mathbf{k})]\}^{-1}\Big). \tag{129}$$

This is the average energy of the crystal where the dispersion of the electromagnetic waves is included explicitly. In the limit when $\beta \to \infty$, Expression (129) is reduced to that of the ground state energy

$$\langle\mathcal{H}\rangle_0^{(0)} = \langle\mathcal{H}_0\rangle_0 + \tfrac{1}{2}\sum_{\mathbf{k}\rho}\Big(\omega_{\rho\lambda}(\mathbf{k}) - E_\mu(\mathbf{k}) + [E_\mu(\mathbf{k}) - ck]\{1 + \Omega_{\mathbf{k}\lambda}^{(0)}[\omega_{\rho\lambda}(\mathbf{k})]\}^{-1}\Big). \tag{130}$$

Expression (130) is identical to that derived in I for the ground state energy of the crystal when only dispersion of the polarization waves is considered. Comparing Expression (127) with that of (129), we see that the second term on the right-hand side of (127) is the renormalized energy of interaction corresponding to the last term of (129) when scattering is neglected, while the last term of (127) depends on the difference $\mathcal{E}_{\mathbf{k}\mu} - E_\mu(\mathbf{k})$, and is therefore much smaller than the second. In the limit when $\beta \to \infty$, $\langle\mathcal{H}\rangle_0^{(1)}$ gives the ground state energy of the crystal.

Thus, the second and last term on the right-hand side of (127) give contributions to the average binding energy of a molecular crystal arising from the dispersion and scattering of the polarization waves as well as from exciton–exciton interactions at finite temperatures. They are correct in the first approximation where terms proportional to N^{-1} are included and are expressed in terms of such quantities as the polarization operator (or the index of refraction) and the energies of excitation of both the dressed and bare exciton $(\mathbf{k}, \mu)$. The discussion of the importance of these terms for actual crystals will be postponed until numerical calculations are performed; this will be the subject of a later publication.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Ground-State Energy of a Finite System of Charged Particles

FREEMAN J. DYSON
The Institute for Advanced Study, Princeton, New Jersey

(Received 13 December 1966)

A trial wavefunction of superconducting type is postulated for the ground state of a system of N positive and N negative charges with Coulomb interactions in the absence of any exclusion principle. The ground-state binding energy is rigorously proved to be greater than $AN^{\frac{7}{5}}$ Ry, where A is an absolute constant. Results of earlier perturbation-theoretic calculations for an infinite system are confirmed. The author, with A. Lenard, has previously proved that the exclusion principle, holding for particles with one sign of charge only, is a sufficient condition for the stability of matter; the present paper shows that the exclusion principle is also necessary for stability.

I. INTRODUCTION

"WE take a piece of metal. Or a stone. When we think about it, we are astonished that this quantity of matter should occupy so large a volume. Admittedly, the molecules are packed tightly together, and likewise the atoms within each molecule. But why are the atoms themselves so big?

"Consider for example the Bohr model of an atom of lead. Why do so few of the 82 electrons run in the orbits close to the nucleus? The attraction of the 82 positive charges in the nucleus is so strong. Many more of the 82 electrons could be concentrated into the inner orbits, before their mutual repulsion would become too large. What prevents the atom from collapsing in this way? Answer: only the Pauli principle, 'No two electrons in the same state.' That is why atoms are so unnecessarily big, and why metal and stone are so bulky.

"You must admit, Pauli, that if you would only partially repeal your prohibition, you could relieve many of our practical worries, for example the traffic problem on our streets."

These words were addressed by Paul Ehrenfest to Pauli in 1931 on the occasion of the award of the Lorentz medal.[1] We have been unable to find in the literature of the 1920's and 1930's any more exact calculation of what would happen to matter if the exclusion principle were abolished. In the present paper we demonstrate that the effects would be even more drastic than those envisaged by Ehrenfest. We show that not only individual atoms but matter in bulk would collapse into a condensed high-density phase. The assembly of any two macroscopic objects would release energy comparable to that of an atomic bomb. It is thus fortunate that Pauli was unwilling to comply with Ehrenfest's well-intentioned proposal.

As a simple model to illustrate the nature of the problem, we consider a system of N positive and N negative charges, all having equal mass m and equal magnitude of the charge e. The Hamiltonian of this system is

$$H = \sum_{j=1}^{2N} \frac{p_j^2}{2m} + \sum\sum_{i<j} \frac{e^2 u_i u_j}{|x_i - x_j|}, \tag{1}$$

where x_j, p_j are position and momentum of particle number j, and $u_j = \pm 1$ accordingly as the particle has positive or negative charge. We suppose that the system obeys the rules of nonrelativistic quantum mechanics without any exclusion principle. We find a wavefunction which we conjecture to be a good approximation to the ground state, and for which the expectation value of H can be calculated exactly. This leads to a rigorous proof of the following statement, which was announced earlier[2] in a tentative way.

Theorem: The ground-state energy E_N of the Hamiltonian (1) satisfies the inequality

$$E_N < -A_1 N^{\frac{7}{5}} \text{ Ry}, \quad \text{Ry} = (me^4/2\hbar^2), \tag{2}$$

where A_1 is an absolute constant.

The binding energy of a macroscopic number of charges ($N \sim 10^{23}$), according to Eq. (2), would be at least of the order of 10^{32} Ry or 1 megaton.

Real matter differs from the model (1) in many respects. First, the masses of nuclei and electrons are not equal. This difference does not weaken our conclusion, because Eq. (2) will continue to hold as long as all particles have masses not less than the value m which can be taken to be the electron mass. Second, nuclei and electrons have different charges. This

[1] P. Ehrenfest, *Collected Scientific, Papers* M. J. Klein, Ed. (North-Holland Publishing Company, Amsterdam, 1959), p. 617. The address appeared originally in Versl. Akad. Amsterdam **40**, 121 (1931). The author is indebted to Dr. H. B. G. Casimir for this reference.

[2] F. J. Dyson and A. Lenard, J. Math. Phys. **8**, 423 (1967).

difference also does not weaken our conclusion, although the details are rather more complicated if the charges are unequal. Provided that the over-all system is approximately neutral, Eq. (2) will hold for unequal charges if each charge is not less than the electron charge e. Third, real matter contains many effects, such as nuclear forces and the finite sizes of nuclei, which are not represented in the model. These differences again do not really weaken our argument. If in a piece of real matter the exclusion principle ceased to operate, the matter would collapse to a density so high that nuclear forces would become dominant. Instead of Eq. (2), the energy available in nuclear reactions would determine the over-all binding. In particular, any material containing hydrogen would be a nuclear explosive. Thus our theorem, although mathematically exact only for a simple model, shows quite generally that matter without the exclusion principle is unstable.

The result (2) is not unexpected, if one looks at perturbation-theory calculations which have been made of the ground-state energy of a Bose gas with Coulomb interactions.[3] These calculations have always considered an infinite uniform system or a system in a box with periodic boundary conditions. They give an energy per particle

$$E = -A_2[\rho a^3]^{\frac{1}{4}}\,\mathrm{Ry}, \tag{3}$$

where ρ is the number of particles per unit volume, a is the Bohr radius, and A_2 is another absolute constant. If one assumes that Eq. (3) also holds in some approximate sense for a finite system of $2N$ particles, then the binding energy of the finite system will be proportional to $N^{\frac{7}{5}}$. However, the step from Eq. (3) to the finite system is not rigorous, and Eq. (3) itself is only the leading term in a perturbation series. The series in inverse powers of $[\rho a^3]^{\frac{1}{4}}$, of which two terms have been calculated,[3] may or may not converge, and the perturbation calculations give no control of the error in Eq. (3). As a by-product of our study of the finite system, we recover the formula (3) for the ground-state energy of the infinite uniform system, thus verifying that our wavefunction preserves the essential features of the perturbation-theory treatment.

In our earlier paper,[2] we proved the lower bound

$$E_N > -A_3 N^{\frac{5}{3}}\,\mathrm{Ry}, \tag{4}$$

which complements Eq. (2). The results of the present paper make it extremely plausible that Eq. (4) actually holds with $\frac{5}{3}$ replaced by $\frac{7}{5}$. There is reason to hope that our trial wavefunction is simultaneously simple enough and accurate enough, so that it may be made the basis of a rigorous proof of an improved lower bound. But the problem of the lower bound is not discussed further in this paper.

There is one philosophical question which is in some sense related to Eq. (2). Suppose that there existed in nature a weakly interacting charged boson (WCB). Then we could construct from N positive and N negative bosons a state with energy satisfying Eq. (2). For large enough N, the binding energy proportional to $N^{\frac{7}{5}}$ would appear to outweigh the rest-energy $2Nmc^2$ necessary to create the particles from the vacuum. The vacuum would be unstable and the world as we know it could not exist. Thus we may claim that our theorem to some extent explains the observed fact that WCB *do not exist in nature*. This argument is unfortunately defective because the densities required to make

$$E_N + 2Nmc^2 < 0 \tag{5}$$

are so high that the nonrelativistic Hamiltonian (1) is inadequate; the stability question ought to be studied within the framework of a fully relativistic theory. A relativistic treatment cannot at present be made rigorous since there exists no fully rigorous relativistic quantum electrodynamics. Our "explanation" for the nonexistence of WCB remains only suggestive and heuristic, not mathematically compelling. Still, it is a striking fact that the known charged bosons, for example the pion and the deuteron, all have strong interactions which, at high densities, would overwhelm their Coulomb binding energy.

II. DIMENSIONAL ARGUMENT

Before beginning the exact analysis, it is useful to explain qualitatively by an elementary argument the origin of the $\frac{7}{5}$ power in Eq. (2). The ground-state wavefunction of a system of $2N$ charges will involve two lengths, the over-all diameter Λ of the system and the range λ of two-particle correlations. The total energy E_N will be a sum of three parts,

$$E_N = E_{K1} + E_{K2} + E_C, \tag{6}$$

E_{K1} being the kinetic energy of the over-all wave packet, E_{K2} being the kinetic energy of short-range correlations, and E_C being the Coulomb energy.

Disregarding numerical factors of the order of unity, we have

$$E_{K1} = N(\hbar^2/m\Lambda^2). \tag{7}$$

[3] L. L. Foldy, Phys. Rev. **124**, 649 (1961); M. Girardeau and R. Arnowitt, *ibid.* **113**, 755 (1959); M. Girardeau, *ibid.* **127**, 1809 (1962); W. H. Bassichis and L. L. Foldy, *ibid.* **133**, A935 (1964); W. H. Bassichis, *ibid.* **134**, A543 (1964); J. M. Stephen, Proc. Phys. Soc. (London) **79**, 994 (1962); D. K. Lee and E. Feenberg, Phys. Rev. **137**, A731 (1965); D. Wright, *ibid.* **143**, 91 (1966); E. H. Lieb, *ibid.* **130**, 2518 (1963); E. H. Lieb and A. Y. Sakakura, *ibid.* **133**, A899 (1964).

The Coulomb energy of the mean charge distribution is zero since the system is, on the average, everywhere neutral. Therefore E_C arises only from the short-range correlations, which produce around each charge a charge cloud containing one unit of charge of the opposite sign distributed over a region of size λ. The interaction of each charge with its charge cloud produces an energy of the order $(-e^2/\lambda)$, while the self-energy of the charge cloud is positive but only half as great. Again discarding a numerical factor, we have

$$E_C = -N(e^2/\lambda). \tag{8}$$

Finally we must estimate E_{K2}, which requires a rather more careful discussion. The charge cloud around each charge is produced by a cooperation of all the particles which are within a volume of the order of λ^3. The number of these particles is

$$\nu = N(\lambda/\Lambda)^3. \tag{9}$$

To produce the net excess or deficiency of one unit of charge within the cloud, each of ν single-particle wavefunctions must be increased or decreased by a factor $[1 \pm \nu^{-1}]^{\frac{1}{2}}$ in passing from the edge to the center of the cloud. This distortion of the wavefunctions produces a kinetic energy of the order of

$$(\hbar^2/m)\,|(\text{grad } \psi)/\psi|^2 \simeq (\hbar^2/m)(\nu\lambda)^{-2} \tag{10}$$

for each particle in each cloud. Altogether then,

$$E_{K2} = N\nu(\hbar^2/m)(\nu\lambda)^{-2} = (\hbar^2\Lambda^3/m\lambda^5). \tag{11}$$

Choosing Λ to minimize the sum of Eqs. (7) and (11), we find

$$\Lambda = N^{\frac{1}{5}}\lambda, \quad E_{K1} + E_{K2} = N^{\frac{3}{5}}(\hbar^2/m\lambda^2). \tag{12}$$

Then, choosing λ to minimize the sum of Eq. (8) and (12),

$$\lambda = N^{-\frac{2}{5}}(\hbar^2/me^2), \quad E_N = -N^{\frac{7}{5}}\,\text{Ry}. \tag{13}$$

This argument shows that the decisive factor in making matter without exclusion principle unstable is the cooperative effect of many particles in screening each other. The charge cloud around each particle is composed not of one or two nearest neighbors, but of a large number $\nu \sim N^{\frac{2}{5}}$ of slightly distorted wavefunctions. This enables the charge clouds to be produced with a very small expenditure of kinetic energy. The exclusion principle makes matter stable by forbidding such a cooperation of many particles with small momentum.

III. DEFINITION OF WAVEFUNCTIONS

Let $\psi_\alpha(x)$ be any complete orthonormal sequence of real functions of the single space point x. A "kinetic energy" integral is defined by

$$T_\alpha = (\hbar^2/2m)\int |\nabla\psi_\alpha(x)|^2\, d_3x, \tag{14}$$

and an "exchange Coulomb energy" integral by

$$e_{\alpha\beta} = e^2\iint \psi_\alpha(x)\psi_\beta(x)\psi_\alpha(y)\psi_\beta(y)\,|x-y|^{-1}\, d_3x\, d_3y. \tag{15}$$

It is convenient to label each particle with a space coordinate x_j and a charge coordinate u_j. Each x_j is a 3-vector, and each u_j takes the values ± 1 to indicate whether the particle is positive or negative. A symmetric two-particle wavefunction is given by

$$G(x_1, u_1, x_2, u_2) = \lambda_0\psi_0(x_1)\psi_0(x_2) - u_1u_2\sum_{\alpha>0}\lambda_\alpha\psi_\alpha(x_1)\psi_\alpha(x_2). \tag{16}$$

This is intended to represent a pair of particles which are mainly distributed independently in the one-particle state ψ_0, but have a short-range correlation which is repulsive for like charges and attractive for unlike charges.

Our basic wavefunction for a $2N$-particle system is

$$\Psi_{2N}(x_1, u_1, \cdots, x_{2N}, u_{2N}) = \sum_P \prod_{j=1}^{N} G(x_{P2j-1}, u_{P2j-1}, x_{P2j}, u_{P2j}), \tag{17}$$

where P is any permutation of the integers $1, \cdots, 2N$, and, for typographical reasons, we write $P2j-1$ instead of P_{2j-1}. This wavefunction is similar in form to the ground-state wavefunction of the Bardeen–Cooper–Schrieffer theory of superconductivity.[4] To obtain a BCS wavefunction, we have only to replace symmetrization by antisymmetrization and replace the charge coordinate by a spin coordinate. The state Ψ_{2N} describes a state in which all the $2N$ particles are correlated in pairs, and each pair has the identical wavefunction G. These boson pairs may be called "Bogoliubov pairs"[5]; they are analogous to the Cooper pairs in the theory of superconductivity.[6]

The state Ψ_{2N} does not have a well-defined total charge. It contains components with N_+ positive and N_- negative particles, where N_+ and N_- are any integers with

$$N_+ + N_- = 2N. \tag{18}$$

We write

$$\Phi_{N_+,N_-} = Q_{N_+,N_-}\Psi_{2N}, \tag{19}$$

where Q_{N_+,N_-} is the projection operator for states of N_+ positive and N_- negative charges. Our trial wavefunction for the ground state of the Hamiltonian (1) is then $\Phi_{N,N}$.

[4] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[5] N. N. Bogoliubov, J. Phys. (USSR) **11**, 23 (1947).
[6] L. N. Cooper, Phys. Rev. **104**, 1189 (1956); M. R. Schafroth, *ibid.* **96**, 1442 (1954).

To make the calculations simple, it is convenient to consider a state of indefinite particle number, namely,

$$\Psi(z) = \sum_{N=0}^{\infty} (w_N)^{\frac{1}{2}} \Psi_{2N} \tag{20}$$

with

$$w_N = (z/16)^N ((2N)!)^{-1} (N!)^{-2}. \tag{21}$$

Here z is a positive real parameter, and $\Psi(z)$ contains components with all possible values of N_+ and N_- such that $(N_+ + N_-)$ is even. Since H commutes with N_+ and N_-, the expectation values are related by

$$H(z) = (M(z)/N(z)), \tag{22}$$

$$M(z) = (\Psi(z), H\Psi(z)) = \sum_N w_N n_{2N} H_{2N}, \tag{23}$$

$$N(z) = |\Psi(z)|^2 = \sum_N w_N n_{2N}, \tag{24}$$

$$n_{2N} H_{2N} = (\Psi_{2N}, H\Psi_{2N}) = \sum_n |\Phi_{n,2N-n}|^2 H_{n,2N-n}, \tag{25}$$

$$n_{2N} = |\Psi_{2N}|^2 = \sum_n |\Phi_{n,2N-n}|^2. \tag{26}$$

Here $H(z)$, H_{2N} and $H_{n,2N-n}$ are the expectation values of H in the states $\Psi(z)$, Ψ_{2N} and $\Phi_{n,2N-n}$. The Rayleigh–Ritz principle asserts that the ground-state energy E_N of Eq. (2) satisfies

$$E_N \leq H_{N,N}.$$

IV. CALCULATION OF EXPECTATION VALUES

We begin by calculating the normalization integral n_{2N} defined by Eq. (26). From this the expectation value of the Hamiltonian can be deduced rather easily. By virtue of Eq. (17), the normalization integral becomes

$$n_{2N} = \sum_P \sum_Q \int \cdots \int d_3x_1 \cdots d_3x_{2N} \times \sum_{u_1} \cdots \sum_{u_{2N}} \prod_{j=1}^{N} \{G(x_{P2j-1}, u_{P2j-1}, x_{P2j}, u_{P2j}) \times G(x_{Q2j-1}, u_{Q2j-1}, x_{Q2j}, u_{Q2j})\}. \tag{27}$$

The symbols P and Q denote independent permutations of the indices $(1, \cdots, 2N)$. For any particular choice of P and Q, the integral on the right of Eq. (27) will break up into a product of cyclic factors of the form

$$K_j = \int \cdots \int dy_1 \cdots dy_{2j} \times \sum_{v_1} \cdots \sum_{v_{2j}} \left\{ \prod_{k=1}^{2j-1} G(y_k, v_k, y_{k+1}, v_{k+1}) \right\} \times G(y_{2j}, v_{2j}, y_1, v_1). \tag{28}$$

Each factor K_j arises from a cycle of $2j$ indices which are connected in pairs by the kernels G in Eq. (27). Thus,

$$n_{2N} = \sum_h \nu(h) \left\{ \prod_{j=1}^{N} (K_j)^{h_j} \right\}, \tag{29}$$

where h_j is the number of cycles of length $2j$, the sum extends over all integers $(h_1, \cdots, h_N)$ consistent with the condition

$$\sum_j j h_j = N, \tag{30}$$

and $\nu(h)$ is the number of pairs of permutations (P, Q) which give rise to the pattern of cycles specified by the h_j.

The evaluation of $\nu(h)$ is a familiar problem of combinatorics. In fact, $\nu(h)$ is a product of three factors, ν', ν'', and ν''', where

$$\nu' = (2N)! \prod_j \{((2j)!)^{-h_j} (h_j!)^{-1}\} \tag{31}$$

is the number of ways of grouping $2N$ indices into the cycles of given lengths, then

$$\nu'' = \prod_j \{(2j)!/(2j)\}^{h_j} \tag{32}$$

is the number of ways of arranging the indices in each group in cyclic order, and

$$\nu''' = (N!)^2 2^{2N} \tag{33}$$

is the number of ways of choosing the labels P_{2j-1}, P_{2j}, Q_{2j-1}, Q_{2j} for the consecutive pairs of indices within each cycle. Thus,

$$\nu(h) = (2N)! (N!)^2 2^{2N} \prod_j \{(2j)^{-h_j} (h_j!)^{-1}\}, \tag{34}$$

which, with Eq. (29), completes the evaluation of n_{2N}.

The utility of the composite state $\Psi(z)$, defined by Eq. (20) and (21), is now apparent. The normalization integral for this state, by Eq. (24), (29), and (34), is

$$N(z) = \sum_h \prod_j \left\{ \left(\frac{z}{4}\right)^{jh_j} \left(\frac{K_j}{2j}\right)^{h_j} (h_j!)^{-1} \right\} = \exp[f(z)] \tag{35}$$

with

$$f(z) = \sum_{j=1}^{\infty} \left(\frac{z}{4}\right)^j \left(\frac{K_j}{2j}\right). \tag{36}$$

The summation over N has disappeared by virtue of Eq. (30). Since K_j is just the trace of the $2j$-fold iterated kernel G^{2j}, Eq. (36) may be written

$$f(z) = -\tfrac{1}{2} \operatorname{Tr} \log [1 - \tfrac{1}{4} z G^2], \tag{37}$$

and Eq. (35) becomes

$$N(z) = \Delta^{-\frac{1}{2}}, \tag{38}$$

where Δ is the Fredholm determinant of the kernel G^2.

With the particular choice of kernel given by Eq. (16),

$$G^2 = 2\sum_\alpha \lambda_\alpha^2 \psi_\alpha(x_1)\psi_\alpha(x_2) \tag{39}$$

is independent of the charge coordinates, and

$$K_j = 2^{2j}\sum_\alpha \lambda_\alpha^{2j}.$$

Equations (36) and (38) then give

$$f(z) = -\tfrac{1}{2}\sum_\alpha \log(1 - z\lambda_\alpha^2), \tag{40}$$

$$N(z) = \prod_\alpha (1 - z\lambda_\alpha^2)^{-\frac{1}{2}}. \tag{41}$$

We take z smaller than the smallest of the λ_α^{-2}, so that the series (36) converges absolutely. The mean particle number in the state $\Psi(z)$ is

$$\begin{aligned} A(z) &= \left(\sum_N 2Nw_N n_{2N}\right)\Big/\left(\sum_N w_N n_{2N}\right) \\ &= 2zN^{-1}(dN/dz) \\ &= 2z(df/dz) \\ &= \sum_\alpha n_\alpha, \end{aligned} \tag{42}$$

with

$$n_\alpha = z\lambda_\alpha^2(1 - z\lambda_\alpha^2)^{-1}. \tag{43}$$

We now calculate the expectation value $H(z)$ given by Eqs. (22) and (23). The Hamiltonian H is a sum of kinetic and Coulomb terms, and $M(z)$ splits correspondingly into a kinetic part M_K and a Coulomb part M_C. The quantity $(H_{2N})_K n_{2N}$ is obtained from n_{2N} according to Eq. (27) by applying the operation

$$(\hbar^2/2m)\sum_k |\nabla_k|^2 \tag{44}$$

to the integrand, with one gradient operator ∇_k acting upon each of the two G kernels which contain the coordinate x_k. When the integral $(H_{2N})_K$ is divided into cyclic factors as in Eq. (29), the gradient operators appear in each factor K_j in turn. As a result we have

$$(H_{2N})_K n_{2N} = \sum_j (L_j(\partial/\partial K_j))n_{2N}, \tag{45}$$

with

$$\begin{aligned} L_j = 2j\int\cdots\int dy_1\cdots dy_{2j} \\ \times \sum_{v_1}\cdots\sum_{v_{2j}}\left\{\prod_{j=2}^{2j-1} G(y_k, v_k, y_{k+1}, v_{k+1})\right\} \\ \times (\hbar^2/2m)(\nabla_1 G(y_{2j}, v_{2j}, y_1, v_1) \\ \cdot \nabla_1 G(y_1, v_1, y_2, v_2)). \end{aligned} \tag{46}$$

By virtue of Eq. (16) and the orthogonality of the $\psi_\alpha(x)$, this becomes

$$L_j = 2j\cdot 2^{2j}\cdot\sum_\alpha \lambda_\alpha^{2j} T_\alpha \tag{47}$$

with T_α given by Eq. (14). When Eq. (45) is multiplied by w_N and summed over N, the result is

$$\begin{aligned} M_K(z) &= \sum_j (L_j(\partial/\partial K_j))N(z) \\ &= N(z)\sum_j (L_j(\partial/\partial K_j))f(z), \\ M_K(z) &= N(z)\sum_j (z/4)^j(L_j/2j) \\ &= N(z)\sum_j\sum_\alpha z^j\lambda_\alpha^{2j}T_\alpha \\ &= N(z)\sum_\alpha T_\alpha n_\alpha \end{aligned} \tag{48}$$

with n_α given by Eq. (42). This result is physically reasonable, since the kinetic energy is an additive property of single particles.

In a similar way, we now consider the Coulomb energy $(H_{2N})_C n_{2N}$, which is obtained from n_{2N} according to Eq. (27) by inserting the factor

$$\sum_{i<j} e^2 u_i u_j |x_i - x_j|^{-1} \tag{49}$$

into the integrand. When the integral is factorized into cycles, two cases arise. Either the points (x_i, x_j) belong to different cycles or to the same cycle. In the first case the summation over the charge coordinates within either cycle gives zero. This expresses the physical fact that the Coulomb energy of the average charge distribution vanishes. The surviving Coulomb terms then factorize just as do the kinetic energy terms, and in analogy with Eq. (48), we obtain

$$\begin{aligned} M_C(z) &= \sum_j (C_j(\partial/\partial K_j))N(z) \\ &= N(z)\sum_j (z/4)^j(C_j/2j), \end{aligned} \tag{50}$$

with

$$\begin{aligned} C_j = 2j\int\cdots\int dx_1\cdots dx_{2j} \\ \times \sum_{u_1}\cdots\sum_{u_{2j}}(\tfrac{1}{2}e^2)\sum_{k=2}^{2j}|x_1 - x_k|^{-1}\, u_1 u_k \\ \times \left\{\prod_{l=1}^{2j-1} G(x_l, u_l, x_{l+1}, u_{l+1})\right\} G(x_{2j}, u_{2j}, x_1, u_1). \end{aligned} \tag{51}$$

When the kernel G from Eq. (16) is substituted into Eq. (51), we obtain a sum of Coulomb exchange integrals $e_{\alpha\beta}$ given by Eq. (15). The sum over the charge coordinates gives a vanishing coefficient to $e_{\alpha\beta}$ if α, β are either both zero or both nonzero. We are then left with

$$C_j = 2j\cdot 2^{2j}\sum_{\alpha>0} e_{\alpha 0}\sum_{k=2}^{2j}(-\lambda_\alpha)^{k-1}\lambda_0^{2j+1-k}. \tag{52}$$

Substituting this into Eq. (50) and using Eq. (43), we find

$$\begin{aligned} M_C(z) &= -N(z)\sum_{\alpha>0} e_{\alpha 0}\left(\frac{n_0\lambda_\alpha + n_\alpha\lambda_0}{\lambda_\alpha + \lambda_0}\right) \\ &= N(z)\sum_{\alpha>0} e_{\alpha 0}(n_0 n_\alpha - [n_0 n_\alpha(n_0+1)(n_\alpha+1)]^{\frac{1}{2}}). \end{aligned} \tag{53}$$

Combining this with Eq. (48), we have for the expectation value of the energy in the state $\Psi(z)$ the exact expression

$$H(z) = \sum_\alpha T_\alpha n_\alpha + \sum_{\alpha>0} e_{\alpha 0}(n_0 n_\alpha - [n_0 n_\alpha (n_0+1)(n_\alpha+1)]^{\frac{1}{2}}). \quad (54)$$

The remarkably simple way in which the Coulomb energy appears in Eq. (54) is the main reason for our choice of trial wavefunction.

V. MINIMUM-ENERGY STATES

Let us hold the $\psi_\alpha(x)$, z, and n_0 fixed, and vary the λ_α for $\alpha > 0$ so as to make the energy $H(z)$ a minimum. There is a unique minimum, which occurs at

$$\lambda_\alpha = z^{-\frac{1}{2}} e^{-\varphi_\alpha}, \quad n_\alpha = [e^{2\varphi_\alpha} - 1]^{-1}, \quad (55)$$

with the exponent φ_α given by

$$\cosh \varphi_\alpha = [n_0(n_0+1)]^{-\frac{1}{2}}[n_0 + (T_\alpha/e_{\alpha 0})], \quad \varphi_\alpha > 0. \quad (56)$$

The minimum of $H(z)$ is then

$$H(n_0) = T_0 n_0 - \tfrac{1}{2}[n_0(n_0+1)]^{\frac{1}{2}} \sum_{\alpha>0} e_{\alpha 0} e^{-\varphi_\alpha}$$
$$= T_0 n_0 - \tfrac{1}{2} \sum_{\alpha>0} [T_\alpha + n_0 e_{\alpha 0} - (T_\alpha^2 + 2T_\alpha n_0 e_{\alpha 0} - n_0 e_{\alpha 0}^2)^{\frac{1}{2}}], \quad (57)$$

while the corresponding mean particle number is, from Eq. (42),

$$A(n_0) = n_0 + \tfrac{1}{2} \sum_{\alpha>0} [(T_\alpha + n_0 e_{\alpha 0}) \times (T_\alpha^2 + 2T_\alpha n_0 e_{\alpha 0} - n_0 e_{\alpha 0}^2)^{-\frac{1}{2}} - 1]. \quad (58)$$

From Eq. (57) and (58) it is easy to chart the relation between energy and particle numbers for any particular choice of the $\psi_\alpha(x)$.

In particular, we may consider the case of a Coulomb system in a finite box of side L with periodic boundary conditions, and use for $\psi_\alpha(x)$ the plane-wave states

$$\psi_j(x) = L^{-\frac{3}{2}} \exp[(2\pi i/L)(j \cdot x)], \quad (59)$$

where $j = (j_1, j_2, j_3)$ is a triplet of integers. Strictly speaking, the $\psi_j(x)$ for $j \neq 0$ should be taken to be sines and cosines rather than running waves, but Eq. (59) gives the correct values for the kinetic and exchange energies,

$$T_j = (2\pi^2 \hbar^2 j^2/mL^2), \quad (60)$$

$$e_{j0} = (e^2/L\pi j^2). \quad (61)$$

We go to the limit of an infinite uniform system of density ρ by letting $L \to \infty$ with

$$n_0 = \rho_0 L^3. \quad (62)$$

We find then from Eqs. (57) and (58),

$$[H(n_0)/n_0] = -(2/\pi)[8\pi\rho_0 a^3]^{\frac{1}{4}} J \text{ Ry}, \quad (63)$$

$$[\rho/\rho_0] = 1 + (2/\pi)[8\pi\rho_0 a^3]^{-\frac{1}{4}} J', \quad (64)$$

with

$$J = \int_0^\infty dx[x^4 + 1 - x^2(x^4+2)^{\frac{1}{2}}], \quad (65)$$

$$J' = \int_0^\infty dx[(x^4+1)(x^4+2)^{-\frac{1}{2}} - x^2], \quad (66)$$

$$J = (2^{\frac{3}{4}}/5)[\Gamma(\tfrac{1}{2})\Gamma(\tfrac{3}{4})/\Gamma(\tfrac{5}{4})] = 0.81\cdots, \quad (67)$$

$$J' = (2^{-\frac{7}{4}}/3)[\Gamma(\tfrac{1}{2})\Gamma(\tfrac{1}{4})/\Gamma(\tfrac{3}{4})] = 0.52\cdots. \quad (68)$$

These results exactly coincide with the leading terms of the perturbation-theory calculations,[3] which are valid in the limit when $\rho_0 a^3$ and ρa^3 are large.

In one respect we have gone beyond the perturbation-theory results. Since Eqs. (63) and (64) are exact for our particular choice of wavefunction, they set upper bounds to the energy per particle in the true ground state. Now, there is a discrepancy between the various calculations of the second term in the perturbation expansion of the ground-state energy. Girardeau[3] obtained a term in $\log[\rho_0 a^3]$ with a positive coefficient, while Lee and Feenberg[3] obtained only a constant term which has recently been confirmed by Brueckner.[7] Our calculation provides additional evidence that no positive logarithmic term can exist.

VI. RIGOROUS UPPER BOUNDS

To prove a rigorous upper bound of the form (2) for a finite system, it is inconvenient to use the optimum choice for the parameters λ_α given by Eqs. (55) and (56). Instead, we make the simple choice

$$\lambda_0 = 1, \quad \lambda_\alpha = \tfrac{1}{2} \quad \text{for} \quad \alpha = 1, \cdots, Q, \quad (69)$$

where Q is an integer to be fixed later, and take $\lambda_\alpha = 0$ for $\alpha > Q$. Then Eqs. (41), (42), (43), and (54) give

$$n_0 = z(1-z)^{-1}, \quad n_\alpha = z(4-z)^{-1}, \quad \alpha = 1, \cdots, Q, \quad (70)$$

$$N(z) = (1-z)^{-\frac{1}{2}}(1 - \tfrac{1}{4}z)^{-\frac{1}{2}Q}, \quad (71)$$

$$A(z) = z(1-z)^{-1} + Qz(4-z)^{-1}, \quad (72)$$

$$H(z) = \left[T_0 - \tfrac{1}{8}\sum_1^Q e_{\alpha 0}\right] z(1-z)^{-1} + \left[\sum_1^Q (T_\alpha - \tfrac{2}{3} e_{\alpha 0})\right] z(4-z)^{-1}. \quad (73)$$

By Eqs. (22), (23), (24), H_{2N} is the coefficient of z^N in $(N(z)H(z))$, divided by the coefficient of z^N in $N(z)$.

[7] K. A. Brueckner, Phys. Rev. 156, 204 (1967).

Now

$$N(z)A(z) = 2z(d/dz)N(z) \tag{74}$$

has coefficients equal to those of $N(z)$ multiplied by $2N$. Therefore

$$H_{2N} = 2N\Big[T_0 - \tfrac{1}{3}\sum_1^Q e_{\alpha 0}\Big] + \beta_N\Big[\sum_1^Q (T_\alpha + \tfrac{1}{3}(Q-2)e_{\alpha 0}) - QT_0\Big], \tag{75}$$

where β_N is the ratio of the coefficients of z^N in $(z(4-z)^{-1}N(z))$ and $N(z)$. For large N and Q we have approximately $\beta_N = \frac{1}{3}$; it is easy to prove rigorously for all N and Q that

$$\beta_N < 1. \tag{76}$$

Hence Eq. (75) implies that

$$H_{2N} < \sum_1^Q T_\alpha + (2N-Q)\Big[T_0 - \tfrac{1}{3}\sum_1^Q e_{\alpha 0}\Big]. \tag{77}$$

For the theorem, we are interested not in H_{2N}, which is the average energy of a mixture of states of different charge, but in $H_{N,N}$ which is the energy of a pure state of charge zero. It is highly plausible that

$$H_{n,2N-n} \geq H_{N,N}, \quad n = 0, 1, \cdots, 2N, \tag{78}$$

so that the neutral state has the lowest energy of all states of $2N$ particles. If Eq. (78) were true, then by Eqs. (25) and (26)

$$H_{2N} \geq H_{N,N} \geq E_N. \tag{79}$$

However, we have not proved Eq. (78) and therefore we do not make use of it. Instead we use the fact that the ground-state energy ϵ_{N_+,N_-} of a system of N_+ positive and N_- negative charges is a decreasing function of both N_+ and N_-. The decreasing property follows from the consideration that any added particle can be placed infinitely far away from those already present without increasing the total energy. Also,

$$\epsilon_{2N,0} = \epsilon_{0,2N} = 0. \tag{80}$$

Therefore Eq. (25) gives

$$n_{2N}H_{2N} \geq \sum_{n=1}^{2N-1} |\Phi_{n,2N-n}|^2 \epsilon_{n,2N-n} \geq n_{2N}\epsilon_{2N-1,2N-1}, \tag{81}$$

and so

$$H_{2N} \geq E_{2N-1} \geq E_{2N}. \tag{82}$$

From Eqs. (77) and (82) we see that, whether N is even or odd,

$$E_N < \sum_1^Q T_\alpha + (N-Q)\Big[T_0 - \tfrac{1}{3}\sum_1^Q e_{\alpha 0}\Big]. \tag{83}$$

It now remains to choose the functions $\psi_\alpha(x)$ so as to make the right side of Eq. (83) large and negative. We need the T_α to be as small and the $e_{\alpha 0}$ to be as large as possible. Fortunately, these two requirements work in the same direction, since T_α and $e_{\alpha 0}$ are approximately reciprocals of each other. There is a strict inequality

$$e_{\alpha\beta}t_{\alpha\beta} > 4\pi e^2 (r_{\alpha\beta})^2, \tag{84}$$

where

$$r_{\alpha\beta} = \int (\psi_\alpha(x)\psi_\beta(x))^2 \, d_3x, \tag{85}$$

$$t_{\alpha\beta} = \int |\nabla(\psi_\alpha(x)\psi_\beta(x))|^2 \, d_3x. \tag{86}$$

Equation (84) follows immediately from Schwarz's inequality when the integrals $e_{\alpha\beta}$, $r_{\alpha\beta}$, $t_{\alpha\beta}$ are written in terms of the Fourier transform of $(\psi_\alpha\psi_\beta)$. Now, $r_{\alpha\beta}$ is just a measure of the mean density of particles, while $t_{\alpha\beta}$ is closely related to $(T_\alpha + T_\beta)$, so that Eq. (84) gives an inverse relationship between $e_{\alpha 0}$ and T_α. We must therefore choose the $\psi_\alpha(x)$ to make the T_α as small as possible for a given over-all particle density.

Guided by the heuristic argument of Sec. II, we choose a length Λ which fixes the mean particle density, and write

$$\psi_j(x) = \left(\frac{2}{\Lambda}\right)^{\frac{3}{2}} \prod_{\mu=1}^{3} \sin\left[\left(\frac{\pi}{\Lambda}\right) j_\mu x_\mu\right] \tag{87}$$

for points x within the cube

$$0 < x_\mu < L, \quad \mu = 1, 2, 3, \tag{88}$$

with $\psi_j(x) = 0$ outside the cube. The index j represents a triplet of strictly positive integers. We then have

$$T_j = \tfrac{1}{2}\pi^2 j^2 (\hbar^2/m\Lambda^2). \tag{89}$$

Instead of j we may use the index α, which labels the triplets j in increasing order of j^2. Equations (84)–(87) give

$$r_{\alpha 0} = \Lambda^{-3}, \tag{90}$$

$$t_{\alpha 0} = \pi^2 \Lambda^{-5}(j^2 + 3), \tag{91}$$

$$e_{\alpha 0}T_\alpha \geq (4\pi/3)(e^2\hbar^2/m\Lambda^3). \tag{92}$$

Without any attempt to find the best numerical coefficients, we write

$$j^2 \leq 6\alpha^{\frac{2}{3}} \quad \text{for} \quad \alpha \geq 1, \tag{93}$$

and deduce from Eqs. (89) and (92)

$$\sum_1^Q T_\alpha \leq 3\pi^2 Q^{\frac{5}{3}}\left(\frac{\hbar^2}{m\Lambda^2}\right), \tag{94}$$

$$\sum_1^Q e_{\alpha 0} \geq \left(\frac{4}{9\pi}\right) Q^{\frac{1}{3}}(e^2/\Lambda). \tag{95}$$

Assembling Eqs. (83), (89), (94), and (95), we find

$$E_N < 3\pi^2(\tfrac{1}{2}N + Q^{\frac{5}{3}})(\hbar^2/m\Lambda^2) - (4/27\pi)(N - Q)Q^{\frac{1}{3}}(e^2/\Lambda). \quad (96)$$

It is of interest to compare this estimate (96) with the considerations of Sec. II. The physical meaning of the estimate is somewhat obscured by the use of Eq. (84) to avoid a direct but tedious calculation of the $e_{\alpha 0}$. The physical basis of Eq. (95) is the fact that

$$\sum_1^Q e_{\alpha 0} = e^2 \iint g(x, y)\,|x - y|^{-1}\,d_3x\,d_3y, \quad (97)$$

where

$$g(x, y) = \psi_0(x)\psi_0(y)\sum_1^Q \psi_\alpha(x)\psi_\alpha(y) \quad (98)$$

is a measure of the short-range two-particle correlations in the wavefunction (17). If we introduce the correlation length

$$\lambda = Q^{-\frac{1}{3}}\Lambda, \quad (99)$$

then the function $g(x, y)$ has a sharp peak of magnitude $(\lambda\Lambda)^{-3}$ extending over the volume $|x - y| \sim \lambda$. The right side of Eq. (97) will be of the order (e^2/λ), which agrees with Eq. (95). Thus the three terms on the right of Eq. (96) correspond exactly to the three terms of Eq. (6), whose magnitudes were estimated in Eqs. (7), (11), and (8). The precise inequality (96) fully confirms the correctness of the earlier heuristic argument.

To complete the proof of Eq. (2), we take in Eq. (96) $Q = (\frac{1}{2}N)^{\frac{3}{5}}$. This gives

$$E_N < 3\pi^2 N(\hbar^2/m\Lambda^2) - (18\pi)^{-1}N^{\frac{8}{5}}(e^2/\Lambda). \quad (100)$$

Finally we take

$$\Lambda = 108\pi^3 N^{-\frac{1}{5}}(\hbar^2/me^2), \quad (101)$$

and Eq. (100) becomes

$$E_N < -[1944\pi^4]^{-1}N^{\frac{7}{5}}\,\mathrm{Ry}, \quad (102)$$

which proves the theorem.

VII. CONCLUDING REMARKS

It would be easy to obtain a more reasonable numerical coefficient in Eq. (102). We know that for all N

$$E_N < -\tfrac{1}{2}N\,\mathrm{Ry}, \quad (103)$$

because N separated positronium atoms form a possible state of N positive and N negative charges. On the other hand, for large N we can use much better numerical estimates than Eqs. (76), (93), (94), (95). The combination of Eq. (103) for small N with improved estimates for large N will give a coefficient in Eq. (102) which is not enormously less than unity.

More interesting than making piecemeal improvements is the determination of the best possible coefficient A_1 in Eq. (2). A plausible conjecture is that $(-A_1)$ should be the minimum of the quantity

$$Q(\varphi) = \int |\nabla\varphi(x)|^2\,d_3x - (2/\pi)(8\pi)^{\frac{1}{4}}J\int |\varphi(x)|^{\frac{5}{2}}\,d_3x, \quad (104)$$

minimized over all functions $\varphi(x)$ with

$$\int |\varphi(x)|^2\,d_3x = 1, \quad (105)$$

the coefficient J being given by Eq. (65). This conjecture results from a Thomas–Fermi type of approximation, in which it is assumed that the total energy is a sum of kinetic and correlation energies, the correlation energy being a function only of the local density according to the perturbation-theory formula (63). The minimization of $Q(\varphi)$ will give a nonlinear equation for $\varphi(x)$, similar to the Thomas–Fermi equation, which could be solved numerically once and for all. The minimizing $|\varphi(x)|^2$ will give the shape of the density distribution in the ground state of $2N$ particles. We conjecture that this recipe will give the best possible coefficient A_1 in our theorem, at least in the limit $N \to \infty$. We have no idea at present how the conjecture might be proved.

The main part of the work of this paper (particularly Secs. III and IV) does not require a Coulomb interaction, but would apply to a system of bosons with any two-particle interaction bilinear in the charges, provided that there is no hard core. For example, if the Coulomb potential in Eq. (1) were replaced by a potential

$$V(r) = cr^{-n}, \quad 0 < n < 2, \quad (106)$$

then the ground-state energy E_N would satisfy

$$E_N < -AN^{(10-3n)/(10-5n)}. \quad (107)$$

ACKNOWLEDGMENT

The author expresses his deep indebtedness to Dr. Andrew Lenard for suggesting the problem and for many fruitful discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Lower Bound on the Mean Kinetic Energy of a System of Particles*

R. M. MAZO
Institute of Theoretical Science, and Department of Chemistry, University of Oregon, Eugene, Oregon

(Received 9 January 1967)

It is shown that $K \geq \frac{3}{5}\mu_0$ for a system of fermions. K is $\langle P^2/2m\rangle$, and μ_0 is the Fermi energy of an ideal gas of fermions of the same mass, and at the same density, as the system under consideration.

THIS paper derives a rigorous lower bound for the mean kinetic energy of a one component system of fermions in thermal equilibrium. The analogous result for bosons is seen to be trivial.

The system under consideration has a Hamiltonian given by

$$H_0 = \sum_{i=1}^{N} \frac{P_i^2}{2m_0} + V. \tag{1}$$

The potential V is completely arbitrary, except for the requirement that the partition function,

$$Z = \mathrm{Tr}\,[\exp(-\beta H)]$$

exists. For the purposes of the proof we consider a comparison system with Hamiltonian

$$H = \sum_{i=1}^{N} \frac{P_i^2}{2m} + V. \tag{2}$$

The potentials V of Eqs. (1) and (2) are the same, and we require that $0 < m < m_0$. We further define

$$H_1 = H - H_0 = \sum_{i=1}^{N} \frac{P_i^2}{2m_1}, \tag{3}$$

$$m_1^{-1} = m^{-1} - m_0^{-1}. \tag{3a}$$

That is, H_1 is the Hamiltonian of an ideal gas of particles with mass m_1.

Let $F(H, \beta)$ denote the Helmholtz free energy of a system with Hamiltonian H and temperature T; $\beta = (kT)^{-1}$. Then the inequality of Gibbs[1] and Bogoliubov[2] states that

$$F(H_0, \beta) + \langle H - H_0\rangle_0 \geq F(H, \beta), \tag{4}$$

where the brackets denote a thermal average with respect to the statistical operator of H_0. In addition, there is an inequality recently derived by Falk[3] which states that

$$F(H, \beta) \geq F(H_0, 2\beta) + F(H_1, 2\beta). \tag{5}$$

Since we know from thermodynamics that

$$(\partial F/\partial T)_{N,\Omega} = -S,$$

the entropy, F must be an increasing function of β. Hence

$$F(H_0, 2\beta) \geq F(H_0, \beta). \tag{6}$$

Combining (4), (5), and (6), we can eliminate all mention of $F(H, \beta)$ and conclude that

$$\langle H - H_0\rangle_0 \geq F(H_1, 2\beta). \tag{7}$$

$F(H_1, 2\beta)$ is the Helmholtz function of an ideal gas of mass m_1 at a temperature $\frac{1}{2}T$. Now

$$\langle H - H_0\rangle_0 = (m_0/m_1)NK, \tag{8}$$

where $K = \langle P^2/2m_0\rangle_0$. Hence

$$K \geq (m_1/m_0N)F(H_1, 2\beta). \tag{9}$$

Equation (9) is a general result, independent of the statistics obeyed by the particles. For the bose case, however, the free energy of the ideal gas is always negative; since $H - H_0$ is a positive semidefinite operator, Eq. (9) does not teach us anything. We henceforth confine ourselves to the Fermi case.

The auxiliary mass m can be chosen to be arbitrarily small; hence m_1 can be made arbitrarily small. Therefore, at any finite temperature and any positive density, the ideal gas described by H_1 can be made extremely degenerate, so that $F(H_1, 2\beta)$ can be replaced by $(\frac{3}{5})N\mu_1$. Here μ_1 is the Fermi energy of an ideal gas of mass m_1. Furthermore, since the Fermi energy is given by

$$\mu = (h^2/8m)(3N/\pi\Omega)^{\frac{2}{3}}, \tag{10}$$

where Ω is the volume of the system, we have

$$(m_1/m_0)\mu_1 = \mu_0, \tag{11}$$

where μ_0 is the Fermi energy of an ideal gas of mass

* Supported by a National Science Foundation Grant GP-6002.

[1] J. W. Gibbs, *The Collected Works* (Yale University Press, New Haven, Connecticut, 1948), Vol. II, p. 131, Theorem III. Naturally, Gibb's proof is only for classical systems.

[2] N. N. Bogoliubov (unpublished); cited in footnote 4 of V. V. Tolmachev, Dokl. Akad. Nauk SSSR **134**, 1324 (1960) [English transl.: Soviet Phys.—Dokl. **5**, 984 (1961)].

[3] H. Falk, J. Math. Phys. **7**, 977 (1966). (Penultimate equation, first column, p. 978, with $\gamma = 0$.)

m_0. Our final result is, then

$$K \geq \tfrac{3}{5}\mu_0, \tag{12}$$

where μ_0 is the Fermi energy of an ideal Fermi gas of particles of the same mass and at the same density as the system under consideration.

The result (12) is somewhat reminiscent of the Hugenholtz–van Hove theorem.[4] It is less powerful since it is an inequality. On the other hand, it is not restricted to zero temperature, and, being independent of the nature of the forces, will also be valid when the forces can produce bound states not accessible by perturbation theory, as in superconductors.

We give one example. For liquid He^3, the zero-point kinetic energy has been estimated to be of the order of 60 cal/mole,[5] although this is probably somewhat high. The right-hand side of our inequality is 5.8 cal/mole. In this case, the lower bound does not yield much of practical interest, but this is perhaps too much to expect for an estimate which utilizes absolutely no information about the dynamics of the system.

[4] N. M. Hugenholtz and L. van Hove, Physica **24**, 363 (1958).

[5] F. London, *Superfluids*, (John Wiley & Sons, Inc., New York, 1954), Vol. II, p. 165.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Empty Space-Times Algebraically Special on a Given World Line or Hypersurface

C. D. COLLINSON
Department of Applied Mathematics, University of Hull, Hull, England

(Received 9 December 1966)

The null tetrad formalism of Newman and Penrose is used to investigate empty space-times which are algebraically special on a given world line. It is found that, when the world line is timelike, the space-time admits a congruence which is geodesic and shear-free on the world line. A similar result is obtained for those empty space-times which are algebraically special on a given space like hypersurface.

1. INTRODUCTION

THE four null directions corresponding to the vectors p_i satisfying the equation

$$p_{[i}C_{j]kl[m}p_{n]}p^k p^l = 0$$

are called principal null directions of the Weyl tensor. Space–times for which two or more principal null directions coincide are algebraically special and the vector p_i corresponding to the repeated principal null direction satisfies the equation

$$C_{jkl[m}p_{n]}p^k p^l = 0.$$

An important geometrical characteristic of algebraically special empty space–times is that the repeated principal null direction is geodesic and shear-free.[1] The purpose of this paper is to investigate those empty space–times which are algebraically special on a given submanifold. The following two theorems are proved.

Theorem 1: Let V be a locally empty space–time (with C^5 metric) which is algebraically special but not flat on a given world line W. Let l be a principal null direction of the curvature tensor which, on W, points in the repeated principal null direction. Then:

(a) if W is null, l is shear-free on W;

(b) if W is timelike, l is geodesic and shear-free on W.

Theorem 2: Let V be a locally empty space–time (with C^5 metric) which is algebraically special but not flat on a given spacelike hypersurface S. Let l be a principal null direction of the curvature tensor which, on S, points in the repeated principal null direction. Then l is geodesic and shear-free on S.

2. NOTATION

Newman and Penrose[2] introduce at each point of space–time a tetrad of null vectors l^i, n^i, m^i, and $\bar{m}^i$

[1] A. Lichnerowicz, Compt. Rend. **246**, 893 (1958).

[2] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

m_0. Our final result is, then

$$K \geq \tfrac{3}{5}\mu_0, \tag{12}$$

where μ_0 is the Fermi energy of an ideal Fermi gas of particles of the same mass and at the same density as the system under consideration.

The result (12) is somewhat reminiscent of the Hugenholtz–van Hove theorem.[4] It is less powerful since it is an inequality. On the other hand, it is not restricted to zero temperature, and, being independent of the nature of the forces, will also be valid when the forces can produce bound states not accessible by perturbation theory, as in superconductors.

We give one example. For liquid He^3, the zero-point kinetic energy has been estimated to be of the order of 60 cal/mole,[5] although this is probably somewhat high. The right-hand side of our inequality is 5.8 cal/mole. In this case, the lower bound does not yield much of practical interest, but this is perhaps too much to expect for an estimate which utilizes absolutely no information about the dynamics of the system.

[4] N. M. Hugenholtz and L. van Hove, Physica **24**, 363 (1958).

[5] F. London, *Superfluids*, (John Wiley & Sons, Inc., New York, 1954), Vol. II, p. 165.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Empty Space-Times Algebraically Special on a Given World Line or Hypersurface

C. D. COLLINSON
Department of Applied Mathematics, University of Hull, Hull, England

(Received 9 December 1966)

The null tetrad formalism of Newman and Penrose is used to investigate empty space-times which are algebraically special on a given world line. It is found that, when the world line is timelike, the space-time admits a congruence which is geodesic and shear-free on the world line. A similar result is obtained for those empty space-times which are algebraically special on a given space like hypersurface.

1. INTRODUCTION

THE four null directions corresponding to the vectors p_i satisfying the equation

$$p_{[i}C_{j]kl[m}p_{n]}p^k p^l = 0$$

are called principal null directions of the Weyl tensor. Space–times for which two or more principal null directions coincide are algebraically special and the vector p_i corresponding to the repeated principal null direction satisfies the equation

$$C_{jkl[m}p_{n]}p^k p^l = 0.$$

An important geometrical characteristic of algebraically special empty space–times is that the repeated principal null direction is geodesic and shear-free.[1] The purpose of this paper is to investigate those empty space–times which are algebraically special on a given submanifold. The following two theorems are proved.

Theorem 1: Let V be a locally empty space–time (with C^5 metric) which is algebraically special but not flat on a given world line W. Let l be a principal null direction of the curvature tensor which, on W, points in the repeated principal null direction. Then:

(a) if W is null, l is shear-free on W;

(b) if W is timelike, l is geodesic and shear-free on W.

Theorem 2: Let V be a locally empty space–time (with C^5 metric) which is algebraically special but not flat on a given spacelike hypersurface S. Let l be a principal null direction of the curvature tensor which, on S, points in the repeated principal null direction. Then l is geodesic and shear-free on S.

2. NOTATION

Newman and Penrose[2] introduce at each point of space–time a tetrad of null vectors l^i, n^i, m^i, and $\bar{m}^i$

[1] A. Lichnerowicz, Compt. Rend. **246**, 893 (1958).

[2] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

satisfying the orthonormality conditions

$$l^i n_i = -m^i \bar{m}_i = 1, \quad l^i m_i = n^i m_i = m^i m_i = 0. \quad (2.1)$$

The propagation of the tetrad is specified by certain linear combinations of the (complex) Ricci rotation coefficients,[3] called the spin coefficients. The spin coefficients τ, κ, ρ, and σ, introduced by the equation

$$l_{i;j} m^i = \tau l_j + \kappa n_j - \rho m_j - \sigma \bar{m}_j,$$

are closely related to the geometry of the null congruence defined by the tetrad vector l^i. In particular the congruence is geodesic if $\kappa = 0$, and shear-free if $\sigma = 0$.[4]

The Weyl tensor of a space–time is specified by five complex tetrad components ψ_0, ψ_1, ψ_2, ψ_3, and ψ_4. The empty space–time Bianchi identities can then be written

$$D\psi_1 - \bar{\delta}\psi_0 = -3\kappa\psi_2 + [2\epsilon + 4\rho]\psi_1 - [-\pi + 4\alpha]\psi_0, \quad (2.2)$$

$$D\psi_2 - \bar{\delta}\psi_1 = -2\kappa\psi_3 + 3\rho\psi_2 - [-2\pi + 2\alpha]\psi_1 - \lambda\psi_0, \quad (2.3)$$

$$D\psi_3 - \bar{\delta}\psi_2 = -\kappa\psi_4 - [2\epsilon - 2\rho]\psi_3 + 3\pi\psi_2 - 2\lambda\psi_1, \quad (2.4)$$

$$D\psi_4 - \bar{\delta}\psi_3 = -[4\epsilon - \rho]\psi_4 + [4\pi + 2\alpha]\psi_3 - 3\lambda\psi_2, \quad (2.5)$$

$$\Delta\psi_0 - \delta\psi_1 = [4\gamma - \mu]\psi_0 - [4\tau + 2\beta]\psi_1 + 3\sigma\psi_2, \quad (2.6)$$

$$\Delta\psi_1 - \delta\psi_2 = \nu\psi_0 + [2\gamma - 2\mu]\psi_1 - 3\tau\psi_2 + 2\sigma\psi_3, \quad (2.7)$$

$$\Delta\psi_2 - \delta\psi_3 = 2\nu\psi_1 - 3\mu\psi_2 + [-2\tau + 2\beta]\psi_3 + \sigma\psi_4, \quad (2.8)$$

$$\Delta\psi_3 - \delta\psi_4 = 3\nu\psi_2 - [2\gamma + 4\mu]\psi_3 + [-\tau + 4\beta]\psi_4. \quad (2.9)$$

Here D, Δ, and δ are the intrinsic derivatives defined by

$$D\phi = \phi_{,i} l^i, \quad \Delta\phi = \phi_{,i} n^i, \quad \text{and} \quad \delta\phi = \phi_{,i} m^i.$$

The intrinsic derivatives do not commute. In particular,

$$\delta D - D\delta = (\bar{\alpha} + \beta - \bar{\pi})D + \kappa\Delta - \sigma\bar{\delta} - (\bar{\rho} + \epsilon - \bar{\epsilon})\delta. \quad (2.10)$$

Only one of the field equations developed by Newman and Penrose is used here, namely

$$D\sigma - \delta\kappa = (\rho + \bar{\rho})\sigma + (3\epsilon - \bar{\epsilon})\sigma - (\tau - \bar{\pi} + \bar{\alpha} + 3\beta)\kappa + \psi_0. \quad (2.11)$$

When space–time admits a preferred direction corresponding to a (normalized) vector v^i the tetrad vectors l^i and n^i can be chosen so that

$$v^i = (n^i + cl^i) \,|\, \sqrt{2}, \quad (2.12)$$

where c is $+1$, -1, or 0 according as v^i is timelike, spacelike, or null. A short calculation shows that those tetrad transformations which leave invariant the orthonormality conditions (2.1) and the vector (2.12) are

$$m^{i'} = e^{i\theta} m^i, \quad (2.13)$$

and

$$l^{i'} = A(l^i + a\bar{a}n^i + \bar{a}m^i + a\bar{m}^i),$$
$$n^{i'} = A(n^i + c^2 a\bar{a}l^i - c\bar{a}m^i - ca\bar{m}^i),$$
$$m^{i'} = A(m^i - ca^2\bar{m}^i - cal^i + an^i)$$
$$\text{with} \quad A = (1 + ca\bar{a})^{-1}. \quad (2.14)$$

Under (2.14) the tetrad components of the Weyl tensor transform as follows:

$$\sum_{I=0}^{N} {}^N C_I (c\bar{a})^{N-I} \psi'_I = A^{2-N} \sum_{I=N}^{4} {}^4 C_I a^I \psi_I, \quad \text{for} \quad N = 0, \cdots, 4. \quad (2.15)$$

Using this transformation the tetrad vector l^i can be chosen to correspond to a principal null direction of the Weyl tensor. With this choice $\psi_0 = 0$. If the space–time is algebraically special, the tetrad vector l^i can be chosen to correspond to the repeated principal null direction and then $\psi_0 = \psi_1 = 0$.

3. PROOF OF THEOREM 1

Let the vector (2.12) be tangent to the given world line W. Let P be an arbitrary point on W. A given function ϕ is zero at all points of W only if the function and its derivatives in the direction $n^i + cl^i$ are zero at point P. In terms of intrinsic derivatives

$$\phi \overset{P}{=} (\Delta + cD)\phi \overset{P}{=} (\Delta + cD)^2\phi \overset{P}{=} 0,$$

where $\overset{P}{=}$ denotes equality at P. Since P is chosen arbitrarily on W the equality

$$\phi \overset{P}{=} 0$$

implies

$$(\Delta + cD)\phi \overset{P}{=} (\Delta + cD)^2\phi \overset{P}{=} 0.$$

This last remark saves a lot of calculation, although the results can also be obtained analytically.

Choosing the tetrad vector l^i to correspond to a principal null direction of the Weyl tensor which, on

[3] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1925).
[4] R. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).

W, points in the repeated principal null direction yields

$$\psi_0 = 0, \tag{3.1}$$

$$\psi_1 \overset{P}{=} (\Delta + cD)\psi_1 \overset{P}{=} (\Delta + cD)^2\psi_1 \overset{P}{=} 0. \tag{3.2}$$

Under (2.13) ψ_1 transforms as

$$\psi_1' = e^{i\theta}\psi_1.$$

This transformation can therefore be used to make

$$\psi_1 = \bar{\psi}_1. \tag{3.3}$$

Obtaining $\delta D\psi_1$ and $D\delta\psi_1$ from (2.2) and (2.6), substituting into the commutator (2.10), using Eqs. (2.3), (2.7), and (2.11) to eliminate $D\psi_2$, $\delta\psi_2$, and $D\sigma - \delta\kappa$, respectively, and putting $\psi_0 = 0$ gives

$$4\sigma\delta\psi_1 - 4\kappa\Delta\psi_1 + 2\psi_1[-5\psi_1 - 2\Delta\kappa + 2\delta\sigma + \kappa(10\gamma + 2\bar{\gamma} - 2\bar{\mu}) + \sigma(-10\alpha - 2\bar{\tau} + 2\bar{\beta})] = 0. \tag{3.4}$$

From Eq. (3.3)

$$\delta\psi_1 = (\overline{\delta\psi_1}),$$

and this can be eliminated using (2.6). Also

$$\Delta\psi_1 = (\Delta + cD)\psi_1 - cD\psi_1 = (\Delta + cD)\psi_1 - c(\overline{D\psi_1}),$$

and $(\overline{D\psi_1})$ can be eliminated using (2.2). Equation (3.4) then becomes

$$-12(\sigma\bar{\sigma} + c\kappa\bar{\kappa})\bar{\psi}_2 - 4\kappa(\Delta + cD)\psi_1 + \psi_1 F = 0, \tag{3.5}$$

where F is a function, not necessarily zero.

Equation (3.5) can be used to prove

$$\sigma\bar{\sigma} + c\kappa\bar{\kappa} \overset{P}{=} 0.$$

The method is by *reductio ad adsurdum*. Suppose hereafter that

$$\sigma\bar{\sigma} + c\kappa\bar{\kappa} \overset{P}{\neq} 0.$$

From (3.5)

$$\psi_2 \overset{P}{=} 0 \tag{3.6}$$

and so

$$(\Delta + cD)\psi_2 \overset{P}{=} (\Delta + cD)^2\psi_2 \overset{P}{=} 0. \tag{3.7}$$

Substituting (3.6) into (2.2) and (2.6) yields

$$D\psi_1 \overset{P}{=} \delta\psi_1 \overset{P}{=} 0$$

and also, from (3.2) and (3.3),

$$\Delta\psi_1 \overset{P}{=} \delta\psi_1 \overset{P}{=} 0.$$

Hence

$$(\Delta + cD)D\psi_1 \overset{P}{=} (\Delta + cD)\delta\psi_1 \overset{P}{=} (\Delta + cD)\Delta\psi_1 \overset{P}{=} (\Delta + cD)\bar{\delta}\psi_1 \overset{P}{=} 0.$$

Since all the first derivatives of ψ_1 are zero at P, the second derivatives commute at P. Differentiating (3.5) now gives

$$-12(\sigma\bar{\sigma} + c\kappa\bar{\kappa})D\bar{\psi}_2 \overset{P}{=} 4\kappa D(\Delta + cD)\psi_1 \overset{P}{=} 4\kappa(\Delta + cD)D\psi_1 \overset{P}{=} 0.$$

Hence

$$D\psi_2 \overset{P}{=} 0. \tag{3.8}$$

Similarly,

$$\Delta\psi_2 \overset{P}{=} \delta\psi_2 \overset{P}{=} \bar{\delta}\psi_2 \overset{P}{=} 0, \tag{3.9}$$

and so

$$(\Delta + cD)D\psi_2 \overset{P}{=} (\Delta + cD)\Delta\psi_2 \overset{P}{=} (\Delta + cD)\delta\psi_2 \overset{P}{=} (\Delta + cD)\bar{\delta}\psi_2 \overset{P}{=} 0. \tag{3.10}$$

Substituting (3.8) and (3.9) into (2.3) and (2.7) yields

$$\kappa\psi_3 \overset{P}{=} \sigma\psi_3 \overset{P}{=} 0.$$

Hence, since both κ and σ cannot vanish at P,

$$\psi_3 \overset{P}{=} 0.$$

Differentiating (2.2) and using (3.8) and (3.9) gives

$$DD\psi_1 \overset{P}{=} \Delta D\psi_1 \overset{P}{=} \delta D\psi_1 \overset{P}{=} 0. \tag{3.11}$$

Differentiating (3.5) twice now gives

$$\begin{aligned} &-12(\sigma\bar{\sigma} + c\kappa\bar{\kappa})DD\bar{\psi}_2 \\ &\overset{P}{=} 4\kappa DD(\Delta + cD)\psi_1 \\ &\overset{P}{=} 4\kappa D(\Delta + cD)D\psi_1 + D\,(\text{first derivatives of } \psi_1) \\ &\overset{P}{=} 4\kappa D(\Delta + cD)D\psi_1 + \text{first derivatives of } D\psi_1 \\ &\overset{P}{=} 4\kappa D(\Delta + cD)D\psi_1. \end{aligned} \tag{3.12}$$

Substituting (2.2) into the right-hand side of (3.12) gives

$$-12(\sigma\bar{\sigma} + c\kappa\bar{\kappa})DD\bar{\psi}_2 \overset{P}{=} -12\kappa^2 D(\Delta + cD)\psi_2 \overset{P}{=} -12\kappa^2(\Delta + cD)D\psi_2 \overset{P}{=} 0.$$

Hence

$$DD\bar{\psi}_2 \overset{P}{=} 0.$$

From (2.3) and (2.4)

$$DD\bar{\psi}_2 \overset{P}{=} -2\bar{\kappa}D\bar{\psi}_3 \overset{P}{=} 2\bar{\kappa}^2\psi_4 \overset{P}{=} 0.$$

Similarly it can be shown that

$$\bar{\delta}\bar{\delta}\bar{\psi}_2 \overset{P}{=} -2\bar{\sigma}\bar{\delta}\bar{\psi}_3 \overset{P}{=} 2\bar{\sigma}^2\psi_4 \overset{P}{=} 0.$$

Since both σ and κ cannot vanish at P,

$$\psi_4 \overset{P}{=} 0.$$

This contradicts the hypothesis that the space–time is not flat at P. Hence

$$\sigma\bar{\sigma} + c\kappa\bar{\kappa} \overset{P}{=} 0.$$

If W is null, $c = 0$ and therefore $\sigma \overset{P}{=} 0$. If W is time-like, $c = 1$ and therefore $\sigma \overset{P}{=} \kappa \overset{P}{=} 0$. Since the point

P was chosen arbitrarily on W, the theorem is proved.

4. PROOF OF THEOREM 2

Let the vector (2.12) be normal to the given hypersurface S. Then S is generated by the three real tangent vectors

$$(n^i - cl^i)/\sqrt{2}, \quad (m^i + \bar{m}^i)/\sqrt{2}, \quad \text{and} \quad i(m^i - \bar{m}^i)/\sqrt{2}.$$

The transformation (2.13) can again be used to make

$$\psi_1 = \bar{\psi}_1. \tag{4.1}$$

Let P be an arbitrary point of S. Choosing the tetrad vector l^i to correspond to a principal null direction of the Weyl tensor which, on S, points in the repeated principal null direction, now yields

$$\psi_0 = 0, \tag{4.2}$$

$$\psi_1 \overset{P}{=} (\Delta - cD)\psi_1 \overset{P}{=} \delta\psi_1 \overset{P}{=} 0, \tag{4.3}$$

and

$$(\Delta - cD)^2\psi_1 \overset{P}{=} (\Delta - cD)\delta\psi_1 \overset{P}{=} \delta^2\psi_1 \overset{P}{=} \delta(\Delta - cD)\psi_1 \overset{P}{=} \delta\delta\psi_1 = 0. \tag{4.4}$$

Following the proof of Theorem 1,

$$\sigma\bar{\sigma} - c\kappa\bar{\kappa} \overset{P}{=} 0. \tag{4.5}$$

If $c = -1$, that is, if the hypersurface S is timelike, σ and κ both vanish on S. The more interesting case is when $c = +1$. Equation (2.6), with the conditions (4.3), yields

$$\sigma\psi_2 \overset{P}{=} 0.$$

Suppose hereafter that

$$\sigma \overset{P}{\neq} 0.$$

Then

$$\psi_2 \overset{P}{=} 0.$$

As in the last proof, because P is chosen arbitrarily on S,

$$(\Delta \overset{\prime}{-} cD)\psi_2 \overset{P}{=} 0 \tag{4.6}$$

and

$$\delta\psi_2 \overset{P}{=} 0. \tag{4.7}$$

Substituting (4.7) into (2.7) gives

$$\Delta\psi_1 \overset{P}{=} 2\sigma\psi_3.$$

But from (4.3) and (2.2),

$$\Delta\psi_1 \overset{P}{=} cD\psi_1 \overset{P}{=} 0,$$

and so

$$\sigma\psi_3 \overset{P}{=} 0.$$

Hence

$$\psi_3 \overset{P}{=} 0,$$

and again

$$\delta\psi_3 \overset{P}{=} 0. \tag{4.8}$$

Substituting (4.8) into (2.8) gives

$$\Delta\psi_2 \overset{P}{=} \sigma\psi_4.$$

From (4.6) and (2.3)

$$\Delta\psi_2 \overset{P}{=} cD\psi_2 \overset{P}{=} 0$$

and so

$$\sigma\psi_4 \overset{P}{=} 0$$

or

$$\psi_4 \overset{P}{=} 0.$$

This contradicts the hypothesis that the space-time is not flat at P. Hence

$$\sigma \overset{P}{=} 0.$$

Substituting this into (4.5), with $c = -1$, yields

$$\kappa \overset{P}{=} 0.$$

Since P was chosen arbitrarily on S, the theorem is proved.

5. REMARKS

The methods employed here can be used to investigate space-times algebraically special on a given two-dimensional submanifold M. The only technical difficulty is that both normals to the submanifold cannot be given a simple canonical form. For this reason an unknown complex function appears in the work. The analysis gives no indication that the congruence l should be geodesic and shear-free on M, except in the trivial case when M contains a timelike world line.

ACKNOWLEDGMENT

The author would like to thank Dr. F. A. E. Pirani for his helpful comments on a previous draft of this paper.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Associative Algebra in the Problem of Mass Formulas

ARNO BÖHM*
Department of Physics, Syracuse University, Syracuse, New York

(Received 30 December 1966)

An associative algebra of continuous operators in a rigged Hilbert space, which contains the enveloping algebra of the Poincaré group and gives rise to a discrete mass spectrum, is studied. In Appendix B some general results on the representation of Lie algebras in a rigged Hilbert space are derived.

1. INTRODUCTION

ATTEMPTS to combine the Poincaré group $\mathfrak{P}$ with an intrinsic symmetry group to obtain a mass formula as a property of an enveloping algebra $\mathfrak{E}(G)$ of a large group G, containing $\mathfrak{P}$ and the intrinsic symmetry group, have been unsuccessful. According to O'Raifeartaigh's theorem,[1] the momentum operators P_μ can not be in the Lie algebra of such a group G. Therefore, keeping the general concept of the dynamical group approach,[2] but dropping the restriction to an enveloping algebra $\mathfrak{E}(G)$ of a group G, might overcome the difficulties. The first attempt in this direction has been undertaken by Werle,[3] who suggested using an algebra that is not an enveloping algebra of a group.

In a recent work[4] it has been suggested to describe a physical system by an algebra $\mathcal{A} \subset L(\phi)$ of continuous operators in a rigged Hilbert space

$$\phi \subset \mathcal{H} \subset \phi^X.$$

In the frame of this program we want to try to construct an associative algebra $\mathcal{A}$ containing the enveloping algebra of the Poincaré group $\mathfrak{E}(\mathfrak{P})$ and giving rise to a mass formula. That this associative algebra $\mathcal{A}$ cannot be the enveloping algebra of a Lie group is a consequence of the O'Raifeartaigh theorem,[1] which is valid for enveloping algebras of continuous operators in a rigged Hilbert space.[4]

2. DEFINING RELATIONS OF THE ALGEBRA

In the present work we study a simple model to learn about the mathematical problems involved; in a forthcoming work we will consider a more realistic case, and compare the results with experimental data. Our model, in which the (noncompact) intrinsic "noninvariance" group[2b] is $SL(2, c)$ and the intrinsic symmetry group is $SU(2)$, will provide a generalization of the mathematical structure of the rotator model.[4-6] The physical interpretation, however, will be different; i.e., the generators of the algebra will represent different physical observables. The associative algebra $\mathcal{A}$, as the mathematical image of this model, is generated by

$$\begin{aligned} &P_\mu, \quad L_{\mu\nu}, \quad M, \quad \mu, \nu = 0, 1, 2, 3, \\ &I_i, \quad F_i, \qquad\qquad i = 1, 2, 3, \end{aligned} \tag{1}$$

in which the multiplication is defined by the relations[7]

$$M^2 = P_\mu P^\mu, \quad [P_\mu, P_\nu] = 0,$$

$$[P_\rho, L_{\mu\nu}] = i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu),$$

$$[L_{\mu\nu}, L_{\rho\sigma}] = i(g_{\mu\rho}L_{\nu\sigma} + g_{\nu\sigma}L_{\mu\rho} - g_{\mu\sigma}L_{\nu\rho} - g_{\nu\rho}L_{\mu\sigma}), \tag{2}$$

$$[I_i, I_j] = i\epsilon_{ij}{}^k I_k, \tag{3}$$

$$[I_i, F_j] = i\epsilon_{ij}{}^k F_k, \quad [F_i, F_j] = -i\epsilon_{ij}{}^k I_k, \tag{4}$$

$$[L_{\mu\nu}, I_i] = 0, \quad [L_{\mu\nu}, F_i] = 0, \tag{5}$$

$$[P_\mu, I_i] = 0, \tag{6}$$

$$[MP_\mu, F_iM] = ib\epsilon_i{}^{kl}\{I_k, F_l\}P_\mu, \tag{7}$$

$$P_\mu F_i M - MF_i P_\mu = 0. \tag{8}$$

[b is a universal constant of the dimension $(\text{MeV})^2$; in the units we use, $\hbar = c = 1$.] We see that the subalgebra generated by P_μ, $L_{\mu\nu}$ is the enveloping algebra of the Poincaré group $\mathfrak{E}(\mathfrak{P})$, so that P_μ, $L_{\mu\nu}$, and M have the usual physical interpretation. The subalgebra generated by I_i is the enveloping algebra of $SU(2)$, which we want to call the isospin of our model; the subalgebra generated by I_i, F_i is

* Supported in part by the U.S. National Science Foundation and the U.S. Atomic Energy Commission.

[1] L. O'Raifeartaigh, Phys. Rev. Letters **14**, 332 (1965).

[2] (a) A. O. Barut, in *Proceedings of Seminar on High Energy Physics*, Trieste (1965) (IAEA, Vienna), and references therein; A. Böhm, in *Proceedings of Seminar on Elementary Particle Physics*, Boulder, Colorado (1966) and references therein. (b) N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan, Phys. Letters **19**, 322 (1965).

[3] J. Werle, preprint I.C./65/48, Trieste (1965).

[4] C. M. Andersen, A. Böhm, and A. M. Bouncristiani, "Rigged Hilbert Space and Mathematical Description of Physical Systems," Boulder Lectures, Mathematical Methods (1966).

[5] A. O. Barut and A. Böhm, Phys. Rev. **139**, 1107 (1965).

[6] A. Böhm, Nuovo Cimento **43**, 665 (1966).

[7] Notation $[A, B] = AB - BA$; $\{A, B\} = AB + BA$.

$\mathfrak{G}(SL(2,c))$, so that the F_i are the step operators transforming the different isospin states into each other.

We always want to assume that M^2 is a positive operator; i.e., $(f, M^2 f) \geq m_0^2(f,f)$, $m_0^2 > 0$ for every f of the representation space, which is a physical assumption excluding zero-mass particles from our consideration. As we see later, this assumption amounts to an appropriate choice of the irreducible component of $\mathcal{A}$, describing our physical system, and is always possible.

3. LIMIT OF ZERO-COUPLING CONSTANT

In addition to the above properties, $\mathcal{A}$ goes into the enveloping algebra of the direct product $\mathfrak{T} \times SL(2,c)$ if the coupling constant $b \to 0$. To prove this, it is sufficient to show that

$$[P_\mu, F_i] \to 0 \quad \text{for} \quad b \to 0.$$

For $b \to 0$, (7) goes into

$$MP_\mu F_i - F_i M P_\mu = 0, \tag{9a}$$

which together with (8) gives

$$\{P_\mu, [M, F_i]\} = 0, \tag{9b}$$

which can only be fulfilled if

$$[M, F_i] = 0. \tag{9c}$$

From this we obtain

$$[P_\mu, F_i] = 0 \tag{9d}$$

because of the equation

$$[P_\mu, F_i]M = [M, F_i]P_\mu, \tag{10}$$

which is an immediate consequence of relation (8).

4. SYSTEM OF COMMUTING OPERATORS

We want to find an appropriate complete system of commuting operators (CSCO). Since we assume that the physically preparable states are eigenstates of I_3 and $\mathbf{I}^2$, we choose the CSCO such that I_3 and $\mathbf{I}^2$ belong to it. The spin operator

$$\Gamma = -(P_\rho P^\rho)^{-1}\Gamma_\mu \Gamma^\mu, \quad \Gamma_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}P^\nu L^{\rho\sigma} \tag{11}$$

is already an invariant operator of $\mathcal{A}$. To prove this, it remains to be shown that

$$[\Gamma, F_i] = 0. \tag{12}$$

From relation (8) we obtain

$$[P_\mu/M, F_i] = 0, \tag{13}$$

an equation "derived" before by Werle,[3,8] Equation (12) is an immediate consequence of (13). The operator of the spin component

$$S_3 = L_3^{\nu}(P)\frac{\Gamma_\nu}{M} = \frac{\Gamma_3}{M} - \frac{\Gamma_0}{M(M+P_0)}P_3$$

also belongs to this system of commuting operators because of relations (5) and (6).

The Casimir operators of $SL(2,c)$

$$\mathfrak{Q} = \mathbf{F}^2 - \mathbf{I}^2, \quad \mathbf{I}\cdot\mathbf{F}$$

are also already invariant operators of $\mathcal{A}$. That $\mathfrak{Q}$ and $\mathbf{I}\cdot\mathbf{F}$ commute with $L_{\mu\nu}$ follows from relation (5); we still have to show that

$$[P_\mu, \mathfrak{Q}] = [P_\mu, \mathbf{F}^2] = 0 \tag{14}$$

and

$$[P_\mu, F_i I^i] = 0. \tag{15}$$

Proof of (14). From (10) we obtain, using (13),

$$[P_\mu, F_i]F^i = [M, F_i]F^i(P_\mu/M),$$

which is

$$P_\mu F_i F^i - F_i F^i P_\mu - F_i[P_\mu, F_i] = (MF_iF^i - F_iF^iM - F_i[M, F^i])(P_\mu/M).$$

Again using (10), we obtain

$$[P_\mu, F_iF^i] = [M, F_iF^i](P_\mu/M). \tag{16}$$

From (21)

$$[M^2, F_i] = ib\epsilon_{ikl}\{I^k, F^l\},$$

which is derived later, it follows that

$$[M^2, F_iF^i] = ib\epsilon_{ikl}\{\{I^k, F^l\}, F^i\}.$$

Using

$$\{\{I_k, F_l\}, F_i\} = \{I_k, \{F_l, F_i\}\} + (\delta_i^s\delta_{kl} - \delta_{il}\delta_k^s)I_s,$$

which follows from relation (4), we obtain

$$[M^2, F_iF^i] = ib\epsilon_{ikl}\{I_k\{F_l, F_i\}\} = 0.$$

Therefore $[M, F_iF^i] = 0$, and (14) follows from (16).

Proof of (15). From (13) we obtain

$$[MP_\mu, F_i] = \{P_\mu, [M, F_i]\} \tag{17}$$

and from (7), (8), and (6),

$$[MP_\mu, F_i]I^i = ib\epsilon_{ikl}\{I^k, F^l\}I^i(P_\mu/M) = 0 \tag{18}$$

because $\epsilon_{ikl}\{I^k, F^l\}I^i = 0$ as a consequence of (4). Thus we obtain from (17) and (18)

$$\{P_\mu, [M, F_iI^i]\} = 0$$

or

$$[M, F_iI^i] = 0. \tag{19}$$

From (10) follows

$$[P_\mu, F_iI^i] = [M, F_i]I^i(P_\mu/M),$$

and therefore from (19) we obtain (15).

[8] Equation (8) is only another way of writing this relation of Werle.

It follows in particular from (14) and (15) that M^2 commutes with $\mathfrak{Q}$ and $\mathbf{I}\cdot\mathbf{F}$, so that M^2 belongs to the system of commuting operators

$$\mathbf{F}^2-\mathbf{I}^2,\quad \mathbf{I}\cdot\mathbf{F},\quad I_3,\quad \mathbf{I}^2,\quad S_3. \tag{20}$$

However, M^2 is not an invariant operator of $\mathcal{A}$ because

$$[M^2, F_i] = ib\mathfrak{E}_{ikl}\{I^k, F^l\} \neq 0. \tag{21}$$

Proof of (21). From (7) we obtain

$$ib\mathfrak{E}_{ikl}\{I^k, F^l\}(P_\mu/M) = [MP_\mu, F_i];$$

multiplying with P_μ/M from the right and using (13), we obtain (21).

Comparing (21) with

$$[\mathbf{I}^2, F_i] = \{I^k[I_k, F_i]\} = -i\epsilon_{ikl}\{I^k, F^l\},$$

we see that the operator

$$Z = M^2 + b\mathbf{I}^2 \tag{22}$$

commutes with F_i.

$$[Z, F_i] = 0, \tag{23}$$

and we see immediately from the relations (6) and (5) that Z also commutes with all the other generators of $\mathcal{A}$; thus Z is a new invariant operator of $\mathcal{A}$.[9]

Following the convention we also adjoin P_i to our commuting system; this is possible because of (6). Thus our maximal commuting system is given by

$$Z, \Gamma, \mathfrak{Q} = \mathbf{F}^2-\mathbf{I}^2,\quad \mathbf{F}\cdot\mathbf{I},\quad I_3,\quad \mathbf{I}^2,\quad S_3,\quad P_i, \tag{24}$$

where the first four operators are invariant operators of $\mathcal{A}$, the eigenvalues of which characterize our physical (model) system.

5. CONSTRUCTION OF THE RIGGED HILBERT SPACE

To obtain the rigged Hilbert space in which $\mathcal{A}$ is an algebra of continuous operators, we first consider the limiting case $b \to 0$. In this case $\mathcal{A} \to \mathfrak{E}(\mathfrak{F} \times SL(2, c)$. We should therefore start with the representation space of $\mathfrak{E}(\mathfrak{F})$ and $\mathfrak{E}(SL(2, c))$.

We denote by $\phi_1 \subset \mathcal{H}_1 \subset \phi_1^X$ the rigged Hilbert space in which $\mathfrak{E}(\mathfrak{F})$ is an algebra of continuous operators. According to Appendix BI and BIII this is indeed a rigged Hilbert space. The topology in ϕ_1 is given by the countable set of scalar products

$$(\varphi \mid \psi)_P = (\varphi, \Delta_{\mathfrak{F}}^P\psi) \tag{25}$$

where $(\cdot,\cdot)$ is the scalar product in $\mathcal{H}_1$, which is the usual Hilbert space for an irreducible unitary representation of $\mathfrak{F}$, $\mathcal{H}(m, s)$, and $\Delta_{\mathfrak{F}}$ is the Nelson operator of $\mathfrak{F}$,[10]

$$\Delta_{\mathfrak{F}} = \frac{1}{b}P_0^2 + \frac{1}{b}\mathbf{P}^2 + \mathbf{N}^2 + \mathbf{M}^2,$$

$$= \frac{1}{b}\sum_\mu P_\mu P_\mu + \frac{1}{2}\sum_{\mu,\nu} L_{\mu\nu}L_{\mu\nu}. \tag{26}$$

In the case we consider here, where M^2 is positive, $\Delta_{\mathfrak{F}}$ is a positive operator, so that $(\varphi\Delta_{\mathfrak{F}}^P\varphi)$ is already a system of norms which is obviously equivalent to the system of (B1).

In $\phi_1 \subset \mathcal{H}(m, s) \subset \phi_1^X$ one conventionally uses the bases $|p_i, s_3\rangle$, which are eigenvectors of p_i and s_3 and elements of ϕ^X and therefore do not represent physically preparable states.[11] The physical states are the antilinear functionals $\psi \in \phi_1$ on ϕ_1^X:

$$\psi = \int \sum_{s_3} d\mu(p)\, |p_i s_3\rangle\langle s_3 p_i \mid \psi\rangle, \tag{27}$$

where[12]

$$\langle s_3, p_i \mid \psi\rangle = \psi(|p_i, s_3\rangle) = \overline{\langle\psi \mid p_i s_3\rangle}.$$

To make the subsequent considerations clearer, it is useful to introduce a basis of elements of ϕ_1 (which should be chosen adequately to the measurement system), though it is for our consideration irrelevant and we could as well work with the states $|p_i s_3\rangle$. We denote this basis by $|\xi, \eta, \zeta, s_3\rangle$, and it is

$$|\xi, \eta, \zeta, s_3\rangle = \int d\mu(p)\, |p_i s_3\rangle\langle p_i s_3 \mid \xi, \eta, \zeta, s_3\rangle. \tag{28}$$

The irreducible representation space of $\mathfrak{E}(SL(2, c))$ characterized by (k_0, a) has been constructed in Sec. IIC of Ref. 4 and is denoted by $\phi_2 \subset \mathcal{H}_2 \subset \phi_2^X$. Then we obtain in

$$\phi_1 \hat{\otimes} \phi_2 \subset \mathcal{H}_1 \bar{\otimes} \mathcal{H}_2 \subset (\phi_1 \hat{\otimes} \phi_2)^X \tag{29}$$

the rigged Hilbert space for $\mathfrak{E}(SL(2c) \times \mathfrak{F})$.[13] Here $\mathcal{H}_1 \bar{\otimes} \mathcal{H}_2$ is the completion of the direct product of linear spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\phi_1 \hat{\otimes} \phi_2$ is the completion of the direct product of the linear spaces $\phi_1 \otimes \phi_2$ with respect to the projective tensor product topology of ϕ_1 and ϕ_2.[13,14] In ϕ_2 we had the basis $|I_3, I, (k_0, a)\rangle$; in ϕ_1 we have the basis (28);

$$|I_3, I, \xi, \eta, \zeta, s_3; (k_0, a), s, m\rangle = |\xi, \eta, \zeta, s_3\rangle \otimes |I, I_3\rangle \tag{30}$$

[9] $\mathcal{A}$ has been constructed such that Z is an invariant operator.

[10] The factor $1/b$ is just a scale factor of dimension (MeV)$^{-2}$ converting the dimension (MeV) of P into the dimension 1 of I_i and L.

[11] Cf. Sec. IID in Ref. 4.

[12] We consider antilinear rather than linear functionals. The usual notation is $\langle s_3 p_i \mid \psi\rangle = \psi(\mathfrak{p}, s_3)$.

[13] K. Maurin and L. Maurin, Studia Math. 23 (1963).

[14] Let $\varphi_j^1 \in \phi_1$, $\varphi_j^2 \in \phi_2$, and $\phi_1 \otimes \phi_2 \ni \varphi = \Sigma\varphi_j^1 \otimes \varphi_2^j$; then $\|\varphi\|_{\nu_1,\nu_2} = \Sigma_j \|\varphi_j^1\|_{\nu_1}\|\varphi_j^2\|_{\nu_2}$ (where $\|\varphi_j^\alpha\|_{\nu\alpha}$ are the norms in ϕ_α) is a norm in $\phi_1 \otimes \phi_2$ and the projective tensor product $\phi_1 \hat{\otimes} \phi_2$ is the completion of $\phi_1 \otimes \phi_2$ with respect to the topology given by these norms.

is therefore a basis of $\phi_1 \otimes \phi_2$.[15] $\phi_1 \hat{\otimes} \phi_2$ is then the completion of the linear envelope of (30) with respect to the projective tensor product topology (in particular all eigenstates of $\mathbf{I}^2$ are in $\phi_1 \hat{\otimes} \phi_2$).

6. RIGGED HILBERT SPACE OF $\mathcal{A}$

We should also expect something like this for the rigged Hilbert space of $\mathcal{A}$. One physical system is described by an irreducible component of $\mathcal{A}$, which is characterized by the eigenvalue of the generators of the center of $\mathcal{A}$:

$$\begin{aligned} \mathbf{F}^2 - \mathbf{I}^2 &= (1 + a^2 - k_0^2)\mathbf{1}, \\ \mathbf{F} \cdot \mathbf{I} &= k_0 a \mathbf{1}, \\ \Gamma &= S(S+1)\mathbf{1}, \quad Z = z\mathbf{1}. \end{aligned} \tag{31}$$

Thus the invariants (k_0, a) and s of $\mathcal{E}(\mathfrak{T} \times SL(2, c))$ are also invariants of $\mathcal{A}$. Therefore an irreducible component of $\mathcal{A}$ contains only one irreducible component of $\mathcal{E}(SL(2, c))$ characterized by (k_0, a) and reduces out with respect to $\mathcal{E}(SU^I(2))$ generated by I_i in the same way as $\mathcal{E}_{(k_0,a)}(SL(2, c))$:

$$\mathcal{H}(k_0, a) = \sum_{I=k_0}^{\infty} \mathcal{H}^I, \tag{32}$$

i.e., $\mathcal{A}$ contains every isospin $I \geq k_0$ exactly once. As s is an invariant of $\mathcal{A}$, it describes only one spin. However, m is no longer an invariant. So we should expect instead of (29) a direct sum or integral of (29). We therefore take instead of $\mathcal{H}_1$ the continuous direct sum

$$\tilde{\mathcal{H}}_1 = \int_{\substack{m^2>0 \\ p_0>0}} \oplus \mathcal{H}(m, s)\, d(m^2). \tag{33}$$

Let us call the corresponding countably Hilbert space $\tilde{\phi}_1$; then we conjecture in an irreducible subspace of

$$\tilde{\phi}_1 \hat{\otimes} \phi_2 \subset \tilde{\mathcal{H}}_1 \bar{\otimes} \mathcal{H}_2 \subset (\tilde{\phi}_1 \hat{\otimes} \phi_2)^X,$$

the rigged Hilbert space for a representation of $\mathcal{A}$.

In Appendix A it is shown that $\mathcal{A}$ is an algebra of continuous operators in the countably Hilbert space

$$\tilde{\phi} = (\tilde{\phi}_1 \otimes \phi_2, (\cdot, \cdot)_p), \tag{34}$$

where

$$(\varphi_1 \psi)_p = (\varphi, \theta^p \psi) \quad \varphi_1 \psi \in \tilde{\phi}_1 \otimes \phi_2 \tag{35}$$

with

$$\theta = \mathbf{I}^2 + \Delta_{\mathfrak{T}} = \mathbf{I}^2 + (1/b)P_0^2 + (1/b)\mathbf{P}^2 + \mathbf{N}^2 + \mathbf{M}^2 \tag{36}$$

and $(\cdot, \cdot)$ the scalar product in $\tilde{\mathcal{H}}_1 \otimes \mathcal{H}_2$. From the definition of the projective tensor product topology one sees that the topologies of $\tilde{\phi}$ and $\tilde{\phi}_1 \otimes \phi_2$ are equivalent.

7. MASS SPECTRUM IN AN IRREDUCIBLE REPRESENTATION

With these preparations it is now easy to reduce $\mathcal{A}$ into its irreducible components $\mathcal{A}_{z,s,(a,k_0)}$ and find the mass spectrum in an irreducible representation space of $\mathcal{A}$.

We take the expectation value (eigenvalue) of Eq. (22) in the states (30) and obtain

$$m^2 = z - bI(I+1). \tag{37}$$

For an irreducible component $\mathcal{A}_{z,s,(a,k_0)}$, z is a number characterizing one physical system together with $s, (k_0, a)$.[16]

For the reducible representation space of $\mathcal{A}$,

$$\begin{aligned} \mathcal{H} &= \left(\int \oplus \mathcal{H}(m, s)\, d(m^2) \bar{\otimes} \mathcal{H}(k_0, a) \right), \\ &= \int \oplus \mathcal{H}(m, s)\, d(m^2) \bar{\otimes} \sum_{I}^{(a,k_0)} \mathcal{H}^I_{(a,k_0)}, \\ &= \int d(m^2) \sum_{I} \oplus (\mathcal{H}(m, s)\ \mathcal{H} \bar{\otimes}^I_{(a,k_0)}); \end{aligned} \tag{38}$$

i.e., the direct sums over I and m are independent such that, for a given I, $\mathcal{H}(m, s) \bar{\otimes} \mathcal{H}^I$ is contained in $\mathcal{H}$ for any m. In the representation space $\mathcal{H}(z, s, k_0, a)$ of the irreducible component $\mathcal{A}_{z,s,(k_0,a)}$, for a given I only that $\mathcal{H}(m, s) \otimes \mathcal{H}_I$ can be contained in $\mathcal{H}(z, s, k_0, a)$ for which $m = m_I$ given by (37). Thus reducing $\mathcal{H}$ out with respect to Z we obtain from (38) by reordering the direct sums according to (37):

$$\begin{aligned} \mathcal{H} &= \int dz \sum_{I} \oplus (\mathcal{H}(m_I, s) \bar{\otimes} \mathcal{H}^I)_{(z,s,k_0,a)}, \\ &= \int dz \oplus \mathcal{H}(z, s, k_0, a). \end{aligned} \tag{39}$$

Thus we obtain the canonical triplet of spaces of an irreducible component of $\mathcal{A}_{z,s,(k_0,a)}$:

$$\phi \subset \mathcal{H}(z, s, k_0, a) \subset \phi^X, \tag{40}$$

where ϕ is the irreducible subspace of $\tilde{\phi}$ with the same topology given by (35). ϕ is even nuclear, so that (40) is the rigged Hilbert space of our system. According to theorems by Grothendieck, the direct product, the direct sum, and the completion of nuclear spaces are again nuclear. By this the nuclearity of ϕ follows from the nuclearity of ϕ_1 and ϕ_2, which is proved in Appendix Secs. BII and BIII. [It also follows directly by Theorem 1 of Ref. 17 if we choose $(\mathbf{I}^2 + 1)$ for the operator A in Sec. BII.]

[15] G. Köthe, *Topologische Lineare Räume* (Springer-Verlag, Berlin, 1960), Vol. I.

[16] Here we see the confirmation of the statement in Sec. 2; for $b < 0$, $m^2 > 0$ in a system characterized by $z > 0$.

[17] J. E. Roberts, Commun. Math. Phys. 3, 98 (1966).

Instead of the basis (28) of physical states (which we do not completely know) we could also have used the conventional basis of generalized eigenstates of the maximal commuting system (24) of self-adjoint operators in $\phi \subset \mathcal{H} \subset \phi^X$. If we assume that (24) is a CSCO, the Dirac spectral theorem applies (cf. Sec. 10 of Ref. 4), and we obtain the basis of generalized eigenvectors:

$$|I_3, I, p_i s_3; (a, k_0), s,\rangle.$$

The expectation (eigen)value of (22) in this basis again gives (37).

Thus we have found in (40) the rigged Hilbert space of an irreducible component of $\mathcal{A}$, which reduces out with respect to $\mathfrak{T} \times SU^I(2)$ [by (39)] according to[18]

$$\phi = \sum_{I \geq k_0} \hat{\oplus} (\mathcal{H}(m_I, s) \bar{\otimes} \mathcal{H}^I), \tag{41}$$

where the spectrum of m_I is given by the mass formula (37).

8. FINITE MASS SPECTRUM

Concluding, we remark that we could have chosen instead of $\mathcal{H}(k_0, a)$ an irreducible representation space $\mathcal{H}(k_0, n)$ of a nonunitary representation of $SL(2, c)$ or a unitary representation of $SO(4)$ (cf. Ref. 6), which would have amounted to considering $SO(4)$ as the intrinsic noninvariance group. The only difference in the foregoing consideration would have been that, instead of (41), we would obtain

$$\phi = \sum_{I=k_0}^{k_0+n-1} (\mathcal{H}(m_I, s) \otimes \mathcal{H}^I), \tag{42}$$

i.e., a finite mass spectrum. In this case too, for $b > 0$, (37) would not lead to unphysical masses and describe a spectrum in which the masses decrease with increasing isospin.

ACKNOWLEDGMENTS

The author gratefully acknowledges valuable discussions on the subject of this paper with Professor N. Mukunda, Professor L. O'Raifeartaigh, Professor F. Rohrlich, and Professor E. C. G. Sudarshan.

APPENDIX A

To prove the continuity of $\mathcal{A}$ in ϕ, where the topology $\mathcal{T}_\phi$ is given by the countable set of norms

$$(\varphi, \varphi)_p = (\varphi, \theta^p \varphi),$$
$$\theta = \mathbf{I}^2 + \frac{1}{b}\sum_\mu P_\mu P_\mu + \frac{1}{2}\sum_{\mu,\nu} L_{\mu\nu} L_{\mu\nu}, \tag{A1}$$

we have to show that from

$$\varphi_n \to 0 \quad \text{for} \quad n \to \infty, \quad \text{i.e.,} \quad (\varphi_n, \varphi_n)_p \to 0 \quad \text{for every} \quad p, \tag{A2}$$

follows

$$A_i \varphi_n \to 0, \quad \text{i.e.,} \quad (A_i \varphi_n, A_i \varphi_n)_p \to 0 \quad \text{for every} \quad p,$$

for every $A_i \in$ generators of $\mathcal{A}$.

I. The Continuity of I_i

From

$$I_i \theta^p I^i = \theta^p I_i I^i \tag{A3}$$

follows

$$\begin{aligned}(I_i\varphi_n, I_i\varphi_n)_p &= (I_i\varphi_n, \theta^p I_i \varphi_n) = (\varphi_n \theta^p \mathbf{I}^2 \varphi_n), \\ &\leq (\varphi_n, \theta^p \theta \varphi_n) = (\varphi_n \theta^{p+1} \varphi_n), \\ &= (\varphi_n, \varphi_n)_{p+1},\end{aligned} \tag{A4}$$

so that from (A2) we obtain

$$(I_i\varphi_n, I^i\varphi_n)_p \to 0 \quad \text{for every} \quad p \quad \text{and} \quad n \to \infty,$$

and from

$$(I_i\varphi_n, I_i\varphi_n)_p \leq (I_i \varphi_n I^i \varphi_n)_p$$

we obtain

$$(I_i\varphi_n, I_i\varphi_n) \to 0 \quad \text{for} \quad n \to \infty \text{ and every } p; \tag{A5}$$

i.e., I_1 is continuous. In the same way it follows that all I_i, $i = 1, 2, 3$, are continuous.

II. Continuity of F_i

We write

$$\theta^p = [I^2 + (1/b)P_0^2 + (1/b)\mathbf{P}^2 + \mathbf{N}^2 + \mathbf{M}^2]^p$$

in the form

$$\theta^p = \sum_{\nu=0}^{p} \binom{p}{\nu} \left(2\mathbf{I}^2 + \frac{1}{b}P_0^2\right)^\nu \left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}\mathbf{Z}\right)^{p-\nu}, \tag{A6}$$

using (22). Then we calculate

$$\begin{aligned}[F_i, \theta^p]F_i &= \sum_{\nu=0}^{p} \binom{p}{\nu} \left[F_i, \left(2\mathbf{I}^2 + \frac{1}{b}P_0^2\right)^\nu\right] \\ &\quad \times F^i \left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}\mathbf{Z}\right)^{p-\nu} \\ &= \sum_{\nu=0}^{p} \sum_{\alpha=0}^{\nu} \binom{p}{\nu}\binom{\nu}{\alpha} \frac{2^\alpha}{b^{(\nu-\alpha)}} \\ &\quad \times ([F_i, \mathbf{I}^{2\alpha}]P_0^{2(\nu-\alpha)} + \mathbf{I}^{2\alpha}[F_i, P_0^{2(\nu-\alpha)}]) \\ &\quad \times F^i \left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}\mathbf{Z}\right)^{p-\nu},\end{aligned} \tag{A7}$$

where we have used the fact that $L_{\mu\nu}$ and Z commute

[18] $\hat{\otimes}$ indicates that the limit is to be taken with respect to the topology of ϕ and is to be contrasted with $\mathcal{H}(z, s, k_0, a) = \Sigma \oplus \mathcal{H}(m_I, s) \bar{\otimes} \mathcal{H}^I$ where the limit is being taken with respect to the Hilbert space topology.

with F_i, with

$$[F_i, P_0^{2(\nu-\alpha)}] = \frac{P_0^{2(\nu-\alpha)}}{M^{2(\nu-\alpha)}}[F_i, M^{2(\nu-\alpha)}],$$
$$= -b^{(\nu-\alpha)}\frac{P_0^{2(\nu-\alpha)}}{M^{2(\nu-\alpha)}}[F_i, \mathbf{I}^{2(\nu-\alpha)}], \quad \text{(A8)}$$

where we have used (13) and (22). We obtain for

$$\text{II} \overset{\text{def}}{=} ([F_i, \mathbf{I}^{2\alpha}]P_0^{2(\nu-\alpha)} + \mathbf{I}^{2\alpha}[F_i, P_0^{2(\nu-\alpha)}])F^i,$$
$$\text{II} = (P_0^{2(\nu-\alpha)}/M^{2(\nu-\alpha)})$$
$$\times ([F_i, \mathbf{I}^{2\alpha}]M^{2(\nu-\alpha)} - b^{(\nu-\alpha)}\mathbf{I}^{2\alpha}[F_i, \mathbf{I}^{2(\nu-\alpha)}])F^i,$$
$$\text{II} = (P_0^{2(\nu-\alpha)}/M^{2(\nu-\alpha)})$$
$$\times ([F_i, \mathbf{I}^{2\alpha}]F^iM^{2(\nu-\alpha)} + [F_i, \mathbf{I}^{2\alpha}][\mathbf{I}^{2(\nu-\alpha)}, F_i]b^{(\nu-\alpha)}$$
$$- b^{(\nu-\alpha)}\mathbf{I}^{2\alpha}[F_i, \mathbf{I}^{2(\nu-\alpha)}]F^i).$$

In the last step we have again used (22). With

$$[F_i, \mathbf{I}^{2\alpha}][I^{2(\nu-\alpha)}, F_i] = F_i[\mathbf{I}^{2\nu}, F^i] + \mathbf{I}^{2\alpha}[\mathbf{I}^{2(\nu-\alpha)}, F_i]F^i$$
$$+ F_i[\mathbf{I}^{2\alpha}, F^i]\mathbf{I}^{2(\nu-\alpha)},$$

which one calculates easily, we obtain

$$\text{II} = \frac{P_0^{2(\nu-\alpha)}}{M^{2(\nu-\alpha)}}([F_i, \mathbf{I}^{2\alpha}]F^iM^{2(\nu-\alpha)}$$
$$+ b^{(\nu-\alpha)}(2\mathbf{I}^{2\alpha}[\mathbf{I}^{2(\nu-\alpha)}, F_i]F^i + F_i[\mathbf{I}^{2\nu}, F^i]$$
$$+ F_i[\mathbf{I}^{2\alpha}, F^i]\mathbf{I}^{2(\nu-\alpha)}. \quad \text{(A9)}$$

We now use

$$[F_i, \mathbf{I}^{2n}]F^i = C_n(k_0, a)\mathbf{I}^{2n}$$
$$+ C_{n-1}(k_0, a)\mathbf{I}^{2(n-1)} + \cdots C_0(k_0, a). \quad \text{(A10)}$$

Here $C_m(k_0, a)$ are constants depending on the invariants

$$\mathbf{F}^2 - \mathbf{I}^2 = (1 + a^2 - k_0)\mathbf{1}, \quad \text{(A11)}$$
$$\mathbf{F}\cdot\mathbf{I} = k_0, a\,\mathbf{1} \quad \text{(A12)}$$

of the system. (A10) can be calculated from the commutation relations of $SL(2, c)$: (3), (4), (A11), and (A12). Inserting (A10) into (A9) we obtain

$$\text{II} = \left(\sum_{\kappa=0}^{\alpha} C_\kappa(k_0a)\mathbf{I}^{2\kappa}\right)P_0^{2(\nu-\alpha)}$$
$$+ \left(\sum_{\kappa=0}^{\nu} a_\kappa(k_0, a)\mathbf{I}^{2\kappa}\right)b^{(\nu-\alpha)}\left(\frac{P_0}{M}\right)^{2(\nu-\alpha)}, \quad \text{(A13)}$$

where the new constants $a_\nu(k_0, a)$ are combinations of various $C_\kappa(k_0, a)$.

Inserting (A13) into (A7) and rearranging the three finite sums, we obtain

$$[F_i, \theta^p]F^i = \left(\sum_{\nu=0}^{p}\tilde{c}_\nu\mathbf{I}^{2\nu}\left(\frac{P_0^2}{b}\right)^{(p-\nu)} + \sum_{\nu=0}^{p}\tilde{a}_\nu\mathbf{I}^{2\nu}\left(\frac{P_0}{M}\right)^{2(p-\nu)}\right)$$
$$\times \left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}Z\right)^{p-\nu}, \quad \text{(A14)}$$

where the new constants $\tilde{c}_\nu$ and $\tilde{a}_\nu$ are combinations of the C_ν and a_ν and binomial coefficients.

Let $\varphi \in \phi$; we then obtain from (A14)

$$(\varphi, [F_i, \theta^p]F^i\varphi)$$
$$\leq \sum_{\nu=0}^{p}\tilde{c}_\nu\left(\varphi, \theta^p\left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}Z\right)^{p-\nu}\varphi\right)$$
$$+ \sum_{\nu=0}^{p}\tilde{a}_\nu\left(\varphi\theta^\nu\frac{\theta^{(p-\nu)}}{M^{2(p-\nu)}}\left(\mathbf{N}^2 + \mathbf{M}^2 - \frac{1}{b}Z\right)^{p-\nu}\varphi\right)$$
$$\leq \sum_{\nu=0}^{p}\tilde{c}_\nu(\varphi, \theta^p\theta^{p-\nu}\varphi)$$
$$+ \sum_{\nu=0}^{p}\tilde{a}_\nu\left(\varphi\theta^p\theta^{(p-\nu)}\frac{1}{M^{2(p-\nu)}}\varphi\right). \quad \text{(A15)}$$

Here we have used the fact that $\mathbf{I}^2$, $\mathbf{N}^2$, $\mathbf{M}^2$, $P_0^2 + \tilde{\mathbf{P}}^2$, $(1/b)Z$ are positive, which follows from the fact that our space $\phi \subset \mathcal{H} \subset \phi^X$ is constructed in such a way that it reduces into representation spaces of the groups $SL(2, c)$ or $\mathfrak{P}$ such that the generators are essentially self-adjoint[19], and from the fact that for our physical systems Z has a positive eigenvalue.[20]

As

$$F_i\theta^pF_i = \theta F_iF^i + [F_i, \theta^p]F^i$$
$$= \theta^p\mathbf{I}^2 + \alpha(k_0, a)\theta^p + [F_i, \theta^p]F^i, \quad \text{(A16)}$$

we obtain

$$(F_i\varphi, \theta^pF^i\varphi)$$
$$= (\varphi F_i\theta^pF^i\varphi) \leq (\varphi, \theta^{(p+1)}\varphi) + \alpha(\varphi, \theta^p\varphi)$$
$$+ \sum_{\nu=0}^{p}\tilde{c}_\nu(\varphi, \theta^{2p-\nu}\varphi)$$
$$+ \sum_{\nu=0}^{p}\tilde{a}_\nu(\varphi\theta^{2p-\nu}\varphi)\left(\varphi\frac{1}{M^{2(p-\nu)}}\varphi\right), \quad \text{(A17)}$$

where

$$[\varphi\theta^{2p-\nu}(1/M^{2(p-\nu)})\varphi] \leq (\varphi\theta^{2p-\nu}\varphi)[\varphi(1/M^2)^{(p-\nu)}\varphi]$$

has been used. With the aid of (A17) the continuity of F_i is easily seen: Let

$$\varphi_n \to 0 \quad \text{for} \quad n \to \infty; \quad \text{i.e.,}$$
$$(\varphi_n\varphi_n)_p = (\varphi_n\theta^p\varphi_n) \to 0 \quad \text{for every } p. \quad \text{(A18)}$$

Then

$$(F_1\varphi_n, F_1\varphi_n)_p = (F_1\varphi_n\theta^pF_1\varphi_n) \leq (F_i\varphi_n\theta^pF^i\varphi_n)$$
$$\leq (\varphi_n\theta^{(p+1)}\varphi_n) + \alpha(\varphi_n\theta^p\varphi_n)$$
$$+ \sum_{\nu=0}^{p}\tilde{c}_\nu(\varphi_n, \theta^{2p-\nu}\varphi_n)$$
$$+ \sum_{\nu=0}^{p}\tilde{a}_\nu(\varphi_n, \theta^{2p-\nu}\varphi_n)(\varphi_n, M^{-2(p-\nu)}\varphi_n). \quad \text{(A19)}$$

M^2 is a continuous operator, as shown in Appendix AIII. Since M^2 is a positive operator, $(M^2)^{-1}$ is defined everywhere in ϕ. Thus M^2 fulfills the conditions

[19] E. Nelson and W. F. Stinespring, Am. J. Math. **81**, 547 (1959).

[20] For $b < 0$ the argumentation has to be changed a bit by going to the absolute value; the result also remains true in this case.

of Banach's theorem on the inverse operator[21] and therefore M^{-2} is also a continuous operator; i.e., from $\varphi_n \to 0$ follows $M^{-2(p-\nu)}\varphi_n \to 0$.

By using (A18), then it follows from (A19) that

$$(F_1\varphi_n, F_1\varphi_n)_p \to 0 \quad \text{for every} \quad p,$$

i.e., F_1, and in the same way F_2 and F_3 are continuous.

III. Continuity of the Generators of $\mathfrak{s}$

To prove the continuity of P_μ in the topology of ϕ given by the scalar products

$$(\varphi, \psi)_p = (\varphi, \theta^p\psi) \quad \text{with} \quad \theta = \Delta_{\mathfrak{s}} + \mathbf{I}^2,$$

we calculate

$$|(\varphi P_\kappa \theta^p P_\kappa \varphi)| = \sum_{n=0}^{p} \binom{p}{n} |(\varphi P_\kappa \Delta_{\mathfrak{s}}^n P_\kappa \mathbf{I}^{2(p-n)}\varphi)|$$
$$\leq \sum_{n=0}^{p} \binom{p}{n} \|P_\kappa \Delta_{\mathfrak{s}}^n P_\kappa \varphi\| \, \|\mathbf{I}^{2(p-n)}\varphi\|.$$

We use, as in Appendix B, Lemma 6.3 of Ref. 22, from which follows

$$\|P_\kappa \Delta_{\mathfrak{s}}^n P_\kappa \varphi\| \leq c \, \|(\Delta - 1)^{n+1}\varphi\|, \quad c = \text{const}, \quad \text{(A20)}$$

so that

$$|(\varphi P_\kappa \theta^p P_\kappa \varphi)| \leq c \sum_{n=0}^{p} \binom{p}{n} \sum_{m=0}^{n+1} \binom{n+1}{m} \|\Delta^{n+1}\varphi\| \, \|\mathbf{I}^{2(p-n)}\varphi\|,$$

$$|(\varphi P_\kappa \theta^p P_\kappa \varphi)| \leq c \sum_{n=0}^{p} \sum_{m=0}^{n+1} \binom{p}{n} \binom{n+1}{m} |(\varphi \Delta^{2(n+1)}\varphi)|^{\frac{1}{2}} \, \|\mathbf{I}^{2(p-n)}\varphi\|. \quad \text{(A21)}$$

We have to show that from $\|\varphi_\nu\|_p \to 0$ for every p and $\nu \to \infty$ follows $\|P_\kappa \varphi_\nu\|_q \to 0$ for every q. Using (A21), we obtain

$$\|P_\kappa \varphi_\nu\|_q^2 = |(\varphi_\nu P_\kappa \theta^p P_\kappa \varphi_\nu)|$$
$$\leq c \sum_{n=0}^{p} \sum_{m=0}^{n+1} \binom{p}{n} \binom{n+1}{m} \|\varphi_\nu\|_{2(n+1)} \, \|\mathbf{I}^{2(p-n)}\varphi_\nu\| \to 0$$

for $\nu \to \infty$ because $\|\varphi_\nu\|_{2(n+1)} \to 0$ for every n and because of the continuity of $\mathbf{I}^2$, which is proved in Sec. AI.

The continuity of $L_{\mu\nu}$ follows in the same way; to prove it one just has to replace P_μ by $L_{\mu\nu}$ in the above proof.

The continuity of M follows from the continuity of $M^2 = P_\mu P^\mu$ and the commutativity of M and θ, and the continuity of M^{-1} follows then from Banach's theorem of the inverse operator.

[21] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1967), Vol. II.
[22] E. Nelson, Ann. Math. **70**, 572 (1959).
[23] We use the notations of Ref. 4.

APPENDIX B

I. Continuity of the Enveloping Algebra with Respect to the Topology in ϕ Given by the Nelson Operator

Let $E(G)$ be the enveloping algebra of the Lie group G of linear operators on a linear space Ψ; let $(\cdot, \cdot)$ be the scalar product in Ψ with respect to which $X \in \mathfrak{L}(G)$ (Lie algebra of G) is symmetric; let $\mathfrak{H}$ be the completion of Ψ with respect to the topology given by the scalar product $(\cdot, \cdot)$. Let X_i be a basis of $\mathfrak{L}(G)$ and $\Delta = \Sigma X_i^2$ the Nelson operator; and let ϕ be the completion of Ψ with respect to the topology given by the countable number of compatible scalar products:

$$(\varphi, \psi)_p = (\varphi, (\Delta + 1)^p\psi). \quad \text{(B1)}$$

Then $E(G)$ is an algebra of continuous operators on the countably normed space $\phi \subset \mathfrak{H}$.

The proof is a simple consequence of a lemma by Nelson: $\phi \subset \mathfrak{H}$ follows from $(\cdot, \cdot)_{p=0} = (\cdot, \cdot)$. It remains to be shown that the generators X_i of E are continuous operators. According to theorem of Sec. IE of Ref. 4 we have to show that from $\varphi_n \to 0$ with respect to $\mathfrak{T}_\phi$, (topology of ϕ) follows $X_i\varphi_n \to 0$ with respect to $\mathfrak{T}_\phi$.

$$\varphi_n \to 0 \quad \text{with respect to} \quad \mathfrak{T}_\phi \Leftrightarrow (\varphi_n, \varphi_n)_p \to 0 \quad \text{(B2)}$$

for every p (cf. Sec. IC, Ref. 4). Therefore it is to be shown that

$$(X_i\varphi_n, X_i\varphi_n)_q \to 0$$

for every q.

From lemma 6.3 of Ref. 22 follows

$$(\varphi X_i(\Delta + 1)^q X_i\varphi) \leq k(\varphi(\Delta - 1)^{q+1}\varphi) \quad \text{(B3)}$$

for every $\varphi \in \Psi$ and all positive integers q, where k is a constant $k < \infty$. From this we obtain

$$(X_i\varphi_n, X_i\varphi_n)_q = (\varphi_n, X_i(\Delta + 1)^q X_i\varphi_n)$$
$$\leq k(\varphi_{nq}(\Delta - 1)^{q+1}\varphi_n)$$
$$\leq \sum_{\nu}^{\leq q+1} c_\nu(\varphi_n(\Delta + 1)^{\nu+1}\varphi_n)$$
$$\leq \sum_{\nu}^{\leq q+1} c_\nu(\varphi_n, \varphi_n)_{\nu+1} \to 0;$$

i.e., $X_i\varphi_n \to 0$.

For the triplet $\phi \subset \mathfrak{H} \subset \phi^\times$ to be a rigged Hilbert space it has still to be shown that ϕ is nuclear.

The proof of the nuclearity for this general case is not known to us. However, in some more special cases of physical importance, where G is semisimple or G is the Poincaré group and $\mathfrak{H}$ is a representation space of an irreducible unitary representation of G, the nuclearity can easily be proved with the aid of a

theorem by Roberts[17] and some results on group representation due to Harish-Chandra and Godement.[24]

II. Nuclearity of ϕ for Semisimple Lie Groups

Let $g \to T_g$ be a unitary irreducible representation of the semisimple Lie group G in $\mathcal{H}$ and ϕ the countably Hilbert space of I. Then ϕ is nuclear and consequently $\phi \subset \mathcal{H} \subset \phi^X$, a rigged Hilbert space.

To prove the nuclearity of ϕ we have to show, according to Theorem 1 of Ref. 17, that there exists an $A \in E(G) \subset L(\phi)$ with $\hat{A}$ self-adjoint in $\mathcal{H}$, whose inverse is nuclear.

Let K be the maximal compact subgroup of G and C the second-order Casimir operator. We consider $A = (C + 1)^n \in E(G)$, where n is a sufficiently large integer. C and C^n are essentially self-adjoint according to a theorem by Nelson and Stinespring[19]; consequently A is also essentially self-adjoint.

It remains to be shown that A^{-1} is nuclear. We first remark that it is sufficient to show that A^{-1} is Hilbert–Schmidt, because the product of two Hilbert–Schmidt operators is nuclear so that A^2 would then fulfill the above conditions. A Hermitian operator B is Hilbert–Schmidt if $B = \Sigma\lambda_k P_k$, where the projections P_k project on finite dimensional spaces $\mathcal{H}_k$ and

$$\Sigma(|\lambda_k| \dim \mathcal{H}_k)^2 < \infty.$$

Let d be an equivalence class of unitary representations of K and let $\mathcal{H}_d$ denote the subspace of $\mathcal{H}$ which transforms according to d when G is restricted to K. Then $\dim \mathcal{H}_d < \infty$ (Harish–Chandra)[24] and the irreducible representation d occurs at most (degree of d) times in the restriction of $g \to T_g$ to K (Godement)[24]; consequently $(\dim \mathcal{H}_d) \leq (\text{degree of } d)^2$. From this and the fact that the irreducible representation d of K is characterized by a finite number of Casimir operators it follows that $\dim \mathcal{H}_i \leq N \dim \mathcal{H}_{d(c_i)}$, where $\mathcal{H}_i$ is the eigenspace of C with eigenvalue c_i and N is a number depending upon the number of independent Casimir operators of K, i.e., depending upon K. Therefore $A^{-1} = \Sigma[1/(c_i - 1)^n]P_i$, where P_i projects on finite dimensional spaces $\mathcal{H}_i$. Further, we calculate

$$\sum_i \left(\frac{1}{(c_i + 1)^n} \dim \mathcal{H}_i\right)^2 \leq N^2 \sum_i \left(\frac{1}{(c_i + 1)^n} \dim \mathcal{H}_{d(c_i)}\right)^2$$
$$\leq N^2 \sum_i \left[\frac{1}{(c_i + 1)^n} (\text{degree of } d(c_i))^2\right]^2 \leq \infty$$

for sufficiently large n. {Comparing Eqs. (117) and (103) of Ref. 25 one finds that [degree of $d(c_i)$] $\leq$ const $(c_i)^m$, where m depends upon the number of roots of K, i.e., upon K.} This completes the proof of the nuclearity of ϕ.

III. Nuclearity of ϕ for the Poincaré Group

The proof of Sec. BII depended on the fact that $\dim \mathcal{H}_d < \infty$, which is true for semisimple Lie groups but not generally so that for a general Lie group the proof of the nuclearity cannot be carried out as in Sec. BII. However, in the case that G is the Poincaré group (or more generally, any inhomogeneous pseudo-orthogonal group) we can reduce the proof to the former case.

Let $H(m, s)$ be a representation space of a unitary irreducible representation of $\mathfrak{P}$. Then the countable Hilbert space ϕ with the norms given by (B1) is nuclear.

To show this, we use the connection between an irreducible representation (m, s) of $\mathfrak{P}$ and the irreducible representation (α, s) of $SO(4, 1)$.[26]

Let P_μ, $L_{\mu\nu}$ be the generators of $E_{m,s}(\mathfrak{P})$; then $L_{\mu\nu}$ and

$$B_\mu = P_\mu + (\lambda/2m)(P^\rho L_{\rho\mu} + L_{\rho\mu}P^\rho) \in E_{m,s}(\mathfrak{P})$$

form a basis of the Lie algebra $SO(4, 1)$ in the irreducible representation

$$\{\alpha = [(m^2/\lambda^2) + 9/4 - S(S + 1)]^{\frac{1}{2}}, S\}$$

of $SO(4, 1)$. The elements B_i and $M_i = \frac{1}{2}\mathcal{E}_{ikl}L^{kl}$ $(i, k, l = 1, 2, 3)$ form a basis of the Lie algebra of the maximal compact subgroup $SO(4)$ of $SO(4, 1)$. Therefore the second-order Casimir operator $\mathbf{B}^2 + \mathbf{M}^2 \in E_{m,s}(\mathfrak{P})$ fulfills all the conditions of the operator A in BII, which proves the nuclearity of ϕ.

[24] E. M. Stein, A Survey of Representations of Non-Compact Groups, Proceedings Trieste Seminar (1965) (IAEA Vienna, 1966) and references therein.

[25] G. Racah, "Group Theory and Spectroscopy," Lecture Notes, Princeton (1951).

[26] E.g., A. Böhm, Phys. Rev. **145**, 1212 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Klein–Gordon Equation for Spinors

EGON MARX*
Clarkson College of Technology, Potsdam, New York

(Received 12 October 1966)

Spinors obeying the Dirac equation also obey the Klein–Gordon equation, but the converse is not true. In this paper we make a systematic study of four-component spinors obeying the Klein–Gordon equation, with special regard for the additional solutions. The starting point is the Lagrangian density $\mathfrak{L} = \bar{\psi}_{,\mu}\gamma_\mu\gamma_\nu\psi_{,\nu} - m^2\bar{\psi}\psi$, and we first develop from it the theory of a classical spinor field. We then proceed to the canonical quantization of this field and are confronted by some anticommutators of creation and annihilation operators equal to -1, and the subsequent need for an indefinite metric in Hilbert space. Quantum electrodynamics can be reformulated, and in spite of a modified fermion propagator, gradient coupling, and vertices with two photon lines, the amplitude for Compton scattering to order e^2 is the usual one. Special problems arising for massless fermions are indicated, and we note that the four-fermion point interaction is now renormalizable. Some interesting variations of strong interactions also become possible.

1. INTRODUCTION

THE usual presentations of the Dirac equation

$$(-i\hat{\partial} + m)\psi(x) = 0 \tag{1}$$

involve, directly or indirectly, the factorization of the Klein–Gordon equation

$$(\Box + m^2)\psi(x) = 0. \tag{2}$$

It is quite obvious that not all solutions of Eq. (2) are solutions of Eq. (1); the purpose of this paper is to study the properties and possible relevance for physical problems of the discarded solutions.

In Sec. 2 we deal with the classical theory of a four-component spinor field (it has the same behavior as the Dirac spinor field under Lorentz transformations) that has Eq. (2) for its equation of motion. We choose an appropriate Lagrangian density, and we find the canonical momenta, the conserved quantities, and their expansions in terms of plane wave solutions. It is observed that the contributions to conserved quantities coming from the "anomalous" solutions have the opposite sign from those pertaining to the "normal" solutions. The separation of the angular momentum tensor into spin and orbital parts shows some interesting features even when the anomalous part is eliminated.

In Sec. 3 we follow the canonical quantization procedure to obtain the anticommutation relations for this field. The anomalous solutions again show an additional minus sign; therefore an indefinite metric in Hilbert space is introduced to avoid contradictions. The space of physical states can then be restricted by means of the Dirac equation applied to the state vectors.

In Sec. 4 we obtain the electromagnetic interaction of the field by means of the usual gauge-invariant substitution. Although the fermion propagator and the vertices are changed, the amplitude for Compton scattering in lowest order remains the same.

In Sec. 5 we present a number of difficulties that arise when the mass of the fermion tends to zero. In Secs. 6 and 7 we discuss briefly the possible forms of weak and strong interactions, respectively.

Finally, in Sec. 8 we mention some of the open questions that can be a subject of further research.

Conventions to be used throughout this paper are that Greek indices run from 0 to 3 and Latin indices from 1 to 3, unless otherwise stated; we use the time-favoring metric, i.e., the nonzero components of the metric tensor $\delta_{\mu\nu}$ are

$$\delta_{00} = -\delta_{11} = -\delta_{22} = -\delta_{33} = 1, \tag{3}$$

so that the scalar product of two four-vectors is

$$a \cdot b = a_\mu b_\mu = a_0 b_0 - \mathbf{a} \cdot \mathbf{b} = a_0 b_0 - a_k b_k, \tag{4}$$

where an explicit representation of the Dirac matrices γ_μ is used; it is

$$\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \tag{5}$$

where the σ_i are the 2×2 Pauli matrices, and we add the matrix

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_0 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{6}$$

We write

$$\hat{a} = a_\mu\gamma_\mu, \tag{7}$$

* Present address: Drexel Institute of Technology, Philadelphia, Pa.

and derivatives are indicated by

$$\partial_\mu = \partial/\partial x^\mu \tag{8}$$

or an index μ following a comma. Summation over repeated indices is implied, modified according to Eq. (4) for Greek subindices, and we set $\hbar = c = 1$.

2. CLASSICAL SPINOR FIELD

The factorization of the Klein–Gordon equation into

$$(i\hat{\partial} + m)(-i\hat{\partial} + m)\psi = 0 \tag{9}$$

shows that either of the equations

$$(\mp i\hat{\partial} + m)\psi = 0 \tag{10}$$

could be chosen for the Dirac equation. Bogoliubov and Shirkov[1] further remark that both are equivalent under the unitary transformation of the Dirac matrices

$$\gamma'_\mu = \gamma_5\gamma_\mu\gamma_5^{-1} = -\gamma_\mu. \tag{11}$$

Nevertheless, when both equations are considered *simultaneously*, they are *not* equivalent.

The Lagrangian and Hamiltonian formulations for fields obeying the Dirac equation appear unsatisfactory for various reasons. We can use the Lagrangian density

$$\mathfrak{L} = i\bar{\psi}\hat{\partial}\psi - m\bar{\psi}\psi, \tag{12}$$

where[2]

$$\bar{\psi} = \psi^\dagger\gamma_0 = \tilde{\psi}^*\gamma_0, \tag{13}$$

and vary ψ and $\bar{\psi}$ (or $\psi^\dagger$) independently to get the equations of motion. Corresponding to these generalized coordinates we have the canonical momenta[3]

$$\Pi^\psi = i\psi^\dagger, \tag{14}$$

$$\Pi^{\bar{\psi}} = 0, \tag{15}$$

a rather unsatisfactory result. A more symmetric answer is obtained from

$$\mathfrak{L}' = \tfrac{1}{2}i(\bar{\psi}\gamma_\mu\psi_{,\mu} - \bar{\psi}_{,\mu}\gamma_\mu\psi) - m\bar{\psi}\psi, \tag{16}$$

whence

$$\Pi^\psi = \tfrac{1}{2}i\psi^\dagger, \tag{17}$$

$$\Pi^{\bar{\psi}} = -\tfrac{1}{2}i\gamma_0\psi, \tag{18}$$

but it is still bothersome to have a factor $\frac{1}{2}$ and to again have the momenta as coordinates. The basic formulation of a variational problem becomes unclear when we have equations of motion of first order, since the generalized coordinates are to be held fixed at *arbitrary* values *both* for the initial and the final times.[4] Other lines of inquiry[5] have also led to the consideration of both Eqs. (10) where the double sign is attached to the mass and the correspondence to particles and antiparticles is made. Therefore, there seems to be ample justification for a closer examination of the solutions of the Klein–Gordon equation.

We can use the projection operators

$$P_\pm = (\pm i\hat{\partial} + m)/2m \tag{19}$$

[operating on the vector space of solutions of Eq. (2)] to separate any solution into two parts

$$\psi = \varphi + \chi, \tag{20}$$

$$\varphi = P_+\psi, \tag{21}$$

$$\chi = P_-\psi, \tag{22}$$

so that φ and χ obey

$$(-i\hat{\partial} + m)\varphi = 0, \tag{23}$$

$$(i\hat{\partial} + m)\chi = 0. \tag{24}$$

A Lagrangian density that is a Lorentz scalar and gives the equations of motion (2) is[6]

$$\mathfrak{L} = \bar{\psi}_{,\mu}\gamma_\mu\gamma_\nu\psi_{,\nu} - m^2\bar{\psi}\psi. \tag{25}$$

The canonical momenta are now

$$\Pi^\psi = \bar{\psi}_{,\mu}\gamma_\mu\gamma_0, \tag{26}$$

$$\Pi^{\bar{\psi}} = \gamma_0\gamma_\nu\psi_{,\nu}, \tag{27}$$

which become, by Eqs. (20), (23), and (24),

$$\Pi^\psi = im(\bar{\varphi} - \bar{\chi})\gamma_0, \tag{28}$$

$$\Pi^{\bar{\psi}} = -im\gamma_0(\varphi - \chi), \tag{29}$$

and the Hamiltonian density is

$$\mathfrak{H} = \Pi^\psi\gamma_0\gamma_k\psi_{,k} + \bar{\psi}_{,k}\gamma_k\gamma_0\Pi^{\bar{\psi}} + \Pi^\psi\Pi^{\bar{\psi}} + m^2\bar{\psi}\psi. \tag{30}$$

Noether's theorem[7] gives us the usual tensors that obey differential conservation laws. They are

$$T_{\mu\nu} = \bar{\psi}_{,\nu}\gamma_\mu\gamma_\lambda\psi_{,\lambda} + \bar{\psi}_{,\lambda}\gamma_\lambda\gamma_\mu\psi_{,\nu} - \mathfrak{L}\delta_{\mu\nu}, \tag{31}$$

$$M_{\mu\nu\rho} = x_\nu T_{\mu\rho} - x_\rho T_{\mu\nu} + \tfrac{1}{2}i(\bar{\psi}_{,\lambda}\gamma_\lambda\gamma_\mu\sigma_{\nu\rho}\psi - \bar{\psi}\sigma_{\nu\rho}\gamma_\mu\gamma_\lambda\psi_{,\lambda}), \tag{32}$$

$$j_\mu = i(\bar{\psi}\gamma_\mu\gamma_\lambda\psi_{,\lambda} - \bar{\psi}_{,\lambda}\gamma_\lambda\gamma_\mu\psi), \tag{33}$$

[1] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publishers, Inc., New York, 1959), Chap. 1, p. 64.

[2] A star indicates complex conjugation; a tilde, the transpose; a dagger, the conjugate transpose for classical fields and the Hermitian conjugate for quantized fields.

[3] We omit the spinor indices whenever the meaning is clear.

[4] This question was first brought to my attention by E. A. Power.

[5] I. Saavedra, Nucl. Phys. 77, 673 (1966).

[6] The more immediate choice $\mathfrak{L} = \bar{\psi}_{,\mu}\psi_{,\mu} - m^2\bar{\psi}\psi$ was rejected because it does not give the same electromagnetic interaction as the Dirac equation by the usual gauge-invariant substitution.

[7] E. Noether, Nachr. Ges. Göttingen, Math.-Phys. Kl., p. 235 (1918); see also Ref. 1, p. 19, Eq. (11), and other books on field theory.

where

$$\sigma_{\nu\rho} = -\tfrac{1}{2}i(\gamma_\nu\gamma_\rho - \gamma_\rho\gamma_\nu). \tag{34}$$

The energy momentum tensor $T_{\mu\nu}$ is not symmetric; the symmetrized form[8,9] is

$$\Theta_{\mu\nu} = \tfrac{1}{4}(\bar\psi_{,\lambda}\gamma_\lambda\gamma_\nu\psi_{,\mu} + \bar\psi_{,\lambda}\gamma_\lambda\gamma_\mu\psi_{,\nu} + \bar\psi_{,\mu}\gamma_\nu\gamma_\lambda\psi_{,\lambda} + \bar\psi_{,\nu}\gamma_\mu\gamma_\lambda\psi_{,\lambda} - \bar\psi_{,\lambda\mu}\gamma_\lambda\gamma_\nu\psi - \bar\psi_{,\lambda\nu}\gamma_\lambda\gamma_\mu\psi - \bar\psi\gamma_\nu\gamma_\lambda\psi_{,\lambda\mu} - \bar\psi\gamma_\mu\gamma_\lambda\psi_{,\lambda\nu}) \tag{35}$$

$$= \tfrac{1}{2}im(\bar\varphi\gamma_\nu\varphi_{,\mu} + \bar\varphi\gamma_\mu\varphi_{,\nu} - \bar\varphi_{,\mu}\gamma_\nu\varphi - \bar\varphi_{,\nu}\gamma_\mu\varphi - \bar\chi\gamma_\nu\chi_{,\mu} - \bar\chi\gamma_\mu\chi_{,\nu} + \bar\chi_{,\mu}\gamma_\nu\chi + \bar\chi_{,\nu}\gamma_\mu\chi). \tag{36}$$

In the latter form, it can be recognized as ($2m$ times) the *difference* of the symmetrized energy-momentum tensors of two Dirac fields φ and χ; this is strongly reminiscent of the way the vector and the scalar fields are combined to form a Lorentz vector field,[10] where the scalar field also comes in with the "wrong" sign.[11] Similarly,

$$j_\mu = 2m(\bar\varphi\gamma_\mu\varphi - \bar\chi\gamma_\mu\chi). \tag{37}$$

Another form of the conserved current is

$$j'_\mu = i(\bar\psi\psi_{,\mu} - \bar\psi_{,\mu}\psi); \tag{38}$$

they are equivalent, since the difference can be expressed in the form $f_{\alpha\mu,\alpha}$, with

$$f_{\alpha\mu} = \bar\psi\sigma_{\alpha\mu}\psi. \tag{39}$$

The asymmetry of $T_{\mu\nu}$ casts serious doubts on the usual separation of the angular momentum density tensor into orbital and spin parts,[12] since then

$$L'_{\mu\nu\rho} = x_\nu T_{\mu\rho} - x_\rho T_{\mu\nu}, \tag{40}$$

$$S'_{\mu\nu\rho} = \tfrac{1}{2}i(\bar\psi_{,\lambda}\gamma_\lambda\gamma_\mu\sigma_{\nu\rho}\psi - \bar\psi\sigma_{\nu\rho}\gamma_\mu\gamma_\lambda\psi_{,\lambda}) \tag{41}$$

are not even conserved. We find that this question is best discussed in terms of momentum-space expansions; they are

$$\varphi(x) = \sum_\lambda \int d^3p(2\pi)^{-\frac{3}{2}}(2E)^{-\frac{1}{2}}[u(\mathbf{p},\lambda)b(\mathbf{p},\lambda)e^{-ip\cdot x} + v(\mathbf{p},\lambda)\,d^*(\mathbf{p},\lambda)e^{ip\cdot x}], \tag{42}$$

$$\chi(x) = \sum_\lambda \int d^3p(2\pi)^{-\frac{3}{2}}(2E)^{-\frac{1}{2}}[v(\mathbf{p},\lambda)f(\mathbf{p},\lambda)e^{-ip\cdot x} + u(\mathbf{p},\lambda)g^*(\mathbf{p},\lambda)e^{ip\cdot x}], \tag{43}$$

where the index λ ranges over the two helicity states, the spinors u and v satisfy

$$(\hat p - m)u = 0, \tag{44}$$

$$(\hat p + m)v = 0, \tag{45}$$

$$\bar uu = -\bar vv = 1, \tag{46}$$

and we always set

$$E = p_0 = +(\mathbf{p}^2 + m^2)^{\frac{1}{2}}. \tag{47}$$

The charge and the energy-momentum vector become

$$Q = \sum_\lambda \int d^3p(b^*b + dd^* - f^*f - gg^*), \tag{48}$$

$$P_\nu = \sum_\lambda \int d^3pp_\nu(b^*b - dd^* - f^*f + gg^*), \tag{49}$$

where the only unusual feature is the additional minus sign for the terms coming from χ. After a lengthy calculation, we also find the angular momentum tensor $M_{\nu\rho}$, which we split into the separately conserved parts $L_{\nu\rho}$ and $S_{\nu\rho}$ according to

$$L_{\nu\rho} = mi\sum_\lambda \int \frac{d^3p}{E}\left\{p_\rho\left[\frac{\partial}{\partial p^\nu}(u^\dagger b^*)ub + \frac{\partial}{\partial p^\nu}(v^\dagger d)vd^* - \frac{\partial}{\partial p^\nu}(v^\dagger f^*)vf - \frac{\partial}{\partial p^\nu}(u^\dagger g)ug^*\right] - (\rho\leftrightarrow\nu)\right\}, \tag{50}$$

$$S_{ij} = m\sum_\lambda \int d^3p(2E)^{-1}(-u^\dagger\sigma_{ij}ub^*b - v^\dagger\sigma_{ij}vdd^* + v^\dagger\sigma_{ij}vf^*f + u^\dagger\sigma_{ij}ugg^*), \tag{51}$$

$$S_{0j} = mi\int d^3p(2\,|\mathbf{p}|)^{-1}\Big[u^\dagger(\tfrac12)\,\Sigma_j\,u(-\tfrac12)b^*(\tfrac12)b(-\tfrac12) + v^\dagger(\tfrac12)\,\Sigma_j\,v(-\tfrac12)d(\tfrac12)d^*(-\tfrac12) - v^\dagger(\tfrac12)\,\Sigma_j\,v(-\tfrac12) \times f^*(\tfrac12)f(-\tfrac12) - u^\dagger(\tfrac12)\,\Sigma_j\,u(-\tfrac12)g(\tfrac12)g^*(-\tfrac12)\Big] + \text{h.c.}, \tag{52}$$

where derivatives with respect to p_0 are to be set equal to zero (as a matter of notation), and in Eq. (52),

$$\Sigma_k = -\tfrac{1}{2}\epsilon_{ijk}\sigma_{ij}, \tag{53}$$

i.e.,

$$\mathbf{\Sigma} = \begin{pmatrix}\boldsymbol{\sigma} & 0\\ 0 & \boldsymbol{\sigma}\end{pmatrix}. \tag{54}$$

[8] G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1949), Appendix I, p. 217.

[9] J. Rzewuski, *Field Theory* (PWN—Polish Scientific Publishers, Warsaw, 1964), Part I, Chap. 2, p. 105.

[10] I. Goldberg and E. Marx, Nuovo Cimento (to be published).

[11] This is certainly not a new problem; it is present in the electromagnetic field, and it has also been discussed for vector fields by T. D. Lee and C. N. Yang, Phys. Rev. **128**, 885 (1962). Negative energies appear also in other contexts; see, for instance, A. Pais and G. E. Uhlenbeck, Phys. Rev. **79**, 145 (1950); R. E. Norton, J. Math. Phys. **6**, 981 (1965). In the theory of quantized fields, the same "wrong" sign affects commutators or anticommutators, and an indefinite metric is usually introduced in the Hilbert space of state vectors to avoid certain inconsistencies due to this sign change. Nevertheless, this does *not* solve the problem of negative values for the expectation of the energy (defined in the usual manner, but in the space with the indefinite metric); it becomes positive only after the subsidiary conditions in the corresponding theory are introduced.

[12] See, for instance, J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1955), Chap. 1, p. 20; Ref. 1, pp. 25 and 80. A different approach to this question is given by J. Hilgevoord and E. A. DeKerf, Physica **31**, 1002 (1965).

This separation of the total angular momenta into orbital and spin parts is different from the usual ones.[12] It is a natural separation insofar as $L_{\nu\rho}$ is an extension of the angular momentum for a scalar particle and S_{ij} has the form which would be expected for the spin vector; on the other hand, S_{0j} contains only terms with mixed helicities, although the contributions from φ and χ are still separate, as throughout. It is also worth noticing that, in going over to $L_{\nu\rho}$ and $S_{\nu\rho}$, not only the time-dependent parts of $L'_{\nu\rho}$ and $S'_{\nu\rho}$ cancel each other, but that some time independent cross terms do likewise. A similar separation is also possible in the case of Lorentz vector fields.[10]

The question whether a reasonable relativistic quantum mechanics for a single spin $\frac{1}{2}$ particle exists might not be really significant. Feshbach and Villars[13] make a very convincing case for the interpretation of a scalar field that obeys the Klein–Gordon equation: they replace the (positive definite) probability density of nonrelativistic quantum mechanics by a (no longer positive) "charge" density in the relativistic theory; the problem of negative energy states is then solved by the introduction of an indefinite metric in a certain "charge" space (*not* in the Hilbert space of the quantized field theory). In their equally relativistic spin $\frac{1}{2}$ theory, they go back to a positive definite probability and the problem of negative energies remains unsolved until the fields become anticommuting operators by a second quantization; *then* the probability density becomes the (indefinite) charge density, and the energy becomes positive. An infinite sea of electrons with negative energies is not a very satisfactory assumption for a single particle theory either. Our charge density j_0, given by Eq. (37), is indefinite, but Eqs. (48) and (49) show that there is no correlation between the signs of the charge and energy, so that an interpretation similar to that given in Ref. 13 for a scalar particle is apparently not possible. The basic difference between anticommuting field operators and commuting classical fields suggests that maybe the interpretation of relativistic quantum mechanics for spin $\frac{1}{2}$ particles should not be pursued too far.

3. QUANTIZATION OF THE FREE FIELD

We can use the canonical quantization procedure and set

$$\{\psi_\alpha(\mathbf{x}, t), \Pi^\psi_\beta(\mathbf{x}', t)\} = i\delta_{\alpha\beta}\delta^3(\mathbf{x} - \mathbf{x}'), \quad (55)$$

$$\{\bar\psi_\alpha(\mathbf{x}, t), \Pi^{\bar\psi}_\beta(\mathbf{x}', t)\} = -i\delta_{\alpha\beta}\delta^3(\mathbf{x} - \mathbf{x}'), \quad (56)$$

where the spinor indices α and β go from 1 to 4 and $\delta_{\alpha\beta}$ is the ordinary Kronecker delta, and we set all other equal-time anticommutators equal to zero. The minus sign in Eq. (56) is necessary to avoid getting only mixed anticommutators of φ and χ different from zero, and it also makes Eq. (56) the Hermitian conjugate of Eq. (55). We obtain from them the nonzero anticommutators

$$\{\varphi_\alpha(\mathbf{x}, t), \bar\varphi_\beta(\mathbf{x}', t)\} = (1/2m)(\gamma_0)_{\alpha\beta}\delta^3(\mathbf{x} - \mathbf{x}'), \quad (57)$$

$$\{\chi_\alpha(\mathbf{x}, t), \bar\chi_\beta(\mathbf{x}', t)\} = -(1/2m)(\gamma_0)_{\alpha\beta}\delta^3(\mathbf{x} - \mathbf{x}'), \quad (58)$$

and but for the minus sign in Eq. (58) and the normalization factor $1/(2m)$, they are the usual ones for Dirac spinor fields.

These anticommutation relations imply

$$\{\varphi_\alpha(x), \bar\varphi_\beta(x')\} = -iS_{\alpha\beta}(x - x')/2m, \quad (59)$$

$$\{\chi_\alpha(x), \bar\chi_\beta(x')\} = -iS'_{\alpha\beta}(x - x')/2m, \quad (60)$$

$$\{\psi_\alpha(x), \bar\psi_\beta(x')\} = i\delta_{\alpha\beta}\Delta(x - x'), \quad (61)$$

where

$$\Delta(x) = -i(2\pi)^{-3}\int d^4p\epsilon(p_0)\delta(p^2 - m^2)e^{-ip\cdot x}, \quad (62)$$

$$S(x) = -(i\hat\partial + m)\Delta(x), \quad (63)$$

$$S'(x) = -(-i\hat\partial + m)\Delta(x). \quad (64)$$

We can also derive from them

$$\{b(\mathbf{p}, \lambda), b^\dagger(\mathbf{p}', \lambda')\} = \{d(\mathbf{p}, \lambda), d^\dagger(\mathbf{p}', \lambda')\} = \delta_{\lambda\lambda'}\delta^3(\mathbf{p} - \mathbf{p}'), \quad (65)$$

$$\{f(\mathbf{p}, \lambda), f^\dagger(\mathbf{p}', \lambda')\} = \{g(\mathbf{p}, \lambda), g^\dagger(\mathbf{p}', \lambda')\} = -\delta_{\lambda\lambda'}\delta^3(\mathbf{p} - \mathbf{p}'). \quad (66)$$

Going over to the discrete language of box normalization, we are faced by anticommutation relations such as

$$\{a, a^\dagger\} = -1, \quad (67)$$

which are clearly in contradiction with the usual definition of Hermitian conjugate. The problem is similar to the one presented by the time component of the electromagnetic potential,[14] and can be solved in the same manner by introducing an indefinite metric in Hilbert space.[15] We can use (a direct product of) state vectors with only two components,[16] with a

[13] H. Feshbach and F. Villars, Rev. Mod. Phys. **30**, 24 (1958).

[14] Reference 1, Chap. 2, p. 130ff.

[15] S. N. Gupta, Proc. Phys. Soc. (London) **A63**, 681 (1950); K. Bleuler, Helv. Phys. Acta **23**, 567 (1950); L. K. Pandit, Suppl. Nuovo Cimento **11**, 159 (1959); K. L. Nagy, Suppl. Nuovo Cimento **17**, 92 (1960).

[16] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill Book Company, Inc., New York, 1965), Chap. 13, p. 47ff.

metric

$$\eta = \eta^{\dagger} = \eta^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{68}$$

The adjoint of the annihilation operator

$$a = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{69}$$

is then

$$a^{\star} = \eta a^{\dagger} \eta = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \tag{70}$$

so that

$$\{a, a^{\star}\} = -1. \tag{71}$$

We consequently have to change the definition of $\bar{\psi}$ to

$$\bar{\psi} = \psi^{\star}\gamma_0, \tag{72}$$

and replace $f^{\dagger}$ and $g^{\dagger}$ by $f^{\star}$ and $g^{\star}$; naturally we have $b^{\star} = b^{\dagger}$, $d^{\star} = d^{\dagger}$.

The vacuum expectation value of Wick's chronological product of $\psi_{\alpha}(x)$ and $\bar{\psi}_{\beta}(x')$ is calculated to be

$$\langle 0|\, T(\psi_{\alpha}(x)\bar{\psi}_{\beta}(x'))\,|0\rangle = \tfrac{1}{2}\delta_{\alpha\beta}\Delta_F(x - x'), \tag{73}$$

where

$$\Delta_F(x) = 2i(2\pi)^{-4}\int d^4p(p^2 - m^2 + i\epsilon)^{-1}e^{-ip\cdot x}. \tag{74}$$

We are now faced with the problem of the physical interpretation of states with negative norm, such as those containing one particle created by $f^{\star}$ or $g^{\star}$; it is always possible to restrict the physical states to those (with positive norm) containing particles created by $b^{\dagger}$ and $d^{\dagger}$. The physical state vectors then satisfy

$$(-i\gamma_{\mu}\psi^{(+)}_{,\mu} + m\psi^{(+)})\,|\,\rangle = 0, \tag{75}$$

$$(i\bar{\psi}^{(+)}_{,\mu}\gamma_{\mu} + m\bar{\psi}^{(+)})\,|\,\rangle = 0, \tag{76}$$

or

$$\chi^{(+)}\,|\,\rangle = 0, \quad \bar{\chi}^{(+)}\,|\,\rangle = 0, \tag{77}$$

where the positive frequency parts $\psi^{(+)}$, $\bar{\psi}^{(+)}$, $\chi^{(+)}$ and $\bar{\chi}^{(+)}$ contain the annihilation operators.

4. ELECTROMAGNETIC INTERACTIONS

By addition of a Lagrangian density for the free electromagnetic field and the usual gauge-invariant substitution

$$\partial_{\mu} \to D_{\mu} = \partial_{\mu} - ieA_{\mu}, \tag{78}$$

the Lagrangian density (25) becomes

$$\mathfrak{L} = (D^{*}_{\mu}\bar{\psi})\gamma_{\mu}\gamma_{\nu}D_{\nu}\psi - m^2\bar{\psi}\psi - \tfrac{1}{4}F_{\mu\nu}F_{\mu\nu}, \tag{79}$$

where the fields $F_{\mu\nu}$ are defined in terms of the potentials A_{μ} by

$$F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}. \tag{80}$$

The equations of motion for the spinor field become

$$(\hat{D}^2 + m^2)\psi = 0, \tag{81}$$

or

$$(D^2 + m^2)\psi + \tfrac{1}{2}eF_{\mu\nu}\sigma_{\mu\nu}\psi = 0, \tag{82}$$

where the so-called magnetic moment term[6] $\frac{1}{2}eF_{\mu\nu}\sigma_{\mu\nu}\psi$ is the same distinctive spin $\frac{1}{2}$ term obtained by operating with $\hat{D}$ on the Dirac equation

$$(-i\hat{D} + m)\psi = 0. \tag{83}$$

It is still possible to define the gauge, covariant fields

$$\varphi^{A} = (i\hat{D} + m)\psi/2m, \tag{84}$$

$$\chi^{A} = (-i\hat{D} + m)\psi/2m, \tag{85}$$

since

$$P^{A}_{\pm} = (\pm i\hat{D} + m)/2m \tag{86}$$

are again projection operators in the space of solutions of Eq. (81).

The action integral

$$I = \int \mathfrak{L}[\psi(x), A(x)]\, d^4x \tag{87}$$

is obviously invariant under proper Lorentz transformations, since the spinor ψ and the vector A transform in the usual way. So it is under space reflection $\mathfrak{P}$, time reversal $\mathfrak{T}$, and charge conjugation $\mathfrak{C}$, characterized by

$$\psi(x) \to \psi'(x') = \mathfrak{P}\psi(x)\mathfrak{P}^{-1} = \gamma_0\psi(-\mathbf{x}', t'), \tag{88}$$

$$A_{\mu}(x) \to A'_{\mu}(x') = \mathfrak{P}A_{\mu}(x)\mathfrak{P}^{-1} = \delta_{\mu\mu}A_{\mu}(-\mathbf{x}', t') \tag{89}$$

(no summation over μ),

$$\psi(x) \to \psi''(x'') = \mathfrak{T}\psi(x)\mathfrak{T}^{-1} = i\gamma_1\gamma_3\psi^{*}(\mathbf{x}'', -t''), \tag{90}$$

$$A_{\mu}(x) \to A''_{\mu}(x'') = \mathfrak{T}A_{\mu}(x)\mathfrak{T}^{-1} = \delta_{\mu\mu}A_{\mu}(\mathbf{x}'', -t''), \tag{91}$$

and

$$\psi(x) \to \psi'''(x''') = \mathfrak{C}\psi(x)\mathfrak{C}^{-1} = -i\gamma_2\psi^{*}(x'''), \tag{92}$$

$$A_{\mu}(x) \to A'''_{\mu}(x''') = \mathfrak{C}A_{\mu}(x)\mathfrak{C}^{-1} = -A_{\mu}(x'''). \tag{93}$$

Then the new ψ fields obey the same Eq. (81) with the new potentials.

The action is also invariant under the chirality transformation

$$\psi(x) \to \gamma_5\psi(x). \tag{94}$$

It should be noted that, contrary to $\mathfrak{P}$, $\mathfrak{T}$, and $\mathfrak{C}$, the Dirac equation is not invariant under a chirality transformation; moreover, if φ^{A} is a solution of Eq. (83), $\gamma_5\varphi^{A}$ belongs to the subspace of the χ^{A}.

The interaction Lagrangian density is defined to be

$$\mathfrak{L}_I = ie(\bar{\psi}\hat{A}\gamma_{\lambda}\psi_{,\lambda} - \bar{\psi}_{,\lambda}\gamma_{\lambda}\hat{A}\psi) + e^2A^2\bar{\psi}\psi, \tag{95}$$

which introduces both gradient coupling and vertices with two photon lines. Nevertheless, we have calculated the amplitude for the Compton effect in lowest order with the above interaction, using[17]

$$S = T\left[\exp\left(i\int \mathfrak{L}_I(x)\, d^4x\right)\right] \tag{96}$$

and omitting "singular expressions arising when differentiating singular functions,"[18] and the result is precisely the same one that is obtained from the usual theory.[19] It should then be of interest to compute radiative corrections and see whether there is a difference that might be tested experimentally.

By examining the vertices and propagators,[20] we come to the conclusion that the theory is still renormalizable.

5. MASSLESS FERMIONS

The theory set forth in Secs. 2 and 3 runs into quite predictable problems when the mass m is allowed to be zero. The normalization conditions for u and v have to be changed if their components are to remain finite; we can take, for instance,[21]

$$\bar{u}u = \bar{v}v = 0, \tag{97}$$

$$u^\dagger u = v^\dagger v = 2E. \tag{98}$$

Moreover, the u and v used before become essentially equal, i.e., the solutions of

$$\hat{p}u = 0 \tag{99}$$

with $p_0 = E > 0$. Consequently, the spinors φ and χ become essentially the same, say $\varphi^{(0)}$, and we have to find the other solutions of

$$\Box\psi = 0 \tag{100}$$

by a limiting process or otherwise. They correspond to $E < 0$ and can be expressed in terms of the spinors $\gamma_0 u$ that complete the set of orthogonal spinors with the u. Completeness now reads

$$\sum_\lambda [u(\mathbf{p}, \lambda)u^\dagger(\mathbf{p}, \lambda) + \gamma_0 u(\mathbf{p}, \lambda)u^\dagger(\mathbf{p}, \lambda)\gamma_0] = 2E \times 1, \tag{101}$$

where the 1 in the right-hand side is the 4×4 unit matrix. We can separate the general solution of Eq. (100) into two parts

$$\psi = \varphi^{(0)} + \chi^{(0)}, \tag{102}$$

where $\varphi^{(0)}$ and $\chi^{(0)}$ satisfy

$$i\varphi\hat{\partial}^{(0)} = 0, \tag{103}$$

$$i\gamma_0\hat{\partial}\gamma_0\chi^{(0)} = 0. \tag{104}$$

Their momentum-space expansions are of the form

$$\varphi^{(0)}(x) = \sum_\lambda \int d^3p(2\pi)^{-\frac{3}{2}}(2E)^{-\frac{1}{2}}[u(\mathbf{p}, \lambda)b(\mathbf{p}, \lambda)e^{-ip\cdot x} + u(\mathbf{p}, \lambda)\, d^*(\mathbf{p}, \lambda)e^{ip\cdot x}], \tag{105}$$

$$\chi^{(0)}(x) = \sum_\lambda \int d^3p(2\pi)^{-\frac{3}{2}}(2E)^{-\frac{1}{2}}[\gamma_0 u(\mathbf{p}, \lambda)f(\mathbf{p}, \lambda)e^{-ip\cdot x} + \gamma_0 u(\mathbf{p}, \lambda)g^*(\mathbf{p}, \lambda)e^{ip\cdot x}]. \tag{106}$$

The projection operators now become integral operators, and we use the matrix

$$F(x) = i(2\pi)^{-3}\int d^3p(2E)^{-2}\sum_\lambda u(\mathbf{p}, \lambda)u^\dagger(\mathbf{p}, \lambda)e^{-ip\cdot x}, \tag{107}$$

$$= i(2\pi)^{-3}\int d^3p(2E)^{-2}\hat{p}\gamma_0 e^{-ip\cdot x} \tag{108}$$

to express

$$\varphi^{(0)}(x) = \int d^3x'\{[F(x - x') + F^\dagger(x - x')]\partial'_0\psi(x') - \partial'_0[F(x - x') + F^\dagger(x - x')]\psi(x')\}, \tag{109}$$

$$\chi^{(0)}(x) = \int d^3x'\{\gamma_0[F(x - x') + F^\dagger(x - x')]\gamma_0\partial'_0\psi(x') - \gamma_0\partial'_0[F(x - x') + F^\dagger(x - x')]\gamma_0\psi(x')\}. \tag{110}$$

We can check that expressions (109) and (110) are independent of t' for functions ψ that satisfy Eq. (100) and vanish sufficiently rapidly at infinity.

If we now try to follow a procedure similar to that used for massive fermions, starting from the Lagrangian density

$$\mathfrak{L} = \bar{\psi}_{,\mu}\gamma_\mu\gamma_\nu\psi_{,\nu}, \tag{111}$$

we are led to unreasonable expressions such as

$$\Pi^\psi = \bar{\psi}_{,\mu}\gamma_\mu\gamma_0 \tag{112}$$

$$= -i\sum_\lambda \int d^3p(2\pi)^{-\frac{3}{2}}(2E)^{\frac{1}{2}}(u^\dagger f^* e^{ip\cdot x} - u^\dagger g e^{-ip\cdot x}), \tag{113}$$

$$P_\mu = \sum_\lambda \int d^3p(2E)p_\mu(b^*f + dg^* + f^*b + gd^*). \tag{114}$$

The canonical quantization procedure also has to be abandoned for this field. The reason an expression such as (114) occurs is clear: The original fields φ and

[17] Ref. 1, Chap. 3, p. 226.

[18] S. S. Schweber, *An Introduction to the Relativistic Quantum Field Theory* (Row, Peterson and Company, Evanston, Illinois, 1961), Chap. 14, p. 482; see also Ref. 1, Chap. 3, p. 230ff.

[19] A similar result is obtained by L. M. Brown, Phys. Rev. **111**, 957 (1958), for the electromagnetic interactions of a two-component spinor that also obeys the Klein–Gordon equation. As far as these theories reformulate ordinary quantum electrodynamics, they should be equivalent.

[20] Ref. 1, Chap. 4, p. 340ff.

[21] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1964), Chap. 10, p. 249. It should be noted that their "projection operators" do not satisfy any of the usual relations $P_1 + P_2 = 1$, $P_1^2 = P_1$, $P_2^2 = P_2$.

χ become equal, but they contributed terms with opposite signs to the conserved quantities; hence, these now vanish unless the new solutions are introduced. These problems do not appear when only one field φ is considered, with a Lagrangian density such as

$$\mathfrak{L}' = \tfrac{1}{2}i(\bar{\varphi}\gamma_\mu\varphi_{,\mu} - \bar{\varphi}_{,\mu}\gamma_\mu\varphi), \tag{115}$$

which gives Eq. (103) for φ.

This peculiar behavior for massless particles is certainly not limited to the spin $\frac{1}{2}$ case. In the much better known spin 1 case, the scalar part of the Lorentz vector field and the longitudinal component of the vector part become equal, while their contributions to conserved quantities also have the opposite sign.[10]

6. WEAK INTERACTIONS

These interactions occur between four fermion fields. Since the range of this interaction is very short, it has been customary to assume a point interaction of the form[22]

$$\mathfrak{L}_I = \sum_i C_i\bar{\psi}^{(1)}O_i\psi^{(2)}\bar{\psi}^{(3)}O_i'\psi^{(4)} + \text{h.c.}, \tag{116}$$

where the operators O_i can be 1, γ_μ, $\sigma_{\mu\nu}$, $\gamma_5\gamma_\mu$, or γ_5. Alternatively, the suggestion has been made that the interaction is mediated by a vector boson. The theory with a point interaction has led to difficulties because it is nonrenormalizable, while the vector boson has not been observed (and it is not clear whether its electromagnetic interactions are renormalizable).

On the other hand, if we use the fermion fields described above, the point interaction becomes renormalizable.[20] The most objectionable feature would be the consideration of the neutrino as a massless particle; it could be given a small mass and then the results should be taken in the limit $m_\nu \to 0$.

Feynman and Gell-Mann[23] have proposed a theory that is a special case of Eq. (116); it is

$$\mathfrak{L}_I = C\bar{\psi}^{(1)}\gamma_\mu a_+\psi^{(2)}\bar{\psi}^{(3)}\gamma_\mu a_+\psi^{(4)}, \tag{117}$$

where

$$a_\pm = \tfrac{1}{2}(1 \pm i\gamma_5). \tag{118}$$

It is introduced by the replacement of the fields $\psi^{(i)}$ by $a_+\psi^{(i)}$. In our theory, the operators $a_\pm$ are projection operators that separate the solutions of the Klein–Gordon equation into two subspaces. But $a_\pm\psi$ is not a solution of the Dirac equation (with either sign) when ψ is, so that the subspaces defined by $P_\pm$ become intermixed by relations of the form

$$a_+\varphi = \varphi' + \chi'. \tag{119}$$

The only case in which this does not happen is that of the field corresponding to massless fermions, as is well known.

The equations of motion obtained from the Lagrangian density (117) in our theory are easily seen not to be invariant under parity transformation or charge conjugation separately, but they are invariant under their product.

7. STRONG INTERACTIONS

The usual interaction Lagrangian density for nucleons and pions is of the form[24]

$$\mathfrak{L}_I' = g\bar{\psi}\gamma_5\psi\phi, \tag{120}$$

where ψ is an ordinary Dirac field and ϕ a pseudoscalar field, so that the equation of motion for ψ is

$$(-i\hat{\partial} + m - g\gamma_5\phi)\psi = 0. \tag{121}$$

Operating on this equation with $(i\hat{\partial} + m + g\gamma_5\phi)$ we obtain

$$(\Box + m^2 - g^2\phi^2)\psi = 0, \tag{122}$$

where the pseudoscalar nature of the field ϕ is expressed by the minus sign of the third term and the absence of a first degree term in ϕ brought in by a scalar coupling. The obvious choice for an interaction Lagrangian density that would give Eq. (122) is

$$\mathfrak{L}_I = g^2\phi^2\bar{\psi}\psi; \tag{123}$$

it would give vertex diagrams with two meson lines, which should give essentially the same results as the original theory.

There is a second possibility, which is to keep $\mathfrak{L}_I'$ for our ψ fields; this should give interesting variations on calculations of transition amplitudes.

In general, we can say that the old equations of interacting fields can be "squared" to bring them into a form similar to the Klein–Gordon equation, but the modified propagator for the fermion fields allows a number of new interactions which are also renormalizable.

8. OPEN QUESTIONS

What we have done so far is to sketch a new approach to spinor fields. We have not gone far enough to come to any conclusions, so instead we briefly point out several considerations that indicate possible areas of further research.

[22] Ref. 18, Chap. 10, p. 295.

[23] R. P. Feynman and M. Gell-Mann, Phys. Rev. **109**, 193 (1958). The same proposal was made by E. C. G. Sudarshan and R. E. Marshak, *Proceedings of the Padua–Venice International Conference on Mesons and Recently Discovered Particles, 1957* Societá (Italiana di Fisica, Padova, 1958); Phys. Rev. **109**, 1860 (1958); and J. J. Sakurai, Nuovo Cimento **7**, 649 (1958).

[24] See, for instance, Ref. 18, Chap. 10, p. 283.

A completely satisfactory interpretation of the quantum mechanics of spin $\frac{1}{2}$ particles has not been achieved, and there are indications that this might not be possible. For the same reasons, the possible usefulness of the theory of a classical spinor field as an approximation to the quantized theory (in the sense the classical electromagnetic field and even the meson field[25] are useful) is not apparent.

The general problem of the separation of the total angular momentum of a field into orbital and spin parts has also not received a satisfactory treatment in the literature, in our opinion. The interpretation of the space–time components of the tensor should also be examined in greater detail.

We have seen that, both in the present theory and in that of Lorentz vector fields[10] (or the electromagnetic field), a part of the field appears in conserved quantities and commutation relations with the "wrong" sign, leading to difficulties such as an indefinite expression for the total energy or inconsistencies that are solved by an indefinite metric in Hilbert space. The origin of these problems can be traced back to the requirement that the Lagrangian density be a Lorentz scalar, and the indefinite nature of the Lorentz metric.

As far as quantum electrodynamics is concerned, either a general proof of equivalence of this theory with the usual one should be found, or transition amplitudes for the different processes and divergent diagrams should be recalculated to find possible changes in the results.

The peculiarities of the limit $m \to 0$ require further attention, and a search for a better formulation is clearly necessary. If this problem is satisfactorily solved, it becomes possible to calculate higher order corrections to the transition amplitudes in weak interactions.

It is also possible to write down a number of new interaction terms for strongly interacting particles, and the corresponding Feynman diagrams can be evaluated.

In general terms, we can say that a further study of the ramifications of this theory is required to reach a better understanding of the significance of the new features it presents.

[25] D. Iwanenko and A. Sokolow, *Klassische Feldtheorie* (Akademie-Verlag, Berlin, 1953).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Connection between Complex Angular Momenta and the Inverse Scattering Problem at Fixed Energy

R. G. NEWTON*
Indiana University, Bloomington, Indiana

(Received 21 November 1966)

The paper contains a number of remarks occasioned by Sabatier's enlargement of the class of solutions to the inverse scattering problem at fixed energy found previously by the author. The implications of the large class of scattering-equivalent potentials for the angular momentum interpolation problem are discussed. The additional angular momenta that appear in the expansion of the potential are directly related to the singularities of its Mellin transform. It is shown that the expansion coefficients must not converge too rapidly to zero unless the first moment of the potential vanishes. Finally we analyze the information contained in an "angular momentum dispersion relation" obeyed by the Jost function which is found as a by-product.

I. INTRODUCTION

THE "inverse scattering problem at fixed energy" is the problem of finding the local potentials which, when inserted in the Schrödinger equation, lead to a prescribed set of phase shifts of all angular momenta. A solution to this problem was given five years ago.[1] In a recent series of papers[2] Sabatier considerably expanded this method and found a very much larger class of solutions. The following remarks are concerned with this general set of scattering-equivalent potentials and the connection of the results with the theory of complex angular momenta. Of particular interest in this connection is an "angular momentum

* Supported in part by the National Science Foundation and by the U.S. Army Research Office—Durham.

[1] R. G. Newton, J. Math. Phys. 3, 75 (1962).

[2] P. C. Sabatier, J. Math. Phys. 7, 1515 and 2079 (1966); 8, 905 (1967).

A completely satisfactory interpretation of the quantum mechanics of spin $\frac{1}{2}$ particles has not been achieved, and there are indications that this might not be possible. For the same reasons, the possible usefulness of the theory of a classical spinor field as an approximation to the quantized theory (in the sense the classical electromagnetic field and even the meson field[25] are useful) is not apparent.

The general problem of the separation of the total angular momentum of a field into orbital and spin parts has also not received a satisfactory treatment in the literature, in our opinion. The interpretation of the space–time components of the tensor should also be examined in greater detail.

We have seen that, both in the present theory and in that of Lorentz vector fields[10] (or the electromagnetic field), a part of the field appears in conserved quantities and commutation relations with the "wrong" sign, leading to difficulties such as an indefinite expression for the total energy or inconsistencies that are solved by an indefinite metric in Hilbert space. The origin of these problems can be traced back to the requirement that the Lagrangian density be a Lorentz scalar, and the indefinite nature of the Lorentz metric.

As far as quantum electrodynamics is concerned, either a general proof of equivalence of this theory with the usual one should be found, or transition amplitudes for the different processes and divergent diagrams should be recalculated to find possible changes in the results.

The peculiarities of the limit $m \to 0$ require further attention, and a search for a better formulation is clearly necessary. If this problem is satisfactorily solved, it becomes possible to calculate higher order corrections to the transition amplitudes in weak interactions.

It is also possible to write down a number of new interaction terms for strongly interacting particles, and the corresponding Feynman diagrams can be evaluated.

In general terms, we can say that a further study of the ramifications of this theory is required to reach a better understanding of the significance of the new features it presents.

[25] D. Iwanenko and A. Sokolow, *Klassische Feldtheorie* (Akademie-Verlag, Berlin, 1953).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Connection between Complex Angular Momenta and the Inverse Scattering Problem at Fixed Energy

R. G. NEWTON*
Indiana University, Bloomington, Indiana

(Received 21 November 1966)

The paper contains a number of remarks occasioned by Sabatier's enlargement of the class of solutions to the inverse scattering problem at fixed energy found previously by the author. The implications of the large class of scattering-equivalent potentials for the angular momentum interpolation problem are discussed. The additional angular momenta that appear in the expansion of the potential are directly related to the singularities of its Mellin transform. It is shown that the expansion coefficients must not converge too rapidly to zero unless the first moment of the potential vanishes. Finally we analyze the information contained in an "angular momentum dispersion relation" obeyed by the Jost function which is found as a by-product.

I. INTRODUCTION

THE "inverse scattering problem at fixed energy" is the problem of finding the local potentials which, when inserted in the Schrödinger equation, lead to a prescribed set of phase shifts of all angular momenta. A solution to this problem was given five years ago.[1] In a recent series of papers[2] Sabatier considerably expanded this method and found a very much larger class of solutions. The following remarks are concerned with this general set of scattering-equivalent potentials and the connection of the results with the theory of complex angular momenta. Of particular interest in this connection is an "angular momentum

* Supported in part by the National Science Foundation and by the U.S. Army Research Office—Durham.

[1] R. G. Newton, J. Math. Phys. 3, 75 (1962).

[2] P. C. Sabatier, J. Math. Phys. 7, 1515 and 2079 (1966); 8, 905 (1967).

dispersion relation" obeyed by the Jost function. It is discussed in Sec. IV.

II. INVERSION AND THE MELLIN TRANSFORM OF THE POTENTIAL

If we define[3]

$$f(x, x') = \sum_l c_l u_l(x) u_l(x'), \tag{1}$$

where $u_l(x) = (\frac{1}{2}\pi x)^{\frac{1}{2}} J_\lambda(x)$, $\lambda = l + \frac{1}{2}$, and letting $K(x, x')$, $x' \leq x$, for fixed x, be the unique solution of the integral equation

$$K(x, x') = f(x, x') - \int_0^x dx'' x''^{-2} K(x, x'') f(x'', x'), \tag{2}$$

then the functions

$$\varphi_l(x) = u_l(x) - \int_0^x dx' x'^{-2} K(x, x') u_l(x') \tag{3}$$

are the regular solutions of the Schrödinger equations

$$\left[-\frac{d^2}{dx^2} + \frac{l(l+1)}{x^2} + V(x)\right] \varphi_l(x) = \varphi_l(x) \tag{4}$$

with the potential

$$V(x) = -2x^{-1}(d/dx)[x^{-1}K(x, x)], \tag{5}$$

which leads to the expression

$$-\tfrac{1}{2}x \int_0^x dx' x' V(x') = \sum_l c_l \varphi_l(x) u_l(x). \tag{6}$$

Furthermore, the Jost functions satisfy the following set of equations[4]:

$$f_l = 1 + \tfrac{1}{2} i \sum_{l'} \frac{e^{i\pi(l'-l)} - 1}{(l' - l)(l' + l + 1)} c_{l'} f_{l'}. \tag{7}$$

Since $f_l = |f_l| e^{-i\delta_l}$, where δ_l is the phase shift of angular momentum l, the imaginary part of (7) multiplied by $e^{i\delta_l}$ reads

$$\sin \delta_l = \sum_{l'} \Gamma^{(1)}_{ll'} c_{l'} |f_{l'}| \tag{8}$$

with

$$\Gamma^{(1)}_{ll'} = \frac{\sin^2 \frac{1}{2}\pi(l' - l) \cos(\delta_l - \delta_{l'}) + \frac{1}{2} \sin \pi(l' - l) \sin(\delta_l - \delta_{l'})}{(l' - l)(l' + l + 1)},$$

and the real part

$$|f_l| = \cos \delta_l + \sum_{l'} \Gamma^{(2)}_{ll'} c_{l'} |f_{l'}| \tag{9}$$

with

$$\Gamma^{(2)}_{ll'} = \frac{\sin^2 \frac{1}{2}\pi(l' - l) \sin(\delta_{l'} - \delta_l) - \frac{1}{2} \sin \pi(l' - l) \cos(\delta_l - \delta_{l'})}{(l' - l)(l' + l + 1)}.$$

In Ref. 1 we allowed l everywhere to assume integral values only. For given phase shifts δ_l, the infinite set of equations (8) was then to be solved for the products $c_{l'} |f_{l'}|$; these were to be inserted in (9), which then gave us the $|f_l|$, and therefore the c_l.[5]

It was pointed out by Sabatier[2] that allowing l to take on integral values only is an unnecessary restriction on the potential. Allowing other values of l leads to a very much larger class of potentials for the same set of phase shifts. This is because for nonintegral l values we may choose the c_l arbitrarily, and still solve Eqs. (8) and (9) for the c_l (l integer) and the phase shifts δ_l (l noninteger). This arbitrariness has important implications for the "interpolation problem" of the physical phase shifts by analytic functions.

The fact is that within the framework of local potentials, we can choose the interpolation of the physical phase shifts largely arbitrarily, except for the restriction that[6,7]

$$d\delta_l/dl < \tfrac{1}{2}\pi. \tag{10}$$

We may choose a discrete set of points l_i of arbitrary density, and choose arbitrary values of δ_l there, keeping the δ_l at integral l fixed, except that we must have

$$\delta_{l_i} - \delta_{l_{i-1}} < \tfrac{1}{2}\pi(l_i - l_{i-1}) \tag{10a}$$

if $l_1 < l_2 < \cdots$. Equations (8) and (9) can then be solved for the c_l and the local potential that produces these phase shifts can be constructed. What is more, we may even choose the phase shifts δ_l as an arbitrary smooth curve satisfying (10) with prescribed values at the integral l. Then the summations in (8)–(9) all become integrals, but we can still find local potentials that produce this phase shift curve. Therefore, as long

[3] Since the energy is fixed in the following, it is convenient to use the variable $x = kr$.

[4] I am using here the same definition of the Jost function as in R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Company, Inc., New York, 1966). For real k and l it is the complex conjugate of the more customary definition, so that its phase is the negative of the phase shift. Equation (7) corrects an error in Eq. (20.62a) on p. 628 of the above-mentioned work.

[5] Sabatier, Ref. 2, has shown that a sufficient condition for this procedure to lead to convergent results is that $c_l = O(l^{\frac{1}{3}})$ as $l \to \infty$.

[6] T. Regge, Nuovo Cimento **14**, 951 (1959); see also Ref. 7, Appendix B.

[7] R. G. Newton, *The Complex j-Plane* (W. A. Benjamin, Inc., New York, 1964).

as we only demand existence of a local potential, the physical phase shifts can be interpolated quite arbitrarily, except for smoothness and (10).[8]

The question naturally arises: What is the significance of the particular l values that enter in the summation (1)? In other words, how can we determine the set Λ such that

$$-\tfrac{1}{2}x\int_0^x dx'x'V(x') = \sum_\Lambda \gamma_\lambda \varphi_{\lambda-\frac{1}{2}}(x)u_{\lambda-\frac{1}{2}}(x), \quad (6a)$$

where $\gamma_\lambda = c_{\lambda-\frac{1}{2}}$? (If Λ contains a continuum then the sum becomes at least partly an integral.) The answer is contained in (7), which reads, with $f(\lambda) = f_{\lambda-\frac{1}{2}}$,

$$f(\lambda) = 1 + \tfrac{1}{2}i\sum_{\lambda'\epsilon\Lambda}\frac{e^{i\pi(\lambda'-\lambda)}-1}{\lambda'^2-\lambda^2}\gamma_{\lambda'}f(\lambda'), \quad (7a)$$

if the potential has the expansion (6a). If the $f_{\lambda'-\frac{1}{2}}$ are solutions of (7), then (7a) defines an analytic function $f(\lambda)$ whose values, when λ assumes values in Λ, are just the numbers $f_{\lambda'-\frac{1}{2}}$ that appear on the right. What is more, this analytic interpolation of f is the "right" one. That is because, for any λ with $\mathrm{Re}\,\lambda > 0$, Eq. (3) defines the regular solution of (4) with the potential (5), and (7a) is obtained from its asymptotic behavior. Therefore the interpolation defined by (7a) is the same as that defined by the "dynamical" interpolation of (4).[9]

Now (7a) tells us that $f(\lambda)$ has singularities exactly when $-\lambda$ takes on values in Λ, except at positive integers, if they occur in Λ.[10] Equation (7a) shows that at the integers

$$f(-n)/f(n) = 1 + (\pi/2n)\gamma_n, \quad n = 1, 2, \cdots. \quad (11)$$

The set Λ is identical with the set of points where $f(-\lambda)$ is singular, augmented by the positive integers. Without loss of generality we may assume that Λ contains *all* the positive integers. In specific instances, of course, some, or all, of the $\gamma_{\lambda'}$ at integral λ' may vanish.

Now the set of singularities of $f(\lambda)$ has been explicitly connected by Froissart[11] to the set of singularities of the Mellin transform of the potential. That situation is as follows.

Define

$$v(\sigma) = \int_0^1 dx x^{\sigma-1}V(x) \quad (12)$$

as an analytic function of σ and continue it to the region $\mathrm{Re}\,\sigma < 0$. Let $\mathcal{S}$ be the set of singular points of $v(\sigma)$. Then $f(\lambda)$ can have singularities only at the points

$$\lambda = 1 - n + \tfrac{1}{2}m(\mathcal{S} - 2), \quad m, n = 1, 2, \cdots \quad (13)$$

if the symbol $m(\mathcal{S}-2)$ denotes the set $\sigma_1 + \sigma_2 + \cdots + \sigma_m - 2m$, where $\sigma_1, \cdots, \sigma_m$ are any m points of $\mathcal{S}$ (not necessarily all different); except that in general $f(\lambda)$ will *not* have poles at the negative integral values of λ.[12] The set Λ therefore consists of the negatives of the numbers described in and below (13), plus the positive integers.[13] Note that Λ may contain complex numbers. But, according to (12), the reality of the potential implies that if λ is in Λ, then so is λ^*. Furthermore, (6) shows that, because for a real potential $\varphi_{l^*} = \varphi_l^*$, we must have

$$c_{l^*} = c_l^*, \quad \text{i.e.,} \quad \gamma_{\lambda^*} = \gamma_\lambda^*. \quad (14)$$

The coefficients γ_λ, corresponding to real points in Λ, must be real.

Let us return now to the interpolation problem. If we are given a set of physical phase shifts, we must in general assume that Λ contains the half-integral values of λ (integral l). Otherwise we have no way of guaranteeing that the phase of $f(\lambda)$ at $\lambda = l + \frac{1}{2}$ (integral l) has the prescribed value.[14] In the course of solving for the constants c_l it may then turn out that some (or all) of them, for integral l, are zero, so that, in effect some, or all, of the half-integral λ values are not in Λ. But

[8] It should be added that, except for Sabatier's general results on the asymptotic behavior of the potentials constructed via (1)–(5), we have no guaranty that any particular one is "well behaved" in any sense of the term. That situation is quite analogous to the one in the inverse scattering problem at fixed angular momentum.

[9] If the potential is well enough behaved, then we know that this "dynamical" interpolation is the unique one in the sense of Carlson's theorem (see, for example, Ref. 7, Chap. 15). If it is not, then the "dynamical" interpolation may not satisfy the hypotheses of Carlson's theorem. In that case there is the possibility that another interpolation exists which does, and which is therefore the "right" one. (No such cases are actually known.) So all we can assert is that the interpolation of the Jost function given by (7a) is identical with the "dynamical" one. This is sufficient for the purpose of identifying Λ.

[10] $f(\lambda)$ can be made singular at $\lambda = -n$ by letting two terms on the right of (7a) coincide while their coefficients approach infinity. In that case (1) and (6) will contain derivatives of the u_l and φ_l with respect to l.

[11] M. Froissart, J. Math. Phys. 3, 922 (1962).

[12] This point was not mentioned in Ref. 11, but it follows from the presence of the gamma function in the denominator of the Bessel function [Eq. (16) of Ref. 11]. The function $f(\lambda)$ has a singularity at a negative integer only if that value of λ in (13) comes from a point in $\mathcal{S}$ that is *not* a simple pole of (12).

[13] Strictly speaking, Λ is a *subset* of the set Λ' that consists of the negatives of the numbers given in (13), plus the positive integers. But the members of Λ' that are not in Λ may be thought of as absent "accidentally." We may handle that situation most conveniently by making Λ equal to Λ' and then setting $\gamma_\lambda = 0$ at those values of λ that are "accidentally" missing.

[14] Sabatier showed in Ref. 2 that if the physical phase shifts are sufficiently small, then there exists a potential for which Λ contains only integers. (This potential is an even function of r.) The Jost function f then contains no singularities in the λ plane. However, it is not known what "sufficiently small" means. [One may conjecture that it has something to do with the violation of (10a) by the phase shifts for half-integral l.] Hence for a prescribed set of physical phase shifts, one cannot assume the existence of an associated even potential.

that cannot be anticipated. However, the remaining members of Λ are at our disposal. This means that, even though the physical phase shifts are given, we are free to prescribe arbitrarily the nonintegral singularities of the Mellin transform (12) of the potential. [According to (13), the singularities of (12) at negative integers give rise to singularities of $f(\lambda)$ at integral and half-integral values of $-\lambda$.] The freedom to choose singularity points of (12) means freedom to choose the exact way the potential behaves at $r \to 0$. For example, a term r^α produces a simple pole of $v(\sigma)$ at $\sigma = -\alpha$. Such a question is independent of bounds on the potential or of whether a term of that nature dominates at small distances. Specifically, a given set of physical phase shifts may always be fitted by a potential that is analytic in a neighborhood of the origin. Such a potential produces simple poles in the Mellin transform (12) at negative integral σ, and no other singularities. Hence Λ then contains integral and half-integral λ only. A special class among these potentials are those for which $c_l = 0$ for all half-integral l values (integral λ). They are the ones discussed in Ref. 1. But the class of analytic potentials that produce a given set of physical phase shifts is enormously much larger. Since all the phase shifts for integral λ are free, this class contains a denumerable infinity of parameters.[15]

III. SLOW DECREASE OF THE c_l

We cannot in general assume that the coefficients c_l vanish rapidly as $l \to \infty$. Such an assumption puts a more severe restriction on the potential than is usually physically acceptable. That can be seen as follows.

Multiply the Schrödinger equation (4) by $u_l(x)$, subtract the same equation with $V = 0$ for $u_l(x)$ multiplied by $\varphi_l(x)$, and integrate from zero to infinity. The result is

$$\int_0^\infty dx \varphi_l(x) u_l(x) V(x) = \operatorname{Im} f_l. \tag{15}$$

Now multiply (15) by c_l and sum over l. Use of (5) then shows that

$$\left[\int_0^\infty dx x V(x)\right]^2 = -4 \sum_l c_l \operatorname{Im} f_l. \tag{16}$$

But multiplying (8) by $c_l |f_l|$ and summing shows that

$$-\sum_l c_l \operatorname{Im} f_l = \sum_l \sum_{l'} \Gamma^{(1)}_{ll'} c_l c_{l'} |f_l f_{l'}|. \tag{17}$$

If the c_l vanish so rapidly as $l \to \infty$ that inversion of the order of summation is allowed in (17), then the right-hand side must be zero because $\Gamma_{ll'}$ is antisymmetric under the interchange of l and l'. Consequently (16) says that

$$\int_0^\infty dx x V(x) = 0. \tag{18}$$

The same conclusion follows from (6) if we assume that as $l \to \infty$ the c_l vanish sufficiently fast that the asymptotic value of the right-hand side of (6) is given by

$$\sum_l c_l |f_l| \sin(x - \tfrac{1}{2}\pi l) \sin(x - \tfrac{1}{2}\pi l - \delta_l)$$

and this is bounded. Then (6) also implies (18).

We must therefore conclude that if the potential does not satisfy (18) (for example, if it has everywhere the same sign), then c_l can at best vanish rather slowly as $l \to \infty$. It cannot even be absolutely summable:

$$\sum_l |c_l| = \infty$$

follows from a violation of (18).

IV. ANGULAR MOMENTUM DISPERSION RELATION

Finally we want to discuss Eq. (7a), which can be considered an *angular momentum dispersion relation* satisfied by the Jost function. If Λ contains discrete points only, (7a) is equivalent to a Mittag-Leffler expansion with a remainder.[16] In order to see how much information it contains, we may try to derive it without any of the machinery contained in Eqs. (1)–(5). That can be done as follows.

The most essential ingredient to be used is the symmetry of the Jost function,[17]

$$f(\lambda) f^*(-\lambda^*) e^{i\pi\lambda} - f(-\lambda) f^*(\lambda^*) e^{-i\pi\lambda} = 2i \sin \pi\lambda, \tag{19}$$

which holds whenever the potential is local and real and well-enough behaved to permit the definition of f. Let us define two functions

$$g_1(\lambda^2) \equiv \tfrac{1}{2}(1 - i \cot \pi\lambda) f(\lambda) + \tfrac{1}{2}(1 + i \cot \pi\lambda) f(-\lambda),$$
$$g_2(\lambda^2) \equiv \tfrac{1}{2} i [f(\lambda) - f(-\lambda)] / \sin \pi\lambda, \tag{20}$$

so that

$$f(\lambda) = g_1(\lambda^2) + e^{-i\pi\lambda} g_2(\lambda^2). \tag{21}$$

Insertion in (19) then leads to

$$g_1(\lambda^2) g_1^*(\lambda^{2*}) - g_2(\lambda^2) g_2^*(\lambda^{2*}) = 1, \tag{22}$$

[15] This was already pointed out by P. C. Sabatier, Ref. 2.

[16] I am indebted to Dr. Sabatier for correspondence on this point.

[17] See, for example, Ref. 7, p. 32, Eq. (5-11). However, note Ref. 4.

which means particularly on the real and imaginary axes,

$$|g_1|^2 - |g_2|^2 = 1. \tag{22a}$$

The other fact to be used is the behavior of the Jost function[18]

$$f(\lambda) \to 1 \quad \text{as} \quad \nu \equiv \operatorname{Im} \lambda \to -\infty. \tag{23}$$

As $\nu \to +\infty$, $e^{i\pi\lambda} f(\lambda)$ is bounded by an inverse power of ν.[18] These statements are most reasonably interpreted by demanding

$$\left.\begin{aligned} g_1(\lambda^2) &\to 1 \\ g_2(\lambda^2) &\to 0 \end{aligned}\right\} \text{ as } \nu \to \pm\infty. \tag{24}$$

Of course, they do not imply (24). But let us now make the additional assumption that g_1 and g_2 possess Mittag–Leffler expansions of the form[19]

$$\begin{aligned} g_1(\lambda^2) &= 1 + \sum_{\Lambda} \frac{a_{\lambda'}}{\lambda'^2 - \lambda^2}, \\ g_2(\lambda^2) &= \sum_{\Lambda} \frac{b_{\lambda'}}{\lambda'^2 - \lambda^2}, \end{aligned} \tag{25}$$

which incorporate (24).[20] This is the only additional assumption we are going to make. Of course, it does not follow from (21) and (23), nor at present are any sufficient conditions on the potential known which would allow us to prove (25) directly. We therefore assume these equations.

Now, according to (21), the singularities of g_1 and g_2 are singularities of f, except that if

$$b_{\lambda'} = -a_{\lambda'} e^{i\pi\lambda'}, \tag{26}$$

then $f(\lambda)$ has no pole at $\lambda = \lambda'$. Therefore, if we demand that $f(\lambda)$ have no poles in the right half-plane, as we must for well-behaved potentials,[21] then (26) must hold for all λ' in Λ. Furthermore, if λ' is in Λ, then so must be λ'^*. That follows from (19) (or from the reality of the potential). Now let us insert (26) and (25) in (22). Then we find, after a bit of algebra, that

$$a_{\lambda'} = -\tfrac{1}{2} i \gamma_{\lambda'} f(\lambda'),$$

where the constants $\gamma_{\lambda'}$ must have the property (14). This means that we have derived (7a).

The reasoning may, of course, be easily inverted. The result is that the "dispersion relation" (7a) is equivalent to the symmetry (19), together with the absence of singularities in f in the right half-plane, *and* the expansions (25). The extent to which the latter incorporate (23) or other statements about the asymptotic behavior of $f(\lambda)$ depends on the distribution of singularities of $f(\lambda)$, i.e., of g_1 and g_2. And that, in turn, depends on the detailed behavior of the potential near $r = 0$, as expressed in the singularities of (12). For example, we cannot conclude from (7a) that $f(\lambda) \to 1$ as $\operatorname{Re} \lambda \to -\infty$, with $\operatorname{Im} \lambda \neq 0$ fixed, unless we know that the singularities of f do not extend infinitely far away from the real axis.

The integers in Λ again require a brief separate discussion. If g_1 and g_2 have simple poles at an integral value of λ, then f is analytic there. Equations (20) show that these poles come from the $\sin \pi\lambda$ in the denominators. In fact, unless $f(n) = f(-n)$, g_1 and g_2 *must* have poles at $\pm n$. According to (11), the exceptional case $f(n) = f(-n)$ does imply $\gamma_n = 0$.[22]

[18] See, for example, Ref. 7, Chap. 6.

[19] There is no loss in generality in summing both over the same set Λ. Λ is the union of the sets of singularities of g_1 and g_2. Some of the coefficients $a_{\lambda'}$ or $b_{\lambda'}$ may, *a priori*, be zero.

[20] Provided that Λ does not extend to infinity in the imaginary directions.

[21] See, for example, Ref. 7, Chap. 5.

[22] In view of the fact that (7a) implies the existence of a real local potential, it must imply another important property of $f(\lambda)$: It can have no zeros in the fourth quadrant (see, for example, Ref. 7, p. 51). However, I have been unable to prove this fact directly from (7a).

Matrix Elements for Irreducible Representations of the $U(6, 6)$ Algebra in Harmonic Function Space

AWELE MADUEMEZIA*
International Center for Theoretical Physics, International Atomic Energy Agency, Trieste, Italy

(Received 29 November 1966)

In this paper we realize the algebra of the group $U(6, 6)$ as an algebra of differential operators acting in the Hilbert space of functions defined on a 23-dimensional pseudosphere. We then calculate the matrix elements of the generators of this algebra between certain harmonic function states.

1. INTRODUCTION

In a complex n-dimensional differentiable manifold C^n, the harmonic functions are defined as the eigenvectors of the so-called Laplace–Beltrami operator[1]

$$\Delta(C^n) = \frac{1}{(|\bar{g}|)^{\frac{1}{2}}} \frac{\partial}{\partial x^i} (|\bar{g}|)^{\frac{1}{2}} g^{ij} \frac{\partial}{\partial x^j} \tag{1.1}$$

$$\bar{g} \equiv \det (g_{ij}); \quad g^{ij} = (g^{-1})_{ij}. \tag{1.2}$$

The metric is defined by the line element

$$ds^2 = g_{ij}\, dx^i\, dx^j; \quad x \in C^n. \tag{1.3}$$

If a group of transformations $\mathfrak{G}$ acts transitively on the manifold C^n, i.e., is such that every two ordinary points of C^n are transformable into one another by one or more transformations of the group, then the harmonic functions defined on C^n form a basis for the irreducible representation of the group $\mathfrak{G}$. Bég and Ruegg[2] have used this idea to construct the harmonic functions for the group $SU(3)$. Rączka[3] and Rączka and Fischer[4] have generalized the method of Bég and Ruegg to an arbitrary noncompact group, $U(p, q)$, and have derived the harmonic functions for this chain of groups on homogeneous spaces of unit rank.

In this paper, we particularize the results of Rączka and Fischer[4] to the $U(6, 6)$ group, which is of interest in elementary particle physics, and we study the special features of these particular results. We then realize the algebra of $U(6, 6)$ as an algebra of differential operators, diagonalized with respect to a subset of the set of compact generators. From the set of functions that provide a basis for an irreducible representation of the $U(6, 6)$ algebra, we pick out what are in effect extreme vectors.[5] We then calculate the matrix elements of the generators between any two such vectors.

In order to meet the particular needs of this paper, we have found it necessary to introduce some minor phase changes in some relations in Ref. 4.

2. HARMONIC FUNCTIONS FOR $U(6, 6)$

The homogeneous (carrier) space is the 23-pseudosphere Z^{23},

$$\bar{\psi}\psi \equiv \psi^\dagger \beta \psi = 1, \tag{2.1}$$

embedded in the 12-dimensional complex space C^{12}, of which ψ are vectors, or in the 24-dimensional flat Minkowski space $M^{12,12}$.

Here the matrix β is given by

$$\begin{aligned} \beta = (\beta_{ij}) = (\delta_{ij}) \quad &\text{for} \quad i, j = 1, 2, \cdots, 6 \\ = -(\delta_{ij}) \quad &\text{for} \quad i, j = 7, 8, \cdots, 12. \end{aligned} \tag{2.2}$$

Z^{23} is homeomorphic to the coset spaces

$$U(6, 6)/U(5, 6), \quad U(6, 6)/U(6, 5),$$

on which $U(6, 6)$ acts transitively.[4]

We parametrize the pseudosphere Z^{23} as follows.

We first construct two identical 11-spheres with bases

$$\xi = \{\xi_1, \xi_2, \cdots, \xi_6\}, \tag{2.3}$$

$$\eta = \{\eta_1, \eta_2, \cdots, \eta_6\}, \tag{2.4}$$

respectively, such that

$$\xi_i = e^{i\phi_i} \prod_{k=i+1}^{6} \sin \theta_k \cos \theta_i; \quad i = 1, \cdots, 6, \tag{2.5}$$

$$\eta_i = e^{-i\hat{\phi}_i} \prod_{k=i+1}^{6} \sin \hat{\theta}_k \cos \hat{\theta}_i; \quad i = 1, \cdots, 6, \tag{2.6}$$

where

$$\begin{aligned} &\hat{\theta}_1 = \theta_1 \equiv 0; \quad 0 \leq \theta_i, \quad \hat{\theta}_i \leq \tfrac{1}{2}\pi, \quad i \neq 1; \\ &0 \leq \phi_i, \quad \hat{\phi}_i \leq 2\pi. \end{aligned} \tag{2.7}$$

* On leave of absence from Department of Physics, University of Nigeria, Nsukka, Nigeria. Address after 1 October 1967: Department of Physics, University of Ghana, Legon, Ghana.

[1] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962), p. 387.

[2] M. A. B. Bég and H. Ruegg, J. Math. Phys. 6, 677 (1965).

[3] R. Rączka, ICTP, Trieste, Preprint IC/65/80.

[4] R. Rączka and J. Fischer, Commun. Math. Phys. 3, 233 (1966).

[5] E. B. Dynkin, Am. Math. Soc. Transl. Ser. 2, 6, 327 (1960).

We denote the set $\{\theta_i, \phi_i\}$ by ω and the set $\{\hat{\theta}_i, \hat{\phi}_i\}$ by $\hat{\omega}$. We note that

$$|\xi|^2 = |\eta|^2 = 1. \tag{2.8}$$

Basis vectors of Z^{23} are then realized as

$$\psi_i = \xi_i \cosh\theta, \quad i = 1, \cdots, 6, \tag{2.9}$$

$$\psi_j = \eta_j \sinh\theta, \quad j = 7, 8, \cdots, 12, \quad j = j(\text{mod } 6), \tag{2.10}$$

$$0 \leq \theta \leq \infty. \tag{2.11}$$

It is convenient to denote the set $\{\omega, \hat{\omega}, \theta\}$ by Ω. From Eq. (2.8) and (1.3), which now becomes

$$ds^2 = \sum_i |d\psi_i|^2 - \sum_j |d\psi_j|^2, \tag{2.12}$$

we obtain the determinant of the metric g_{ij}:

$$\tilde{g}^{\frac{1}{2}} = (\cosh\theta \sinh\theta)^{11} \prod_{j=2}^{6} \sin^{2j-3}\theta_j \times \cos\theta_j \prod_{i=2}^{6} \sin^{2i-3}\theta_i \cos\theta_i.$$

This provides us with an invariant measure in Z^{23}, namely,

$$d\Omega = \tilde{g}^{\frac{1}{2}}\, d\phi_1\, d\hat{\phi}_1 \prod_{j=2}^{6} d\theta_j\, d\hat{\theta}_j\, d\phi_j\, d\hat{\phi}_j. \tag{2.13}$$

The Laplace–Beltrami operator reduces to

$$\Delta(Z^{23}) = \frac{1}{\cosh^{11}\theta \sinh^{11}\theta} \frac{\partial}{\partial\theta}\left(\cosh^{11}\theta \sinh^{11}\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{\cosh^2\theta}\Delta(\xi^{11}) - \frac{1}{\sinh^2\theta}\Delta(\eta^{11}), \tag{2.14}$$

where $\Delta(\xi^{11})$ is the Laplace–Beltrami operator for the 11-dimensional sphere on which the compact group $U(6)$ acts transitively. $\Delta(\eta^{11})$ is interpreted similarly. The eigenvalue equation

$$(\Delta(Z^{23}) - \lambda) Y^{\lambda}(\Omega) = 0 \tag{2.15}$$

is completely separable.

If we write

$$Y^{\lambda}(\Omega) = \Psi^{\lambda}(\theta)\Phi(\omega)\hat{\Phi}(\hat{\omega}) \tag{2.16}$$

and demand that $Y^{\lambda}(\Omega)$ also be an eigenvector of the operator

$$M' = \sum_{i=1}^{6} \frac{\partial}{\partial\phi_i} + \sum_{j=1}^{6} \frac{\partial}{\partial\hat{\phi}_j} \tag{2.17}$$

belonging to the center of the algebra, then we find that the eigensolution (2.16), regular at $\theta = 0$, is given by[4]

$$\psi^{\lambda}(\theta) = \frac{1}{(N_{\lambda})^{\frac{1}{2}}} \frac{\tanh^{\hat{J}_6}\theta}{\cosh^{\alpha}\theta} \times {}_2F_1\left(\tfrac{1}{2}(\hat{J}_6 - J_6 + \alpha - 10), \tfrac{1}{2}(\hat{J}_6 + J_6 + \alpha); \hat{J}_6 + 6; \tanh^2\theta\right), \tag{2.18}$$

$$\Phi(\omega) = \frac{1}{(N_{\omega})^{\frac{1}{2}}} \prod_{l=1}^{6} e^{im_l\phi_l} \prod_{k=2}^{6} \sin^{2-k}\theta_k\, d^{j_k}_{a_k,\beta_k}(2\theta_k), \tag{2.19}$$

and

$$\hat{\Phi}(\hat{\omega}) = \frac{1}{(N_{\hat{\omega}})^{\frac{1}{2}}} \prod_{l=1}^{6} e^{-i\hat{m}_l\hat{\phi}_l} \prod_{k=2}^{6} \sin^{2-k}\hat{\theta}_k\, d^{\hat{j}_k}_{\hat{a}_k,\hat{\beta}_k}(2\hat{\theta}_k), \tag{2.20}$$

where m_i, $\hat{m}_j$ are zero or integers, $-J_k(J_k + 2k - 2)$ $(J_k = 0, 1, 2, \cdots)$ are the eigenvalues of the Laplace–Beltrami operator, $\Delta(\xi^{2k-1})$, on a $(2k-1)$-sphere, and $\hat{J}_k$ replaces J_k in η space. We also have

$$j_k = \tfrac{1}{2}(J_k + k - 2), \quad k = 2, 3, \cdots, 6, \quad J_1 = m_1, \tag{2.21}$$

$$\alpha_k = \tfrac{1}{2}(m_k + J_{k-1} + k - 2) \equiv \tfrac{1}{2}(m_k + m_k^*), \tag{2.22}$$

$$\beta_k = \tfrac{1}{2}(m_k - J_{k-1} - k + 2) \equiv \tfrac{1}{2}(m_k - m_k^*), \tag{2.23}$$

$$N_{\omega} = (2\pi)^6 \prod_{k=2}^{6} \frac{1}{J_k + k - 1},$$

$$N_{\hat{\omega}} = (2\pi)^6 \prod_{k=2}^{6} \frac{1}{\hat{J}_k + k - 1},$$

$$\alpha = 11 + (121 - \lambda)^{\frac{1}{2}} > 0. \tag{2.24}$$

N_{λ} is a normalization coefficient for $\psi(\theta)$, and the angles θ, θ_i, $\hat{\theta}_i$, ϕ_i, $\hat{\phi}_i$ have the ranges given by Eqs. (2.7) and (2.11). ${}_2F_1(\alpha, \beta; \gamma; \delta)$ is a hypergeometric function. The d functions are defined in Edmonds[6] in terms of Jacobi polynomials. For a square integrable solution regular at $\theta \to \infty$, we must have

$$\tfrac{1}{2}(\hat{J}_6 - J_6 + \alpha - 10) = -n, \tag{2.25}$$

where n is an arbitrary nonnegative integer. This in effect makes the hypergeometric function a Jacobi polynomial.

We then obtain a discrete spectrum for λ:

$$\lambda = \lambda_n = 121 - (\hat{J}_6 - J_6 + 2n + 1)^2. \tag{2.26}$$

The normalization coefficient for $\psi(\theta)$, with respect to the measure

$$\cosh^{11}\theta \sinh^{11}\theta\, d\theta, \tag{2.27}$$

[6] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957), p. 57.

is readily shown to be given by[4]

$$N_\lambda = \frac{\Gamma(1+n)\Gamma^2(\hat{J}_6+6)\Gamma(J_6-\hat{J}_6-n)}{2(J_6-\hat{J}_6-2n-1)\Gamma(J_6+5-n)\Gamma(\hat{J}_6+6-n)}. \tag{2.28}$$

Let us write $-\hat{J}_6 + J_6 = \Delta$, say. If we assume that $\hat{J}_6$ and J_6 are free to assume any positive integer values, then an examination of the structure of the coefficient N_λ [Eq. (2.28)] shows that a restriction is imposed on the possible values of n, by the following relations:

$$\Delta - n \neq 0, \tag{2.29}$$

$$\Delta - 2n - 1 > 0, \tag{2.30}$$

$$J_6 + 5 - n \neq 0, \tag{2.31}$$

but

$$\Delta - 2n - 1 < 0 \quad \text{at} \quad |\tanh^2\theta| = 1, \tag{2.32}$$

if n is nonintegral. The relation (2.32) is necessary for the absolute convergence of the series for ${}_2F_1$ when n is nonintegral. It therefore does not strictly apply here.

The irreducible representations given by the harmonic functions are characterized by the numbers λ and M:

$$M = \sum_{i=1}^{6} m_i + \sum_{j=1}^{6} \hat{m}_j. \tag{2.33}$$

Suppose we consider the case $\hat{J}_6 = J_6 = J$, say. This is forbidden by (2.30) for any $n > 0$. The case $\Delta = 1$ is also forbidden. For $\Delta = 2$, we have

$$n < \tfrac{1}{2}, \quad \text{i.e.,} \quad n = 0.$$

There is therefore, in this case, only one member of the discrete series labeled by λ_n, and this corresponds to $n = 0$. We then have

$$N_\lambda = 1/2(J+4),$$
$$\lambda = \lambda_0 = 120,$$
$$\alpha = 12,$$

$$\psi^{\lambda_0}(\theta) = [2(J+4)]^{\frac{1}{2}}(\tanh^{J-2}\theta/\cosh^{12}\theta). \tag{2.34}$$

In this paper we study only this case, reserving the more general case for subsequent papers.

We turn now to the functions $\Phi(\omega)$, $\Phi(\hat{\omega})$ defined by Eqs. (2.19) and (2.20). We notice that since the d functions are polynomials, we must have

$$\alpha_k \leq j_k, \tag{2.35}$$

i.e.,

$$m_k \leq J_k - J_{k-1}; \quad k = 2, \cdots, 6, \tag{2.36}$$

$$m_1 \equiv J_1 \geq 0, \tag{2.37}$$

and

$$M_6 \equiv \sum_{i=1}^{6} m_i \leq J_6.$$

We must also have

$$\beta_k \geq -j_k, \tag{2.38}$$

which gives

$$M_6 \geq -J_6.$$

Hence

$$|M_6| \leq J_6. \tag{2.39}$$

Similarly

$$|\hat{M}_6| \leq \hat{J}_6. \tag{2.40}$$

Now, since the operators

$$\frac{\partial}{\partial\hat{\phi}_j}, \quad \frac{\partial}{\partial\phi_i}, \quad i,j = 1, 2, \cdots, 6$$

mutually commute, any 11 linearly independent combinations of them form a basis for the Cartan subalgebra of $U(6,6)$. We can then choose the linear combinations so that the numbers

$$m_k = J_k - J_{k-1}, \quad k = 2, \cdots, 6, \tag{2.41}$$

$$\hat{m}_k = \hat{J}_k - \hat{J}_{k-1}, \tag{2.42}$$

and

$$m_1 \equiv J_1; \quad \hat{m}_1 \equiv \hat{J}_1 \tag{2.43}$$

correspond to the highest weight of the representation.[7] The corresponding eigenfunction is an extreme vector.[5]

We are interested in the matrix elements of the generators of the algebra between any two such extreme vectors. This restriction is justified from the following analogy with the three-dimensional rotation group O_3. For, suppose

$$|jm\rangle, \quad -j \leq m \leq j$$

are the usual basis vectors for an irreducible representation of O_3 in spin space. Let O be any operator-valued function of the operators of O_3. The following equation is true:

$$\langle jm'|\, O\, |jm\rangle = \langle jj|\, (J_+)^{j-m'} O (J_-)^{j-m}\, |jj\rangle, \tag{2.44}$$

where J_+, J_- are operators that raise and lower weights, respectively. By commuting the J_+'s and J_-'s successively with O, eventually making use of the relations

$$0 = \langle jj|\, J_-, \quad J_+\, |jj\rangle = 0, \tag{2.45}$$

we reduce the matrix element between the states $|jm'\rangle, |jm\rangle$ to one between highest weight states $|jj\rangle$.

Our argument then is that the harmonic function states for which

$$m_k = J_k - J_{k-1}$$

would be highest weight states for a proper choice

[7] The existence of such a highest weight does not imply that the representation is finite-dimensional. This is because the m's are given by differences of arbitrary numbers.

of basis in the Cartan subalgebra. With this simplification, we now have

$$\Phi(\omega) = \frac{1}{(2\pi)^3}\prod_{l=1}^{6} e^{im_l\phi_l}\prod_{k=2}^{6} N_k(\cos\theta_k)^{m_k}(\sin\theta_k)^{m_k^*+2-k}, \tag{2.46}$$

where

$$N_k^2 = 2\Gamma(m_k + m_k^* + 2)/\Gamma(m_k + 1)\Gamma(m_k^* + 1). \tag{2.47}$$

3. LIE ALGEBRA

We may write the generators A_{ij} of $U(6, 6)$ as

$$\{A_{\alpha\beta}\} = \begin{pmatrix} A_{ij} & A_{i\hat{j}} \\ A_{\hat{i}j} & A_{\hat{i}\hat{j}} \end{pmatrix}, \quad \begin{matrix} j, i = 1, 2, \cdots, 6, \\ \hat{i}, \hat{j} = 7, 8, \cdots, 12, \end{matrix} \tag{3.1}$$

in which the elements A_{ij} generate the compact subgroup $U(6)$, as also do the elements $A_{\hat{i}\hat{j}}$. The elements $A_{\hat{i}j}$, $A_{i\hat{j}}$ generate the noncompact completion to the group $U(6, 6)$.

The generators by definition satisfy the Hermiticity condition[8]

$$\{A_{\alpha\beta}^\dagger\} = \beta\{A_{\beta\alpha}\}\beta, \tag{3.2}$$

where β is given by Eq. (2.2) for the space C^2. Hence

$$\begin{aligned} A_{ij}^\dagger &= A_{ji}, \qquad A_{\hat{i}\hat{j}}^\dagger = A_{\hat{j}\hat{i}}, \\ A_{i\hat{j}}^\dagger &= -A_{\hat{j}i}, \quad A_{\hat{i}j} = -A_{j\hat{i}}. \end{aligned} \tag{3.3}$$

The generators $A_{\alpha\beta}$ also satisfy the commutation relations

$$[A_{\alpha\beta}, A_{\gamma\delta}] = \beta_{\beta\gamma}A_{\alpha\delta} - \beta_{\delta\alpha}A_{\gamma\beta}. \tag{3.4}$$

We choose the operators $A_{ii}A_{\hat{j}\hat{j}}$ diagonal, with eigenvalues m_i and $\hat{m}_j$, respectively. In terms of differential operators,

$$H_i \equiv A_{ii} = \frac{1}{(-1)^{\frac{1}{2}}}\frac{\partial}{\partial\phi_i} = \psi_i\frac{\partial}{\partial\psi_i} - \psi_i^*\frac{\partial}{\partial\psi_i^*}, \tag{3.5}$$

$$H_j \equiv A_{\hat{j}\hat{j}} = \frac{1}{(-1)^{\frac{1}{2}}}\frac{\partial}{\partial\hat{\phi}_j} = \psi_j\frac{\partial}{\partial\psi_j} - \psi_j^*\frac{\partial}{\partial\psi_j^*}, \tag{3.6}$$

$$A_{mn} = \psi_m\partial/\partial\psi_n - \psi_n^*\partial/\partial\psi_m^*, \tag{3.7}$$

$$A_{\hat{m}\hat{n}} = \psi_{\hat{m}}\partial/\partial\psi_{\hat{n}} - \psi_{\hat{n}}^*\partial/\partial\psi_{\hat{m}}^*, \tag{3.8}$$

$$A_{ij} = \psi_i\partial/\partial\psi_j + \psi_j^*\partial/\partial\psi_i^*, \tag{3.9}$$

$$A_{\hat{i}\hat{j}} = \psi_{\hat{i}}\partial/\partial\psi_{\hat{j}} + \psi_{\hat{j}}^*\partial/\partial\psi_{\hat{i}}^*. \tag{3.10}$$

The relations (3.3) are satisfied because the operators

$$\partial/\partial\psi$$

satisfy the Hermiticity condition

$$(\partial/\partial\psi)^\dagger = -\partial/\partial\psi^*. \tag{3.11}$$

This follows from the invariance under $\partial/\partial\psi$ of the scalar product (f, g) of two functions defined on Z^{23}:

$$(\partial/\partial\psi)(f, g) = 0.$$

Hence

$$\left(f, \frac{\partial}{\partial\psi}g\right) = -\left(\frac{\partial}{\partial\psi^*}f, g\right) \equiv -\left(f, \left(\frac{\partial}{\partial\psi^*}\right)^\dagger g\right).$$

In order to express the operators in terms of θ, θ_i, ϕ_i, $\hat{\theta}_i$, and $\hat{\phi}_i$, we use the following inversions of Eqs. (2.5), (2.6), (2.9), and (2.10):

$$\tan^2\theta_j = \sum_{k=1}^{j-1}\frac{\psi_k\psi_k^*}{\psi_j\psi_j^*}, \tag{3.12}$$

$$\tan^2\hat{\theta}_j = \sum_{k=7}^{j-1}\frac{\psi_{\hat{k}}\psi_{\hat{k}}^*}{\psi_{\hat{j}}\psi_{\hat{j}}^*}, \tag{3.13}$$

$$\cosh^2\theta = \sum_i \psi_i\psi_i^*, \tag{3.14}$$

$$\sinh^2\theta = \sum_j \psi_{\hat{j}}\psi_{\hat{j}}^*, \tag{3.15}$$

$$e^{2i\phi_j} = \psi_j/\psi_j^*, \tag{3.16}$$

$$e^{-2i\hat{\phi}_i} = \psi_{\hat{j}}/\psi_{\hat{j}}^*. \tag{3.17}$$

We thus obtain[9]

$$\begin{aligned} A_{mn} = e^{i(\phi_m-\phi_n)}\Bigg\{ & \frac{\sum_{k=1}^{m-1}\prod_{\substack{i=k+1\\ j=n+1}}^{6}\sin^2\theta_i\cos^2\theta_k\sin\theta_j\cos\theta_n}{2\sin\theta_m\prod_{i=m+1}^{6}\sin^3\theta_i}\cdot\frac{\partial}{\partial\theta_m} - \frac{\sum_{k=1}^{n-1}\prod_{\substack{i=k+1\\ j=m+1}}^{6}\sin^2\theta_i\cos^2\theta_k\sin\theta_j\cos\theta_m}{2\sin\theta_n\prod_{i=n+1}^{6}\sin^3\theta_i}\cdot\frac{\partial}{\partial\theta_n} \\ & + \left(\sum_{j=n+1}^{6} - \sum_{j=m+1}^{6}\right)\left(\frac{\prod_{\substack{i=n+1\\ r=j+1\\ k=m+1}}^{6}\sin\theta_i\sin\theta_r\sin\theta_k\cos\theta_n\cos\theta_j\cos\theta_m}{2\sin\theta_j\prod_{r=j+1}^{6}\sin^3\theta_r}\cdot\frac{\partial}{\partial\theta_j}\right) \\ & + \frac{\prod_{i=m+1}^{6}\sin\theta_i\cos\theta_m}{2i\prod_{i=n+1}^{6}\sin\theta_i\cos\theta_n}\cdot\frac{\partial}{\partial\phi_n} + \frac{\prod_{i=n+1}^{6}\sin\theta_i\cos\theta_n}{2i\prod_{i=m+1}^{6}\sin\theta_i\cos\theta_m}\cdot\frac{\partial}{\partial\phi_m}\Bigg\}. \end{aligned} \tag{3.18}$$

[8] We adopt the following convention with respect to subscripts: (a) Greek subscripts take the values 1, 2, $\cdots$, 12. (b) Latin subscripts *without* a caret take values 1, 2, $\cdots$, 6. (c) Latin subscripts *with* a caret take values 7, 8, $\cdots$, 12. This convention is assumed forthwith *without further mention.*

[9] We note that for any functions A_α $\sum_{\alpha=x}^{y}A_\alpha = 0$ for $y < x$, while $\prod_{\alpha=x}^{y}A_\alpha = 1$ for $y < x$,

We obtain a similar expression for $A_{\hat{m}\hat{n}}$, with the following replacements:

$$\theta_m \to \hat{\theta}_m; \quad \phi \to \hat{\phi}, \quad e^{i\phi_m} \to e^{-i\hat{\phi}_m}, \quad m = 1, 2, \cdots, 6. \tag{3.19}$$

Furthermore,

$$A_{m\hat{n}} = e^{i(\phi_m + \hat{\phi}_n)} \Bigg\{ \prod_{\substack{i=n+1 \\ j=m+1}}^{6} \sin\hat{\theta}_i \cos\hat{\theta}_n \sin\theta_j \cos\theta_m \cdot \frac{\partial}{\partial\theta} + \sum_{j=n+1}^{6} \frac{\prod_{\substack{r=n+1 \\ s=j+1 \\ t=m+1}}^{6} \sin\hat{\theta}_r \sin\hat{\theta}_s \sin\theta_t \cos\hat{\theta}_n \cos\hat{\theta}_j \cos\theta_m}{2\sin\hat{\theta}_j \prod_{r=j+1}^{6} \sin^3\hat{\theta}_r} \cdot \frac{\partial}{\partial\theta_j}$$

$$+ \sum_{j=m+1}^{6} \frac{\prod_{\substack{r=m+1 \\ s=j+1 \\ t=n+1}}^{6} \sin\theta_r \sin\theta_s \sin\hat{\theta}_t \cos\theta_m \cos\theta_j \cos\hat{\theta}_n}{2\sin\theta_j \prod_{i=j+1}^{6} \sin^3\theta_i} \cdot \frac{\partial}{\partial\theta_j} - \frac{\sum_{k=1}^{m-1} \prod_{\substack{i=k+1 \\ j=n+1}}^{6} \sin^2\theta_i \sin\hat{\theta}_j \cos^2\theta_k \cos\hat{\theta}_n}{2\sin\theta_m \prod_{i=m+1}^{6} \sin^3\theta_i} \cdot \frac{\partial}{\partial\theta_m}$$

$$- \frac{\sum_{k=1}^{n-1} \prod_{\substack{i=k+1 \\ j=m+1}}^{6} \sin^2\hat{\theta}_i \sin\theta_j \cos^2\hat{\theta}_k \cos\theta_m}{2\sin\hat{\theta}_n \prod_{i=n+1}^{6} \sin^3\hat{\theta}_i} \cdot \frac{\partial}{\partial\theta_n} + \frac{\prod_{i=m+1}^{6} \sin\theta_i \cos\theta_m}{2i \prod_{i=n+1}^{6} \sin\hat{\theta}_i \cos\hat{\theta}_n} \cdot \frac{\partial}{\partial\hat{\phi}_n} - \frac{\prod_{i=n+1}^{6} \sin\hat{\theta}_i \cos\hat{\theta}_n}{2i \prod_{i=m+1}^{6} \sin\theta_i \cos\theta_m} \cdot \frac{\partial}{\partial\phi_m} \Bigg\}. \tag{3.20}$$

We obtain $A_{\hat{m}n}$ from Eq. (3.20) simply by putting carets where there were no carets, and removing them where there were, i.e., by the interchanges

$$\hat{\theta}_m \leftrightarrow \theta_m, \quad \hat{\phi}_m \leftrightarrow \phi_m, \quad m = 1, \cdots, 6. \tag{3.21}$$

4. MATRIX ELEMENTS OF THE GENERATORS OF $U(6, 6)$ ALGEBRA

The problem we have to solve is the evaluation of the integrals

$$D(\alpha, \beta) \equiv D^{\lambda_0 J M}_{m,m';\hat{m},\hat{m}'}(A_{\alpha\beta}),$$

$$\equiv \int d\Omega (Y^{\lambda_0 J M}_{m',\hat{m}'}(\Omega))^* A_{\alpha\beta} Y^{\lambda_0 J M}_{m,\hat{m}}(\Omega) \tag{4.1}$$

for $\alpha, \beta = 1, 2, \cdots, 12$, where $d\Omega$ is given by Eq. (2.13) and

$$\begin{aligned} J &= J_6 = \hat{J}_6 + 2 = J'_6 = \hat{J}'_6 + 2, \\ m_k &= J_k - J_{k-1}; \quad k = 2, 3, \cdots, 6, \\ m_1 &= J_1, \end{aligned} \tag{4.2}$$

$$M = \sum_i m_i + \sum_j \hat{m}_j. \tag{4.3}$$

It is convenient at this point to introduce the notation

$$\begin{aligned} \mu_i &\equiv m_i + m'_i, \quad \nu_i \equiv m^*_i + m^{*\prime}_i, \\ \hat{\mu}_i &\equiv \hat{m}_i + \hat{m}'_i, \quad \hat{\nu}_i \equiv \hat{m}^*_i + \hat{m}^{*\prime}_i. \end{aligned} \tag{4.4}$$

We proceed now to evaluate $D(m, n)$, $D(\hat{m}, \hat{n})$, $D(m\hat{n})$, and $D(\hat{m}, n)$ in that order. The matrices $D(m, m)$, $D(\hat{m}, \hat{m})$ are trivially given by

$$m_m \delta_{MM'}, \quad \hat{m}_m \delta_{MM'}, \tag{4.5}$$

respectively.

For $D(m, n)$, we substitute Eq. (3.18) in (4.1) and carry out the differentiations, making use of the relations

$$\frac{\partial}{\partial\theta_j} \prod_{k=2}^{6} (\cos\theta_k)^{m_k} (\sin\theta_k)^{m_k^* + 2 - k}$$

$$= (-m_j \tan\theta_j + J_{j-1} \cot\theta_j)$$

$$\times \prod_{k=2}^{6} (\cos\theta_k)^{m_k} (\sin\theta_k)^{m_k + 2 - k}, \tag{4.6}$$

$$\sum_{l=1}^{m-1} \prod_{j=l+1}^{6} \sin^2\theta_j \cos^2\theta_l = (1 - \delta_{m1}) \prod_{k=m}^{6} \sin^2\theta_k, \qquad \theta_1 \equiv 0, \tag{4.7}$$

$$\Omega_k \equiv N_k N'_k$$

$$= \left[\frac{\Gamma(m_k + m_k^* + 2)\Gamma(m'_k + m_k^{*\prime} + 2)}{\Gamma(m_k + 1)\Gamma(m'_k + 1)\Gamma(m_k^* + 1)\Gamma(m_k^{*\prime} + 1)} \right]^{\frac{1}{2}}. \tag{4.8}$$

We then obtain:

For $m > n$,

$$D(m,n) = \delta_{m_m+1,m_m'}\delta_{m_n-1,m_n'}\prod_{k=1}^{6}{}^{m,n}\,\delta_{m_k,m_k'}\prod_{i=1}^{6}\delta_{m_i,\hat{m}_i'}\int_0^{\pi/2}\prod_{k=2}^{6}\Omega_k\,d\theta_k(\cos\theta_k)^{\mu_k+1}(\sin\theta_k)^{\nu_k+1}$$

$$\times\Bigg\{\tfrac{1}{2}\sin\theta_m\cos\theta_n(J_{m-1}\cot\theta_m - m_m\tan\theta_m)(1-\delta_{m1})\prod_{i=n+1}^{m}\sin\theta_i + \tfrac{1}{2}\cos\theta_m\cos\theta_n$$

$$\times\left(J_n\prod_{i=n+1}^{m}\sin^{-1}\theta_i - J_m\prod_{r=n+1}^{m}\sin\theta_r\right) - \frac{1}{2}\frac{\sin\theta_n\cos\theta_m}{\prod_{i=n+1}^{m}\sin\theta_i}(J_{n-1}\cot\theta_n - m_n\tan\theta_n)(1-\delta_{n1})$$

$$+\frac{m_n\cos\theta_m}{2\prod_{i=n+1}^{m}\sin\theta_i\cos\theta_n} + \frac{m_m\prod_{i=n+1}^{m}\sin\theta_i\cos\theta_n}{2\cos\theta_m}\Bigg\}, \tag{4.9}$$

where $\prod_{k=1}^{6}{}^{m,n}$ indicates that the terms $k = m, n$ are omitted from the product. This product of 12 delta functions occurs often in this work, and so we denote it by a new symbol. We define

$$\delta_{MM'}^{m^+n^-} \equiv \delta_{m_m+1,m_n'}\delta_{m_n-1\,m_n'}\prod_{k=1}^{6}{}^{m,n}\delta_{m_k,m_k'}\prod_{i=1}^{6}\delta_{\hat{m}_i,\hat{m}_i'}. \tag{4.10}$$

For $m < n$,

$$D(m,n) = \delta_{MM'}^{m^+n^-}\int_0^{\pi/2}\prod_{k=2}^{6}\Omega_k\,d\theta_k(\cos\theta_k)^{\mu_k+1}(\sin\theta_k)^{\nu_k+1}$$

$$\times\Bigg\{\frac{\sin\theta_m\cos\theta_n}{2\prod_{i=m+1}^{n}\sin\theta_i}(J_{m-1}\cos\theta_m - m_m\tan\theta_m)(1-\delta_{m1}) - \tfrac{1}{2}\sin\theta_n\cos\theta_m(J_{n-1}\cot\theta_n - m_n\tan\theta_n)$$

$$\times(1-\delta_{n1})\prod_{i=m+1}^{n}\sin\theta_i + \frac{m_n\prod_{i=m+1}^{n}\sin\theta_i\cos\theta_m}{2\cos\theta_n} + \frac{m_m\cos\theta_n}{2\prod_{i=m+1}^{n}\sin\theta_i\cos\theta_m}$$

$$+\tfrac{1}{2}\cos\theta_m\cos\theta_n\left(J_n\prod_{i=m+1}^{n}\sin\theta_i - J_m\prod_{r=m+1}^{n}\sin^{-1}\theta_r\right)\Bigg\}. \tag{4.11}$$

In carrying out the θ integrations in Eq. (4.10), we notice that the integrals for $k < n$ and for $k > m$ give unity, since the Φ functions are properly normalized with respect to the measure

$$\prod_{j=2}^{6}\sin^{2j-3}\theta_j\,d\theta_j.$$

For the same reasons, the θ integrations in (4.11) give unity for $k > n$ and for $k < m$

We finally obtain:

(a) For $m > n > 1$,

$$D(m,n) = \frac{1}{2}\prod_{i=n}^{m}\left(\frac{\Gamma(m_i+m_i^*+2)\Gamma(m_i'+m_i^{*\prime}+2)}{\Gamma(m_i+1)\Gamma(m_i^*+1)\Gamma(m_i'+1)\Gamma(m_i^{*\prime}+1)}\right)^{\frac{1}{2}}\delta_{M,M'}^{m^+n^-}$$

$$\times\Bigg\{\Omega_m\Omega_n\frac{\Gamma(\tfrac{1}{2}(\mu_n+3))\Gamma(\tfrac{1}{2}(\nu_n+2))\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+1))}{2\Gamma(\tfrac{1}{2}(\mu_n+\nu_n+5))\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+6))}$$

$$\times \prod_{r=n+1}^{m-1} \Omega_r \frac{\Gamma(\tfrac{1}{2}(\mu_r+2))\Gamma(\tfrac{1}{2}(\nu+1))}{\Gamma(\tfrac{1}{2}(\mu_r+\nu_r+3))}\left(J_n(\mu_m+\nu_m+4)-J_m(\nu_m+1)\prod_{r=n+1}^{m-1}\frac{\nu_r+1}{\mu_r+\nu_r+3}\right)$$

$$+J_m(m_m+1)\frac{\Gamma(m_n+1)\Gamma(m_m+1)\Gamma(\tfrac{1}{2}(\nu_m+3))\Gamma(\tfrac{1}{2}(\nu_n+2))}{\Gamma(m_m+\tfrac{1}{2}(\nu_m+7))\Gamma(m_n+\tfrac{1}{2}(\nu_n+4))}\prod_{i=n+1}^{m-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+5))}$$

$$+m_n(m_n+\nu_n-J_{n-1}+2)\frac{\Gamma(m_m+2)\Gamma(m_n)\Gamma(\tfrac{1}{2}(\nu_n+2))\Gamma(\tfrac{1}{2}(\nu_m+1))}{\Gamma(m_n+\tfrac{1}{2}(\nu_n+4))\Gamma(m_m+\tfrac{1}{2}(\nu_m+5))}\prod_{i=n+1}^{m-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+1))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+3))}\Bigg\}. \tag{4.12}$$

(b) For $n > m > 1$,

$$D(m,n)=\frac{1}{2}\prod_{i=m}^{n}\left(\frac{\Gamma(m_i+m_i^*+2)\Gamma(m_i'+m_i^{*\prime}+2)}{\Gamma(m_i+1)\Gamma(m_i'+1)\Gamma(m_i^*+1)\Gamma(m_i^{*\prime}+1)}\right)^{\frac{1}{2}}\delta_{MM'}^{m^+n^-}$$

$$\times\Bigg\{\Omega_m\Omega_n\frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))\Gamma(\tfrac{1}{2}(\nu_n+1))\Gamma(\tfrac{1}{2}(\mu_n+3))}{2\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))\Gamma(\tfrac{1}{2}(\mu_n+\nu_n+6))}$$

$$\times\prod_{r=m+1}^{n-1}\Omega_r\frac{\Gamma(\tfrac{1}{2}(\mu_r+2))\Gamma(\tfrac{1}{2}(\nu_r+1))}{\Gamma(\tfrac{1}{2}(\mu_r+\nu_r+3))}\left(J_n(\nu_n+1)\prod_{r=m+1}^{n-1}\frac{\nu_r+1}{\mu_r+\nu_r+3}-J_m(\mu_n+\nu_n+4)\right)$$

$$+J_m(m_m+1)\frac{\Gamma(m_n+1)\Gamma(m_m+1)\Gamma(\tfrac{1}{2}(\nu_m+3))\Gamma(\tfrac{1}{2}(\nu_n+1))}{\Gamma(m_m+\tfrac{1}{2}(\nu_m+6))\Gamma(m_n+\tfrac{1}{2}(\nu_n+3))}\prod_{i=m+1}^{n-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+1))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+3))}$$

$$+m_n(m_n+\nu_n-J_{n-1}+3)\frac{\Gamma(m_m+2)\Gamma(m_n)\Gamma(\tfrac{1}{2}(\nu_n+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(m_m+\tfrac{1}{2}(\nu_m+6))\Gamma(m_n+\tfrac{1}{2}(\nu_n+5))}\prod_{i=m+1}^{n-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+5))}\Bigg\}. \tag{4.13}$$

The numbers μ, ν are defined by Eqs. (4.4) and the symbol $\delta_{MM'}$ by Eq. (4.10).

(c) For $m > n = 1$,

$$D(m,1)=\frac{1}{2}\prod_{i=2}^{m}\left(\frac{\Gamma(m_i+m_i^*+2)\Gamma(m_i'+m_i^{*\prime}+2)}{\Gamma(m_i+1)\Gamma(m_i^*+1)\Gamma(m_i'+1)\Gamma(m_i^{*\prime}+1)}\right)^{\frac{1}{2}}\delta_{MM'}^{m^+1^-}$$

$$\times\Bigg\{J_m(m_m+1)\frac{\Gamma(m_m+1)\Gamma(\tfrac{1}{2}(\nu_m+3))}{\Gamma(m_m+\tfrac{1}{2}(\nu_m+7))}\prod_{i=2}^{m-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+5))}$$

$$+m_1\frac{\Gamma(m_m+2)\Gamma(\tfrac{1}{2}(\nu_m+1))}{\Gamma(m_m+\tfrac{1}{2}(\nu_m+5))}\prod_{i=2}^{m-1}\frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+1))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+3))}$$

$$+\Omega_m\frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+1))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+6))}$$

$$\times\prod_{r=2}^{m-1}\Omega_r\frac{\Gamma(\tfrac{1}{2}(\mu_r+2))\Gamma(\tfrac{1}{2}(\nu_r+1))}{\Gamma(\tfrac{1}{2}(\mu_r+\nu_r+3))}\left(m_1(\mu_m+\nu_m+4)-J_m(\nu_m+1)\prod_{r=2}^{m-1}\frac{\nu_r+1}{\mu_r+\nu_r+3}\right)\Bigg\}. \tag{4.14}$$

(d) For $n > m = 1$,

$$D(1,n) = \frac{1}{2}\prod_{i=2}^{n}\left(\frac{\Gamma(m_i+m_i^*+2)\Gamma(m_i'+m_i^{*\prime}+2)}{\Gamma(m_i+1)\Gamma(m_i^*+1)\Gamma(m_i'+1)\Gamma(m_i^{*\prime}+1)}\right)^{\frac{1}{2}}\delta^{1^+n^-}_{MM'}\Bigg\{\Omega_n\frac{\Gamma(\frac{1}{2}(\mu_n+3))\Gamma(\frac{1}{2}(\nu_n+1))}{\Gamma(\frac{1}{2}(\mu_n+\nu_n+6))}$$

$$\times\prod_{r=2}^{n-1}\Omega_r\frac{\Gamma(\frac{1}{2}(\mu_r+2))\Gamma(\frac{1}{2}(\nu_r+1))}{\Gamma(\frac{1}{2}(\mu_r+\nu_r+3))}\left(J_n(\nu_n+1)\prod_{r=2}^{n-1}\frac{\nu_r+1}{\mu_r+\nu_r+3}-m_1(\mu_n+\nu_n+4)\right)$$

$$+(m_n+\nu_n-J_{n-1}+3)m_n\frac{\Gamma(m_n)\Gamma(\frac{1}{2}(\nu_n+3))}{\Gamma(m_n+\frac{1}{2}(\nu_n+5))}\prod_{i=2}^{n-1}\frac{\Gamma(\frac{1}{2}(\mu_i+2))\Gamma(\frac{1}{2}(\nu_i+3))}{\Gamma(\frac{1}{2}(\mu_i+\nu_i+5))}$$

$$+m_1\frac{\Gamma(m_n+1)\Gamma(\frac{1}{2}(\nu_n+1))}{\Gamma(m_n+\frac{1}{2}(\nu_n+3))}\prod_{i=2}^{n-1}\frac{\Gamma(\frac{1}{2}(\mu_1+2))\Gamma(\frac{1}{2}(\nu_i+1))}{\Gamma(\frac{1}{2}(\mu_i+\nu_i+3))}\Bigg\}. \tag{4.15}$$

In order to find the matrix elements of $A_{m\hat{n}}$, we substitute Eq. (3.20) in Eq. (4.1) and make use of Eqs. (4.6), (4.7), and (4.8).

Noting that

$$-m_j+J_{j-1}\cot^2\theta_j=-J_j+(J_{j-1}/\sin^2\theta_j), \tag{4.16}$$

so that

$$\sum_{j-m+1}^{6}(-m_j+J_{j-1}\cot^2\theta_j)\prod_{i=2}^{j}\sin^2\theta_i=J_m\prod_{i=2}^{m}\sin^2\theta_i-J\prod_{i=2}^{6}\sin^2\theta_i \tag{4.17}$$

and

$$\sum_{j=n+1}^{6}(-\hat{m}_j+\hat{J}_{j-1}\cot^2\hat{\theta}_j)\prod_{i=2}^{j}\sin^2\hat{\theta}_i=\hat{J}_n\prod_{i=2}^{n}\sin^2\theta_i-(J-2)\prod_{i=2}^{6}\sin^2\hat{\theta}_i, \tag{4.18}$$

and carrying out the θ integrations we finally obtain:

(a) For $m, n > 1$,

$$D(m,\hat{n})=\delta_{m_{m+1},m_m'}\delta_{\hat{m}_n-1,\hat{m}_n'}\prod_{k=1}^{6^m}\delta_{m_k,m_k'}\prod_{l=1}^{6^n}\delta_{\hat{m}_l,\hat{m}_l'}$$

$$\times\frac{1}{2}\Bigg\{-\frac{44(J+4)^2}{(2J+7)(2J+9)}\prod_{i=n+1}^{6}\hat{\Omega}_i\frac{\Gamma(\frac{1}{2}(\hat{\nu}_i+3))\Gamma(\frac{1}{2}(\hat{\mu}_i+3))}{\Gamma(\frac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))}\prod_{j=m+1}^{6}\Omega_j\frac{\Gamma(\frac{1}{2}(\mu_j+3))\Gamma(\frac{1}{2}(\nu_j+3))}{\Gamma(\frac{1}{2}(\mu_j+\nu_j+6))}$$

$$+\Bigg\{J_m\Omega_m\frac{\Gamma(\frac{1}{2}(\mu_m+3))\Gamma(\frac{1}{2}(\nu_m+2))}{\Gamma(\frac{1}{2}(\mu_m+\nu_m+5))}\prod_{j=2}^{m-1}\Omega_j\frac{\Gamma(\frac{1}{2}(\mu_j+2))\Gamma(\frac{1}{2}(\nu_j+2))}{\Gamma(\frac{1}{2}(\mu_j+\nu_j+4))}\prod_{k=m+1}^{6}\Omega_k\frac{\Gamma(\frac{1}{2}(\mu_k+2))\Gamma(\frac{1}{2}(\nu_k+1))}{\Gamma(\frac{1}{2}(\mu_k+\nu_k+3))}$$

$$-J\Omega_m\frac{\Gamma(\frac{1}{2}(\mu_m+3))\Gamma(\frac{1}{2}(\nu_m+2))}{\Gamma(\frac{1}{2}(\mu_m+\nu_m+5))}\prod_{k=m+1}^{6}\Omega_k\frac{\Gamma(\frac{1}{2}(\mu_k+2))\Gamma(\frac{1}{2}(\nu_k+3))}{\Gamma(\frac{1}{2}(\mu_k+\nu_k+5))}\Bigg\}\prod_{i=n+1}^{6}\hat{\Omega}_i\frac{\Gamma(\frac{1}{2}(\hat{\mu}_i+3))\Gamma(\frac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\frac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))}$$

$$+\Bigg\{\hat{J}_n\hat{\Omega}_n\frac{\Gamma(\frac{1}{2}(\hat{\mu}_n+3))\Gamma(\frac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\frac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))}\prod_{j=2}^{n-1}\hat{\Omega}_j\frac{\Gamma(\frac{1}{2}(\hat{\mu}_j+2))\Gamma(\frac{1}{2}(\hat{\nu}_j+2))}{\Gamma(\frac{1}{2}(\hat{\mu}_j+\hat{\nu}_j+4))}\prod_{k=n+1}^{6}\hat{\Omega}_k\frac{\Gamma(\frac{1}{2}(\hat{\mu}_k+2))\Gamma(\frac{1}{2}(\hat{\nu}_k+1))}{\Gamma(\frac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+3))}$$

$$-(J-2)\frac{\Gamma(\frac{1}{2}(\hat{\mu}_n+3))\Gamma(\frac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\frac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))}\hat{\Omega}_n\prod_{k=n+1}^{6}\hat{\Omega}_k\frac{\Gamma(\frac{1}{2}(\hat{\mu}_k+2))\Gamma(\frac{1}{2}(\nu_k+3))}{\Gamma(\frac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+5))}\Bigg\}$$

$$\times\prod_{i=m+1}^{6}\Omega_i\frac{\Gamma(\frac{1}{2}(\mu_i+3))\Gamma(\frac{1}{2}(\nu_i+3))}{\Gamma(\frac{1}{2}(\mu_i+\nu_i+6))}-J_m\Omega_m\frac{\Gamma(\frac{1}{2}(\mu_m+3))\Gamma(\frac{1}{2}(\nu_m+2))}{\Gamma(\frac{1}{2}(\mu_m+\nu_m+5))}$$

$$\times\prod_{k=2}^{m-1}\Omega_k\frac{\Gamma(\frac{1}{2}(\mu_k+2))\Gamma(\frac{1}{2}(\nu_k+2))}{\Gamma(\frac{1}{2}(\mu_k+\nu_k+4))}\prod_{i=m+1}^{6}\Omega_i\frac{\Gamma(\frac{1}{2}(\mu_i+2))\Gamma(\frac{1}{2}(\nu_i+1))}{\Gamma(\frac{1}{2}(\mu_i+\nu_i+3))}$$

$$\times \hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))} \prod_{j=n+1}^{6} \hat{\Omega}_j \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_j+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+\hat{\nu}_j+5))}$$

$$- \hat{J}_n \hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}+\hat{\nu}_n+5))} \prod_{k=2}^{n-1} \hat{\Omega}_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+4))} \prod_{i=n+1}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+3))}$$

$$\times \Omega_m \frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))} \prod_{j=m+1}^{6} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+5))} \Bigg\}. \tag{4.19}$$

(b) For $n > 1$,

$$D(1,\hat{n}) = \delta_{m_1+1,m_1'}\delta_{\hat{m}_n-1,\hat{m}_n'} \prod_{k=2}^{6} \delta_{m_k,m_k'} \prod_{l=1}^{6} {}^{n}\delta_{\hat{m}_l\hat{m}_l'}$$

$$\times \frac{1}{2}\Bigg\{ \frac{-44(J+4)^2}{(2J+7)(2J+9)} \prod_{i=n+1}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))} \prod_{j=2}^{6} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+3))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+6))}$$

$$- \prod_{i=n+1}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))} \Bigg(-\prod_{j=2}^{6} \Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+1))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+3))} m_1$$

$$+ J \prod_{j=2}^{6} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+5))} \Bigg) - \Bigg(-\hat{J}_n \hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))}$$

$$\times \prod_{j=2}^{n-1} \hat{\Omega}_j \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_j+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+\hat{\nu}_j+4))} \prod_{k=n+1}^{6} \hat{\Omega}_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+3))}$$

$$+ (J-2)\hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))} \prod_{k=n+1}^{6} \hat{\Omega}_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+5))} \Bigg) \prod_{i=2}^{6} \Omega_i \frac{\Gamma(\tfrac{1}{2}(\mu_i+3))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+6))}$$

$$- m_1 \hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))} \prod_{k=2}^{6} \Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma(\tfrac{1}{2}(\nu_k+1))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+3))} \prod_{l=n+1}^{6} \hat{\Omega}_l \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_l+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_l+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_l+\hat{\nu}_l+5))}$$

$$- \hat{J}_n \prod_{k=2}^{n-1} \hat{\Omega}_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+4))} \times \hat{\Omega}_n \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_n+2))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_n+\hat{\nu}_n+5))}$$

$$\times \prod_{i=n+1}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+3))} \prod_{j=m+1}^{6} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+5))} \Bigg\}. \tag{4.20}$$

(c) For $m > 1$,

$$D(m,\hat{1}) = \delta_{m_m+1,m_m'}\delta_{\hat{m}_1-1,\hat{m}_1'} \prod_{k=1}^{6} {}^{m}\delta_{m_k m_k'} \prod_{l=2}^{6} \delta_{\hat{m}_l,\hat{m}_l'}$$

$$\times \frac{1}{2}\Bigg\{ \frac{-44(J+4)^2}{(2J+7)(2J+9)} \prod_{i=2}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))} \prod_{j=m+1}^{6} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+3))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+6))}$$

$$- \Bigg(-J_m \Omega_m \frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))} \prod_{j=2}^{m-1} \Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+2))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+4))} \prod_{k=m+1}^{6} \Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma(\tfrac{1}{2}(\nu_k+1))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+3))}$$

$$+ J\Omega_m \frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))} \prod_{k=m+1}^{6} \Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma(\tfrac{1}{2}(\nu_k+3))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+5))} \Bigg) \prod_{i=2}^{6} \hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))}$$

$$+ \prod_{i=m+1}^{6} \Omega_i \frac{\Gamma(\tfrac{1}{2}(\mu_i+3))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+6))} \Bigg(\hat{m}_1 \prod_{k=2}^{6} \Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma((\tfrac{1}{2}\nu_k+1))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+3))} -$$

$$- (J-2)\prod_{k=2}^{6}\Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+5))}\Bigg)$$

$$- \hat{m}_1\Omega_m \frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))}\prod_{j=m+1}^{6}\Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+5))}\prod_{k=2}^{6}\hat{\Omega}_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+3))}$$

$$- J_m\Omega_m \frac{\Gamma(\tfrac{1}{2}(\mu_m+3))\Gamma(\tfrac{1}{2}(\nu_m+2))}{\Gamma(\tfrac{1}{2}(\mu_m+\nu_m+5))}\times\prod_{k=2}^{m-1}\Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma(\tfrac{1}{2}(\nu_k+2))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+4))}$$

$$\times\prod_{j=2}^{6}\hat{\Omega}_j \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_j+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+\hat{\nu}_j+5))}\times\prod_{i=m+1}^{6}\Omega_i \frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+1))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+3))}\Bigg\}. \tag{4.21}$$

(d) Finally

$$D(1,\hat{1}) = \delta_{m_1+1,m_1'}\delta_{\hat{m}_1-1,m_1'}\prod_{k=2}^{6}\delta_{m_k m_k'}\prod_{l=2}^{6}\delta_{\hat{m}_l\hat{m}_l'}$$

$$\times\frac{1}{2}\Bigg\{\frac{-44(J+4)^2}{(2J+7)(2J+9)}\prod_{i=2}^{6}\hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))}\prod_{j=2}^{6}\Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+3))\Gamma(\tfrac{1}{2}(\nu_j+3))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+6))}$$

$$+\prod_{i=2}^{6}\hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+3))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+6))}\Bigg(m_1\prod_{j=2}^{6}\Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_j+2))\Gamma(\tfrac{1}{2}(\nu_j+1))}{\Gamma(\tfrac{1}{2}(\mu_j+\nu_j+3))}$$

$$- J\prod_{k=2}^{6}\Omega_k \frac{\Gamma(\tfrac{1}{2}(\mu_k+2))\Gamma(\tfrac{1}{2}(\nu_k+3))}{\Gamma(\tfrac{1}{2}(\mu_k+\nu_k+5))}\Bigg)+\prod_{i=2}^{6}\Omega_i \frac{\Gamma(\tfrac{1}{2}(\mu_i+3))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+6))}$$

$$\times\Bigg(\hat{m}_1\prod_{j=2}^{6}\hat{\Omega}_j \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_j+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+\hat{\nu}_j+3))} - (J-2)\prod_{k=2}^{6}\Omega_k \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_k+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_k+\hat{\nu}_k+5))}\Bigg)$$

$$- m_1\prod_{i=2}^{6}\hat{\Omega}_i \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+3))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+5))}\prod_{j=2}^{6}\Omega_j \frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+1))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+3))}$$

$$- \hat{m}_1\prod_{i=2}^{6}\Omega_i \frac{\Gamma(\tfrac{1}{2}(\mu_i+2))\Gamma(\tfrac{1}{2}(\nu_i+3))}{\Gamma(\tfrac{1}{2}(\mu_i+\nu_i+5))}\prod_{j=2}^{6}\hat{\Omega}_j \frac{\Gamma(\tfrac{1}{2}(\hat{\mu}_j+2))\Gamma(\tfrac{1}{2}(\hat{\nu}_i+1))}{\Gamma(\tfrac{1}{2}(\hat{\mu}_i+\hat{\nu}_i+3))}\Bigg\}. \tag{4.22}$$

The matrices $D(\hat{m}, n)$ are obtained from Eqs. (4.19)–(4.22) by interchanging symbols with carets with symbols without carets, at the same time interchanging the factors J and $J-2$ that appear in these formulas.

5. CONCLUSION

We have derived and explicitly displayed the matrices of the generators of the algebra of $U(6, 6)$ between harmonic function states which may, under certain circumstances, be considered highest weight states.

It is now in principle possible to calculate the matrix elements of any element of the $U(6, 6)$ algebra between these states.

If we proceed in a similar manner to find the matrix elements of the generators between any two harmonic function states, then it is again in principle possible to find the matrix elements of the finite transformations of the groups generated by the algebra. It is clear, however, from the complicated structure of the results obtained above that such a task would be prohibitively complicated, at least within the framework of the formalism used here. We should, nevertheless, like to find compact expressions that describe the matrices for finite transformations—expressions that are analogous to the Wigner d functions for the rotation group. This would require the decomposition of $U(6, 6)$ into an ordered product of a finite number of its one-parameter subgroups. This problem is being studied.

ACKNOWLEDGMENTS

The author is grateful to the IAEA for providing him with financial support. He is also grateful to Professor A. Salam, Professor P. Budini, and the IAEA for hospitality extended to him at the International Center for Theoretical Physics, Trieste. He deeply appreciates illuminating discussions he has had with Professor R. L. Ingraham over certain aspects of this work.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Superconductivity in One and Two Dimensions. II. Charged Systems

T. M. RICE*
Department of Physics, University of California, San Diego, La Jolla, California

(Received 10 October 1966)

In an earlier paper the Ginzburg–Landau free energy functional was used to calculate the effect of thermodynamic fluctuations on the off-diagonal correlation function and we found no off-diagonal long-range order in one- and two-dimensional systems. It has been pointed out that, for a charged system, the use of the Ginzburg–Landau free energy functional is in error for arbitrary nonequilibrium values of the order parameter since the electrostatic energy of the charge fluctuations associated with an arbitrary order parameter is not included in the free energy functional. We have not succeeded in a direct generalization of the free energy functional so we are forced to proceed by inference from the generalized random phase approximation (RPA). We find that, for uncharged systems, the RPA gives a linearization of the results obtained earlier using the Ginzburg–Landau theory. For charged systems we find in the RPA results similar to those obtained for uncharged systems. From this we conclude that it is very likely that, as in uncharged systems, there will be no ODLRO in charged infinite one- and two-dimensional systems.

I. INTRODUCTION

THERE has been considerable interest recently in the possibility of superconductivity in one- and two-dimensional systems. The original impetus for this work came from a suggestion by Little[1] that certain organic macromolecules might be superconductors with very high transition temperatures. Little[1] used as a criterion for superconductivity that the Bardeen–Cooper–Schrieffer equations have a solution. Ferrell[2] examined the effects of low-lying collective modes on a one-dimensional superconductor and found that the expectation value of the gap function was zero and therefore no superconductivity. Recently Bychkov, Gorkov, and Dyaloshinski[3] have argued that the macromolecules proposed by Little[1] would not have low-lying collective modes because of the Coulomb interaction and that such macromolecules could be superconductors. In a previous paper[4] (hereafter referred to as I) we examined the effects of thermodynamic fluctuations on the existence of off-diagonal long-range order (ODLRO) in one- and two-dimensional systems. The concept of ODLRO was first introduced by Penrose and Onsager[5] for bosons in the superfluid state. Yang[6] has shown that ODLRO in Fermi systems implies flux quantization. In I we calculated the off-diagonal correlation function by first assuming the existence of an order parameter and then averaging over all possible choices of the order parameter weighting each according to its thermodynamic probability in the Ginzburg–Landau theory.[7] We found that, whereas in three dimensions the resulting correlation function exhibited ODLRO, in one and two dimensions this was not so.

At first sight it would seem that it is possible to apply the arguments given in I to charged or uncharged superconductors since they depend only on the Ginzburg–Landau theory which is known to work well for real conductors. Kohn[8] has pointed out that the use of the Ginzburg–Landau free energy functional for arbitrary values of the Ginzburg–Landau order parameter $\Psi(\mathbf{x})$ is highly suspect for charged systems. An arbitrary $\Psi(\mathbf{x})$ has a current associated with it $\mathbf{J} \sim \Psi^*\nabla\Psi - (\nabla\Psi^*)\Psi$. Further, since in general $\operatorname{div}\mathbf{J} \neq 0$, an arbitrary $\Psi(\mathbf{x})$ will have a charge fluctuation associated with it. Charge fluctuations require an electrostatic energy which is positive and this tends to inhibit them, but this fact is not contained in the Ginzburg–Landau theory. This criticism does not apply to the equilibrium values of $\Psi(\mathbf{x})$, which satisfy the Ginzburg–Landau equations, (since $\operatorname{div}\mathbf{J} = 0$ for all such values). The criticism only applies to the use of the Ginzburg–Landau free energy functional to describe an arbitrary nonequilibrium value of $\Psi(\mathbf{x})$ in a charged system.

We have not succeeded in generalizing the Ginzburg–Landau free energy functional to include these effects and thus we cannot give a direct generalization of the results in I to a charged system. Thus we are forced to proceed by inference from the generalized random

* Present address: Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] W. A. Little, Phys. Rev. **134**, A1416 (1964).

[2] R. A. Ferrell, Phys. Rev. Letters **13**, 330 (1964).

[3] Yu. A. Bychkov, L. P. Gorkov, and I. E. Dyaloshinski, JETP Pismav Redaktsiyu **2**, 146 (1965) [English transl.: Soviet Phys.—JETP Letters **2**, 92 (1965)].

[4] T. M. Rice, Phys. Rev. **140**, A1889 (1965).

[5] O. Penrose, Phil. Mag. **42**, 1373 (1951); O. Penrose and L. Onsager, Phys. Rev. **104**, 576 (1956).

[6] C. N. Yang, Rev. Mod. Phys. **34**, 694 (1962).

[7] V. L. Ginzburg and L. D. Landau, Zh. Eksperim. i Teor. Fiz. **20**, 1064 (1950).

[8] W. Kohn (private communication).

phase approximation (RPA). We calculate the off-diagonal correlation function in the RPA for both the uncharged and the charged systems. We find that, in an uncharged system, the form of the off-diagonal correlation function obtained from the RPA is a linearization of that obtained in I using the Ginzburg–Landau theory. We then calculate the off-diagonal correlation function in the RPA for a charged system and we find a form similar to that obtained for the uncharged system. From this we infer that it is very likely that a more correct calculation would give results similar to that obtained in I—namely, that there could be no ODLRO in an infinite one- or two-dimensional system. This is rather remarkable since the question of ODLRO is closely related to the long-wavelength collective modes of the system. These modes are strongly influenced by the Coulomb interaction.

We use the functional integral formulation of the theory of superconductivity developed by Hubbard[9] and Langer[10] since it facilitates the comparison of the Ginzburg–Landau theory and the RPA. These two approaches appear in this formulation as two different approximations to the full functional integral. In Appendix A we show how Werthamer's[11] generalization of the Ginzburz–Landau functional to all temperatures can be derived readily from this formulation.

II. FUNCTIONAL INTEGRATION TECHNIQUE

We start by discussing the functional integral formulation of the theory of superconductivity. As we are interested in the effects of the Coulomb interaction on the off-diagonal correlation function, we use a Hamiltonian which contains a short-range attractive force and a Coulomb repulsion. The functional integration technique has been applied to superconductivity by Hubbard,[9] Muhlschlegel,[12] and Langer.[10] Hubbard and Langer use just a short-range attractive force while Muhlschlegel considers only the truncated BCS Hamiltonian. We wish to extend the Hubbard formulation to include the Coulomb repulsion. The Hamiltonian of the system is given by

$$H-\mu N=\sum_{\mathbf{p},\sigma}\mathcal{E}(\mathbf{p})a^{\dagger}_{\mathbf{p},\sigma}a_{\mathbf{p},\sigma}-\tfrac{1}{2}\sum_{\substack{\mathbf{p},\mathbf{p}',\mathbf{Q}\\ \sigma}}g_{\mathbf{Q}}a^{\dagger}_{\mathbf{p}+\mathbf{Q},\sigma}a^{\dagger}_{-\mathbf{p},-\sigma}a_{-\mathbf{p}',-\sigma}a_{\mathbf{p}'+\mathbf{Q},\sigma}+\tfrac{1}{2}\sum_{\substack{\mathbf{p},\mathbf{p}',\mathbf{Q}\\ \sigma,\sigma'}}v(\mathbf{Q})a^{\dagger}_{\mathbf{p}+\mathbf{Q}',\sigma}a_{\mathbf{p},\sigma}a^{\dagger}_{\mathbf{p}'-\mathbf{Q},\sigma'}a_{\mathbf{p}',\sigma'}. \quad (1)$$

Here the $a_{\mathbf{p},\sigma}$ and $a^{\dagger}_{\mathbf{p},\sigma}$ are the annihilation and creation operators for Fermions of momentum $\mathbf{p}$ and spin σ and $\mathcal{E}(\mathbf{p})$ in the energy measured from the chemical potential μ. The momentum dependence of the attractive interaction is a purely formal device which, as we see, enables us to calculate easily the off-diagonal correlation function. We always set $g_{\mathbf{Q}}\equiv g$ at the end of our calculations. The Coulomb interaction is given by $v(\mathbf{Q})=4\pi e^2/Q^2$.

Let Z denote the grand partition function

$$Z=\mathrm{Tr}\exp\{-\beta(H-\mu N)\}, \quad (2)$$

where β is the inverse temperature. Following Hubbard we introduce a Feynman[13] ordering label s and write

$$Z=\mathrm{Tr}\exp\left\{-\frac{\beta}{N}\left[\sum_{\mathbf{p},\sigma,s}\mathcal{E}(\mathbf{p})a^{\dagger}_{\mathbf{p},\sigma,s}a_{\mathbf{p},\sigma,s}-\sum_{\mathbf{Q},s}g_{\mathbf{Q}}b^{\dagger}_{\mathbf{Q},s}b_{\mathbf{Q},s}+\tfrac{1}{2}\sum_{\mathbf{Q},s}v(\mathbf{Q})\rho^{\dagger}_{\mathbf{Q},s}\rho_{\mathbf{Q},s}\right]\right\}, \quad (3)$$

where we have introduced the operators

$$b^{\dagger}_{\mathbf{Q},s}=\sum_{\mathbf{p}}a^{\dagger}_{\mathbf{p}+\mathbf{Q},\sigma,s}a^{\dagger}_{-\mathbf{p},-\sigma,s}$$

and

$$\rho^{\dagger}_{\mathbf{Q},s}=\sum_{\mathbf{p},\sigma}a^{\dagger}_{\mathbf{p}+\mathbf{Q},\sigma,s}a_{\mathbf{p},\sigma,s}.$$

The sum over s runs from 1 to N. With the introduction of the ordering label, the quantities now can be treated as c numbers with an error which goes to zero as $N\to\infty$. Using the identity

$$\exp\{|a|^2\}=\frac{1}{\pi}\int_{-\infty}^{+\infty}dx_1\,dx_2\times\exp\{-|x|^2+ax+a^*x^*\},$$

where $x=x_1+ix_2$, we can express the partition function as a double functional integral

$$Z=\int_{-\infty}^{+\infty}\prod_{\mathbf{Q},i\,s}\left\{\frac{\beta}{\pi N}\,dx_{\mathbf{Q},s,i}\,dy_{\mathbf{Q},s,i}\right\}\times\exp\left\{-\frac{\beta}{N}\sum_{\mathbf{Q},s}(|x_{\mathbf{Q},s}|^2+|y_{\mathbf{Q},s}|^2)\right\}\mathfrak{L}[x_{\mathbf{Q},s},y_{\mathbf{Q},s}]. \quad (4)$$

The functional $\mathfrak{L}$ is given by

$$\mathfrak{L}[x_{\mathbf{Q},s},y_{\mathbf{Q},s}]=\mathrm{Tr}\exp\left\{-\frac{\beta}{N}\left[\sum_{\mathbf{p},\sigma,s}\mathcal{E}(\mathbf{p})a^{\dagger}_{\mathbf{p},\sigma,s}a_{\mathbf{p},\sigma,s}+\sum_{\mathbf{Q},s}(g_{\mathbf{Q}})^{\frac{1}{2}}(x^*_{\mathbf{Q},s}b^{\dagger}_{\mathbf{Q},s}+\mathrm{h.c.})+\sum_{\mathbf{Q},s}(-\tfrac{1}{2}v(\mathbf{Q}))^{\frac{1}{2}}(y^*_{\mathbf{Q},s}\rho^{\dagger}_{\mathbf{Q},s}+\mathrm{h.c.})\right]\right\}. \quad (5)$$

$\mathfrak{L}$ is the partition function of noninteracting electrons moving in an arbitrary "scattering" potential

[9] J. Hubbard, Phys. Rev. Letters 3, 77 (1959).
[10] J. Langer, Phys. Rev. 134, A553 (1964).
[11] N. R. Werthamer, Phys. Rev. 132, 663 (1963); L. Tewordt, *ibid.* 132, 595 (1963).
[12] B. Muhlschlegel, J. Math. Phys. 3, 522 (1962).
[13] R. P. Feynman, Phys. Rev. 84, 108 (1951).

and an arbitrary "pairing" potential. It can be evaluated by the usual procedures of many-body perturbation theory. If now we replace the sum by an integral, i.e., $(\beta/N)\sum_s \to \int_0^\beta ds$, and expand all quantities depending on s in a Fourier series as is usually done in finite temperature many-body theory, we get that

$$Z = \int \prod_{\substack{\alpha \\ i=1,2}} \left(\frac{1}{\pi} dx_{\alpha,i}\, dy_{\alpha,i}\right) \times \exp\left\{-\sum_\alpha (|x_\alpha|^2 + |y_\alpha|^2)\right\} \mathfrak{L}[x_\alpha, y_\alpha], \quad (6)$$

where $\alpha = (\mathbf{Q}, m)$ and m runs over all integers and

$$\begin{aligned}\mathfrak{L}[x_\alpha, y_\alpha] = \operatorname{Tr} \exp\Big\{ &- \int_0^\beta ds \sum_{\mathbf{p},\sigma} \mathcal{E}(\mathbf{p}) a^\dagger_{\mathbf{p},\sigma,s} a_{\mathbf{p},\sigma,s} \\ &- \beta^{-1} \int_0^\beta ds \sum_\alpha (g_{\mathbf{Q}})^{\frac{1}{2}} (x^*_\alpha b^\dagger_{\mathbf{Q},s} \exp(2\pi i m s/\beta) + \text{h.c.}) \\ &- \beta^{-1} \int_0^\beta ds \sum_\alpha (-\tfrac{1}{2} v(\mathbf{Q}))^{\frac{1}{2}} \\ &\times (y^*_\alpha \rho^\dagger_{\mathbf{Q},s} \exp(2\pi i m s/\beta) + \text{h.c.})\Big\}. \end{aligned} \quad (7)$$

The calculation of $\mathfrak{L}$ reduces to the calculation of the "partition" function of noninteracting particles moving in arbitrary "scattering" and "pairing" potentials. An expansion of $\mathfrak{L}$ in powers of the "pairing" potentials $\{x\}$ and "scattering" potential $\{y\}$ is generated by perturbation theory. The "partition" function then is given by the exponential of the sum of all connected electronlike graphs. The contribution from each graph may be written down using the following rules:

(1) For each solid electron line, a factor

$$[\beta(\mathcal{E}(\mathbf{p}) - \pi i l/\beta)]^{-1}.$$

Label each interaction line with a momentum and frequency which must be conserved at all vertices.

(2) For each ingoing "pairing" interaction, denoted by a wavy line, a factor $(g_{\mathbf{Q}}\beta)^{\frac{1}{2}} x^*_\alpha$; for each outgoing "pairing" interaction, denoted by a wavy line, a factor $-(g_{\mathbf{Q}}\beta)^{\frac{1}{2}} x_\alpha$.

(3) For each ingoing "scattering" interaction, denoted by a broken line, a factor $(-\frac{1}{2}v(\mathbf{Q})\beta)^{\frac{1}{2}} y^*_\alpha$; for each outgoing "scattering" interaction, denoted by a broken line, a factor $(-\frac{1}{2}v(\mathbf{Q})\beta)^{\frac{1}{2}} y_\alpha$.

(4) $(-1)/w$ for each closed loop where w is the rotational symmetry of the loop, and a factor $(\frac{1}{2})$ for each closed loop involving a "pairing" interaction for overcounting the spin sum.

(5) Sum over $\mathbf{p}$, σ and l odd.

With these rules we can evaluate formally the functional $\mathfrak{L}[x_\alpha, y_\alpha]$ and can write the ratio of the partition functions of the interacting and noninteracting systems Z' as

$$Z' = \int \prod_{\alpha,i} \left(\frac{1}{\pi} dx_{\alpha i}\, dy_{\alpha i}\right) \exp\{-Y[x_\alpha, y_\alpha]\}, \quad (8)$$

where

$$Y[x_\alpha, y_\alpha] = -\langle S(\beta)\rangle_{\text{linked}} + \sum_\alpha (|x_\alpha|^2 + |y_\alpha|^2),$$

and

$$S(\beta) = T_s \exp\left(-\int_0^\beta H_{\text{int}}(s)\, ds\right). \quad (9)$$

$H_{\text{int}}(s)$ is the interaction Hamiltonian Eq. (7) in the interaction representation.

In the next section we discuss the evaluation of Y for the uncharged system in various approximations and the calculation of the off-diagonal correlation function.

III. UNCHARGED SYSTEM

We begin by considering the uncharged system, i.e., no Coulomb interactions between the particles $v(\mathbf{Q}) \equiv 0$. The Hamiltonian is now the same as that used by Langer.[10] Hubbard[9] and Langer[10] showed that the superconducting transition manifests itself in a shift of the minimum of $Y[x_\alpha]$ away from the origin. The position of the minimum is at $|x_{0,0}|^2 = s_0$, where s_0 is given by the BCS equation

$$1 = \frac{g_0}{2} \sum_{\mathbf{p}} \frac{\tanh\{\frac{1}{2}\beta[\mathcal{E}(\mathbf{p})^2 + g_0 s_0/\beta]^{\frac{1}{2}}\}}{[\mathcal{E}(\mathbf{p})^2 + g_0 s_0/\beta]^{\frac{1}{2}}}. \quad (10)$$

There are two alternative approximations we can now make,

(1) We can expand $Y[x_\alpha]$ in powers of x_α about the origin, i.e., $x_\alpha = 0$, all α. If we drop the "time"-dependent terms—ignore the s dependence of the variables $\{x_{\mathbf{Q},s}\}$—then we need only include $\alpha = \{\mathbf{Q}, 0\}$ terms and, expanding the coefficients in powers of $\mathbf{Q}$, we get the Ginzburg–Landau functional. This procedure is valid only near the transition temperature but it can be extended to obtain the Werthamer[11] function, which is valid for all temperatures, as indicated in Appendix A.

(2) Alternatively, we can expand $Y[x_\alpha]$ about the new minimum $x^2_{0,0} = s_0$; $x_\alpha = 0$, $\alpha \neq (0, 0)$. This expansion, which is a generalized RPA, has been carried out by Langer.[10] We calculate the off-diagonal correlation function in both approximations and show how the results are related.

Ginzburg–Landau Approximation

We begin by discussing the first approach. We neglect all "time"-dependent graphs and consider

just the two lowest-order graphs shown in Fig. 1; then we have

$$Y[x_Q] = \sum_Q (1 + K_Q^{(1)}) x_Q^* x_Q + \sum_{Q_1 Q_2 Q_3} K^{(2)}_{Q_1,Q_2,Q_3} x^*_{Q_1} x^*_{Q_2} x_{Q_3} x_{Q_1+Q_2-Q_3}, \quad (11)$$

where

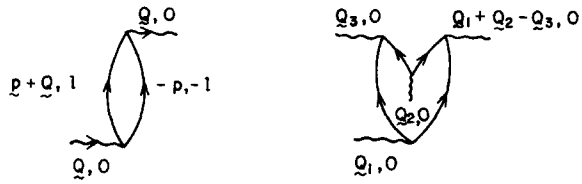

FIG. 1. The graphs included in the Ginzburg–Landau approximation.

$$K_Q^{(1)} = -\frac{1}{2\beta} \sum_{p,l,\sigma} \frac{g_Q}{(\mathcal{E}(\mathbf{p}+\mathbf{Q}) - \pi i l/\beta)(\mathcal{E}(\mathbf{p}) + \pi i l/\beta)}, \quad (12)$$

$$K^{(2)}_{Q_1,Q_2,Q_3} = \frac{1}{2\beta^2} \sum_{p,l,\sigma} \frac{(g_{Q_1} g_{Q_2} g_{Q_3} g_{Q_1+Q_2-Q_3})^{\frac{1}{2}}}{(\mathcal{E}(\mathbf{p}+\mathbf{Q}_1) - \pi i l/\beta)(\mathcal{E}(\mathbf{p}+\mathbf{Q}_1-\mathbf{Q}_3) + \pi i l/\beta)(\mathcal{E}(\mathbf{p}+\mathbf{Q}_1+\mathbf{Q}_2-\mathbf{Q}_3) - \pi i l/\beta)} \times \frac{1}{(\mathcal{E}(\mathbf{p}) + \pi i l/\beta)}. \quad (13)$$

Expanding in powers of $\mathbf{Q}$ and setting $g_Q \equiv g$, we write

$$K_Q^{(1)} = K_0 + K_1 Q^2 + \cdots; \quad K^{(2)}_{Q_1,Q_2,Q_3} = K_2 + \cdots.$$

The evaluation of the coefficients K_0, K_1, and K_2 has been given by Gorkov.[14] Writing $Y[x_Q]$ as a functional in position space, we get

$$Y[\Delta(\mathbf{R})] = \beta a' \int |\Delta(\mathbf{R})|^2 \, d\mathbf{R} + \tfrac{1}{2}\beta b' \int |\Delta(\mathbf{R})|^4 \, d\mathbf{R} + \beta c' \int |\nabla \Delta(\mathbf{R})|^2 \, d\mathbf{R}, \quad (14)$$

where

$$\Delta(\mathbf{R}) = (g/\beta)^{\frac{1}{2}} \sum_Q x_Q \exp(i\mathbf{Q}\cdot\mathbf{R}).$$

$\Delta(\mathbf{R})$ is Gorkov gap function[14] and is proportional to the Ginzburg–Landau order parameter $\Psi(\mathbf{R})$:

$$\Psi(\mathbf{R}) = (7\zeta(3)N/8\pi^2 T_c^2)^{\frac{1}{2}} \Delta(\mathbf{R}). \quad (15)$$

The coefficients a', b', c' are given by

$$a' = (1 + K_0)/g; \quad b' = 2\beta K_2/g^2; \quad c' = -K_1/g^2.$$

If we write the functional Y in Eq. (14) as a functional of $\Psi(\mathbf{R})$ rather than $\Delta(\mathbf{R})$, then we see at once that it has the same functional form as the Ginzburg–Landau free energy functional. Comparing coefficients, we find, as shown in detail in Appendix A, that $Y = \beta F$ where F is the Ginzburg–Landau free energy functional. In I we calculated the off-diagonal correlation function by averaging over all possible choices of the order parameter and using the Ginzburg–Landau free energy functional to weight each choice.

The Generalized Random Phase Approximation

In this section we calculate the off-diagonal correlation function in the generalized RPA. The off-diagonal correlation function can be obtained by differentiating the grand partition function,

$$\frac{\partial}{\partial g_Q} \ln Z[g_Q]\Big|_{g_Q = g} = \frac{\mathrm{Tr}\left\{\frac{1}{2}\beta \sum_{p,p',\sigma} a^\dagger_{p+Q\sigma} a^\dagger_{-p,-\sigma} a_{-p',-\sigma} a_{p'+Q,\sigma} \exp(-\beta(H-\mu N))\right\}}{\mathrm{Tr}\{\exp(-\beta(H-\mu N))\}} = \beta G^{OD}(\mathbf{Q}), \quad (16)$$

and we have used the cyclic property of the trace to write the derivative in this form.

The grand partition function is obtained in the RPA by expanding the functional about its minimum and evaluating the functional integral by using a saddle-point approximation. As Langer[10] has shown, the saddle-point approximation is at best a physical one, which is mathematically unjustified. We follow Langer's analysis, extending it to a momentum-dependent pairing interaction and we find

$$Z'[g_Q] = \prod_{\substack{\alpha \\ i=1,2}} \left(\frac{1}{\pi}\int dx_{\alpha,i}\right) \exp(-Y[g_Q, x_\alpha]),$$

where Z' is ratio of partition functions of the interacting and noninteracting systems, and

$$Y[g_Q, x_\alpha] = Y(g_0, s_0(g_0)) + \tfrac{1}{2}Y''(g_0, s_0(g_0))(|x_0|^2 - s_0^2(g_0)) + \sum_{\alpha \neq (0,0)} \{(1 - \gamma_\alpha(g_Q, g_0)) x^*_\alpha x_\alpha + \tfrac{1}{2}\delta_\alpha(g_Q, g_0)(x^*_\alpha x^*_{-\alpha} + \text{c.c.})\}. \quad (17)$$

[14] L. I. Gorkov, Zh. Eksperim i Teor. Fiz. 36, 1918 (1959) English transl.: Soviet Phys.—JETP 9, 1364 (1959)].

The first term is

$$Y(g_0, s_0(g_0)) = s_0 - 2\sum_{\mathbf{p}} \ln\left(\frac{\cosh[(\frac{1}{2}\beta)(\mathcal{E}(p)^2 + g_0 s_0/\beta)^{\frac{1}{2}}]}{\cosh \frac{1}{2}\beta\mathcal{E}(\mathbf{p})}\right), \quad (18)$$

where the minimum value of $Y(|x_{0,0}|)$ occurs at

$$|x_{0,0}|^2 = s_0,$$

i.e., $\partial Y/\partial s\,|_{s=s_0} = 0$.

The coefficients γ_α, δ_α can be calculated by summing all diagrams which may be obtained by inserting all possible numbers of $\alpha = (0, 0)$ vertices between two $\alpha \neq (0, 0)$ vertices. It is straightforward to show that all such diagrams are included when one replaces the bare propagator by the Gorkov propagator.[10]

$$G(\mathbf{p}, l) = \frac{1}{\beta}\frac{\mathcal{E}(\mathbf{p}) + \pi i l/\beta}{\mathcal{E}(\mathbf{p})^2 + \Delta^2 + \pi^2 l^2/\beta^2},$$
$$F(\mathbf{p}, l) = \frac{1}{\beta}\frac{\Delta}{\mathcal{E}(\mathbf{p})^2 + \Delta^2 + \pi^2 l^2/\beta^2}, \quad (19)$$

with $\Delta^2 = g_0 s_0(g_0)/\beta$. Note $F^\dagger = -F$ because of rule (2).

Using these propagators, the coefficients can be expressed by the diagrams in Fig. 2, and we have

$$\gamma_\alpha = g_{\mathbf{Q}}\beta\sum_{\mathbf{p},l} G(\mathbf{p}+\mathbf{Q}, l+m)G(-\mathbf{p}, -l),$$
$$\delta_\alpha = (g_{\mathbf{Q}}g_{-\mathbf{Q}})^{\frac{1}{2}}\beta\sum_{\mathbf{p},l} F(\mathbf{p}+\mathbf{Q}, l+m)F(-\mathbf{p}, -l). \quad (20)$$

Integrating over the $\{x_\alpha\}$ we obtain

$$Z'[g_{\mathbf{Q}}] = \exp(-Y(g_0, s_0))\left(\frac{2\pi}{Y''(g_0, s_0)}\right)^{\frac{1}{2}} \times \prod_{\alpha\neq(0,0)}[(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2]^{-\frac{1}{2}}. \quad (21)$$

We can now compute the off-diagonal correlation function by differentiating with respect to $g_{\mathbf{Q}}$:

$$G^{OD}(\mathbf{Q}) = \frac{1}{\beta}\frac{\partial}{\partial g_{\mathbf{Q}}}\{\ln Z[g_{\mathbf{Q}}]\}_{g_{\mathbf{Q}}=g} \quad (22)$$
$$= \frac{-1}{\beta g}\sum_m \frac{-\gamma_\alpha(1-\gamma_{-\alpha}) - \delta_\alpha^2}{(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2}; \quad \mathbf{Q} \neq 0. \quad (23)$$

The behavior of the off-diagonal correlation function for large separation in position space is determined by the character of $G^{OD}(\mathbf{Q})$ at small Q. Using particle–hole symmetry, we find that γ_α is real and further $\gamma_{-\alpha} = \gamma_\alpha$. If we expand for small Q and m, we get

$$1 - \gamma_\alpha - \delta_\alpha = c_0 Q^2 + c_1 m^2 + \cdots$$
$$G^{OD}(\mathbf{Q}) = \frac{1}{2g\beta}\sum_m (c_0 Q^2 + c_1 m^2 + \cdots)^{-1}. \quad (24)$$

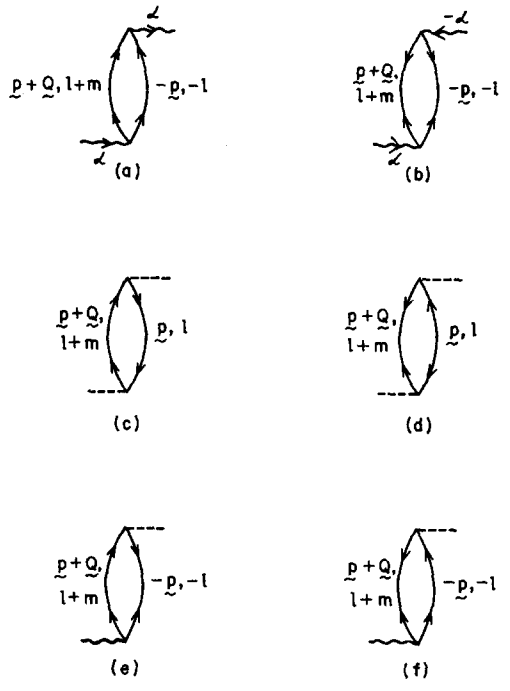

FIG. 2. The graphs included in the generalized RPA. A wavy line denotes a "pairing" interaction and a broken line a "scattering" interaction.

As we stressed above, we are only interested in the dominant term of small Q. Thus we may keep only the $m = 0$ since, for finite temperatures, the sum on m is over discrete points. Thus we find

$$G^{OD}(\mathbf{Q}) = 1/2\beta c_0 g Q^2; \quad \mathbf{Q} \neq 0, \quad (25)$$

where

$$c_0 = \frac{g}{6m^2\beta}\sum_{\mathbf{p},l}\frac{p^2}{(\mathcal{E}(\mathbf{p})^2 + \Delta^2 + \pi^2 l^2/\beta^2)^2}. \quad (26)$$

Turning now to the calculation of the $\mathbf{Q} = 0$ term, we see that differentiating Eq. (21) with respect to g_0 gives two contributions, one from differentiating the exponential and one from differentiating the product, since the coefficients γ_α, δ_α are functions of Δ. Taking the latter first we use the result that

$$\left\{\frac{d}{dg_0} + \sum_{\mathbf{k}\neq 0}\frac{d}{dg_{\mathbf{k}}}\right\} \times \ln\prod_{\alpha\neq(0,0)}[(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2]^{-\frac{1}{2}}\Big|_{g_{\mathbf{k}}=g_0=g}$$
$$= \frac{d}{dg}\left\{\ln\prod_{\alpha\neq(0,0)}[(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2]^{-\frac{1}{2}}\Big|_{g_{\mathbf{Q}}=g_0=g}\right\}. \quad (27)$$

Now the right-hand side of Eq. (27) is easily evaluated, and by Eq. (24) gives

$$\sum_\alpha \frac{c_0' Q^2 + c_1' m^2 + \cdots}{c_0 Q^2 + c_1 m^2 + \cdots}, \quad (28)$$

where

$$c_0' = dc_0/dg, \quad c_1' = dc_1/dg.$$

Thus, using Eqs. (25) and (22), we get

$$\frac{d}{dg_0}\left(\ln \prod_{\alpha\neq(0,0)} [(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2]^{-\frac{1}{2}}\right)\Big|_{g_Q=g_0=g} = -\sum_{\mathbf{k}} \left\{\frac{1}{2c_0 g k^2} + F(\mathbf{k})\right\}, \quad (29)$$

where $F(\mathbf{k})$ is regular as $\mathbf{k} \to 0$. The contribution from differentiating the exponential in Eq. (21) is easily calculated and combining the results we get

$$G^{OD}(\mathbf{Q}) = \frac{\Delta^2}{g^2}\left[\left(1 - \sum_{\mathbf{k}} \frac{g}{2\beta c_0 \Delta^2 k^2}\right)\delta_{\mathbf{Q},0} + \frac{g}{2\beta c_0 \Delta^2 Q^2}\right]. \quad (30)$$

Thus in position space at large separation

$$G^{OD}(X) = \frac{\Delta^2}{g^2}\left[1 - \frac{g}{2\beta c_0 \Delta^2} \sum_{\mathbf{k}} \frac{1 - \cos \mathbf{k}\cdot\mathbf{X}}{k^2}\right]. \quad (31)$$

In I we found that the functional integral over the phase which controls the nature of the long-range correlation gave

$$G^{OD}(X) = \frac{\Delta^2}{g^2} \exp\left\{-\frac{1}{2\beta c \Psi_0^2} \sum_{\mathbf{k}} \frac{1 - \cos \mathbf{k}\cdot\mathbf{x}}{k^2}\right\}. \quad (32)$$

Thus, comparing coefficients for $T \sim T_c$ using Eqs. (A8) and (A9), we see at once that the RPA result, Eq. (31), is just a linearization of the more exact result found in I.

IV. CHARGED SUPERCONDUCTORS

We now consider the effects of including the Coulomb repulsion between the particles. As we stressed in the Introduction, it would appear at first sight that the treatment in I is sufficiently general to cover the charged case since it depends only on the Ginzburg–Landau theory which is known to work well for real superconductors. However, as shown in the Introduction, if we examine the free energy functional more carefully, then we see that it does not include the energy due to charge fluctuations which must, by the conservation laws, accompany an arbitrary $\Psi(\mathbf{x})$.[8] The generalized RPA as developed by Anderson[15] and Thouless[16] does not suffer from this defect since the charge conservation condition is built into this approximation explicitly. We now develop the generalized RPA in our formulation. The approximations made are entirely equivalent to those of Anderson and Thouless, although formally they seem different.

[15] P. W. Anderson, Phys. Rev. **112**, 1900 (1958).
[16] D. J. Thouless, Ann. Phys. (N.Y.) **10**, 553 (1960).

The generalized RPA for a charged superconductor is generated by expanding the functional $Y[x_\alpha, y_\alpha]$ about its minimum position. For temperatures below the transition temperature, the minimum is at

$$|x_{00}|^2 = s_0; \quad x_\alpha = 0, \quad \alpha \neq (0,0); \quad y_\alpha = 0; \quad (33)$$

and we write

$$\begin{aligned} Y[x_\alpha, y_\alpha] &= Y(g_0, s_0) + \tfrac{1}{2}Y''(g_0, s_0)(|x_{0,0}|^2 - s_0)^2 \\ &+ \sum_{\alpha\neq(0,0)} \{(1-\gamma_\alpha)x_\alpha^* x_\alpha + \tfrac{1}{2}\delta_\alpha(x_\alpha^* x_{-\alpha}^* + x_\alpha x_{-\alpha}) \\ &+ (1+\Pi_\alpha)y_\alpha^* y_\alpha + \tfrac{1}{2}\Pi_\alpha(y_\alpha^* y_{-\alpha}^* + y_\alpha y_{-\alpha}) \\ &+ \Phi_\alpha(x_\alpha^*(y_\alpha + y_{-\alpha}^*) + x_\alpha(y_{-\alpha} + y_\alpha^*))\}. \quad (34) \end{aligned}$$

The functions γ_α, δ_α, Π_α, and Φ_α can be written down from Fig. 2 using the rules given above and are most simply expressed using the Gorkov G and F functions. The functions γ_α, δ_α, and $Y(g_0, s_0)$ are the same as before [see Eq. (20)]. We obtain from Fig. 2(c) and (d)

$$\Pi_\alpha = -\beta v(\mathbf{Q}) \sum_{\mathbf{p},l} \{G(\mathbf{p}+\mathbf{Q}, l+m)G(\mathbf{p}, l) + F^\dagger(\mathbf{p}+\mathbf{Q}, l+m)F(\mathbf{p}, l)\}, \quad (35)$$

and from Fig. 2(e) and (f)

$$\Phi_\alpha = -(-\tfrac{1}{2}g_Q v(\mathbf{Q}))^{\frac{1}{2}}\beta \sum_{\mathbf{p},l} \{G(\mathbf{p}+\mathbf{Q}, l+m)F(\mathbf{p}, l) + F(\mathbf{p}+\mathbf{Q}, l+m)G(\mathbf{p}, l)\}. \quad (36)$$

The functional integrals over the $\{x_\alpha\}$, $\{y_\alpha\}$ can be done readily, and we obtain

$$Z[g_Q] = \exp(-Y(g_0, s_0))\left(\frac{2\pi}{Y''(g_0, s_0)}\right)^{\frac{1}{2}} \prod_{\alpha\neq(0,0)} A_\alpha^{-\frac{1}{2}}, \quad (37)$$

where A_α is the determinant

$$A_\alpha = \begin{vmatrix} 1-\gamma_\alpha & \delta_\alpha & \Phi_\alpha & \Phi_\alpha \\ \delta_\alpha & 1-\gamma_{-\alpha} & \Phi_{-\alpha} & \Phi_{-\alpha} \\ \Phi_\alpha & \Phi_{-\alpha} & 1+\Pi_\alpha & \Pi_\alpha \\ \Phi_\alpha & \Phi_{-\alpha} & \Pi_\alpha & 1+\Pi_\alpha \end{vmatrix}. \quad (38)$$

Using Eq. (22) we find for the off-diagonal correlation function

$$G^{OD}(\mathbf{Q}) = \frac{-1}{\beta} \sum_m \left\{\frac{(\partial/\partial g_Q)A_\alpha[g_Q]}{A_\alpha}\right\}_{g_Q=g}; \quad \mathbf{Q} \neq 0. \quad (39)$$

Evaluating the determinant gives

$$A_\alpha = \{(1-\gamma_\alpha)(1-\gamma_{-\alpha}) - \delta_\alpha^2\}(1+2\Pi_\alpha) - 2\Phi_{-\alpha}^2(1-\gamma_\alpha) - 2\Phi_\alpha^2(1-\gamma_{-\alpha}) + 4\delta_\alpha\Phi_\alpha\Phi_{-\alpha}\}. \quad (40)$$

Substituting in Eq. (39) gives, for $\mathbf{Q} \neq 0$,

$$\begin{aligned} G^{OD}(\mathbf{Q}) &= \frac{-1}{\beta g} \sum_m A_\alpha^{-1}\{(-\gamma_\alpha(1-\gamma_{-\alpha}) - \delta_\alpha^2)(1+2\Pi_\alpha) \\ &+ 2\gamma_\alpha\Phi_{-\alpha}^2 - 2\Phi_\alpha^2(1-\gamma_{-\alpha}) + 4\delta_\alpha\Phi_\alpha\Phi_{-\alpha}\}. \quad (41) \end{aligned}$$

It remains to calculate the quantities, γ_α, δ_α, Φ_α, and Π_α. These quantities are in fact the same as those evaluated by Thouless and Tilley[17] and we obtain with

$$
\begin{aligned}
&z = \pi i m/\beta \\
&\gamma_\alpha = \gamma_{-\alpha} = \frac{g}{4}\sum_{\mathbf{p}}\left\{t^{+}(\mathbf{p},z)(E'+E)\left(1+\frac{\mathcal{E}'\mathcal{E}}{E'E}\right)\right. \\
&\qquad\left. + t^{-}(\mathbf{p},z)(E'-E)\left(1-\frac{\mathcal{E}'\mathcal{E}}{E'E}\right)\right\}, \\
&\delta_\alpha = \frac{g}{4}\sum_{\mathbf{p}}\left\{t^{+}(\mathbf{p},z)(E'+E)\frac{\Delta^2}{E'E}\right. \\
&\qquad\left. - t^{-}(\mathbf{p},z)(E'-E)\frac{\Delta^2}{E'E}\right\}, \\
&\Pi_\alpha = \frac{v(\mathbf{Q})}{4}\sum_{\mathbf{p}}\left\{t^{+}(\mathbf{p},z)(E'+E)\left(1-\frac{\mathcal{E}'\mathcal{E}-\Delta^2}{E'E}\right)\right. \\
&\qquad\left. + t^{-}(\mathbf{p},z)(E'-E)\left(1+\frac{\mathcal{E}'\mathcal{E}-\Delta^2}{E'E}\right)\right\}, \\
&\Phi_\alpha = -\Phi_{-\alpha} = -\frac{z\Delta}{4}(-\tfrac{1}{2}v(\mathbf{Q})g)^{\frac{1}{2}} \\
&\times\sum_{\mathbf{p}}\left\{t^{+}(\mathbf{p},z)\left(\frac{1}{E'}+\frac{1}{E}\right)+t^{-}(\mathbf{p},z)\left(\frac{1}{E'}-\frac{1}{E}\right)\right\},
\end{aligned}
\tag{42}
$$

where we have written

$$
t^{\pm}(\mathbf{p},z) = \frac{\tanh(\frac{1}{2}\beta E') \pm \tanh(\frac{1}{2}\beta E)}{(E'\pm E)^2 - z^2} \tag{43}
$$

and

$$
E'^2 = E(\mathbf{p}+\mathbf{Q})^2 = \mathcal{E}(\mathbf{p}+\mathbf{Q})^2 + \Delta^2.
$$

Substituting into Eq. (41) gives us $G^{OD}(\mathbf{Q})$. We discuss the evaluation of $G^{OD}(\mathbf{Q})$ separately for each number of dimensions.

Three-Dimensional System

The sum over m in Eq. (41) can be transformed in the usual manner to a contour integral. This contour integral picks up contributions from the zeros of A_α. The zeros of A_α give the position of the collective modes. In the uncharged system it was not necessary to do this transformation since there are low-lying collective modes for small Q and m

$$
A_\alpha \sim c_0 Q^2 + c_1 m^2 + \cdots .
$$

The $m = 0$ term dominates the sum as $Q \to 0$ since it alone diverges in this limit. All $m \neq 0$ terms were finite as $Q \to 0$. Thus the behavior of $G^{OD}(Q)$ for the uncharged system can be connected directly to the presence of low-lying modes. However, for the three-dimensional charged superconductor, Anderson[15,18] has shown that there are no low-lying modes at $T = 0$. We assume that this is so at finite temperatures and it can be shown that temperature dependence of the plasmon mode that Anderson found is negligible. In this case the behavior of A_α for small Q and m is now drastically changed and it is necessary to evaluate the contour integral. The dominant contribution at small Q comes from the zero-point motion of the plasmon mode. There is also a contribution from the collective modes due to oscillations of the magnitude of gap, but this is finite as $Q \to 0$. Evaluating the residue of the plasmon modes, we find the dominant contribution at small Q comes from the $2\Phi^2$ term in the numerator of Eq. (41) and, neglecting terms of $O(\omega_D/\omega_P)$ and $O((\beta\omega_P)^{-1})$, we get

$$
G^{OD}(\mathbf{Q}) = \frac{\Delta^2}{g^2}\left[\left(1-\sum_{\mathbf{k}}\frac{8\pi e^2}{\omega_P k^2}\right)\delta_{\mathbf{Q},0} + \frac{8\pi e^2}{\omega_P Q^2}\right], \tag{44}
$$

where the first term is calculated in a similar manner to Eqs. (27)–(30). ω_P is the plasma frequency.

Comparing this form for $G^{OD}(\mathbf{Q})$ with that found for the uncharged system in the RPA Eq. (26), we see that it has the same behavior at small Q. There we found that the RPA linearized the more correct answer found from Ginzburg–Landau. Thus we expect a more correct theory to give

$$
G^{OD}(X) = \frac{\Delta^2}{g^2}\exp\left\{-\frac{8\pi e^2}{\omega_P}\sum_{\mathbf{Q}}^{Q<Q_M}\frac{1-\cos\mathbf{Q}\cdot\mathbf{X}}{Q^2}\right\} \tag{45}
$$

$$
\sim \frac{\Delta^2}{g^2}\exp\left[-(Q_M/2\pi - 1/X)2e^2/\omega_P\right]; \qquad X \to \infty, \tag{46}
$$

where Q_M is an upper cut-off whose magnitude is determined by the terms omitted in the expansion in Eq. (44). This form for $G^{OD}(X)$ of course exhibits ODLRO. The behavior of $G^{OD}(\mathbf{Q})$ for small Q in the charged system is similar to that in the uncharged system, although we have no low-lying modes. The divergence at small Q for the charged system comes from the residue of the plasmon mode rather than the modes' dispersion relation.

In order to understand physically what is going on, it is instructive to ask how the phase of the gap parameter varies in a superconductor in the presence of a charge fluctuation. Consider a localized charge

[17] D. J. Thouless and D. R. Tilley, Proc. Phys. Soc. (London) 77, 1175 (1961); 80, 320 (1962).

[18] *Note added in proof.* Quite recently P. C. Hohenberg [Phys. Rev. 158, 383 (1967)] has given a very elegant and general proof of the absence of ODLRO in one- and two-dimensional systems at finite temperatures. This proof is applicable to both charged and uncharged superconductors.

oscillation at the origin which we represent by an electromagnetic potential

$$\varphi_{\text{ext}}(\mathbf{k}, \omega) = (4\pi e^2/k^2)\,\delta(\omega - \omega_v). \tag{47}$$

Then the Fourier transform of the phase of the order parameter $W(\mathbf{k}, \omega)$ is given by [see Ambegaokar and Kadanoff,[19] Eq. (4.9)]

$$W(\mathbf{k}, \omega) = 4\pi e^2\omega\delta(\omega - \omega_0)/k^2(\omega_P^2 - \omega_0^2). \tag{48}$$

Thus the phase difference is

$$W(\mathbf{X}) - W(0) = \sum_{\mathbf{k}} - \frac{4\pi e^2\omega_0(1 - \cos \mathbf{k}\cdot\mathbf{X})}{k^2(\omega_P^2 - \omega_0^2)} \tag{49}$$

$$\sim - \frac{e^2\omega_0}{\omega_P^2 - \omega_0^2}\left(\frac{k_{\max}}{2\pi} - \frac{1}{X}\right). \tag{50}$$

Thus we see that, whereas in an uncharged system a localized density fluctuation gives rise only to a localized phase fluctuation, in a charged system the phase fluctuation falls off slowly. We attribute the behavior of $G^{OD}(\mathbf{Q})$ for small Q in the charged system to the zero-point motion of the plasmon modes. We see that the presence of a long-range force does not alter the character of $G^{OD}(X)$ at large X because, although such an interaction forces the density modes up to the plasmon frequency, it also causes localized density fluctuations to have a long-range effect on the phase.

One- and Two-Dimensional Charged Systems

We begin by defining more exactly what we mean by a one- or two-dimensional charged system. For an uncharged system it sufficed to define the dimensionality of the system purely in terms of the dimensionality of the sum over $\mathbf{Q}$; but this prescription breaks down for a charged system since the one-dimensional Fourier transform of the Coulomb potential is not well defined. It is necessary to take a specific model, and we take as our model in one dimension a system of electrons constrained to move inside a long, very thin cylinder, and for a two-dimensional system a very thin film. Thus, in one dimension, consider the electrons as moving in a potential

$$V(x, y, z) = \tfrac{1}{2}m\omega^2(x^2 + y^2). \tag{51}$$

The energy eigenvalues of a system of noninteracting electrons moving in such a potential are given by

$$\mathcal{E}_{n,k_z} = (k_z^2/2m) + (n + \tfrac{1}{2})\omega. \tag{52}$$

We assume that ω is very large, so large in fact that we need never consider the higher transverse states, i.e., we keep only the $n = 0$ term in all sums over the transverse wavefunctions. The wavefunctions associated with this system are

$$\Psi_{0\,k_z}(\mathbf{r}) = \varphi_0(x, y)L^{-\frac{1}{2}}\exp(ik_z z), \tag{53}$$

where L is the length of the system and φ_0 is lowest oscillator function

$$\varphi_0(x, y) = (\nu/\pi)\exp\{-\tfrac{1}{2}\nu^2(x^2 + y^2)\} \tag{54}$$

with $\nu = m\omega$.

We now switch on two different interactions between the electrons, an attractive δ function interaction and a repulsive Coulomb interaction. Migdal[20] has extended the Gorkov theory to nonuniform systems. In this case, the extension is particularly easy since we have only one transverse level. The derivation of the functional integral in this case is straightforward and the rules for evaluating the graphs are the same, except that the three-dimensional momentum vector $\mathbf{p}$ is replaced by the one-dimensional vector p_z, the transverse label $n = 0$ being understood throughout. The interaction matrix elements $g_{\mathbf{Q}}$ and $v(\mathbf{Q})$ are replaced by the matrix elements g'_{Q_z}, $v'(Q_z)$, where we again introduce a momentum dependence to the coupling constant g to facilitate the calculation of the off-diagonal correlation function. The new matrix elements are given by

$$g' = g\int dx\,dy\,\varphi_0^4(x, y) \tag{55}$$

and

$$v'(Q_z) = \int d^3r\,d^3r'\Psi^*_{0,k_z}(\mathbf{r}')\Psi^*_{0,Q_z+k_z}(\mathbf{r}) \times \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}\Psi_{0,k_z+Q_z}(\mathbf{r}')\Psi_{0,k_z}(\mathbf{r}), \tag{56}$$

$$= -(e^2/\pi L)\log(Q_z/\sqrt{2}\,\nu) + \cdots, \tag{57}$$

where we keep the dominant term for small Q_z. Below the "transition" temperature we again introduce the Gorkov G and F functions,

$$G_0(k_z, l) = \frac{1}{\beta}\frac{\mathcal{E}_0(k_z) + \pi i l/\beta}{\mathcal{E}_0^2(k_z) + \Delta_0^2 + \pi^2 l^2/\beta^2} \tag{58}$$

and

$$F_0(k_z, l) = \frac{1}{\beta}\frac{\Delta_0}{\mathcal{E}_0^2(k_z) + \Delta_0^2 + \pi^2 l^2/\beta^2}, \tag{59}$$

where $\mathcal{E}_0(k_z) = \mathcal{E}_{0,k_z} - \mu$ and Δ_0 is determined by the equation

$$\Delta_0 = \frac{g'}{2}\sum_{k_z}\frac{\Delta_0\tanh[\frac{1}{2}\beta E_0(k_z)]}{E_0(k_z)}. \tag{60}$$

We now examine the collective modes of the system. In the normal phase it is shown in Appendix B that this system has low-lying modes. The dispersion

[19] V. Ambegaokar and L. P. Kadanoff, Nuovo Cimento 22, 914 (1961).

[20] A. B. Migdal, Nucl. Phys. 13, 655 (1959).

relation in the limit as $k_z \to 0$ is

$$\omega_{k_z} = k_z\left\{\frac{N}{L}\frac{e^2}{m}[-\log(k_z/\sqrt{2}\,v)]\right\}^{\frac{1}{2}} + \cdots. \quad (61)$$

These modes are similar to those discussed by Ferrell.[2] Below the "transition" temperature the zeros of A_α determine the modes and we show in Appendix B that these modes persist. To calculate the off-diagonal correlation function we need to evaluate the sum over m in Eq. (41). Because of the low-lying modes the dominant term at small Q_z in the sum for $T \neq 0$ will, as in the uncharged case, be the $m = 0$ term. But $\Phi_{Q_z,m=0} = 0$ from Eq. (42), so that in the one-dimensional system

$$G^{OD}(Q_z) = \frac{1}{\beta g}\frac{-\gamma'(Q_z,0)(1-\gamma'(Q_z,0)) - \delta'^2(Q_z(0)}{(1-\gamma'(Q_z,0))^2 - \delta'^2(Q_z,0)}, \quad (62)$$

and all effects of the Coulomb interaction have canceled out in this approximation. The result is identical to that which we found in the RPA for the uncharged system. The RPA as we discussed above gives a linearization of the more exact results found in I. Thus we argue that, in view of the equivalence of the results found in the RPA for the charged and uncharged system, a more correct theory of the charged system would give the results found in I.

The two-dimensional charged system is similar. There are low-lying collective modes which persist below the "transition" temperature. Because of these, the terms which couple in the Coulomb interaction again do not contribute to the dominant term at small Q_z and the result for $G^{OD}(Q_z)$ in the RPA is identical to the uncharged system.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor W. Kohn for suggesting the problem and for his advice and to Dr. K. Sawada for useful conversations.

This work was supported by the U.S. Office of Naval Research Contract NONR-2216(11) and U.S. Air Force Grant No. AF-AFOSR-610-64.

APPENDIX A

In this Appendix we show how the procedure used in the text to obtain the Ginzburg–Landau functional for temperatures near the transition temperature can be generalized to all temperatures to obtain the Werthamer[11] functional. The functional $Y[x_{Q,m}]$ is written first as a functional in position space. The "time"-independent approximation is made and all terms with $m \neq 0$ are discarded. Then a gradient expansion is carried out. We confine our attention to the field-free case, although our results can be generalized to include a static external vector potential. In order to demonstrate that the resulting functional is the Werthamer functional, it is sufficient to consider just the nearly uniform situation. Werthamer's functional for the free energy for the field-free case can be written as

$$\mathcal{F}[\Delta(\mathbf{R})] = \int d^3\mathbf{R}N(0)\Big[w(|\Delta(\mathbf{R})|^2) + \tfrac{1}{6}v_F^2\left|\frac{\partial}{\partial\mathbf{R}}\Delta(\mathbf{R})\right|^2 w''(|\Delta(\mathbf{R})|^2) + \frac{v_F^2}{36}\left(\frac{\partial}{\partial\mathbf{R}}|\Delta(\mathbf{R})|^2\right)^2 w'''(|\Delta(\mathbf{R})|^2)\Big)\Big], \quad (A1)$$

where $N(0)$ is the density of states at the Fermi surface, and

$$w(x) = -\int_{-\infty}^{+\infty} d\epsilon\left[\frac{1}{\beta}\ln\frac{1+\cosh\beta(\epsilon^2+x^2)^{\frac{1}{2}}}{1+\cosh\beta\epsilon} - x\frac{\tanh\frac{1}{2}\beta_c\epsilon}{2\epsilon}\right]. \quad (A2)$$

We could now proceed to check that the functional obtained by a gradient expansion of $Y[\Delta(\mathbf{R})]$ is identical to $\beta\mathcal{F}$; but, in fact, it is not necessary to calculate this expansion directly since Langer's[10] expansion of Y in momentum space about the minimum value can be used to derive the coefficients. Langer [Eq. (4.14)] has shown that

$$Y[\Delta_q] = Y(|\Delta_0|) + \left(\frac{\beta}{g}\right)\sum_{q\neq 0}\{(1-\gamma_q(|\Delta_0|))\Delta_q^*\Delta_q + \tfrac{1}{2}\delta_q(|\Delta_0|)(\Delta_0^2\Delta_q^*\Delta_{-q}^* + \text{c.c.})/|\Delta_0|^2\}, \quad (A3)$$

where we have written

$$\Delta_q = (g/\beta)^{\frac{1}{2}}t_q\exp(i\varphi_q)$$

and the coefficients are given by

$$\gamma_q(|\Delta_0|) = \frac{g}{2\beta}\sum_{\mathbf{p},l,\sigma}\frac{(\mathcal{E}(\mathbf{p}) - \pi il/\beta)(\mathcal{E}(\mathbf{p}+\mathbf{q}) + \pi il/\beta)}{(\mathcal{E}(\mathbf{p})^2 + \pi^2l^2/\beta^2 + |\Delta_0|^2)(\mathcal{E}(\mathbf{p}+\mathbf{Q})^2 + \pi^2l^2/\beta^2 + |\Delta_0|^2)} \quad (A4)$$

and

$$\delta_q(|\Delta_0|) = \frac{g|\Delta_0|^2}{2\beta}\sum_{\mathbf{p},l,\sigma}\frac{1}{(\mathcal{E}(\mathbf{p})^2 + \pi^2l^2/\beta^2 + |\Delta_0|^2)(\mathcal{E}^2(\mathbf{p}+\mathbf{Q}) + \pi^2l^2/\beta^2 + |\Delta_0|^2)}. \quad (A5)$$

This can be rewritten as

$$Y[\Delta_q] = Y(|\Delta_0|) + \frac{\beta}{g}\sum_{q\neq 0}\{(1-\gamma_q(|\Delta_0|)-\delta_q(|\Delta_0|))\Delta_q^*\Delta_q + \tfrac{1}{2}\delta_q(|\Delta_0|)(\Delta_0^2\Delta_q^*\Delta_{-q}^* + 2|\Delta_0|^2\Delta_q^*\Delta_q + \Delta_0^{*2}\Delta_q\Delta_{-q})/|\Delta_0|^2\}. \quad \text{(A6)}$$

Expanding the coefficient in powers of $\mathbf{q}$, we write

$$1-\gamma_q(|\Delta_0|)-\delta_q(|\Delta_0|) = \frac{g}{4\beta}\sum_{p,l,\sigma}\frac{(\mathcal{E}(\mathbf{p}+\mathbf{Q})-\mathcal{E}(\mathbf{p}))^2}{(\mathcal{E}(\mathbf{p})^2+\pi^2l^2/\beta^2+|\Delta_0|^2)(\mathcal{E}(\mathbf{p}+\mathbf{Q})^2+\pi^2l^2/\beta^2+|\Delta_0|^2)} = c_0(|\Delta_0|)q^2+\cdots, \quad \text{(A7)}$$

$$\delta_q(|\Delta_0|) = \delta_0(|\Delta_0|)+\delta_2(|\Delta_0|)q^2+\cdots. \quad \text{(A8)}$$

Transforming to position space, we obtain

$$Y[\Delta(\mathbf{R})] = Y(|\Delta_0|)+\frac{\beta}{g}c_0(|\Delta_0|)\int\left|\frac{\partial}{\partial\mathbf{R}}\Delta(\mathbf{R})\right|^2 d\mathbf{R} + \frac{\beta\delta_0(|\Delta_0|)}{2g|\Delta_0|^2}\int(\Delta_0^2\Delta^{*2}(\mathbf{R}) + 2|\Delta_0|^2|\Delta(\mathbf{R})|^2+\Delta_0^{*2}\Delta^2(\mathbf{R}))\,d\mathbf{R} + \frac{\beta}{2g}\frac{\delta_2(|\Delta_0|)}{|\Delta_0|^2}\int\left\{\Delta_0^2\left(\frac{d}{d\mathbf{R}}\Delta^*(\mathbf{R})\right)^2 + 2|\Delta_0|^2\left|\frac{d}{d\mathbf{R}}\Delta(\mathbf{R})\right|^2\right\}d\mathbf{R}. \quad \text{(A9)}$$

Comparing with the free energy functional $\mathcal{F}$, we see that Y has the same form as $\beta\mathcal{F}$ in the nearly uniform case, i.e., one writes $\Delta(\mathbf{R}) = \Delta_0 + \Delta'(\mathbf{R})$ and expands in $\Delta'(\mathbf{R})$. It remains to show that the coefficients are the same. After some algebra we find

$$Y(|\Delta_0|) = \frac{\beta}{g}|\Delta_0|^2 - 2\sum_{\mathbf{p}}\ln\left\{\frac{\cosh[(\beta/2)(\mathcal{E}(\mathbf{p})^2+|\Delta_0|^2)^{\frac{1}{2}}]}{\cosh\frac{1}{2}\beta\mathcal{E}(\mathbf{p})}\right\} = \beta w(|\Delta_0|^2),$$

$$\frac{\beta}{g}c_0(|\Delta_0|) = \frac{1}{4}\sum_{p,l,\sigma}\frac{(\mathbf{v}\cdot\hat{\mathbf{q}})^2}{(\mathcal{E}(p)^2+\pi^2l^2/\beta^2+|\Delta_0|^2)^2} = \tfrac{1}{6}\beta v_F^2 w''(|\Delta_0|^2),$$

$$\frac{\beta\delta_0(|\Delta_0|)}{2g|\Delta_0|^2} = \frac{1}{4}\sum_{p,l,\sigma}\frac{1}{(\mathcal{E}(\mathbf{p})^2+\pi^2l^2/\beta^2+|\Delta_0|^2)^2} = \tfrac{1}{2}\beta w''(|\Delta_0|^2),$$

$$\frac{\beta\delta_2(|\Delta_0|)}{2g|\Delta_0|^2} = \frac{1}{4}\sum_{p,l,\sigma}\frac{(\mathbf{v}\cdot\hat{\mathbf{q}})^2(3\mathcal{E}^2(p)-\pi^2l^2/\beta^2-|\Delta_0|^2)}{(\mathcal{E}(\mathbf{p})^2+\pi^2l^2/\beta^2+|\Delta_0|^2)^4} = \beta v_F^2 w'''(|\Delta_0|^2)/36. \quad \text{(A10)}$$

Thus we have shown that the two functionals are equivalent and that a gradient expansion of $Y[\Delta(\mathbf{R})]$ yields Werthamer's generalization of the Ginzburg–Landau free energy functional.

APPENDIX B. COLLECTIVE MODES OF THE ONE- AND TWO-DIMENSIONAL CHARGED SYSTEM

In this Appendix we calculate the collective density modes of one- and two-dimensional model systems. Consider first the normal phase of the one-dimensional system. Then, in the RPA, the dispersion relation is determined by the eigenvalue equation

$$1 = (v'(k_z)+g')s_0(k_z,\omega), \quad \text{(B1)}$$

where

$$s_0(k_z,\omega) = \sum_{q_z,\sigma}\int\frac{d\epsilon}{2\pi i}G_0(k_z+q_z,\epsilon+\omega)G_0(q_z,\epsilon)$$

and $G_0(q_z,\epsilon)$ is the single-particle Green's function. At small k_z the dominant contribution is from the logarithmic term in $v'(k_z)$. Thus, using Eq. (57), we get

$$1 = -\frac{e^2}{\pi L}\log\left(\frac{k_z}{\sqrt{2}\,\nu}\right)\sum_{q_z,\sigma}\frac{n_0(k_z+q_z)-n_0(q_z)}{\omega-\mathcal{E}_0(k_z+q_z)+\mathcal{E}_0(q_z)} \quad \text{(B2)}$$

$$= -\frac{e^2}{\pi^2}\log\left(\frac{k_z}{\sqrt{2}\,\nu}\right)\frac{v_Fk_z^2}{\omega^2-v_F^2k_z^2}, \quad \text{(B3)}$$

where v_F is the Fermi velocity. Solving for ω, we get

$$\omega^2 = v_F^2k_z^2\{(e^2/\pi^2v_F)[-\log(k_z/\sqrt{2}\,\nu)]+1\}. \quad \text{(B4)}$$

Thus at small k_z the first term dominates, and we obtain

$$\omega = k_z\{(N/L)(e^2/m)[-\log(k_z/\sqrt{2}\,\nu)]\}^{\frac{1}{2}}+\cdots. \quad \text{(B5)}$$

Next we consider the two-dimensional charged system. In this system we again consider only one transverse energy level, and it is straightforward to generalize the treatment given above for the one-dimensional system. The new matrix elements are g'' and $v''(k_\parallel)$ where g'' differs from g only by a constant and

$$v''(k_\parallel) = 2\pi e^2/L^2k_\parallel. \quad \text{(B6)}$$

Then we find, on solving the eigenvalue Eq. (B1) for

the normal two-dimensional system, that

$$\omega(k_{\|}) = [N/L^2)(e^2k_{\|}/m)]^{\frac{1}{2}} + \cdots. \tag{B7}$$

We now turn to the calculation of the dispersion relation below the "transition" temperature. The dispersion relation is determined by the zeros of the determinant A_α. We consider only $T = 0$ since the additional terms in Eq. (40) which enter at nonzero temperature are nonanalytic and cannot be expanded about $\omega = 0$, $k_{\|} = 0$. At $T = 0$ we have to solve for the zeros of Eq. (40), namely,

$$A_\alpha = (1 - \gamma'_\alpha + \delta'_\alpha)\{(1 - \gamma'_\alpha - \delta'_\alpha)(1 + 2\prod'_\alpha) - 4\Phi'^2_\alpha\}, \tag{B8}$$

where the prime denotes that we have replaced the interaction g and $v(\mathbf{Q})$ by the appropriate primed matrix elements given in Eqs. (55), (57), and (B6). The zeros of $1 - \gamma'_\alpha + \delta'_\alpha$ correspond to oscillations in the magnitude of the gap and these oscillations all start at a finite frequency, namely, at twice the gap. Using Eq. (42), we get, in one-dimension with $\omega = \pi i m/\beta$,

$$1 - \gamma'(\omega, k_z) - \delta'(\omega, k_z) = -\frac{g'\omega^2}{8}\sum_{p_z}\frac{1}{E^3_{p_z}} + \frac{g'}{8}\sum_{p_z}\frac{(\mathcal{E}(p_z + k_z) - \mathcal{E}_0(p_z))^2}{E^3_{p_z}} + \cdots,$$

$$1 + 2\prod'(\omega, k_z) = 1 + \Delta_0^2 v'(k_z)\sum_{p_z}\frac{1}{E^3_{p_z}} + \cdots, \tag{B9}$$

$$4\Phi^2(\omega, k_z) = -\frac{\omega^2\Delta_0^2 v'(k_z)g'}{8}\left(\sum_{p_z}\frac{1}{E^3_{p_z}}\right)^2 + \cdots.$$

Solving for the dispersion relation, we get

$$\omega^2 = v_F^2k_z^2 + k_z^2v_F^2v'(k_z)\Delta_0^2\left(\sum_{p_z}\frac{1}{E^3_{p_z}}\right) + \cdots. \tag{B10}$$

Using Eq. (58) for $v'(k_z)$, we see that the second term dominates at small k_z, but nevertheless there are low-lying collective modes. We have not been able to demonstrate explicitly the existence of these modes for temperatures $0 < T < T_c$, but it is reasonable to assume that they are present at these temperatures. Similarly, one finds low-lying modes for the two-dimensional system below the "transition" temperature.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Description of Extended Bodies by Multipole Moments in Special Relativity*

W. G. Dixon†
Department of Physics and Astronomy, University of Maryland, College Park, Maryland

(Received 23 September 1966)

The description of an extended charged body in a given electromagnetic field in flat space-time is considered, and it is shown that such a body may be completely specified by a certain set of multipole moments of the energy-momentum tensor $T^{\alpha\beta}$ and the charge-current vector J^α. These moments include the momentum vector, spin tensor, and total charge of the body, and they completely determine $T^{\alpha\beta}$ and J^α. It is shown that the only relations between the moments due to the "generalized conservation equations" $\partial_\beta T^{\alpha\beta} = -F^{\alpha\beta}J_\beta$ and $\partial_\alpha J^\alpha = 0$ are the constancy of the total charge and equations of motion for the momentum vector and spin tensor, in contrast to previous descriptions by moments, such as that of Mathisson, which have an infinite number of such relations. The equations of motion are given exactly, as infinite series in the moments, without assuming the applied electromagnetic field to be analytic, and an approximation procedure is developed, based on the smallness of the body compared with a typical length scale for the external field.

1. INTRODUCTION

In both the special and the general theories of relativity, the basic description of charged matter moving in an electromagnetic field $F^{\alpha\beta}$ is given by an energy-momentum tensor $T^{\alpha\beta}$ and a charge-current vector J^α satisfying

$$\nabla_\beta T^{\alpha\beta} = -F^{\alpha\beta}J_\beta \tag{1.1}$$

and

$$\nabla_\alpha J^\alpha = 0. \tag{1.2}$$

But, for describing the motion of a body, it is more convenient to use a set of multipole moments of $T^{\alpha\beta}$ and J^α rather than these tensors themselves. In the present paper we first discuss various definitions that have been proposed for such moments of extended bodies, and then for the case of flat space-time we define a new set of multipole moments that avoids some of the defects of the earlier definitions. A

* Work supported by the National Aeronautical and Space Administration under Contract NsG-436.

† Present address: Churchill College, Cambridge, England.

the normal two-dimensional system, that

$$\omega(k_{\|}) = [N/L^2)(e^2k_{\|}/m)]^{\frac{1}{2}} + \cdots. \tag{B7}$$

We now turn to the calculation of the dispersion relation below the "transition" temperature. The dispersion relation is determined by the zeros of the determinant A_α. We consider only $T = 0$ since the additional terms in Eq. (40) which enter at nonzero temperature are nonanalytic and cannot be expanded about $\omega = 0$, $k_{\|} = 0$. At $T = 0$ we have to solve for the zeros of Eq. (40), namely,

$$A_\alpha = (1 - \gamma'_\alpha + \delta'_\alpha)\{(1 - \gamma'_\alpha - \delta'_\alpha)(1 + 2\prod'_\alpha) - 4\Phi'^2_\alpha\}, \tag{B8}$$

where the prime denotes that we have replaced the interaction g and $v(\mathbf{Q})$ by the appropriate primed matrix elements given in Eqs. (55), (57), and (B6). The zeros of $1 - \gamma'_\alpha + \delta'_\alpha$ correspond to oscillations in the magnitude of the gap and these oscillations all start at a finite frequency, namely, at twice the gap. Using Eq. (42), we get, in one-dimension with $\omega = \pi i m/\beta$,

$$1 - \gamma'(\omega, k_z) - \delta'(\omega, k_z) = -\frac{g'\omega^2}{8}\sum_{p_z}\frac{1}{E^3_{p_z}} + \frac{g'}{8}\sum_{p_z}\frac{(\mathcal{E}(p_z + k_z) - \mathcal{E}_0(p_z))^2}{E^3_{p_z}} + \cdots,$$

$$1 + 2\prod'(\omega, k_z) = 1 + \Delta_0^2 v'(k_z)\sum_{p_z}\frac{1}{E^3_{p_z}} + \cdots, \tag{B9}$$

$$4\Phi^2(\omega, k_z) = -\frac{\omega^2\Delta_0^2 v'(k_z)g'}{8}\left(\sum_{p_z}\frac{1}{E^3_{p_z}}\right)^2 + \cdots.$$

Solving for the dispersion relation, we get

$$\omega^2 = v_F^2k_z^2 + k_z^2v_F^2v'(k_z)\Delta_0^2\left(\sum_{p_z}\frac{1}{E^3_{p_z}}\right) + \cdots. \tag{B10}$$

Using Eq. (58) for $v'(k_z)$, we see that the second term dominates at small k_z, but nevertheless there are low-lying collective modes. We have not been able to demonstrate explicitly the existence of these modes for temperatures $0 < T < T_c$, but it is reasonable to assume that they are present at these temperatures. Similarly, one finds low-lying modes for the two-dimensional system below the "transition" temperature.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Description of Extended Bodies by Multipole Moments in Special Relativity*

W. G. Dixon†
Department of Physics and Astronomy, University of Maryland, College Park, Maryland

(Received 23 September 1966)

The description of an extended charged body in a given electromagnetic field in flat space-time is considered, and it is shown that such a body may be completely specified by a certain set of multipole moments of the energy-momentum tensor $T^{\alpha\beta}$ and the charge-current vector J^α. These moments include the momentum vector, spin tensor, and total charge of the body, and they completely determine $T^{\alpha\beta}$ and J^α. It is shown that the only relations between the moments due to the "generalized conservation equations" $\partial_\beta T^{\alpha\beta} = -F^{\alpha\beta}J_\beta$ and $\partial_\alpha J^\alpha = 0$ are the constancy of the total charge and equations of motion for the momentum vector and spin tensor, in contrast to previous descriptions by moments, such as that of Mathisson, which have an infinite number of such relations. The equations of motion are given exactly, as infinite series in the moments, without assuming the applied electromagnetic field to be analytic, and an approximation procedure is developed, based on the smallness of the body compared with a typical length scale for the external field.

1. INTRODUCTION

In both the special and the general theories of relativity, the basic description of charged matter moving in an electromagnetic field $F^{\alpha\beta}$ is given by an energy-momentum tensor $T^{\alpha\beta}$ and a charge-current vector J^α satisfying

$$\nabla_\beta T^{\alpha\beta} = -F^{\alpha\beta}J_\beta \tag{1.1}$$

and

$$\nabla_\alpha J^\alpha = 0. \tag{1.2}$$

But, for describing the motion of a body, it is more convenient to use a set of multipole moments of $T^{\alpha\beta}$ and J^α rather than these tensors themselves. In the present paper we first discuss various definitions that have been proposed for such moments of extended bodies, and then for the case of flat space-time we define a new set of multipole moments that avoids some of the defects of the earlier definitions. A

* Work supported by the National Aeronautical and Space Administration under Contract NsG-436.

† Present address: Churchill College, Cambridge, England.

summary of the notational conventions used in the present section is given in Sec. 2.

Consider first an extended body in a curved space–time with $F^{\alpha\beta} = 0$. Then $T^{\alpha\beta}$ is a function of class C^1 whose support, i.e., the closure of the set of points on which it is nonzero, is a world-tube W whose intersection with every spacelike geodesic is bounded. This property of W, for brevity, will be called *spacewise boundedness*. Choose any timelike world-line L lying in W and let its parametric equation be $x^\alpha = z^\alpha(s)$, where s is the proper time measured along it. Let $\Sigma(s)$ be the hypersurface generated by all geodesics through $z^\alpha(s)$ orthogonal to $v^\alpha(s) \overset{\text{def}}{=} dz^\alpha/ds$. Then, as was shown by Bielecki, Mathisson, and Weyssenhoff,[1] there exist unique tensor fields $t^{\gamma_1\cdots\gamma_n\alpha\beta}(s)$ defined along L for each $n \geq 0$ and satisfying

$$t^{\gamma_1\cdots\gamma_n\alpha\beta} = t^{(\gamma_1\cdots\gamma_n)(\alpha\beta)} \tag{1.3}$$

and

$$v_{\gamma_1} t^{\gamma_1\cdots\gamma_n\alpha\beta} = 0, \tag{1.4}$$

which are such that

$$\int T^{\alpha\beta} p_{\alpha\beta} (-g)^{\frac{1}{2}}\, d^4x = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{1}{n!}\, t^{\gamma_1\cdots\gamma_n\alpha\beta}(s) [\nabla_{\gamma_1\cdots\gamma_n} p_{\alpha\beta}]_{z(s)}\, ds \tag{1.5}$$

for any C^∞ tensor field $p_{\alpha\beta}$ of compact support whose Taylor series expansion about $z^\alpha(s)$ is valid on $\Sigma(s) \cap W$. It will be shown in a later paper that the coefficients $t^{\gamma_1\cdots\gamma_n\alpha\beta}(s)$ are identical with the quantities having the same notation which were defined by Dixon[2] as explicit integrals of $T^{\alpha\beta}$ over $\Sigma(s)$. In the present section we call these the multipole moments of the body, but in subsequent sections we reserve this name for other quantities defined later. They can be shown to determine $T^{\alpha\beta}$ completely, as we see for the case of flat space–time in Sec. 5. This is contrary to a statement made in Ref. 2, and I am indebted to Dr. J. Ehlers for this correction.

As $F^{\alpha\beta} = 0$, we have

$$\nabla_\beta T^{\alpha\beta} = 0. \tag{1.6}$$

So, given any vector field k_α satisfying the same conditions as imposed on $p_{\alpha\beta}$, if we choose

$$p_{\alpha\beta} = \nabla_\alpha k_\beta + \nabla_\beta k_\alpha, \tag{1.7}$$

then the left-hand side of (1.5) vanishes identically. We thus have the identity

$$\int \sum_{n=0}^{\infty} \frac{1}{n!}\, t^{\gamma_1\cdots\gamma_n\alpha\beta}(s) [\nabla_{\gamma_1\cdots\gamma_n\alpha} k_\beta]_{z(s)}\, ds \equiv 0, \tag{1.8}$$

due to Mathisson,[3,4] who calls it the *variational equation of dynamics*. It is equivalent to an infinite system of ordinary differential equations relating the moments, which is usually treated by the following process of approximation. We suppose that the diameter d of the body is small compared with a characteristic length scale R of the external gravitational (i.e., metric) field. If we assume that the 2^n-pole moment $t^{\gamma_1\cdots\gamma_n\alpha\beta}$ interacts only with n-fold and higher derivatives of the field, this interaction should be of order $(d/R)^n$, and thus negligible for large n. It then should be a good approximation to treat all but a finite number of the moments as zero, and thus to terminate (1.8) at, say, $n = N$. This gives a system of $(N + 1)$ equations relating the $(N + 1)$ nonvanishing moments, which may be obtained by the method of Mathisson.[3] It is there applied to the cases $N = 1$ and $N = 2$, but with restrictions on the form of the moments in both cases.

Tulczyjew[5] simplified Mathisson's procedure by noting that the truncated variational equation is equivalent to requiring the tensor distribution

$$*T^{\alpha\beta}(x) \overset{\text{def}}{=} \sum_{n=0}^{N} \frac{(-1)^n}{n!}\, \nabla_{\gamma_1\cdots\gamma_n} \int t^{\gamma_1\cdots\gamma_n\alpha\beta}(s) \delta(z(s), x)\, ds \tag{1.9}$$

to satisfy

$$\nabla_\beta *T^{\alpha\beta} = 0. \tag{1.10}$$

He treats the pole–dipole approximation ($N = 1$) with no additional restrictions, and, from the form of the resulting equations, identifies certain quantities constructed from $t^{\alpha\beta}$ and $t^{\alpha\beta\gamma}$ with the momentum vector p^α and internal angular momentum (spin) tensor $S^{\alpha\beta}$ of the body, showing that they agree with the usual definitions in the case of flat space–time. Such an identification is open to ambiguity, especially when higher approximations are considered, and this procedure makes p^α and $S^{\alpha\beta}$ depend on the order of approximation used. This is unsatisfactory if p^α and $S^{\alpha\beta}$ are used in defining the center of mass of the body, as is further discussed below. This difficulty is avoided by the author's treatment in Ref. 2 by using the equivalent integral expressions for the moments, and separately defining p^α and $S^{\alpha\beta}$ as integrals of $T^{\alpha\beta}$ which can be expressed in terms of the moments to any required degree of approximation. This is not so in the earlier theories that use explicit integral expressions for the moments, due to Papapetrou,[6] to Urich and

[1] A. Bielecki, M. Mathisson, and J. W. Weyssenhoff, Bull. Intern. Acad. Polon. Sci. Lett., Cl. Sci. Math. Nat. Sér. A: Sci. Math. p. 22 (1939).

[2] W. G. Dixon, Nuovo Cimento **34**, 317 (1964).

[3] M. Mathisson, Acta Phys. Polon. **6**, 163 (1937).

[4] M. Mathisson, Proc. Cambridge Phil. Soc. **36**, 331 (1940).

[5] W. Tulczyjew, Acta Phys. Polon. **18**, 393 (1959).

[6] A. Papapetrou, Proc. Roy. Soc. (London) **A209**, 248 (1951).

Papapetrou,[7] and to Tulczyjew and Tulczyjew.[8] The treatment in Ref. 2 also allows for the presence of an electromagnetic field (as also does Ref. 7) and defines moments of J^α as well as of $T^{\alpha\beta}$. The total charge q, which is the electromagnetic analog of p^α and $S^{\alpha\beta}$, is similarly separately defined.

A feature common to all the above-mentioned theories is that (1.1) and (1.2), where applicable, imply an infinite number of relations between the moments. This suggests that these relations could be used to extract a subset of the moments that is still sufficient to determine $T^{\alpha\beta}$ and J^α completely, but upon which (1.1) and (1.2) impose only a finite number of restrictions. We might expect p^α and $S^{\alpha\beta}$ to be sufficient to describe the monopole and dipole structure of $T^{\alpha\beta}$ instead of needing the full $t^{\alpha\beta}$ and $t^{\alpha\beta\gamma}$, respectively, and similarly the charge q to suffice for the description of the monopole structure of J^α. (1.1) should then be expected to impose laws of motion on p^α and $S^{\alpha\beta}$ of the form

$$\delta p^\alpha/ds = F^\alpha \quad \text{and} \quad \delta S^{\alpha\beta}/ds = G^{\alpha\beta}, \tag{1.11}$$

where $\delta/ds \overset{\text{def}}{=} v^\alpha \nabla_\alpha$ and F^α, $G^{\alpha\beta}$ are, respectively, a force and a couple constructed from the curvature and electromagnetic field tensors and the moments of $T^{\alpha\beta}$ and J^α, while (1.2) should require charge conservation

$$dq/ds = 0. \tag{1.12}$$

We might hope that, with suitable definitions for the higher moments, (1.11) and (1.12) would be all of the consequences of (1.1) and (1.2), so that the variation of the other moments with time would be determined entirely by the equations of state of the body. Our main result in this paper is to show that, in the case of flat space-time, a set of moments of $T^{\alpha\beta}$ and J^α can be defined—as explicit integrals of these tensors—which has all these properties. F^α and $G^{\alpha\beta}$ are given explicitly as infinite series in the moments. The case of curved space-time will be treated in a later paper.

Here we should also mention theories of point particles. Although in many respects theories of point particles and of extended bodies overlap, there are important differences between them. If no restrictions other than (1.3) and (1.4) are placed on the moments, then (1.1) and (1.2) impose no restrictions on the world-line L. Thus, to obtain equations of motion, additional conditions must be used. For an extended body we take L to be the world line of a suitably defined center of mass of the body. The problem of defining such a mass center has been discussed by Tulczyjew[5] and by Dixon,[2] who conjecture without proof that, imposing

$$p_\beta S^{\alpha\beta} = 0, \tag{1.13}$$

L is suitably and uniquely determined. A more rigorous discussion has been given by Beiglböck,[9] who proves this uniqueness under very general conditions, using the definitions of p^α and $S^{\alpha\beta}$ given in Ref. 2, and who also suggests an alternative definition of the center of mass. In the case of point particles, however, we are not free to define a center of mass, since the position of the particle is well defined; any additional conditions we impose are physical restrictions on the particle. This has been discussed in more detail by Dixon.[10] Another difference is that the energy-momentum tensor of a point particle has exactly the form (1.9) and satisfies (1.10), so that no approximation is involved in limiting oneself to only a finite number of nonvanishing moments. This is not as trivial as it might seem at first, for we later have cause to question the validity of the multipole approximation procedure in the form in which it is explained above. Multipole theories of point particles have been given by Mathisson,[3] by Lubański,[11] by Hönl and Papapetrou,[12,13] and by Taub.[14,15]

2. NOTATION AND CONVENTIONS

Space–time is treated as a pseudo-Riemannian manifold $\mathcal{M}$ with metric tensor $g_{\alpha\beta}$ and signature $+ - - -$, tensor indices running from 0 to 3 and the summation convention being used throughout. We write $g \overset{\text{def}}{=} \det g_{\alpha\beta}$. Partial and covariant differentiation with respect to x^α is denoted by ∂_α and ∇_α, respectively, with the kernel ∂ or ∇ being written only once in repeated differentiations, thus $\nabla_{\alpha\beta} \equiv \nabla_\alpha \nabla_\beta$. We say that a function on $\mathcal{M}$ is of *class* C^r if it has continuous derivatives of all orders $\leq r$, and of *class* C^∞ if it has continuous derivatives of all orders.

Following Lichnerowicz,[16] we treat the Dirac δ function as a biscalar distribution on $\mathcal{M}$ defined by

$$\int f(x)\delta(x, y)(-g)^{\frac{1}{2}}\, d^4x = f(y) \tag{2.1}$$

for any C^∞ scalar function f of compact support. If

[7] W. Urich and A. Papapetrou, Z. Naturforsch. **10a**, 109 (1955).
[8] B. Tulczyjew and W. Tulczyjew, in *Recent Developments in General Relativity* (Pergamon Press, Inc., New York, 1962), p. 465.
[9] W. D. Beiglböck, Commun. Math. Phys. **5**, 106 (1967).
[10] W. G. Dixon, Nuovo Cimento **38**, 1616 (1965).
[11] J. Lubański, Acta Phys. Polon. **6**, 356 (1937).
[12] H. Hönl and A. Papapetrou, Z. Physik **112**, 512 (1939).
[13] H. Hönl and A. Papapetrou, Z. Physik **114**, 478 (1939).
[14] A. H. Taub, J. Math. Phys. **5**, 112 (1964).
[15] A. H. Taub, Proc. Galileo IV Centario Conference, Florence (1964), p. 77.
[16] A. Lichnerowicz, in *Relativity, Groups and Topology*, C. M. DeWitt and B. S. DeWitt, Eds. (Gordon and Breach, Science Publishers, Inc., New York, 1964), p. 821.

$T^{\alpha_1\cdots\alpha_n}$ is a C^∞ tensor field of compact support, then we also write

$$\int T^{\alpha_1\cdots\alpha_n}(x)\delta(x, y)(-g)^{\frac{1}{2}}\,d^4x = T^{\alpha_1\cdots\alpha_n}(y), \quad (2.2)$$

although then $\delta(x, y)$ is acting strictly as a bitensor distribution.

We follow Schouten[17] in denoting the symmetrization and antisymmetrization of any number of tensor indices by ( ) and [ ], respectively, e.g.,

$$A_{[\alpha\beta\gamma]} = \tfrac{1}{6}(A_{\alpha\beta\gamma} + A_{\beta\gamma\alpha} + A_{\gamma\alpha\beta} - A_{\alpha\gamma\beta} - A_{\beta\alpha\gamma} - A_{\gamma\beta\alpha}), \quad (2.3)$$

and in writing for four indices

$$A_{[\alpha[\beta\gamma]\delta]} \stackrel{\text{def}}{=} \tfrac{1}{4}(A_{\alpha\beta\gamma\delta} - A_{\gamma\beta\alpha\delta} - A_{\alpha\delta\gamma\beta} + A_{\gamma\delta\alpha\beta}). \quad (2.4)$$

Indices enclosed in vertical lines are excluded from these operations, e.g.,

$$A_{(\alpha|\beta\gamma|\delta)} = \tfrac{1}{2}(A_{\alpha\beta\gamma\delta} + A_{\delta\beta\gamma\alpha}). \quad (2.5)$$

When considering flat space-time $\mathcal{E}$, we use Minkowskian coordinates with metric tensor

$$g_{\alpha\beta} = \text{diag}\,(1, -1, -1, -1), \quad (2.6)$$

and denote the scalar product of vectors a^α, b^α by $a\cdot b \stackrel{\text{def}}{=} a^\alpha b_\alpha$. The sign of the electromagnetic field tensor $F^{\alpha\beta}$ is such that

$$\mathbf{E} \stackrel{\text{def}}{=} (F^{01}, F^{02}, F^{03}) \quad \text{and} \quad \mathbf{H} \stackrel{\text{def}}{=} (F^{23}, F^{31}, F^{12}) \quad (2.7)$$

are the electric and magnetic 3-vectors, respectively, in the 3-space $x^0 = \text{const}$.

Following Gel'fand and Shilov,[18] we denote by K the set of all C^∞ complex-valued scalar functions on $\mathcal{E}$ of compact support, and by Z the set of slowly increasing entire analytic functions on $\mathcal{E}$. Specifically, this means that $f\in Z$ if, in any Minkowskian coordinate system, (1) it may be extended to all complex values of its arguments so as to be everywhere analytic in each variable, and (2) for this extension, there exist positive constants $C_{q_0\cdots q_3}$ for all positive integers q_r, and positive constants $a_0, \cdots, a_3$, such that

$$|z_0^{q_0}\cdots z_3^{q_3} f(z_0, \cdots, z_3)| < C_{q_0\cdots q_3}\exp(a_0|y_0| + \cdots + a_3|y_3|) \quad (2.8)$$

for all complex z_r, where $z_r = x_r + iy_r$ with x_r and y_r real. Then, if $f\in K$ or Z, its Fourier transform $\tilde{f}$, defined by

$$\tilde{f}(k) = \int f(x)e^{ik\cdot x}\,d^4x, \quad (2.9)$$

is an element of Z or K, respectively. The inverse of (2.9) is

$$f(x) = \frac{1}{(2\pi)^4}\int \tilde{f}(k)e^{-ik\cdot x}\,d^4k. \quad (2.10)$$

3. STRUCTURE OF THE CHARGE-CURRENT VECTOR

In the subsequent sections we consider an extended charged body in a flat space-time $\mathcal{E}$ and use Minkowskian coordinates throughout. The charge-current vector J^α is then a C^1 vector field satisfying

$$\partial_\alpha J^\alpha = 0, \quad (3.1)$$

whose support W is a spacewise bounded world tube. Choose any C^1 timelike world-line L, with parametric equation $x^\alpha = z^\alpha(s)$, where s is the proper time measured along it, and let $\Sigma(s)$ be the hyperplane through $z^\alpha(s)$ orthogonal to

$$v^\alpha(s) \stackrel{\text{def}}{=} dz^\alpha/ds. \quad (3.2)$$

Our exposition is worded as if L lies in W, but no real restriction is involved in this. On $\Sigma(s)$ define

$$r^\alpha(x) \stackrel{\text{def}}{=} x^\alpha - z^\alpha(s) \quad (3.3)$$

and

$$w^\alpha(x) \stackrel{\text{def}}{=} v^\alpha(s)[1 - \dot{v}^\beta(s)r_\beta(x)], \quad (3.4)$$

where $\dot{v}^\alpha \stackrel{\text{def}}{=} dv^\alpha/ds$. Then

$$v_\alpha(s)r^\alpha(x) = 0, \quad \text{for} \quad x\in\Sigma(s). \quad (3.5)$$

And if f is any C^1 function on $\mathcal{E}$ of bounded support,

$$\int f(x)\,d^4x = \int ds\int_{\Sigma(s)} f(x)w^\alpha\,d\Sigma_\alpha \quad (3.6)$$

and

$$\frac{d}{ds}\int_{\Sigma(s)} f\,d\Sigma_\alpha = \int_{\Sigma(s)}\partial_\alpha f w^\beta\,d\Sigma_\beta, \quad (3.7)$$

as is shown in Refs. 4 and 2.

We need the following integrals of J^α:

$$j^{\alpha_1\cdots\alpha_n\beta}(s) \stackrel{\text{def}}{=} \int_{\Sigma(s)} r^{\alpha_1}\cdots r^{\alpha_n}J^\beta w^\gamma\,d\Sigma_\gamma \quad (3.8)$$

and

$$q^{\alpha_1\cdots\alpha_n}(s) \stackrel{\text{def}}{=} \int_{\Sigma(s)} r^{\alpha_1}\cdots r^{\alpha_n}J^\beta\,d\Sigma_\beta, \quad \text{for} \quad n\geq 0. \quad (3.9)$$

Using (3.5), we see that they satisfy

$$j^{\alpha_1\cdots\alpha_n\beta} = j^{(\alpha_1\cdots\alpha_n)\beta}, \quad q^{\alpha_1\cdots\alpha_n} = q^{(\alpha_1\cdots\alpha_n)} \quad (3.10)$$

and

$$v_{\alpha_1}j^{\alpha_1\cdots\alpha_n\beta} = 0, \quad v_{\alpha_1}q^{\alpha_1\cdots\alpha_n} = 0. \quad (3.11)$$

Let $d(s)$ be the diameter of the compact section $\Sigma(s)\cap W$. Then, since J^β is continuous, each of its

[17] J. A. Schouten, *Ricci Calculus. An Introduction to Tensor Analysis and its Geometrical Applications* (Springer-Verlag, Berlin, 1954), 2nd ed.

[18] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. 1.

components is bounded on $\Sigma(s) \cap W$, and so there exist continuous positive scalar functions $A(s)$ and $B(s)$ such that

$$|j^{\alpha_1 \cdots \alpha_n \beta}(s)| < A(s)\, d^n \quad \text{and} \quad |q^{\alpha_1 \cdots \alpha_n}(s)| < B(s)\, d^n \tag{3.12}$$

for all n, these inequalities holding for each component separately.

Now let $f \in K$, and write

$$\langle J^\alpha, f \rangle \stackrel{\text{def}}{=} \int J^\alpha(x) f(x)\, d^4x. \tag{3.13}$$

With (3.6) and (2.10) this gives

$$\langle J^\alpha, f \rangle = \frac{1}{(2\pi)^4} \int ds \int_{\Sigma(s)} w^\beta\, d\Sigma_\beta \int d^4k \tilde{f}(k) J^\alpha(x) e^{-ik\cdot x}. \tag{3.14}$$

But $\tilde{f} \in Z$, so that $|\tilde{f}(k_0, \cdots, k_3)| < C\,|k_0, \cdots, k_3|^{-2}$ for some constant $C > 0$, and hence the k-space integration is uniformly convergent for $x \in \Sigma(s)$. Thus, since $\Sigma(s) \cap W$ is compact, the order of integration over $\Sigma(s)$ and k space may be interchanged. Doing this and expanding the exponential in a power series about $z(s)$ gives

$$\langle J^\alpha, f \rangle = \frac{1}{(2\pi)^4} \int ds \int d^4k \int_{\Sigma(s)} w^\beta\, d\Sigma_\beta \tilde{f}(k) J^\alpha(x) e^{-ik\cdot z(s)} \times \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} (k \cdot r)^n. \tag{3.15}$$

From now on, we will not give the details of the arguments involving uniform convergence which justify our operations with infinite series and integrals, unless some special point is involved, since they are all very simple. So, on integrating the series in (3.15) term by term over $\Sigma(s)$ and using (3.8), we get

$$\langle J^\alpha, f \rangle = \frac{1}{(2\pi)^4} \int ds \int d^4k f(k) e^{-ik\cdot z(s)} \times \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} k_{\beta_1} \cdots k_{\beta_n} j^{\beta_1 \cdots \beta_n \alpha}(s). \tag{3.16}$$

This shows that the $j^{\cdots}$'s determine $\langle J^\alpha, f \rangle$ for every $f \in K$, and hence they completely specify J^α itself.

In particular, this shows that the $j^{\cdots}$'s determine the $q^{\cdots}$'s. This can also be seen directly from (3.8) and (3.9), for, as $\Sigma(s)$ is orthogonal to $v^\alpha(s)$, we have

$$J^\beta\, d\Sigma_\beta = J^\beta v_\beta v^\gamma\, d\Sigma_\gamma.$$

Putting this into (3.9) and using (3.4), we see that

$$q^{\alpha_1 \cdots \alpha_n} - \dot{v}_\beta q^{\alpha_1 \cdots \alpha_n \beta} = j^{\alpha_1 \cdots \alpha_n \beta} v_\beta \quad \text{for} \quad n \geq 0.$$

This may be iterated to give

$$q^{\alpha_1 \cdots \alpha_n} = v_\gamma \sum_{p=0}^{\infty} j^{\alpha_1 \cdots \alpha_n \beta_1 \cdots \beta_p \gamma} \dot{v}_{\beta_1} \cdots \dot{v}_{\beta_p}, \quad \text{for} \quad n \geq 0,$$

which, by (3.12), is convergent for sufficiently small $\dot{v}_\alpha$. However, for larger $\dot{v}_\alpha$ we cannot give an explicit expression of this sort, and we would have to resort to (3.16) to determine the $q^{\cdots}$'s from the $j^{\cdots}$'s. We need the integral expression (3.9) for the $q^{\cdots}$'s in our development of the theory given below.

If we could perform the k-space integration in (3.16) term by term, we would get

$$\langle J^\alpha, f \rangle = \int ds \sum_{n=0}^{\infty} \frac{1}{n!} j^{\beta_1 \cdots \beta_n \alpha}(s) [\partial_{\beta_1} \cdots \partial_{\beta_n} f]_{z(s)}, \tag{3.17}$$

or equivalently

$$J^\alpha(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \partial_{\beta_1} \cdots \partial_{\beta_n} \int j^{\beta_1 \cdots \beta_n \alpha}(s) \delta(z(s), x)\, ds. \tag{3.18}$$

(3.17) is the analog in flat space–time for J^α of (1.5) for $T^{\alpha\beta}$, and is only valid if f is analytic on $\Sigma(s) \cap W$ for each s. It clearly cannot be true for a general $f \in K$ since the derivatives of f along L which appear on the right-hand side do not determine f throughout W, and hence cannot be sufficient to determine the left-hand side. This failure is reflected in the form of the right-hand side of (3.18), since this, which formally appears to be a distribution of support L, does not exist within the framework of the theory of distributions unless the series terminates. This can be seen from (3.17), which would be its value at $f \in K$, since in general the right-hand side of (3.17) need not be convergent. The structure of a general distribution whose support is a C^∞ submanifold has been determined by Schwartz,[19] whose result is applicable to our case if L is of class C^∞.

Nevertheless, any finite number of terms of (3.16) can be integrated term by term, giving

$$\langle J^\alpha, f \rangle = \int ds \sum_{n=0}^{N} \frac{1}{n!} j^{\beta_1 \cdots \beta_n \alpha}(s) [\partial_{\beta_1} \cdots \partial_{\beta_n} f]_{z(s)} + \frac{1}{(2\pi)^4} \int ds \int d^4k \tilde{f}(k) e^{-ik\cdot z} \times \sum_{n=N+1}^{\infty} \frac{(-i)^n}{n!} k_{\beta_1} \cdots k_{\beta_n} j^{\beta_1 \cdots \beta_n \alpha}. \tag{3.19}$$

From this we get the unexpected result that, if the moments $j^{\cdots}$ are known for all $n > N$, they completely determine J^α and thus all the lower moments! For, by (3.19), the moments for $n > N$ determine $\langle J^\alpha, f \rangle$ for any $f \in K$ which vanishes in some neighborhood of L, and thus they determine J^α at all points not

[19] L. Schwartz, *Théorie des distributions*. (Hermann et Cie., Paris, 1957), Vol. I, Chap. III, Théorème XXXVII.

on L. By continuity this determines J^α everywhere, as stated. We discuss this result further in Sec. 8.

We now use (3.1) to obtain a new and more convenient set of moments that also determines J^α. Using (3.7) and remembering that r^α also depends explicitly on s, we get from (3.9) that

$$dq/ds = 0 \tag{3.20}$$

and

$$(d/ds)q^{\alpha_1\cdots\alpha_n} = nj^{(\alpha_1\cdots\alpha_n)} - nv^{(\alpha_1}q^{\alpha_2\cdots\alpha_n)} \quad \text{for} \quad n \geq 1. \tag{3.21}$$

This is a system of differential equations determining the dependence of the $q^{\cdots}$'s on s in terms of the completely symmetric part of the $j^{\cdots}$'s. Since, as we saw above, the $j^{\cdots}$'s determine the $q^{\cdots}$'s, this is equivalent to an infinite system of differential relations between the $j^{\cdots}$'s as a consequence of the conservation equation (3.1).

Let us put

$$a^{\alpha_1\cdots\alpha_n\beta} \overset{\text{def}}{=} j^{\alpha_1\cdots\alpha_n\beta} - j^{(\alpha_1\cdots\alpha_n\beta)} \quad \text{for} \quad n \geq 1. \tag{3.22}$$

With (3.21) this gives

$$\left.\begin{aligned} j^\alpha &= qv^\alpha + dq^\alpha/ds, \\ j^{\alpha_1\cdots\alpha_n\beta} &= a^{\alpha_1\cdots\alpha_n\beta} + q^{(\alpha_1\cdots\alpha_n}v^{\beta)} \\ &\quad + \frac{1}{n+1}\frac{d}{ds}q^{\alpha_1\cdots\alpha_n\beta} \quad \text{for} \quad n \geq 1, \end{aligned}\right\} \tag{3.23}$$

from which

$$\begin{aligned} e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}j^{\beta_1\cdots\beta_n\alpha} &= e^{-ik\cdot z}\left[qv^\alpha + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}m^{\beta_1\cdots\beta_n\alpha}\right] \\ &\quad + \frac{d}{ds}\left[e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{(n+1)!}k_{\beta_1}\cdots k_{\beta_n}q^{\beta_1\cdots\beta_n\alpha}\right], \end{aligned} \tag{3.24}$$

where

$$m^{\beta_1\cdots\beta_n\alpha} \overset{\text{def}}{=} a^{\beta_1\cdots\beta_n\alpha} + q^{(\beta\cdots\beta_n}v^{\alpha)} - v^{(\beta_1}q^{\beta_2\cdots\beta_n)\alpha} \quad \text{for} \quad n \geq 1. \tag{3.25}$$

(3.25), (3.22), and (3.12) show that there exists a continuous positive function $C(s)$ such that

$$|m^{\beta_1\cdots\beta_n\alpha}(s)| < C(s)\,d^n(s) \quad \text{for all} \quad n \geq 1, \tag{3.26}$$

and from (3.25) and (3.23) we get

$$m^{\beta_1\cdots\beta_n\alpha} = j^{\beta_1\cdots\beta_n\alpha} - v^{(\beta_1}q^{\beta_2\cdots\beta_n)\alpha} - \frac{1}{n+1}\frac{d}{ds}q^{\beta_1\cdots\beta_n\alpha} \quad \text{for} \quad n \geq 1. \tag{3.27}$$

(3.12) and (3.26) may be used to verify the convergence of the series in (3.24), and in particular (3.12) has been used above to justify the term-by-term differentiation of the second series on the right-hand side of (3.24). This series may be summed explicitly on using (3.9) for

$$\begin{aligned} \sum_{n=0}^{\infty}\frac{(-i)^n}{(n+1)!}k_{\beta_1}\cdots k_{\beta_n}q^{\beta_1\cdots\beta_n\alpha} &= \int_{\Sigma(s)}\left[\sum_{n=0}^{\infty}\frac{(-i)^n}{(n+1)!}(k\cdot r)^n\right]r^\alpha J^\beta\,d\Sigma_\beta \\ &= \int_{\Sigma(s)}\left[\int_0^1 du\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}(uk\cdot r)^n\right]r^\alpha J^\beta\,d\Sigma_\beta \\ &= \int_{\Sigma(s)}d\Sigma_\beta\int_0^1 du r^\alpha J^\beta \exp\,[-iuk\cdot r]. \end{aligned} \tag{3.28}$$

So, on substituting (3.24) into (3.16), the contribution from the last term of (3.24) is

$$\begin{aligned} &\frac{1}{(2\pi)^4}\int ds\int d^4k\tilde{f}(k)\frac{d}{ds}\int_{\Sigma(s)}d\Sigma_\beta \\ &\qquad \times\int_0^1 du r^\alpha J^\beta \exp\,[-ik\cdot(z(s)+ur)] \\ &= \int ds\frac{d}{ds}\int_{\Sigma(s)}d\Sigma_\beta\int_0^1 du r^\alpha J^\beta f(z(s)+ur), \end{aligned} \tag{3.29}$$

where the k-space integration has been brought through the d/ds and two other integrations and performed first with the help of (2.10). Since f has bounded support, we see that (3.29) vanishes, so that only the first term on the right-hand side of (3.24) contributes to (3.16). We thus have

$$\langle J^\alpha, f\rangle = \frac{1}{(2\pi)^4}\int ds\int d^4k\tilde{f}(k)Q^\alpha(k,s), \tag{3.30}$$

where

$$Q^\alpha(k,s) \overset{\text{def}}{=} e^{-ik\cdot z}\left[qv^\alpha + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}m^{\beta_1\cdots\beta_n\alpha}\right], \tag{3.31}$$

which shows that J^α is completely determined by q and the set of $m^{\cdots}$'s.

Consequently, q and the $m^{\cdots}$'s must completely determine the $j^{\cdots}$'s. For small $\dot{v}^\alpha$ we can, from (3.23) and (3.27), demonstrate this explicitly. For on multiplying (3.27) by v_{β_1} and using (3.11), we get

$$v_\alpha m^{\alpha\beta_1\cdots\beta_n} = -\frac{1}{n}q^{\beta_1\cdots\beta_n} + \frac{1}{n+1}\dot{v}_\alpha q^{\alpha\beta_1\cdots\beta_n} \quad \text{for} \quad n \geq 1. \tag{3.32}$$

By iteration, this yields

$$q^{\beta\cdots\beta_n} = -nv_\alpha\sum_{p=0}^{\infty}\dot{v}_{\gamma_1}\cdots\dot{v}_{\gamma_p}m^{\alpha\gamma_1\cdots\gamma_p\beta_1\cdots\beta_n} \quad \text{for} \quad n \geq 1,$$

which by (3.26) converges for sufficiently small $\dot{v}^\alpha$.

This then gives the $q^{\alpha_1\cdots\alpha_n}$ for $n \geq 1$ in terms of the $m^{\cdots}$'s, and q itself is assumed given. (3.27) together with the first of equations (3.23) then give all the $j^{\cdots}$'s. However, just as in the general case we must resort to (3.16) to evaluate the $q^{\cdots}$'s when the $j^{\cdots}$'s are known, so also for larger $\dot{v}^\alpha$ we must use (3.30) with (3.31) to determine the $j^{\cdots}$'s given q and the $m^{\cdots}$'s.

We thus see that q and the $m^{\cdots}$'s describe the arbitrariness in the $j^{\cdots}$'s left after the conservation equation (3.1), or equivalently the system (3.20) and (3.21) of differential relations between them has been taken into account. For later use, we also note here that (3.32) and (3.10) imply

$$v_\alpha m^{\alpha\beta_1\cdots\beta_n} = v_\alpha m^{\alpha(\beta_1\cdots\beta_n)} \quad \text{for} \quad n \geq 1. \tag{3.33}$$

$Q^\alpha(k, s)$ may also be given in an integral form that gives a useful upper bound for it. By summing the left-hand side of (3.24) [cf. the transition from (3.14) to (3.16)] and using (3.31) and (3.28), we obtain

$$Q^\alpha(k, s) = \int_{\Sigma(s)} J^\alpha(x) e^{-ik\cdot x} w^\beta \, d\Sigma_\beta - \frac{d}{ds}\int_{\Sigma(s)} d\Sigma_\beta \int_0^1 du r^\alpha J^\beta(x) \exp\left[-ik \cdot \big(z(s) + ur\big)\right]. \tag{3.34}$$

On explicitly performing the differentiation of the second term on the right-hand side, we easily see that there exist continuous positive functions $D(s)$ and $E(s)$ such that

$$|Q^\alpha(k, s)| < D(s) + E(s)(|k_0| + \cdots + |k_3|). \tag{3.35}$$

We need this result later.

From (3.25) and (3.22) we see that

$$m^{\beta_1\cdots\beta_n\alpha} = m^{(\beta_1\cdots\beta_n)\alpha} \quad \text{and} \quad m^{(\beta_1\cdots\beta_n\alpha)} = 0. \tag{3.36}$$

We now verify that as a consequence of these symmetry conditions, the only condition imposed on q and the $m^{\cdots}$'s by (3.1) is the conservation of charge (3.20). For if $f \in K$, we have

$$\langle \partial_\alpha J^\alpha, f\rangle = -\langle J^\alpha, \partial_\alpha f\rangle \tag{3.37}$$

and

$$\widetilde{\partial_\alpha f}(k) = -ik_\alpha \tilde{f}(k) \tag{3.38}$$

so that $\langle \partial_\alpha J^\alpha, f\rangle$ is obtained by replacing $\tilde{f}(k)$ in (3.30) by $ik_\alpha \tilde{f}(k)$. Then in virtue of (3.36), only the contribution from the first term in (3.31) survives, giving

$$\langle \partial_\alpha J^\alpha, f\rangle = \frac{i}{(2\pi)^4}\int ds \int d^4k \tilde{f}(k) e^{-ik\cdot z} q k_\alpha v^\alpha = -\frac{1}{(2\pi)^4} q \int ds \frac{d}{ds}\int d^4k \tilde{f}(k) e^{-ik\cdot z}$$

by (3.20)

$$= -q\int ds \frac{d}{ds}\big(z(s)\big) \tag{3.39}$$

by (2.10), which vanishes as f has bounded support. This gives $\partial_\alpha J^\alpha = 0$ as a consequence of (3.20) when J^α is given by (3.30) with (3.31) and (3.36) satisfied, as required.

The orthogonality conditions (3.11) also impose conditions on the $m^{\cdots}$'s, namely (3.33). These can be put in a simpler form if we describe the body, not by the $m^{\cdots}$'s, but instead, by a new set of tensors defined by

$$Q^{\alpha_1\cdots\alpha_n\beta\gamma} = m^{\alpha_1\cdots\alpha_n[\beta\gamma]} \quad \text{for} \quad n \geq 0. \tag{3.40}$$

Using (3.36) we see that

$$\begin{aligned} Q^{\alpha_1\cdots\alpha_n\beta\gamma} &= Q^{(\alpha_1\cdots\alpha_n)[\beta\gamma]} \quad \text{for} \quad n \geq 0, \\ Q^{\alpha_1\cdots\alpha_{n-1}[\alpha_n\beta\gamma]} &= 0 \quad \text{for} \quad n \geq 1, \end{aligned} \tag{3.41}$$

and

$$m^{\alpha_1\cdots\alpha_n\beta} = [2n/(n+1)]Q^{(\alpha_1\cdots\alpha_n)\beta} \quad \text{for} \quad n \geq 1. \tag{3.42}$$

(3.40) and (3.42) together show that the $m^{\cdots}$'s and the $Q^{\cdots}$'s are equivalent for describing the body. Moreover, from (3.41) and (3.42) we can deduce (3.40) and (3.36), so that the symmetry conditions (3.36) and (3.41) are equivalent. But from (3.40) and (3.33) we get

$$v_{\alpha_1} Q^{\alpha_1\cdots\alpha_n\beta\gamma} = 0 \quad \text{for} \quad n \geq 1, \tag{3.43}$$

which are the orthogonality conditions satisfied by the $Q^{\cdots}$'s and which are simpler than the equivalent (3.33).

From (3.40), (3.25), and (3.22) we have

$$Q^{\alpha_1\cdots\alpha_n\beta\gamma} = j^{\alpha_1\cdots\alpha_n[\beta\gamma]} + [1/(n+1)]q^{\alpha_1\cdots\alpha_n[\beta}v^{\gamma]} \quad \text{for} \quad n \geq 0, \tag{3.44}$$

which, when expressed in terms of J^α through (3.8) and (3.9), is simpler than the corresponding expression for the $m^{\cdots}$'s. Together with the simplicity of the symmetry and orthogonality conditions (3.41) and (3.43), this makes the $Q^{\cdots}$'s seem to be the most satisfactory description of the multipole structure of the charge distribution of the body. In virtue of (3.41) and (3.43), $Q^{\alpha_1\cdots\alpha_n\beta\gamma}$ has $\frac{1}{2}(n+3)(3n+4)$ linearly independent components, and we call it the 2^{n+1}-pole electromagnetic moment tensor of the body. Together with q, the total charge of the body, which describes the monopole structure, these moments completely determine J^α through (3.30), (3.31), and (3.42), whilst the only condition on them due to (3.1) is the conservation of charge, (3.20).

4. LORENTZ FORCE

Our next task is to express the Lorentz force

$$F^\alpha \overset{\text{def}}{=} -F^{\alpha\beta}J_\beta \tag{4.1}$$

in terms of the moments of J^α in a form similar to (3.30) and (3.31) for J^α itself. Because we have only proved (3.30) for functions of class C^∞, we need to assume that $F^{\alpha\beta}$ is also of class C^∞, but almost certainly this can be weakened—possibly even to taking $F^{\alpha\beta}$ to be of class C^1.

Let $f \in K$ have support $S(f)$, and let ${}^*F^{\alpha\beta}$ be a C^∞ tensor field of compact support that equals $F^{\alpha\beta}$ on $S(f) \cap W$. Then

$$\langle F^{\alpha\beta} J_\beta, f\rangle = \langle J^\beta, {}^*F^\alpha_{\cdot\beta} f\rangle. \tag{4.2}$$

Also, we have

$$\widetilde{{}^*F^\alpha_{\cdot\beta} f}(k) = \int d^4x\, {}^*F^\alpha_{\cdot\beta}(x) f(x) e^{ik\cdot x} = \frac{1}{(2\pi)^4} \int d^4x \int d^4l\, {}^*F^\alpha_{\cdot\beta}(x) \tilde{f}(l) e^{i(k-l)\cdot x} \tag{4.3}$$

on using (2.10), and since $\tilde{f} \in Z$ and ${}^*F^{\alpha\beta}$ has compact support, we may change the order of integration and so perform the x-integration first using (2.9), obtaining

$$\widetilde{{}^*F^\alpha_{\cdot\beta} f}(k) = \frac{1}{(2\pi)^4} \int d^4l \tilde{f}(l)\, \widetilde{{}^*F}^\alpha_{\cdot\beta}(k - l). \tag{4.4}$$

Together with (4.1), (4.2), and (3.30) this gives

$$\langle F^\alpha, f\rangle = -\frac{1}{(2\pi)^8} \int ds \int d^4k \times \int d^4l \tilde{f}(l)\, \widetilde{{}^*F}^\alpha_{\cdot\beta}(k - l) Q^\beta(k, s). \tag{4.5}$$

Remembering that $\tilde{f} \in Z$ and $\widetilde{{}^*F}^\alpha_{\cdot\beta} \in Z$ and using (3.35), we see that we may invert the order of the k and l integrations. After doing so, we then change the variables in the k integration from k_α to $u_\alpha \overset{\text{def}}{=} k_\alpha - l_\alpha$, and then we relabel l_α as k_α and u_α as l_α to give

$$\langle F^\alpha, f\rangle = -[1/(2\pi)^8] \int ds \int d^4k \times \int d^4l \tilde{f}(k)\, \widetilde{{}^*F}^\alpha_{\cdot\beta}(l) Q^\beta(k + l, s). \tag{4.6}$$

The infinite series for $Q^\beta(k + l, s)$ given by (3.31) is seen, on using (3.26), to be absolutely convergent when regarded as a double (or rather, octuple) series in k_α and l_α. It may thus be rearranged in any order, in particular as a series in k_α in the form

$$Q^\alpha(k + l, s) = e^{-ik\cdot z} \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} k_{\beta_1} \cdots k_{\beta_n} \psi^{\beta_1 \cdots \beta_n \alpha}(l, s), \tag{4.7}$$

where

$$\left.\begin{aligned} &\psi^\alpha(l, s) = Q^\alpha(l, s), \\ &\text{given by (3.31), and} \\ &\psi^{\beta_1 \cdots \beta_n \alpha}(l, s) \\ &\quad = e^{-il\cdot z} \sum_{p=0}^{\infty} \frac{(-i)^p}{p!} l_{\gamma_1} \cdots l_{\gamma_p} m^{\gamma_1 \cdots \gamma_p \beta_1 \cdots \beta_n \alpha}(s) \\ &\qquad\qquad \text{for } n \geq 1. \end{aligned}\right\} \tag{4.8}$$

We now want to substitute (4.7) into (4.6) and perform the l integration term by term, but to justify doing so we need to obtain an upper bound for $|\psi^{\beta_1 \cdots \beta_n \alpha}|$. This involves explicitly summing the series in (4.8). We first substitute for the $m^{\cdots}$'s from (3.27), and then we use (3.8) and (3.9). Proceeding in an analogous manner to the derivation of (3.28), we then find that

$$\psi^{\beta_1 \cdots \beta_n \alpha} = K^{\beta_1 \cdots \beta_n \alpha} - n v^{(\beta_1} H^{\beta_2 \cdots \beta_n)\alpha} - (d/ds) H^{\beta_1 \cdots \beta_n \alpha} \quad \text{for } n \geq 1, \tag{4.9}$$

where we have defined for $n \geq 0$

$$K^{\beta_1 \cdots \beta_n \alpha}(l, s) \overset{\text{def}}{=} \int_{\Sigma(s)} e^{-il\cdot x} r^{\beta_1} \cdots r^{\beta_n} J^\alpha w^\gamma \, d\Sigma_\gamma \tag{4.10}$$

and

$$H^{\beta_1 \cdots \beta_n \alpha}(l, s) \overset{\text{def}}{=} \int_{\Sigma(s)} d\Sigma_\gamma \int_0^1 du u^n r^{\beta_1} \cdots r^{\beta_n} r^\alpha J^\gamma \times \exp[-il \cdot (z + ur)], \tag{4.11}$$

while (3.34) and (4.8) give

$$\psi^\alpha = K^\alpha - (d/ds) H^\alpha. \tag{4.12}$$

On explicitly performing the differentiation in the last terms of (4.9) and (4.12), we see that there exist continuous positive functions $F(s)$, $G(s)$ such that

$$|\psi^{\beta_1 \cdots \beta_n \alpha}(l\, s,)| < d^n(s)[F(s) + (|l_0| + \cdots + |l_3|)G(s)] \tag{4.13}$$

for all n, l, and s. This enables us to put a uniform bound on the partial sums of (4.7):

$$\sum_{n=0}^{N} \left| \frac{(-i)^n}{n!} k_{\beta_1} \cdots k_{\beta_n} \psi^{\beta_1 \cdots \beta_n \alpha}(l, s) \right| < [F + (|l_0| + \cdots + |l_3|)G] \exp[(|k_0| + \cdots + |k_3|)\, d], \tag{4.14}$$

from which we see that the l integration in (4.6), when performed on the partial sums of N terms of (4.7) instead of on Q^α itself, converges uniformly in N, and that the series (4.7) converges uniformly in any bounded region of l space. Together, these two uniformities of the convergence of (4.6) validate the term-by-term integration referred to above, which gives

$$\langle F^\alpha, f\rangle = -\frac{1}{(2\pi)^4} \int ds \int d^4k \tilde{f}(k) e^{-ik\cdot z} \times \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} k_{\beta_1} \cdots k_{\beta_n} \phi^{\beta_1 \cdots \beta_n \alpha}(s), \tag{4.15}$$

where

$$\phi^{\beta_1 \cdots \beta_n \alpha}(s) \overset{\text{def}}{=} \frac{1}{(2\pi)^4} \int d^4l\, \widetilde{{}^*F}^\alpha_{\cdot\gamma}(l) \psi^{\beta_1 \cdots \beta_n \gamma}(l, s) \quad \text{for } n \geq 0. \tag{4.16}$$

The coefficients $\phi^{\cdots}$ are given in terms of the moments $m^{\cdots}$ through (4.16) and (4.8), but we also need the integral forms for them that follow from (4.16) with (4.9). Using (2.10) with (4.10) and (4.11), we find for $n \geq 0$ that

$$\frac{1}{(2\pi)^4}\int d^4l\, {}^*\widetilde{F}^{\alpha}_{\cdot\gamma}(l)K^{\beta_1\cdots\beta_n\gamma}(l,s) = f^{\beta_1\cdots\beta_n\alpha}(s) \quad (4.17)$$

and

$$\frac{1}{(2\pi)^4}\int d^4l\, {}^*\widetilde{F}^{\alpha}_{\cdot\gamma}(l)H^{\beta_1\cdots\beta_n\gamma}(l,s) = h^{\beta_1\cdots\beta_n\alpha}(s), \quad (4.18)$$

where

$$f^{\beta_1\cdots\beta_n\alpha}(s) \stackrel{\text{def}}{=} \int_{\Sigma(s)} r^{\beta_1}\cdots r^{\beta_n}\, {}^*F^{\alpha}_{\cdot\gamma}J^{\gamma}w^{\delta}\, d\Sigma_{\delta} \quad (4.19)$$

and

$$h^{\beta_1\cdots\beta_n\alpha}(s) \stackrel{\text{def}}{=} \int_{\Sigma(s)} d\Sigma_{\delta} \times \int_0^1 du\, u^n r^{\beta_1}\cdots r^{\beta_n} r^{\gamma}J^{\delta}\, {}^*F^{\alpha}_{\cdot\gamma}(z+ur). \quad (4.20)$$

Equation (4.16) with Eqs. (4.9) and (4.12) then gives

$$\left.\begin{aligned} \phi^{\alpha} &= f^{\alpha} - dh^{\alpha}/ds, \\ \text{and} \\ \phi^{\beta_1\cdots\beta_n\alpha} &= f^{\beta_1\cdots\beta_n\alpha} - nv^{(\beta_1}h^{\beta_2\cdots\beta_n)\alpha} \\ &\quad - (d/ds)h^{\beta_1\cdots\beta_n\alpha} \quad \text{for} \quad n \geq 1, \end{aligned}\right\} \quad (4.21)$$

which are the required integral forms. We now use (4.21) in the infinite series in (4.15) to give

$$\begin{aligned} & e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}\phi^{\beta_1\cdots\beta_n\alpha} \\ &= e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}f^{\beta_1\cdots\beta_n\alpha} \\ &\quad - \frac{d}{ds}\left[e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}h^{\beta_1\cdots\beta_n\alpha}\right] \\ &= \int_{\Sigma(s)} e^{-ik\cdot x}\, {}^*F^{\alpha}_{\cdot\beta}J^{\beta}w^{\gamma}\, d\Sigma_{\gamma} - \frac{d}{ds}\int_{\Sigma(s)} d\Sigma_{\gamma}\int_0^1 du r^{\beta} \\ &\quad \times {}^*F^{\alpha}_{\cdot\beta}(z+ur)J^{\gamma}(x)\exp\left[-ik\cdot(z+ur)\right], \end{aligned} \quad (4.22)$$

which enables the k space integration in (4.15) to be performed explicitly, giving

$$\begin{aligned} & \frac{1}{(2\pi)^4}\int d^4k\tilde{f}(k)e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\beta_1}\cdots k_{\beta_n}\phi^{\beta_1\cdots\beta_n\alpha} \\ &= \int_{\Sigma(s)} f(x)\, {}^*F^{\alpha}_{\cdot\beta}(x)J^{\beta}(x)w^{\gamma}\, d\Sigma_{\gamma} - \frac{d}{ds}\int_{\Sigma(s)} d\Sigma_{\gamma} \\ &\quad \times \int_0^1 du f(z+ur)r^{\beta}\, {}^*F^{\alpha}_{\cdot\beta}(z+ur)J^{\gamma}(x). \end{aligned} \quad (4.23)$$

Since f has compact support, we see that (4.23) vanishes outside a finite range of s, say the interval $[s_1, s_2]$, this interval depending on f but not on the particular ${}^*F^{\alpha}_{\cdot\beta}$ used. This incidentally shows that, on performing the final integration in (4.15) over s, the contribution from the second term on the right-hand side of (4.23) vanishes, while on using (3.6), the contribution from the first term is seen to be $\langle F^{\alpha}, f\rangle$ as required. However, it shows more than this; it shows that the left-hand side of (4.23) vanishes for s outside the interval $[s_1, s_2]$, independently of the choice of the tensor field ${}^*F^{\alpha}_{\cdot\beta}$ used in (4.16) in defining the $\phi^{\cdots}$'s, provided only that it has class C^{∞} and compact support. But let U be a bounded convex set containing both $S(f)$ and the segment of W lying between $\Sigma(s_1)$ and $\Sigma(s_2)$, so that we may choose ${}^*F^{\alpha}_{\cdot\beta}$ to be equal to $F^{\alpha}_{\cdot\beta}$ in U. Then (4.19)–(4.21) show that for any $s \in [s_1, s_2]$, the value of the $\phi^{\cdots}(s)$'s depends only on the value of ${}^*F^{\alpha}_{\cdot\beta}$ in the neighborhood of the convex hull of $\Sigma(s) \cap W$, and the convex hull of $\Sigma(s) \cap W$ has a neighborhood lying in U in which ${}^*F^{\alpha}_{\cdot\beta} = F^{\alpha}_{\cdot\beta}$. Hence, altogether, we see that (4.15) remains valid if, for each s, the $\phi^{\cdots}(s)$'s are evaluated using (4.16) with any ${}^*F^{\alpha}_{\cdot\beta}$ (of class C^{∞} and compact support) that equals $F^{\alpha}_{\cdot\beta}$ in some neighborhood of the convex hull of $\Sigma(s) \cap W$. The $\phi^{\cdots}$'s thus obtained are seen to be independent of this choice of ${}^*F^{\alpha}_{\cdot\beta}$, and they are given by (4.21) when ${}^*F^{\alpha}_{\cdot\beta}$ in (4.19) and (4.20) is replaced by $F^{\alpha}_{\cdot\beta}$. In the future, references to $\phi^{\cdots}$, $f^{\cdots}$, and $h^{\cdots}$ refer to these modified functions.

The modified definition of the coefficients $\phi^{\cdots}$ used in (4.15) thus defines them uniquely and independently of f. It is to achieve this independence from f that we have gone through these final stages of the argument, for with the previous definition by (4.16) as it stood, the ${}^*F^{\alpha}_{\cdot\gamma}$ used was independent of s but could not be chosen independently of f, and thus it made the $\phi^{\cdots}$'s depend on f. We cannot use $F^{\alpha}_{\cdot\beta}$ itself in (4.16) since in general its Fourier transform will not exist. Our final expression for the Lorentz force is thus (4.15) with the modified definition of the $\phi^{\cdots}$'s explained above.

5. STRUCTURE OF THE ENERGY–MOMENTUM TENSOR

The energy–momentum tensor $T^{\alpha\beta}$ is a C^1 tensor field satisfying

$$\partial_{\beta}T^{\alpha\beta} = -F^{\alpha\beta}J_{\beta}, \quad (5.1)$$

whose support is a spacewise-bounded world-tube W'. We now analyze its structure following the procedure developed in Sec. 3 for J^{α}, but the details are rather more complicated than for that case. We first define for $n \geq 0$ the integrals

$$t^{\alpha_1\cdots\alpha_n\beta\gamma}(s) = \int_{\Sigma(s)} r^{\alpha_1}\cdots r^{\alpha_n}T^{\beta\gamma}w^{\delta}\, d\Sigma_{\delta} \quad (5.2)$$

and

$$p^{\alpha_1\cdots\alpha_n\beta}(s)=\int_{\Sigma(s)} r^{\alpha_1}\cdots r^{\alpha_n}T^{\beta\gamma}\,d\Sigma_\gamma, \tag{5.3}$$

which satisfy

$$t^{\alpha_1\cdots\alpha_n\beta\gamma}=t^{(\alpha_1\cdots\alpha_n)(\beta\gamma)},\quad p^{\alpha_1\cdots\alpha_n\beta}=p^{(\alpha_1\cdots\alpha_n)\beta}, \tag{5.4}$$

$$v_{\alpha_1}t^{\alpha_1\cdots\alpha_n\beta\gamma}=0 \quad\text{and}\quad v_{\alpha_1}p^{\alpha_1\cdots\alpha_n\beta}=0, \tag{5.5}$$

and also inequalities similar to (3.12). Then for $f\in K$, as the analog of (3.16), we obtain

$$\langle T^{\alpha\beta},f\rangle=\frac{1}{(2\pi)^4}\int ds\int d^4k\tilde{f}(k)e^{-ik\cdot z}\times\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\gamma_1}\cdots k_{\gamma_n}t^{\gamma_1\cdots\gamma_n\alpha\beta}(s), \tag{5.6}$$

showing that the integrals (5.2) completely determine $T^{\alpha\beta}$ (as stated in Sec. 1) and thus also the $p^{\cdots}$'s. This can be shown more explicitly for small $\dot{v}^\alpha$, as in Sec. 3.

Using (5.2) and (4.19), we now obtain from (5.3), (3.7), and (5.1)

$$dp^\alpha/ds=-f^\alpha \tag{5.7}$$

and

$$\frac{d}{ds}p^{\alpha_1\cdots\alpha_n\beta}=-nv^{(\alpha_1}p^{\alpha_2\cdots\alpha_n)\beta}+nt^{(\alpha_1\cdots\alpha_n)\beta}-f^{\alpha_1\cdots\alpha_n\beta}\quad\text{for}\quad n\geq 1. \tag{5.8}$$

Using (4.21) and setting

$$\pi^{\alpha_1\cdots\alpha_n\beta}\overset{\text{def}}{=}p^{\alpha_1\cdots\alpha_n\beta}+h^{\alpha_1\cdots\alpha_n\beta}\quad\text{for}\quad n\geq 0, \tag{5.9}$$

these may be written as

$$d\pi^\alpha/ds=-\phi^\alpha \tag{5.10}$$

and

$$(d/ds)\pi^{\alpha_1\cdots\alpha_n\beta}=-nv^{(\alpha_1}\pi^{\alpha_2\cdots\alpha_n)\beta}+nt^{(\alpha_1\cdots\alpha_n)\beta}-\phi^{\alpha_1\cdots\alpha_n\beta}\quad\text{for}\quad n\geq 1. \tag{5.11}$$

Now define as the analog of (3.22)

$$b^{\alpha_1\cdots\alpha_n\beta}\overset{\text{def}}{=}\pi^{(\alpha_1\cdots\alpha_n\beta)}\quad\text{for}\quad n\geq 1, \tag{5.12}$$

$$c^{\alpha_1\cdots\alpha_n\beta}\overset{\text{def}}{=}\pi^{\alpha_1\cdots\alpha_n\beta}-\pi^{(\alpha_1\cdots\alpha_n\beta)}\quad\text{for}\quad n\geq 1 \tag{5.13}$$

and

$$d^{\alpha_1\cdots\alpha_n\beta\gamma}\overset{\text{def}}{=}t^{\alpha_1\cdots\alpha_n\beta\gamma}-[(n+1)/n]\times[t^{(\alpha_1\cdots\alpha_n\beta)\gamma}+t^{(\alpha_1\cdots\alpha_n\gamma)\beta}]+[(n+2)/n]t^{(\alpha_1\cdots\alpha_n\beta\gamma)}\quad\text{for}\quad n\geq 2, \tag{5.14}$$

so that

$$\left.\begin{aligned}c^{\alpha_1\cdots\alpha_n\beta}&=c^{(\alpha_1\cdots\alpha_n)\beta} & c^{(\alpha_1\cdots\alpha_n\beta)}&=0,\\ d^{\alpha_1\cdots\alpha_n\beta\gamma}&=d^{(\alpha_1\cdots\alpha_n)(\beta\gamma)} \text{ and } & d^{(\alpha_1\cdots\alpha_n\beta)\gamma}&=0,\end{aligned}\right\} \tag{5.15}$$

and, for later convenience, put

$$S^{\alpha\beta}\overset{\text{def}}{=}2c^{\alpha\beta}=2\pi^{[\alpha\beta]}. \tag{5.16}$$

We use these with (5.11) to obtain the analogs of (3.23), but the cases $n=1$ and $n=2$ of (5.11) must be treated separately from the rest. For $n=1$ we find

$$t^{\alpha\beta}=v^\alpha\pi^\beta+\tfrac{1}{2}(d/ds)(S^{\alpha\beta}+2b^{\alpha\beta})+\phi^{\alpha\beta}, \tag{5.17}$$

which we separate into its symmetric and antisymmetric parts thus:

$$t^{\alpha\beta}=\pi^{(\alpha}v^{\beta)}+(d/ds)b^{\alpha\beta}+\phi^{(\alpha\beta)} \tag{5.18}$$

and

$$(d/ds)S^{\alpha\beta}=2\pi^{[\alpha}v^{\beta]}-2\phi^{[\alpha\beta]}. \tag{5.19}$$

For $n=2$, (5.11) gives

$$2t^{(\alpha\beta)\gamma}=2v^{(\alpha}b^{\beta)\gamma}+v^{(\alpha}S^{\beta)\gamma}+(d/ds)(b^{\alpha\beta\gamma}+c^{\alpha\beta\gamma})+\phi^{\alpha\beta\gamma}, \tag{5.20}$$

which, with the identity

$$t^{\alpha\beta\gamma}=t^{(\alpha\beta)\gamma}+t^{(\alpha\gamma)\beta}-t^{(\beta\gamma)\alpha}, \tag{5.21}$$

gives

$$t^{\alpha\beta\gamma}=v^\alpha b^{\beta\gamma}+S^{\alpha(\beta}v^{\gamma)}+(d/ds)(\tfrac{1}{2}b^{\alpha\beta\gamma}-c^{\beta\gamma\alpha})+\phi^{\alpha(\beta\gamma)}-\tfrac{1}{2}\phi^{\beta\gamma\alpha}. \tag{5.22}$$

Finally, using (5.11) for $n\geq 3$ together with (5.14), we get

$$t^{\alpha_1\cdots\alpha_n\beta\gamma}=d^{\alpha_1\cdots\alpha_n\beta\gamma}+v^{(\alpha_1}b^{\alpha_2\cdots\alpha_n)\beta\gamma}+2v^{(\alpha_1}c^{\alpha_2\cdots\alpha_n)(\beta\gamma)}+\frac{2}{n}c^{\alpha_1\cdots\alpha_n(\beta}v^{\gamma)}+\frac{d}{ds}\left[\frac{1}{n+1}b^{\alpha_1\cdots\alpha_n\beta\gamma}+\frac{2}{n}c^{\alpha_1\cdots\alpha_n(\beta\gamma)}\right]+\frac{2}{n}\phi^{\alpha_1\cdots\alpha_n(\beta\gamma)}-\frac{n+2}{n(n+1)}\phi^{(\alpha_1\cdots\alpha_n\beta\gamma)}\quad\text{for } n\geq 2. \tag{5.23}$$

We now use (5.18), (5.22), and (5.23) to put the infinite series in (5.6) in the form

$$\begin{aligned}&e^{-ik\cdot z}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}k_{\gamma_1}\cdots k_{\gamma_n}t^{\gamma_1\cdots\gamma_n\alpha\beta}\\&=e^{-ik\cdot z}\Big\{v^{(\alpha}\pi^{\beta)}+\phi^{(\alpha\beta)}+(-ik_\gamma)\\&\times[S^{\gamma(\alpha}v^{\beta)}+\phi^{\gamma(\alpha\beta)}-\tfrac{1}{2}\phi^{\alpha\beta\gamma}]+\sum_{n=2}^{\infty}\frac{(-i)^n}{n!}k_{\gamma_1}\cdots k_{\gamma_n}\\&\times\left[I^{\gamma_1\cdots\gamma_n\alpha\beta}+\frac{2}{n}\phi^{\gamma_1\cdots\gamma_n(\alpha\beta)}-\frac{n+2}{n(n+1)}\phi^{(\gamma_1\cdots\gamma_n\alpha\beta)}\right]\Big\}\\&+\frac{d}{ds}\left[e^{-ik\cdot z}\left[b^{\alpha\beta}+\sum_{n=1}^{\infty}\frac{(-i)^n}{n!}k_{\gamma_1}\cdots k_{\gamma_n}\right.\right.\\&\left.\left.\times\left(\frac{1}{n+1}b^{\gamma_1\cdots\gamma_n\alpha\beta}+\frac{2}{n}c^{\gamma_1\cdots\gamma_n(\alpha\beta)}\right)\right]\right],\end{aligned} \tag{5.24}$$

where

$$I^{\alpha_1\cdots\alpha_n\beta\gamma} \overset{\text{def}}{=} d^{\alpha_1\cdots\alpha_n\beta\gamma} - [2/(n-1)] \times v^{(\alpha_1}c^{\alpha_2\cdots\alpha_n)(\beta\gamma)} + (2/n)c^{\alpha_1\cdots\alpha_n(\beta}v^{\gamma)} \quad (5.25)$$

for $n \geq 2$. The term in (5.24) involving a total derivative with respect to s may be explicitly summed, as before, and shown to give a zero contribution to the right-hand side of (5.6). We thus have

$$\langle T^{\alpha\beta}, f\rangle = \frac{1}{(2\pi)^4}\int ds \int d^4k \tilde{f}(k)e^{-ik\cdot z}\Big\{v^{(\alpha}\pi^{\beta)} + \phi^{(\alpha\beta)} + (-ik_\gamma) \times [S^{\gamma(\alpha}v^{\beta)} + \phi^{\gamma(\alpha\beta)} - \tfrac{1}{2}\phi^{\alpha\beta\gamma}] + \sum_{n=2}^{\infty}\frac{(-i)^n}{n!}k_{\gamma_1}\cdots k_{\gamma_n} \times \left[I^{\gamma_1\cdots\gamma_n\alpha\beta} + \frac{2}{n}\phi^{\gamma_1\cdots\gamma_n(\alpha\beta)} - \frac{n+2}{n(n+1)}\phi^{(\gamma_1\cdots\gamma_n\alpha\beta)}\right]\Big\}, \quad (5.26)$$

which shows that to determine $T^{\alpha\beta}$ we need, in addition to the $\phi^{\cdots}$ given by (4.16) and (4.8) in terms of the electromagnetic moments, the quantities π^α, $S^{\alpha\beta}$ and the $I^{\cdots}$'s, which are seen to be defined as explicit integrals over $\Sigma(s)$ involving both $T^{\alpha\beta}$ and J^α. The contributions from J^α vanish when $F^{\alpha\beta} = 0$, and may be interpreted as due to the potential energy of the charge distribution in the electromagnetic field.

From (5.15) and (5.25), we see that the $I^{\cdots}$'s have the symmetry properties

$$I^{\alpha_1\cdots\alpha_n\beta\gamma} = I^{(\alpha_1\cdots\alpha_n)(\beta\gamma)} \quad \text{and} \quad I^{(\alpha_1\cdots\alpha_n\beta)\gamma} = 0. \quad (5.27)$$

But as in the electromagnetic case, there are additional restrictions on the $I^{\cdots}$'s due to the orthogonality conditions (5.5), and to find these it is more convenient to use, instead of the $I^{\cdots}$'s, new tensors defined by

$$J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} \overset{\text{def}}{=} I^{\alpha_1\cdots\alpha_n[\beta[\delta\gamma]\epsilon]} \quad \text{for} \quad n \geq 0. \quad (5.28)$$

Using (5.25), (5.14), and (5.13) we find that

$$J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} = I^{\alpha_1\cdots\alpha_n[\beta[\delta\gamma]\epsilon]} - [1/(n+1)] \times [v^{[\beta}\pi^{\gamma]\alpha_1\cdots\alpha_n[\delta\epsilon]} + v^{[\delta}\pi^{\epsilon]\alpha_1\cdots\alpha_n[\beta\gamma]}], \quad (5.29)$$

and, as a result of (5.27), we see that the $J^{\cdots}$'s have the symmetry properties

$$\left.\begin{aligned} &J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} = J^{(\alpha_1\cdots\alpha_n)[\beta\gamma][\delta\epsilon]}, \quad J^{\alpha_1\cdots\alpha_n\beta[\gamma\delta\epsilon]} = 0, \\ &\text{and} \\ &J^{\alpha_1\cdots\alpha_{n-1}[\alpha_n\beta\gamma]\delta\epsilon} = 0 \quad \text{for} \quad n \geq 1, \end{aligned}\right\} \quad (5.30)$$

and that

$$I^{\alpha_1\cdots\alpha_n\beta\gamma} = [4(n-1)/(n+1)]J^{(\alpha_1\cdots\alpha_{n-1}|\beta|\alpha_n)\gamma} \quad \text{for} \quad n \geq 2. \quad (5.31)$$

Moreover, (5.27) and (5.28) are consequences of (5.30) and (5.31), so that the $J^{\cdots}$'s are equivalent to the $I^{\cdots}$'s in their information content. Incidentally, (5.30) also implies that

$$J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} = J^{\alpha_1\cdots\alpha_n\delta\epsilon\beta\gamma}. \quad (5.32)$$

From (4.20) we see that

$$v_{\alpha_1}h^{\alpha_1\cdots\alpha_n\beta} = 0 \quad \text{for} \quad n \geq 1, \quad (5.33)$$

which, with (5.9) and (5.5), shows that the $\pi^{\cdots}$'s satisfy

$$v_{\alpha_1}\pi^{\alpha_1\cdots\alpha_n\beta} = 0 \quad \text{for} \quad n \geq 1. \quad (5.34)$$

Together with (5.5) and (5.29), this shows that the required orthogonality conditions for the $J^{\cdots}$'s are

$$v_{\alpha_1}J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} = 0 \quad \text{for} \quad n \geq 1. \quad (5.35)$$

In virtue of (5.30) and (5.35), $J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon}$ has $(n+4)\times(3n+5)$ linearly independent components. We call it the 2^{n+2}-pole inertial moment tensor of the body. Together with π^α and $S^{\alpha\beta}$, the $J^{\cdots}$'s seem to be the most convenient description of the multipole structure of $T^{\alpha\beta}$, and with the electromagnetic moments they completely determine $T^{\alpha\beta}$ through (5.26) and (5.31). π^α describes the monopole structure and we identify it with the momentum vector of the body, while $S^{\alpha\beta}$ describes the dipole structure and is identified with the spin tensor of the body. Both depend on the choice of L, and are thus fully determined only when a center of mass has been chosen. These identifications are discussed further in Sec. 7.

We finally show that as a consequence of the symmetry conditions (5.27), or equivalently (5.30), the only relations between the moments imposed by (5.1) are (5.10) and (5.19). For, on using (3.38) and the analog of (3.37), we see that we may evaluate $\langle \partial_\beta T^{\alpha\beta}, f\rangle$ by replacing $\tilde{f}(k)$ in (5.26) by $ik_\beta\tilde{f}(k)$. The contributions from the $I^{\cdots}$'s then vanish on account of (5. 27), and the contributions from the $\phi^{\cdots}$'s under the summation sign are seen to be equal to the corresponding terms of (4.15). We thus obtain

$$\langle \partial_\beta T^{\alpha\beta} + F^{\alpha\beta}J_\beta, f\rangle = \frac{1}{(2\pi)^4}\int ds\int d^4k\tilde{f}(k)e^{-ik\cdot z} \times [ik_\beta v^{(\alpha}\pi^{\beta)} + \tfrac{1}{2}k_\beta k_\gamma S^{\gamma\alpha}v^\beta + \phi^\alpha + ik_\beta\phi^{[\alpha\beta]}]. \quad (5.36)$$

On using (5.10) and (5.19), the right-hand side of this

simplifies to

$$\frac{1}{(2\pi)^4}\int ds\,\frac{d}{ds}\int d^4k\tilde{f}(k)e^{-ik\cdot z}[-\pi^\alpha + \tfrac{1}{2}ik_\gamma S^{\gamma\alpha}]$$
$$= -\int ds\,\frac{d}{ds}[\pi^\alpha f(z(s)) + \tfrac{1}{2}S^{\beta\alpha}[\partial_\beta f]_{z(s)}], \quad (5.37)$$

which vanishes as required, since f has bounded support.

6. APPROXIMATION PROCEDURES

We now study what approximations can be made when the body is small in comparison to a typical length-scale of the external field, but, before so doing, it is necessary to say more precisely what we mean by such a length-scale. Suppose we are interested in the electromagnetic field $F^{\alpha\beta}$ in a bounded region V of space–time and in some particular Lorentz frame. Let $^*F^{\alpha\beta}$ be any extension having compact support and class C^∞ of $F^{\alpha\beta}$ in V to the whole of space–time and let m be a positive integer. Suppose that there exists an $R > 0$ such that if Ω is the region of l-space given by

$$|l_0| + \cdots + |l_3| > R^{-1}, \quad (6.1)$$

then

$$\int_\Omega (|l_0| + \cdots + |l_3|)^m\,|\widetilde{^*F^{\alpha\beta}}|\,d^4l$$
$$\ll \int_{\text{all } l \text{ space}} (|l_0| + \cdots + |l_3|)^m\,|\widetilde{^*F^{\beta\alpha}}|\,d^4l. \quad (6.2)$$

Then, if there is not a much larger R also satisfying this, R is a measure of the distances and times over which $^*F^{\alpha\beta}$ and its derivatives up to order m vary appreciably in that Lorentz frame, which we abbreviate to "R is a typical mth-order length scale for $^*F^{\alpha\beta}$." As such variation may be due either to its value in V, the region of interest, or to the particular extension we have made, we say that R is a typical mth-order length scale for $F^{\alpha\beta}$ in V if this is a length scale as explained above for its "best possible" extension—meaning "best possible" in the sense of maximizing R. We note that we cannot just cut off $F^{\alpha\beta}$ at the boundary of V and take it as zero outside, as in general the Fourier transform of the resulting function (which will be discontinuous on the boundary of V) will tend to zero at infinity no faster than $|l_0 \cdots l_3|^{-1}$.

Now, we saw in Sec. 4 that the $\phi^{\cdots}$'s depend only on $F^{\alpha\beta}$ in the neighborhood of the convex hull of $\Sigma(s) \cap W$. Let $m \geq 1$, $R(s)$ be a typical mth-order length scale for $F^{\alpha\beta}$ in such a neighborhood, and $d(s)$ be, as before, the diameter of the section $\Sigma(s) \cap W$. Then assuming $d(s) \ll R(s)$, we investigate the validity of using only a finite number of terms in the series (4.8) for the $\psi^{\cdots}$'s in (4.16). For this purpose, the $^*F^{\alpha\beta}$ in (4.16) is to be chosen to have length scale $R(s)$ also. On defining the partial sums

$$\left.\begin{aligned}&\psi^\alpha_N(l,s)\\ &\quad \overset{\text{def}}{=} e^{-il\cdot z}\left[qv^\alpha + \sum_{p=1}^{N}\frac{(-i)^p}{p!}\,l_{\beta_1}\cdots l_{\beta_p}m^{\beta_1\cdots\beta_p\alpha}\right]\\ &\text{and}\\ &\psi^{\beta_1\cdots\beta_n\alpha}_N(l,s)\\ &\quad \overset{\text{def}}{=} e^{-il\cdot z}\sum_{p=0}^{N}\frac{(-i)^p}{p!}\,l_{\gamma_1}\cdots l_{\gamma_p}m^{\gamma_1\cdots\gamma_p\beta_1\cdots\beta_n\alpha}\\ &\qquad\qquad \text{for}\quad n \geq 1,\end{aligned}\right\} \quad (6.3)$$

(3.26) shows that

$$|\psi^{\beta_1\cdots\beta_n\alpha} - \psi^{\beta_1\cdots\beta_n\alpha}_N|$$
$$< C(s)\sum_{p=N+1}^{\infty}\frac{1}{p!}(|l_0| + \cdots + |l_3|)^p\,d^{n+p}$$
$$< \frac{C(s)\,d^n}{(N+1)!}\,\frac{(d/D)^{N+1}}{1 - d/D} \quad \text{for} \quad n \geq 0, \quad (6.4)$$

where

$$D(l) \overset{\text{def}}{=} (|l_0| + \cdots + |l_3|)^{-1}. \quad (6.5)$$

Now for fixed s, let ω be the region of l space given by

$$|l_0| + \cdots + |l_3| < R^{-1}(s), \quad (6.6)$$

i.e., $D(l) > R(s)$. Then in virtue of (6.2) and (4.13), we see that unless the $\psi^{\cdots}$'s have an exceptional behavior, e.g., being extremely small in ω, the contribution to the integral in (4.16) from the region Ω of l space will be very small compared with the contribution from ω and we may neglect it, taking the integral only over ω. But in ω, $D(l) > R(s)$, and so the assumed smallness of d/R together with (6.4) shows that if N is sufficiently large, to a good approximation we may use $\psi^{\cdots}_N$ instead of $\psi^{\cdots}$ in this integral over ω. Provided $N \leq m$, we may now use (6.2) to extend the region of integration back from ω to the whole of l space. Since we now have only a finite number of terms in our series, the integration may be performed term by term, and since

$$\frac{1}{(2\pi)^4}\int d^4l\,\widetilde{^*F}^\alpha_{\cdot\beta}(l)\,\frac{(-i)^p}{p!}\,e^{-il\cdot z}l_{\gamma_1}\cdots l_{\gamma_p}$$
$$= \frac{1}{p!}[\partial_{\gamma_1\cdots\gamma_q}F^\alpha_{\cdot\beta}]_{z(s)}, \quad (6.7)$$

to our approximation we have

$$\left.\begin{aligned}&\phi^\alpha = qv^\beta F^\alpha_{\cdot\beta} + \sum_{p=1}^{N}\frac{1}{p!}\,m^{\beta_1\cdots\beta_p\gamma}[\partial_{\beta_1\cdots\beta_p}F^\alpha_{\cdot\gamma}]_{z(s)}\\ &\text{and}\\ &\phi^{\beta_1\cdots\beta_n\alpha}\\ &\quad = \sum_{p=0}^{N}\frac{1}{p!}\,m^{\beta_1\cdots\beta_n\gamma_1\cdots\gamma_p\delta}[\partial_{\gamma_1\cdots\gamma_p}F^\alpha_{\cdot\delta}]_{z(s)}\\ &\qquad\qquad \text{for}\quad n \geq 1.\end{aligned}\right\} \quad (6.8)$$

We see that the approximation improves as N increases, provided $N \leq m$, but that in general we cannot take the limit $N \to \infty$, for as we go to higher orders, R has to be a typical length scale for higher derivatives of $F^{\alpha\beta}$, and thus R decreases and d/R increases, worsening the stage in the approximation procedure that uses (6.4). Only by making some analyticity assumptions for $F^{\alpha\beta}$ can we guarantee the convergence of (6.8) as $N \to \infty$. We note that if the results of Sec. 4 are valid for an $F^{\alpha\beta}$ of class C^r rather than of class C^∞, we cannot be led to a contradiction by having $m > r$ and thus allowing $N > r$ in (6.8), since $\widetilde{*F}^{\alpha\beta}(l)$ would only be of order $|l_0 \cdots l_3|^{-(r+1)}$ at infinity, and the integrals in (6.2) would not converge for $m \geq r$. Due to the crudeness of our limits, our procedure would not even validate the case $N = r$, although (6.8) is then meaningful.

We thus see that without any assumptions of analyticity for $F^{\alpha\beta}$, we can justify the approximations (6.8) provided that the size of the body is small compared with a typical Nth-order length scale for the external field, and by following the above procedure step by step in a particular case, one could obtain an estimate for the error. But in general the series in (6.8) will not converge as $N \to \infty$ to the exact expressions given in Sec. 4.

Using (6.8) in (5.10) and (5.19) and retaining terms involving up to the 2^rth electromagnetic moment gives what the earlier theories would call "equations of motion in the 2^r-pole approximation." But in contrast to these earlier theories, our neglect of the higher moments has been made only in the interaction terms, whereas they unjustifiably also neglect the higher moments in terms not involving the external fields. This is inevitable in any procedure that starts from a truncated form of (1.8) or an equivalent integral formalism, as the higher moments are neglected from the very beginning, wherever they should occur. This is conveniently illustrated from the author's treatment in Ref. 2. For a body in flat space–time with $F^{\alpha\beta} = 0$, the equations of motion obtained there reduce to

$$dp^\alpha/ds = 0 \quad \text{and} \quad dS^{\alpha\beta}/ds = 2p^{[\alpha}v^{\beta]}, \tag{6.9}$$

the notation agreeing with that of Sec. 5 if the surfaces $\Sigma(s)$ in Ref. 2. are taken to be hyperplanes orthogonal to v^α, as in the present treatment. As written, these equations are exact, but we may also deduce, from (6.22) of Ref. 2, that

$$S^{\alpha\beta} = 2t^{[\alpha\beta]\gamma}v_\gamma, \tag{6.10}$$

which is in general true only in the pole–dipole approximation if L is taken as an arbitrary world line. If this is used in (6.9), they thus become only approximate equations of motion. However, if L is taken as the world line of the center of mass of the body so that $\dot{v}^\alpha = 0$, then (6.10) *is* exact. This suggests that the neglect of the higher moments throughout, as in the earlier theories, may be justifiable when used in conjunction with a suitable definition of the center of mass, but this is by no means obvious and needs to be more thoroughly investigated to validate these methods.

The validity of the present procedure is independent of any definition of the center of mass, but difficulties occur if we try to use

$$\pi_\beta S^{\alpha\beta} = 0, \tag{6.11}$$

which is the most natural candidate, because of our defining π^α and $S^{\alpha\beta}$ as integrals over a hyperplane orthogonal to v^α, rather than π^α as in Ref. 2. This makes it more difficult to see whether or not (6.11) determines a unique world line. But we do not discuss this point further here. (*Note added in proof.* The present treatment can be modified to use hyperplanes orthogonal to π^α, and then this difficulty disappears.)

7. INTERPRETATION OF THE MOMENTS

From (5.9), (5.3), and (4.20) we see that

$$\pi^\alpha(s) = \int_{\Sigma(s)} (T^{\alpha\beta} + A^\alpha J^\beta)\, d\Sigma_\beta, \tag{7.1}$$

where

$$A^\alpha(x) \overset{\text{def}}{=} \int_0^1 F^\alpha_{\cdot\beta}(z + ur)r^\beta\, du. \tag{7.2}$$

If the term in (7.1) involving J^α were absent, π^α would be the usual definition of the momentum vector, e.g., as given by Aharoni.[20] Let us look at this extra term in the "local rest-frame," i.e., in coordinates in which $v^\alpha = \delta^\alpha_0$ at the point of interest, where δ^α_β is the Kronecker δ symbol. Then using (2.7),

$$A^0(x) = -\int_z^x \mathbf{E} \cdot \mathbf{ds}, \tag{7.3}$$

where the integral is taken along the straight line joining $z(s)$ to x. This may be interpreted as the "electrostatic potential" of x relative to $z(s)$, and is precisely this in the static case when the integral is path-independent. The electromagnetic contribution to π^0 being

$$\int A^0 J^0\, d^3x, \tag{7.4}$$

[20] J. Aharoni, *The Special Theory of Relativity* (Oxford University Press, New York, 1959), Chap. 4.

it may thus be interpreted as the excess potential energy of the charge distribution due to its being spread out over what it would be if all the charge were concentrated at $z(s)$. Thus for π^0, this extra term has a simple interpretation that makes it seem very natural. For the space components $\boldsymbol{\pi} = (\pi^1, \pi^2, \pi^3)$, we do not have as natural an interpretation, but for them (7.1) is reminiscent of the relation

$$\boldsymbol{\pi} = m\mathbf{v} + e\mathbf{A} \tag{7.5}$$

between the canonical momentum and velocity for a charged particle, although our expression (7.1) is gauge-invariant, whereas (7.5) is not.

We also have, from (5.16), (5.9), (5.3), and (4.20), that

$$S^{\alpha\beta} = \int_{\Sigma(s)} [r^\alpha(T^{\beta\gamma} + B^\beta J^\gamma) - r^\beta(T^{\alpha\gamma} + B^\alpha J^\gamma)]\, d\Sigma_\gamma\,, \tag{7.6}$$

where

$$B^\alpha(x) \overset{\text{def}}{=} \int_0^1 u F^\alpha{}_\beta(z + ur) r^\beta\, du. \tag{7.7}$$

If the terms in J^α were absent, this would be the usual definition of the spin tensor, e.g., as given by Aharoni,[20] which supports our terminology, but this time the electromagnetic contribution seems to have no simple interpretation.

Before considering the other inertial moments, the $J^{\cdots}$'s, we discuss the electromagnetic moments. From (3.9) we see directly that

$$q = \int J^\alpha\, d\Sigma_\alpha, \tag{7.8}$$

which agrees with the usual definition of the total charge of the body. To interpret the $Q^{\cdots}$'s, we decompose them with respect to v^α, setting

$$E^{\alpha_1\cdots\alpha_{n+1}} \overset{\text{def}}{=} Q^{\alpha_1\cdots\alpha_{n+1}\beta} v_\beta \tag{7.9}$$

and

$$M^{\alpha_1\cdots\alpha_n\beta\gamma} \overset{\text{def}}{=} B^{\beta\gamma}_{\delta\epsilon} Q^{\alpha_1\cdots\alpha_n\delta\epsilon} \tag{7.10}$$

for $n \geq 0$, where

$$B^\alpha_\beta \overset{\text{def}}{=} \delta^\alpha_\beta - v^\alpha v_\beta \tag{7.11}$$

is the operator projecting orthogonally to v^α and we write the kernel B only once in a repeated product of projection operators. Then the $E^{\cdots}$'s are totally symmetric, the $M^{\cdots}$'s have the symmetries (3.41) of the $Q^{\cdots}$'s, and both are orthogonal to v^α on all indices. In the local rest frame, we find from (3.39), (3.8), (3.9), and (3.4) that

$$E^{a_1\cdots a_{n+1}} = \frac{1}{2}\int\left[\frac{n+2}{n+1} - r\cdot\dot{v}\right] r^{a_1}\cdots r^{a_{n+1}} J^0\, d^3x \tag{7.12}$$

and

$$M^{a_1\cdots a_n bc} = \int r^{a_1}\cdots r^{a_n} r^{[b} J^{c]}(1 - r\cdot\dot{v})\, d^3x, \tag{7.13}$$

where Latin indices run from 1 to 3. When our reference line L is straight, i.e., $\dot{v}^\alpha = 0$, these agree, respectively, apart from numerical factors, with the electric and magnetic moment tensors defined by Bloch,[21] and so our $Q^{\cdots}$'s are a natural relativistic generalization of Bloch's moments into a single electromagnetic moment tensor. We may call (7.9) and (7.10) the electric and magnetic parts of $Q^{\alpha_1\cdots\alpha_n\beta\gamma}$, and they have, respectively, $\frac{1}{2}(n+2)(n+3)$ and $(n+1)(n+3)$ linearly independent components.

Finally, the $J^{\cdots}$'s may be interpreted by a decomposition similar to that used for the $Q^{\cdots}$'s, in which we set

$$\begin{aligned}
\rho^{\alpha_1\cdots\alpha_{n+2}} &\overset{\text{def}}{=} J^{\alpha_1\cdots\alpha_n\alpha_{n+1}\beta\alpha_{n+2}\gamma} v_\beta v_\gamma\,,\\
\mu^{\alpha_1\cdots\alpha_{n+1}\beta\gamma} &\overset{\text{def}}{=} B^{\beta\gamma}_{\kappa\lambda} J^{\alpha_1\cdots\alpha_n\alpha_{n+1}\delta\kappa\lambda} v_\delta\,,\\
\sigma^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} &\overset{\text{def}}{=} -B^{\beta\gamma\delta\epsilon}_{\kappa\lambda\mu\nu} J^{\alpha_1\cdots\alpha_n\kappa\lambda\mu\nu}.
\end{aligned} \tag{7.14}$$

The $\rho^{\cdots}$'s are totally symmetric, the $\mu^{\cdots}$'s have the symmetries (3.41) of the $Q^{\cdots}$'s, the $\sigma^{\cdots}$'s have the symmetries (5.30) of the $J^{\cdots}$'s, and again, they are all orthogonal to v^α on all indices. If $J^\alpha = 0$ and we choose a world-line L with $\dot{v}^\alpha = 0$, they are given in the (now global) rest frame by

$$\begin{aligned}
\rho^{a_1\cdots a_{n+2}} &= \frac{1}{4}\frac{n+3}{n+1}\int r^{a_1}\cdots r^{a_{n+2}} T^{00}\, d^3x,\\
\mu^{a_1\cdots a_{n+1}bc} &= \frac{1}{2}\frac{n+2}{n+1}\int r^{a_1}\cdots r^{a_{n+1}} r^{[b} T^{c]0}\, d^3x,\\
\sigma^{a_1\cdots a_n bcde} &= \int r^{a_1}\cdots r^{a_n} r^{[b} T^{c][d} r^{e]}\, d^3x,
\end{aligned} \tag{7.15}$$

and they have, respectively, $\frac{1}{2}(n+3)(n+4)$, $(n+2)\times(n+4)$, and $\frac{3}{2}(n+1)(n+4)$ linearly independent components. As T^{00}, T^{a0}, and T^{ab} are the mass, momentum, and stress parts of $T^{\alpha\beta}$, we call $\rho^{\cdots}$, $\mu^{\cdots}$, and $\sigma^{\cdots}$ the mass, momentum, and stress parts of $J^{\cdots}$, respectively. The electromagnetic contributions that occur when $J^\alpha \neq 0$ again seem to have no simple interpretation, but from our considerations for π^α, they seem to be associated with the potential energy of the body in the electromagnetic field.

8. SUMMARY AND DISCUSSION

We have shown that a charged body moving in an electromagnetic field $F^{\alpha\beta}$ in a flat space–time may be completely described by (i) a set of inertial moments

$$\pi^\alpha, S^{\alpha\beta}, \text{ and } J^{\alpha_1\cdots\alpha_n\beta\gamma\delta\epsilon} \text{ for } n \geq 0, \tag{8.1}$$

[21] F. Bloch, in *Werner Heisenberg und die Physik unserer Zeit* (Vieweg und Sohn, Braunschweig, 1961), p. 93.

where π^α and $S^{\alpha\beta}$ are the momentum vector and spin tensor, respectively, and (ii) a set of electromagnetic moments

$$q, \text{ and } Q^{\alpha_1\cdots\alpha_n\beta\gamma} \quad \text{for} \quad n \geq 0, \tag{8.2}$$

where q is the total charge, which satisfies the symmetry and orthogonality conditions (5.30), (5.35), (3.41), and (3.43). These moments are defined with respect to an arbitrary world-line L, and they are given as explicit integrals of the energy-momentum tensor $T^{\alpha\beta}$ and the charge-current vector J^α through (5.29), (5.9), (5.3), (5.2), (4.20), and (3.4) for the inertial moments and (3.44), (3.9), (3.8), and (3.4) for the electromagnetic moments. If L is chosen to be the world line of a suitably defined center of mass of the body, the moments become uniquely determined. We have shown that $T^{\alpha\beta}$ and J^α are completely determined by these moments, and that the only relations between them imposed by the "generalized conservation equations"

$$\partial_\beta T^{\alpha\beta} = -F^{\alpha\beta}J_\beta \quad \text{and} \quad \partial_\alpha J^\alpha = 0 \tag{8.3}$$

are the conservation of charge

$$dq/ds = 0 \tag{8.4}$$

and the equations of motion

$$d\pi^\alpha/ds = -\phi^\alpha \quad \text{and} \quad dS^{\alpha\beta}/ds = 2\pi^{[\alpha}v^{\beta]} - 2\phi^{[\alpha\beta]}, \tag{8.5}$$

where ϕ^α and $2\phi^{[\alpha\beta]}$ are, respectively, a force and a couple given explicitly through (4.16) and (4.8) in terms of the applied electromagnetic field $F^{\alpha\beta}$ and the electromagnetic moments (8.2). When the body is small compared with a typical length scale for the external field, we have obtained the approximate expressions

$$\phi^\alpha = qv^\beta F^\alpha_{\cdot\beta} + \sum_{p=1}^{N} \frac{1}{p!} m^{\beta_1\cdots\beta_p\gamma}[\partial_{\beta_1}\cdots_{\beta_p}F^\alpha_{\cdot\gamma}]_{z(s)} \tag{8.6}$$

and

$$\phi^{\alpha\beta} = \sum_{p=0}^{N} \frac{1}{p!} m^{\gamma_1\cdots\gamma_p\alpha\delta}[\partial_{\gamma_1}\cdots_{\gamma_p}F^\beta_{\cdot\delta}]_{z(s)} \tag{8.7}$$

for ϕ^α and $\phi^{\alpha\beta}$, where

$$m^{\beta_1\cdots\beta_n\alpha} = [2n/(n+1)]Q^{(\beta_1\cdots\beta_n)\alpha}. \tag{8.8}$$

One important question, however, remains unanswered: What conditions other than the symmetry and orthogonality properties and the equations (8.4) and (8.5) must be satisfied by a set of tensors for them to be the moments of an extended body containing no singularities? We saw in Sec. 3 that for the moments $j^{\cdots}$ there defined, if all but a finite number of them are known, the rest are uniquely determined, and clearly the same argument can be applied to the more complicated expressions for $T^{\alpha\beta}$ and J^α in terms of the moments (8.1) and (8.2). We have obtained en route certain necessary conditions, e.g., that the sum of the infinite series (3.31) must have an asymptotic form given by (3.35), and there will be a similar condition for the inertial moments, but it is probable that stronger conditions than this are required for sufficiency. We must leave this question open, but we close with the conjecture that, given any *finite* set of tensors with the necessary symmetry and orthogonality properties, we *can* find a $T^{\alpha\beta}$ and a J^α that has these among its moments.

ACKNOWLEDGMENTS

The author is indebted to Dr. J. Ehlers and Dr. G. L. Hinds for several stimulating discussions that initiated this work, and to Dr. J. Weber and the Department of Physics and Astronomy of the University of Maryland for the award of a Research Associateship, during the tenure of which this work was carried out.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Finiteness of the Number of Positive-α Landau Surfaces in Bounded Portions of the Physical Region*

HENRY P. STAPP
Lawrence Radiation Laboratory, University of California, Berkeley, California

(Received 29 November 1966)

It is shown that if the spectrum of physical particle rest masses contains neither accumulation points nor the zero point, then the number of different positive-α Landau surfaces that enter any bounded portion of the physical region of any multiple-particle scattering process is finite. This implies that if the physical-region singularities of scattering functions are confined to the closure of the set of points lying on positive-α Landau surfaces, then the scattering functions are analytic at almost all points of the physical region. The proof is made by proving an equivalent property of systems of classical point particles scattering via point interactions.

I. INTRODUCTION

THERE are a number of reasons for believing that the physical-region singularities of scattering functions are confined to the closure $\mathcal{M}^+$ of the set of points lying on positive-α Landau surfaces. This restriction holds for the terms of the perturbation expansion in field theory.[1] It also follows directly from a macroscopic causality condition on the mass-shell S matrix.[2] And recent works have shown how, in simple cases, the singularities on positive-α Landau surfaces are precisely the ones that emerge from the assumption that the only physical region singularities of scattering functions are those generated by the unitarity equations.[3]

The supposition that the physical-region singularities of scattering functions are confined to $\mathcal{M}^+$ does not immediately ensure that the scattering functions are anywhere analytic; the conceivable alternative is that the positive-α Landau surfaces are everywhere dense in the physical region.

For the simplest case of the scattering of two initial particles into two final particles, each positive-α Landau surface is a normal threshold manifold, which is a manifold lying at a value of the total center-of-mass energy E that equals the sum of the rest masses of a set of the physical particles. If the spectrum of the physical-particle rest masses does not include the value zero and has no accumulation points, then the number of these manifolds entering any bounded portion of the physical region is finite. This ensures that the set of points not lying on $\mathcal{M}^+$ is everywhere dense; almost every physical-region point has a neighborhood that contains no point lying on any positive-α Landau surface.

The object of the present work is to show that this result carries over to reactions of arbitrary numbers of particles. It is shown that if the spectrum of physical-particle rest masses contains neither the value zero nor accumulation points, then the number of different positive-α Landau surfaces entering any bounded portion of the physical region is finite. Since the complement of the closure of any single positive-α Landau surface is everywhere dense in the physical region,[2] the same is true of any finite sum of such surfaces. Thus the assumption that the physical-region singularities are confined to $\mathcal{M}^+$ entails that each scattering function be analytic at almost every point of the physical region.

The result just stated was used in a recent work of the author on the crossing properties of the S matrix.[4] It has also been a tacit assumption in many other works in analytic S-matrix theory.

Coleman and Norton[5] have recently emphasized that the set of physical-region points lying on the positive-α Landau surface corresponding to a Landau diagram D is precisely the set of points such that the classical point-particle multiple-scattering process pictured by D is dynamically possible. By definition, each point on a positive-α Landau surface of a given process is a point (in the space of the external energy-momentum vectors of this process) such that the Landau equations[1] associated with a corresponding Landau diagram D are satisfied. But the Landau loop

* This work was done under the auspices of the U.S. Atomic Energy Commission.

[1] L. D. Landau, Nucl. Phys. **13**, 181 (1959).

[2] C. Chandler and H. P. Stapp, "S-Matrix Causality Conditions and Physical-Region Analyticity Properties" (to be published).

[3] P. V. Landshoff and D. I. Olive, J. Math. Phys. **7**, 1464 (1966); M. J. W. Bloxham, Nuovo Cimento **44**, 794 (1966); P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, J. Math. Phys. **7**, 1600 (1966). J. Storrow, Nuovo Cimento **48A**, 593 (1967): J. Coster and H. P. Stapp, "Physical-Region Discontinuity Equations" (submitted to J. Math. Phys.).

[4] H. P. Stapp, "Crossing, Hermitian Analyticity and the Connection Between Spin and Statistics," Lawrence Radiation Laboratory Report UCRL-16816 (1966).

[5] S. Coleman and R. Norton, Nuovo Cimento **38**, 438 (1965).

equations are precisely the requirement that the Landau diagram D can be converted into a space–time diagram $\bar{D}$ of the same structure by simply replacing each line L_i of D by the "space–time" vector $\alpha_i q_i$, where q_i and α_i are the momentum-energy vector and parameter α associated with line L_i. The parameter α_i is considered to be a proper time divided by the mass of particle i, and the positive-α condition corresponds to the requirement that the particles move forward in time.[5] The other two Landau equations ensure that the energy-momentum conservation laws are satisfied at each vertex and that each particle has the correct mass value. The space–time locations of the classical particles are not restricted, but those of the external particles are in fact determined by the gradient to the Landau surface.[2]

By virtue of this Coleman–Norton correspondence the number of Landau surfaces that enters a given portion of the physical region is the same as the number of classical point-particle multiple-particle scattering processes that are dynamically possible in this portion of the physical region. However, it may happen that several different multiple-scattering processes give Landau surfaces that exactly coincide with one another. Such Landau surfaces are, in our terminology, not "different" and need be counted only once.

As an example, suppose the sum of the masses of a set of physical particles S_1 is equal to the sum of the masses of a set of physical particles S_2. If the particles of the set S_1 are all relatively at rest, then this set can convert into the set of particles S_2, all relatively at rest. An unlimited number of conversions back and forth between these two sets of relatively-at-rest particles can evidently take place without affecting the kinematical situation. This permits an unlimited number of different Landau diagrams to be compatible with certain fixed points in momentum space. However, the Landau surfaces corresponding to these different diagrams all lie exactly on top of one another, and hence are not different. Since it is the number of different Landau surfaces that must be shown finite, various Landau diagrams (or their corresponding physical processes) that differ only by conversions of this type can be considered equivalent, and collisions effecting such conversions (called trivial collisions) can be disregarded. This fact is used continually, without explicit mention, in the following proofs.

The possibility of selection rules can be ignored, as they would only decrease the number of possible processes. Each particle can therefore be identified by its mass alone; an interchange of the identities of particles of the same mass would not alter a Landau surface.

II. PROOF FOR A ONE-DIMENSIONAL WORLD

A proof is given first for a world of one space and one time dimension. Let the space and time coordinates be called x and t and consider a plot of the particle trajectories on a two-dimensional x–t diagram. The trajectory lying at the most positive value of x is called the first trajectory and its slope is called the velocity of the first particle. This trajectory will generally have a number of straight line segments joined at points called the collisions of the first particle. The trajectory lying at the second highest value of x is called the second trajectory, and its slope is called the velocity of the second particle. There may, of course, be several particles that trace out a segment of the first trajectory. Their velocities are all equal to the velocity called the "velocity of the first particle," etc.

At each collision of the first particle the velocity of the first particle increases. This fact is obvious in the collision center-of-mass frame, and is carried to the general frame by a Lorentz transformation.

Let $S(E)$ be the set of all multiple-scattering processes (of point particles with point interactions) possible with a total center-of-mass energy less than E, and let $N(E)$ be the least upper bound on the number of collisions of the first particle for processes in $S(E)$. Our main problem is to show that $N(E)$ is finite for finite E; the remainder of the proof is then easy.

Let $m > 0$ be the mass of smallest-mass particle. For $2m < E < 3m$ only two-particle processes are contained in $S(E)$. In this case we have $N(E) = 1$; a system consisting of just two particles can evidently have at most one collision of the first particle. Let E_1 be the least upper bound on the values of E such that $N(E)$ is finite. Then the requirement that $N(E)$ be finite for finite E is equivalent to the requirement that E_1 be infinite.

Suppose E_1 is finite. Then for any positive integer n there must be a reaction R_N at center-of-mass energy $E < E_1 + (m/n)$ with at least $N = n^2$ collisions of the first particle. In this reaction the total change in v_1, the velocity of the first particle, cannot be more than

$$\Delta V = 2[(E_1 + m)^2 - 4^2]^{\frac{1}{2}}/(E_1 + m),$$

which is the change it would have if both an initial and final energy of $E_1 + m$ were divided between two minimal mass particles. Pick out the n collisions of the first particle of R_N that give the greatest change in v_1, or, more generally, such that the smallest change

δv_1 in these n collisions is not smaller than the largest change δv_1 in the remaining collisions. These n collisions divide the first trajectory of R_N into $n+1$ segments at least one of which has at least $n-1$ collisions. The change δv_1 at each collision of the first particle occurring in the interior of any of these $n+1$ segments must be less than $\Delta V/n$, since otherwise the total variation of v_1 would be more than ΔV. Thus we have shown that for any positive integer n there is in $S[E_1 + (m/n)]$ a reaction R_{n-1} with $n-1$ collisions of the first particle such that the δv_1 at each of these collisions is less than $\Delta V/n$.

Any given reaction may have a portion in which the particles can be divided into groups such that the particles in each group collide only with each other. For any reaction in $S(E_1 + m)$ the number of collisions of the first particle in any such portion must be bounded, since otherwise the least upper bound E_1 could be lowered. Thus for sufficiently large n it is not possible that the particles of the reaction R_{n-1} just constructed are divided into groups of particles that interact only among themselves; all the particles of R_{n-1} are connected to one another by collisions, for sufficiently large n.

By virtue of the above arguments, if E_1 is finite, there must be, for every positive integer n, a reaction R_{n-1} at a center of mass energy $E < E_1 + (m/n)$ having $n-1$ collisions of the first particle, and such that at each of these $n-1$ collisions the change δv_1 of v_1 is less than $\Delta V/n$. In this reaction R_{n-1}, let δ_i be the maximum magnitude of the difference between the velocity of the first particle and the velocity of the second particle at the ith collision of the first particle. And let δ be the greatest of the δ_i. Because δv_1 is less than $\Delta V/n$ it follows that δ must be less than $b\Delta V/n$, where b is $(E_1 + m)/m$. This limit comes from the optimal case in which the second particle has the least possible mass m, and the first particle has mass E_1, which is an upper bound.

Between collisions with the first particle the velocity of the second particle must increase monotonically. It follows from this that the difference in velocities of the first and second particles must always be less than δ, and hence also less than $b\Delta V/n$.

In order that the velocity of the second particle always differ by less than $b\Delta V/n$ from the velocity of the first particle, the maximum change of velocity of the second particle in a collision with a third particle must be less than $2b\Delta V/n$. This means, in turn, that the velocity of the third particle can differ from that of the second by at most $2b^2\Delta V/n$. (In the special case where the third and second particles can simultaneously collide with the first particle, the condition on the velocity of the third particle is more stringent.)

Reapplication of the same arguments shows that the difference between the velocities of the third and fourth particles is always less than $4b^3\Delta V/n$, and so on. Since the total number of particles in the reaction must be less than $b = (E_1 + m)/m$, it follows that the velocities of all the particles of R_{n-1} must, in the center of mass frame, be less than $C\Delta V/n$, where $C = \frac{1}{2}(2b)^b$ is a constant determined by E_1/m. From this limit on the center-of-mass velocities one obtains as an upper bound on the center-of-mass kinetic energy the value[6] $\frac{1}{2}(E_1 + m)(C\Delta V/n)^2$.

According to the above result, the kinetic energy of the particles of R_{n-1} approaches zero as n approaches infinity. This requires that E_1 be equal to the sum of rest masses of some set of physical particles, and that for n larger than some finite value L, the sum of the rest masses of the particles of R_{n-1} be equal to E_1. The limit L is the greater of the two values L_1 and L_2, where L_1 is defined by $[\tilde{E}_1 - E_1 - (m/L_1)] = 0$, $\tilde{E}_1$ being the smallest sum of rest masses that is greater than E_1, and L_2 is the least upper bound on values of n that satisfy

$$E(n^2) - \tilde{E}_1 \leq \tfrac{1}{2}(E_1 + m)(C\Delta V/n)^2,$$

where $\tilde{E}_1$ is the greatest sum of particle rest masses that is less than E_1, and $E(N)$ is the least upper bound on values of E' such that $N(E')$ is less than or equal to N. That L_2 is finite is ensured by the fact that $E(N)$ is a nondecreasing function of N that approaches E_1 as N becomes infinite. For n larger than L_1, the condition $E < E_1 + m/n$ ensures that the sum of the rest masses of particles of R_{n-1} is not larger than E_1. For n larger than L_2, the bound on the kinetic energy ensures that the sum of the rest masses of particles of R_{n-1} is not less than E_1. Thus for n larger than L, the sum of the rest energies of particles of R_{n-1} is precisely E_1.

Take n greater than L. Then the kinetic energy is $\epsilon = E - E_1$. But then the total variation of the velocity of the first particle in R_{n-1} is bounded by the value Δv defined by[6] $\frac{1}{2}m(\Delta v/2)^2 = \epsilon$.

We now repeat the arguments given before, but with $R_{n-1} = R_{r^2}$ in place of R_N and with Δv in place of ΔV. In place of the earlier bound $\frac{1}{2}(E_1 + m) \times (C\Delta V/n)^2$ on the kinetic energy, we now get $\frac{1}{2}(E_1 + m)(C\Delta v/r)^2$. That is, we have, for n larger than L,

$$\epsilon < \tfrac{1}{2}(E_1 + m)\left(\frac{C\Delta v}{r}\right)^2 = \frac{4(E_1 + m)C^2\epsilon}{m(n-1)}.$$

This gives $n - 1 < (2b)^{2b+1}$, which requires n, and hence $N(E)$, to be finite for finite E.

[6] Relativistic formulas are used throughout.

Since $N(E)$ is finite, the finite number of collisions of the first particle divides any reaction into a finite number of subreactions. In each of these subreactions the first trajectory is separated from the other trajectories. Thus the previous argument can be applied to the subreaction, but with the second trajectory in place of the first, etc. One concludes, then, that the number of collisions of the second particle is bounded by $[N(E) + 1]^2$.

Proceeding in this way one concludes that the total number of collisions is bounded, if the center-of-mass energy E is bounded. Since also the number of different types of particles that can enter into reactions in a bounded center-of-mass energy region is bounded, the total number of different types of reactions that can occur in such a region is bounded, in the one-dimensional case.

III. EXTENSION OF ONE-DIMENSION PROOF TO THE THREE-DIMENSIONAL CASE

The foregoing proof for the one-dimensional case can be easily generalized to the three-dimensional case. To do this we consider simultaneously the first x trajectory, the first y trajectory, and the first z trajectory. These are the trajectories lying at the largest values of x, y, and z, respectively.

A "triple" is a sequence of collisions (in their natural order) that includes at least one collision of the first x particle, at least one collision of the first y particle, and at least one collision of the first z particle. A "sequence of triples" is a sequence of collisions (in their natural order) that are separated into an ordered set of triples such that the final collision of any triple is earlier than the earliest collision of the next triple.

Let $N(E)$ be the least upper bound on the number of triples in sequences of triples for reactions in $S(E)$. Let E_1 be the least upper bound on the values of E such that $N(E)$ is finite. Suppose (contrary to fact) that E_1 is finite. Then for any positive integer n there must be a reaction R_{3N} in $S[E_1 + (m/n)]$ with a sequence of at least $3N = 3n^2$ triples. Pick out those n triples that contain the collisions of the first x particle with the n largest values of δv_{ix}. Here δv_{ix} is the change in the x component of velocity of the first x particle. Pick out also those n triples that contain the n collisions of the first y particle with the largest δv_{iy}. Do the same also for z. These $3n$ triples separate the reaction R_{3N} into at most $3n + 1$ (sub)reactions, at least one of which, called R_{n-1}, must contain at least $n - 1$ triples.

In the reaction R_{n-1} the maximum possible change of δv_{ix} at any collision of the first x particle is less than $\Delta V/n$ for exactly the same reasons as before. This bound holds also for the changes δv_{iy} and δv_{iz}. Thus the same arguments as in the one-dimensional case now give $\frac{3}{2}(E_1 + m)(C\Delta V/n)^2$ as an upper bound on the center-of-mass kinetic energy of R_{n-1}, where C is again $\frac{1}{2}(2b)^b$.

Continuing as before, one concludes that E_1 must be equal to the sum of rest masses of some set of physical particles, and that for sufficiently large n, the sum of the rest masses of the particles of R_{n-1}, must be precisely E_1. Thus the center-of-mass kinetic energy for R_{n-1} is again $\epsilon = E - E_1$ for sufficiently large n.

If the kinetic energy is ϵ, then the total variation of v_{ix} in R_{n-1} is no more than Δv defined $(\frac{1}{2}m)(\frac{1}{2}\Delta v)^2 = \epsilon$. From this one concludes, by the same argument as before, that $N(E)$ cannot become infinite at finite E.

From the fact that $N(E)$, the maximum number of triples in $S(E)$, is finite, it follows that the maximum number of collisions of reactions in $S(E)$ is finite. To show this we proceed as follows: For any reaction in $S(E)$ let t_1 be the greatest time such that the subreaction consisting of the portion of the reaction occurring at $t < t_1$ contains no triple. Then let t_2 be the largest time such that the portion of the reaction occurring in the interval $t_1 < t < t_2$ contains no triple. Let $t_3, t_4, \cdots, t_n$ be defined in the analogous way. This sequence of times must terminate at a time t_n with $n < N(E)$, where $N(E)$ is the maximum number of triples for reactions in $S(E)$.

The times t_i divide the original reaction into a set of no more than $[N(E) + 1]$ subreactions each of which contains no triple. But a subreaction that contains no triple must have a first x, y, or z trajectory that is disjoint from the other x, y, or z trajectories, respectively. This implies that the particles of each of the various subreactions must separate into groups such that the particles of each group interact only with each other. And at least one of these groups must be confined to a single x, y, or z trajectory. The number of these independent groups into which a subreaction divides is evidently no greater than E/m, which is an upper bound on the number of particles in reactions in $S(E)$.

The above analysis takes the original reaction into no more than $(E/m)[N(E) + 1]$ independent (i.e., self-interacting) new reactions. The same analysis is next applied to each of these independent new reactions. For reactions that are confined to a single x, y, or z trajectory, one uses, however, "doubles" instead of "triples." Doubles are the two-dimensional analog of triples; one eliminates the x, y, or z coordinate if the reaction is confined to an x, y, or z trajectory, respectively. The number of doubles (and analogously,

of singles) in reactions in $S(E)$ are shown to be bounded by essentially the same argument as was just given for triples. In fact, the bound $N(E)$ obtained previously for the number of triples is also a bound on the number of doubles (and also on the number of singles).

The original reaction is separated at the first stage of the analysis into no more than $(E/m)[N(E) + 1]$ independent new reactions. Each of these is separated at the second stage into no more than $(E/m)[N(E) + 1]$ new independent reactions. One continues in this way until the stage at which no further decomposition is obtained. This must occur after no more than E/m iterations, since the energy of a part that decomposes at a given stage must have been reduced at every earlier stage by at least m, due to the separation into independent parts. Thus the total number of independent reactions that are picked out altogether is no more than

$$\lambda(E) \equiv \{(E/m)[N(E) + 1]\}^{(E/m)+1}.$$

Also, the total number of times t_i singled-out in the entire course of the analysis is bounded by $\lambda(E)$.

Each (nontrivial) collision of the original reaction occurs at one of the times t_i singled-out in the above analysis. For, on the one hand, the only collisions that are eliminated at any stage of the analysis are those occurring at one of the singled-out times. On the other hand, the analysis does not terminate as long as any nontrivial collision remains. In the first place any independent part that is not confined to a single x, y, or z trajectory must contain a triple, since otherwise a trajectory can be separated out. Thus the analysis cannot terminate as long as there are still independent parts not confined to an x, y, or z trajectory. Similarly, all parts confined to a single x, y, or z trajectory must be reduced, before the analysis terminates, to parts lying on at least two trajectories. These parts lying on two trajectories are confined to one-dimensional subspaces. For such parts the analysis proceeds until the first trajectory becomes the same as the last trajectory, since otherwise the first trajectory of an independent part must have a collision. One is left, finally, with only trivial collisions.

The conclusion from the above arguments is that the collisions of any reaction in $S(E)$ are confined to a set of times t_i, the number of which is no more than $\lambda(E)$. Furthermore, the number of different types of particles that can participate in reactions in $S(E)$ is also finite, due to the spectral conditions on the particle rest masses. (Particles of the same mass can be identified, as mentioned in the Introduction.) But a finite number of different particles colliding only at a bounded number of different times can give only a finite number of different types of reactions. In particular, if $mq(E)/E$ is the number of different kinds of particles with rest energy less than E, then the number of different possible collisions at a single given time t_i is no more than $[2q(E)]^{2E/m}$. Thus the total number of different types of multiple-scattering reactions involving collisions at no more than $\lambda(E)$ different instants of time is no more than $(2q)^{2E\lambda/m}$. This upper bound on the number of different types of collisions possible in the portion of the physical region lying at center-of-mass energy less than E could be lowered with but little extra effort, should the need arise.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Convergence of the Bremmer Series for the Spatially Inhomogeneous Helmholtz Equation

H. L. BERK,* D. L. BOOK,† AND D. PFIRSCH‡
International Atomic Energy Agency, International Center for Theoretical Physics, Trieste, Italy

(Received 7 July 1966)

The convergence of the Bremmer series expansion for the solution of the one-dimensional Helmholtz equation with varying wave number is investigated. It is proved that the series converges provided the quantity $\epsilon\sigma$ is sufficiently small, where ϵ measures the relative rate of change $1/k^2(dk/dx)$ and σ the relative total change of the wave number k. An exactly soluble example is discussed in order to show that convergence fails if the above criterion is significantly relaxed. It is shown that if the Bremmer method is applied to the calculation of the time-dependent response to an externally imposed signal, the series is convergent at any finite time after the signal is turned on.

I. INTRODUCTION

THE Bremmer series was originally constructed[1] in order to obtain an expression for the reflection coefficient of the one-dimensional Helmholtz equation,

$$d^2\phi/dx^2 + k^2(x)\phi = 0, \qquad (1)$$

in the interval $-\infty < x < \infty$, where $k^2(x)$ is greater than zero and approaches constant values at infinity. The virtue of this series is that it can be derived from a simple physical picture based on Huygens' principle. At every point in the medium where $dk/dx \neq 0$, a wave is partly reflected and partly transmitted. The lowest order term $\phi_1^{(0)}$ is calculated by neglecting the further history of all the reflected "wavelets." $\phi_1^{(0)}$ is then found to be the first-order WKB solution for forward propagation. The next term, $\phi_2^{(1)}$, is composed of a coherent superposition of the reflected wavelets propagating in the negative direction, where again the wavelets generated by additional reflections are ignored in this order. The next term, $\phi_1^{(2)}$, represents a correction to the transmitted wave $\phi_1^{(0)}$ and gives the contribution from wavelets that have undergone just two reflections and then propagate to $+\infty$, and so on. Thus the principle of the Bremmer series is that the total field is composed of a coherent superposition of multiply reflected wavelets, where the reflection coefficient is determined by local gradient parameters.

This principle has been used by others in equations more complicated than Eq. (1). For example, Bellman and Kalaba[2] have used a variation of the method to study $2N$th-order differential equations. Aamodt and Book[3] have studied a fourth-order differential equation for a plasma fluid and Berk, Horton, Rosenbluth, and Sudan[4] apply the Bremmer method to obtain wave reflection arising from the Vlasov equation. We see that the Bremmer method can be applied to a large class of problems where mathematical proofs and alternative descriptions[5,6] are difficult to obtain, and it is therefore important to establish the regime of validity of the Bremmer method for the simplest case of application, the Helmholtz equation.

It has been established that the Bremmer series is a solution of the Helmholtz equation (or in the paper of Aamodt and Book, the fourth-order fluid equation) if the series converges.[2,3] However, previous convergence proofs have used very general properties of $k(x)$ and as a result have obtained sufficiency criteria for convergence only if the total change of k is of order k itself. For example, the convergence criteria obtained by Bellman and Kalaba[2] for continuous k are $k^2 > a^2 > 0$ and $\int_{-\infty}^{\infty} |k'|\, dx < 2a$, and by Atkinson[7] for continuously differentiable $k(x)$, $k^2 > 0$, and $\int_{-\infty}^{\infty} |k'/k|\, dx < \pi$. Atkinson also shows that there exists continuously differentiable $k(x)$ for which the Bremmer series diverges if $\int_{-\infty}^{\infty} |k'/k|\, dx > \pi$. However, since physically interesting forms of k arise that violate the above convergence criteria, we investigate if stronger convergence statements exist when more information is given about $k(x)$ and $k'(x)$ than just their continuity.

* University of California, San Diego, California.
† General Atomic Div. of General Dynamics Corp., San Diego, California.
‡ Max Planck Institut für Physik und Astrophysik, Munich, Germany.

[1] H. Bremmer, *The Theory of Electromagnetic Waves, A Symposium* (Interscience Publishers, Inc., New York, 1951), pp. 169–179.
[2] R. Bellman and R. Kalaba, J. Math. Mech. **8**, 683 (1959).
[3] R. E. Aamodt and D. L. Book, Phys. Fluids **9**, 143 (1966).
[4] H. L. Berk, C. W. Horton, M. N. Rosenbluth, and R. N. Sudan, ICTP preprint, Trieste (1966).
[5] F. L. Ince, *Ordinary Differential Equations* (Dover Publications, Inc., New York, 1956), p. 520.
[6] G. Knorr and D. Pfirsch, Max-Planck-Institut für Physik und Astrophysik, MPI-PAE/PI15/65, 1965.
[7] D. Atkinson, J. Math. Anal. Appl. **1**, 255 (1960).

Specifically, it is known that the first term of the Bremmer series is the lowest-order WKB solution, and hence we might expect that the Bremmer series converges when the WKB approximation

$$k'(x)/k^2(x) \ll 1$$

is applicable even when the criteria of Bellman, Kalaba, and Atkinson are violated. Here we present a convergence proof that when applied specifically to the case when the WKB approximation is valid, has a much greater range of convergence than previously obtained. However, we can show by means of a specific example that if the change of k is sufficiently large, the Bremmer series diverges.

It is perhaps surprising that a method apparently based on physical construction should break down for a range of parameters that is physically sensible. To throw some light on this problem, we show that the Bremmer method does not break down if we treat the Helmholtz equation as a Laplace transform of the wave equation. We then show that the inverse Laplace transform, which depends on time, converges absolutely.

II. CONVERGENCE OF THE BREMMER SERIES

Bremmer has shown that if $k(x)$ is finite, nonvanishing, and piecewise smooth in the interval $-\infty < x < +\infty$, and

$$k(x) \to k_- = \text{const} \ (x \to -\infty),$$
$$k(x) \to k_+ = \text{const} \ (x \to +\infty),$$

then a solution of the Helmholtz equation can be written as

$$\phi(x) = \phi_1(x) + \phi_2(x), \tag{2}$$

where ϕ_1 and ϕ_2 satisfy the relations

$$\phi_1(x) = \phi_1^{(0)}(x) + \int_{-\infty}^{x} dx' \, \frac{k'(x')}{2k(x')} P^*(x', x)\phi_2(x'), \tag{3}$$

$$\phi_2(x) = -\int_{x}^{\infty} dx' \, \frac{k'(x')}{2k(x')} P(x', x)\phi_1(x'). \tag{4}$$

Here the P^* denotes the complex conjugate and

$$P(x', x) = \left[\frac{k(x')}{k(x)}\right]^{\frac{1}{2}} \exp\left[-i\int_{x'}^{x} k(y)\, dy\right]. \tag{5}$$

In (3), $\phi_1^{(0)}$ is the first-order WKB solution

$$\phi_1^{(0)}(x) = \left[\frac{k_-}{k(x)}\right]^{\frac{1}{2}} \exp\left[i\int_0^x k(y)\, dy\right].$$

The boundary conditions satisfied by this solution are that $\phi(x)$ asymptotically approach the forms

$$\phi(x) = e^{i(k^- x - \beta)} + re^{-ik^- x} \quad (x \to -\infty),$$
$$\phi(x) = e^{ik^+ x} \quad (x \to \infty).$$

Here $\beta = \int_{-\infty}^{0} (k^- - k)\, dx$, $\exp[i(k^- x - \beta)]$ is an incident wave of unit amplitude and zero phase at $x = 0$, and r and t are the reflection and transmission coefficients.

The solution with different boundary conditions is obtained by varying the amplitude of the incident wave or introducing the first-order WKB solution for propagation in the negative direction, $\phi_2^{(0)}$. An outline of the derivation of the Bremmer Eqs. (3) and (4) is given in Appendix A.

When (3) and (4) are solved by the obvious iteration, the Bremmer series results. However, first it is convenient to define the new variables

$$\xi(x) = 2\int_0^x k(y)\, dy,$$

$$u_1(x) = \phi_1(x)\left(\frac{k(x)}{k_-}\right)^{\frac{1}{2}} \exp\left[-i\int_0^x k(y)\, dy\right],$$

$$u_2(x) = \phi_2(x)\left(\frac{k(x)}{k_-}\right)^{\frac{1}{2}} \exp\left[i\int_0^x k(y)\, dy\right],$$

In terms of these, we have

$$u_1(\xi) = 1 + \int_{-\infty}^{\xi} d\eta \delta(\eta) e^{-i\eta} u_2(\eta), \tag{7}$$

$$u_2(\xi) = -\int_{\xi}^{\infty} d\eta \delta(\eta) e^{i\eta} u_1(\eta), \tag{8}$$

where

$$\delta(\xi) = \frac{1}{2k}\frac{dk}{d\xi} = \frac{1}{2k^2}\frac{dk}{dx}.$$

Equations (7) and (8) are solved by iteration if we put

$$u_1^{(0)}(\xi) \equiv 1 \tag{9}$$

and define

$$u_1^{(2n)}(\xi) = \int_{-\infty}^{\xi} d\eta \delta(\eta) e^{-i\eta} u_2^{(2n-1)}(\eta) \tag{10}$$

and

$$u_2^{(2n+1)}(\xi) = -\int_{\xi}^{\infty} d\eta \delta(\eta) e^{i\eta} u_1^{(2n)}(\eta). \tag{11}$$

Then formally

$$u_1(\xi) = \sum_{n=0}^{\infty} u_1^{(2n)}(\xi), \tag{12}$$

$$u_2(\xi) = \sum_{n=0}^{\infty} u_n^{(2n+1)}(\xi) \tag{13}$$

satisfy Eqs. (7) and (8). Equations (10) and (11) may be combined to connect successive orders in the expansion of u_1 alone:

$$u_1^{(2n+2)}(\xi) = -\int_{-\infty}^{\xi} d\eta \delta(\eta) e^{-i\eta} \int_{\eta}^{\infty} d\eta' \delta(\eta') e^{i\eta'} u_1^{(2n)}(\eta')$$
$$= \int_{-\infty}^{\infty} d\eta' \delta(\eta') e^{i\eta'} G(\xi, \eta') u_1^{(2n)}(\eta'), \tag{14}$$

where

$$G(\xi, \eta') = \begin{cases} F(\xi), & \xi < \eta', \\ F(\eta'), & \xi \geq \eta', \end{cases} \tag{14a}$$

$$F(\xi) = \int_{-\infty}^{\xi} d\eta e^{-i\eta} \delta(\eta). \tag{14b}$$

Suppose temporarily that $k(x)$ increases monotonically from k_- to k_+ and define a new independent variable

$$s = \int_{-\infty}^{\xi} d\eta \delta(\eta) = \frac{1}{2} \int_{-\infty}^{\xi} \frac{dk}{k}.$$

Thus $s(\xi)$ is just $\frac{1}{2} \ln [k(\xi)/k_-]$; $s(-\infty) = 0$ and $0 \leq s \leq \frac{1}{2} \ln (k_+/k_-) = \sigma$. Equation (14) becomes

$$u_1^{(2n+2)}(s) = -\int_0^{\sigma} dt K(s, t) u_1^{(2n)}(t), \tag{15}$$

where

$$K(s, t) = e^{i\xi(t)} G[\xi(s), \xi(t)]. \tag{16}$$

By a well-known theorem,[8] the series (12) converges if the operator k has the following norm less than 1:

$$\int_0^{\sigma} ds \int_0^{\sigma} dt \, |K(s, t)|^2 < 1. \tag{17}$$

By (16) this condition is

$$\int_0^{\sigma} ds \int_0^{\sigma} dt \, |F[\xi(t)]|^2 + \int_0^{\sigma} ds \, |F[\xi(s)]|^2 (\sigma - s) < 1. \tag{18}$$

Now let $\delta(\xi)$ take on values less than zero so that $\delta = 0$ at a discrete set of points. We now define the total logarithmic variation of k from $-\infty$ to ξ as twice the quantity

$$s(\xi) = \int_{-\infty}^{\xi} d\eta \, |\delta(\eta)|.$$

Thus we have a one-to-one correspondence between ξ and s. Equation (14) now takes the form

$$u_1^{(2n+2)}(s) = -\int_0^{\sigma} dt K(s, t) u_1^{(2n)}(t), \tag{15'}$$

where

$$K(s, t) = \alpha(t) e^{i\xi(t)} G[\xi(s), \xi(t)]. \tag{16'}$$

Here $\alpha(t)$ is the sign of $\delta[\xi(t)]$ and $\sigma = s(\infty)$. The remaining arguments are as before and the general convergence criterion is given by (18) with s, σ, and $K(s, t)$ as redefined above.

We now investigate the implications of (18). If we have a bound $|F|^2_{\max}$ for $|F(t)|^2$, then (18) becomes

$$\tfrac{3}{2} \sigma^2 |F|^2_{\max} < 1. \tag{19}$$

By the definition (14b) of F it is clear that

$$|F(s)| \leq s \leq \sigma.$$

So we have the sufficiency criteria.

(A) If $\delta(x)$ is finite and piecewise smooth, then

$$\sigma < (\tfrac{2}{3})^{\frac{1}{2}}$$

implies the convergence of the Bremmer series (12). It is clear that the same criterion applies to (13).

Criterion (A) is slightly more restrictive but essentially the same as that given by Atkinson.[7] However, it is important to determine whether the Bremmer series converges in cases where the relative change in k is large. Notice that if k is sufficiently slowly varying, the usual WKB approximation is valid, and we would expect the Bremmer series to converge by its construction. Hence, we assume $\delta(\xi) \ll 1$ and introduce the following definitions in order to classify $k(\xi)$. A function $f(x)$ is defined to be gentle in the small quantity ϵ_0 if $f'(x)/f(x) = \epsilon_0 \alpha(x)$, where $\alpha(x)$ is order 1. Similarly $f(x)$ is said to be very gentle in ϵ_0 if $f''(x)/f(x) = \epsilon_0^2 \beta(x)$, $\beta(x) = O(1)$ and so on.

Now suppose $|\delta(\xi)| \leq \epsilon \ll 1$ for all ξ. Let $\delta(x)$ be gentle in ϵ_0. Then integrating (14b) by parts, we have

$$F(\xi) = i\delta(\xi) e^{-i\xi} - i \int_{-\infty}^{\xi} d\eta \delta'(\eta) e^{-i\eta},$$

$$|F(\xi)| \leq |\delta(\xi)| + \left| \epsilon \int_{-\infty}^{\xi} d\eta \delta(\eta) \alpha(\eta) e^{-i\eta} \right|$$

$$\leq \epsilon + \epsilon\alpha \int_{-\infty}^{\xi} d\eta \, |\delta(\eta)|$$

$$< \epsilon[1 + \alpha\sigma].$$

Then (19) yields the criterion

(B) If $\delta(\xi)$ is gentle in $\epsilon = \max |\delta(\xi)|$, then

$$(\tfrac{3}{2})^{\frac{1}{2}} \sigma\epsilon(1 + \alpha\sigma) < 1$$

implies convergence of the Bremmer series. Continuing to integrate by parts, we find the following.

(C) If $\delta(\xi)$ is very gentle in ϵ, then

$$(\tfrac{3}{2})^{\frac{1}{2}} \sigma\epsilon(1 + \beta\sigma\epsilon) < 1$$

implies convergence.

(D) If $\delta(\xi)$ is very very gentle in ϵ, then

$$(\tfrac{3}{2})^{\frac{1}{2}} \epsilon\sigma < 1$$

implies convergence.

It is thus clear that for "sufficiently gentle" wavenumbers k, the natural measure of convergence is $\epsilon\sigma$. Note that the gentleness of any order criterion can be slightly violated without changing the effective convergence criteria. For example, if in a small region of ξ, say $\Delta\xi$, δ changes rapidly but the total change $\Delta\delta$ in this region is much less than 1, then the convergence criterion (B) is still valid.

Before going on, it is interesting to compare the

[8] F. E. Riez and B. Sz-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 147.

Bremmer formulation with that of Knorr and Pfirsch.[6] They show that Eq. (1) is equivalent to

$$w(\xi) = w_0(\xi) + \tfrac{1}{2}\int_{-\infty}^{\xi} d\xi \sin \tfrac{1}{2}(\xi - \xi')f(\xi')w(\xi'), \tag{20}$$

where $w(\xi) = k^{\frac{1}{2}}(x)\phi(x)$ and

$$f(\xi) = \frac{2}{k}\frac{d^2k}{d\xi^2} - \left(\frac{1}{k}\frac{dk}{d\xi}\right)^2.$$

At $\xi \sim -\infty$, $w(\xi) \sim w_0(\xi) = \exp(i\xi/2)$; this boundary condition corresponds to making the upper limit in (8) $-\infty$ instead of ∞. Then writing

$$f(\xi) = 2k^{-\frac{1}{2}}\frac{d}{d\xi}\left(\frac{1}{k^{\frac{1}{2}}}\frac{dk}{d\xi}\right)$$

and integrating by parts, we obtain

$$w(\xi) = e^{i\xi/2} - \int_{-\infty}^{\xi} d\xi' \left[\frac{k'}{k}\sin \tfrac{1}{2}(\xi - \xi')w(\xi') - \frac{k'}{2k}\cos \tfrac{1}{2}(\xi - \xi')w(\xi') - \frac{1}{2}\left(\frac{k'}{k}\right)^2 \sin \tfrac{1}{2}(\xi - \xi')w(\xi')\right]. \tag{21}$$

From the Bremmer equations with boundary conditions modified as noted above, we get

$$w_B(\xi) = u_1 e^{i\xi/2} + u_2 e^{-i\xi/2},$$

$$= e^{i\xi/2} + \int_{-\infty}^{\xi} d\xi' \delta(\xi')\,[w_B(\xi')\cos \tfrac{1}{2}(\xi - \xi') + i\phi(\xi')\sin \tfrac{1}{2}(\xi - \xi')]$$

where

$$\phi(\xi) = u_1 e^{i\xi/2} - u_2 e^{-i\xi/2}.$$

But

$$i\phi(\xi) = 2[w_B'(\xi) - \delta(\xi)w_B(\xi)],$$

so w and w_B both satisfy Eq. (21); and thus $w_B = w$. In Ref. 6 it is shown that the series obtained by iterating (20), starting with w_0, converges absolutely provided only that

$$\int_{-\infty}^{\infty} |f(\xi)|\, d\xi < \infty.$$

As shown by the example of Sec. III, this is certainly not true for the Bremmer series obtained in the reflection problem. The difference is traceable to the definition of the boundary conditions in the two cases. Because the Bremmer problem includes specification of the behavior of the solution at $\pm\infty$, each order of the iteration process (14) involves integration over the whole range of variation of $k(\xi)$ and effects can accumulate from the entire ξ axis; in the series of Knorr and Pfirsch, the value of $w(\xi)$ depends on an iterated integral from the boundary at $-\infty$ to ξ only. Thus it is possible to majorize the series with an exponential series, implying absolute convergence. We return to this point in Sec. IV.

III. SPECIAL EXAMPLE

Let us consider as a special case the following example that can be solved exactly and readily expressed in terms of the Bremmer series. Consider

$$k(\xi) = \begin{cases} k_0, & -\infty < \xi < 0, \\ k_0 e^{\epsilon\xi}, & 0 \le \xi \le \xi_0, \\ k_0 e^{\epsilon\xi_0}, & \xi_0 < \xi < \infty. \end{cases}$$

For this example, (7) and (8) become

$$u_1(\xi) = 1 + \frac{\epsilon}{2}\int_0^{\xi} d\eta e^{-i\eta}u_2(\eta), \tag{22}$$

$$u_2(\xi) = -\frac{\epsilon}{2}\int_0^{\xi_0} d\eta e^{i\eta}u_1(\eta). \tag{23}$$

From the first derivatives of (22) and (23), we obtain the relations

$$u_1'(\xi) = \tfrac{1}{2}\epsilon e^{-i\xi}u_2(\xi), \tag{24a}$$

$$u_2'(\xi) = \tfrac{1}{2}\epsilon e^{i\xi}u_1(\xi). \tag{24b}$$

Using these relations we derive the following differential equation for $u_1(\xi)$:

$$u_1''(\xi) + iu_1'(\xi) - \tfrac{1}{4}\epsilon^2 u_1(\xi) = 0. \tag{25}$$

[This is a special case of the equation

$$u_1''(\xi) + \left(\frac{k'}{k} - \frac{k''}{k'} + i\right)u_1'(\xi) - \frac{k'}{4k^2}u_1(\xi) = 0$$

that would be derived for arbitrary $k(\xi)$.]

We obtain the appropriate boundary conditions from (22), (23), and (24a):

$$u_1(0) = 1, \quad u_1'(\xi_0) = 0. \tag{26}$$

The exact solution for $u_1(\xi)$ is then found to be

$$u_1(\xi) = \frac{e^{-\frac{1}{2}i\xi}[(\gamma + 1)e^{-\frac{1}{2}i\gamma(\xi - L)} + (\gamma - 1)e^{\frac{1}{2}i\gamma(\xi - L)}]}{2[\gamma\cos\phi - i\sin\phi]}, \tag{27}$$

where $\gamma = (1 - \epsilon^2)^{\frac{1}{2}}$ and $\phi = \frac{1}{2}\gamma\xi_0$.

If we substitute this solution into (23) and integrate, we find that $u_2(\xi)$ is given by

$$u_2(\xi) = \frac{-\epsilon e^{i\xi/2}\sin[\frac{1}{2}\gamma(\xi_0 - \xi)]}{[\gamma\cos\phi - i\sin\phi]}. \tag{28}$$

We obtain the Bremmer series either by expanding (27) and (28) in powers of ϵ or directly iterating (22)

and (23). The first few terms are

$$u_1^{(0)}(\xi) = 1,$$
$$u_2^{(1)}(\xi) = -(\epsilon/2i)[e^{i\xi_0} - e^{i\xi}],$$
$$u_1^{(2)}(\xi) = (\epsilon^2/4i)[\xi - i(e^{i(\xi_0-\xi)} - e^{i\xi_0})],$$
$$u_2^{(3)}(\xi) = -(\epsilon^3/8i)[-2i\xi_0 e^{i\xi_0} + i\xi(e^{i\xi_0} + e^{i\xi}) + e^{i\xi_0} - e^{i\xi} + e^{2i(\xi_0+\xi)}].$$

Notice that $u_2^{(1)}(0)$ is a good approximation to the reflected wave if $\epsilon^2\xi_0 \ll 1$. The next approximation to the reflection is significant only if $\xi_0 = 2\pi n + \delta$ where $\delta \lesssim \epsilon^2\xi_0$.

We see that the Bremmer series is a power series in ϵ in this example. The radius of convergence of this power series is given by the value of ϵ closest to zero for which $u_1(\xi)$ or $u_2(\xi)$ is singular. Both the points $\epsilon = 1$ and the zero of $\gamma \cos\phi - i \sin\phi$ seem likely candidates. However, the point $\epsilon = 1$ can be eliminated since, near $\epsilon = 1$, $u_1(\xi)$ and $u_2(\xi)$ can be expressed as a convergent power series in $\gamma = (1 - \epsilon^2)^{\frac{1}{2}}$. However, both (27) and (28) are even in γ and hence the power series is in $\gamma^2 = 1 - \epsilon^2$, which is also a power series in ϵ.

We now try to approximate the root of the equation

$$(1 - \epsilon^2)^{\frac{1}{2}} \cos[\tfrac{1}{2}(1 - \epsilon^2)^{\frac{1}{2}}\xi_0] = i \sin[\tfrac{1}{2}(1 - \epsilon^2)^{\frac{1}{2}}\xi_0].$$

By squaring this equation and using trigonometric identities, this equation can be written in the form

$$\cos^2\{[\tfrac{1}{2}(1 - \epsilon^2)^{\frac{1}{2}}]\xi_0\} = 1/\epsilon^2. \tag{29}$$

Let us assume the root occurs where $\epsilon^2\xi_0 \gg 1$ but $\epsilon^4\xi_0 \ll 1$. The square root of (29) can then be expressed approximately by

$$\cos(\tfrac{1}{2}\xi_0)\cos(\tfrac{1}{4}\epsilon^2\xi_0) + \sin(\tfrac{1}{2}\xi_0)\sin(\tfrac{1}{4}\epsilon^2\xi_0) = \pm(1/\epsilon).$$

If we define $p = -i\epsilon^2$ and assume that $\xi_0 = \frac{1}{2}\pi(2n + 1)$, where n is an integer, we find

$$e^{p^2\xi_0/4} = 1/p,$$

and therefore,

$$\xi_0 = (4/p^2)\ln(1/p). \tag{30}$$

Hence, if ϵ is small, the above approximations can be satisfied, and consequently the Bremmer series for our example diverges for a sufficiently large change in k.

The range of convergence of our special example should be compared with the criterion (D) of Sec. II. For our example, the total maximum logarithmic variation of k is given by $\sigma = \epsilon\xi_0$ so that from (30) the convergence criterion of our example is

$$\epsilon\sigma < \ln(1/|\epsilon|),$$

while the general criterion is $(3/2)^{\frac{1}{2}}\epsilon\sigma < 1$. This shows that the bounds of the previous section are almost optimal.

IV. TIME-DEPENDENT REFLECTION PROBLEM

It is of interest to consider how the Bremmer series can be applied to wave equations having both space and time dependence. We consider the simplest form of a one-dimensional wave equation

$$\left\{\frac{\partial^2}{\partial t^2} - c^2(x)\frac{\partial^2}{\partial x^2}\right\}\psi(x, t) = 0. \tag{31}$$

We investigate the response of this equation to a localized source turned on at time $t = 0$. Using a technique similar to one employed by Nyquist[9] in circuit theory, we show that the time-dependent Bremmer series for ψ converges absolutely.

If $c(x)$ is constant, we find that the normal mode solutions described by (31) are waves propagating in either direction without dispersion, i.e., ω/k constant.

Suppose now that $1/c(x)$ satisfies the restrictions placed on $k(x)$ in Sec. II (positive, finite, piecewise smooth). In addition, we assume that $c(x)$ is constant from $-\infty$ to some point on the real axis, and choose x_0 inside this interval. At $t = 0$, we begin to generate a time-dependent signal at the point x_0. Thus on the right-hand side of (31) there is a source term in the form

$$S(x, t) = \delta(x - x_0)\Theta(t)f(t). \tag{32}$$

[$\Theta(t)$ is the Heaviside step-function.] We assume that $f(t)$ is integrable,

$$\left|\int_0^\infty f(t)\,dt\right| < \infty, \tag{33}$$

and impose the initial conditions

$$\psi(x, 0) = (\partial\psi/\partial t)(x, 0) = 0.$$

Equation (31) with the source term (32) can now be solved by means of Laplace transformations. Defining

$$\phi(x, \omega) = \int_0^\infty dt e^{i\omega t}\psi(x, t), \tag{34}$$

etc., with the inversion

$$\psi(x, t) = \frac{1}{2\pi}\int_I d\omega e^{-i\omega t}\phi(x, \omega), \tag{35}$$

where I is a contour in the complex ω-plane above all singularities of $\phi(x, \omega)$, there results

$$-\omega^2\phi(x, \omega) - c^2(x)(d^2\phi/dx^2)(x, \omega) = \delta(x - x_0)f(\omega)$$

[9] H. Nyquist, Bell System Tech. J. 11, 126 (1932).

or

$$(d^2\phi/dx^2) + k^2(x)\phi = S(\omega)\delta(x - x_0). \quad (36)$$

Here

$$k(x) = \omega/c(x),$$
$$S(\omega) = -c^{-2}(x_0)f(\omega),$$

and $f(\omega)$ is the Laplace transform of $f(t)$.

The solution of (36) is

$$\phi = \phi_G + \phi_H, \quad (37)$$

where ϕ_G is the Green's function for a source localized at x_0 and ϕ_H is a solution of the homogeneous equation. The former is readily found in a neighborhood of x_0 using the fact that at x_0, $k(x) = k_0$ is a constant:

$$\phi_G(x) = Be^{-ik_0(x-x_0)} \quad (x < x_0), \quad (38a)$$
$$= Be^{ik_0(x-x_0)} \quad (x > x_0), \quad (38b)$$

where

$$B = S(\omega)/2ik_0 = -[c(x_0)/2i]f(\omega). \quad (39)$$

Causality requires that ϕ_H be the solution of the reflection problem for a wave of amplitude B incident from the left at $x = x_0$ of the form (38b). [By transforming back from ω to t, we see that if ϕ has a term like $\exp(-ikx)$ at $+\infty$, ψ does not vanish for $t < 0$.]

We can now use the Bremmer techniques of Sec. II to solve for ϕ_H:

$$\phi = \phi_1 + \phi_2,$$

where ϕ_1 and ϕ_2 satisfy (3) and (4) except that $\phi_1^{(0)}$ is multiplied by a factor B. We define $\tau(x) = \int_{x_0}^{x} dx/c(x)$, the time required for a wave to propagate from x_0 to x. In terms of the new variable, (3) and (4) become

$$u(\tau, \omega) = B + \int_0^\tau d\tau' \mathcal{E}(\tau') e^{-2i\omega\tau'} v(\tau'), \quad (40)$$

$$v(\tau, \omega) = -\int_\tau^\infty d\tau' \mathcal{E}(\tau') e^{2i\omega\tau} u(\tau'), \quad (41)$$

where

$$\mathcal{E}(\tau) = (1/2k)(dk/d\tau).$$

As before, an infinite series of functions u_n, v_n in (ω, τ) space is obtained by iteration of (40) and (41.) These yield a corresponding series $\Sigma(\psi_1^{(2n)}, \psi_2^{(2n+1)})$ in (τ, t) [or (x, t)] space on application of the Laplace inversion formula (35). Thus, using (39),

$$\psi_1^{(0)}(\tau, t) = \int_I \frac{d\omega}{2\pi} e^{-i\omega t}\phi^{(0)}(\tau, \omega),$$
$$= \int_I \frac{d\omega}{2\pi} e^{-i\omega t}\left[\frac{k_0}{k(\tau)}\right]^{\frac{1}{2}} e^{i\omega\tau} B(\omega), \quad (42)$$
$$= \frac{c^{-1}(x_0)}{2}\left[\frac{k_0}{k(\tau)}\right]^{\frac{1}{2}} \int_0^t dt' f(t).$$

By the assumption (33) and the conditions on $c(x)$, we can bound $\psi_1^{(0)}$:

$$|\psi_1^{(0)}(\tau, t)| \leq M. \quad (43)$$

Note that $[k_0/k(\tau)]^{\frac{1}{2}} = [c(\tau)/c_0]^{\frac{1}{2}}$ is independent of ω because of the absence of dispersion. For $t < \tau(x)$, $\psi_1^{(0)}$ vanishes, indicating that the signal does not have time to propagate from x_0 to x.

Likewise, applying (35) to the recursion relation for successive terms in the expansion of u,

$$u_{2(n+1)} = \int_0^\tau d\tau' \mathcal{E}(\tau') e^{-2i\omega\tau'} \int_{\tau'}^\infty d\tau'' \mathcal{E}(\tau'') e^{2i\omega\tau''} u_{2n}(\tau''), \quad (44)$$

we have

$$\psi_1^{2(n+1)}(\tau, t) = -\int_I \frac{d\omega}{2\pi} e^{-i\omega t}\left[\frac{k_0}{k(\tau)}\right]^{\frac{1}{2}} e^{i\omega\tau}$$
$$\times \int_0^\tau d\tau' \mathcal{E}(\tau') e^{-2i\omega\tau'} \int_{\tau'}^\infty d\tau'' \mathcal{E}(\tau'') e^{2i\omega\tau''}$$
$$\times \left[\frac{k(\tau'')}{k_0}\right]^{\frac{1}{2}} e^{-i\omega\tau''} \phi_1^{(2n)}(\tau'', \omega)$$
$$= -\int_0^\tau d\tau' \mathcal{E}(\tau') \int_{\tau'}^\infty d\tau'' \mathcal{E}(\tau'')$$
$$\times \left[\frac{k(\tau'')}{k(\tau)}\right]^{\frac{1}{2}} \psi_1^{(2n)}(\tau'', t - \tau - \tau'' + 2\tau'). \quad (45)$$

The following assumptions are now made:

$$|\mathcal{E}(\tau)| \leq E, \quad (46)$$

$$[k(\tau'')/k(\tau)]^{\frac{1}{2}} \leq K \quad (47)$$

for all τ, τ''; both (46) and (47) certainly hold for physically reasonable choices of $c(x)$.

We show by induction that the series

$$\sum_{n=0}^{\infty} \psi_1^{(2n)}(\tau, t) = \psi_1(x, t)$$

is majorized by a series which converges as an exponential series, and therefore itself converges absolutely for any finite t, τ.

Suppose

$$\psi_1^{(2n)}(\tau, t) = 0, \quad \tau > t, \quad (48)$$

$$|\psi_1^{(2n)}(\tau, t)| \leq (E^{2n}K^n/n!\, 2^{2n})M(t - \tau)^n(t + \tau)^n, \quad (t > \tau). \quad (49)$$

By (48), the integrand vanishes in (45) unless the variable τ'' satisfies

$$\tau' \leq \tau'' \leq \tau' + \tfrac{1}{2}(t - \tau). \quad (50)$$

But (50) can only hold if $t > \tau$, so that (48) is true

also for $\psi_1^{2(n+1)}$. We are left with

$$
\begin{aligned}
&|\psi_1^{2(n+1)}(\tau, t)| \\
&\quad = \left| \int_0^{\tau} d\tau' \delta(\tau') \int_{\tau'}^{\tau'+\frac{1}{2}(t-\tau)} d\tau'' \delta(\tau'') \left[\frac{k(\tau'')}{k(\tau)}\right]^{\frac{1}{2}} \right. \\
&\qquad \left. \times \psi_1^{(2n)}(\tau'', t - \tau - \tau'' + 2\tau') \right| \\
&\quad \leq \int_0^{\tau} d\tau' \int_{\tau'}^{\tau'+\frac{1}{2}(t-\tau)} d\tau'' E^2 K \frac{E^{2n} K^n}{n!\, 2^{2n}} M(t - \tau + 2\tau')^n \\
&\qquad \times (t - \tau - 2\tau'' + 2\tau')^n \\
&\quad \leq \frac{E^{2(n+1)} K^{n+1}}{n!\, 2^{2n}} M(t + \tau)^n \int_0^{\tau} d\tau' \\
&\qquad \times \int_{\tau'}^{\tau'+\frac{1}{2}(t-\tau)} d\tau''(t - \tau' - 2\tau'' + 2\tau')^n, \\
&\quad = \frac{E^{2(n+1)} K^{n+1}}{n!\, 2^{2n}} M \frac{(t + \tau)^n \tau}{n + 1} \left(\frac{t - \tau}{2}\right)^{n+1} \\
&\quad \leq \frac{E^{2(n+1)} K^{n+1}}{(n + 1)!\, 2^{2(n+1)}} M(t + \tau)^{n+1}(t - \tau)^{n+1}.
\end{aligned}
$$

Clearly, (43) is a special case of (49) and, for $\psi_1^{(0)}$, (48) follows from (42). So the induction is complete and

$$
\left| \sum_{n=0}^{\infty} \psi_1^{(2n)}(\tau, t) \right| \leq M \exp\left[\tfrac{1}{4} E^2 K (t + \tau)(t - \tau)\right]. \quad (51)
$$

In entirely analogous fashion we show that

$$
\psi_2^{(2n+1)}(\tau, t) = 0, \quad \tau > t,
$$

and

$$
\begin{aligned}
\psi_2^{(2n+1)}(\tau, t) &= \left| \int_{\tau}^{\frac{1}{2}(t+\tau)} d\tau' \delta(\tau') \left(\frac{k(\tau')}{k(\tau)}\right)^{\frac{1}{2}} \right. \\
&\qquad \left. \times \psi_1^{(2n)}(\tau', t - \tau' + \tau) \right| \\
&\leq EK \int_{\tau}^{\frac{1}{2}(t+\tau)} d\tau' M \frac{(E^2 K)^n}{2^{2n} n!} \\
&\qquad \times (t + \tau)^n (t - 2\tau' + \tau)^n, \\
&= \frac{MEK}{n + 1} \frac{(E^2 K)^n}{n!\, 2^n} (t + \tau)^n \left(\frac{t - \tau}{2}\right)^{n+1}.
\end{aligned}
$$

So

$$
\begin{aligned}
&\left| \sum_{n=0}^{\infty} \psi_2^{(2n+1)}(\tau, t) \right| \\
&\quad \leq \frac{2ME^{-1}}{t + \tau} \{\exp\left[\tfrac{1}{4} E^2 K (t + \tau)(t - \tau)\right] - 1\}. \quad (52)
\end{aligned}
$$

Together (51) and (52) comprise the result sought.

It is clear from the proof just presented that convergence is a consequence of the finite speed of propagation. The region in coordinate space which can contribute wavelets to the total Bremmer wave in any given order is limited by the condition $\tau(x) < t$, in contrast to the situation of Sec. II. In that case there is no initial point in time; waves have had effectively an infinitely long time in which to traverse all of coordinate space, and wavelets of any order of multiple reflection come from all points in space where $k' \neq 0$. Thus it is possible to find the terms in the Bremmer series increasing with order, so that the series diverges.

Because the idea of a finite speed of propagation is central to the above proof, it should be possible to generalize the argument to cases where the wave equation describes propagation in dispersive media, provided that the dispersion relation is such that ω/k is bounded.

We see that the Bremmer series for the Helmholtz equation is simply a formal mathematical expression. Only when time is brought into the picture can the Bremmer series be viewed strictly as a physical construction. Hence, only in the time domain is the Bremmer series guaranteed to be absolutely convergent, and it is not surprising that, for the Helmholtz equation, the Bremmer series can diverge.

V. CONCLUSION

We have presented a proof that greatly improves the convergence criterion for the Bremmer series of the Helmholtz equation in the case that

$$
1 \gg \delta(\xi) \equiv (2k)^{-1}\, dk/d\xi = (2k^2)^{-1}\, dk/dx
$$

and $\delta(\xi)$ is sufficiently gentle. In principle the method of proof should be applicable to higher-order differential equations and similar convergence criteria should be obtained.

It is important to determine if the first term for the reflected wave $\phi_2^{(1)}$ is a good approximation to the exact solution. It is clear from our equations that if $\delta(\xi)$ formally can be ordered in some small dimensionless parameter, say ϵ, then the Bremmer series is asymptotic in small ϵ. Thus if the reflection is not exponentially small, we have $\phi_2 = \phi_2^{(1)} + \theta(\epsilon^3)$. In fact, the example of Sec. III demonstrated this relationship explicitly. However, if $\delta(\xi)$ is sufficiently gentle and smooth, the reflection coefficient is indeed exponentially small. Then each term in the Bremmer expansion for ϕ_2 is exponentially small, and it is not clear if the first term dominates the remaining ones.

It is interesting to observe that the higher-order WKB solution can be obtained from the Bremmer series. Each term of the Bremmer series in x space is

explicitly given by

$$\phi_1^{(2n)}(x) = \left(\frac{k_-}{k}\right)^{\frac{1}{2}} e^{i\xi(x)/2} \int_{-\infty}^{\xi(x)} d\xi_{2n}\,\delta(\xi_{2n}) e^{-i\xi_{2n}} \times \int_{\xi_{2n}}^{\infty} d\xi_{2n-1}\,\delta(\xi_{2n-1}) e^{i\xi_{2n-1}} \int_{-\infty}^{\xi_{2n-1}} \cdots \times \int_{-\infty}^{\xi_2} d\xi_2 e^{-i\xi_2}\delta(\xi_2) \int_{\xi_2}^{\infty} d\xi_1 e^{i\xi_1}\delta(\xi_1),$$

$$\phi_2^{(2n-1)}(x) = \left(\frac{k_-}{k}\right)^{\frac{1}{2}} e^{-i\xi(x)/2} \int_{\xi(x)}^{\infty} d\xi_{2n-1} e^{i\xi_{2n-1}}\delta(\xi_{2n-1}) \times \int_{\xi_{2n-1}}^{\infty} \cdots \int_{\xi_2}^{\infty} d\xi_1 e^{i\xi_1}\delta(\xi_1).$$

If $\delta(\xi) \ll 1$ and is arbitrarily gentle, these integrals are asymptotically evaluated by integrating by parts. Now the WKB expansion is an asymptotic power series in ϵ when the derivatives satisfy $d^r\delta/d\xi^r \sim \epsilon^{r+1}$. Hence the nth-order WKB term is obtained by collecting the terms obtained by performing n partial integrations of the above integrals. Notice that all terms calculated in this way will be proportional to $\exp[i\xi(x)/2]$ and hence appear to be forward propagating waves, even though ϕ_2, the backward propagating wave, contributes to the result. An exponentially small term proportional to $\exp[-i\xi(x)/2]$ is obtained from ϕ_2 only if the integral is evaluated with greater precision. For large negative ξ, where $\delta \approx 0$, the partial integration of ϕ_2 yields arbitrarily small terms, and only then does the exponentially small term dominate.

ACKNOWLEDGMENTS

The authors are grateful to the IAEA for hospitality at the International Centre for Theoretical Physics, Trieste, Italy.

One of the authors (D. L. B.) had the work performed under a joint General Atomic–Texas Atomic Research Foundation program on controlled thermonuclear reactions, and another (H. L. B.) was supported by a U.S. Atomic Energy Commission contract.

APPENDIX A. DERIVATION OF THE BREMMER SERIES

We begin with Eq. (1) and review a derivation of the Bremmer equation.[1] Suppose $k(x) > k_{\min} > 0$ for all x, and suppose

$$k(x) = k_-, \quad x \le a,$$
$$k(x) = k_+, \quad x > b,$$

with $a < b$, where k_- and k_+ are constant. For p, a positive integer, define $\Delta x = (b-a)/p$ and $x_j = a + j\Delta x$, $0 \le j \le p$. Let $k_i \equiv k(x_i)$, and define the stepwise approximation $k^*(x)$

$$k^*(x) = k_-, \quad x \le a,$$
$$= k_i, \quad x_{i-1} < x \le x_i, \qquad \text{(A1)}$$
$$= k_+, \quad x > b.$$

For a "wave" propagating from left to right [i.e., a solution that looks locally like $\exp(ikx)$], there is a reflection coefficient at the step at $x = x_i$:

$$r_i = (k_i - k_{i+1})/(k_i + k_{i+1}) \qquad \text{(A2)}$$

and a transmission coefficient

$$t_i = 2k_i/(k_i + k_{i+1}) = 1 + r_i. \qquad \text{(A3)}$$

Similarly, for waves propagating from right to left we have r_i' and t_i' obtained by interchanging k_i and $k_i + 1$.

Let $\phi(a) = 1$, so as to satisfy the boundary condition of unit amplitude in the incident wave at $x < a$; then the pure transmitted (unreflected) wave is given by

$$\phi_1^{(0)} = \phi_1(a) e^{ik_1\Delta x} t_1 e^{ik_j\Delta x} t_2 \cdots e^{ik_j(x-x_j)} \qquad \text{(A4)}$$

for $x_j < x \le x_{j+1}$. Also, for the first-order reflected wave,

$$\phi_2^{(1)}(x) = \sum_i r_i \phi_1^{(0)}(x_i) P_2(x_i, x) \qquad \text{(A5)}$$

where the summation is carried out over $x_i > x$. The propagators P_1 and P_2 are given by

$$P_1(x_i, x) = \prod_{l=i}^{j} e^{ik_l\Delta x} t_l e^{ik_j(x-x_j)}, \qquad \text{(A6)}$$

where $x_i < x_j < x \le x_{j+1}$, and

$$P_2(x_i, x) = \prod_{l=j}^{i} e^{-ik_l\Delta x} t_l' e^{-ik_{j+1}(x_j-x)}, \qquad \text{(A7)}$$

where $x_{j-1} < x \le x_j < x_i$ and

$$t_l' = (2k_{i+1})^{-1}(k_i + k_{i+1}).$$

In general,

$$\phi_1^{(2n)}(x) = \sum_{x_i' \le x} r_i' \phi_2^{(2n-1)}(x_i) P_1(x_i, x), \qquad \text{(A8)}$$

$$\phi_2^{(2n+1)}(x) = \sum_{x_i' > x} r_i \phi_1^{(2n)}(x_i) P_2(x_i, x), \qquad \text{(A9)}$$

where

$$r_i' = (k_{i+1} - k_i)/(k_i + k_{i+1}).$$

Now let p, the number of steps, become very large. In this limit $k^*(x) \to k(x)$ and the sums go over into integrals, and in place of (A4) ($k' \equiv dk/dx$) we get

$$\phi_1^{(0)}(x) = \phi_1(a) \exp(ik_1\Delta x)\left[1 - \frac{k_1'\Delta x}{2k_1}\right] \times e^{ik_2\Delta x}\left[1 - \frac{k_2'\Delta x}{k_2}\right] \cdots$$
$$= \phi_1(a) \exp\left[i\sum k_i\Delta x - \sum \frac{k_i'\Delta x}{2k_i}\right]$$
$$\to \phi_1(a)\left[\frac{k_-}{k(x)}\right]^{\frac{1}{2}} \exp\left(i\int_a^x k\,dx'\right) \qquad \text{(A4')}$$

(the lowest-order WKB solution). Similarly,

$$\phi_2^{(1)}(x) = -\int_a^b dx' \frac{k'(x')}{2k} P(x', x)\phi_1^{(0)}(x'), \quad \text{(A5')}$$

$$P(x_1, x_2) = \left[\frac{k(x_1)}{k(x_2)}\right]^{\frac{1}{2}} \exp\left(-i\int_{x_1}^{x_2} k\, dx\right), \quad \text{(A6')}$$

$$\phi_1^{(2p)}(x) = \int_a^x dx' \frac{k'}{2k} P^*(x', x)\phi_2^{(2p-1)}(x'),$$

$$= [k(x)]^{-\frac{1}{2}} \int_a^x dx' \frac{k'(x')}{2k^{\frac{1}{2}}} \phi_2^{(2p-1)} \times \exp\left(i\int_{x'}^x k(y)\, dy\right), \quad \text{(A8')}$$

$$\phi_2^{(2p+1)}(x) = -\int_x^b dx' \frac{k'(x')}{2k} \phi_1^{(2p)}(x')P(x', x),$$

$$= -[k(x)]^{-\frac{1}{2}} \int_x^b dx' \frac{k'(x')}{2k^{\frac{1}{2}}} \phi_1^{(2p)}(x') \times \exp\left(i\int_x^{x'} k(y)\, dy\right). \quad \text{(A9')}$$

If we formally sum these series, we get the Bremmer integral equations

$$\phi_1(x) = \phi_1^{(0)}(x) + \int_a^x dx \frac{k'(x)}{2k} P^*(x', x)\phi_2(x'), \quad \text{(A10)}$$

$$\phi_2(x) = -\int_x^b dx' \frac{k'(x')}{2k} P(x', x)\phi_1(x'). \quad \text{(A11)}$$

One can easily verify by substitution that the solution of (A10) and (A11) satisfies (1).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Solutions to the Neutron Transport Equation for a Critical Slab by Perturbation Theory*

W. L. HENDRY
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico

(Received 18 October 1966)

A perturbation-theoretic method is developed by which solutions to the monoenergetic neutron transport equation for an unreflected critical slab can be obtained in the case of anisotropic scattering and/or nonconstant cross sections. The method rests upon the use of a generalized Green's function to solve the perturbation equations. The Green's function is derived by an eigenfunction expansion technique and as the solution to two coupled singular integral equations, and the results are compared numerically.

1. INTRODUCTION

SOLUTIONS to the monoenergetic neutron transport equation for a critical bare slab reactor have been obtained independently by Mitsis[1] and Zelazny.[2] Both authors used the normal mode method of Case[3] to cast the solution for the neutron distribution function into the form of a singular eigenfunction expansion, and found a Fredholm integral equation for the continuum coefficient. The analysis of both authors was based on the assumption of isotropic scattering in the laboratory system of coordinates and constant cross sections throughout the medium. The extension of Case's method to arbitrary anisotropic scattering has been carried out by Mika.[4]

In principle, any problem that can be solved subject to the assumption of isotropic scattering can also be solved using Mika's approach, which is to expand the scattering kernel in Legendre polynomials. Unfortunately, this technique loses most of its practical utility if more than the first few terms of the expansion are retained. This is due to two reasons: (1) certain identities which are very useful in simplifying the final analytical results in the case of isotropic scattering no longer apply for anisotropic scattering;

* Based on a Ph.D. thesis submitted to the Department of Nuclear Engineering, Massachusetts Institute of Technology. This paper was prepared at the Los Alamos Scientific Laboratory, under the auspices of the U.S. Atomic Energy Commission.

[1] G. J. Mitsis, Nucl. Sci. Eng. **17**, 55 (1963).

[2] R. Zelazny, J. Math. Phys. **2**, 538 (1961).

[3] K. M. Case, Ann. Phys. (N.Y.) **9**, 1 (1960).

[4] J. R. Mika, Nucl. Sci. Eng. **11**, 415 (1961).

(the lowest-order WKB solution). Similarly,

$$\phi_2^{(1)}(x) = -\int_a^b dx' \frac{k'(x')}{2k} P(x', x)\phi_1^{(0)}(x'), \quad \text{(A5')}$$

$$P(x_1, x_2) = \left[\frac{k(x_1)}{k(x_2)}\right]^{\frac{1}{2}} \exp\left(-i\int_{x_1}^{x_2} k\, dx\right), \quad \text{(A6')}$$

$$\phi_1^{(2p)}(x) = \int_a^x dx' \frac{k'}{2k} P^*(x', x)\phi_2^{(2p-1)}(x'),$$

$$= [k(x)]^{-\frac{1}{2}} \int_a^x dx' \frac{k'(x')}{2k^{\frac{1}{2}}} \phi_2^{(2p-1)} \times \exp\left(i\int_{x'}^x k(y)\, dy\right), \quad \text{(A8')}$$

$$\phi_2^{(2p+1)}(x) = -\int_x^b dx' \frac{k'(x')}{2k} \phi_1^{(2p)}(x')P(x', x),$$

$$= -[k(x)]^{-\frac{1}{2}} \int_x^b dx' \frac{k'(x')}{2k^{\frac{1}{2}}} \phi_1^{(2p)}(x') \times \exp\left(i\int_x^{x'} k(y)\, dy\right). \quad \text{(A9')}$$

If we formally sum these series, we get the Bremmer integral equations

$$\phi_1(x) = \phi_1^{(0)}(x) + \int_a^x dx \frac{k'(x)}{2k} P^*(x', x)\phi_2(x'), \quad \text{(A10)}$$

$$\phi_2(x) = -\int_x^b dx' \frac{k'(x')}{2k} P(x', x)\phi_1(x'). \quad \text{(A11)}$$

One can easily verify by substitution that the solution of (A10) and (A11) satisfies (1).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Solutions to the Neutron Transport Equation for a Critical Slab by Perturbation Theory*

W. L. HENDRY
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico

(Received 18 October 1966)

A perturbation-theoretic method is developed by which solutions to the monoenergetic neutron transport equation for an unreflected critical slab can be obtained in the case of anisotropic scattering and/or nonconstant cross sections. The method rests upon the use of a generalized Green's function to solve the perturbation equations. The Green's function is derived by an eigenfunction expansion technique and as the solution to two coupled singular integral equations, and the results are compared numerically.

1. INTRODUCTION

SOLUTIONS to the monoenergetic neutron transport equation for a critical bare slab reactor have been obtained independently by Mitsis[1] and Zelazny.[2] Both authors used the normal mode method of Case[3] to cast the solution for the neutron distribution function into the form of a singular eigenfunction expansion, and found a Fredholm integral equation for the continuum coefficient. The analysis of both authors was based on the assumption of isotropic scattering in the laboratory system of coordinates and constant cross sections throughout the medium. The extension of Case's method to arbitrary anisotropic scattering has been carried out by Mika.[4]

In principle, any problem that can be solved subject to the assumption of isotropic scattering can also be solved using Mika's approach, which is to expand the scattering kernel in Legendre polynomials. Unfortunately, this technique loses most of its practical utility if more than the first few terms of the expansion are retained. This is due to two reasons: (1) certain identities which are very useful in simplifying the final analytical results in the case of isotropic scattering no longer apply for anisotropic scattering;

* Based on a Ph.D. thesis submitted to the Department of Nuclear Engineering, Massachusetts Institute of Technology. This paper was prepared at the Los Alamos Scientific Laboratory, under the auspices of the U.S. Atomic Energy Commission.

[1] G. J. Mitsis, Nucl. Sci. Eng. **17**, 55 (1963).

[2] R. Zelazny, J. Math. Phys. **2**, 538 (1961).

[3] K. M. Case, Ann. Phys. (N.Y.) **9**, 1 (1960).

[4] J. R. Mika, Nucl. Sci. Eng. **11**, 415 (1961).

and (2) the roots (diffusion lengths) of a complicated transcendental expression must be found numerically in the case of anisotropic scattering.

In this paper we develop a perturbation-theoretic method with which solutions to the neutron transport equation can be obtained for systems which display small departures from homogeneity and/or small amounts of anisotropic scattering.

The plan of the succeeding sections is as follows. We begin by showing how the problems of anisotropic scattering and nonconstant cross sections can be formulated in such a way that perturbation theory is applicable. Perturbation theory is then applied to recast the problems into a form where the perturbation is introduced as inhomogeneous terms in an infinite set of transport equations. It is then shown that these equations may be solved by means of a generalized Green's function. The Green's function is derived in two ways: by eigenfunction expansion, and by converting to singular integral equations. The two forms of the Green's function that result from these procedures are compared numerically.

2. STATEMENT OF THE PROBLEM

For an unreflected source-free slab the monoenergetic steady-state transport equation has the form

$$\mu \frac{\partial \psi(x,\mu)}{\partial x} + \psi(x,\mu) = \frac{c(x)}{2}\int_{-1}^{+1} f(\mu' \to \mu)\psi(x,\mu')\,d\mu'. \tag{2.1}$$

Here ψ is the neutron distribution function, μ the direction cosine of the neutron velocity vector, x the optical distance from the center of the slab, $c(x)$ the mean number of secondary neutrons emitted from a collision event, and $f(u' \to \mu)$ is the scattering kernel. If the slab is taken to be of width $2b$ centered about the origin, the appropriate boundary conditions are

$$\psi(b,\mu) = 0, \quad \mu < 0, \tag{2.2}$$

and

$$\psi(x,\mu) = \psi(-x,-\mu). \tag{2.3}$$

In the case in which the mean number of secondaries $c(x)$ is not a constant, but the scattering is isotropic, it is easy to cast the problem into a form in which the inhomogeneity can be handled by perturbation theory. We write

$$c(x) = c + \epsilon g(x), \tag{2.4}$$

where c is a constant and ϵ a small parameter. The transport equation takes the form

$$\mu \frac{\partial \psi(x,\mu)}{\partial x} + \psi(x,\mu) = \frac{c}{2}\int_{-1}^{+1}\psi(x,\mu')\,du' + \epsilon\frac{g(x)}{2}\int_{-1}^{+1}\psi(x,\mu')\,d\mu'. \tag{2.5}$$

When the mean number of secondaries is constant, but the scattering is anisotropic, we can obtain a formulation suitable for the application of perturbation theory by assuming that we have a mixture of scatterers, some of which scatter isotropically and some of which scatter anisotropically. The transport equation becomes

$$\mu \frac{\partial \psi(x,\mu)}{\partial x} + \psi(x,\mu) = \frac{\Sigma_s^{\text{iso}} + \nu\Sigma_f}{2\Sigma_t}\int_{-1}^{+1}\psi(x,\mu')\,d\mu' + \frac{\Sigma_s^{\text{aniso}}}{2\Sigma_t}\int_{-1}^{+1} f(\mu' \to \mu)\psi(x,\mu')\,d\mu', \tag{2.6}$$

where Σ_s, Σ_f, and Σ_t are the macroscopic scattering, fission, and total cross sections, respectively, and ν is the mean number of neutrons released from a fission event. If we define

$$c = (\Sigma_s^{\text{iso}} + \Sigma_s^{\text{aniso}} + \nu\Sigma_f)/\Sigma_t, \tag{2.7}$$

$$\epsilon = \Sigma_s^{\text{aniso}}/(\Sigma_s^{\text{iso}} + \Sigma_s^{\text{aniso}} + \nu\Sigma_f), \tag{2.8}$$

$$F(\mu';\mu) = f(\mu' \to \mu) - 1, \tag{2.9}$$

then Eq. (2.6) takes the form

$$\mu \frac{\partial \psi(x,\mu)}{\partial x} + \psi(x,\mu) = \frac{c}{2}\int_{-1}^{+1}[1 + \epsilon F(\mu';\mu)]\psi(x,\mu')\,d\mu'. \tag{2.10}$$

The parameter ϵ lies in the range (0, 1) and is small when the fraction of secondary neutrons due to anisotropic scattering is small. It is therefore suitable as the basis of a perturbation expansion.

Of course, both the inhomogeneous medium problem and the anisotropic scattering problem can be solved simultaneously; it is only necessary to use the same expansion parameter for each perturbation. This is possible because the parameter ϵ is arbitrary in the inhomogeneous medium case.

3. APPLICATION OF THE PERTURBATION METHOD

Both Eqs. (2.5) and (2.10) may be reduced to a more tractable set of equations by making the substitutions

$$\psi(x,\mu) = \sum_{m=0}^{\infty}\psi_m(x,\mu)\epsilon^m, \tag{3.1}$$

$$c = \sum_{l=0}^{\infty} c_l\epsilon^l, \tag{3.2}$$

and equating equal powers of ϵ. It follows from the arbitrariness of ϵ that the ψ_m must satisfy the same boundary conditions as ψ. When this procedure is

applied to Eq. (2.5) the result is

$$B_0\psi_0(x,\mu) = 0, \tag{3.3}$$

$$B_0\psi_m(x,\mu) = R_m(x), \tag{3.4}$$

where we have introduced the operator B_0 defined such that

$$B_0\psi = \mu\frac{\partial\psi}{\partial x} + \psi - \frac{c_0}{2}\int_{-1}^{+1}\psi(x,\mu')\,d\mu' \tag{3.5}$$

and

$$R_m(x) = \tfrac{1}{2}g(x)\int_{-1}^{+1}\psi_{m-1}(x,\mu')\,d\mu' + \tfrac{1}{2}\sum_{l=0}^{m-1}c_{m-l}\int_{-1}^{+1}\psi_l(x,\mu')\,d\mu'. \tag{3.6}$$

Similarly, when the procedure is applied to Eq. (2.10), we obtain Eq. (3.3) plus

$$B_0\psi_m(x,\mu) = Q_m(x,\mu), \tag{3.7}$$

where

$$Q_m(x,\mu) = \tfrac{1}{2}\sum_{l=1}^{m}c_l\int_{-1}^{+1}\psi_{m-l}(x,\mu')\,d\mu' + \tfrac{1}{2}\sum_{l=0}^{m-1}c_l\int_{-1}^{+1}\psi_{m-l-1}(x,\mu')F(\mu,\mu')\,d\mu'. \tag{3.8}$$

Both cases are seen to result in the same type of problem. Equation (3.3) is the unperturbed problem, solved by Mitsis and Zelazny.

4. SOLUTIONS OF THE UNPERTURBED PROBLEM

For easy reference, we present solutions to Eq. (3.3), subject to the boundary conditions, Eqs. (2.2) and (2.3), as found by Mitsis,[5] with several minor extensions. To avoid confusion with the expansion coefficients in the perturbation series we denote the nth eigenfunction and eigenvalue of Eq. (3.3) as $\chi_n(x,\mu)$ and $2\gamma_n$ ($n = 0, 1, \cdots$). Thus

$$B_0\chi_n(x,\mu) = \mu\frac{\partial\chi_n(x,\mu)}{\partial x} + \chi_n(x,\mu) - \gamma_n\int_{-1}^{+1}\chi_n(x,\mu')\,d\mu' = 0. \tag{4.1}$$

Mitsis was interested only in the everywhere-positive solution $\chi_0(x,\mu)$ and the corresponding mean number of secondaries required for criticality, γ_0 (given the half-thickness b). However, all eigenfunctions obeying the symmetry condition, Eq. (2.3), were implicit in his results.

In addition to the symmetric eigenfunctions, a denumerably infinite set of antisymmetric eigenfunctions exist which satisfy, instead of Eq. (2.3), the condition

$$\psi(x,\mu) = -\psi(-x,-\mu). \tag{4.2}$$

Mitsis found these solutions also, in the context of the critical sphere problem, where they play the role of a "density transform."

Using the notation of Case,[3] the *elementary* solutions of Eq. (4.1) are

$$\psi_\pm(x,\mu) = [\gamma_n\nu_0/(\nu_0 \mp \mu)]e^{\mp x/\nu_0}, \tag{4.3}$$

$$\psi_\nu(x,\mu) = \{P[\gamma_n\nu/(\nu-\mu)] + \lambda(\nu)\delta(\nu-\mu)\}e^{-x/\nu}, \tag{4.4}$$

where the symbol P indicates that the principal value is to be taken upon integration, and $\lambda(\nu) = P\Lambda(\nu)$, where

$$\Lambda(z) = 1 - \gamma_n z\int_{-1}^{+1}\frac{d\mu}{(z-\mu)}. \tag{4.5}$$

The numbers $\pm\nu_0$ are the (pure imaginary for $2\gamma_n > 1$) roots of the transcendental equation $\Lambda(z) = 0$. In addition, the following notation is needed.

$$X(z) = \frac{1}{1-z}\exp\left\{\frac{1}{2\pi i}\int_0^1 \ln\left[\frac{\Lambda^+(\mu)}{\Lambda^-(\mu)}\right]\frac{d\mu}{\mu - z}\right\}, \tag{4.6}$$

$$g(2\gamma_n,\mu) = 1/\Lambda^+(\mu)\Lambda^-(\mu), \tag{4.7}$$

where the functions $\Lambda^\pm(\mu)$ are the boundary values of the function $\Lambda(z)$ as $z \to \mu\epsilon(-1,+1)$ from the upper and lower half-planes, respectively.

The eigenfunctions χ_{2n} are given by

$$\chi_{2n}(x,\mu) = a_{2n}[\psi_+(x,\mu) + \psi_-(x,\mu)] + \int_{-1}^{+1}A_{2n}(\nu)\psi_\nu(x,\mu)\,d\nu, \tag{4.8}$$

where a_{2n} is an arbitrary constant and $A_{2n}(\nu)$ is the solution to the following Fredholm equation.

$$A_{2n}(\mu)e^{b/\mu} = -(\nu_0^2 - \mu^2)(1-2\gamma_{2n})X(-\mu)g(2\gamma_{2n},\mu) \times\left\{a_{2n}[\psi_+(-b,\mu)X(\nu_0) + \psi_-(-b,\mu)X(-\nu_0)] + \int_0^1\frac{\gamma_{2n}\nu X(-\nu)A_{2n}(\nu)e^{-b/\nu}}{\nu+\mu}\,d\nu\right\}. \tag{4.9}$$

The constants γ_{2n} must satisfy the auxiliary condition

$$a_{2n}e^{b/\nu_0}\nu_0[X(\nu_0) - e^{-2b/\nu_0}X(-\nu_0)] = \int_0^1\nu X(-\nu)A_{2n}(\nu)e^{-b/\nu}\,d\nu, \tag{4.10}$$

given the half-thickness b.

The antisymmetric eigenfunctions χ_{2n+1} are given by

$$\chi_{2n+1}(x,\mu) = a_{2n+1}[\psi_+(x,\mu) - \psi_-(x,\mu)] + \int_{-1}^{+1}A_{2n+1}(\nu)\psi_\nu(x,\mu)\,d\nu, \tag{4.11}$$

[5] G. J. Mitsis, Argonne National Laboratory Report No. ANL-6787 (1963).

where a_{2n+1} is an arbitrary constant and $A_{2n+1}(\nu)$ is the solution to the following Fredholm equation.

$$A_{2n+1}(\mu)e^{b/\mu} = -(\nu_0^2 - \mu^2)(1 - 2\gamma_{2n+1}) \times X(-\mu)g(2\gamma_{2n+1}, \mu) \times \left\{a_{2n+1}[\psi_+(-b, \mu)X(\nu_0) - \psi_-(-b, \mu)X(-\nu_0)] - \int_0^1 \frac{\gamma_{2n+1}\nu X(-\nu)A_{2n+1}(\nu)e^{-b/\nu}}{\nu + \mu}\, d\nu\right\}. \quad (4.12)$$

The constants γ_{2n+1} must satisfy the auxiliary condition

$$a_{2n+1}e^{b/\nu_0}\nu_0[X(\nu_0) + e^{-2b/\nu_0}X(-\nu_0)] = -\int_0^1 \nu X(-\nu)A_{2n+1}(\nu)e^{-b/\nu}\, d\nu,$$

given the half-thickness b.

5. SOLUTIONS TO THE PERTURBATION EQUATIONS: THE GENERALIZED GREEN'S FUNCTION

The above reduction of both Eqs. (2.5) and (2.10) by the perturbation method led to a system of equations having the general form

$$B_0\psi_0(x, \mu) = 0, \quad (5.1)$$

$$B_0\psi_m(x, \mu) = S_m(x, \mu), \quad m > 0. \quad (5.2)$$

In both cases, S_m had the following two properties: (1) S_m is a functional containing $\psi_{m-1}, \psi_{m-2}, \cdots, \psi_0$, but not ψ_m; and (2) S_m is a function of $c_m, c_{m-1}, \cdots, c_0$. Thus, if we solve for $\psi_0, \psi_1, \cdots, \psi_{m-1}$ successively, then S_m will contain one degree of freedom: c_m.

To derive an expression for c_m, we introduce the operator $B_0^\dagger$, adjoint to B_0, and the function $\psi_0^\dagger$, adjoint to ψ_0, satisfying

$$B_0^\dagger\psi_0^\dagger = -\mu\frac{\partial\psi_0^\dagger(x, \mu)}{\partial x} + \psi_0^\dagger(x, \mu) - \frac{c_0}{2}\int_{-1}^{+1}\psi_0^\dagger(x, \mu')\, d\mu' = 0. \quad (5.3)$$

The function $\psi_0^\dagger$ satisfies the adjoint boundary condition

$$\psi_0^\dagger(b, \mu) = 0, \quad \mu > 0, \quad (5.4)$$

and the symmetry condition

$$\psi_0^\dagger(x, \mu) = \psi_0^\dagger(-x, -\mu). \quad (5.5)$$

As usual,

$$\langle\psi_0, B_0^\dagger\psi_0^\dagger\rangle = \langle\psi_0^\dagger, B_0\psi_0\rangle, \quad (5.6)$$

where we use the Dirac notation for the scalar product:

$$\langle f, g\rangle = \int_{-b}^{+b} dx \int_{-1}^{+1} d\mu(fg). \quad (5.7)$$

We are interested in the everywhere positive solutions to Eqs. (5.1) and (5.3); thus,

$$\psi_0(x, \mu) = \chi_0(x, \mu), \quad \tfrac{1}{2}c_0 = \gamma_0, \quad (5.8)$$

and, as is easily seen by comparing the equations and boundary conditions for ψ_0 and $\psi_0^\dagger$,

$$\psi_0^\dagger(x, \mu) = \psi_0(x, -\mu) \quad (5.9)$$

and

$$\chi_n^\dagger(x, \mu) = \chi_n(x, -\mu). \quad (5.10)$$

We obtain a necessary condition for the solubility of Eq. (5.2) by noting that since $\langle\psi_0^\dagger, B_0\psi_m\rangle = \langle\psi_m, B_0^\dagger\psi_0^\dagger\rangle = 0$, we have

$$\langle\psi_0^\dagger, S_m\rangle = 0. \quad (5.11)$$

This expression fixes the constant c_m, and thus specifies S_m completely. With S_m known, we turn to the solution of Eq. (5.2) for ψ_m. The fact that this equation is inhomogeneous suggests the use of a Green's function. The Green's function in the ordinary sense does not exist for this problem, however, due to the existence of a nontrivial solution to the corresponding homogeneous equation. We therefore employ the notion of a generalized Green's function.[6]

We define the generalized Green's function for our problem as the solution to the equation

$$B_0^\dagger K^\dagger(x, \mu \mid x_0) = \delta(x - x_0) - \rho_0(x)\rho_0(x_0), \quad (5.12)$$

subject to the (adjoint) boundary conditions

$$K^\dagger(-b, \mu \mid x_0) = K^\dagger(b, -\mu \mid x_0) = 0, \quad \mu < 0. \quad (5.13)$$

We have defined

$$\rho_0(x) = \int_{-1}^{+1}\psi_0(x, \mu)\, d\mu, \quad (5.14)$$

and chosen the normalization of $\rho_0(x)$ such that

$$\int_{-b}^{b}\rho_0^2(x)\, dx = 1. \quad (5.15)$$

The definition of the generalized Green's function differs from that of the ordinary Green's function in the term $\rho_0(x)\rho_0(x_0)$. This term renders the right-hand side of Eq. (5.12) orthogonal to $\psi_0(x, \mu)$, which is required for precisely the same reason that $S_m(x, \mu)$ must be orthogonal to $\psi_0^\dagger(x, \mu)$.

To show that $K^\dagger$ suffices to solve the perturbation equations, take the scalar product of $K^\dagger$ with Eq. (5.2), with the result

$$\rho_m(x) = \langle K^\dagger(x', \mu' \mid x), S_m(x', \mu')\rangle + \langle\psi_m(x', \mu'), \rho_0(x')\rangle\rho_0(x), \quad (5.16)$$

[6] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. I, p. 356.

where

$$\rho_m(x) = \int_{-1}^{+1} \psi_m(x, \mu)\, d\mu. \tag{5.17}$$

Equation (5.16) reflects the fact that $\psi_m\,(\rho_m)$ is arbitrary to within an additive multiple of $\psi_0\,(\rho_0)$. Given the normalization of the solution to the unperturbed equation, $\psi_0(x, \mu)$, and of the solution to the perturbed equation, $\psi(x, \mu)$, all the ψ_m are specified uniquely. If, for example, we require that

$$\int_{-b}^{+b} \rho_0(x)\rho(x)\, dx = 1, \tag{5.18}$$

then it follows immediately from Eqs. (5.15) and (3.1) that

$$\langle \psi_m, \rho_0 \rangle = \int_{-b}^{b} \rho_m(x)\rho_0(x)\, dx = 0, \quad m > 0. \tag{5.19}$$

Since $K^\dagger$ is arbitrary to within an additive multiple of $\psi_0^\dagger$, we can demand that

$$\int_{-b}^{b} \rho_0(x)\langle K^\dagger(x', \mu' \mid x), S_m(x', \mu')\rangle\, dx = 0, \tag{5.20}$$

from which it follows that

$$\rho_m(x) = \langle K^\dagger(x', \mu' \mid x), S_m(x', \mu')\rangle. \tag{5.21}$$

Given $\rho_m(x)$, Eq. (5.2) becomes an ordinary differential equation in x for ψ_m. The solution is

$$\psi_m(x, \mu) = \frac{1}{\mu} \int_{\mp b}^{x} e^{(y-x)/\mu} \left[\frac{c_0}{2} \rho_m(y) + S_m(y, \mu) \right] dy, \quad \mu \gtrless 0. \tag{5.22}$$

We have therefore shown that the generalized Green's function $K^\dagger$, if it can be found, suffices to solve the perturbation equations, Eq. (5.2).

In discussing the construction of the generalized Green's function in the sequel, we demand that the orthogonality condition $\langle \psi_0, K^\dagger \rangle = 0$ hold. Although this will not in general satisfy Eq. (5.20), we can, having calculated $\rho_m(x)$ from Eq. (5.21), always add on that multiple of $\rho_0(x)$ which will satisfy Eq. (5.19). Also, from a practical point of view, we are generally more interested in the change of shape in the neutron density that will result from a given perturbation than in $O(\epsilon)$ changes in normalization.

6. CONSTRUCTION OF THE GENERALIZED GREEN'S FUNCTION: METHOD OF EIGENFUNCTION EXPANSIONS

It is known that the functions $\chi_n^\dagger(x, \mu)$ do not form a complete set in the variables x and μ.[7] We can, nevertheless, express the generalized Green's function as an infinite sum of these functions by first converting the equation for $K^\dagger$ to a Fredholm integral equation, and then applying well known theorems. We define

$$\phi_n(x) = \int_{-1}^{+1} \chi_n(x, \mu)\, d\mu, \tag{6.1}$$

$$\phi_n^\dagger(x) = \int_{-1}^{+1} \chi_n^\dagger(x, \mu)\, d\mu = \phi_n(x), \tag{6.2}$$

$$G^\dagger(x \mid x_0) = \int_{-1}^{+1} K^\dagger(x, \mu \mid x_0)\, d\mu. \tag{6.3}$$

Of course, $\phi_0(x) = \rho_0(x)$. Substituting from Eq. (6.3) into Eq. (5.12) and taking $G^\dagger$ as known, we obtain an ordinary differential equation in x, with the solution

$$K^\dagger(x, \mu \mid x_0) = \frac{c_0}{2\mu} \int_x^{\pm b} \exp\left(\frac{x - x'}{\mu}\right) G^\dagger(x' \mid x_0)\, dx' \pm \frac{u(\pm x_0 \mp x)}{\mu} \exp\left(\frac{x - x_0}{\mu}\right) - \frac{2}{c_0} \rho_0(x_0)\psi_0^\dagger(x, \mu), \quad \mu \gtrless 0, \tag{6.4}$$

where $u(x) = 0$, $x < 0$, and $u(x) = 1$, $x > 0$. Now integrating over μ, we obtain the desired Fredholm equation

$$G^\dagger(x \mid x_0) = \frac{c_0}{2} \int_{-b}^{b} E_1(|x - y|) G^\dagger(y \mid x_0)\, dy + E_1(|x_0 - x|) - \frac{2}{c_0} \rho_0(x_0)\rho_0(x), \tag{6.5}$$

where E_1 is the first exponential integral,

$$E_n(z) = \int_0^1 \mu^{n-2} e^{-z/\mu}\, d\mu. \tag{6.6}$$

In the same way, the Boltzmann equation for $\chi_n^\dagger$ may be reduced to the following (well known) Fredholm equation for ϕ_n:

$$\phi_n(x) = \gamma_n \int_{-b}^{b} E_1(|x - y|)\phi_n(y)\, dy. \tag{6.7}$$

Noting that $E_1(|x - y|)$ is a symmetric, nondegenerate, quadratically integrable kernel; it follows immediately[8] that the γ_n form a denumerably infinite set of real eigenvalues, with a corresponding set of eigenfunctions ϕ_n.

It is not known that the ϕ_n form a complete set on the interval $(-b, +b)$. However, by virtue of the properties listed above, the Hilbert–Schmidt theorem applies [Ref. 8, p. 110] and the resolvent kernel may

[7] B. Davison, *Neutron Transport Theory* (Oxford University Press, London, 1958), p. 436.

[8] F. G. Tricomi, *Integral Equations* (Interscience Publishers, Inc., New York, 1957), p. 105.

be written [Ref. 8, p. 115]

$$H(x, y; \gamma_0) = -E_1(|x - y|) + \gamma_0 \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(x_0)}{\gamma_n(\gamma_0 - \gamma_n)}. \tag{6.8}$$

The solution for $G^\dagger$ is

$$G^\dagger(x \mid x_0) = E_1(|x - x_0|) - \frac{2}{c_0}\rho_0(x)\rho_0(x_0) + \gamma_0 \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(x_0)}{\gamma_n(\gamma_n - \gamma_0)}, \tag{6.9}$$

where we have made use of the equality

$$\int_{-b}^{b} E_1(|x_0 - y|)E_1(|x - y|)\,dy = \sum_{n=0}^{\infty} \frac{\phi_n(x)\phi_n(x_0)}{\gamma_n^2}. \tag{6.10}$$

If we assume the validity of the bilinear formula (which holds if the ϕ_n form a complete set)

$$E_1(|x - x_0|) = \sum_{n=0}^{\infty} \frac{\phi_n(x)\phi_n(x_0)}{\gamma_n}, \tag{6.11}$$

then Eq. (6.9) takes the elegant form

$$G^\dagger(x \mid x_0) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(x_0)}{\gamma_n - \gamma_0}. \tag{6.12}$$

The numerical results presented below indicate that this is in fact a valid representation.

Finally, we may substitute from Eqs. (6.9) and (6.12) into Eq. (6.4) to obtain

$$K^\dagger(x, \mu \mid x_0) = \mu^{-1}[u(\mu)u(x_0 - x) - u(-\mu)u(x - x_0)] \times \exp\left(\frac{x - x_0}{\mu}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0^\dagger(x, \mu) + \gamma_0 \sum_{n=1}^{\infty} \frac{\chi_n^\dagger(x, \mu)\phi_n(x_0)}{\gamma_n(\gamma_n - \gamma_0)}, \tag{6.13}$$

and

$$K^\dagger(x, \mu \mid x_0) = \sum_{n=1}^{\infty} \frac{\chi_n^\dagger(x, \mu)\phi_n(x_0)}{\gamma_n - \gamma_0}. \tag{6.14}$$

Equations (6.12) and (6.14) are the simplest forms for numerical computation, but Eqs. (6.9) and (6.13) would be expected to converge more rapidly, due to the explicit representation of the singularity of $K^\dagger$ and because γ_n increases with n.

The idea of solving perturbation equations by eigenfunction expansions is, of course, well known. The advantage of the generalized Green's function formulation is that it allows one to do the analysis once and for all, and obtain results by means of the simple formula, Eq. (5.21). The objection to this analysis is that now, in addition to truncating the perturbation expansion in ϵ, we must also truncate the eigenfunction expansion for the generalized Green's function. It is desirable to have a closed form expression for $K^\dagger$, even if it is too complex to be used easily in applications, in order to determine how much error is introduced in the truncation of the eigenfunction expansion in selected cases, and to estimate the number of terms of Eqs. (6.13) or (6.14) that should be retained. This expression is derived in the next section.

7. THE METHOD OF SINGULAR INTEGRAL EQUATIONS

In this section we obtain an expression for the generalized Green's function by solving two coupled singular integral equations. Briefly, the method used is to convert the integro-differential equation into the coupled singular integral equations, and then convert these into a pair of coupled Fredholm equations. The latter may be solved by iteration. The generalized Green's function is given by a closed form expression involving two functions related to its boundary values. The solution to the Fredholm equations yields these functions. The general methods used here were developed by Leonard and Mullikin[9] and by Mitsis,[5] whose work related to the unperturbed critical problem.

Instead of solving Eq. (5.12), we solve for the function K satisfying the equation

$$B_0K(x, \mu \mid x_0) = \delta(x - x_0) - \rho_0(x_0)\rho_0(x), \tag{7.1}$$

subject to

$$K(-b, \mu \mid x_0) = K(b, -\mu \mid x_0) = 0, \quad \mu > 0. \tag{7.2}$$

We can always recover $K^\dagger$ by using the relation $K^\dagger(x, \mu \mid x_0) = K(x, -\mu \mid x_0)$. Defining

$$G(x \mid x_0) = \int_{-1}^{+1} K(x, \mu \mid x_0)\,d\mu, \tag{7.3}$$

we see that $G(x \mid x_0) = G^\dagger(x \mid x_0)$.

We obtain dual singular integral equations from which the function K can be determined by starting from a pair of equations similar to the pair, Eq. (6.4). Using Eq. (7.3) in Eq. (7.1), we obtain

$$K(x, \mu \mid x_0) = \frac{c_0}{2\mu}\int_{\mp b}^{x} G(y \mid x_0) \exp\left(\frac{y - x}{\mu}\right) dy \pm \frac{u(\mp x_0 \pm x)}{\mu} \exp\left(\frac{x_0 - x}{\mu}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x, \mu), \quad \mu \gtrless 0. \tag{7.4}$$

Consider the analytic continuation of Eq. (7.4) onto

[9] A. Leonard and T. W. Mullikin, Rand Corporation Report No. RM-3256-PR (1962).

the complex μ plane. Since the integrand in the integral term of Eq. (7.4) is an analytic function of μ in every closed contour not enclosing the origin, and is a continuous function of the integration variable y, it follows that the integral itself is an analytic function of μ within every closed contour not enclosing the origin.[10] Referring to Eq. (5.22) (putting $m = 0$), we see that ψ_0 is an analytic function of $\mu \neq 0$ for the same reason. The second term on the right-hand side of Eq. (7.4) is also an analytic function of $\mu \neq 0$. It follows that Eq. (7.4) with $\mu > 0$ generates a function $K_1(x, z \mid x_0)$ analytic for all $z \neq 0$, and with $\mu < 0$ generates a function $K_2(x, z \mid x_0)$ also analytic for $z \neq 0$. In no case shall we evaluate K_1 on the negative real axis of the z plane or K_2 on the positive real axis. So expressions involving integrals of K over the range $\mu\epsilon(-1, +1)$ are unambiguous, and we need not append subscripts for real values of z. With these conventions, we write

$$K_k(x, z \mid x_0) = \frac{c_0}{2z}\int_{\mp b}^{\infty} G(y \mid x_0) \exp\left(\frac{y-x}{z}\right) dy \pm \frac{u(\pm x \mp x_0)}{z} \exp\left(\frac{x_0 - x}{z}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x, z), \quad k = \begin{matrix}1\\2\end{matrix}. \tag{7.5}$$

Equation (7.5) can be converted into two singular integral equations by the following procedure. Substitute for G on the right-hand side from the expression

$$G(x \mid x_0) = \int_0^1 [K(x, \nu \mid x_0) + K(x, -\nu \mid x_0)]\, d\nu, \tag{7.6}$$

and invert the y and ν integrations. The y integration may be performed explicitly by substituting for K from Eq. (7.4). After the y integration has been performed, it is possible to identify certain groupings of terms in such a way that G is again eliminated in favor of K. The details are omitted here since the analysis is very similar to that performed by Mitsis[5] in finding a singular integral equation satisfied by ψ_0. The result of the above manipulations is the equation

$$\Lambda(z)K_k(x, z \mid x_0) - \int_{-1}^{+1} \frac{(\frac{1}{2}c_0)\nu K(x, \nu \mid x_0)}{\nu - z}\, d\nu = -\exp\left(\mp \frac{b \pm x}{z}\right)\int_0^1 \frac{(\frac{1}{2}c_0)\nu K(\mp b, \mp|\nu\ x_0)}{\nu \pm z}\, d\nu \pm \frac{u(\pm x \mp x_0)}{z} \exp\left(\frac{x_0 - x}{z}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x, z), \quad k = \begin{matrix}1\\2\end{matrix}. \tag{7.7}$$

In the above we have made use of (the analytic continuation of) an equation for ψ_0 found by Mitsis:

$$\lambda(\mu)\psi_0(x, \mu) - P\int_{-0}^{+1} \frac{(\frac{1}{2}c_0)\nu\psi_0(x, \nu)}{\nu - \mu}\, d\nu = -\exp\left(\mp \frac{b \pm x}{\mu}\right)\int_0^1 \frac{(\frac{1}{2}c_0)\nu\psi(b, \nu)}{\nu \pm \mu}\, d\nu, \quad \mu \gtrless 0. \tag{7.8}$$

Using the Plemelj formulas,[11] we obtain the equation corresponding to Eq. (7.7) for $z = \mu$ in the real interval $(-1, +1)$. The result is

$$\lambda(\mu)K(x, \mu \mid x_0) - P\int_{-1}^{+1} \frac{(\frac{1}{2}c_0)\nu K(x, \nu \mid x_0)}{\nu - \mu}\, d\nu = f(x, \mu \mid x_0), \tag{7.9}$$

where

$$f(x, \mu \mid x_0) = -\exp\left(\mp \frac{b \pm x}{\mu}\right) \times \int_0^1 \frac{(\frac{1}{2}c_0)\nu K(\mp b, \mp\nu \mid x_0)}{\nu \pm \mu}\, d\nu \pm \frac{u(\pm x \mp x_0)}{\mu} \times \exp\left(\frac{x_0 - x}{\mu}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x, \mu), \quad \mu \gtrless 0. \tag{7.10}$$

Equation (7.9) amounts to dual singular integral equations on the partial ranges $(-1, 0)$, $(0, 1)$. We may, however, obtain an explicit representation of K in terms of its boundary values by treating Eq. (7.9) as a single equation.

To solve Eq. (7.9) we introduce the complex transformation

$$N(x, z) = \frac{1}{2\pi i}\int_{-1}^{+1} \frac{(\frac{1}{2}c_0)\nu K(x, \nu \mid x_0)}{\nu - z}\, d\nu. \tag{7.11}$$

Since $K(x, z \mid x_0)$ is analytic for $z \neq 0$ and continuous at $z = 0$, it follows that we may proceed in the usual fashion to conclude that

(a) $N(x, z)$ is an analytic function of z in the plane cut from -1 to $+1$,

(b) $N(x, z) \sim 1/z$ as $z \to \infty$,

(c) $N < \text{const}/|z \mp 1|^{\gamma}$, $\gamma < 1$, as $z \to \mp 1$.

Using the Plemelj formulas, for $\mu \in (-1, +1)$,

$$N^+(x, \mu) + N^-(x, \mu) = \frac{1}{\pi i} P\int_{-1}^{+1} \frac{(\frac{1}{2}c_0)\nu K(x, \nu \mid x_0)}{\nu - \mu}\, d\nu, \tag{7.12}$$

$$N^+(x, \mu) - N^-(x, \mu) = \tfrac{1}{2}c_0\mu K(x, \mu \mid x_0). \tag{7.13}$$

Substituting the above into Eq. (7.9), we obtain the

[10] E. T. Copson, *Theory of Functions of a Complex Variable* (Oxford University Press, London, 1935), p. 108.

[11] N. I. Muskhelishvili, *Singular Integral Equations* (P. Noordhoff, Ltd., Groningen, The Netherlands, 1953), p. 27.

following Hilbert boundary value problem

$$N^+(x,\mu) - \frac{\Lambda^+(\mu)}{\Lambda^-(\mu)} N^-(x,\mu) = \frac{c_0\mu}{2} \frac{f(x,\mu \mid x_0)}{\Lambda^-(\mu)}, \quad (7.14)$$

where we have used the identity

$$\Lambda^{\pm}(\nu) = \lambda(\nu) \pm (\tfrac{1}{2}i\pi)c\nu, \quad (7.15)$$

which follows from an application of the Plemelj formulas to Eq. (4.5). Equation (7.14) has been solved by Mitsis[5] with a different inhomogeneous term. The result is

$$N(x,z) = \frac{X_1(z)}{2\pi i}\left[a_0(x) + a_1(x)z + \int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu f(x,\nu \mid x_0)}{X_1^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu\right], \quad (7.16)$$

where

$$X_1(z) = \Lambda(z)/[(\nu_0^2 - z^2)(1 - c_0)]; \quad (7.17)$$

$a_0(x)$ and $a_1(x)$ are at this point arbitrary.

To evaluate the functions $a_0(x)$ and $a_1(x)$ we take advantage of the analytic properties of K in the complex μ plane. Note that

$$N(x, \pm\nu_0) = \frac{1}{2\pi i}\int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu K(x,\nu \mid x_0)}{\nu \mp \nu_0}\, d\nu. \quad (7.18)$$

Using the analyticity of K and the definition of ν_0, we have from Eq. (7.9) that

$$\int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu K(x,\nu \mid x_0)}{\nu \mp \nu_0}\, d\nu = -f(x, \pm\nu_0 \mid x_0), \quad (7.19)$$

so

$$N(x, \pm\nu_0) = -(1/2\pi i) f(x, \pm\nu_0 \mid x_0). \quad (7.20)$$

Since there are two forms of f, one for $\mu > 0$ and one for $\mu < 0$, a question arises as to which is the appropriate form to use for $\mu = \nu_0$ or $\mu = -\nu_0$. The answer is that it makes no difference: both forms are equal to the left-hand side of Eq. (7.19) and hence to each other. In what follows we consistently use the expression valid for $\mu \gtrless 0$ when substituting $\pm\nu_0$. Of course, the above consideration implies a constraint on the allowed form of f, which in turn must relate to the necessity of the inhomogeneous term of Eq. (7.1) being orthogonal to $\psi_0^\dagger$. This statement has not been proven, however.

Combining Eqs. (7.20) and (7.16), we obtain the result

$$f(x, \pm\nu_0 \mid x_0) = -X_1(\pm\nu_0)\left[a_0(x) \pm a_1(x)\nu_0 + \int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu f(x,\nu \mid x_0)\, d\nu}{X_1^+(\nu)\Lambda^-(\nu)(\nu \mp \nu_0)}\right]. \quad (7.21)$$

This provides us with two equations for the two unknowns a_0 and a_1. We may find K by use of Eqs. (7.13) and (7.16). The result is

$$K(x,\mu \mid x_0) = \frac{a_0(x) + a_1(x)\mu}{(\nu_0^2 - \mu^2)(1 - c_0)} + \lambda(\mu)g(c_0,\mu)f(x,\mu \mid x_0) + \frac{1}{(\nu_0^2 - \mu^2)(1 - c_0)} \times P\int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu f(x,\nu \mid x_0)}{X_1^+(\nu)\Lambda^-(\nu)(\nu - \mu)}\, d\nu. \quad (7.21')$$

This result expresses K in terms of its boundary values through f. It remains to find f.

We now derive a pair of coupled Fredholm equations, the solutions of which will yield the function $f(x,\mu \mid x_0)$. Putting $x = \pm b$ in Eq. (7.9) we have

$$\lambda(\mu)K(\pm b, \pm\mu \mid x_0) - P\int_0^1 \frac{(\tfrac{1}{2}c_0)\nu K(\pm b, \pm\nu \mid x_0)}{\nu - \mu}\, d\nu = f(\pm b, \pm\mu \mid x_0), \quad \mu \gtrless 0. \quad (7.22)$$

Introducing the transforms

$$N_k(b,z) = \frac{1}{2\pi i}\int_0^1 \frac{(\tfrac{1}{2}c_0)\nu K(\pm b, \pm\nu \mid x_0)}{\nu - z}\, d\nu, \quad k = \begin{cases}1,\\ 2,\end{cases} \quad (7.23)$$

we proceed as before. Equation (7.22) is reduced to a Hilbert problem which has been solved by Mitsis with a different inhomogeneous term. The result is

$$N_k(b,z) = \frac{X(z)}{2\pi i}\left[\alpha_k(b) + \int_0^1 \frac{(\tfrac{1}{2}c_0)\nu f(\pm b, \pm\nu \mid x_0)}{X^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu\right], \quad k = \begin{cases}1.\\ 2.\end{cases} \quad (7.24)$$

The numbers $\alpha_1(b)$ and $\alpha_2(b)$ are at this point arbitrary. Now put

$$f_k(b,z \mid x_0) = e^{-b/z}\int_0^1 \frac{(\tfrac{1}{2}c_0)\nu K(\pm b, \pm\nu \mid x_0)}{\nu + z}\, d\nu, \quad k = \begin{cases}1.\\ 2.\end{cases} \quad (7.25)$$

Then from Eq. (7.10)

$$f(x,\mu \mid x_0) = -\exp\left(-\frac{x}{\mu}\right) f_2(b,\mu \mid x_0) + \frac{u(x - x_0)}{\mu}\exp\left(\frac{x_0 - x}{\mu}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x,\mu), \quad \mu > 0, \quad (7.26)$$

$$f(x,\mu \mid x_0) = -\exp\left(-\frac{x}{\mu}\right) f_1(b, -\mu \mid x_0) - \frac{u(x_0 - x)}{\mu}\exp\left(\frac{x_0 - x}{\mu}\right) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x,\mu), \quad \mu < 0. \quad (7.27)$$

Thus, given $f_1(b,\mu \mid x_0)$ and $f_2(b,\mu \mid x_0)$, the function

$f(x, \mu \mid x_0)$ is determined. These functions are the solutions to the Fredholm equations which are obtained by noting that

$$f_k(b, z \mid x_0) = 2\pi i e^{-b/z} N_k(b, -z), \quad k = 1, 2, \quad (7.28)$$

and substituting into Eq. (7.24). The result is the following pair of coupled Fredholm equations:

$$f_1(b, \mu \mid x_0) = e^{-b/\mu} X(-\mu) \left[\alpha_1(b) + F_1(b, \mu) - \int_0^1 \frac{(\frac{1}{2}c_0)\nu e^{-b/\nu} f_2(b, \nu \mid x_0)}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)} \, d\nu \right], \quad (7.29)$$

$$f_2(b, \mu \mid x_0) = e^{-b/\mu} X(-\mu) \left[\alpha_2(b) + F_2(b, \mu) - \int_0^1 \frac{(\frac{1}{2}c_0)\nu e^{-b/\nu} f_1(b, \nu \mid x_0)}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)} \, d\nu \right], \quad (7.30)$$

where $\mu > 0$. The functions F_1 and F_2 are defined by the equation

$$F_k(b, \mu) = \int_0^1 \left\{ e^{\pm(x_0 \mp b)/\nu} - \frac{2}{c_0} \rho_0(x_0)\nu\psi_0(b, \nu) \right\} \times \frac{(\frac{1}{2}c_0)\, d\nu}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)}, \quad k = 1, 2. \quad (7.31)$$

The constant $\alpha_1(b)$ may be determined by substituting $\mu = \nu_0$ in Eq. (7.22) to obtain

$$\int_0^1 \frac{(\frac{1}{2}c_0)\nu K(b, \nu \mid x_0)}{\nu - \nu_0} \, d\nu = -f(b, \nu_0 \mid x_0). \quad (7.32)$$

Using Eq. (7.26) to evaluate the right-hand side of this expression and Eq. (7.24) for the left-hand side, we obtain an equation expressing $\alpha_1(b)$ in terms of the constant $f_2(b, \nu_0 \mid x_0)$ and an integral of the function $f_2(b, \nu \mid x_0)$. Similarly, we may put $\mu = -\nu_0$ in Eq. (7.22) and derive an expression for $\alpha_1(b)$ in terms of $f_1(b, \nu_0 \mid x_0)$ and an integral of $f_1(b, \nu \mid x_0)$.

We now examine the Fredholm equations, Eqs. (7.29) and (7.30). We introduce the function φ,

$$\varphi(\mu) = f_1(b, \mu \mid x_0) e^{-b/\mu}, \quad (7.33)$$

and the integral operator J defined such that

$$J\varphi = e^{-2b/\mu} X(-\mu) \int_0^1 \frac{(\frac{1}{2}c_0)\nu\varphi(\nu)\, d\nu}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)}. \quad (7.34)$$

Then the result of substituting for f_2 into Eq. (7.29) from Eq. (7.30) is, formally,

$$\varphi = G + J^2\varphi, \quad (7.35)$$

with an obvious definition of G. The Neumann series solution to Eq. (7.35) is

$$\varphi = G + J^2G + J^4G + \cdots. \quad (7.36)$$

To examine the convergence of this expansion we obtain a bound on the right-hand side:

$$\begin{aligned} |\varphi| &= |G + J^2G + J^4G + \cdots| \\ &\leq |G| + |J^2G| + |J^4G| + \cdots \\ &\leq \max|G| \, [1 + |J^2 1| + |J^4 1| + \cdots] \\ &\leq \max|G| \, [1 + (\max J1)^2 + (\max J1)^4 + \cdots]. \end{aligned}$$

The Neumann series for φ (hence f_1) converges if $\max J1 < 1$. Leonard and Mullikin[12] have shown that the required inequality

$$\frac{c_0}{2} \max_{0 \leq \mu \leq 1} \left[X(-\mu) e^{-2b/\mu} \int_0^1 \frac{\nu\, d\nu}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)} \right] < 1 \quad (7.37)$$

holds whenever $1 < 2be$, i.e., $2b > 0.368$. This corresponds to a c_0 in excess of 2.6, and therefore includes all cases of practical interest. Obviously, the same procedure could be followed to show that f_2 may also be expressed in a convergent Neumann series.

It is apparent that the foregoing results are rather highly implicit, necessitating an iterative procedure in addition to that required to solve the dual Fredholm equations. This does, of course, place a limitation on their practical usefulness. It should be noted that a similar situation exists for the "solution" to the unperturbed problem. The auxiliary condition in the latter case depends upon the continuum coefficient $A_0(\nu)$, which is the solution to a Fredholm equation itself depending upon the eigenvalue γ_0, but the eigenvalue is determined from the auxiliary condition (given b). Thus an "outer iteration" is also required in this case.

The reason that only a single Fredholm equation needs to be solved in the case of the unperturbed problem is that the duality which is introduced as a result of having a separate relation between ψ_0 and ρ_0 [see Eq. (5.22)] for the two partial ranges of μ can be removed by a consideration of the symmetry of ψ_0. It is shown below that it is possible to define a symmetric generalized Green's function which has the same property of depending on only a single Fredholm equation.

8. THE SYMMETRIC PART OF THE GENERALIZED GREEN'S FUNCTION

In the interest of both simplicity and clarity the numerical results in the next section relate to the asymptotic part (as $b \to \infty$) of the function

$$G_+(x \mid x_0) = \int_{-1}^{+1} K_+^\dagger(x, \mu \mid x_0)\, d\mu, \quad (8.1)$$

where $K_+^\dagger$ is the symmetric part of $K^\dagger$.

[12] Ref. 9, p. 17.

The symmetric part of $K^\dagger$ is given by

$$K_+^\dagger(x, \mu \mid x_0) = \tfrac{1}{2}[K^\dagger(x, \mu \mid x_0) + K^\dagger(-x, -\mu \mid x_0)]. \tag{8.2}$$

An equation for $K_+^\dagger$ may be derived by writing Eq. (5.12) once for x and μ, once for $-x$ and $-\mu$, and adding the results, making use of Eq. (8.2). The result is

$$B_0 K_+^\dagger(x, \mu \mid x_0) = \tfrac{1}{2}\delta(x - x_0) + \tfrac{1}{2}\delta(x + x_0) - \rho_0(x)\rho_0(x_0). \tag{8.3}$$

It is easy to see that if ψ_m is symmetric, Eq. (5.16) still holds. ψ_m will be symmetric if the term S_m in Eq. (5.2) is symmetric. This is always true for the case in which the perturbation is due to anisotropic scattering. It will also be true in the case of non-constant cross sections, if the perturbation is distributed symmetrically. It is also true that if S_m is antisymmetric, then the antisymmetric part of $K^\dagger$ suffices to determine the (antisymmetric) ψ_m. This fact is of little practical value, however, since an antisymmetric perturbation of the cross section does not imply an antisymmetric S_m.

If it is only G_+ that is desired, it is convenient to separate the function $K_+^\dagger$ out at the level of singular integral equations. Thus, if we write Eq. (7.9) twice, once for x and μ and once for $-x$ and $-\mu$, and add the results, we obtain

$$\lambda(\mu)K_+(x, \mu \mid x_0) - P\int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu K_+(x, \nu \mid x_0)}{\nu - \mu}\, d\nu = -e^{-x/\mu} f_+(b, \pm\mu \mid x_0) + H(\pm x, \pm\mu \mid x_0), \quad \mu \gtrless 0. \tag{8.4}$$

We have introduced the definitions

$$f_+(b, \mu \mid x_0) = e^{-b/\mu}\int_0^1 \frac{(\tfrac{1}{2}c_0)\nu K_+(b, \nu \mid x_0)}{\nu + \mu}\, d\nu, \tag{8.5}$$

$$H(x, \mu \mid x_0) = \frac{1}{2\mu}\exp\left(\frac{x_0 - x}{\mu}\right)u(x - x_0) + \frac{1}{2\mu}\exp\left(-\frac{x_0 + x}{\mu}\right)u(x_0 + x) - \frac{2}{c_0}\rho_0(x_0)\psi_0(x, \mu). \tag{8.6}$$

This equation may be treated in precisely the same way as Eq. (7.9). We introduce the complex transformation

$$N_+(x, z) = \frac{1}{2\pi i}\int_{-1}^{+1} \frac{(\tfrac{1}{2}c_0)\nu K_+(x, \nu \mid x_0)}{\nu - z}\, d\nu, \tag{8.7}$$

and obtain

$$N_+(x, z) = \frac{X_1(z)}{2\pi i}\left[\beta_0(x) + \beta_1(x)z - \int_{-1}^{0} \frac{(\tfrac{1}{2}c_0)\nu f_+(b, -\nu \mid x_0)e^{-x/\nu}}{X_1^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu - \int_0^1 \frac{(\tfrac{1}{2}c_0)\nu f_+(b, \nu \mid x_0)e^{-x/\nu}}{X_1^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu + \int_{-1}^{0} \frac{(\tfrac{1}{2}c_0)\nu H(-x, -\nu \mid x_0)}{X_1^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu + \int_0^1 \frac{(\tfrac{1}{2}c_0)\nu H(x, \nu \mid x_0)}{X_1^+(\nu)\Lambda^-(\nu)(\nu - z)}\, d\nu\right]. \tag{8.8}$$

To find β_0 and β_1 we again put $\mu = \pm\nu_0$ in Eq. (8.4) and obtain

$$2\pi i N_+(x, \pm\nu_0) = e^{\mp x/\nu_0} f_+(b, \nu_0 \mid x_0), - H(\pm x, \nu_0 \mid x_0), \tag{8.9}$$

We have used the same convention in making this substitution as before [see discussion following Eq. (7.20)]. Equation (8.9) may be substituted into Eq. (8.8) to obtain two equations in the two unknowns β_0 and β_1. The solution for β_0 is

$$\beta_0(x) = \frac{f_+(b, \nu_0 \mid x_0)}{X_1(\nu_0)}\cosh\left(\frac{x}{\nu_0}\right) - \frac{1}{2X_1(\nu_0)}[H(x, \nu_0 \mid x_0) + H(-x, \nu_0 \mid x_0)] + c_0(c_0 - 1)\int_0^1 \nu^2 g(c_0, \nu)\cosh\left(\frac{x}{\nu}\right) f_+(b, \nu \mid x_0)\, d\nu - \frac{c_0(c_0 - 1)}{2}\int_0^1 \nu^2 g(c_0, \mu) \times [H(x, \nu \mid x_0) + H(-x, \nu \mid x_0)]\, d\nu. \tag{8.10}$$

We do not write down the corresponding expression for $\beta_1(x)$, since it is not needed in the derivation of G_+.

From the definition of G_+, it is clear that

$$G_+(x \mid x_0) = 2\pi i(2/c_0)N(x, 0). \tag{8.11}$$

Putting $z = 0$ in Eq. (8.8) and using Eq. (8.10), we obtain, after some simplification,

$$G_+(x \mid x_0) = \frac{2X_1(0)}{c_0}\left[\frac{f_+(b, \nu_0 \mid x_0)}{X_1(\nu_0)}\cosh\left(\frac{x}{\nu_0}\right) - \frac{1}{2X_1(\nu_0)}\{H(x, \nu_0 \mid x_0) + H(-x, \nu_0 \mid x_0)\} + c_0(c_0 - 1)\nu_0^2\int_0^1 g(c_0, \nu)\cosh\left(\frac{x}{\nu}\right) f_+(b, \nu \mid x_0)\, d\nu - \frac{c_0(c_0 - 1)}{2}\nu_0^2\int_0^1 g(c_0, \nu) \times \{H(x, \nu \mid x_0) + H(-x, \nu \mid x_0)\}\, d\nu\right]. \tag{8.12}$$

This is an exact expression for G_+ in terms of the yet to be determined function $f_+(b, \nu \mid x_0)$. Comparison of the definitions of f_+, f_1, and f_2, Eqs. (8.5) and (8.25), and use of Eq. (8.2) leads to the result

$$f_+(b, \nu \mid x_0) = \tfrac{1}{2}[f_1(b, \nu \mid x_0) + f_2(b, \nu \mid x_0)]. \quad (8.13)$$

We can therefore obtain a single Fredholm equation for f_+ by adding Eqs. (7.29) and (7.30). Putting

$$\alpha_+(b) = \tfrac{1}{2}[\alpha_1(b) + \alpha_2(b)], \quad (8.14)$$

$$F_+(b, \mu) = \tfrac{1}{2}[F_1(b, \mu) + F_2(b, \mu)], \quad (8.15)$$

we have

$$f_+(b, \mu \mid x_0) = e^{-b/\mu}X(-\mu)\Big[\alpha_+(b) + F_+(b, \mu) - \int_0^1 \frac{(\tfrac{1}{2}c_0)\nu e^{-b/\nu} f_+(b, \nu \mid x_0)}{X^+(\nu)\Lambda^-(\nu)(\nu + \mu)}\, d\nu\Big]. \quad (8.16)$$

This determines G_+ completely. From Eq. (8.16) f_+ is $O(e^{-b/\mu})$ as $b \to \infty$. It follows that the first integral on the right-hand side of Eq. (8.12) can be neglected far from boundaries, since the integrand is $O(e^{-(b-x)/\mu})$. It is in this sense that we compute the "asymptotic" part of G_+. Note that our results will be asymptotically exact no matter how small the quantity $|x - x_0| > 0$, so long as $|x - b|$ is sufficiently large. Thus our results are not comparable with diffusion theory without further simplification, and may be expected to be very good for thick enough slabs.

The simplification of Eq. (8.12) for the asymptotic part of G_+ ($f_+ = 0$) is straightforward with the following exception: it is necessary to evaluate the asymptotic part of $\psi_0(x, \mu)$ at $\mu = \pm\nu_0$. Examination of Eqs. (4.8) and (4.3) ($n = 0$) shows that if we consider the asymptotic part of ψ_0 to be given by the right-hand side of Eq. (4.8) exclusive of the integral term, then it has simple poles at $\pm\nu_0$. This difficulty can be circumvented by considering the asymptotic part of ψ_0 to be given by Eq. (5.22) ($m = 0$, $S_0 = 0$) with ρ_0 replaced by the asymptotic part of ρ_0. Since the latter function is continuous, the right-hand side of Eq. (5.22) is analytic for all $\mu \neq 0$. Subject to this convention, Eq. (8.12) becomes

$$G_+(x \mid x_0) \simeq A \cos\left(\frac{x}{|\nu_0|}\right) - \frac{X_1(0)}{2\,|\nu_0|\, c_0 X_1(\nu_0)} \times \left\{\sin\left(\frac{|x + x_0|}{|\nu_0|}\right) + \sin\left(\frac{|x - x_0|}{|\nu_0|}\right) - 8a_+^2 \cos\left(\frac{x_0}{|\nu_0|}\right) x \sin\left(\frac{x}{|\nu_0|}\right)\right\} + \frac{1}{2}\int_0^1 \frac{g(c_0, \nu)}{\nu} \times \left\{\exp\left(-\frac{|x + x_0|}{\nu}\right) + \exp\left(-\frac{|x - x_0|}{\nu}\right)\right\} d\nu. \quad (8.17)$$

The notation $\simeq$ means asymptotically exact as $b \to \infty$. Since the asymptotic part of ρ_0 is proportional to $\cos(x/|\nu_0|)$, the constant A is arbitrary.

9. NUMERICAL COMPARISON OF THE THREE FORMS OF THE GREEN'S FUNCTION

The even part of Eq. (6.9) is

$$G_+(x \mid x_0) = \tfrac{1}{2}E_1(|x - x_0|) + \tfrac{1}{2}E_1(|x + x_0|) - \frac{2}{c_0}\rho_0(x)\rho_0(x_0) + \gamma_0 \sum_{n=1}^{\infty} \frac{\phi_{2n}(x)\phi_{2n}(x_0)}{\gamma_{2n}(\gamma_{2n} - \gamma_0)}. \quad (9.1)$$

The even part of Eq. (6.12) is

$$G_+(x \mid x_0) = \sum_{n=1}^{\infty} \frac{\phi_{2n}(x)\phi_{2n}(x_0)}{\gamma_{2n} - \gamma_0}. \quad (9.2)$$

In the following the above expressions for $G_+^\dagger$ are referred to as A and B, respectively. Equation (8.17) is referred to as C. In calculating with forms A and B, we use the asymptotic parts of the ϕ_{2n} (putting the integral terms equal to zero). The following calculations were done for a slab of half-thickness $b = 4.196$. The corresponding c_0 is 1.034. We choose $x_0 = 2.0$.

Figure 1 shows a comparison of forms B and C. Shown are one-, two-, and four-term expansions (the three-term expansion differed very little from the two-term). The agreement between the two forms is seen to be rather good for the four-term expansion (we are trying to represent a logarithmic singularity with continuous functions). It should be remarked that the curves resulting from the eigenfunction

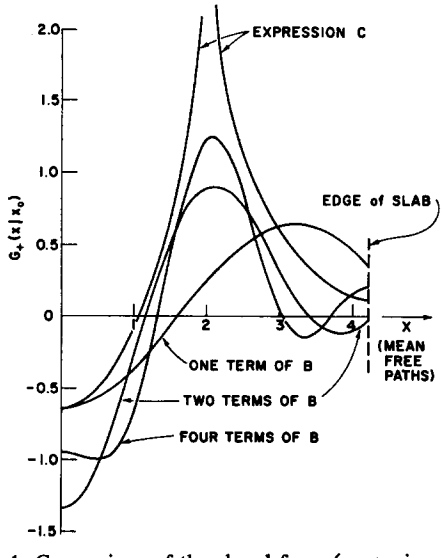

FIG. 1. Comparison of the closed form (expression C) with an eigenfunction expanded form (expression B) of the generalized Green's function.

expansion are not orthogonal to $\rho_0(x)$ because only the asymptotic parts were taken, and the correct extrapolation distances used. Since we are free to add an arbitrary multiple of $\rho_0(x)$ to form B, it is clear that we could shift the curves upward to agree better with form C. In Fig. 2 we compare forms A and C. The general agreement between the four-term expansion and form C is seen to be excellent. As mentioned above, we can improve the agreement by adding on a multiple of $\rho_0(x)$ to form A. This is shown in Fig. 3.

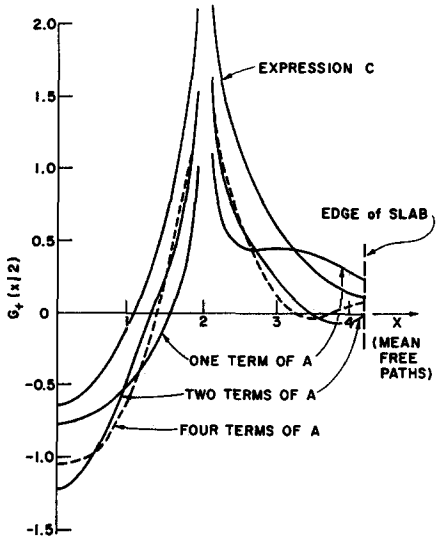

FIG. 2. Comparison of the closed form (expression C) with an eigenfunction expanded form (expression A) of the generalized Green's function.

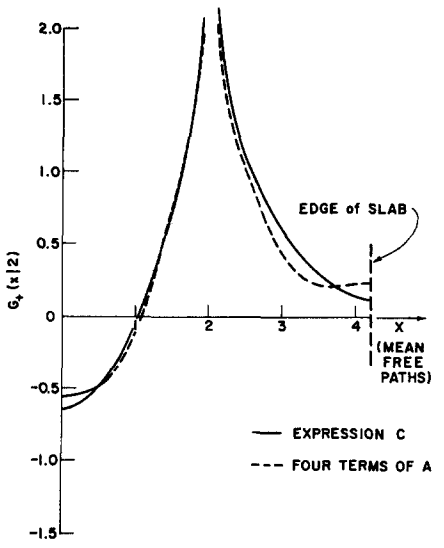

FIG. 3. Comparison of the closed form (expression C) with a renormalized four-term eigenfunction expansion (expression A) of the generalized Green's function.

ACKNOWLEDGMENTS

The author is pleased to acknowledge the advice and encouragement of his thesis supervisors, Professor K. F. Hansen and Professor T. H. Dupree. The author was supported at MIT by the U.S. Atomic Energy Commission through its Special Fellowship Program in Nuclear Science and Engineering, which is administered by the Oak Ridge Institute of Nuclear Studies. Dr. K. D. Lathrop read an early draft of this paper, and made a number of helpful comments and criticisms.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Drukarev Transformation of Dirac Equation*

HENRY BRYSK
Union Carbide Corporation, Space Science and Engineering Laboratory, White Plains, New York

(Received 16 January 1967)

Drukarev has transformed the Fredholm equation that is the Green's function formulation of the solution of the Schrödinger equation into a Volterra equation. The present paper exhibits the corresponding result for the Dirac equation. The advantages of this technique in the numerical evaluation of phase shifts (and wavefunctions) are discussed.

1. GREEN'S FUNCTION

THE Green's function formulation of the Dirac equation for a central potential was given by Rose.[1] Denoting as usual by F/r and G/r the radial functions, by F_0/r and G_0/r the free-space functions (regular solution), and by $\bar{F}_0/r$ and $\bar{G}_0/r$ the corresponding irregular solution,

$$F = F_0\left[1 + \int_r^\infty (F\bar{F}_0 + G\bar{G}_0)V\,dr'\right] + \bar{F}_0\int_0^r (FF_0 + GG_0)V\,dr', \quad (1a)$$

$$G = F_0\left[1 + \int_r^\infty (F\bar{F}_0 + G\bar{G}_0)V\,dr'\right] + \bar{G}_0\int_0^r (FF_0 + GG_0)V\,dr', \quad (1b)$$

where explicitly,

$$F_0 = S(\kappa)[(E-1)/k]^{\frac{1}{2}}krj_{l(-\kappa)}(kr), \quad (2a)$$

$$G_0 = [(E+1)/k]^{\frac{1}{2}}krj_{l(\kappa)}(kr), \quad (2b)$$

and the irregular solutions are obtained by replacing the Bessel functions by Neumann functions (i.e., $j_l \to y_l$). More generally (as Rose points out), F_0 and G_0 could be wavefunctions of a potential V_0 (typically, a Coulomb potential), in which case V in Eqs. (1) would be replaced by $V - V_0$.

2. DRUKAREV TRANSFORMATION

Following Drukarev's approach,[2] Eqs. (1) are rewritten as

$$F = F_0\left[1 + \int_0^\infty (F\bar{F}_0 + G\bar{G}_0)V\,dr'\right] - F_0\int_0^r (F\bar{F}_0 + G\bar{G}_0)V\,dr' + \bar{F}_0\int_0^r (FF_0 + GG_0)V\,dr', \quad (3a)$$

$$G = F_0\left[1 + \int_0^\infty (F\bar{F}_0 + G\bar{G}_0)V\,dr'\right] - G_0\int_0^r (F\bar{F}_0 + G\bar{G}_0)V\,dr' + \bar{G}_0\int_0^r (FF_0 + GG_0)V\,dr'. \quad (3b)$$

The change of variable

$$\varphi = CF, \quad \Gamma = CG \quad (4)$$

is then introduced, where

$$C = \left[1 + \int_0^\infty (F\bar{F}_0 + G\bar{G}_0)V\,dr'\right]^{-1} = \left[1 + C^{-1}\int_0^\infty (\varphi\bar{F}_0 + \Gamma\bar{G}_0)V\,dr'\right]^{-1}, \quad (5)$$

or, on inverting,

$$C = 1 - \int_0^\infty (\varphi\bar{F}_0 + \Gamma\bar{G}_0)V\,dr'. \quad (6)$$

Equations (3) then reduce to a coupled pair of Volterra equations. With the convenient notation

$$C(r) = 1 - \int_0^r (\varphi\bar{F}_0 + \Gamma\bar{G}_0)V\,dr', \quad (7a)$$

$$S(r) = -\int_0^r (\varphi F_0 + \Gamma G_0)V\,dr', \quad (7b)$$

the Volterra equations are

$$\varphi(r) = C(r)F_0(r) - S(r)\bar{F}_0(r), \quad (8a)$$

$$\Gamma(r) = C(r)G_0(r) - S(r)\bar{G}_0(r). \quad (8b)$$

Asymptotically,

$$\Gamma \sim [(E+1)/k]^{\frac{1}{2}}[C\sin(kr - l(\kappa)\pi/2) - S\cos(kr - l(\kappa)\pi/2)], \quad (9)$$

$$G = C^{-1}\Gamma \sim [(E+1)/k]^{\frac{1}{2}}[\sin(kr - l(\kappa)\pi/2) - (S/C)\cos(kr - l(\kappa)\pi/2)], \quad (10)$$

leading to the identification

$$\tan\delta_\kappa = \frac{S(\infty)}{C(\infty)} = -\frac{\int_0^\infty (\varphi F_0 + \Gamma G_0)V\,dr'}{1 - \int_0^\infty (\varphi\bar{F}_0 + \Gamma\bar{G}_0)V\,dr'}. \quad (11)$$

* Work supported by the National Aeronautics and Space Administration under Contract No. NASw-1235.

[1] M. E. Rose, Phys. Rev. **82**, 389 (1951).

[2] G. F. Drukarev, Zh. Eksperim. i Teor. Fiz. **25**, 139 (1953).

For the alternative normalization convention

$$G \sim [(E+1)/k]^{\frac{1}{2}} \sin [kr - l(\kappa)(\pi/2) + \delta_\kappa]$$
$$= [(E+1)/k]^{\frac{1}{2}} [\cos \delta_\kappa \sin (kr - l(\kappa)\pi/2) + \sin \delta_\kappa \cos (kr - l(\kappa)\pi/2)], \quad (12)$$

there results

$$\cos \delta_\kappa = C(\infty)/[C^2(\infty) + S^2(\infty)]^{\frac{1}{2}},$$
$$\sin \delta_\kappa = S(\infty)/[C^2(\infty) + S^2(\infty)]^{\frac{1}{2}}. \quad (13)$$

3. COMPUTATIONAL CONSIDERATIONS

The formal properties of this approach have been fully discussed previously for the Schrödinger equation,[3] and the arguments carry over. Some remarks pertinent to numerical solution are presented here.

(1) There is no normalization problem, as the solutions start out with the free-space functions.

(2) The direct numerical solution of the radial equations requires numerical integration for the wavefunctions which (away from the origin) are oscillatory—a delicate procedure. Here, the numerical integration is for the slowly varying functions $C(r)$ and $S(r)$, the oscillatory behavior appearing in terms of analytic functions (the spherical Bessel and Neumann functions, expressible as sines and cosines times polynomials in $1/r$).

(3) The integration procedure directly yields $S(r)$ and $C(r)$, integrals whose limits are $\sin \delta$ and $\cos \delta$. This integral representation of the phase shift converges faster and more dependably than does the determination of the phase shift by matching the wave to its asymptotic form.

(4) If the potential cuts off at $r = a$, then for $r > a$ Eqs. (8) become

$$\varphi(r) = C(a)F_0(r) - S(a)\bar{F}_0(r), \quad (14a)$$
$$\Gamma(r) = C(a)G_0(r) - S(a)\bar{G}_0(r), \quad (14b)$$

yielding directly the appropriate analytical form for the wavefunctions, not only their numerical value at the cutoff.

(5) The last observation leads into a simple iterative device for improving on a foreshortened solution. From Eqs. (7),

$$C(\infty) = C(a) - \int_a^\infty (\varphi \bar{F}_0 + \Gamma \bar{G}_0) V \, dr', \quad (15a)$$
$$S(\infty) = S(a) - \int_a^\infty (\varphi F_0 + \Gamma G_0) V \, dr'. \quad (15b)$$

Suppose now that the numerical integration has been stopped at $r = a$, but the potential extends beyond. The contribution from the tail to the phase shift can be evaluated approximately by carrying out the integral in Eq. (15) with φ and Γ represented by Eqs. (14). If the potential is given analytically (or is fitted to an analytic expression for $r > a$), the integral may be carried out analytically. If the error in the phase shift upon truncation at $r = a$ is of order ϵ, after this approximate evaluation of the tail contribution there will be an error of order ϵ^2 only.

(6) Suppose the Dirac equation has been solved by some other numerical technique up to $r = a$, yielding the unnormalized wavefunctions

$$F_u = NF, \quad G_u = NG, \quad (16)$$

with N the normalization constant. Define

$$C_u(r) = CNC(r), \quad S_u(r) = CNS(r). \quad (17)$$

Then, in view of Eq. (4), multiplication of Eqs. (8) by CN gives

$$F_u(r) = C_u(r)F_0(r) - S_u(r)\bar{F}_0(r), \quad (18a)$$
$$G_u(r) = C_u(r)G_0(r) - S_u(r)\bar{G}_0(r). \quad (18b)$$

From Eqs. (18) at $r = a$, the values of $C_u(a)$ and $S_u(a)$ are known. Then the counterparts of Eqs. (15),

$$C_u(\infty) = C_u(a) - \int_a^\infty (F_u \bar{F}_0 + G_u \bar{G}_0) V \, dr', \quad (19a)$$
$$S_u(\infty) = S_u(a) - \int_a^\infty (F_u F_0 + G_u G_0) V \, dr', \quad (19b)$$

permit the integration to be carried forward. The phase shifts are still obtained from Eq. (13), which applies equally well to the unnormalized (subscripted) C's and S's. The wavefunctions (if desired) now require normalization. This possibility of switching integration schemes is of particular interest in the case of singular potentials (such as a screened Coulomb potential) for which the Volterra equation formulation is troublesome at the origin.

[3] H. Brysk, Phys. Rev. 133, B1625 (1964). The reader is referred to this article for extensive references.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Relativistic and Causal Theories with Four-Fermion Interaction Hamiltonians*

F. A. BLOOD, JR.
Brookhaven National Laboratory, Upton, New York

(Received 12 December 1966)

Possible ways of constructing field-theoretic operators satisfying the commutation relations of the inhomogeneous Lorentz group are investigated along lines laid down by Dirac. They can be satisfied in both the instant form, in which operators representing rotations and translations in space remain unchanged, and the point form, in which operators representing the homogeneous Lorentz group G_4^1 remain unchanged, provided one can find a causal Hamiltonian density such that $[H(t, \mathbf{x}), H(t, \mathbf{y})]$ is proportional to $\delta(\mathbf{x} - \mathbf{y})$ and which transforms as a scalar under G_4^1. Less restrictive sufficient conditions in the instant form are found, similar to those of Dirac. The commutator can be proportional to derivatives of $\delta(\mathbf{x} - \mathbf{y})$ if the coefficients on the derivatives satisfy a certain condition. The only way found to satisfy these conditions for an interaction Hamiltonian constructed from fields for identical spin $\frac{1}{2}$ particle (in the interaction picture) is to have the commutator proportional to $\delta(\mathbf{x} - \mathbf{y})$, which implies local coupling with no derivations. The possibility of having relativistic theories in the instant form without causality is also investigated for the case of a four-fermion interaction Hamiltonian constructed from creation and annihilation operators for a spin $\frac{1}{2}$ particle, but no definite conclusion is arrived at.

INTRODUCTION

IT is a well-known fact that the local four-fermion (4F) interaction Hamiltonian

$$H' = \int d^3x \tilde{\psi}_P(\mathbf{x})\theta^\mu \psi_N(\mathbf{x})\tilde{\psi}_e(\mathbf{x})\theta_\mu \psi_\nu(\mathbf{x}), \qquad \theta^\mu = \gamma^\mu(1 + i\gamma_5) \tag{1}$$

leads to a nonrenormalizable S matrix. One might, therefore, attempt to modify the Hamiltonian in some fashion so that the resulting S matrix will become renormalizable. For this reason, and also because of the intrinsic interest that 4F Hamiltonians have, we have initiated a search for 4F Hamiltonians which give a causal and relativistic theory.[1] In the first two sections we consider rather general methods of obtaining a relativistic and causal theory for any type of field theoretical Hamiltonian. Then, in the third section, we specialize to a Hamiltonian constructed from the creation and annihilation operators of a single type of spin $\frac{1}{2}$ fermion.

We do not find any startling new results in this paper, in particular no 4F Hamiltonians giving a renormalizable S matrix are found, but rather consider it to be a guide to the difficulties and possibilities of constructing a relativistic theory with a 4F interaction Hamiltonian.

If a physical theory is to be relativistic, then there must be a set of unitary operators[2-4] $U(\Lambda, a)$ obeying the group multiplication law of the inhomogeneous Lorentz groups (IHLG), i.e.,

$$U(\Lambda', a')U(\Lambda, a) = U(\Lambda'\Lambda, \Lambda' a + a'), \tag{2}$$

$$U(\Lambda^{-1}, -\Lambda^{-1}a) = U^{-1}(\Lambda, a) = U^*(\Lambda, a), \tag{2a}$$

where $U(\Lambda, a)$ is the operator representing the transformation Λ from the homogeneous Lorentz group G_4^1 followed by the translation a. The method normally employed for finding operators satisfying Eq. (2) is to find operators satisfying the commutation relations which the infinitesimal generators (IG's) of the IHLG must satisfy, and this is essentially the method we use. That is, we suppose that when Λ is very near the identity I and a_μ is nearly 0, then[5]

$$U(\Lambda, a) = U(I, 0) + i\{a_\mu P_\mu + \theta_i J_i + v_i k_i\}, \tag{3}$$

where θ_i, v_i are very small numbers, P_i is the IG of translations along the ith axis, $P_0 = H$ is the IG of time translations, J_i is the IG of rotations about the ith axis, and K_i is the IG of "velocity boosts" along the ith axis. The multiplication laws (2) then imply that these IG's must satisfy certain commutation relations [Eq. (9)].

The implication when one uses IG's is that the physical systems will be specified on some surface, called the initial surface, and then the IG's will be used to integrate forward and find the state of the system on a different surface. The usual initial surface

* Work performed under auspices of the U.S. Atomic Energy Commission.

[1] We consider only the *proper* inhomogeneous Lorentz group and thus do not deal with parity or time reversal.

[2] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949).

[3] P. A. M. Dirac, Rev. Mod. Phys. **34**, 592 (1962).

[4] L. L. Foldy, Phys. Rev. **122**, 275 (1961).

[5] A word about our summation convention: Repeated μ and ν subscripts indicate a covariant sum, i.e., $a \cdot b = a_0 b_0 - \mathbf{a} \cdot \mathbf{b} = a_\mu b_\mu = a^\mu b_\mu = a^\mu b^\mu$; repeated α and β subscripts indicate a normal sum from 1 to 4 while repeated i and j subscripts indicate a sum from 1 to 3.

is a constant-time surface and one uses the Hamiltonian to find the state of the system on a later constant-time surface. Other less familiar choices of initial surfaces are possible, however; we use the hyperboloids $x_0^2 - \mathbf{x}^2 = d^2$, where d^2 is a positive constant.

The question of what initial surface to use is tied in with how one tries to satisfy the commutation relations. If we assume that some subset, which must also be a subgroup, of the $U(\Lambda, a)$ remains the same as the free operators[6] when the interaction is "turned on," then the IG's of this subgroup also remain the same and hence some of the commutation relations are automatically satisfied. Dirac has suggested,[2] and we follow his suggestion, that the subgroup left the same should be a subgroup, called an initial subgroup, which leaves an initial surface invariant because the initial subgroup is not involved in integrating the system forward and thus does not involve the interactions. If the initial subgroup is one which leaves the initial surface $t = \text{const}$ invariant (the set of all rotations and translations in 3-space), then $\mathbf{J}$ and $\mathbf{P}$ remain the same while H and $\mathbf{K}$ are changed. This gives the *instant* form of Dirac. We have a somewhat more symmetrical-looking situation in the case where the initial surface is one for which x^2 is a constant, for there $\mathbf{J}$, $\mathbf{K}$ remain the same while H, $\mathbf{P}$ are changed to give the *point* form of Dirac.

In both the instant and the point form, three sets of commutation relations are automatically satisfied because the initial subgroup is unchanged, and four more can be satisfied by assuming the interaction Hamiltonian is the integral, over the initial surface, of a Hamiltonian density having certain transformation properties under the initial subgroup. The remaining two commutation relations are quadratic in the Hamiltonian density and it is here that all the difficulties are concentrated. These quadratic relations can be satisfied in the instant form provided the density $H(x)$ transforms as a scalar under $U^{(0)}(\Lambda, 0)$, where Λ belongs to G_4^1, and

$$\int d^3x\, d^3y (x_i - y_i)[H(\mathbf{x}), H(\mathbf{y})] = 0, \tag{4a}$$

$$\int d^3x\, d^3y (x_i y_j - x_j y_i)[H(\mathbf{x}), H(\mathbf{y})] = 0. \tag{4b}$$

And they can be satisfied in the point form, provided $H(x)$ is again a scalar under $U^{(0)}(\Lambda, 0)$ and

$$\int d^3x\, d^3y \left(\frac{x_i}{x_0} - \frac{y_i}{x_0}\right)[H(x), H(y)] = 0,$$

$$x_0 = (\mathbf{x}^2 + d^2)^{\frac{1}{2}}, \quad y_0 = (\mathbf{y}^2 + d^2)^{\frac{1}{2}}, \quad d^2 > 0. \tag{5}$$

There are various ways to proceed in attempting to satisfy Eqs. (4). The most general method is to write $H(x)$ as an integral in momentum space over products of four creation and/or annihilation operators times a function of the four momenta [Eq. (38)] and then obtain conditions on the function implied by Eq. (4). This is done in Sec. 3. The general conditions obtained are too difficult to work with, however, and so we are led to try a somewhat simpler approach.

Since the difference between any two points on either of the initial surfaces we have used is never timelike, it would seem that one could try to satisfy the commutation relations by letting

$$[H(x), H(y)] = 0, \quad (x - y)^2 < 0, \tag{6a}$$

$$[H(t, \mathbf{x}), H(t, \mathbf{y})] = a\delta^{(3)}(\mathbf{x} - \mathbf{y}) + b_j \partial_j \delta^{(3)}(\mathbf{x} - \mathbf{y}) + c_{ij}\partial_i \partial_j \delta^{(3)}(\mathbf{x} - \mathbf{y}) + \cdots, \tag{6b}$$

where the series is a finite one in $\delta^{(3)}$, and its derivatives and a, b_i, c_{ij}, $\cdots$ are functions of $\mathbf{x}$, $\mathbf{y}$, t.[7]

In Sec. 1, we discuss possible ways of satisfying the commutation relations in the instant form. The general line of discussion follows Dirac[3] and holds for any interaction Hamiltonian. The interaction Hamiltonian is written as

$$H' = \int d^3x H(\mathbf{x})$$

with $H(x)$ a scalar under G_4^1 and Eqs. (6) are assumed to hold. A sufficient condition on the coefficients $a, b, c, \cdots$ of Eq. (6b) is found such that the Hamiltonian density will give a relativistic (and causal) theory. The only 4F densities we find which satisfy the equation, however, have $b = c = \cdots = 0$ and are the usual local ones:

$$H(x) = \sum_{m=1}^{5} G_m H_m(x),$$

$$H_m(x) = \bar{\psi}(x) A_i^{(m)} \psi(x) \bar{\psi}(x) A_i^{(m)} \psi(x), \tag{7}$$

$$A^{(1)} = I, \quad A^{(2)} = i\gamma_5, \quad A_\mu^{(3)} = \gamma_\mu,$$

$$A_\mu^{(4)} = i\gamma_5\gamma_\mu, \quad A_{\mu\nu}^{(5)} = \sigma_{\mu\nu},$$

where the G_m are arbitrary real constants.

Possible methods of satisfying the commutation

[6] Free operators are those which transform free states (i.e., states before the interaction Hamiltonian is "turned on") from one coordinate system to another. The free Hamiltonian is just the kinetic energy part of the Hamiltonian with the interaction part equal to 0. Free operators will be denoted by a 0 superscript.

[7] The requirement that Eq. (6a) holds is called (microscopic) causality because one should be able to simultaneously measure the energy density at points of spacelike separation. If this is to be so, then evidently the energy-density operators at spacelike separations must commute.

relations in the point form are discussed in Sec. 2. Again all of the discussion holds for any interaction Hamiltonian. The only way found to obtain a relativistic theory in this form was to suppose that

$$H' = \int d^4x\theta(x_0)\delta(x^2 - m^2)H(x)x_0 \tag{8}$$

with $H(x)$, a scalar under G_4^1, satisfying Eqs. (6a) and (6b) with $b = c = \cdots = 0$; i.e., the density must be that described in Eq. (7).

Instead of using a causal Hamiltonian in the instant form and obtaining conditions on the coefficients a, b, c, $\cdots$, one can go directly to Eqs. (4) and see what conditions they imply on the Hamiltonian. This is done in Sec. 3 for 4F interaction Hamiltonians, but the conditions derived there are too complicated to yield any further information.

1. THE INSTANT FORM

If we are to have operators satisfying Eq. (2), then the 10 IG's of Eq. (3) must satisfy the commutation relations

$$[P_i, P_j] = 0, \tag{9a}$$
$$[J_i, P_j] = i\epsilon_{ijk}P_k, \tag{9b}$$
$$[J_i, J_j] = i\epsilon_{ijk}J_k, \tag{9c}$$
$$[H, P_i] = 0, \tag{9d}$$
$$[H, J_i] = 0, \tag{9e}$$
$$[P_i, K_j] = -i\delta_{ij}H, \tag{9f}$$
$$[J_i, K_j] = i\epsilon_{ijk}K_k, \tag{9g}$$
$$[H, K_i] = -iP_i, \tag{9h}$$
$$[K_i, K_j] = -i\epsilon_{ijk}J_k. \tag{9i}$$

The assumption of the instant form is that $U(R, \mathbf{a}) = U^{(0)}(R, \mathbf{a})$, where R belongs to $O(3)$, and thus $\mathbf{J}$, $\mathbf{P}$ remain the same as in the free case. This implies that the first three commutation relations are satisfied.

Equations (9d)–(9g) are the infinitesimal forms of the transformation rules of H, K_i under $U^{(0)}(R, \mathbf{a})$. For example, (9d) and (9e) can be derived from $U(R, \mathbf{a})U(0, a_0)U^{-1}(R, \mathbf{a}) = U(0, a_0)$ for very small $\mathbf{a}$, a_0, and R near the identity. It is more convenient to use the fully integrated forms of Eqs. (9d)–(9g) rather than having to calculate $\mathbf{J}$, $\mathbf{P}$ and then calculate their commutators with H, $\mathbf{K}$, so, for these equations, we substitute

$$U^{(0)}(I, \mathbf{a})HU^{(0)^{-1}}(I, \mathbf{a}) = H, \tag{10a}$$
$$U^{(0)}(R, 0)HU^{(0)^{-1}}(R, 0) = H, \tag{10b}$$
$$U^{(0)}(I, \mathbf{a})K_iU^{(0)^{-1}}(I, \mathbf{a}) = K_i + a_iH, \tag{10c}$$
$$U^{(0)}(R, 0)K_iU^{(0)^{-1}}(R, 0) = R_{ij}^{-1}K_j. \tag{10d}$$

We are assuming that H is the sum of a free plus an interaction part, and so K_i will also be the sum of a free plus an "interaction" part, i.e.,

$$H = H^{(0)} + H', \tag{11a}$$
$$K_i = K_i^{(0)} + K_i'. \tag{11b}$$

Since the free operators evidently satisfy Eq. (10), we must have

$$U^{(0)}(I, \mathbf{a})H'U^{(0)^{-1}}(I, \mathbf{a}) = H', \tag{12a}$$
$$U^{(0)}(R, 0)H'U^{(0)^{-1}}(R, 0) = H', \tag{12b}$$
$$U^{(0)}(I, \mathbf{a})K_i'U^{(0)^{-1}}(I, \mathbf{a}) = K_i' + a_iH', \tag{12c}$$
$$U^{(0)}(R, 0)K_i'U^{(0)^{-1}}(R, 0) = R_{ij}^{-1}K_j'. \tag{12d}$$

The commutation relations (9h) and (9i) are the difficult ones to satisfy in the instant form, so we defer a discussion of them until after we have shown how to satisfy Eqs. (12).

We start out with the general form

$$H' = \int d^4x f(x_0\mathbf{x})H'(x), \tag{13}$$

where

$$U^{(0)}(I, a)H'(0)U^{(0)}(I, -a) = H'(-a), \tag{14}$$

and derive

$$U^{(0)}(I, \mathbf{a})H'U^{(0)}(I, -\mathbf{a}) = \int d^4x f(x_0, \mathbf{x} + \mathbf{a})H'(x).$$

If the right-hand side is to equal H', then we must have f independent of $\mathbf{x}$ so that $f = f(x_0)$. We can then integrate over x_0 and replace $\int dx_0 f(x_0)H'(x)$ by $H(\mathbf{x})$ to obtain

$$H' = \int d^4x\delta(x_0)H'(x) = \int d^3xH(\mathbf{x}) \tag{15}$$

as the most general form for H' allowed by Eq. (10a).

A necessary and sufficient condition that the H' of Eq. (15) satisfy Eq. (12b) is that $H(\mathbf{x})$ be a scalar under rotations from $O(3)$ i.e.,

$$U^{(0)}(R, 0)H(\mathbf{x})U^{(0)^{-1}}(R, 0) = H(\mathbf{Rx}), \quad R \in O(3). \tag{16}$$

It is not difficult to find the most general solution to (12c), for we see that the difference between any two K_i' satisfying it commutes with $\mathbf{P}$. Thus the most general solution is any particular solution plus any operator commuting with $\mathbf{P}$. One can show, by using Eq. (14), that a particular solution is

$$K_i' = \int d^3x x_iH(\mathbf{x}). \tag{17}$$

The particular solution of Eq. (17) is the only one we use (see, however, footnote 8); we do not add onto it any operator commuting with $\mathbf{P}$.

From (16), we see that the K_i' of Eq. (17) satisfies Eq. (12d). This leaves us with the two difficult commutation relations (9h) and (9i) to satisfy.

After subtracting off the free parts of these equations, they read

$$[H^{(0)}, K_i'] + [H', K_i^{(0)}] + [H', K_i'] = 0, \quad (18a)$$

$$[K_i^{(0)}, K_j'] + [K_i', K_j^{(0)}] + [K_i', K_j'] = 0. \quad (18b)$$

All the difficulties in satisfying these equations lie in the terms $[H', K_i']$, $[K_i', K_j']$, which are quadratic in the Hamiltonian density. If we assume that they are both 0 (it is not necessary that they be 0, but they are 0 for the solutions we find), then we must have

$$[H^{(0)}, K_i'] + [H', K_i^{(0)}] = 0, \quad (19a)$$

$$[K_i^{(0)}, K_j'] + [K_i', K_j^{(0)}] = 0. \quad (19b)$$

Equation (19) can be made to hold if we require that $H(x)$ transform as a scalar[8] under the whole of G_4^1 [and not just the subgroup $O(3)$ as in Eq. (16)], i.e.,

$$U^{(0)}(\Lambda, 0)H(x)U^{(0)^{-1}}(\Lambda, 0) = H(\Lambda x), \quad (20)$$

for then

$$[K_i^{(0)}, H(x)] = x_i[H^{(0)}, H(x)] - x_0[P_i, H(x)],$$

[in analogy with the IG $K_i = i(x_i\partial t + t\partial x_i)$ in space time] and thus (setting $x_0 = 0$),

$$[H^{(0)}, K_i'] - [K_i^{(0)}, H'] = \int d^3x[x_i[H^{(0)}, H(x)] - x_i[H^{(0)}, H(x)]] = 0,$$

$$[K_i^{(0)}, K_j] - [K_i^{(0)}, K_j'] = \int d^3x[x_ix_j - x_jx_i][H^{(0)}, H(x)] = 0.$$

Finally, we must consider the terms $[H', K_i']$ and $[K_i', K_j']$ which are quadratic in the Hamiltonian density. Although these terms do not have to be 0, there will be certain parts of them which must vanish for a 4F density. That is, $[H', K_i']$ and $[K_i', K_j']$ will contain terms proportional to $b^{*3}\, d^{*3}$ (for example) and the coefficients on these terms must be 0 if Eqs. (18a) and (18b) are to hold. This argument can be used to find necessary conditions [see Sec. 3, Eqs. (45) and (46)] on the Hamiltonian. These conditions, however, appear to be too complicated to work with, so we drop this general approach and retreat to a somewhat simpler one.

Under the assumptions of Eqs. (15) and (17), we have

$$[H', K_i'] = \int d^3x\, d^3y\, y_i[H(\mathbf{x}), H(\mathbf{y})] = -\tfrac{1}{2}\int d^3x\, d^3y(x_i - y_i)[H(\mathbf{x}), H(\mathbf{y})], \quad (21)$$

$$[K_i', K_j'] = \tfrac{1}{2}\int d^3x\, d^3y(x_iy_j - x_jy_i)[H(\mathbf{x}), H(\mathbf{y})], \quad (22)$$

which suggests that the use of a causal Hamiltonian density might provide a method of obtaining

$$[H', K_i'] = [K_i', K_j'] = 0.$$

We therefore assume that Eq. (6b) holds and obtain for $[H', K_i']$

$$[H', K_i'] = -\tfrac{1}{2}\int d^3x\, d^3y(x_i - y_i) \times (a + b_j\partial_j + c_{j_1j_2}\partial_{j_1j_2} + d_{j_1j_2j_3}\partial_{j_1j_2j_3} + \cdots)\delta^3(\mathbf{x}, \mathbf{y}),$$

where, as in Eq. (6b), a, b, $\cdots$ are functions of $\mathbf{x}$, $\mathbf{y}$ and the partials are with respect to $x - y$ holding $x + y$ constant. If we now integrate by parts, we obtain

$$[H', K_i'] = -\tfrac{1}{2}\int d^3x\, d^3y\delta^3(x_i - y_i) \times [(x_i - y_i)(a - \partial_jb_j + \partial_{j_1j_2}b_{j_1j_2} + \cdots) + (-b_i + \partial_{j_2}(c_{ij_2} + c_{j_2i}) - \partial_{j_2j_3} \times (d_{ij_2j_3} + d_{j_2ij_3} + d_{j_2j_3i}) + \cdots)].$$

The first term, proportional to $(x_i - y_i)$, integrates to 0 since $x\delta(x) = 0$, but the rest of the integral is not automatically 0. If we require that $[H', K_i'] = 0$, then this introduces the condition on the coefficients $b, c, \cdots$ that

$$[b_i - \partial_{j_2}(c_{ij_2} + c_{j_2i}) + \partial_{j_2j_3}(d_{ij_2j_3} + d_{j_2ij_3} + d_{j_2j_3i}) + \cdots]_{\mathbf{x}=\mathbf{y}} = 0. \quad (23)$$

We can deal with $[K_i', K_j']$ in a similar manner and find that Eq. (23) also guarantees that it is 0.

Thus we have found sufficient conditions on a Hamiltonian density (4F or any other kind) so that it will give a relativistic theory. The conditions are that it transform as a scalar under G_4^1 and that it satisfy Eqs. (6b) and (23). This is similar to the conditions that Dirac obtains for a relativistic theory.

The only densities we have found which satisfy these conditions are those of Eq. (7), which have $b = c = \cdots = 0$.

2. THE POINT FORM

In the point form, we assume $U(\Lambda, 0) = U^{(0)}(\Lambda, 0)$ for all Λ in G_4^1. This implies that the commutation relations (9c), (9g), and (9i) are satisfied. The commutation relations (9b), (9e), (9f), and (9h) are the infinitesimal forms of

$$U^{(0)}(\Lambda, 0)U(0, a)U^{(0)^{-1}}(\Lambda, 0) = U(0, \Lambda a),$$

which states that the P_μ (with $H = P_0$) transform like

[8] We could be slightly more general here by adding a term L_i onto the K_i' of Eq. (17), where $[L_i, \mathbf{P}] = 0$ and $\mathbf{L}$ transforms as K_i in Eq. (10d) under rotations. Equation (18) can then be satisfied provided $[H', K_i'] = [K_i', K_j'] = 0$, $\delta(x_0)[K_i^{(0)} + L_i, H(x)] = \delta(x_0)x_i[H^{(0)}, H(x)]$, and $[L_i, K_j^0] + [K_i^{(0)}, L_j] + [L_i, L_j] = 0$.

a vector under $U^{(0)}(\Lambda, 0)$, i.e.,

$$U^{(0)}(\Lambda, 0)P_\mu U^{(0)^{-1}}(\Lambda, 0) = \Lambda^{-1}_{\mu\nu}P_\nu. \quad (24)$$

We again assume that (11a) holds so that we must have

$$P_\mu = P^{(0)}_\mu + P'_\mu, \quad (25)$$

where the $P^{(0)}_\mu$ are the IG's of $U^{(0)}(0, a)$. Since the $P^{(0)}_\mu$ satisfy Eq. (24), we must have

$$U^{(0)}(\Lambda, 0)P'_\mu U^{(0)^{-1}}(\Lambda, 0) = \Lambda^{-1}_{\mu\nu}P'_\nu. \quad (26)$$

There are, of course, a large number of ways to satisfy Eq. (26). One way is to modify the Hamiltonian of Eq. (43) (with f, g functions of k^2) by replacing the first θ^i by γ_μ and the second by I to obtain P_μ. Another possibility, which gives a Hamiltonian more nearly like that of Eq. (15), is to suppose that

$$P'_0 = \int d^4x x_0 \delta(x^2 - d^2)\theta(x_0)H(x), \quad (27a)$$

$$P'_i = \int d^4x x_i \delta(x^2 - d^2)\theta(x_0)H(x), \quad (27b)$$

where $H(x)$ is a scalar under G^1_4 obeying Eq. (14), d^2 is a positive real number, and $\theta(x_0)$ is the step function which is 0 for $x_0 < 0$ but 1 for $x_0 > 0$. We restrict our attention to interaction momenta of the form of Eq. (27).

The two commutation relations left to satisfy are (9a) and (9d). Since (9a) can be derived from (9d), (9f), (9h), we need only to satisfy (9d). After subtracting $[H^{(0)}, P^{(0)}_j]$, Eq. (9d) reads

$$[H^{(0)}, P'_j] - [P^{(0)}_j, H'] + [H', P'_j] = 0.$$

From Eq. (27), we obtain

$$\begin{aligned}[H^{(0)}, P'_j] - [P^{(0)}_j, H'] &= \int d^4x \theta(x_0)\delta(x^2 - d^2) \\ &\quad \times (x_j[H^{(0)}, H(x)] - x_0[P^{(0)}_j, H(x)]) \\ &= \left[K^{(0)}_j, \int d^4x \theta(x_0)\delta(x^2 - d^2)H(x)\right] = 0,\end{aligned}$$

and so we must evidently have

$$[H', P'_j] = \tfrac{1}{2}\int d^3x\, d^3y \left(\frac{y_j}{y_0} - \frac{x_j}{x_0}\right)[H(x), H(y)] = 0,$$
$$x_0 = (\mathbf{x}^2 + d^2)^{\frac{1}{2}}, \quad y_0 = (\mathbf{y}^2 + d^2)^{\frac{1}{2}}. \quad (28)$$

One might possibly try to do the same sort of analysis for a general causal Hamiltonian density in this case as is done for the instant form in Sec. 3. But the resulting conditions on b, c, $\cdots$ are considerably more complicated than in the instant form, so we will not pursue this possibility.

We see that one solution to Eq. (28) is to have $H(x)$ satisfy Eq. (6) with $b = c = \cdots = 0$. This, coupled with the scalarity of $H(x)$ under G^1_4, gives sufficient conditions for the construction of a Hamiltonian,

$$H' = \int d^3x H(x), \quad x_0 = (\mathbf{x}^2 + d^2)^{\frac{1}{2}}, \quad (29)$$

which gives a relativistic theory in the point form[9] (for a 4F or any other kind of interaction Hamiltonian). As before, the densities of Eq. (7) satisfy these conditions.

3. FOUR-FERMION HAMILTONIANS AND NECESSARY CONDITIONS ON THE HAMILTONIAN IN THE INSTANT FORM

There are two aims in this section. One is to show how to construct densities from particle creation and annihilation operators which transform in a simple manner under $U^{(0)}(\Lambda, a)$ and the other is to give general conditions on the Hamiltonian implied by Eqs. (4).

The interaction Hamiltonian is to be made up of sums of products of creation and annihilation operators for a spin $\frac{1}{2}$ fermion. These are, respectively, for a particle with momentum $\mathbf{p}$ in spin state r ($r = 1, 2$), $b^*_r(\mathbf{p})$, and $b_r(\mathbf{p})$, and for an antiparticle $d^*_r(\mathbf{p})$, $d_r(\mathbf{p})$. The star denotes Hermitian adjoint and b, b^*, d, d^* satisfy the usual anticommutation relations, i.e.,

$$[b, b]_+ = [d, d]_+ = [b, d]_+ = [b, d^*]_+ = 0,$$
$$[b_r(\mathbf{p}), b^*_s(\mathbf{p}')]_+ = [d_r(\mathbf{p}), d^*_s(\mathbf{p}')]_+ = \delta_{rs}\delta(\mathbf{p} - \mathbf{p}'),$$

plus the Hermitian adjoint relations. If we N-order the interaction Hamiltonian, it will consist of a part which is the sum of products of four creation and/or annihilation operators (the 4F part), a part which is the sum of products of two creation and/or annihilation operators, and perhaps a constant. We give only the 4F part in momentum space because that is the only part one uses in finding necessary conditions that the Hamiltonian must satisfy. In both the instant and point forms, it is necessary that we know the transformation properties of the Hamiltonian under the free operators $U^{(0)}(\Lambda, a)$, so we discuss these properties for the creation and annihilation operators before giving the general form for the 4F part of the Hamiltonian.

The transformation properties of a one particle state (without interaction) under $U^{(0)}$ are

$$U^{(0)}(\Lambda, a)\,|\mathbf{p}, r\rangle = e^{i\Lambda p\cdot a}|E(\Lambda\mathbf{p})/E(\mathbf{p})|^{\frac{1}{2}} \times D_{rs}(R(\mathbf{p}, \Lambda)\,|\Lambda\mathbf{p}, s\rangle,$$

[9] We note that we do not need to assume Eq. (11a) for H'_j; we can drop the free part and still obtain a relativistic theory.

where $p_0 = E(p) = (\mathbf{p}^2 + m^2)^{\frac{1}{2}}$ (with m the mass of the base particle), $R(\mathbf{p}, \Lambda)$ is a rotation belonging to $O(3)$, and $D_{rs}(R)$ is a matrix representation of R in SU_2. Since $|\mathbf{p}, r\rangle = b_r^*(\mathbf{p})|0\rangle$, we evidently have

$$U^{(0)}(\Lambda, a)b_r^*(p)U^{(0)^{-1}}(\Lambda, a) = e^{i\Lambda p\cdot a}|E(\Lambda\mathbf{p})/E(\mathbf{p})|^{\frac{1}{2}}D_{rs}(R(\mathbf{p}, \Lambda))b_s^*(\Lambda\mathbf{p}). \quad (30a)$$

The transformation properties of $b_r(\mathbf{p})$ are obtained by taking the Hermitian adjoint of (30a):

$$U^{(0)}(\Lambda, a)b_r(\mathbf{p})U^{(0)^{-1}}(\Lambda, a) = e^{-i\Lambda p\cdot a}|E(\Lambda\mathbf{p})/E(\mathbf{p})|^{\frac{1}{2}}\bar{D}_{rs}(R(\mathbf{p}, \Lambda))b_s(\Lambda\mathbf{p}), \quad (30b)$$

where the bar over D_{rs} stands for complex conjugate. The transformation properties of the antiparticle creation and annihilation operators are, similarly,

$$U^{(0)}(\Lambda, a)d_r^*(\mathbf{p})U^{(0)^{-1}}(\Lambda, a) = e^{i\Lambda p\cdot a}\bar{D}'_{rs}(R(\mathbf{p}, \Lambda))d_s^*(\Lambda\mathbf{p}), \quad (30c)$$

$$U^{(0)}(\Lambda, a)d_r(\mathbf{p})U^{(0)^{-1}}(\Lambda, a) = e^{-i\Lambda p\cdot a}\bar{D}'_{rs}(R(\mathbf{p}, \Lambda))d_s^*(\Lambda\mathbf{p}), \quad (30d)$$

where $D'_{rs}(R)$ is another representation of R in SU_2.

The quantities $D(R(\mathbf{p}, \Lambda))$, $D'(R(\mathbf{p}, \Lambda))$ depend upon $\mathbf{p}$ and Λ in a complicated way, and if we had to deal directly with them, life would be difficult. Fortunately, however, it is possible to find linear combinations,

$$\begin{aligned}\psi_\alpha^{(+)}(\mathbf{p}) &= U_{\alpha,r}^{(+)}(\mathbf{p})b_r(\mathbf{p}),\\ \psi_\alpha^{(-)}(\mathbf{p}) &= U_{\alpha,r}^{(-)}(-\mathbf{p})d_r^*(-\mathbf{p}),\\ \tilde{\psi}_\alpha^{(+)}(\mathbf{p}) &= \tilde{U}_{\alpha,r}^{(+)}(\mathbf{p})b_r^*(\mathbf{p}),\\ \tilde{\psi}_\alpha^{(-)}(\mathbf{p}) &= \tilde{U}_{\alpha,r}^{(-)}(-\mathbf{p})d_r(-\mathbf{p}),\end{aligned} \quad (31)$$

of the creation and annihilation operators which transform in a simple way, and we therefore construct our Hamiltonian from the ψ's rather than directly from the creation and annihilation operators. The U's in Eq. (31) are column vectors of length 4 (the component being designated by α) which satisfy the Dirac equation

$$\begin{aligned}(\gamma_\mu p_\mu - \epsilon m)U_r^{(\epsilon)}(\mathbf{p}) &= 0,\\ \tilde{U}_r^{(\epsilon)}(\mathbf{p})(\gamma_\mu p_\mu - \epsilon m) &= 0,\end{aligned} \quad \epsilon = \pm 1,$$

where $\tilde{U}_r^{(\epsilon)}(\mathbf{p}) = U_r^{(\epsilon)*}(\mathbf{p})$. They also satisfy the conditions

$$\epsilon U_{\alpha,r}^{(\epsilon)}(\mathbf{p})U_{\beta,r}^{(\epsilon)}(\mathbf{p}) = [(m + \epsilon p_\mu\gamma_\mu)/2m]_{\alpha\beta}, \quad (32)$$

$$\tilde{U}_{\alpha,r}^{(\epsilon)}(\mathbf{p})U_{\alpha,s}^{(\epsilon')}(\mathbf{p}) = \epsilon\delta_{\epsilon\epsilon'}\delta_{rs}. \quad (33)$$

If the proper choice is made for the D, D', then the transformation properties of the ψ's are [with $U^{(0)} = U^{(0)}(\Lambda, a)$]

$$\begin{aligned}U^{(0)}\psi_\alpha^{(+)}(\mathbf{p})U^{(0)^{-1}} &= [E(\Lambda\mathbf{p})/E(\mathbf{p})]^{\frac{1}{2}}S_{\alpha\beta}^{-1}(\Lambda)\psi_\beta^{(+)}(\Lambda\mathbf{p})e^{-i\Lambda p\cdot a},\\ U^{(0)}\psi_\alpha^{(-)}(\mathbf{p})U^{(0)^{-1}} &= [E(\Lambda\mathbf{p})/E(\mathbf{p})]^{\frac{1}{2}}S_{\alpha\beta}^{-1}(\Lambda)\psi_\beta^{(-)}(\Lambda\mathbf{p})e^{i(\Lambda p)_0 a_0 + i\Lambda p\cdot a},\\ U^{(0)}\tilde{\psi}_\alpha^{(+)}(\mathbf{p})U^{(0)^{-1}} &= [E(\Lambda\mathbf{p})/E(\mathbf{p})]^{\frac{1}{2}}\tilde{\psi}_\beta^{(+)}(\Lambda\mathbf{p})S_{\beta\alpha}(\Lambda)e^{i\Lambda p\cdot a},\\ U^{(0)}\tilde{\psi}_\alpha^{(-)}(\mathbf{p})U^{(0)^{-1}} &= [E(\Lambda\mathbf{p})/E(\mathbf{p})]^{\frac{1}{2}}\tilde{\psi}_\beta^{(-)}(\Lambda\mathbf{p})S_{\beta\alpha}(\Lambda)e^{-i(\Lambda p)_0 a_0 - i\Lambda p\cdot a},\end{aligned} \quad (34)$$

where $S(\Lambda)$ is a 4×4 matrix satisfying

$$S^{-1}(\Lambda)\gamma_\mu S(\Lambda) = \Lambda_{\mu\nu}\gamma_\nu. \quad (35)$$

We can now build our Hamiltonian out of products of ψ's and $\tilde{\psi}$'s. The 4F part in both the instant and point form is

$$H'_{4F} = \int d^4x f(x)H(x), \quad (36)$$

$$H(x) = \sum_{i=1}^{9} H(x)^{(i)}, \quad (37)$$

with

$$\begin{aligned}H(x)^{(1)} &= \int d\tau_{1234}\tilde{\psi}_1^{(+)}(1)\tilde{\psi}_2^{(+)}(2)\psi_3^{(-)}(+3)\psi_4^{(-)}\\ &\quad\times(+4)h_{\alpha_{1234}}^{(1)}(1234)e^{ix\cdot(1+2-3-4)-ix_0\cdot(1+2+3+4)},\\ H(x)^{(2)} &= \int d\tau_{1234}\tilde{\psi}_1^{(+)}(1)\tilde{\psi}_2^{(+)}(2)\psi_3^{(-)}(3)\psi_4^{(+)}\\ &\quad\times(4)h_{\alpha_{1234}}^{(2)}(1234)e^{ix\cdot(1+2-3-4)-ix_0(1+2+3-4)},\\ H(x)^{(3)} &= \int d\tau_{1234}\tilde{\psi}_1^{(+)}(1)\psi_2^{(-)}(2)\psi_3^{(-)}(3)\tilde{\psi}_4^{(-)}\\ &\quad\times(4)h_{\alpha_{1234}}^{(3)}(1234)e^{ix\cdot(1-2-3+4)-ix_0(1+2+3-4)},\\ H(x)^{(4)} &= \int d\tau_{1234}\tilde{\psi}_1^{(+)}(1)\tilde{\psi}_2^{(+)}(2)\psi_3^{(+)}(3)\psi_4^{(+)},\\ &\quad\times(4)h_{\alpha_{1234}}^{(4)}(1234)e^{ix\cdot(1+2-3-4)-ix_0(1+2-3-4)},\\ H(x)^{(5} &= \int d\tau_{1234}\tilde{\psi}_1^{(+)}(1)\psi_2^{(-)}(2)\tilde{\psi}_3^{(-)}(3)\psi_4^{(+)}\\ &\quad\times(4)h_{\alpha_{1234}}^{(5)}(1234)e^{ix\cdot(1-2+3-4)-ix_0(1+2-3-4)},\\ H(x)^{(6)} &= \int d\tau_{1234}\psi_1^{(-)}(1)\psi_2^{(-)}(2)\tilde{\psi}_3^{(-)}(3)\tilde{\psi}_4^{(-)}\\ &\quad\times(4)h_{\alpha_{1234}}^{(6)}(1234)e^{ix\cdot(-1-2+3+4)-ix_0(1+2-3-4)},\\ H(x)^{(7)} &= H(x)^{(3)*},\quad H(x)^{(8)} = H(x)^{(2)*},\\ H(x)^{(9)} &= H(x)^{(1)*},\end{aligned} \quad (38)$$

where we have made use of the Hermiticity of the Hamiltonian to obtain the last three equations. The

somewhat abbreviated notation used in Eqs. (38) is

$$d\tau_1 = d^3p_1[m/E(\mathbf{p}_1)]^{\frac{1}{2}}, \quad d\tau_{1234} = d\tau_1\, d\tau_2\, d\tau_3\, d\tau_4,$$
$$\psi_1(1) = \psi_{\alpha_1}(\mathbf{p}_1),$$
$$h_{\alpha_{1234}}(1234) = h_{\alpha_1\alpha_2\alpha_3\alpha_4}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4),$$
$$\exp\{i[\mathbf{x}\cdot(1+2-3-4) - x_0(1+2+3+4)]\}$$
$$= \exp\{i[x\cdot(\mathbf{p}_1+\mathbf{p}_2-\mathbf{p}_3-\mathbf{p}_4) - x_0(p_{10}+p_{20}+p_{30}+p_{40})]\}. \quad (39)$$

We note that the $h^{(i)}$ can be assumed to be asymmetric in certain variables. If we look at $H(x)^{(1)}$, for example, interchange $\tilde{\psi}_1^{(+)}(1)$ with $\tilde{\psi}_2^{(+)}(2)$, and then exchange variables 1 and 2, we find that one can substitute $\frac{1}{2}(h_{\alpha_{1234}}(1234) - h_{\alpha_{2134}}(2134))$ for $h_{\alpha_{1234}}(1234)$. Actually, in $H(x)^{(1)}$, one can substitute

$$\tfrac{1}{4}(h_{1234} - h_{2134} - h_{1243} + h_{2143})$$

for h_{1234}.

We can now show how the densities transform under $U^{(0)}(\Lambda, a)$. It follows from Eqs. (34) and (38) that Eq. (14) is satisfied, i.e.,

$$U^{(0)}(I, a)H(x)^{(i)}U^{(0)}(I, a) = H^{(i)}(x - a) \quad (40)$$

and

$$U^{(0)}(\Lambda, 0)H(x)^{(i)}U^{(0)^{-1}}(\Lambda, 0) = H'^{(i)}(\Lambda x), \quad (41a)$$

where $H'^{(i)}(x)$ has the same form as $H^{(i)}(x)$, except $h^{(i)}$ is changed to $h'^{(i)}$. If we look at the $i = 1$ case, for example,

$$h'^{(1)}_{\alpha_{1234}}(\mathbf{q}_1, \cdots, \mathbf{q}_4) = S_{\alpha_1\beta_1}(\Lambda)S_{\alpha_2\beta_2}(\Lambda)S^{-1}_{\beta_3\alpha_3}(\Lambda)S^{-1}_{\beta_4\alpha_4}(\Lambda)$$
$$h^{(1)}_{\beta_{1234}}(\Lambda^{-1}q_1, \cdots, \Lambda^{-1}q_4). \quad (41b)$$

The equations for the other $h'^{(i)}$ are the same except for a permutation of the α, β indices. Equations (41) were derived by changing Λp_i to q_i in

$$U^{(0)}(\Lambda, 0)H^{(i)}(x)U^{(0)^{-1}}(\Lambda, 0).$$

If $H^{(i)}(x)$ is to transform as a scalar under G_4^1, i.e., if

$$U^{(0)}(\Lambda, 0)H^{(i)}(x)U^{(0)^{-1}}(\Lambda, 0) = H^{(i)}(\Lambda x), \quad (42a)$$

then we must have

$$h'^{(i)}_{\alpha_{1234}}(1234) = h^{(i)}_{\alpha_{1234}}(1234). \quad (42b)$$

It would perhaps be useful to correlate certain forms of the Hamiltonian in the more familiar configuration space with their corresponding $h^{(i)}$ in Eq. (38). As a rather general example, we suppose that

$$H(x) = \int d^4y J^i(x + y)F(y)J^i(x - y), \quad (43a)$$

where

$$J^i(z) = \int d^4w\psi(z + w)\,\theta^i\psi(z - w)G(w), \quad (43b)$$

$$F(y) = \int d^4k e^{-ik\cdot y}f(k), \quad (43c)$$

$$G(w) = \int d^4k e^{-ik\cdot w}g(k), \quad (43d)$$

$$\theta^i = aI + ib\gamma_5,\ c\gamma_\mu + i\,d\gamma_5\gamma_\mu,\ \theta_{\mu\nu},$$
$$a, b, c, d \text{ real}, \quad (43e)$$

$$\psi(x) = \int d\tau[\psi^{(+)}(\mathbf{p})e^{-ip\cdot x} + \psi^{(-)}(-\mathbf{p})e^{ip\cdot x}], \quad (43f)$$

$$\tilde{\psi}(x) = \psi^*(x)\gamma_0$$
$$= \int d\tau[\tilde{\psi}^{(+)}(\mathbf{p})e^{ip\cdot x} + \tilde{\psi}^{(-)}(-\mathbf{p})e^{-ip\cdot x}]. \quad (43g)$$

[All of these densities are scalars under G_4^1, provided the f and g of Eqs. (43c) and (43d), respectively, are functions of k^2 only.] We then find that

$$h^{(1)}_{\alpha_{1234}} = -\theta^i_{\alpha_1\alpha_3}\theta^i_{\alpha_2\alpha_4}g(k_1)g(k_2)f(k_3), \quad (44a)$$

with

$$k_{10} = p_{10} - p_{30}, \qquad \mathbf{k}_1 = \mathbf{p}_1 + \mathbf{p}_3,$$
$$k_{20} = p_{20} - p_{40}, \qquad \mathbf{k}_2 = \mathbf{p}_2 + \mathbf{p}_4,$$
$$k_{30} = p_{10} - p_{20} + p_{30} - p_{40}, \qquad \mathbf{k}_3 = \mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3 + \mathbf{p}_4$$

$$h^{(2)}_{\alpha_{1234}}(1234) = -\theta^i_{\alpha_1\alpha_3}\theta^i_{\alpha_2\alpha_4}g(k_1)g(k_2)f(k_3)$$
$$+ \theta^i_{\alpha_1\alpha_4}\theta^i_{\alpha_2\alpha_3}g(k'_1)g(k'_2)f(k'_3) \quad (44b)$$

with

$$k_{10} = p_{10} - p_{30}, \qquad \mathbf{k}_1 = \mathbf{p}_1 + \mathbf{p}_3,$$
$$k_{20} = p_{20} + p_{40}, \qquad \mathbf{k}_2 = \mathbf{p}_2 + \mathbf{p}_4,$$
$$k_{30} = p_{10} - p_{20} + p_{30} + p_{40}, \qquad \mathbf{k}_3 = \mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3 + \mathbf{p}_4,$$
$$k'_{10} = p_{10} + p_{40} \qquad \mathbf{k}'_1 = \mathbf{p}_1 + \mathbf{p}_4,$$
$$k'_{20} = p_{20} - p_{30}, \qquad \mathbf{k}'_2 = \mathbf{p}_2 + \mathbf{p}_3,$$
$$k'_{30} = p_{10} - p_{20} - p_{30} - p_{40}, \qquad \mathbf{k}'_3 = \mathbf{p}_1 - \mathbf{p}_2 + \mathbf{p}_3 - \mathbf{p}_4.$$

We also deal with Hamiltonian densities made up of products of $\psi(x)$, $\tilde{\psi}(x)$ (Eq. 43f, g), and their derivatives. The transformation properties of a few selected operators are

$$U^{(0)}(\Lambda)\psi(x)U^{(0)^{-1}}(\Lambda) = S^{-1}(\Lambda)\psi(\Lambda x),$$
$$U^{(0)}(\Lambda)\tilde{\psi}(x)U^{(0)^{-1}}(\Lambda) = \tilde{\psi}(\Lambda x)S(\Lambda),$$
$$U^{(0)}(\Lambda)\partial_{x_\mu}\psi(x)U^{(0)^{-1}}(\Lambda) = \Lambda^{-1}_{\mu\nu}\partial_{y_\nu}\psi(y)y_\mu = \Lambda_{\mu\nu}x_\nu,$$
$$U^{(0)}(\Lambda)\tilde{\psi}(x)\gamma_\mu\psi(x)U^{(0)^{-1}}(\Lambda) = \Lambda^{-1}_{\mu\nu}\tilde{\psi}(x)\gamma_\nu\psi(x).$$

We see from these that the densities of Eq. (7) are scalars under G_4^1. Two more scalar densities are

$$\tilde{\psi}(x)\partial_\mu\psi(x)\tilde{\psi}(x)\gamma_\mu\psi(x) + \text{h.a.},$$
$$\tilde{\psi}(x)\partial_\mu\psi(x)\tilde{\psi}(x)\partial_\mu\psi(x) + \text{h.a.}$$

The reason it is useful to work with densities constructed from products of ψ, $\bar{\psi}$ and their derivatives is that it is a simple matter to construct causal densities from them. In fact, any finite linear combination of the densities

$$D_{\alpha_{1234}\beta_{1234}}(x) = (\theta_{\beta_1}\bar{\psi}_{\alpha_1}(x))(\theta_{\beta_1}\psi_{\alpha_2}(x)) \times (\theta_{\beta_3}\bar{\psi}_{\alpha_3}(x))(\theta_{\beta_4}x_{\alpha_4}(x))$$

is causal, where θ_β is a finite product of derivatives.

Now we turn to the second goal of this section, namely, to find conditions on the $h^{(i)}$ of Eq. (38) implied by Eqs. (18a, b). We find conditions only on $h^{(1)}$, $h^{(2)}$, $h^{(3)}$, but the same methods can be used to find conditions on the other $h^{(i)}$. The quantity $[H(\mathbf{x}), H(\mathbf{y})]$ occurs in both $[H', K_i']$ and $[K_i', K_j']$. The commutator of the densities will contain terms proportional to $b^{*3}\, d^{*3}$ and, after a lengthy calculation, we find the sum of all such terms to be [using the notation of Eqs. (38) and (39)]

$$M_{33}(\mathbf{x}, \mathbf{y}) = \int d\tau_{123567} b_1^*(1) b_2^*(2) b_3^*(3) \times d_5^*(-5)\, d_6^*(-6)\, d_7^*(-7) \int d^3p_4 f_{\tau_{123567}}(123567; 4) \times [e^{ix\cdot(2+3-7-4)+iy\cdot(1+4-5-6)} - e^{ix\cdot(1+4-5-6)+iy\cdot(2+3-7-4)}],$$

where $b_1^*(1) = b_{r_1}^*(\mathbf{p}_1)$ and f is the function

$$f = \tilde{U}_{1,1}^{(+)}(1)\tilde{U}_{2,2}^{(+)}(2)\tilde{U}_{3,3}^{(+)}(3)U_{5,5}^{(-)}(5)U_{6,6}^{(-)}(-6)U_{7,7}^{(-)}(-7) \times \left[h_{1456}^{(1)}(1456)h_{2378}^{(2)}(2378)\left[\frac{m + p_{40}\gamma_0 - \mathbf{p}_4\cdot\mathbf{y}}{2p_{40}}\right]_{84} + h_{1654}^{(3)}(1654)h_{2378}^{(1)}(3278)\left[\frac{-m + p_{40}\gamma_0 - \mathbf{p}_4\cdot\mathbf{y}}{2p_{40}}\right]_{84}\right]$$

with $U_{1,1}(1) = U_{\alpha_1,r_1}(\mathbf{p}_1)$. If Eq. (18a) is to hold, then we must have

$$N_{33} = \int d^3x\, d^3y\, y_i M_{33}(\mathbf{x}, \mathbf{y}) = 0.$$

If we do the integration on $\mathbf{x}$ and then on $\mathbf{p}_4$, we obtain (to within multiples of i and 2π)

$$N_{33} = \int d^3y y_i \int d\tau_{123567} b_1^*(1)\cdots d_7^*(7) e^{iy\cdot(1+2+3-5-6-7)} \times [f(4 = 2 + 3 - 7) - f(4 = 5 + 6 - 1)],$$

where $f(4 = 2 + 3 - 7)$ means f evaluated at $\mathbf{p}_4 = \mathbf{p}_2 + \mathbf{p}_3 - \mathbf{p}_7$. Since N_{33} is to be 0, it must have 0 matrix element between all possible states and thus between $|0\rangle$ and

$$\langle 0| \int d^3p_{123567} b_1(1) b_2(2) b_3(3) \times d_5(-5)\, d_6(-6)\, d_7(-7) g_{r_{123567}}(123567)$$

in particular, where g is an arbitrary function. This implies that

$$\int d^3y y_i \int d\tau_{123567} g e^{iy\cdot(1+2+3-5-6-7)} A(123; 567) [f(4 = 2 + 3 - 7) - f(4 = 5 + 6 - 1)] = 0,$$

where $A(123; 567)$ indicates that the quantity in brackets is to be antisymmetric in the variables 1, 2, 3 and separately in the variables 5, 6, 7. Evidently the only way this can hold is if

$$A(123; 567)[f(4 = 2 + 3 - 7) - f(4 = 5 + 6 - 1)] = 0. \quad (45)$$

Equation (42) is, therefore, a necessary condition on $h^{(1)}$, $h^{(2)}$, $h^{(3)}$, implied by Eq. (18a).

A similar necessary condition imposed by Eq. (18b) can be found in an analogous manner. The only difference is that, when the integration on $\mathbf{x}$ is done, one obtains a derivative, with respect to $\mathbf{p}_4$, of a δ function rather than just a δ function. The necessary condition is

$$A(123; 567) \times \left[\frac{\partial f(4 = 2 + 3 - 7)}{\partial_{p_4}} - \frac{\partial f(4 = 5 + 6 - 1)}{\partial_{p_4}}\right] = 0. \quad (46)$$

Equations (42) and (43) are only one of a set of ten similar conditions on the $h^{(i)}$. The other nine arise because the coefficients on $b^{*3}\, d^{*2}b$, $b^{*2}\, d^{*3}\, d$, etc., must also be 0. The author has not been able to obtain any simple conditions on the density which these equations imply, but would conjecture that they imply the density must be causal.

4. SUMMARY

In summary, we point out the various possible methods which have been suggested for satisfying the commutation relations of IHLG. The first point one must decide upon is what subgroup of IHLG to leave as free operators. If one would like a causal theory, the most logical choice is one which leaves an initial surface invariant, provided any two points x and y on the initial surface satisfy $(x - y)^2 \leq 0$. The two initial surfaces we have worked with are $t = \text{const}$ and $x^2 = d^2$. (Another possibility suggested by Dirac[2] is the surface $t - x_3 = \text{const}$, but we have not examined that case.)

Once one has decided on what subgroup to leave as free operators, there are still a large number of possibilities. Let us examine the instant form first.

There, one is forced to write the interaction Hamiltonian as

$$H' = \int d^4x\delta(x_0)H(x) = \int d^3xH(\mathbf{x}),$$

and $H(\mathbf{x})$ must transform as a scalar under rotations in space. Then for the IG of velocity boosts, one has the general formula

$$K_i = K_i^{(0)} + K_i' + L_i,$$
$$K_i' = \int d^3x x_i H(\mathbf{x}),$$
$$U^{(0)}(R, 0)L_i U^{(0)^{-1}}(R, 0) = R_{ij}^{-1}L_j,$$

where R belongs to $O(3)$. The least complicated method of proceeding from this point is to assume $\mathbf{L} = 0$, so the burden of satisfying the commutation relations falls entirely on $H(\mathbf{x})$.

If one assumes a 4F Hamiltonian density, then the equations

$$[H', K_i^{(0)}] + [H^{(0)}, K_i'] + [H', K_i'] = 0,$$
$$[K_i', K_j^{(0)}] + [K_i^{(0)}, K_j'] + [K_i', K_j'] = 0,$$

imply certain conditions on it. They are too complicated to deal with, however, and one is thus forced to fall back on causality. The above equations then give the condition (23) on the commutator of the densities. This condition, plus that of the scalarity of $H(x)$ under the free G_4^1, is sufficient for a relativistic theory.

In the point form, H' must be one of four quantities which transforms like a vector under the free G_4^1. This gives a very wide possibility of choices for H'. The only one we examine, however, is the one in closest analogy to the instant form, namely

$$H' = \int d^4x x_0\theta(x_0)\delta(x^2 - d^2)H(x).$$

If $[H(x, \mathbf{x}), H(x, \mathbf{y})]$ is proportional to $\delta^{(3)}(\mathbf{x}, \mathbf{y})$, then this H' yields a relativistic theory.

It would thus appear that the only simple way to have a relativistic theory is to have a causal Hamiltonian density.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Some Theorems Concerning the Phase Problem of Coherence Theory*

D. DIALETIS
Department of Physics and Astronomy, University of Rochester, Rochester, New York

(Received 4 August 1966)

The spectral density of a fluctuating light beam may be determined from the knowledge of both the modulus and the phase of the complex degree of self-coherence $\gamma(\tau)$ of the beam. The phase itself may be determined from the modulus and from the location of the zeros of the analytic continuation of $\gamma(\tau)$ in the lower half of the complex τ plane. In the present paper results of an investigation are presented which show that the determination of the zeros is equivalent to the solution of a certain inhomogeneous eigenvalue problem of the Sturm–Liouville type on a semi-infinite frequency range. This eigenvalue problem is found to be equivalent to a certain stability problem in mechanics. Although no general technique for the solution of this type of an eigenvalue problem appears to be known, the new formulation may be used to determine spectral profiles for which the associated degree of self-coherence has zeros at prescribed points in the complex τ plane. Some illustrative examples are given.

1. INTRODUCTION

IT is known[1,2] that the determination of the spectral density $g(\omega)$ $(0 \leq \omega < \infty)$ of a fluctuating light beam is possible from the knowledge of the absolute value of the complex correlation function $\gamma(\tau)$ $(-\infty < \tau < \infty)$ and the knowledge of the zeros of the analytic continuation of γ in the lower half of the complex τ plane.

In the following we prove that the determination of the zeros is equivalent to the solution of a certain eigenvalue problem of the Sturm–Liouville type.

It is well-known that in terms of the spectral density

* This work was supported by the Air Force Cambridge Research Laboratories. A preliminary account of some of the main results was published in a note by D. Dialetis and E. Wolf, Nuovo Cimento, 47, 113 (1967).

[1] E. Wolf, Proc. Phys. Soc. (London) **80**, 1269 (1962).

[2] H. M. Nussenzveig, J. Math. Phys. **8**, 561 (1967).

There, one is forced to write the interaction Hamiltonian as

$$H' = \int d^4x\delta(x_0)H(x) = \int d^3xH(\mathbf{x}),$$

and $H(\mathbf{x})$ must transform as a scalar under rotations in space. Then for the IG of velocity boosts, one has the general formula

$$K_i = K_i^{(0)} + K_i' + L_i,$$
$$K_i' = \int d^3x x_i H(\mathbf{x}),$$
$$U^{(0)}(R, 0)L_i U^{(0)^{-1}}(R, 0) = R_{ij}^{-1}L_j,$$

where R belongs to $O(3)$. The least complicated method of proceeding from this point is to assume $\mathbf{L} = 0$, so the burden of satisfying the commutation relations falls entirely on $H(\mathbf{x})$.

If one assumes a 4F Hamiltonian density, then the equations

$$[H', K_i^{(0)}] + [H^{(0)}, K_i'] + [H', K_i'] = 0,$$
$$[K_i', K_j^{(0)}] + [K_i^{(0)}, K_j'] + [K_i', K_j'] = 0,$$

imply certain conditions on it. They are too complicated to deal with, however, and one is thus forced to fall back on causality. The above equations then give the condition (23) on the commutator of the densities. This condition, plus that of the scalarity of $H(x)$ under the free G_4^1, is sufficient for a relativistic theory.

In the point form, H' must be one of four quantities which transforms like a vector under the free G_4^1. This gives a very wide possibility of choices for H'. The only one we examine, however, is the one in closest analogy to the instant form, namely

$$H' = \int d^4x x_0\theta(x_0)\delta(x^2 - d^2)H(x).$$

If $[H(x, \mathbf{x}), H(x, \mathbf{y})]$ is proportional to $\delta^{(3)}(\mathbf{x}, \mathbf{y})$, then this H' yields a relativistic theory.

It would thus appear that the only simple way to have a relativistic theory is to have a causal Hamiltonian density.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Some Theorems Concerning the Phase Problem of Coherence Theory*

D. DIALETIS
Department of Physics and Astronomy, University of Rochester, Rochester, New York

(Received 4 August 1966)

The spectral density of a fluctuating light beam may be determined from the knowledge of both the modulus and the phase of the complex degree of self-coherence $\gamma(\tau)$ of the beam. The phase itself may be determined from the modulus and from the location of the zeros of the analytic continuation of $\gamma(\tau)$ in the lower half of the complex τ plane. In the present paper results of an investigation are presented which show that the determination of the zeros is equivalent to the solution of a certain inhomogeneous eigenvalue problem of the Sturm–Liouville type on a semi-infinite frequency range. This eigenvalue problem is found to be equivalent to a certain stability problem in mechanics. Although no general technique for the solution of this type of an eigenvalue problem appears to be known, the new formulation may be used to determine spectral profiles for which the associated degree of self-coherence has zeros at prescribed points in the complex τ plane. Some illustrative examples are given.

1. INTRODUCTION

IT is known[1,2] that the determination of the spectral density $g(\omega)$ $(0 \leq \omega < \infty)$ of a fluctuating light beam is possible from the knowledge of the absolute value of the complex correlation function $\gamma(\tau)$ $(-\infty < \tau < \infty)$ and the knowledge of the zeros of the analytic continuation of γ in the lower half of the complex τ plane.

In the following we prove that the determination of the zeros is equivalent to the solution of a certain eigenvalue problem of the Sturm–Liouville type.

It is well-known that in terms of the spectral density

* This work was supported by the Air Force Cambridge Research Laboratories. A preliminary account of some of the main results was published in a note by D. Dialetis and E. Wolf, Nuovo Cimento, 47, 113 (1967).

[1] E. Wolf, Proc. Phys. Soc. (London) **80**, 1269 (1962).

[2] H. M. Nussenzveig, J. Math. Phys. **8**, 561 (1967).

function $g(\omega)$ the correlation function $\gamma(\tau)$ is given by[3]

$$\gamma(\tau) = \int_0^\infty g(\omega') e^{-i\tau\omega'} \, d\omega', \tag{1.1}$$

where $\operatorname{Im} \tau \leq 0$.

The spectral density function $g(\omega)$ is real, nonnegative, and its integral is normalized to unity, i.e.,

(a) $$g(\omega) = g^*(\omega), \tag{1.2}$$

(b) $$g(\omega) \geq 0 \quad \text{for} \quad \omega > 0, \tag{1.3}$$

(c) $$\int_0^\infty g(\omega') \, d\omega' = 1. \tag{1.4}$$

The correlation function has the following properties:

(a) $$\gamma(0) = 1. \tag{1.5}$$

(b) It is analytic and regular in the lower half of the complex τ plane[4] ($\operatorname{Im} \tau < 0$).

(c) $$\gamma(-\tau^*) = \gamma^*(\tau), \tag{1.6}$$

where τ is real or complex with $\operatorname{Im} \tau < 0$.

(d) It does not vanish for any value of τ on the negative imaginary axis, i.e., for any $\tau = -i\beta$, where $\beta > 0$.

Because of the property (c) satisfied by the correlation function $\gamma(\tau)$ we conclude that if

$$\gamma(\tau_0) = 0, \tag{1.7a}$$

then also

$$\gamma(-\tau_0^*) = 0. \tag{1.7b}$$

We are interested only in the zeros of the correlation function in the lower complex τ plane and on the real axis. We see that, because of property (b), they occur in pairs and, because of property (d), there are no zeros on the imaginary axis of the τ plane. Hence we may restrict our discussion to values of $\tau_0 = \alpha - i\beta$, which lie in the fourth quadrant and on the positive real axis of the complex τ plane; i.e., $\alpha > 0$, $\beta \geq 0$, and for which $\gamma(\tau_0) = 0$, $\gamma(-\tau_0^*) = 0$. We confine our attention to this quadrant.

From Eq. (1.1) we see that once the spectral density $g(\omega)$ is defined for $\omega \geq 0$, the correlation function $\gamma(\tau)$ is completely determined and so are the values τ_0 (with $\operatorname{Re} \tau_0 > 0$, $\operatorname{Im} \tau_0 \leq 0$) for which $\gamma(\tau)$ vanishes.

2. THEOREM CONCERNING THE ZEROS OF THE CORRELATION FUNCTION

We prove that the following statements are equivalent:

(i) The correlation function $\gamma(\tau)$ vanishes for $\tau = \tau_0 = \alpha - i\beta$, ($\alpha > 0$, $\beta \geq 0$), i.e.

$$\gamma(\tau_0) = 0. \tag{2.1}$$

(ii) $$\lim_{\omega \to \infty} e^{i\tau_0\omega} \gamma(\tau_0, \omega) = 0, \tag{2.2}$$

where

$$\gamma(\tau, \omega) = \int_0^\omega g(\omega') e^{-i\tau\omega'} \, d\omega'. \tag{2.3}$$

(iii) There exists a unique real function $\psi(\omega; \tau_0)$, defined for $0 \leq \omega < \infty$, which satisfies the differential equation

$$\psi''(\omega; \tau_0) - i(\tau_0 - \tau_0^*)\psi'(\omega; \tau_0) + \tau_0\tau_0^* \psi(\omega; \tau_0) = g(\omega) \tag{2.4a}$$

and the boundary conditions

$$\psi(0; \tau_0) = \psi'(0; \tau_0) = 0, \tag{2.4b}$$

$$\psi(\infty; \tau_0) = \psi'(\infty; \tau_0) = 0. \tag{2.4c}$$

Here

$$\psi'(\omega; \tau_0) = \partial\psi(\omega; \tau_0)/\partial\omega, \quad \text{etc.} \tag{2.5}$$

Before proving the theorem, several observations may be of interest:

(a) Statement (i) is equivalent to the equation

$$\lim_{\omega \to \infty} \gamma(\tau_0, \omega) = \gamma(\tau_0) = 0, \tag{2.6}$$

as we see from the definition (2.3) of the function $\gamma(\tau, \omega)$. Also $e^{i\tau_0\omega} = e^{\beta\omega + i\alpha\omega}$, which oscillates as $\omega \to \infty$ and its absolute value increases beyond any bound. (If $\beta = 0$, it oscillates and its absolute value is unity.) From Eqs. (2.6) and (2.2) it may appear that the statement (ii) is stronger than the statement (i), but as we show below, they are actually equivalent.

(b) In Eq. (2.3), let τ be a value for which $\gamma(\tau)$ does not vanish. Then

$$\lim_{\omega \to \infty} \gamma(\tau, \omega) = \gamma(\tau) \neq 0. \tag{2.7}$$

Since $e^{i\tau\omega}$ (with $\operatorname{Re} \tau > 0$, $\operatorname{Im} \tau \leq 0$) does not tend to zero as $\omega \to \infty$, we conclude that the limit in statement (ii) is zero only for values τ_0 for which $\gamma(\tau_0) = 0$ and for these alone.

(c) Using the fact that

$$\gamma^*(\tau, \omega) = \gamma(-\tau^*, \omega) \tag{2.8}$$

[3] A similar relation exists in the quantum theory of coherence between the quantum mechanically defined coherence function and the power spectrum. [Cf. C. L. Mehta and E. Wolf, Phys. Rev. 157, 1188 (1967).] Hence our results apply both within the domain of classical and quantum mechanical theory of optical coherence.

[4] This is true under the assumption that

$$\int_{-\infty}^{\infty} |\gamma(\tau)|^2 \, d\tau < \infty.$$

[Cf. E. C. Titchmarch, *Theory of the Fourier Integral* (Oxford University Press, New York, 1948), 2nd ed., p. 128.]

and taking the complex conjugate of the Eq. (2.2), we obtain

$$\lim_{\omega\to\infty} e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega) = 0. \tag{2.9}$$

Equations (2.2) and (2.9) are equivalent.

(d) The relation between the function $\psi(\omega; \tau_0)$ and the correlation function $\gamma(\tau)$ is

$$\gamma(\tau) = (\tau - \tau_0)(\tau + \tau_0^*)\int_0^\infty [-\psi(\omega'; \tau_0)]e^{-i\tau\omega'}\, d\omega'. \tag{2.10}$$

This relation is established later. We note that the integral on the right-hand side of (2.10) exists because of Eq. (2.4c).

(e) The function $\psi(\omega, \tau_0)$ is uniquely determined by the second-order differential equation (2.4a) and the boundary conditions (2.4b). It satisfies the boundary conditions (2.4c), too, if and only if τ_0 is such that $\gamma(\tau_0) = 0$. This last remark is closely related to the statement (b) [Eq. (2.7)].

We see then that, for each τ_0 for which $\gamma(\tau_0) = 0$, there is an eigenfunction $\psi(\omega, \tau_0)$ and an eigenvalue solution $\tau_0 = \alpha - i\beta$ of the system (2.4a)–(2.4c) of the Sturm–Liouville type.

3. PROOF OF THE THEOREM

First we show that the statement (i) implies the statement (ii).

We have from (1.1) and (2.3)

$$\gamma(\tau) = \gamma(\tau, \omega) + \int_\omega^\infty g(\omega')e^{-i\tau\omega'}\, d\omega', \tag{3.1}$$

so that

$$e^{i\tau\omega}\gamma(\tau) = e^{i\tau\omega}\gamma(\tau, \omega) + \int_\omega^\infty g(\omega')e^{-i\tau(\omega'-\omega)}\, d\omega'. \tag{3.2}$$

Hence, if $\tau_0 = \alpha - i\beta$ $(\alpha > 0, \beta \geq 0)$ is a zero of $\gamma(\tau)$,

$$0 = e^{i\tau_0\omega}\gamma(\tau_0, \omega) + \int_\omega^\infty g(\omega')e^{-i\tau_0(\omega'-\omega)}\, d\omega'. \tag{3.3}$$

Also for ω' such that $\omega' \geq \omega$, we have

$$|e^{-i\tau_0(\omega'-\omega)}| \leq e^{-\beta(\omega'-\omega)} \leq 1, \tag{3.4}$$

and

$$\int_\omega^\infty |g(\omega')|\, d\omega' = \int_\omega^\infty g(\omega')\, d\omega' \xrightarrow[\omega\to\infty]{} 0, \tag{3.5}$$

because of the properties (a), (b), and (c) of $g(\omega)$. [See Eqs. (1.2), (1.3), and (1.4).]

From Eq. (3.3) we obtain then the inequality

$$\begin{aligned} |e^{i\tau_0\omega}\gamma(\tau_0, \omega)| &= \left| -\int_\omega^\infty g(\omega')e^{-i\tau_0(\omega'-\omega)}\, d\omega' \right| \\ &\leq \int_\omega^\infty |g(\omega')|\, |e^{-i\tau_0(\omega'-\omega)}|\, d\omega' \\ &\leq \int_\omega^\infty |g(\omega')|\, d\omega'. \end{aligned} \tag{3.6}$$

In the last inequality we used Eq. (3.4). Finally, using Eq. (3.5), it follows that the last expression in Eq. (3.6) tends to zero as $\omega \to \infty$, so that

$$\lim_{\omega\to\infty} e^{i\tau_0\omega}\gamma(\tau_0, \omega) = 0.$$

Hence the statement (ii) is established.

Now let us prove the converse statement, i.e., the fact that (ii) implies statement (i).

According to (ii), for each $\epsilon > 0$ there is an $\mathcal{N}(\epsilon)$ such that for every $\omega > \mathcal{N}(\epsilon)$

$$|e^{i\tau_0\omega}\gamma(\tau_0, \omega)| < \epsilon. \tag{3.7}$$

Equation (3.7) implies that

$$|\gamma(\tau_0, \omega)|\, |e^{i\tau_0\omega}| < \epsilon,$$

or, since $|e^{i\tau_0\omega}| = e^{\beta\omega}$,

$$|\gamma(\tau_0, \omega)|\, e^{\beta\omega} < \epsilon,$$

i.e.,

$$|\gamma(\tau_0, \omega)| < \epsilon e^{-\beta\omega} \leq \epsilon,$$

because

$$e^{-\beta\omega} \leq 1 \quad \text{for} \quad \beta \geq 0, \quad \omega \geq 0.$$

Hence

$$\gamma(\tau_0) = \lim_{\omega\to\infty} \gamma(\tau_0, \omega) = 0,$$

which is what was to be proved.

This completes the proof that the statements (i) and (ii) are equivalent.

Next we show that the statement (iii) implies the statement (i). [Because of the equivalence of statements (i) and (ii), we conclude that statement (iii) will also imply statement (ii).]

For this purpose we make use of the identities

$$\begin{aligned} &\int_0^\omega \psi'(\omega'; \tau_0)e^{-i\tau\omega'}\, d\omega' \\ &= \psi(\omega; \tau_0)e^{-i\tau\omega} - \psi(0; \tau_0) + i\tau\int_0^\omega \psi(\omega'; \tau_0)e^{-i\tau\omega'}\, d\omega', \end{aligned} \tag{3.8}$$

$$\begin{aligned} &\int_0^\omega \psi''(\omega'; \tau_0)e^{-i\tau\omega'}\, d\omega' \\ &= \psi'(\omega; \tau_0)e^{-i\tau\omega} - \psi'(0; \tau_0) + i\tau[\psi(\omega; \tau_0)e^{-i\tau\omega} \\ &\quad - \psi(0; \tau_0)] - \tau^2\int_0^\omega \psi(\omega'; \tau_0)e^{-i\tau\omega'}\, d\omega', \end{aligned} \tag{3.9}$$

from which we obtain the identity

$$\begin{aligned} &\int_0^\omega [\psi''(\omega'; \tau_0) - i(\tau_0 - \tau_0^*)\psi'(\omega'; \tau_0) \\ &\qquad + \tau_0\tau_0^*\psi(\omega'; \tau_0)]e^{-i\tau\omega'}\, d\omega' \\ &\qquad - [\psi'(\omega, \tau_0) + i\{\tau - (\tau_0 - \tau_0^*)\}\psi(\omega; \tau_0)]e^{-i\tau\omega} \\ &\qquad + \psi'(0; \tau_0) + i\{\tau - (\tau_0 - \tau_0^*)\}\psi(0; \tau_0) \\ &= (\tau - \tau_0)(\tau + \tau_0^*)\int_0^\omega [-\psi(\omega'; \tau_0)]e^{-i\tau\omega'}\, d\omega'. \end{aligned} \tag{3.10}$$

Here $\psi(\omega, \tau_0)$ is the function satisfying Eq. (2.4a)–(2.4c).

First we must show that these identities are meaningful if $\psi(\omega, \tau_0)$ is chosen as stated. The solution of Eq. (2.4a), subject to the boundary conditions (2.4b), is

$$\psi(\omega; \tau_0) = \frac{1}{i(\tau_0 + \tau_0^*)} [e^{i\tau_0\omega}\gamma(\tau_0, \omega) - e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega)] \tag{3.11}$$

defined for every ω $(0 \leq \omega < \infty)$. Because of Eq. (2.8), we see that $\psi(\omega; \tau_0)$ is real. From Eq. (3.11) it follows that

$$\psi'(\omega; \tau_0) = \frac{1}{\tau_0 + \tau_0^*} [\tau_0 e^{i\tau_0\omega}\gamma(\tau_0, \omega) + \tau_0^* e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega)]. \tag{3.12}$$

With the help of these two expressions and from the formula (2.4a) written in the form

$$\psi''(\omega; \tau_0) - g(\omega) = i(\tau_0 - \tau_0^*)\psi'(\omega; \tau_0) - \tau_0\tau_0^*\psi(\omega; \tau_0), \tag{3.13}$$

we conclude that $\psi(\omega, \tau_0)$, $\psi'(\omega, \tau_0)$, $\psi''(\omega, \tau_0) - g(\omega)$ are continuous functions of ω in the range $0 \leq \omega < \infty$. From the continuity of these functions and from the existence of the integral $\int_0^\infty g(\omega')e^{-i\tau\omega'}d\omega'$, which has been assumed in Eq. (1.1), we conclude that also the three integrals

$$\int_0^\omega \psi(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega', \quad \int_0^\omega \psi'(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega', \quad \int_0^\omega \psi''(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega', \tag{3.14}$$

which occur in Eqs. (3.8), (3.9), (3.10), exist for all pairs of complex values τ, τ_0 with $\operatorname{Im}\tau \leq 0$ and $(\operatorname{Re}\tau_0 > 0, \operatorname{Im}\tau_0 \leq 0)$.

Now we assert that the integrals in Eq. (3.14) exist even when $\omega \to \infty$. From the identities (3.8), (3.9) we obtain, as $\omega \to \infty$,

$$\int_0^\infty \psi'(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega' = i\tau\int_0^\infty \psi(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega', \tag{3.15a}$$

$$\int_0^\infty \psi''(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega' = -\tau^2\int_0^\infty \psi(\omega'; \tau_0)e^{-i\tau\omega'}\,d\omega', \tag{3.15b}$$

where (2.4b), (2.4c) were used. All we have to show then is that the integral $\int_0^\infty \psi(\omega'; \tau_0)e^{-i\tau\omega'}d\omega'$ exists for τ and τ_0 as chosen above. Using Eq. (2.4a), the identity (3.10) becomes

$$\int_0^\omega g(\omega')e^{-i\tau\omega'}\,d\omega' - [\psi'(\omega; \tau_0) + i\{\tau - (\tau_0 - \tau_0^*)\}\psi(\omega; \tau_0)]e^{-i\tau\omega} + [\psi'(0; \tau_0) + i\{\tau - (\tau_0 - \tau_0^*)\}\psi(0; \tau_0)] = (\tau - \tau_0)(\tau + \tau_0^*)\int_0^\omega [-\psi(\omega; \tau_0)]e^{-i\tau\omega}\,d\omega, \tag{3.16}$$

and using (2.4b) and (2.4c), we obtain, as $\omega \to \infty$,

$$\gamma(\tau) \equiv \int_0^\infty g(\omega')e^{-i\tau\omega'}\,d\omega' = (\tau - \tau_0)(\tau + \tau_0^*)\int_0^\infty [-\psi(\omega'; \tau_0)]e^{-i\tau\omega'}\,d\omega'. \tag{3.17}$$

The left-hand side of Eq. (3.17) is the correlation function $\gamma(\tau)$, which exists. Hence also the integral $\int_0^\infty \psi(\omega'; \tau_0)e^{-i\tau\omega'}d\omega'$ exists for τ, τ_0 such that $\operatorname{Im}\tau \leq 0$ and $\operatorname{Re}\tau_0 > 0$, $\operatorname{Im}\tau_0 \leq 0$. Moreover,

$$\gamma(\tau_0) = \gamma(-\tau_0^*) = 0, \tag{3.18}$$

which was to be shown. [What we tried to show above was that each separate term in the identities (3.8), (3.9), (3.10) exists, so that no cancellation of any meaningless expressions occurs.] Equation (3.17), which is identical to Eq. (2.10), proves the assertion (d).

Finally let us prove that the statement (ii) implies the statement (iii). [Because of the equivalence of statements (i) and (ii) we conclude then that statement (i) also implies the statement (iii).]

For this purpose we define the function

$$\psi(\omega; \tau_0) = \frac{1}{i(\tau_0 + \tau_0^*)} [e^{i\tau_0\omega}\gamma(\tau_0, \omega) - e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega)] \tag{3.19}$$

in the range $0 \leq \omega < \infty$. This function is real because of Eq. (2.8) and is uniquely determined once $g(\omega)$ is given $(0 \leq \omega < \infty)$. The first and second derivatives of this function are

$$\psi'(\omega; \tau_0) = \frac{1}{\tau_0 + \tau_0^*} [\tau_0 e^{i\tau_0\omega}\gamma(\tau_0, \omega) + \tau_0^* e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega)], \tag{3.20}$$

$$\psi''(\omega; \tau_0) = g(\omega) + \frac{1}{i(\tau_0 + \tau_0^*)} [-\tau_0^2 e^{i\tau_0\omega}\gamma(\tau_0, \omega) + \tau_0^{*2} e^{-i\tau_0^*\omega}\gamma(-\tau_0^*, \omega)], \tag{3.21}$$

from which we conclude that the function $\psi(\omega, \tau_0)$ satisfies the second-order differential equation

$$\psi''(\omega; \tau_0) - i(\tau_0 - \tau_0^*)\psi'(\omega; \tau_0) + \tau_0\tau_0^*\psi(\omega; \tau_0) = g(\omega). \tag{3.22}$$

Moreover it satisfies the boundary conditions

$$\psi(0, \tau_0) = \psi'(0, \tau_0) = 0 \tag{3.23}$$

since

$$\gamma(\tau_0, 0) = \gamma(-\tau_0^*, 0) = 0. \tag{3.24}$$

Equations (3.22) and (3.23) completely define the function $\psi(\omega, \tau_0)$, and so they are equivalent to Eq. (3.19) by means of which this function is defined.

Now assuming the statement (ii) [which is equivalent to Eqs. (2.2) and (2.9)] and taking into account the expressions (3.19) and (3.20) for $\psi(\omega, \tau_0)$ and $\psi'(\omega, \tau_0)$, we find that

$$\psi(\infty; \tau_0) = \psi'(\infty; \tau_0) = 0. \tag{3.25}$$

The reason why we must consider the function $\psi(\omega, \tau_0)$ as well as its derivative as $\omega \to \infty$ is that $\psi(\omega, \tau_0)$ is a real function while the limit in Eq. (2.2) implies in fact two conditions, since the expression $e^{i\tau_0\omega}\gamma(\tau_0, \omega)$ is generally complex. The function $\psi(\omega, \tau_0)$ is proportional to the imaginary part of this expression, while the function $\psi'(\omega, \tau_0)$ is proportional to a linear combination of its real and imaginary parts. Also if τ_0 is not a zero of $\gamma(\tau)$, then certainly $\psi(\omega, \tau_0)$ and/or $\psi'(\omega, \tau_0)$ do not tend to zero as $\omega \to \infty$, as deduced from Eqs. (3.19), (3.20) and the statement (b).

The results (3.22), (3.23), and (3.25) taken together are the statement (iii), which was to be proved.

We have thus demonstrated that the statements (i), (ii), and (iii) are equivalent.

From Eq. (3.17) we see what is the mathematical meaning of the function $\psi(\omega, \tau_0)$. Let $\gamma(\tau)$ have a zero at τ_0 ($\mathrm{Re}\,\tau_0 > 0$, $\mathrm{Im}\,\tau_0 \leq 0$) (and consequently also at $-\tau_0^*$) and let us set

$$\gamma(\tau) = (\tau - \tau_0)(\tau + \tau_0^*)\gamma(\tau, \tau_0). \tag{3.26}$$

Comparing this expression to Eq. (3.17), it follows that

$$\gamma(\tau, \tau_0) = \int_0^\infty [-\psi(\omega'; \tau_0)]e^{-i\tau\omega'}\, d\omega'. \tag{3.27}$$

If the correlation function $\gamma(\tau)$ vanishes for $\tau = \tau_0$, then the function $-\psi(\omega, \tau_0)$ exists with the properties (iii). Then also Eq. (3.17) is valid [consequence of (iii)], and so is Eq. (3.27). We see then that the integral $\int_0^\infty [-\psi(\omega'; \tau_0)]e^{-i\tau\omega'}d\omega'$ of the function $-\psi(\omega, \tau_0)$, defined in the range $0 \leq \omega < \infty$, is the correlation function divided by $(\tau - \tau_0)(\tau + \tau_0^*)$. We note that the form of this integral is exactly the same as the one defining the correlation function in Eq. (1.1), except that the spectral density $g(\omega)$ has been replaced by the function $-\psi(\omega, \tau_0)$.

We also conclude that the existence of a zero τ_0 of the correlation function is equivalent to the solution of an eigenvalue problem, since the function $\psi(\omega, \tau_0)$ with the properties (2.4) exists only if a zero of the correlation function $\gamma(\tau)$ exists.

The statement (iii) does not make it in any way easier to find the zeros of $\gamma(\tau)$ in any particular case from the knowledge of the spectral density $g(\omega)$. But it does solve the converse problem of allowing the determination of spectra $g(\omega)$ such that the corresponding correlation function $\gamma(\tau)$ has zeros at prescribed points τ_0. This is done by choosing an appropriate function $\psi(\omega, \tau_0)$ satisfying conditions (2.4b) and (2.4c) and then constructing $g(\omega)$ according to the formula (2.4a).

So far we have confined our attention to the case of a single zero. Our analysis can, however, be extended to the case of several zeros and such a situation will be considered in another publication.

4. PROPERTIES OF THE FUNCTION $\psi(\omega; \tau_0)$

If we set $\tau_0 = \alpha - i\beta$ ($\alpha > 0$, $\beta \geq 0$), the system (2.4) becomes

$$\psi''(\omega; \tau_0) - 2\beta\psi'(\omega; \tau_0) + (\alpha^2 + \beta^2)\psi(\omega; \tau_0) = g(\omega), \tag{4.1a}$$

$$\psi(0; \tau_0) = \psi'(0; \tau_0) = 0, \tag{4.1b}$$

$$\psi(\infty; \tau_0) = \psi'(\infty; \tau_0) = 0. \tag{4.1c}$$

Let us multiply Eq. (4.1a) by $\psi'(\omega, \tau_0)$. Then

$$\tfrac{1}{2}[\psi'^2(\omega; \tau_0)]' - 2\beta\psi'^2(\omega; \tau_0) + \tfrac{1}{2}(\alpha^2 + \beta^2) \times [\psi^2(\omega; \tau_0)]' = g(\omega)\psi'(\omega; \tau_0), \tag{4.2}$$

or, after integrating over the ω range from 0 to ω and on taking into account the condition (4.1b), we find that

$$\tfrac{1}{2}\psi'^2(\omega; \tau_0) + \tfrac{1}{2}(\alpha^2 + \beta^2)\psi^2(\omega; \tau_0) = \int_0^\omega g(\omega')\psi'(\omega'; \tau_0)\, d\omega' + 2\beta\int_0^\omega \psi'^2(\omega'; \tau_0)\, d\omega'. \tag{4.3}$$

The left-hand side of this equation tends to zero as $\omega \to \infty$, and the integral $\int_0^\omega \psi'^2(\omega'; \tau_0)\, d\omega'$ will be larger than some positive number M if $\omega > \omega_0$, so that the expression

$$\frac{\tfrac{1}{2}\psi'^2(\omega; \tau_0) + \tfrac{1}{2}(\alpha^2 + \beta^2)\psi^2(\omega; \tau_0)}{\int_0^\omega \psi'^2(\omega'; \tau_0)\, d\omega'}$$

tends to zero as $\omega \to \infty$. From Eq. (4.3) it then follows that

$$\beta = -\tfrac{1}{2}\lim_{\omega\to\infty} \frac{\int_0^\omega g(\omega')\psi'(\omega'; \tau_0)\, d\omega'}{\int_0^\omega \psi'^2(\omega'; \tau_0)\, d\omega'}. \tag{4.4}$$

If the function $\psi'(\omega;\tau_0)$ is square integrable, then

$$\beta = -\frac{1}{2}\frac{\int_0^\infty g(\omega')\psi'(\omega';\tau_0)\,d\omega'}{\int_0^\infty \psi'^2(\omega';\tau_0)\,d\omega'}. \tag{4.5}$$

Another property of the function $\psi(\omega;\tau_0)$ can be obtained from Eq. (2.10) if we set $\tau = 0$. We then have

$$\int_0^\infty \psi(\omega';\tau_0)\,d\omega' = \frac{1}{\tau_0\tau_0^*} = \frac{1}{\alpha^2+\beta^2}. \tag{4.6}$$

We see then that the function $\psi(\omega;\tau_0)$ is integrable.

Let us solve the differential equation (4.1a) subject to conditions (4.1b). We have

$$\psi(\omega;\tau_0) = \frac{1}{\alpha}\int_0^\omega g(\omega')e^{\beta(\omega-\omega')}\sin\alpha(\omega-\omega')\,d\omega', \tag{4.7a}$$

or, changing the variable of integration,

$$\psi(\omega;\tau_0) = \frac{1}{\alpha}\int_0^\omega g(\omega-\omega')e^{\beta\omega'}\sin\alpha\omega'\,d\omega'. \tag{4.7b}$$

Let us restrict ω to the interval $0 \le \omega \le \pi/\alpha$. Since ω' lies in the interval $0 \le \omega' \le \omega$, we have $0 \le \omega' \le \pi/\alpha$ or $0 \le \alpha\omega' \le \pi$ and $0 \le \sin\alpha\omega' \le 1$. Recalling that the spectral density $g(\omega)$ is nonnegative, we see that if $0 \le \omega \le \pi/\alpha$, the integrand in Eq. (4.7b) is also nonnegative, and hence

$$\psi(\omega;\tau_0) \ge 0 \quad \text{for} \quad 0 \le \omega \le \pi/\alpha. \tag{4.8}$$

The functions $\psi(\omega;\tau_0)$ and $\psi'(\omega;\tau_0)$ are continuous and they tend to zero as $\omega \to 0$. Let us assume that $\psi(\omega;\tau_0)$ becomes negative for some range of ω values with $\omega > \omega_0$ (certainly $\omega_0 > \pi/\alpha$). Since the function $\psi(\omega;\tau_0)$ is continuous and tends to zero as $\omega \to \infty$, it will certainly have a minimum for some $\omega_m > \omega_0$ where

$$\psi(\omega_m;\tau_0) < 0,$$
$$\psi'(\omega_m;\tau_0) = 0,$$
$$\psi''(\omega_m;\tau_0) = g(\omega_m) - (\alpha^2+\beta^2)\psi(\omega_m;\tau_0) > 0.$$

Now for some range of values of ω with $\omega > \omega_m$, the function $\psi'(\omega;\tau_0)$ will be positive (it is equal to zero when $\omega = \omega_m$) and monotonically increasing. Since this function $\psi'(\omega;\tau_0)$ is continuous and tends to zero as $\omega \to \infty$, it will certainly reach a maximum value at $\omega = \omega_i > \omega_m$, where

$$\psi'(\omega_i;\tau_0) > 0,$$
$$\psi''(\omega_i;\tau_0) = 0,$$

and

$$\psi(\omega_i;\tau_0) = [1/(\alpha^2+\beta^2)][g(\omega_i) + 2\beta\psi'(\omega_i;\tau_0)] \ge 0,$$

because $\beta \ge 0$ and $g(\omega_i) \ge 0$. But ω_i is the first inflection point of the function $\psi(\omega_i;\tau_0)$ which is closest to its minimum at $\omega = \omega_m$, if we consider only values of $\omega \ge \omega_m$. We have then shown that:

(a) The function $\psi(\omega;\tau_0)$ must have an inflection point that follows a minimum at $\omega = \omega_m$ where $\psi(\omega_m;\tau_0) < 0$;

(b) At the inflection point $\omega = \omega_i > \omega_m$, we must have $\psi(\omega_i;\tau_0) > 0$, or $\psi(\omega_i;\tau_0) = 0$, if both $\beta = 0$ and $g(\omega_i) = 0$;

(c) In the interval $\omega_m \le \omega \le \omega_i$, the function $\psi'(\omega;\tau_0)$ is monotonically increasing and positive, except at $\omega = \omega_m$, where $\psi'(\omega_m;\tau_0) = 0$. Hence the function

$$\psi(\omega;\tau_0) = \psi(\omega_m;\tau_0) + \int_{\omega_m}^\omega \psi'(\omega';\tau_0)\,d\omega'$$

is also monotonically increasing in this interval.

An immediate consequence of the results (a), (b), and (c) is that, if there is an ω for which $\psi(\omega;\tau_0) < 0$, then there is certainly an $\omega' > \omega$ for which $\psi(\omega';\tau_0) > 0$; i.e., if $\psi(\omega;\tau_0)$ becomes negative for some value of ω, then there is a value ω' greater than ω for which it becomes positive again. In other words, there can be no value ω_0 for which $\psi(\omega;\tau_0) < 0$ for all $\omega > \omega_0$.

5. EQUIVALENCE OF THE ZERO PROBLEM TO A STABILITY PROBLEM IN MECHANICS

In the system of Eqs. (4.1) let us make the substitutions

$$\omega = t,$$
$$\psi(\omega;\tau_0) = X(t), \tag{5.1}$$
$$g(\omega) = F(t).$$

Then the system becomes

$$X''(t) - 2\beta X'(t) + (\alpha^2+\beta^2)X(t) = F(t), \tag{5.2a}$$
$$X(0) = X'(0) = 0, \tag{5.2b}$$
$$X(\infty) = X'(\infty) = 0. \tag{5.2c}$$

Now if we rewrite (5.2a) in the form

$$X''(t) = F(t) + 2\beta X'(t) - (\alpha^2+\beta^2)X(t), \tag{5.3}$$

we see that this is the equation of motion of a particle with unit mass upon which the following forces are acting:

(a) The force $F(t) = g(t)$, which is nonnegative for all values of $t \ge 0$;

(b) A force $+2\beta X'(t)$ proportional to the velocity of the particle, not opposing the motion but in the *same* direction to it (consequently, *not* a friction force of the medium surrounding the particle);

(c) A force $-(\alpha^2+\beta^2)X(t)$ proportional to its

displacement, i.e., an elastic force, with elastic constant $\alpha^2+\beta^2$, which tends to move the particle towards the origin.

The forces $F(t)$ and $+2\beta X'(t)$ have the tendency to move the particle to infinity while the force $-(\alpha^2+\beta^2)X(t)$ has the tendency to bring the particle back to the origin.

The particle starts at $t=0$ from the origin with zero velocity. The requirement is that it returns to the origin with zero velocity at $t=\infty$. This cannot be done with an arbitrary set of the parameters $(2\beta, \alpha^2+\beta^2)$ or (α,β) once the force $F(t)$ is given. As a matter of fact there may be no values (α,β) for which $X(\infty)=X'(\infty)=0$. In this case the particle will oscillate around the origin with increasing amplitude, i.e. its position is *unstable*.

Let us define as *stable* a solution of the system of Eqs. (5.2a), (5.2b) which satisfies also Eq. (5.2c). In this case the particle returns back to the origin with zero velocity at $t=\infty$, i.e., it does not oscillate indefinitely in time with a finite amplitude. In this case the particle will eventually return to the origin—its position is a stable one at $t=\infty$. This situation arises when the parameters (α,β) are properly chosen; whether this is in fact possible depends on the time-dependent force $F(t)$. It is obvious then that our problem as to the existence and location of zeros of the correlation function in the lower half and on the real axis of the complex τ plane is equivalent to the stability problem in mechanics, which we just discussed.

6. EXAMPLES

Example 1

As an illustration of the equivalence of statements (i) and (iii), we first give an example which can be solved explicitly. Consider the case when the power spectrum is given by the expression

$$g(\omega)=\frac{1}{6}\frac{\mu^4}{\lambda^2+\mu^2}[(\lambda^2+4\mu^2)\omega^3-12\mu\omega^2+6\omega]e^{-\mu\omega}, \tag{6.1}$$

where $\lambda>0$, $\mu>0$; if $g(\omega)$ is to be nonnegative, we must also assume that $\lambda^2\geq 2\mu^2$.

Evaluating the integral (1.1), we find that

$$\gamma(\tau)=-\frac{\mu^4}{\lambda^2+\mu^2}\frac{1}{(\mu+i\tau)^4}(\tau-\lambda+i\mu)(\tau+\lambda+i\mu). \tag{6.2}$$

We see then that the correlation function vanishes for $\tau_0=\lambda-i\mu$ and $-\tau_0^*=-\lambda-i\mu$.

Next let us solve the differential Eq. (2.4a) subject to conditions (2.4b). We set $\tau_0=\alpha-i\beta$ $(\alpha>0, \beta\geq 0)$, where α,β are unknown quantities that may be determined by using condition (2.4c). We obtain the solution

$$\begin{aligned}\psi(\omega;\tau_0)=\frac{1}{6}\frac{\mu^4}{\lambda^2+\mu^2}\Bigg[&\frac{6}{\alpha\{\alpha^2+(\beta+\mu)^2\}^2}\\ &\times Ae^{\beta\omega}\sin\alpha\omega-\frac{12}{\{\alpha^2+(\beta+\mu)^2\}^2}\\ &\times Be^{\beta\omega}\cos\alpha\omega+\frac{12}{\{\alpha^2+(\beta+\mu)^2\}^2}Ce^{-\mu\omega}\\ &+\frac{6}{\alpha^2+(\beta+\mu)^2}D\omega e^{-\mu\omega}+\frac{6}{\alpha^2+(\beta+\mu)^2}\\ &\times E\omega^2e^{-\mu\omega}+\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\omega^3e^{-\mu\omega}\Bigg],\end{aligned} \tag{6.3}$$

where

$$\begin{aligned}A=&(\beta+\mu)^2-\alpha^2-\frac{4\mu(\beta+\mu)}{\alpha^2+(\beta+\mu)^2}((\beta+\mu)^2-3\alpha^2)\\ &+\frac{\lambda^2+4\mu^2}{\{\alpha^2+(\beta+\mu)^2\}^2}((\beta+\mu)^4-6\alpha^2(\beta+\mu)^2+\alpha^4),\end{aligned} \tag{6.4a}$$

$$\begin{aligned}B=&\beta+\mu+\frac{2\alpha^2}{\alpha^2+(\beta+\mu)^2}\\ &\times\left(\mu-(\beta+\mu)\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\right)+\frac{2(\beta+\mu)^2}{\alpha^2+(\beta+\mu)^2}\\ &\times\left(\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}(\beta+\mu)-3\mu\right),\end{aligned} \tag{6.4b}$$

$$\begin{aligned}C=&\beta+\mu+\frac{2\alpha^2}{\alpha^2+(\beta+\mu)^2}\\ &\times\left(\mu-(\beta+\mu)\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\right)+\frac{2(\beta+\mu)^2}{\alpha^2+(\beta+\mu)^2}\\ &\times\left(\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}(\beta+\mu)-3\mu\right),\end{aligned} \tag{6.4c}$$

$$\begin{aligned}D=&\frac{\alpha^2}{\alpha^2+(\beta+\mu)^2}\left(1-\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\right)\\ &+\frac{\beta+\mu}{\alpha^2+(\beta+\mu)^2}\left(4\cdot\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}(\beta+\mu)-8\mu\right),\end{aligned} \tag{6.4d}$$

$$E=\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}(\beta+\mu)-2\mu. \tag{6.4e}$$

If the function $\psi(\omega;\tau_0)$ and its derivative are to tend to zero as $\omega\to\infty$ [condition (2.4c)], we must have

$$A=B=0, \tag{6.5}$$

or

$$A \equiv (\beta-\mu)^2+\lambda^2-\alpha^2+\frac{8\alpha^2(\beta+\mu)}{\alpha^2+(\beta+\mu)^2} \times\left(2\mu-(\beta+\mu)\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\right)=0, \quad (6.6)$$

$$B \equiv \beta-\mu+\frac{2\cdot(\alpha^2-(\beta+\mu)^2)}{\alpha^2+(\beta+\mu)^2} \times\left(2\mu-(\beta+\mu)\frac{\lambda^2+4\mu^2}{\alpha^2+(\beta+\mu)^2}\right)=0. \quad (6.7)$$

To solve Eqs. (6.6) and (6.7) for α and β, we note that the expression $4\alpha^2(\beta+\mu)B-(\alpha^2-(\beta+\mu)^2)A$ must identically vanish. Hence

$$4\alpha^2(\beta+\mu)(\beta-\mu)-(\alpha^2-(\beta+\mu)^2) \times[(\beta-\mu)^2+\lambda^2-\alpha^2]=0, \quad (6.8)$$

or

$$\alpha^4+[4(\beta^2-\mu^2)-(\beta+\mu)^2-(\beta-\mu)^2-\lambda^2]\alpha^2 +(\beta+\mu)^2[(\beta-\mu)^2+\lambda^2]=0. \quad (6.9)$$

Now Eq. (6.7) is equivalent to the equation

$$(\beta+3\mu)\alpha^4+\{2\mu^2(\beta+\mu)^3-(\lambda^2+4\mu^2) \times(\beta+\mu)-2(\lambda^2+4\mu^2)\}\alpha^2+(\beta+\mu)^2 \times[(\beta-\mu)((\beta+\mu)^2-4\mu^2)+\beta+3\mu]=0. \quad (6.10)$$

If we multiply Eq. (6.9) by $(\beta+3\mu)$ and subtract it from Eq. (6.10), we obtain the equation

$$(\beta-\mu)[\{4\mu(\beta+\mu)+4\mu^2+\lambda^2\}\alpha^2+(\beta+\mu)^2 \times\{4\mu(\beta+\mu)-4\mu^2-\lambda^2\}]=0. \quad (6.11)$$

This equation shows that either

$$\beta=\mu \quad (6.12)$$

or

$$\alpha^2=-(\beta+\mu)^2\frac{4\mu(\beta+\mu)-4\mu^2-\lambda^2}{4\mu(\beta+\mu)+4\mu^2+\lambda^2}. \quad (6.13)$$

First let us consider the case $\beta=\mu$. Then from Eq. (6.7) we obtain either

$$\alpha=\lambda \quad (6.14)$$

or

$$\alpha=2\mu. \quad (6.15)$$

Now for the pair of values $\alpha=\lambda$, $\beta=\mu$ we obtain, from Eq. (6.6), $A=0$, while for the pair of values $\alpha=2\mu$, $\beta=\mu$ we obtain [again from Eq. (6.6)] $A=-\lambda^2+4\mu^2$, and this quantity is zero only if $\lambda=2\mu$, in which case the pair $\alpha=2\mu$, $\beta=\mu$ becomes $\alpha=\lambda$, $\beta=\mu$. Then, if both Eqs. (6.6) and (6.7) are to be satisfied, we must have $\alpha=\lambda$, $\beta=\mu$.

Next let us consider the case when α is given by Eq. (6.13). Substituting for α^2 from Eq. (6.13) into Eq. (6.7), we obtain $B=\beta+3\mu=0$; this condition can never be satisfied since $\beta\geq 0$, $\mu>0$. We see then that if both Eqs. (6.6), (6.7) are to be satisfied, Eq. (6.11) must also be satisfied. But if $\beta\neq\mu$, $\beta\geq 0$, $\mu>0$, and Eq. (6.11) is satisfied, then Eq. (6.7) cannot be satisfied. We have then the final result that the *only* pair of values α, β which satisfies both Eqs. (6.6), (6.7) is $\alpha=\lambda$, $\beta=\mu$, so that $\tau_0=\lambda-i\mu$. For this pair of values we obtain the following eigensolution of $\psi(\omega;\tau_0)$:

$$\psi(\omega;\tau_0=\lambda-i\mu)=\frac{1}{6}\frac{\mu^4}{\lambda^2+\mu^2}\omega^3e^{-\mu\omega}. \quad (6.16)$$

The above example illustrates that the statement (iii), which is equivalent to the statement (i), does not seem to help appreciably in the determination of the zeros of the correlation function, but it does make it easy to determine spectra $g(\omega)$, which are such that the corresponding correlation function $\gamma(\tau)$ has zeros at prescribed points τ_0.

Example 2

Next, we give an example of a quasi-monochromatic spectral density for which the corresponding correlation function has no zeros anywhere in the complex τ plane. Such an example is furnished by the spectral density

$$g(\omega)=(q^{p+1}/p!)\,\omega^p e^{-q\omega}, \quad (6.17)$$

where $p\geq 0$, $q>0$, and $p\gg 1$. The corresponding correlation function is readily found to be

$$\gamma(\tau)=q^{p+1}/(q+i\tau)^{p+1}. \quad (6.18)$$

We see that $\gamma(\tau)$ does not vanish for any complex value of τ.

Next we show that the spectral density $g(\omega)$, given by Eq. (6.17), is quasi-monochromatic if $p\gg 1$. We define

$$\Delta\omega=[\overline{\omega^2}-\bar{\omega}^2]^{\frac{1}{2}}, \quad (6.19)$$

where the average, denoted by a bar, is defined by

$$\overline{f(\omega)}=\frac{\int_0^\infty f(\omega)g(\omega)\,d\omega}{\int_0^\infty g(\omega)\,d\omega}. \quad (6.20)$$

The spectral density will be said to be quasi-monochromatic if $\Delta\omega/\bar{\omega}\ll 1$.

For the spectral density $g(\omega)$ given by Eq. (6.17), we have

$$\overline{\omega^2}=(q)^{-2}(p+1)(p+2), \quad (6.21)$$

$$\bar{\omega}=(q)^{-1}(p+1), \quad (6.22)$$

so that, according to Eq. (6.19),

$$\Delta\omega = (q)^{-1}(p+1)^{\frac{1}{2}}. \tag{6.23}$$

We see that the ratio

$$\Delta\omega/\bar{\omega} = 1/(p+1)^{\frac{1}{2}} \tag{6.24}$$

is much less than unity if $p \gg 1$.

Example 3

Now using some results derived in the previous example, we determine a spectral density $g(\omega)$ with its corresponding correlation function having a *single* pair of zeros in the complex τ plane. For this purpose let us define the function

$$\psi(\omega; \tau_0 = \alpha - i\beta) = \frac{1}{\alpha^2+\beta^2}\frac{q^{p+1}}{p!}\omega^p e^{-q\omega}, \tag{6.25}$$

where $p > 1$, $q > 0$, $\alpha > 0$, $\beta \geq 0$. This function is similar to the spectral density given by Eq. (6.17). Then, according to Eq. (6.18),

$$\int_0^\infty [-\psi(\omega; \tau_0)]e^{-i\tau\omega}\,d\omega = -\frac{1}{\alpha^2+\beta^2}\frac{q^{p+1}}{(q+i\tau)^{p+1}}. \tag{6.26}$$

The function $\psi(\omega; \tau_0)$ satisfies the conditions (2.4b), (2.4c) and

$$\psi'(\omega; \tau_0) = \frac{1}{\alpha^2+\beta^2}\frac{q^{p+1}}{p!}(p - q\omega)\omega^{p-1}e^{-q\omega}, \tag{6.27}$$

$$\psi''(\omega; \tau_0) = \frac{1}{\alpha^2+\beta^2}\frac{q^{p+1}}{p!} \times [(p-1)p - 2pq\omega + q^2\omega^2]\omega^{p-2}e^{-q\omega}. \tag{6.28}$$

Substituting for $\psi(\omega; \tau_0)$ and its derivatives in Eq. (2.4a), we obtain the following spectral density function:

$$g(\omega) = \frac{1}{\alpha^2+\beta^2}\frac{q^{p+1}}{p!} \times [\{\alpha^2 + (\beta+q)^2\}\omega^2 - 2p(\beta+q)\omega + (p-1)p] \times \omega^{p-2}e^{-q\omega}. \tag{6.29}$$

This spectral density function is nonnegative if $(p-1)\alpha^2 > (\beta+q)^2$.

From Eqs. (6.25) and (6.29) we see that in the solution of the system (2.4) different spectral densities $g(\omega)$ may lead to the same eigensolution $\psi(\omega; \tau_0)$. In this example the different spectral densities $g(\omega)$ correspond to different values of the set α, β with $\alpha^2+\beta^2$ equal to a given constant, and $\psi(\omega; \tau_0)$ depends only on $\alpha^2+\beta^2$ and not on the particular values α and β.

Making use of Eqs. (2.10) and (6.26), we find that the corresponding correlation function is

$$\gamma(\tau) = -\frac{1}{\alpha^2+\beta^2}(\tau - \alpha + i\beta) \times (\tau + \alpha + i\beta)\frac{q^{p+1}}{(q+i\tau)^{p+1}}. \tag{6.30}$$

We see then that $\gamma(\tau)$ has a *single* pair of zeros at the points $\tau_0 = \alpha - i\beta$ and $-\tau_0^* = -\alpha - i\beta$, and $\gamma(0) = 1$, so that the integral of the spectral density $g(\omega)$, given by Eq. (6.29), is normalized to unity [in accordance to Eq. (1.4)].

ACKNOWLEDGMENTS

The author wishes to thank Professor E. Wolf for suggesting the problem to him and for helpful discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Green's Functions and Double-Time Distribution Functions in Classical Statistical Mechanics

JOHN C. HERZEL
School of Physics, Melbourne University, Victoria, Australia.

(Received 10 January 1967)

This paper rederives the Bogoliubov/Sadovnikov classical-equilibrium-correlation Green's function hierarchy by using the double-time theory of Rostoker. Thus the traditional variational technique is avoided.

1. INTRODUCTION

ONE need only refer to Zubarev's[1] review paper on quantum statistical Green's functions to realize the importance of the concept of these functions not only in quantum field theory but also in quantum statistical mechanics.

In classical statistical mechanics Balescu demonstrated the utility of causal and anticausal Green's functions in the initial-value solution of the Liouville equation[2] and also in the study of transport coefficients.[2,3] This technique was also used by Severne[4] in a study of systems in uniform external fields.

The question of extending the quantum statistical Green's approach to classical statistics has been raised by Bogoliubov and Sadovnikov.[5] These authors established a hierarchy of classical equilibrium Green's functions using a variational technique on the single-time distribution functions of the Bogoliubov–Born–Green–Kirkwood–Yvon hierarchy.[6] Recently Sadovnikov[7] employed this hierarchy in a study of the Boltzmann equation.

In the present paper the causal Green's functions are used to relate the double-time system function[8] of Rostoker[6,9] to the well-known single-time system function. Reduced double-time distribution functions are defined in which no particles are singled out as test particles. These reduced functions, which obey the BBGKY equations in their later-time variables, are directly related to the correlation functions of the B–S theory with the result that obtention of the Green's hierarchy is straightforward.

2. THE SYSTEM

The system considered is conservative, contained in a box of volume V, and consists of N identical particles interacting via central two-body forces. All results are stated in the thermodynamic limit, $N \to \infty$, $V \to \infty$, for which it is assumed $N/V \to n$. Integrations extend over the entire phase space of the variables concerned, and when performed by parts assume that all distribution functions and Green's functions obey periodic boundary conditions in positional space and homogeneous boundary conditions in momentum space such that they approach zero exponentially as any momentum variable on which they depend approaches infinity.

3. SYSTEM FUNCTIONS AND THE LIOUVILLE EQUATION

The single-time system distribution function $D_1(\Gamma, t)$,[10] the probability density that, at a particular time t, the system's particles occupy the phase points Γ, satisfies the Liouville equation

$$\frac{\partial D_1}{\partial t}(\Gamma, t) = [H_N(\Gamma);\, D_1(\Gamma, t)], \tag{1}$$

where

$$[A; B] \equiv \sum_{1 \le j \le N} \left(\frac{\partial A}{\partial \mathbf{x}_j} \cdot \frac{\partial B}{\partial \mathbf{p}_j} - \frac{\partial A}{\partial \mathbf{p}_j} \cdot \frac{\partial B}{\partial \mathbf{x}_j} \right), \tag{2}$$

and

$$H_N(\Gamma) = \sum_{1 \le j \le N} \frac{\mathbf{p}_j^2}{2m} + \sum_{1 \le i < j \le N} \varphi_{ij}, \tag{3}$$

φ_{ij} being the scalar potential between particles i and

[1] D. N. Zubarev, Usp. Fiz. Nauk **71**, 71 (1960) [English transl.: Soviet Phys.—Usp. **3**, 320 (1960)].
[2] R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience Publishers, Inc., New York, 1963).
[3] R. Balescu, Physica **27**, 693 (1961).
[4] G. Severne, Physica **30**, 1365 (1964).
[5] N. N. Bogoliubov and B. I. Sadovnikov, Zh. Eksperim. i Teor. Fiz. **43**, 677 (1962) [English transl.: Soviet Phys.—JETP **16**, 482 (1963)].
[6] D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory* (McGraw-Hill Book Company, Inc., New York, 1964).
[7] B. I. Sadovnikov, Physica **32**, 858 (1966).
[8] The term "system function" is used as synonymous with N-particle function.
[9] N. Rostoker, Nucl. Fusion **1**, 101 (1961).
[10] The symbol Γ denotes the entire set of phase points $(\mathbf{x}_1, \mathbf{p}_1, \mathbf{x}_2, \mathbf{p}_2, \cdots, \mathbf{x}_N, \mathbf{p}_N)$ and $(\mathbf{x}_j, \mathbf{p}_j)$ is denoted briefly as $\mathbf{X}_j$. Where reference is made to particular points in the system rather than particular particles, the symbol $\mathbf{R}_j \equiv (i_j, \not{p}_j)$ is used.

j. More explicitly the Liouville equation is

$$\frac{\partial D_1}{\partial t} = \sum_{1\le j\le N}\left(\frac{\partial \varphi_{ij}}{\partial \mathbf{x}_j}\cdot\frac{\partial D_1}{\partial \mathbf{p}_j} - \frac{\mathbf{p}_j}{m}\cdot\frac{\partial D_1}{\partial \mathbf{x}_j}\right), \tag{4}$$

which may be abbreviated by appropriately defining a Liouville operator L such that

$$\partial D_1/\partial t = -iLD_1 . \tag{5}$$

This equation may be solved as an initial-value problem by defining causal and anticausal Green's functions.[2] The causal function is defined to be the solution of

$$(\partial/\partial t + iL)g(\Gamma t/\Gamma' t') = \prod_{j=1}^{N}\delta(\mathbf{X}_j - \mathbf{X}_j')\delta(t-t') \tag{6}$$

$$= \delta(\Gamma - \Gamma')\delta(t-t'), \tag{6a}$$

and to satisfy the causality condition[11]

$$g(t/t') = 0 \quad \text{for} \quad t < t'. \tag{7}$$

The unprimed differential operators in (6) only act on the unprimed variables of g.

One then has

$$D_1(\Gamma, t) = \int d\Gamma' g(\Gamma t/\Gamma' t_0) D_1(\Gamma', t_0). \tag{8}$$

This solution shows $g(\Gamma t/\Gamma' t_0)$ as a transition probability that the system in the state Γ' at time t_0 moves to the state Γ at a later time t. Thus one must have

$$g(\Gamma t/\Gamma' t') \xrightarrow[t\to t']{} \delta(\Gamma - \Gamma'). \tag{9}$$

One may now define the double-time system function of Rostoker's fluctuation theory[9] as

$$D_2(\Gamma t/\Gamma' t') = g(\Gamma t/\Gamma' t') D_1(\Gamma' t'); \tag{10}$$

that is to say, $D_2(\Gamma t/\Gamma' t')$ is the probability density that the system is at Γ' at time t' and at Γ at a later time t. In view of (9) the initial value of D_2 is given by

$$D_2(\Gamma t/\Gamma' t') \xrightarrow[t\to t']{} \delta(\Gamma - \Gamma') D_1(\Gamma' t'). \tag{11}$$

Also since g obeys the Liouville equation in its later-time variables when $t > t'$, the double-time system function also obeys it under this condition.

4. REDUCED DISTRIBUTION FUNCTIONS

Instead of following Rostoker and defining a reduced system function $\psi(\Gamma t/\mathbf{X}_1' t')$ such that $\psi(\Gamma t/\mathbf{X}_1' t')/V$ is the probability density that particle 1 is at $\mathbf{X}_1'$ at time t' and the system moves into the state Γ at time t, a symmetrical function $\Psi(\Gamma t/\mathbf{X}' t')$ is defined.

$$\Psi(\Gamma t/\mathbf{X}' t') \equiv \sum_{1\le j\le N}\int d\Gamma' D_2(\Gamma t/\Gamma' t')\delta(\mathbf{X}' - \mathbf{X}_j'). \tag{12}$$

No particle is singled out at the earlier time, that is, $\Psi(\Gamma t/\mathbf{X}' t')$ is the probability density that at t' there is a particle at $\mathbf{X}'$ and at time t the system occupies the point Γ. Reduced functions are defined as

$$\mathcal{F}_s(\mathbf{X}_1\mathbf{X}_2\cdots\mathbf{X}_s t/\mathbf{X}' t') = V^s\int d\mathbf{X}_{s+1}\, d\mathbf{X}_{s+2}\cdots d\mathbf{X}_N \Psi(\Gamma t/\mathbf{X}' t'). \tag{13}$$

As an example, the fluctuation of density in phase space is given by

$$\langle n(\mathbf{R}t) n(\mathbf{R}' t')\rangle = \int d\Gamma\, d\Gamma' D_2(\Gamma t/\Gamma' t') \sum_{\substack{1\le j\le N\\ 1\le n\le N}} \delta(\mathbf{R} - \mathbf{X}_j)\delta(\mathbf{R}' - \mathbf{X}_n') \tag{14}$$

$$= n\mathcal{F}_1(\mathbf{R}t/\mathbf{R}' t'). \tag{15}$$

The generalized class of dynamical variables defined in the Bogoliubov–Sadovnikov theory is

$$A_{\mathbf{R}_1,\mathbf{R}_2,\cdots\mathbf{R}_s}(t) = \sum_{1\le i_1<\cdots<i_s\le N}\delta(\mathbf{R}_1 - \mathbf{X}_{i_1}) \times \delta(\mathbf{R}_2 - \mathbf{X}_{i_2})\cdots\delta(\mathbf{R}_s - \mathbf{X}_{i_s}), \tag{16}$$

so that, instead of (15), one may write

$$\langle A_{\mathbf{R}}(t) A_{\mathbf{R}'}(t')\rangle = n\mathcal{F}_1(\mathbf{R}t/\mathbf{R}' t'). \tag{17}$$

In general,

$$\langle A_{\mathbf{R}_1\mathbf{R}_2\cdots\mathbf{R}_s}(t) A_{\mathbf{R}'}(t')\rangle = \frac{n^s}{s!}\mathcal{F}_s(\mathbf{R}_1\mathbf{R}_2\cdots\mathbf{R}_s t/\mathbf{R}' t'). \tag{18}$$

In terms of particle probabilities, this is thus the probability density that there is a particle at $(\mathbf{R}', t')$ and an arbitrary set of s particles occupying the phase points $\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_s$ in a particular array at time t. The $\mathcal{F}$'s obey the BBGKY hierarchy equations.

5. EQUILIBRIUM CORRELATION GREEN'S FUNCTIONS AND THEIR HIERARCHY

The classical analogs of the quantum statistical double-time retarded and advanced Green's functions are, respectively,[12]

$$G_r(t/t') = \mathcal{H}(t-t')\langle[A(t); B(t')]\rangle, \tag{19}$$

$$G_a(t/t') = -\mathcal{H}(t'-t)\langle[A(t); B(t')]\rangle, \tag{20}$$

where $\mathcal{H}$ is the Heaviside function.

[11] The notation used here treats primed variables as corresponding to earlier times and conforms with that of Balescu (see Ref. 2).

[12] From the definition of double-time averages in terms of D_2, one can show that these averages may be expressed in terms of classical Heisenberg operators $A_H(t) = e^{iLt}A(0)e^{-iLt}$, highlighting the analogy with quantum statistics.

For brevity, only retarded functions are considered here.

In the equilibrium case

$$D_1^0 = Q_N^{-1} e^{-H_N/\theta}, \tag{21}$$

where $\theta = kT$ and Q_N is the partition function. Using (21) one can show on integrating by parts that[13]

$$\langle [A(t); B(t')] \rangle = \theta^{-1} \langle B(t') [A(t); H_N] \rangle. \tag{22}$$

The dynamical quantities are considered to depend on time only via the time dependence of phase variables, so that they satisfy

$$(d/dt) A\{\Gamma(t)\} = [A; H_N]. \tag{23}$$

Combining (22) and (23) gives

$$\langle [A(t); B(t')] \rangle = (1/\theta)(d/dt) \langle A(t) B(t') \rangle. \tag{24}$$

Now

$$\langle A_\Gamma(t)\, A_{\mathbf{X}'}(t') \rangle = \Psi(\Gamma t/\mathbf{X}' t') \tag{25}$$

so that one may write[14]

$$(\partial/\partial t) \langle A_\Gamma(t) A_{\mathbf{X}'}(t') \rangle = [H_N; \langle A_\Gamma(t) A_{\mathbf{X}'}(t') \rangle]. \tag{26}$$

Equations (23), (24), and the fact that the system is conservative enable one to rewrite (26) as

$$(\partial/\partial t) \langle [A_\Gamma(t); A_{\mathbf{X}'}(t')] \rangle = [H_N; \langle [A_\Gamma(t); A_{\mathbf{X}'}(t')] \rangle. \tag{27}$$

Using (13), (18), (24), and (25), one can show that

$$\langle [A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}(t); A_{\mathbf{X}'}(t')] \rangle = V^s \frac{n^s}{s!} \int d\Gamma^{N-s} \langle [A_\Gamma(t); A_{\mathbf{X}'}(t')] \rangle. \tag{28}$$

Therefore multiplication of (27) by $V^s n^s/s!$ and integration with respect to $\mathbf{X}_{s+1}, \mathbf{X}_{s+2}, \cdots, \mathbf{X}_N$ yields a hierarchy of Green's functions.

$$\begin{aligned} \partial/\partial t \, \langle [A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}(t); A_{\mathbf{X}'}(t')] \rangle &= [H_s(\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s); \langle [A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}(t); A_{\mathbf{X}'}(t')] \rangle] \\ &\quad + (s+1) \sum_{1 \le i \le s} d\mathbf{X}_{s+1} \\ &\quad \times [\varphi_{i,s+1}; \langle [A_{\mathbf{X}_1 \cdots \mathbf{X}_s \mathbf{X}_{s+1}}(t); A_{\mathbf{X}'}(t')] \rangle]. \end{aligned} \tag{29}$$

H_s is the Hamiltonian of s interacting particles.

The hierarchy equation (25) of the B–S paper can be shown to be the one-sided Laplace transform of the above equation.[15]

Since the system is in equilibrium, D_1 does not depend explicitly on the time; and noting that $g(\Gamma t/\Gamma' t')$ is the solution to (6), one may write it as $g(\Gamma/\Gamma';\ t - t')$. Hence, according to (10), one may write

$$D_2(\Gamma t/\Gamma' t') = D_2(\Gamma/\Gamma'; t - t'), \tag{30}$$

and therefore the functional form of the fluctuations is

$$\langle A(t) B(t') \rangle = \varphi(t - t'). \tag{31}$$

If one now takes the Fourier transform of (29) with respect to $\tau = t - t'$, one in fact has a one-sided Laplace transform since $g(t/t')$ is zero for $t < t'$, so that

$$\begin{aligned} &-iE \langle\langle A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}; A_{\mathbf{X}'} \rangle\rangle_E \\ &= [H_s(\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s); \langle\langle A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}; A_{\mathbf{X}'} \rangle\rangle_E] \\ &\quad + (s+1) \sum_{1 \le i \le s} \int d\mathbf{X}_{s+1} [\varphi_{i,s+1}; \langle\langle A_{\mathbf{X}_1 \cdots \mathbf{X}_s \mathbf{X}_{s+1}}; A_{\mathbf{X}'} \rangle\rangle_E] \\ &\quad + \langle [A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}(t'); A_{\mathbf{X}'}(t')] \rangle / 2\pi, \end{aligned} \tag{32}$$

where

$$\langle\langle A; B \rangle\rangle_E \equiv \frac{1}{2\pi} \int_0^\infty d\tau e^{iE\tau} \langle [A(t); B(t')] \rangle, \tag{33}$$

and, in deriving (32), it has been assumed that the imaginary part of E is positive definite.

The initial-value term may be expressed in terms of the reduced equilibrium single-time distribution function F_s^0 defined by

$$F_s^0(\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s) = V^s \int d\Gamma^{n-s} D_1^0. \tag{34}$$

One can show that [the primed time has been removed in respect of (30)]

$$\begin{aligned} \langle [A_{\mathbf{R}_1 \mathbf{R}_2 \cdots \mathbf{R}_s}; A_{\mathbf{R}'}] \rangle &= \frac{n^s}{s!} \sum_{1 \le j \le s} \int d\mathbf{X}_1\, d\mathbf{X}_2 \cdots d\mathbf{X}_s \\ &\quad \times \prod_{\substack{i=1 \\ \neq j}}^{s} \delta(\mathbf{R}_i - \mathbf{X}_i) [\delta(\mathbf{R}_j - \mathbf{X}_j); \delta(\mathbf{R}' - \mathbf{X}_j)] F_s^0. \end{aligned} \tag{35}$$

Now using

$$\int_{-\infty}^{\infty} dx f(x) \frac{\partial}{\partial x} \delta(x - x') = -\frac{\partial}{\partial x'} f(x), \tag{36}$$

one finds that

$$\langle [A_{\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_s}; A_{\mathbf{X}'}] \rangle = \frac{n^s}{s!} \sum_{1 \le j \le s} [\delta(\mathbf{X}' - \mathbf{X}_j); F_s^0]. \tag{37}$$

On substitution of (37) in (32), the hierarchy presented as equation (25) of the paper by Bogoliubov and Sadovnikov is recovered.

ACKNOWLEDGMENTS

The author thanks N. E. Frankel for reading the manuscript.

This work was done with the assistance of a Melbourne University Research Grant.

[13] One can show that the double-time averages are expressible in terms of dynamical quantities which are solutions of (23), that is, $A(t) = e^{iLt} A(0)$, and which therefore are not operators. See J. Weinstock, Phys. Rev. 139, 388 (1965).

[14] Note that a hierarchy for the correlation functions may be written down valid for both the nonequilibrium and equilibrium cases.

[15] See Ref. 2, APP. 1, for a concise discussion of such one-sided Laplace transforms.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Existence and Uniqueness in the Large for Boundary Value Problems in Kinetic Theory

CARLO CERCIGNANI
*Massachusetts Institute of Technology, Cambridge, Massachusetts**

(Received 29 November 1966)

The boundary-value problem for the linearized Boltzmann equation is shown to have a unique solution for a bounded domain (two walls separated by an arbitrary distance). The proof applies to a general class of models with finite collision frequency and appears to be easily extendible to similar problems in two and three dimensions. It differs essentially from previously known proofs because no limitations are put on the distance between the walls.

1. INTRODUCTION

THE solution of boundary value problems in kinetic theory has been treated by many approximate methods, but until now the question of proving the existence of the solution has remained substantially untouched—even for linearized flows. In fact, either only trivial boundary conditions have been taken into account,[1] or the domain has been assumed to be sufficiently small (of the order of a mean free path).[2,3]

Rigorous existence proofs have only been presented for models of the Boltzmann equation, either directly by the method of integral iteration[4] or, indirectly, by the method of "elementary solutions."[5]

In the present paper we present a rather simple proof of existence and uniqueness for the linearized Boltzmann equation with finite collision frequency. Therefore, the proof applies to rigid sphere molecules as well as molecules interacting with any cutoff intermolecular potential. Any linearized molecular model, including the recently studied model with velocity-dependent collision frequency,[6,7] is also contained in the class of equations to which the proof applies.

The domain is assumed to be a slab and the boundary conditions have a rather simple but realistic form, i.e., the distribution function is given for emerging molecules at each wall. The problems which are considered are steady in a general sense, i.e., a dependence on time through an exponential factor e^{st} with $\mathrm{Re}\, s \geq 0$ is allowed. This implies the existence and uniqueness of Laplace transformable solutions in time-dependent problems, as well as the existence and uniqueness of the solution for forced wave problems ($s = i\omega$).

2. BASIC EQUATIONS AND OPERATORS

Consider a one-dimensional problem between two parallel plates for the separated time equation:

$$sf + c_x(\partial f/\partial x) = Lf, \tag{2.1}$$

where f is the perturbation of a basic Maxwellian f_0. We suppose that the emerging distribution $f_+ = \bar{f}_+$ is given at $x = -\frac{1}{2}d\,(c_x > 0)$ and $f = \bar{f}_-$ is given at $x = \frac{1}{2}d\,(c_x < 0)$. We restrict to suitable collision operators L by requiring that they can be split into two parts as follows:

$$Lf = Kf - \nu(c)f, \tag{2.2}$$

where $\nu(c)$ is a multiplication operator and K is such that $\nu^{-\frac{1}{2}}K\nu^{-\frac{1}{2}}$ is a completely continuous operator in the Hilbert space $\mathcal{H}$ of square summable functions with respect to the weight f_0. It is well-known[1,8] that the rigid spheres interaction and a general intermolecular potential with angular cutoff satisfy the above conditions. It is also very likely (and worthwhile to be proved rigorously) that more general cutoff operators (with radial rather than angular cutoff) enjoy the above properties.

As a consequence of these properties one can show (Appendix A) that a real number λ exists such that, if

$$Hf = Kf + \lambda\nu f, \tag{2.3}$$

then

$$0 < (f, Hf) \leq (\lambda + 1)(\nu f, f), \tag{2.4}$$

* On leave of absence from Applicazioni e Ricerche Scientifiche, Milano and Universita di Milano, Milano (Italy).

[1] H. Grad, in *Rarefied Gas Dynamics*, J. A. Laurmann, Ed. (Academic Press Inc., New York, 1963), Vol. I, p. 26.

[2] D. R. Willis, in *Rarefied Gas Dynamics*, L. Talbot, Ed. (Academic Press Inc., New York, 1961), p. 423.

[3] H. Grad, "High Frequency Sound According to the Boltzmann Equation", Joint AFOSR-NSF Report, AFOSR-66-0145, MF-49 (New York University, 1966)—to appear in the SIAM Journal.

[4] D. R. Willis, Princeton Univ. Aeronautical Eng. Rept. 440 (1958).

[5] C. Cercignani, "Elementary Solutions and Boundary Value Problems in the Kinetic Theory of Gases"—Brown University Report (1965).

[6] C. Cercignani, Ann. Phys. (N.Y.) **40**, 454 (1966).

[7] C. Cercignani, Ann. Phys. (N.Y.) **40**, 469 (1966).

[8] L. Finkelstein, Ph.D. thesis, Hebrew University, Jerusalem (1962).

where round brackets denote, as usual, inner product in the $\mathcal{H}$ space. Of course, if λ is a number such that (2.3) and (2.4) are satisfied, then any $\lambda > \lambda_0$ is also such a number. It could also turn out that $\min \lambda = 0$, in which case H can be taken to be equal to K.

As a consequence of Eqs. (2.2) and (2.3), we have

$$Lf = Hf - (\lambda + 1)\nu f, \tag{2.5}$$

and consequently one can rewrite Eq. (2.1) as an "integral" equation:

$$f = f_0 + UHf, \tag{2.6}$$

where

$$f_0 = \bar{f} \exp\left[- \frac{(\lambda + 1)\nu + s}{|c_x|} \left(\frac{d}{2} + x \operatorname{sgn} c_x \right) \right], \tag{2.7}$$

$$Ug = \frac{1}{c_x} \int_{-\frac{1}{2}d \operatorname{sgn} c_x}^{x} \exp\left\{ - \frac{[(\lambda + 1)\nu + s]}{|c_x|} |x - y| \right\} \times g(y)\, dy, \tag{2.8}$$

$$\bar{f} = \begin{cases} \bar{f}_+, & \text{for } c_x > 0, \\ \bar{f}_-, & \text{for } c_x < 0. \end{cases} \tag{2.9}$$

For a steady problem we have merely to set $s = 0$. In the following it is useful to consider a Hilbert space $\mathcal{K}$ of square summable functions with respect to both x and $\mathbf{c}$ and with the weight $\rho(c)f^{(0)}(c)$, where

$$\rho(\mathbf{c}) \equiv \{(\lambda + 1)^2[\nu(c)]^2 + (\pi^2/4d^2)c_x^2\}^{\frac{1}{2}}. \tag{2.10}$$

The norm in such a space is given by

$$\|f\|_{\mathcal{K}}^2 = \int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx \int d\mathbf{c}\rho(c)f^{(0)}(c)|f(x, \mathbf{c})|^2, \tag{2.11}$$

and because of the Fubini theorem,

$$\|f\|^2 = \| \|f\|_{\mathcal{C}} \|_x^2 = \| \|f\|_x \|_{\mathcal{C}}^2. \tag{2.12}$$

$\|f\|_{\mathcal{C}}^2$ and $\|f\|_x^2$ denote the partial norms

$$\|g\|_{\mathcal{C}}^2 = \int d\mathbf{c}\rho(\mathbf{c})f^{(0)}(c)|f(\mathbf{c})|^2, \tag{2.13}$$

$$\|g\|_x^2 = \int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx\, |g(x)|^2, \tag{2.14}$$

where $\mathcal{C}$ is the three-dimensional velocity space.

3. EXISTENCE AND UNIQUENESS IN THE LARGE

We present here a proof of the following:

Theorem: The integral equation form of the Boltzmann equation, Eq. (2.6), has one and only one solution $f \in \mathcal{K}$, for any given $\bar{f}$ such that $\|\bar{f}\|_{\mathcal{C}}^2$ is finite. This solution f can be obtained in principle by a convergent iteration procedure.

The above theorem is an immediate consequence of the contraction mapping theorem, the fact that f_0 as given by Eq. (2.6) belongs to $\mathcal{K}$, and the following:

Lemma: The operator UH is a contraction operator when acting on functions belonging to $\mathcal{K}$.

To prove this lemma, we first note that for any square summable $g(x)$,

$$\|Ug\|_x \leq [1/\rho(\mathbf{c})]\, \|g\|_x. \tag{3.1}$$

This inequality is a matter of a rather elementary exercise; details are given in Appendix B. From Eq. (3.1) it follows that

$$\begin{aligned} \|UHf\|_{\mathcal{K}}^2 &= \| \|UHf\|_x \|_{\mathcal{C}}^2 \leq \| \rho^{-1} \|Hf\|_x \|_{\mathcal{C}}^2 \\ &= \int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx \int d\mathbf{c}\rho^{-1}(\mathbf{c})\, |Hf|^2 f^{(0)}(c) \\ &= \int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx\, \|Jg\|_{\mathcal{H}}^2 = \| \|Jg\|_{\mathcal{H}} \|_x^2, \end{aligned} \tag{3.2}$$

where $\| \ \|_{\mathcal{H}}$ denotes the norm in the more usual Hilbert space $\mathcal{H}$, while

$$J \equiv \rho^{-\frac{1}{2}} H \rho^{-\frac{1}{2}};\ g = \rho^{\frac{1}{2}} f. \tag{3.3}$$

But it can be easily proved (Appendix C) that

$$\|Jg\|_{\mathcal{H}}^2 \leq \alpha\, \|g\|_{\mathcal{H}}^2 = \alpha\, \|f\|_{\mathcal{C}}^2, \tag{3.4}$$

where

$$0 < \alpha < 1.$$

Therefore

$$\|UHf\|_{\mathcal{K}}^2 \leq \alpha\, \| \|f\|_{\mathcal{C}} \|_x^2 = \alpha\, \|f\|_{\mathcal{H}}^2 \quad (0 < \alpha < 1); \tag{3.5}$$

i.e., UH is a contraction operator, as was to be shown.

4. CONCLUDING REMARKS

A theorem of existence and uniqueness for the boundary value problem of a rarefied gas enclosed between two parallel plates has been given. With respect to the previously known results,[2,3] we stress the fact that the proof is not for existence in the small but in the large; i.e., the separation between the walls is arbitrary.

The proof applies to a large class of collision operators, including all the cutoff interactions considered by Grad[3] and the rigid sphere interaction considered by Willis.[2] It is also very likely, although a formal proof has not been presented here, that more general cutoff operators (with radial rather than angular cutoff) are also included in the same class. This circumstance appears to be important since a full proof of this statement would give a sound mathematical basis to the theory of collision operators

with radial cutoff, i.e., with a much more realistic cutoff than the angular one. We notice that the question of the kind of cutoff is not merely a matter of convenience, but influences essentially predictions of (at least in principle) experimentally verifiable results, as, e.g., the far-field disturbance produced by an oscillating wall. Grad[3] presented results for any angular cutoff intermolecular potential, but the same potentials with a radial cutoff would yield essentially different results.

The present proof should be easily extended to two-dimensional and three-dimensional internal problems, while for external flows a preliminary study of the far field should prove expedient.

Finally we remark that, although we proved convergence of the successive approximations method in the $\mathcal{K}$ space, convergence in more usual norms, weighted with $\nu(c)f^{(0)}$, $|c_x|\ f^{(0)}$ or simply $f^{(0)}$ immediately follows, since these norms are bounded by a constant times the norm in $\mathcal{K}$.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation, Grant No. GP-6084. Earlier stages were developed at Applicazioni e Ricerche Scientifiche, Milano, Italy, and were sponsored by the Air Force Office of Scientific Research under contract AF 61(052)-881, through the European Office of Aerospace Research, OAR, U.S. Air Force.

APPENDIX A

Here we want to prove that any collision operator, which can be split according to Eq. (2.2) with K such that $\nu^{-\frac{1}{2}}K\nu^{-\frac{1}{2}}$ is a completely continuous operator in $\mathcal{H}$, can also be split according to Eq. (2.2), where $\lambda \geq 0$ and H satisfies Eq. (2.4).

In fact the complete continuity of $\nu^{-\frac{1}{2}}K\nu^{-\frac{1}{2}}$ implies its boundedness and consequently

$$|(f, Kf)| \leq k(\nu f, f) \tag{A1}$$

for some positive k. Besides it is well-known that

$$(f, Lf) \leq 0, \tag{A2}$$

i.e,

$$(f, Kf) \leq (\nu f, f). \tag{A3}$$

Equations (A1) and (A3) imply that a nonnegative value of λ exists such that

$$-\lambda(\nu f, f) < (f, Kf) \leq (\nu f, f). \tag{A4}$$

Hence, if H is defined by (2.3), Eqs. (2.4) and (2.5) follow.

APPENDIX B

Here we want to prove that

$$\|Ug\|_x \leq [1/\rho(\mathbf{c})]\ \|g\|_x, \tag{B1}$$

where the norms are defined by Eq. (2.14) and $\rho(\mathbf{c})$ by Eq. (2.10).

We have, since $\operatorname{Re} s \geq 0$:

$$\|Ug\|_x^2 = \frac{1}{|c_x|^2}\int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx \left| \int_{-\frac{1}{2}L \operatorname{sgn} c_x}^{x} \exp\left[-\frac{s + (\lambda+1)\nu(c)}{|c_x|}|x-y|\right] g(y)\, dy \right|^2$$
$$\leq \frac{1}{|c_x|^2}\int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx \left| \int_{-\frac{1}{2}d}^{x} \exp\left[-\frac{(\lambda+1)\nu(c)}{|c_x|}|x-y|\right] |g(y)|\, dy \right|^2. \tag{B2}$$

We can assume that $c_x > 0$, since no essential differences arise. Then we can put

$$\alpha = (\lambda+1)\nu(c)/|c_x|, \tag{B3}$$

$$\|Ug\|_x^2 \leq \frac{1}{|c_x|^2}\int_{-\frac{1}{2}d}^{\frac{1}{2}d} dx \int_{-\frac{1}{2}d}^{x} dy \int_{-\frac{1}{2}L}^{x} dz \exp[-\alpha|x-y|]$$
$$\times \exp[-\alpha|x-z|]\,|g(y)|\,|g(2)|$$
$$= (1/|c_x|^2)(|g|, A\,|g|)_x, \tag{B4}$$

where $(\ ,\ )_x$ denotes the inner product which generates the norm given by Eq. (2.14) and A is an integral operator with the kernel

$$A(x,y) = \int_{M(x,y)}^{\frac{1}{2}d} dz \exp[-\alpha(2z - x - y)]$$
$$= (1/2\alpha)[e^{-\alpha|y-z|} - e^{-\alpha(L-y-z)}], \tag{B5}$$

where $M(x,y) = \frac{1}{2}(x + y + |x-y| = \max(x,y)$. A is compact and self-adjoint.

Therefore

$$\|Ug\|_x^2 \leq (1/|c_x|^2)\,|\mu|\ \|g\|_x^2, \tag{B6}$$

where μ is that eigenvalue of A which is largest in absolute value. To find μ we consider the eigenvalue equation

$$A\varphi = \mu\varphi, \tag{B7}$$

which, by double differentiation, is seen to be equivalent to

$$(d^2\varphi/dx^2) + \beta^2\varphi = 0, \tag{B8}$$

$$\varphi(\tfrac{1}{2}d) = 0, \tag{B9}$$

$$\varphi(\tfrac{1}{2}d) = \alpha\varphi(-\tfrac{1}{2}d), \tag{B10}$$

where

$$\beta^2 = (1/\mu) - \alpha^2. \tag{B11}$$

The solutions of Eq. (B8) satisfying condition (B9)

have the form

$$\varphi = C \sin [\beta(\tfrac{1}{2}d - x)], \tag{B12}$$

where C is an arbitrary constant. In order that Eq. (B10) be also satisfied, we must have

$$\beta \cos \beta d + \alpha \sin \beta d = 0; \tag{B13}$$

i.e., since $\beta \neq 0$;

$$\tan \beta d/\beta d = -1/\alpha d. \tag{B14}$$

Since $\alpha d > 0$, it is easily seen (e.g., by sketching a graphic representation of $\tan x$ and $-Kx$ with $K > 0$) that there are always two real values of β in any strip $\lambda K\pi < \beta L < 2(K + 1)\pi$ $(K = \cdots, -2, -1, 0, 1, 2, \cdots)$, and the nearest to the origin are one between $-\pi$ and $-\frac{1}{2}\pi$ and another one between $\frac{1}{2}\pi$ and π. Therefore all the real solutions of Eq. (B14) satisfy

$$|\beta d| > \tfrac{1}{2}\pi. \tag{B15}$$

Now, it is easy to show that there are no complex values for β since Eq. (B14) cannot have more than two complex solutions in any strip of width 2π parallel to the imaginary axis of the βL plane. This is proved by writing

$$-\tan^{-1}(\beta d/\alpha d) = \beta d \tag{B16}$$

with some fixed determination for $\tan^{-1}$ and applying the argument principle. Therefore Eq. (B15) applies to any β and, because of (B11),

$$|\mu| = \frac{1}{\alpha^2 + \beta^2} \leq \frac{1}{\alpha^2 + \pi^2/(4d^2)}. \tag{B17}$$

Then, thanks to Eq. (B6) and (B3), we can write

$$\|Ug\|_x^2 \leq \frac{1}{|c_x|^2} \frac{1}{\dfrac{(\lambda + 1)^2[\nu(c)]^2}{|c_x|^2} + \dfrac{\pi^2}{4d^2}} \|g\|_x^2$$

$$= \frac{1}{[\rho(\mathbf{c})]^2} \|g\|_x^2, \tag{B18}$$

as was to be proved.

APPENDIX C

Here we have to prove that

$$\|Jg\|_{\mathcal{H}}^2 \leq \alpha \|g\|_{\mathcal{H}}^2 \quad (0 < \alpha < 1), \tag{C1}$$

where $J = \rho^{-\frac{1}{2}} H \rho^{-\frac{1}{2}}$. Since J is a self-adjoint operator in $\mathcal{H}$, its norm is the largest of the absolute values of the lower and upper bounds of its spectrum. Since

$$J = \rho^{-\frac{1}{2}} K \rho^{-\frac{1}{2}} + \lambda[\nu(c)/\rho(\mathbf{c})], \tag{C2}$$

if $\lambda \neq 0$, J has both a continuous and a discrete spectrum. The continuous spectrum is constituted by the values α taken by $\lambda\nu(c)/\rho(\mathbf{c})$ and therefore

$$0 \leq \alpha \leq \lambda/(\lambda + 1) < 1; \tag{C3}$$

i.e., the continuous spectrum is bounded away from 1. If we now prove that any $\alpha \geq 1$ is not an eigenvalue, i.e., we cannot have

$$\rho^{-\frac{1}{2}} K(\rho^{-\frac{1}{2}}\varphi) + \lambda[\nu(c)/\rho(\mathbf{c})]\varphi(\mathbf{c}) = \alpha\varphi(\mathbf{c}), \quad (\alpha > 1) \tag{C4}$$

for any $\varphi \in \mathcal{H}$ different from zero, then (C1) will be proved. Let us take the scalar product in $\mathcal{H}$ of φ times both sides of Eq. (C4); we have

$$(\rho^{-\frac{1}{2}}\varphi, K(\rho^{-\frac{1}{2}}\varphi)) + \lambda((\nu/\rho)\varphi, \varphi) = \alpha \|\varphi\|_{\mathcal{H}}^2. \tag{C5}$$

Now, because of Eq. (A3), the first term in the left-hand side is smaller than $(\nu\varphi/\rho, \varphi)$ and consequently

$$(\lambda + 1)((\nu/\rho)\varphi, \varphi) \geq \alpha \|\varphi\|_{\mathcal{H}}^2, \tag{C6}$$

i.e.,

$$\|\varphi\|_{\mathcal{H}}^2 \leq \frac{\lambda + 1}{\alpha}\left(\frac{\nu}{\rho}\varphi, \varphi\right) \leq (\lambda + 1)\left(\frac{\nu}{\rho}\varphi, \varphi\right) \quad (0 < \alpha \leq 1), \tag{C7}$$

where α is positive because of Eq. (2.4) of the main text and not smaller than 1 by assumption. But Eq. (C17) cannot be true, even with the equality sign, because, apart from a zero measure set, $(\lambda + 1)\nu/\rho < 1$. Therefore, for $\alpha \geq 1$, Eq. (C4) does not have any solution $\varphi \in \mathcal{H}$ which is different from zero in a set of nonzero measure, as was to be proved.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

States of Classical Statistical Mechanics

D. RUELLE
State University of New York at Stony Brook, New York and Institut des Hautes Études Scientifiques, 91, Bures-sur-Yvette, France

(Received 6 November 1966)

A state of an infinite system in classical statistical mechanics is usually described by its correlation functions. We discuss here other descriptions in particular: as (1) a state on a B^* algebra; (2) a collection of density distributions; (3) a field theory; (4) a measure on a "space of configurations of infinitely many particles." We consider the relations between these various descriptions and prove, under very general conditions, an integral representation of a state as superposition of "extremal invariant states" corresponding to pure thermodynamical phases.

1. INTRODUCTION

THE idea of representing states of physical systems by states on B^* algebras has been present for some time.[1] Actually the word "state," used to describe a normalized positive linear functional on a B^* algebra, is borrowed from physics. Recently, a number of nontrivial results of general nature have been obtained[2-7] concerning the use of B^* algebras in physics. In the present paper we apply some of these results to classical statistical mechanics.

The use of states in statistical mechanics is not new. The well-known ensembles of Gibbs correspond to states both in classical and quantum statistical mechanics. They describe, however, only systems with—essentially—a finite number of degrees of freedom. If one takes the limit of an infinite system (the thermodynamic limit), another description is needed, and is given by the correlation functions or the reduced density matrices. The correlation functions in classical statistical mechanics and the reduced density matrices in quantum statistical mechanics may be considered as "states" on algebras of unbounded operators.[8,9] Dealing with unbounded operators is, however, a serious mathematical drawback, and we here make use of a B^* algebra which may be thought of as generated by bounded functions of these unbounded operators.

We proceed by giving some motivation for the definitions and assumptions which we make below.

The systems considered in classical statistical mechanics are formed by a large number of "particles." These particles may be points in R^ν or on a lattice, or points in R^ν with a velocity vector, or more complicated objects like continuous mappings of the interval [0, 1] into R^ν. Furthermore, a system may be composed of several species of particles. There is thus a naturally defined one-particle space T. In general there is also a natural group G acting on T. Typically, G might be the Euclidean or the translation group in ν dimensions or a lattice group.

Let $T^{\hat{n}}$ be the symmetrized product of n copies of T. The sum $\hat{\mathfrak{T}}$ of the $T^{\hat{n}}$ is the space of configurations of an arbitrary finite number of particles. If we have a locally compact topology on T, we may now define states "with an essentially finite number of particles" as probability measures on $\hat{\mathfrak{T}}$. It would be natural to represent the states of statistical mechanics, which have typically an infinite number of particles, by probability measures on a new space $\mathfrak{X}$ of configurations of an infinite number of particles.

We consider, as states of classical statistical mechanics, states which are invariant under the action of G and have the property that their restriction to a compact region has an essentially finite number of particles. These states are exhibited as states on a B^* algebra $\mathfrak{A}$, actually an algebra of functions on $\hat{\mathfrak{T}}$. A large part of the paper consists in obtaining equivalent characterizations of these states and in connecting them with the correlation functions. The space $\mathfrak{X}$ mentioned above appears as a subset of the set of pure states on $\mathfrak{A}$, and this imbedding yields a natural compactification of $\mathfrak{X}$.

2. ASSUMPTIONS

It is convenient, for reasons of conciseness and generality, to axiomate that part of classical statistical

[1] See (a) I. E. Segal, *Mathematical Problems of Relativistic Physics* (American Mathematical Society, Providence, Rhode Island, 1963), and references quoted there; (b) R. Haag and D. Kastler, J. Math. Phys. 5, 848 (1964).
[2] S. Doplicher, Commun. Math. Phys. 1, 1 (1965).
[3] D. Ruelle, Commun. Math. Phys. 3, 133 (1966).
[4] S. Doplicher, D. Kastler, and D. Robinson, Commun. Math. Phys. 3, 1 (1966).
[5] D. Kastler and D. Robinson, Commun. Math. Phys. 3, 151 (1966).
[6] D. Robinson and D. Ruelle, "Extremal Invariant States," Ann. Inst. H. Poincaré (to be published)
[7] O. Lanford and D. Ruelle, J. Math. Phys., 8, 1460 (1967).
[8] D. Ruelle, J. Math. Phys. 6, 201 (1965).
[9] D. Ruelle, *Quantum Statistical Mechanics and Canonical Commutation Relations*, F. Lurcat, Ed. (Gordon and Breach Science Publishers, New York, 1967).

mechanics (CSM) in which we are interested. We call CSM theory a triple (T, G, τ) satisfying the following properties:

$(T)\cdot T$ is a locally compact space called one-particle space;

$(G)\cdot G$ is a topological group;

$(\tau)\cdot \tau$ is a homomorphism of G into the homeomorphisms of T such that the mapping $(g, x)\to \tau_g x$ of $G \times T$ onto T is continuous.

A number of the results which we derive do not depend upon the existence of a topology on G. Our assumptions are, however, not restrictive since, when G is given, the discrete topology, the continuity of $\tau: G \times T \to T$, is ensured by the fact that $\tau_g: T \to T$ is a homeomorphism. For some results we need a stronger assumption than (T), namely:

$(T')\cdot T$ is locally compact with countable basis. This means that the topology of T is generated by a countable family of open sets, and is equivalent to requiring that T is countable at infinity and that its compacts are metrizable (see Ref. 10, Secs. 8.3 and 8.19). Condition (T') is, for instance, satisfied for $T = R^\nu$.

The above axiomatic setup is, of course, more general than required by the applications, but seems to contain precisely those assumptions which we later need, so that a particularization would lead to no simplification of the arguments.

3. BASIC DEFINITIONS

In what follows, we make constant use of the topological sum $\mathfrak{T}$ of the powers T^n of T

$$\mathfrak{T} = \sum_{n\geq 0} T^n, \tag{3.1}$$

where T^0 is by definition reduced to a point. It would be natural to consider instead of T^n its quotient $T^{\hat{n}}$ by permutations and to define $\hat{\mathfrak{T}} = \sum_{n\geq 0} T^{\hat{n}}$ as indicated in the Introduction. It is, however, more convenient to work with $\mathfrak{T}$.

Let $\mathcal{K}(\mathfrak{T})$ be the space of real continuous functions with compact support on $\mathfrak{T}$, i.e., the direct sum of the spaces $\mathcal{K}(T^n)$ of real continuous functions with compact support on T^n. Let $f_1, f_2 \in \mathcal{K}(\mathfrak{T})$ and $f_1 = (f_1^n), f_2 = (f_2^n)$; we define $f_1 * f_2 \in \mathcal{K}(\mathfrak{T})$ by

$$(f_1 * f_2)^n(x_1, \cdots, x_n) = \sum_{m=0}^{n} f_1^m(x_1, \cdots, x_m) \times f_2^{n-m}(x_{m+1}, \cdots, x_n). \tag{3.2}$$

With respect to this multiplication $\mathcal{K}(\mathfrak{T})$ becomes a noncommutative algebra which we call $\mathcal{K}_*$ to avoid confusion with the structure of algebra defined on $\mathcal{K}(\mathfrak{T})$ by the usual multiplication of functions. We call **1** the identity in $\mathcal{K}_*$.

Let $f \in \mathcal{K}(T^n)$, $n > 0$, and let ω be a partition of the set $\{1, 2, \cdots, n\}$ into r subsets $S_1 = \{i_{11}, i_{12}, \cdots\}, \cdots, S_r = \{i_{r1}, i_{r2}, \cdots\}$. We may suppose that $i_{jk} < i_{jk'}$ if $k < k'$ and $i_{j1} < i_{j'1}$ if $j < j'$. For all

$$y = (y_1, \cdots, y_r) \in T^r,$$

let $x_i^w(y) = y_j$ if $i \in S_j$. We define $f_w \in \mathcal{K}(T^r)$ by

$$f_w(y_1, \cdots, y_r) = f_w(y) = f[x_1^w(y), \cdots, x_n^w(y)]. \tag{3.3}$$

The sum of the f_w over all partitions of $\{1, 2, \cdots, n\}$ is an element Δf of $\mathcal{K}_*$. For $f^0 \in \mathcal{K}(T^0)$, we write $\Delta f^0 = f^0$. Δ extends then to a linear mapping of $\mathcal{K}_*$ into itself, and one sees readily that this mapping has an inverse Δ^{-1}.

Let $\mathcal{C}(\mathfrak{T})$ be the algebra of complex continuous functions on $\mathfrak{T}$ (for the usual multiplication of functions), i.e., the product over $n \geq 0$ of the spaces $\mathcal{C}(T^n)$ of complex continuous functions on T_n. If $F = (F^n) \in \mathcal{C}(\mathfrak{T})$ and $g \in G$, we define $\tau_g F \in \mathcal{C}(\mathfrak{T})$ by

$$(\tau_g F)^n(x_1, \cdots, x_n) = F^n(\tau_{g^{-1}} x_1, \cdots, \tau_{g^{-1}} x_n). \tag{3.4}$$

In particular, it is seen that if $f_1, f_2 \in \mathcal{K}(\mathfrak{T}) \subset \mathcal{C}(\mathfrak{T})$, then $\tau_g f_1 \in \mathcal{K}(\mathfrak{T})$, $\tau_g \Delta f_1 = \Delta \tau_g f_1$, $\tau_g(f_1 * f_2) = \tau_g f_1 * \tau_g f_2$.

If $f = (f^n) \in \mathcal{K}_*$, we define $Sf \in \mathcal{C}(\mathfrak{T})$ by

$$(Sf)^0 = f^0,$$

$$(Sf)^n(x_1, \cdots, x_n) = f^0 + \sum_{m=1}^{\infty} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} f^m(x_{i_1}, \cdots, x_{i_m}) \quad (n > 0). \tag{3.5}$$

One checks readily that S is a homomorphism of $\mathcal{K}_*$ into $\mathcal{C}(\mathfrak{T})$ such that $S\tau_g f = \tau_g Sf$. Furthermore, it is easily seen that

$$(S\Delta^{-1} f)^n(x_1, \cdots, x_n) = f^0 + \sum_{m=1}^{n} {\sum_{i_1, \cdots, i_m}}' f^m(x_{i_1}, \cdots, x_{i_m}), \tag{3.6}$$

where the summation $\sum'$ extends over the $n!/(n-m)!$ sequences $(i_1, \cdots, i_m)$ of m different integers i, with $1 \leq i \leq n$.

We define $\mathfrak{A}$ to be the closure, with respect to the uniform norm, of the subalgebra of $\mathcal{C}(\mathfrak{T})$ constituted by the elements of the form $\phi(Sf_1, \cdots, Sg_q)$ for all integers $q \geq 0$, $f_1, \cdots, f_q \in \mathcal{K}_*$ and ϕ bounded continuous complex function on R^q. With respect to complex conjugation and the uniform norm, $\mathfrak{A}$ is an Abelian B^* algebra with identity 1. We note $E \subset \mathfrak{A}'$ the set of all states on $\mathfrak{A}$.

[10] N. Bourbaki, *Topologie générale, fascicule de résultats* (Hermann & Cie, Paris, 1953).

If $g \in G$, τ_g is an automorphism of $\mathfrak{A}$. If $\rho \in E$, we define $\tau'_g\rho$ by $(\tau'_g\rho)(A) = \rho(\tau_{g-1}A)$ for all $A \in \mathfrak{A}$. Let $\mathfrak{L}_G$ be the subspace of $\mathfrak{A}$ generated by the elements of the form $A - \tau_g A$ and

$$\mathfrak{L}_G^{\perp} = \{f \in \mathfrak{A}' : A \in \mathfrak{L}_G \Rightarrow f(A) = 0\}. \quad (3.7)$$

Then, the set of τ'_g-invariant states on $\mathfrak{A}$ is $E \cap \mathfrak{L}_G^{\perp}$. The sets E and $E \to \mathfrak{L}_G^{\perp}$ are both convex and weakly compact.

4. THE SET $\mathcal{F}$ OF STATES

We concern ourselves in what follows with the set $\mathcal{F} \subset E$ consisting of those states ρ on $\mathfrak{A}$ which satisfy the condition:

($\mathcal{F}$) if $f_1, \cdots, f_q \in \mathcal{K}_*$, then the functional on the bounded continuous complex functions on R^q defined by

$$\phi \to \rho(\phi(Sf_1, \cdots, Sf_q)) \quad (4.1)$$

is a measure on R^q.

To understand ($\mathcal{F}$), we interpret $\rho(\phi(Sf_1, \cdots, Sf_q))$ as an expectation value of the function $\phi(Sf_1, \cdots, Sf_q)$ on $\mathfrak{X}$. If we restrict ϕ to tend to zero at infinity, the positive linear functional $\phi \to \rho(\phi(Sf_1, \cdots, Sf_q))$ defines a positive measure on R^q which represents the probability distribution of finite values of $Sf_1, \cdots, Sf_q$. Condition ($\mathcal{F}$) means that this distribution has total mass 1, in other words, that the functions $Sf_1, \cdots, Sf_q$ take the value ∞ with probability 0.

As an example, take $q = 1$, $f_1^0 = 0$, $f_1^n = 0$ for $n > 1$ and $f_1^1 \geq 0$ not vanishing identically. Since f_1^1 has compact support and

$$(Sf_1)^n(x_1, \cdots, x_n) = \sum_{i=1}^{n} f_1^1(x_i), \quad (4.2)$$

it is seen that, for ($\mathcal{F}$) to hold, it is necessary (in fact it is also sufficient) that the probability of finding simultaneously an infinite number of particles on a compact subset of T vanishes.

We now express condition ($\mathcal{F}$) in a manner better suited to later purposes (see Proposition 4.3.2 below). Let $0 \leq h \in \mathcal{K}_*$. We denote by $\mathfrak{A}_{(h)}$ the sub-B^* algebra of $\mathfrak{A}$ generated by the elements of the form $\phi(Sh, Sf_1, \cdots, Sf_q)$ where supp $f_1, \cdots,$ supp f_q are contained in $\{X \in \mathfrak{X} : h(X) > 0\}$. We denote by $\mathfrak{I}_{(h)}$ the closed ideal of $\mathfrak{A}_{(h)}$ generated by those elements for which ϕ tends to zero at infinity.

Proposition 4.1: If supp h is metrizable, then $\mathfrak{I}_{(h)}$ is separable.

If ϕ tends to zero at infinity, we may approximate $\phi(Sh, Sf_1, \cdots, Sf_q)$ by $\tilde{\phi}(Sh, Sf_1, \cdots, Sf_q)$, where $\tilde{\phi}$ belongs to a countable family of functions with compact support on R^{q+1}. The space $\{f \in \mathcal{K}_* : \text{supp } f \subset \text{supp } h\}$ is separable (see Ref. 10, Sec. 13.27). and therefore, given $\epsilon > 0$, one may find $\tilde{f}_1, \cdots, \tilde{f}_q \in \mathcal{K}_*$ such that $|f_1 - \tilde{f}_1| < \epsilon h, \cdots, |f_q - \tilde{f}_q| < \epsilon h$, where $\tilde{f}_1, \cdots, \tilde{f}_q$ belong to a countable family of elements of $\mathcal{K}_*$ with support, in $(X \in \mathfrak{X} : h(X) > 0)$. Since $\tilde{\phi}$ has compact support, one may approximate $\tilde{\phi}(Sh, Sf_1, \cdots, Sf_q)$ by $\tilde{\phi}(Sh, S\tilde{f}_1, \cdots, S\tilde{f}_q)$.

Proposition 4.2: Let $\rho \in E$; then $\rho \in \mathcal{F}$ if and only if, for all h, the restriction of ρ to $\mathfrak{I}_{(h)}$ has norm 1.

If $\rho \in \mathcal{F}$, then, according to the comment following the statement of condition ($\mathcal{F}$), the restriction of ρ to $\mathfrak{I}_{(h)}$ has norm 1.

Conversely let the restriction of $\rho \in E$ to $\mathfrak{I}_{(h)}$ have norm 1. Given $\epsilon > 0$, there exists then $\phi(Sh, Sf'_1, \cdots, Sf'_{q'}) \in \mathfrak{I}_{(h)}$ such that ϕ has compact support, $0 \leq \phi \leq 1$ and $\rho(\phi(Sh, Sf'_1, \cdots, Sf'_{q'})) > 1 - \epsilon$. One may then choose $\phi'(Sh) \in \mathfrak{I}_{(h)}$ such that ϕ has compact support, $0 \leq \phi \leq 1$ and $\rho(\phi'(Sh)) > 1 - \epsilon$. Given $f_1, \cdots, f_q \in \mathcal{K}_*$, let $h \in \mathcal{K}_*$ be such that $h \geq 0$ and supp $f_1, \cdots,$ supp $f_q \subset (X \in \mathfrak{X} : h(X) > 0)$. One may choose $\phi''(Sf_1, \cdots, Sf_q)$ such that ϕ'' has compact support, $0 \leq \phi'' \leq 1$ and $\phi'(Sh)\phi''(Sf_1, \cdots, Sf_q) = \phi'(Sh)$; we have then

$$\rho(\phi''(Sf_1, \cdots, Sf_q)) \geq \rho(\phi'(Sh)\phi''(Sf_1, \cdots, Sf_q)) = \rho(\phi'(Sh)) \geq 1 - \epsilon. \quad (4.3)$$

Proposition 4.3: Let $(h_\iota)_{\iota \in I}$ be a family such that

(i) for each $\iota \in I$: $0 \leq h_\iota \in \mathcal{K}_*$;

(ii) for each compact $K \subset \mathfrak{X}$, there exists $\iota \in I$ such that $K \subset \{X \in \mathfrak{X} : h(X) > 0\}$. We then write $\mathfrak{A}_\iota = \mathfrak{A}_{(h_\iota)}$, $\mathfrak{I}_\iota = \mathfrak{I}_{(h_\iota)}$. We have:

(1) the union $\bigcup_\iota \mathfrak{A}_\iota$ is dense in $\mathfrak{A}$;

(2) let $\rho \in E$, then $\rho \in \mathcal{F}$ if and only if, for all $\iota \in I$, the restriction of ρ to $\mathfrak{I}_\iota$ has norm 1;

(3) if T is countable at infinity, then one can choose for (h_ι) a countable family. (1) follows from the fact that, by (ii), every $\phi(Sf_1, \cdots, Sf_q)$ belongs to some $\mathfrak{A}_\iota$.

The proof of (2) is identical to that of Proposition 4.2, except that everywhere one has to take h in the family (h_ι).

In (3) one can even choose for $(\mathfrak{A}_\iota)$ an increasing sequence (see Ref. 10, Sec. 8.19).

Corollary 4.4: If T has a countable basis, one can choose for (h_ι) a countable family and the $\mathfrak{I}_\iota$ are separable.

This follows from Proposition 4.1 and Proposition 4.3.3.

5. G SYSTEMS OF DENSITY DISTRIBUTIONS

The discussion of $(\mathcal{F})$ in Sec. 4 suggests to describe a state $\rho \in \mathcal{F}$ by giving for each relatively compact open set $\Lambda \subset T$ and each integer n the probability of finding n particles in Λ and the probability distribution of their positions.

For every relatively compact open set $\Lambda \subset T$, and integer $n \geq 0$, let $\mu_\Lambda^n \geq 0$ be a measure on $\Lambda^n \subset T^n$. We assume that μ_Λ^n is invariant under permutation of the n factors of T^n. We say that the μ_Λ^n form a G system of density distributions if they satisfy the following conditions:

(D1)
$$\mu_\phi^0(T^0) = 1. \tag{5.1}$$

(D2) Let $\Lambda \subset \Lambda'$ and $\chi_{\Lambda'-\Lambda}$ be the characteristic function of $\Lambda' - \Lambda$.

If $f^n \in \mathcal{K}(\sum_{n\geq 0} \Lambda^n) \subset \mathcal{K}_*$, then

$$\mu_\Lambda^n(f^n) = \sum_{m=0}^{\infty} \frac{(n+m)!}{n!\,m!} \mu_{\Lambda'}^{n+m}(f^n \otimes x_{\Lambda'-\Lambda}^{\otimes m}). \tag{5.2}$$

(D3) If $f^n \in \mathcal{K}(\Lambda^n)$ and $g \in G$, then

$$\mu_\Lambda^n(f^n) = \mu_{\tau_g\Lambda}^n(\tau_g f^n). \tag{5.3}$$

Notice that from (D1) and (D2) we obtain

$$\sum_n \mu_\Lambda^n(\Lambda^n) = 1. \tag{5.4}$$

Theorem 5.1: Given $\rho \in \mathcal{F} \cap \mathcal{L}_G^\perp$, there exists a unique G-system (μ_Λ^n) of density distributions such that if $f_1, \cdots, f_q \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n) \subset \mathcal{K}_*$ and ϕ is a bounded continuous complex function on R^q, then

$$\rho\big(\phi(Sf_1, \cdots, Sf_q)\big) = \sum_{n\geq 0} \int d\mu_\Lambda^n(x_1, \cdots, x_n) \times \phi\big(Sf_1(x_1, \cdots, x_n), \cdots, Sf_q(x_1, \cdots, x_n)\big). \tag{5.5}$$

The mapping $\rho \to (\mu_\Lambda^n)$ is one-to-one from $\mathcal{F} \cap \mathcal{L}_G^\perp$ onto the G systems of density distributions.

We prove here only that (5.5) determines a unique mapping of the G systems of density distributions into $\mathcal{F} \cap \mathcal{L}_G^\perp$.

Given a G system of density distributions, (5.5) defines a linear functional $\tilde{\rho}$ on expressions of the form $\phi(Sf_1, \cdots, Sf_q)$ [because of (D2) the definition does not depend on Λ]. It follows from (5.4) that $\tilde{\rho}$ is continuous with respect to the uniform norm of functions on $\mathfrak{T}$, and thus extends uniquely to a continuous linear functional ρ on $\mathfrak{A}$. Clearly ρ is positive and, by (D1), normalized, hence $\rho \in E$. By (D3), $\rho \in \mathcal{L}_G^\perp$. On the other hand, one checks easily from (5.4) and (5.5) that $(\mathcal{F})$ is satisfied. The other half of the theorem is proved in Sec. 7.

6. G-FIELD THEORIES

By definition we call G-field theory a quadruple $(\mathfrak{H}, Q, U, \Omega)$ satisfying the following conditions.

$(\mathfrak{H})$ $\mathfrak{H}$ is a complex Hilbert space.

(Q) Q is a mapping of $\mathcal{K}_*$ to self-adjoint operators in $\mathfrak{H}$ such that:

(Q1) For all $f_1, f_2 \in \mathcal{K}_*$, the spectral projections of $Q(f_1)$, $Q(f_2)$ commute;

(Q2) Q is a homomorphism in the sense that, for all $f_1, f_2 \in \mathcal{K}_*$ and $\lambda \in R$, we have

$$Q(\mathbf{1}) = 1, \tag{6.1}$$
$$Q(\lambda f_1) \supset \lambda Q(f_1), \tag{6.2}$$
$$Q(f_1 + f_2) \supset Q(f_1) + Q(f_2), \tag{6.3}$$
$$Q(f_1 * f_2) \supset Q(f_1) \cdot Q(f_2). \tag{6.4}$$

(Q3) If $0 \leq f \in \mathcal{K}_*$, then $Q(\Delta^{-1}f) \geq 0$.

(U) U is a unitary representation of G in $\mathfrak{H}$ such that, for all $g \in G$, $f \in \mathcal{K}_*$

$$U(g)Q(f)U(g^{-1}) = Q(\tau_g f). \tag{6.5}$$

(Ω) Ω is an element of $\mathfrak{H}$ such that $\|\Omega\| = 1$ and

(Ω1) for all $g \in G$, $U(g)\Omega = \Omega$.

(Ω2) Ω is cyclic with respect to Q in the sense that if $\tilde{\mathfrak{A}}$ is the C^* algebra generated by the bounded continuous complex functions of the $Q(f)$, then $\tilde{\mathfrak{A}}\Omega$ is dense in $\mathfrak{H}$.

Theorem 6.1: Given $\rho \in \mathcal{F} \cap \mathcal{L}_G^\perp$, there exists a G-field theory $(\mathfrak{H}, Q, U, \Omega)$ unique up to unitary equivalence, such that if $f_1, \cdots, f_q \in \mathcal{K}_*$ and if ϕ is a bounded continuous complex function on R^q, then

$$(\Omega, \phi\big(Q(f_1), \cdots, Q(f_q)\big)\Omega) = \rho\big(\phi(Sf_1, \cdots, Sf_q)\big). \tag{6.6}$$

The mapping $\rho \to (\mathfrak{H}, Q, U, \Omega)$ is one-to-one from $\mathcal{F} \cap \mathcal{L}_G^\perp$ onto the G-field theories defined up to unitary equivalence.

We prove here only that (6.6) determines a unique mapping from $\mathcal{F} \cap \mathcal{L}_G^\perp$ to G-field theories defined up to unitary equivalence.

Let $\rho \in \mathcal{F} \cap \mathcal{L}_G^\perp$. The Gel'fand–Segal construction yields a complex Hilbert space $\mathfrak{H}$, a * homomorphism π_ρ of $\mathfrak{A}$ into the bounded operators on $\mathfrak{H}$, a unitary representation U of G in $\mathfrak{H}$, and a vector $\Omega \in \mathfrak{H}$ such that $\|\Omega\| = 1$ and the following conditions are satisfied for all $A \in \mathfrak{A}$, $g \in G$:

$$\big(\Omega, \pi_\rho(A)\Omega\big) = \rho(A); \tag{6.7}$$
(U') $$U(g)\pi_\rho(A)U(g^{-1}) = \pi_\rho(\tau_g A); \tag{6.8}$$
(Ω1) $$U(g)\Omega = \Omega; \tag{6.9}$$
(Ω'2) $$\pi_\rho(\mathfrak{A})\Omega \text{ is dense in } \mathfrak{H}. \tag{6.10}$$

Furthermore, the Gel'fand–Segal construction is unique within unitary equivalence.

Given $f \in \mathcal{K}_*$, there exists, according to $(\mathcal{F})$ and the Appendix of Ref. 3, a self-adjoint operator $Q(f)$ such that

$$\phi(Q(f)) = \pi_\rho(\phi(Sf)) \tag{6.11}$$

for all continuous complex functions ϕ tending to zero at infinity on R. It is obvious that (Q1) is satisfied.

Let $f_1, \cdots, f_q \in \mathcal{K}_*$. Using the simultaneous spectral decomposition of $Q(f_1), \cdots, Q(f_q)$, and (6.11), we see that

$$\psi(Q(f_1), \cdots, Q(f_q)) = \pi_\rho(\psi(Sf_1, \cdots, Sf_q)) \tag{6.12}$$

if ψ is a complex continuous function tending to zero at infinity of R^q. The properties (Q2) are seen to hold on vectors of the form

$$\psi(Q(f_1), \cdots, Q(f_q))\Psi, \tag{6.13}$$

where $\Psi \in \mathfrak{H}$ and ψ has compact support. This actually proves (Q2) because the operators involved are the closure of their restriction to such vectors.

Let P be a complex polynomial on R^q; then

$$P(Q(f_1), \cdots, Q(f_q))\psi(Q(f_1), \cdots, Q(f_q))\Psi = \pi_\rho(P(Sf_1, \cdots, Sf_q)\psi(Sf_1, \cdots, Sf_q))\Psi. \tag{6.14}$$

If ϕ is a bounded continuous complex function on R^q, it can be approximated uniformly on compacts by polynomials so that P may be replaced by ϕ in the above equation, yielding

$$\phi(Q(f_1), \cdots, Q(f_q)) = \pi_\rho(Q(Sf_1, \cdots, Sf_q)). \tag{6.15}$$

Using (6.15), (6.8), and (6.10), one checks readily (U) and $(\Omega 2)$. Property (Q3) follows from the fact that, if $f \geq 0$, then $S\Delta^{-1}f \geq 0$ by (3.6). This concludes the verification of the conditions defining a G-field theory.

Finally, (6.6) follows from (6.7) and (6.15). Given $\rho \in \mathcal{F}$, we have thus proved that there exists a G-field theory satisfying (6.6). This theory is unique within unitary equivalence because of the uniqueness of the Gel'fand–Segal construction and of the uniqueness of the construction of $Q(f)$ when the $\varphi(Q(f))$ are given. The other half of the theorem is proved in Sec. 7.

7. PROOF OF THEOREMS 5.1 AND 6.1

To conclude the proof of Theorems 5.1 and 6.1, we have to show that, given a G-field theory $(\mathfrak{H}, Q, U, \Omega)$, there is a unique G-system of density distributions (μ_Λ^n) such that if $f_1, \cdots, f_q \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$ and ϕ is a bounded continuous complex function on R^q, then

$$\sum_n \int d\mu_\Lambda^n(x_1, \cdots, x_n) \times \phi(Sf_1(x_1, \cdots, x_n), \cdots, Sf_q(x_1, \cdots, x_n)) = (\Omega, \phi(Q(f_1), \cdots, Q(f_q))\Omega). \tag{7.1}$$

Let $(\mathfrak{H}, Q, U, \Omega)$ be a G-field theory and Λ be an open relatively compact subset of T. Let $h \in \mathcal{K}(T^1) \subset \mathcal{K}_*$ be such that $h \geq 0$ and $h(x) = 1$ if $x \in \Lambda$. If $f^n \in \mathcal{K}(\Lambda^n)$, the reader will check that

$$\Delta[(\Delta^{-1}f^n) * h] = f^n * h + nf^n. \tag{7.2}$$

We assume that $f^n \geq 0$; then

$$\Delta[(\Delta^{-1}f^n) * h] - nf^n = f^n * h \geq 0. \tag{7.3}$$

Hence, by (Q2), (Q3),

$$Q(\Delta^{-1}f^n)[Q(h) - n] \geq 0. \tag{7.4}$$

For every integer $p > 0$, let $\alpha_p \geq 0$ be a continuous real function with support in the closed interval $[-1, p+1] \subset R$ and such that $\alpha_p(t) = 1$ if $0 \leq t \leq p$. We assume also that $\alpha_p \leq \alpha_{p+1}$. From (Q1), (Q3) and (7.4) we obtain

$$Q(\Delta^{-1}f^n)\alpha_p(Q(h)) = 0, \quad \text{if} \quad n > p. \tag{7.5}$$

Furthermore, for all n, $Q(\Delta^{-1}f)\alpha_p(Q(h))$ is bounded because f^n is bounded by a multiple Ch^{*n} of h^{*n}, and therefore

$$0 \leq Q(\Delta^{-1}f^n) \leq CQ(\Delta^{-1}h^{*n}) \leq CQ(h^{*n}) = CQ(h)^n. \tag{7.6}$$

This shows that $[\Omega, Q(\Delta^{-1}f)\alpha_p(Q(h))\Omega]$ is a positive linear functional of $(f^0, f^1, \cdots, f^p)$. Thus there exist bounded measures $\nu^n \geq 0$ on Λ^n for $n = 0, 1, \cdots, p$ such that

$$(\Omega, Q(\Delta^{-1}f)\alpha_p(Q(h))\Omega) = \sum_{n=0}^{p} \nu^n(f^n), \tag{7.7}$$

and (Q1) implies that ν^n is symmetric in its n arguments

If f^n is assumed to be symmetric in its n arguments for $n = 0, 1, \cdots, p$ $(f^0, f^1, \cdots, f^p)$ is uniquely determined by the restrictions of $S\Delta^{-1}f$ to $\Lambda^0, \Lambda^1, \cdots, \Lambda^p$ of $S\Delta^{-1}f$; and the correspondence is such that there exist bounded measures $\mu_{(p)}^n$ on Λ^n for $n = 0, 1, \cdots, p$, symmetric in their arguments and for which

$$(\Omega, Q(\Delta^{-1}f)\alpha_p(Q(h))\Omega) = \sum_{n=0}^{p} \mu_{(p)}^n((S\Delta^{-1}f)^n). \tag{7.8}$$

We define further $\mu_{(p)}^n = 0$ for $n \geq p$. We then have, writing f instead of $\Delta^{-1}f$,

$$(\Omega, Q(f)\alpha_p(Q(h))\Omega) = \sum_{n=0}^{\infty} \int d\mu_{(p)}^n(x_1, \cdots, x_n)Sf(x_1, \cdots, x_n) = \mu_{(p)}(Sf). \tag{7.9}$$

This formula is valid for all $f \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$. Now let $f_1, \cdots, f_q \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$ and let P be a complex polynomial on R^q, we have

$$(\Omega, P(Q(f_1), \cdots, Q(f_q))\alpha_p(Q(h))\Omega) = \mu_{(p)}(P(Sf_1, \cdots, Sf_q)). \tag{7.10}$$

If P tends uniformly on the compacts to a bounded continuous complex function ϕ on R^q, this gives

$$(\Omega, \phi(Q(f_1), \cdots, Q(f_q))\alpha_p(Q(h))\Omega) = \mu_{(p)}(\phi(Sf_1, \cdots, Sf_q)). \quad (7.11)$$

We take $f^m \in \mathcal{K}(\Lambda^m)$ symmetric in its arguments, $k \in \mathcal{K}(\Lambda)$. We assume that k has values in $[0, 1]$ and that $k(x) = 1$ if $x \in K$, where K is a compact such that $\operatorname{supp} f^m \subset K^m \subset \Lambda^m$. We have

$$(\Omega, Q(\Delta^{-1}f^m)\alpha_m(Q(k))\alpha_p(Q(h))\Omega) = \sum_{n=0}^{p} \mu_{(p)}^n((S\Delta^{-1}f^m)^n\alpha_m(Sk)^n). \quad (7.12)$$

Clearly $(S\Delta^{-1}f^m)^n\alpha_m(Sk)^n$ is zero if $n < m$ by (3.6), is $m!f^m$ if $n = m$, and vanishes on K^n if $n > m$. By taking K adequately large, the terms with $n > m$ are made arbitrarily small and we obtain

$$f^m \geq 0 \Rightarrow \mu_{(p)}^m(f^m) \geq 0, \quad (7.13)$$

$$f^m \geq 0 \Rightarrow \mu_{(p+1)}^m(f^m) - \mu_{(p)}^m(f^m) \geq 0. \quad (7.14)$$

Equations (7.13) and (7.14) show that $\mu_{(p)} \geq 0$, $\mu_{(p+1)} \geq \mu_{(p)}$, and therefore

$$\lim_{p,p'\to\infty} \|\mu_{(p')} - \mu_{(p)}\| = \lim_{p,p'\to\infty} |\mu_{(p')}^{(1)} - \mu_{(p)}^{(1)}| = 0. \quad (7.15)$$

Let

$$\mu_\Lambda = \lim_{p\to\infty} \mu_{(p)};$$

(7.1) then follows from (7.11), (D1) follows from (6.1), (Ω), and (D3) follows from (U), (Ω1). Let $\Lambda' \supset \Lambda$ and $h(x) = 1$ if $x \in \Lambda'$. We define $\mu_{(p)'}$ like $\mu_{(p)}$ except for the replacement of Λ by Λ'; then, if $f \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$, we have

$$\begin{aligned}
&\sum_n \mu_{(p)}^n(Sf)^n \\
&= \sum_{n'} \mu_{(p)'}^{n'}(Sf)^{n'} \\
&= \sum_{n'} \sum_{m=0}^{n'} \frac{n'!}{(n'-m)!\,m!} \mu_{(p)'}^{n'}(\chi_\Lambda^{\otimes(n'-m)}(Sf)^{n'-m} \otimes \chi_{\Lambda'-\Lambda}^{\otimes m}) \\
&= \sum_n \sum_m \frac{(n+m)!}{n!\,m!} \mu_{(p)'}^{(n+m)}(\chi_\Lambda^{\otimes n}(Sf)^n \otimes \chi_{\Lambda'-\Lambda}^{\otimes m}). \quad (7.16)
\end{aligned}$$

Taking successively $f = \Delta^{-1}f^n$ with $f^n \in \mathcal{K}(\Lambda^n)$ and $n = p, \cdots, 1, 0$ yields

$$\mu_{(p)}^n(f^n) = \sum_m \frac{(n+m)!}{n!\,m!} \mu_{(p)'}^{n+m}(f^n \otimes \chi_{\Lambda'-\Lambda}^{\otimes m}), \quad (7.17)$$

hence (D2) when $p \to \infty$.

Finally we show that the system (μ_Λ^n) is uniquely determined by $(\mathfrak{H}, Q, U, \Omega)$. Let indeed f^m and k be as above (7.12); then

$$(\Omega, Q(\Delta^{-1}f^m)\alpha_m(Q(k)\Omega) - m!\,\mu_\Lambda^m(f^m)) = \sum_{n>m} \mu_\Lambda^m((S\Delta^{-1}f^m)^n\alpha_m(Sk)^n) \quad (7.18)$$

and the absolute value of the right hand side can be made arbitrarily small by taking K adequately large so that, if $f^m \geq 0$,

$$\mu_\Lambda^m(f^m) = (m!)^{-1} \sup_k (\Omega, Q(\Delta^{-1}f^m)\alpha_m(Q(k))\Omega) \quad (7.19)$$

8. CORRELATION FUNCTIONS

Let $(\mathfrak{H}, Q, U, \Omega)$ be a G-field theory. If Ω is contained in the intersection of the domains of all the $Q(f)$, $f \in \mathcal{K}_*$, there exist positive measures $\bar{\rho}^n$ on T^n, $n \geq 0$, such that

$$\begin{aligned}
&(\Omega, Q(\Delta^{-1}f)\Omega) \\
&= \sum_{n=0}^{\infty} \bar{\rho}^n(f^n) \\
&= \sum_{n=0}^{\infty} \int d\bar{\rho}^n(x_1, \cdots, x_n) f^n(x_1, \cdots, x_n). \quad (8.1)
\end{aligned}$$

$\bar{\rho}^n$ is invariant under permutations of its arguments and is called the *n-body correlation function*[11] of the G-field theory.

The correlation functions, when they exist, may or may not determine the G-field theory (up to unitary equivalence). It is of interest to know when they do because the information about a CSM system is usually given in terms of its correlation functions. The following criterion may be useful for this purpose.

Proposition 8.1: Suppose that for every $f^1 \in \mathcal{K}(T)$, $f^1 \geq 0$, there exists $C > 0$ such that

$$\int d\bar{\rho}^n(x_1, \cdots, x_n) f^1(x_1) \cdots f^1(x_n) \leq C^n. \quad (8.2)$$

Then, the G-field theory is determined up to unitary equivalence by $(\bar{\rho}^n)$.

We have assumed that

$$[\Omega, Q(\Delta^{-1}(f^1)^{*n})\Omega] \leq C^n. \quad (8.3)$$

Since the number of partitions of $\{1, \cdots, n\}$ is $\leq n!$ (use the cycle representation of permutations), there exists $C' > 0$ such that

$$(\Omega, Q(f^1)^n\Omega) \leq n!\,C'^n. \quad (8.4)$$

If $f \in \mathcal{K}_*$, we thus have, for some $C'' > 0$,

$$\begin{aligned}
(1/n!)\,\|Q(f^1)^nQ(f)\Omega\| &= (1/n!)(Q(f^1)^{2n}\Omega, Q(f)^2\Omega)^{\frac{1}{2}} \\
&\leq (1/n!)(\Omega, Q(f^1)^{4n}\Omega)^{\frac{1}{4}}\,\|Q(f)^2\Omega\|^{\frac{1}{2}} \\
&\leq (1/n!)[(4n)!]^{\frac{1}{4}}C'^n\,\|Q(f)^2\Omega\|^{\frac{1}{2}} < C''^n, \quad (8.5)
\end{aligned}$$

[11] This terminology originates from the situation where $T = R_\nu$ and $\bar{\rho}^n$ is absolutely continuous with respect to the Lebesgue measure on T^n, in this case $\bar{\rho}^n$ is identified to a locally integrable function. If this function is bounded by a constant C^n, where C is independent of n, then Proposition 8.1 applies.

which shows that $Q(f)\Omega$ is an analytic vector for $Q(f^1)$ (see Ref. 12), hence that $Q(f^1)$ is essentially self-adjoint on the complex Hilbert space generated by $Q(\mathcal{K}_*)\Omega$ (see Ref. 12, Lemma 5.1., and for a similar application see Ref. 13).

For arbitrary $\tilde{f} \in \mathcal{K}_*$ one may find $f^1 \in \mathcal{K}(T)$, $f^1 \geq 0$, such that $|S\tilde{f}|$ is bounded by a polynomial in Sf^1. In this case, the vectors $\alpha(Q(f^1))Q(Sf)\Omega$, where $f \in \mathcal{K}_*$ and α is complex continuous with compact support, are analytic for $Q(S\tilde{f})$, and $Q(S\tilde{f})$ is thus essentially self-adjoint on these vectors.

Therefore, if the $\bar{\rho}^n$ are known, the $Q(f)$ are known as self-adjoint operators on the complex Hilbert space generated by $Q(\mathcal{K}_*)\Omega$, and it follows from the cyclicity ($\Omega 2$) of Ω that this Hilbert space coincides with $\mathfrak{H}$, which concludes the proof.

9. PROPERTIES OF G-FIELD THEORIES

Proposition 9.1: In a G-field theory, the representation U of G is strongly continuous.

Let $f_1, \cdots, f_q \in \mathcal{K}_*$ and ϕ be a bounded continuous complex function on R^q; we write

$$\psi_g = [\tau_g\phi^*(Sf_1, \cdots, Sf_q)]\phi(Sf_1, \cdots, Sf_q) - |\phi(Sf_1, \cdots, Sf_q)|^2 \quad (9.1)$$

for $g \in G$. We have to prove that, if $\rho \in \mathcal{F}$, then

$$\lim_{g \to e} \rho(\psi_g) = 0, \quad (9.2)$$

where e is the identity of G.

We choose Λ open relatively compact in T such that $f_1, \cdots, f_q \in \mathcal{K}(\sum_{n \geq 0} \Lambda^n)$, then, given $\epsilon > 0$, there exists p such that

$$\sum_{n > p} \|\mu_\Lambda^n\| < (3\,\|\phi(Sf_1, \cdots, Sf_q)\|^2)^{-1}\epsilon. \quad (9.3)$$

Because of the continuity of $(g, x) \to \tau_g x$, there exists a neighborhood $\mathcal{N}$ of e in G, such that if $g \in \mathcal{N}$ then $\tau_g f_1, \cdots, \tau_g f_q \in \mathcal{K}(\sum_{n \geq 0} \Lambda^n)$ and

$$\max_{n \leq p} \max_{x_1, \cdots, x_n \in \Lambda} |\psi_g^n(x_1, \cdots, x_n)| < \tfrac{1}{3}\epsilon. \quad (9.4)$$

If $g \in \mathcal{N}$, (9.3) and (9.4) give $|\rho(\psi_g)| < \epsilon$, which proves (9.2) and the proposition.

Proposition 9.2: If the condition (T') is satisfied (the topology of T has a countable basis), then the Hilbert space $\mathfrak{H}$ of a G-field theory is separable.

Let ρ be the state on $\mathfrak{A}$ corresponding to the G-field theory. Let $(\mathfrak{I}_\iota)$, $(\mathfrak{J}_\iota)$ be as in Proposition 4.3. According to Corollary 4.4, we may choose these families countable and the $\mathfrak{J}_\iota$ are separable. Let now $\mathfrak{H}_\iota$ be the closed subspace of $\mathfrak{H}$ spanned by $\pi_\rho(\mathfrak{A}_\iota)\Omega$. By the uniqueness of the Gel'fand–Segal construction, the representation of $\mathfrak{A}_\iota$ in $\mathfrak{H}_\iota$ is unitarily equivalent to the representation π_{ρ_ι} constructed from the restriction ρ_ι of ρ to $\mathfrak{A}_\iota$. Since the restriction of ρ_ι to $\mathfrak{T}_\iota$ has norm 1 by 4.3.2, $\pi_\rho(\mathfrak{T}_\iota)\Omega$ is dense in $\mathfrak{H}_\iota$ [see Ref. 14, Sec. 2.11.7], hence $\mathfrak{H}_\iota$ is separable. Since the $\mathfrak{H}_\iota$ form a countable family and span $\mathfrak{H}$ by 4.3.1, $\mathfrak{H}$ is separable.

10. INTEGRAL REPRESENTATIONS ON $\mathcal{F} \cap \mathfrak{L}_G^\perp$

In this section we apply to the states in $\mathcal{F} \cap \mathfrak{L}_G^\perp$ some recent general results[3–5,7] which are summarized in the Appendix for the convenience of the reader. If $A \in \mathfrak{A}$, we denote by $\hat{A}$ the function $\rho \to \hat{A}(\rho) = \rho(A)$ on E. If K is a convex set in a topological vector space, $\mathcal{E}(K)$ denotes the set of its extremal points.

Theorem 10.1: Let (T') hold. Given $\rho \in \mathcal{F} \cap \mathfrak{L}_G^\perp$, there exists a unique measure μ_ρ on $E \cap \mathfrak{L}_G^\perp$ concentrated on $\mathcal{E}(E \cap \mathfrak{L}_G^\perp)$ such that, for all $A \in \mathfrak{A}$,

$$\rho(A) = \mu_\rho(\hat{A}). \quad (10.1)$$

The mapping $\rho \to \mu_\rho$ is one-to-one from $\mathcal{F} \cap \mathfrak{L}_G^\perp$ onto the positive measures of norm 1 on $E \cap \mathfrak{L}_G^\perp$ which are concentrated on $\mathcal{F} \cap \mathcal{E}(E \cap \mathfrak{L}_G^\perp)$.

This follows from Proposition A3.1 and Theorem A3.2 of the Appendix, using the fact (Theorem A2.3) that, since $\mathfrak{A}$ is Abelian, it is G-Abelian.

Proposition 10.2: Let $\rho \in \mathcal{F} \cap \mathfrak{L}_G^\perp$, $\mathfrak{U}$ be the unitary representation of G in the Hilbert space $\mathfrak{H}$ of the corresponding G-field theory. If P_ρ is the projection on the subspace of $\mathfrak{H}$ formed by the vectors invariant under $U(G)$, then:

(i) the measure μ_ρ introduced in Theorem 10.1 is determined by

$$\mu_\rho(\hat{A}_1, \cdots, \hat{A}_l) = (\Omega, \pi_\rho(A_1)P_\rho\pi_\rho(A_2)P_\rho \cdots \cdots P_\rho\pi_\rho(A_l)\Omega); \quad (10.2)$$

(ii) $\rho \in \mathcal{E}(E \cap \mathfrak{L}_G^\perp) \Leftrightarrow P_\rho$ is one-dimensional.

(i) and (ii) follow, respectively, from Theorem A2.2 and Proposition A2.4 of the Appendix.

Let G now be a locally compact group. A directed

[12] E. Nelson, Ann. Math. **70**, 572 (1959).

[13] H. J. Borchers and W. Zimmermann, Nuovo Cimento **31**, 1047 (1964).

[14] J. Dixmier, *Les C*—algèbres et leurs représentations* (Gauthier-Villars, Paris, 1964).

set (χ_α) of functions on G is called an $\mathcal{M}$-directed set[5] if

$$\text{(i)}\quad \chi_\alpha \geq 0, \tag{10.3}$$

$$\text{(ii)}\quad \int dg\chi_\alpha(g) = 1, \tag{10.4}$$

$$\text{(iii)}\quad \lim_{\alpha\to\infty} \int dg\, |\chi_\alpha(gg_1) - \chi_\alpha(g)| = 0, \tag{10.5}$$

where the integrations are with respect to the right Haar measure. The existence of $\mathcal{M}$-directed sets is insured if G has an invariant mean [see (A4), Appendix]. This is true in general for the groups of interest in classical statistical mechanics (the Abelian groups, the Euclidean group, etc.) and the $\mathcal{M}$-directed sets may be taken to be sequences.

Proposition 10.3[3,5]*:* With the notations of (10.2), we have

$$\text{(i)}\quad \mu_\rho(\hat{A}_1 \cdots \hat{A}_l) = \lim_{\alpha_1,\cdots,\alpha_l\to\infty} \int dg_1 \cdots \int dg_l \chi_{\alpha_1}(g_1)\cdots\chi_{\alpha_l}(g_l) \times \rho(\tau_{g_1}A_1 \cdots \tau_{g_l}A_l); \tag{10.6}$$

$$\text{(ii)}\quad \rho \in \mathfrak{E}(E \cap \mathfrak{L}_G^\perp) \Leftrightarrow \lim_{\alpha\to\infty} \int dg\chi_\alpha(g)\rho(A_1 \cdot \tau_g A_2) = \rho(A_1)\rho(A_2),$$

for all $A_1, A_2 \in \mathfrak{A}$

The representation U of G is strongly continuous by Proposition 9.1; therefore, Theorem A4.1 of the Appendix yields

$$\lim_{\alpha\to\infty} \int dg\chi(g)U(g) = P_\rho \tag{10.7}$$

strongly. In view of Proposition 10.2 (i), we prove (i) by showing that

$$\lim_{\alpha_1,\cdots,\alpha_l\to\infty} \left\| \int dg_1\chi_{\alpha_1}(g_1)\pi_\rho(\tau_{g_1}A_1)\cdots\int dg_l\chi_{\alpha_l}(g_l) \times \pi_\rho(\tau_{g_l}A_l)\Omega - P_\rho\pi_\rho(A_1)P_\rho\cdots P_\rho\pi_\rho(A_l)\Omega \right\| = 0. \tag{10.8}$$

The norm to be evaluated is majorized by a sum of l terms of the form

$$\left\| \int dg_1\chi_{\alpha_1}(g_1)\pi_\rho(\tau_{g_1}A_1)\cdots\int dg_{m-1} \times \chi_{\alpha_{m-1}}(g_{m-1})\pi_\rho(\tau_{g_{m-1}}A_{m-1}) \times \left[\int dg_m\chi_{\alpha_m}(g_m)\pi_\rho(\tau_{g_m}A_m) - P_\rho\pi_\rho(A_m)\right] \times P_\rho\pi_\rho(A_{m+1})\cdots P_\rho\pi_\rho(A_l)\Omega \right\| \leq \left(\prod_{i=1}^{m-1}\|A_i\|\right) \left\| \left[\int dg_m\chi_{\alpha_m}(g_m)U(g_m) - P_\rho\right] \times \pi_\rho(A_m)P_\rho\pi_\rho(A_{m+1})\cdots P_\rho\pi_\rho(A_l)\Omega \right\|, \tag{10.9}$$

which tend to zero in view of (10.7) when $\alpha_1, \cdots, \alpha_l \to \infty$.

(ii) results from (10.7) and Proposition 10.2 (ii).

Remark: The interpretation of the integral representation in Theorem 10.1 as a decomposition of a state into phases has been discussed in Ref. 8.

11. PURE STATES

Let X be a function from T to the integers ≥ 0 such that, for every compact $K \subset T$, X restricted to T vanishes except at a finite number of points. We call $\mathfrak{X}$ the set of all such X. Given Λ, relatively compact open set in T, and $X \in \mathfrak{X}$, there exist $n \geq 0$ and $(x_1, \cdots, x_n) \subset \Lambda^n$ such that, for all $x \in \Lambda$, $X(x)$ is the number of elements of $(x_1, \cdots, x_n)$ which are equal to x. If $f \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$, we define

$$Sf(X) = f^0 + \sum_{m=1}^{\infty}\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} f^m(x_{i_1}, \cdots, x_{i_m}). \tag{11.1}$$

Proposition 11.1: If $\sigma \in \mathcal{F} \cap \mathfrak{E}(E)$, there exists a unique $X_\sigma \in \mathfrak{X}$ such that if $f_1, \cdots, f_q \in \mathcal{K}_*$ and ϕ is a bounded continuous complex function on R^q, then

$$\sigma(\phi(Sf_1, \cdots, Sf_q)) = \phi(Sf_1(X_\sigma), \cdots, Sf_q(X_\sigma)). \tag{11.2}$$

The mapping $\sigma \to X_\sigma$ is one-to-one from $\mathcal{F} \cap \mathfrak{E}(E)$ onto $\mathfrak{X}$.

Notice that $\mathfrak{E}(E)$ is the set of extremal, i.e., pure states on the Abelian B^* algebra $\mathfrak{A}$, and is thus also the set of homomorphisms of $\mathfrak{A}$ onto the complex field. On the other hand, we may write $\mathcal{F} = \mathcal{F} \cap \mathfrak{L}_{G_0}^\perp$, where G_0 is reduced to the identity.

Let Λ_0 be a relatively compact open subset of T and $\mathfrak{A}(\Lambda)_0$ be the sub $-$ B^* algebra of $\mathfrak{A}$ generated by the bounded continuous complex functions of the Sf for $f \in \mathcal{K}(\sum_{n\geq 0}\Lambda^n)$. If $\sigma \in \mathcal{F} \cap \mathfrak{E}(E)$, then the restriction of σ to $\mathfrak{A}(\Lambda_0)$ is a homomorphism onto the complex field. Therefore, if (μ_Λ^n) is the G_0 system of density distributions associated to σ, there exists n_0 such that $\mu_{\Lambda_0}^n = 0$ for $n \neq n_0$ and $\mu_{\Lambda_0}^{n_0}$ is obtained by symmetrizing the measure $\delta_{x_1} \otimes \cdots \otimes \delta_{x_{n_0}}$ on $\Lambda_0^{n_0}$ for some $x_1, \cdots, x_{n_0} \in \Lambda$. Let $X_{(\Lambda_0)}$ be the function on Λ_0 defined by

$$X_{(\Lambda_0)}(x) = \sum_{i=1}^{n_0} f_i(x), \tag{11.3}$$

where $f_i(x_i) = 1$ and $f_i(x) = 0$ if $x \neq x_i$. We define X_σ to be such that its restriction to each Λ is X_Λ. It is then easily checked that (11.2) is satisfied; furthermore it is clear that (11.2) determines X_σ uniquely.

Let now $X \in \mathfrak{X}$, the state σ_X on $\mathfrak{A}$ defined by

$$\sigma_X(\phi(Sf_1, \cdots, Sf_q)) = \phi(Sf_1(X), \cdots, Sf_q(X)) \tag{11.4}$$

is in $\mathcal{F}$ and is a homomorphism onto the complex numbers; hence $\sigma_X \in \mathcal{F} \cap \mathcal{E}(E)$ which shows that $\sigma \to X_\sigma$ is one-to-one onto, concluding the proof.

Since $\mathcal{E}(E)$ is identical to the spectrum of $\mathfrak{A}$, the Gel'fand isomorphism associates to any $\rho \in E$ a measure $\nu_\rho \geq 0$ on $\mathcal{E}(E)$. This measure is actually the unique measure on $\mathcal{E}(E)$ with resultant ρ. If (T') holds[15] and if $\rho \in \mathcal{F}$, ν_ρ is thus identical to the measure μ_ρ of Theorem 10.1 for the case $G = G_0$; in particular, ν_ρ is concentrated on $\mathcal{F} \cap \mathcal{E}(E)$. The set $\mathfrak{X}$, identified to $\mathcal{F} \cap \mathcal{E}(E)$, is the "space of configurations of an infinite number of particles" promised in the introduction, this interpretation being justified by Proposition 11.1.

Proposition 11.2: Let (T') hold[15] and $\rho \in \mathcal{F} \cap \mathfrak{L}_G^\perp$.

(i) The mappings $\sigma \to \tau_g' \sigma$ define homeomorphisms of $\mathfrak{X}$ which leave the measure ν_ρ invariant.

(ii) If $A \in \mathfrak{A}$, let $\tilde{A}$ be the function on s' defined by $\tilde{A}(X_\sigma) = \sigma(A)$; then $L^2(\nu_\rho)$ is identified to the Hilbert space $\mathfrak{H}$ by $\tilde{A} \to \pi_\rho(A)\Omega$.

(iii) The group G acts ergodically on $(\mathfrak{X}, \nu_\rho)$ if and only if $\rho \in \mathcal{E}(E \cap \mathfrak{L}_G^\perp)$.

(i) is obvious; the identification (ii) goes via the Gel'fand isomorphism; and $\rho \in \mathcal{E}(E \cap \mathfrak{L}_G^\perp)$ is equivalent to the fact that ν_ρ has no nontrivial decomposition into two invariant measures ≥ 0, proving (iii).

12. AVERAGES OVER TRANSLATIONS

In this section we consider the case where $G = R^\nu$ is the group of translations in ν dimensions. This allows us to use a pointwise ergodic theorem.

Theorem 12.1: Let (T') hold, $G = R^\nu$, and let χ_α be the characteristic function of the cube

$$\Lambda_\alpha = \{g : 0 \leq g^i < \alpha \quad \text{for} \quad i = 1, \cdots, \nu\} \tag{12.1}$$

divided by α^ν. If $\rho \in \mathcal{E}(E \cap \mathfrak{L}_G^\perp) \cap \mathcal{F}$, then ν_ρ is concentrated on those $X \in \mathfrak{X}$ such that, for all $A \in \mathfrak{A}$,

$$\lim_{\alpha \to \infty} \int dg \chi_\alpha(g) \tau_g \tilde{A}(X) = \rho(A). \tag{12.2}$$

If a system $(\bar{\rho}^n)$ of correlation functions is associated with ρ, ν_ρ is concentrated on those $X \in \mathfrak{X}$ such that, for all $f \in \mathcal{K}_*$,

$$\lim_{\alpha \to \infty} \int dg \chi_\alpha(g) S\Delta^{-1} \tau_g f(X) = \sum_n \bar{\rho}^n(f^n). \tag{12.3}$$

[15] It is in fact sufficient to assume that T is countable at infinity.

(i) Using the identifications made in Sec. 11 (in particular, Proposition 11.2), one sees that the functions $X \to \hat{A}(X)$ (with $A \in \mathfrak{A}$) and (if the $\bar{\rho}^n$ exist) also $X \to Sf(X)$ (with $f \in \mathcal{K}_*$) are in $L^1(\nu_\rho)$. The continuity and ergodicity of the representation U of G in $L^2(\nu_\rho)$ (see Proposition 9.1 and Proposition 11.2) and a pointwise ergodic theorem (see Ref. 16, VIII. 7.17) imply then that ν_ρ is concentrated on those X such that (12.2) [or (12.3) if the $\bar{\rho}^n$ are defined] holds for a given $A \in \mathfrak{A}$ (or $f \in \mathcal{K}_*$).

(ii) Let $(\mathfrak{I}_\iota)$, $(\mathfrak{J}_\iota)$ be countable families as in Proposition 4.3. We let $(A_{\iota\lambda})$ be a countable dense set in $\mathfrak{I}_\iota$ and $(U_{\iota\lambda})$ be a countable increasing approximate identity in $\mathfrak{J}_\iota$ [see Ref. 14, Sec. 1.7]. Let $\mathfrak{U}$ be the set of those $X \in \mathfrak{X}$ such that (12.2) holds for $A = A_{\iota\lambda}$ and $A = U_{\iota\lambda}$; then ν_ρ is concentrated on $\mathfrak{U}$. Since $(A_{\iota\lambda})$ is dense in $\mathfrak{I}_\iota$, (12.2) holds for $X \in \mathfrak{U}$ and all $A \in \mathfrak{I}_\iota$. We have $\lim_{\lambda \to \infty} \rho(1 - U_{\iota\lambda}) = 0$, and every $A \in \mathfrak{A}$ may be written as $AU_{\iota\lambda} + A(1 - U_{\iota\lambda})$, where $AU_{\iota\lambda} \subset \mathfrak{J}_\lambda$ and $|\hat{A}(1 - \hat{U}_{\iota\lambda})| \leq \|A\|\ (1 - \hat{U}_{\iota\lambda})$ so that (12.2) holds for $X \in U$ and all $A \in \mathfrak{A}_\iota$. Finally, since $v_\iota \mathfrak{A}_\iota$ is dense in $\mathfrak{A}$, ν_ρ is concentrated on those $X \in \mathfrak{X}$ such that (12.2) holds for all $A \in \mathfrak{A}$.

(iii) Let (f_λ) be a countable family of elements of $\mathcal{K}_*$ such that any $f \in \mathcal{K}_*$ may be approximated uniformly on some compact by elements of (f_λ) (such a family exists because T is countable at infinity and its compacts are metrizable). Let (h_ι) be a countable family as in Proposition 4.3. If $\mathfrak{U}$ is the set of those $X \in \mathfrak{X}$ such that (12.3) holds for $f = f_\lambda$ and $f = h_\iota$, then ν_ρ is concentrated on $\mathfrak{U}$. Given $f \in \mathcal{K}_*$ and $\epsilon > 0$, there exist λ, ι such that $|f - f_\lambda| < \epsilon h_\iota$. Therefore

$$\left| \sum_n \bar{\rho}^n(f^n - f_\lambda^n) \right| < \epsilon \sum_n \bar{\rho}^n(h_\iota)$$

may be chosen arbitrarily small and (12.3) holds for all $X \in \mathfrak{U}$, $f \in \mathcal{K}_*$, concluding the proof.

Remark: Theorem 12.1 shows, in particular, that if (T') holds and $\rho \in \mathcal{E}(E \cap \mathfrak{L}_G^\perp) \cap \mathcal{F}$, there exists $X \in \mathfrak{X}$ such that, for all $f_1, \cdots, f_q \in \mathcal{K}_*$ and ϕ a bounded continuous complex function on R^q,

$$\rho(\phi(Sf_1, \cdots, Sf_q)) = \lim_{\alpha \to \infty} \int dg \chi_\alpha(g) \phi(S\tau_g f_1(X), \cdots, S\tau_g f_q(X)). \tag{11.4}$$

This is precisely the statement made in Ref. 17 except for the replacement of the condition $\rho \in \mathcal{F} \cap \mathfrak{L}_G^\perp$ by

[16] N. Dunford and J. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Part I.

[17] D. Ruelle, "A Field Theory Like Axiom System," in *Endicott House Conference (1965)*, R. Goodman and I. Segal, Eds. (M.I.T. Press, Cambridge, Mass., 1966).

the more stringent condition $\rho \in \mathcal{E}(E \cap \mathcal{L}_G^{\perp}) \cap \mathcal{F}$. The proof alluded to in Ref. 17 is rather different from that given here.

13. REMARKS AND QUESTIONS

(i) *Hardcores.* Let $T = R^\nu$; one often imposes the condition that, for some $a > 0$, the Euclidean distance between two particles be always $\geq a$. One easily sees that to impose such a condition on a state ρ on α is equivalent to requiring that ρ vanishes on a family of positive elements of $\mathfrak{A}$. If $E \cap \mathcal{U}$ is the set of states satisfying this condition, then $E \cap \mathcal{U} \subset \mathcal{F}$ and $E \cap \mathcal{L}_G^{\perp} \cap \mathcal{U}$ is a simplex (see part A5 of the Appendix). If $\rho \in E \cap \mathcal{U}$, the correlation functions of $\mathcal{U}$ are defined and describe ρ completely.

Similar remarks hold for the case where T is a lattice and two particles are forbidden to occupy the same lattice point.

(ii) *Example: state of a language.* A language with N symbols may be idealized as a state of classical statistical mechanics with $G = Z$ (the additive group of integers) and T:N copies of Z. A "hard-core" type condition must be introduced to avoid the occupation of a site by more than one symbol. The symbols may be letters, the corresponding correlation functions are well known in cryptography, or they may be words.

(iii) *Entropy per unit volume.* Let $T = R^\nu$ and $G = R^\nu$; then an entropy per unit volume $s(\rho)$ can be defined for $\rho \in \mathcal{F} \cap \mathcal{L}_G^{\perp}$ along the lines indicated in Ref. 8. It has been proved by the author (unpublished) that s is an affine upper semicontinuous function on $\mathcal{F} \cap \mathcal{L}_G^{\perp}$. It would be interesting to have a simple and more general proof of this fact, and to prove the equivalence of various definitions of the entropy per unit volume (for another definition see Ref. 9). One should be able to prove that the equilibrium state of statistical mechanics is the solution of a variational problem (involving s) under more general conditions than those given in Ref. 8. One should be able to prove the Gibbs phase rule [for almost every interaction, temperature, chemical potential, the equilibrium state is in $\mathcal{E}(E \cap \mathcal{L}_G^{\perp})$].

(iv) *The problems of evolution.* We may describe the positions and momenta of point particles by taking $T = R^\nu \times (R^\nu)^\cdot$, where $(R^\nu)^\cdot$ is the one-point compactification of R^ν. The first factor is the one-particle position space, the second factor the one-particle momentum space [the use of $(R^\nu)^\cdot$ corresponds to the fact that we want to restrict the number of particles to be finite on compacts of position space].

No nontrivial existence theorem seems to be known for the evolution of a realistic system of infinitely many particles. Probably the evolution of states can be discussed for suitable interactions[18] (cf. the stability conditions in Ref. 18 and references quoted there) and suitable states (states having finite energy per unit volume with respect to the interaction). In particular, an equilibrium state would be a fixed point for the evolution of states.

It is unclear to the author whether the evolution of an infinite system should increase its entropy per unit volume. Another possibility is that, when the time tends to $+\infty$, a state has a limit with strictly larger entropy.

(v) *The situation described in Secs. 11 and 12:* A group G acting on a space $\mathfrak{X}$ with an invariant measure ν_ρ, is the natural set up for ergodic theory; we have used only the mean and pointwise ergodic theorems, but much more could probably be done.

In quantum statistical mechanics, problems similar to those considered in this paper arise. For instance, a decomposition theorem analogous to Theorem 10.1 can be proved (see part A6 of the Appendix).

APPENDIX

(1) *Integral representations on convex compact sets.*[19-21] Let K be a convex compact set in a locally convex topological vector space. We denote by $\mathcal{C}(K)$ the space of complex continuous functions on K and by $\mathcal{E}(K)$ the set of extremal points of K.

An order $\prec$ is defined on the set $\mathcal{M}_+$ of positive measures on K by $\mu_1 \prec \mu_2 \Leftrightarrow \mu_1(\phi) \leq \mu_2(\phi)$ for all convex $\phi \in \mathcal{C}(K)$. If $\mu_1 \prec \mu_2$, then μ_1 and μ_2 have the same norm and (if this norm is 1) the same resultant. If δ_ρ is the unit mass at $\rho \in K$, $\delta_\rho \prec \mu$ means that ρ is the resultant of μ.

A measure $\mu \in \mathcal{M}_+$ is called maximal if it is maximal for the order $\prec$. For every $\mu \in \mathcal{M}_+$ there exists a maximal μ' such that $\mu \prec \mu'$.

If $\mu \in \mathcal{M}_+$ is concentrated on $\mathcal{E}(K)$, then μ is maximal. Conversely, if K is metrizable and μ is maximal, then μ is concentrated on $\mathcal{E}(K)$.

The set K is called a *simplex* if for every $\rho \in K$ there is a *unique* maximal measure $\mu_\rho \succ \delta_\rho$.

In particular, if K is metrizable and a simplex, there is a unique mapping $\rho \rightarrow \mu_\rho$ of K to the probability measures concentrated on $\mathcal{E}(K)$ such that $\delta_\rho \prec \mu_\rho$. This mapping is one-to-one onto, and μ_ρ may be considered as an integral representation of ρ on $\mathcal{E}(K)$.

(2) *G-Abelian B* algebras.* Let $\mathfrak{A}$ be a B^* algebra with identity, aut($\mathfrak{A}$) the group of its * automorphisms,

[18] M. E. Fisher and D. Ruelle, J. Math. Phys. 7, 260 (1966).
[19] N. Bourbaki, *Intégration* (Hermann & Cie, Paris, 1965), 2nd ed., Chaps. 1–4, epecially Chap. 4, Sec. 7.
[20] G. Choquet and P. A. Meyer, Ann. Inst. Fourier 13, 139 (1963).
[21] R. Phelps, *Lectures on Choquet's Theorem* (D. Van Nostrand Company, Princeton, New Jersey, 1966).

G a group, and τ a (group)homomorphism $G \rightarrow$ aut $(\mathfrak{A})$. If $g \in G$, we denote by $\tau_g : A \rightarrow \tau_g A$ the corresponding automorphism. Let $\mathfrak{L}_G$ be the subspace of $\mathfrak{A}$ generated by the elements of the form $A - \tau_g A$ with $g \in G$, $A \in \mathfrak{A}$ and let

$$\mathfrak{L}_G^{\perp} = \{f \in \mathfrak{A}' : A \in \mathfrak{L}_G \Rightarrow f(A) = 0\}.$$

$\mathfrak{L}_G^{\perp}$ is thus the space of continuous linear forms on $\mathfrak{A}$ which are invariant under the action of G. If E is the set of states on $\mathfrak{A}$, $E \cap \mathfrak{L}_G^{\perp}$ is the set of G-invariant states on $\mathfrak{A}$.

For $\rho \in E$, the Gel'fand–Segal construction yields a complex Hilbert space $\mathfrak{H}_\rho$, a representation π_ρ of $\mathfrak{A}$ in $\mathfrak{H}_\rho$, and a normalized vector $\Omega_\rho \in \mathfrak{H}_\rho$, cyclic with respect to $\pi_\rho(\mathfrak{A})$ and such that for all $A \in \mathfrak{A}$

$$\rho(A) = (\Omega_\rho , \pi_\rho(A)\Omega_\rho).$$

If $\rho \in E \cap \mathfrak{L}_G^{\perp}$, there is a unique unitary representation U_ρ of G in $\mathfrak{H}_\rho$ such that for all $g \in G$, $A \in \mathfrak{A}$

$$U_\rho(g)\Omega_\rho = \Omega_\rho , \quad U_\rho(g)\pi_\rho(A)U_\rho(g^{-1}) = \pi_\rho(\tau_g A).$$

Let P_ρ be the projection on the subspace of $\mathfrak{H}_\rho$ formed by the vectors invariant under $U_\rho(G)$.

Definition A2.1: $\mathfrak{A}$ is said to be G-Abelian if, for all $\rho \in E \cap \mathfrak{L}_G^{\perp}$, the von Neumann algebra generated by $P_\rho \pi_\rho(\mathfrak{A}) P_\rho$ is Abelian (in other words, if A_1, $A_2 \in \mathfrak{A}$, then $[P_\rho \pi_\rho(A_1)P_\rho , P_\rho \pi_\rho(A_2)P_\rho] = 0$).

Theorem A2.2: If $\mathfrak{A}$ has an identity and is G-Abelian, then $E \cap \mathfrak{L}_G^{\perp}$ is a simplex and the unique maximal measure μ_ρ with resultant $\rho \in E \cap \mathfrak{L}_G^{\perp}$ is determined by

$$\mu_\rho(\hat{A}_1 \cdots \hat{A}_l) = (\Omega_\rho , \pi_\rho(A_1)P_\rho \, \pi_\rho(A_2)P_\rho \cdots P_\rho \, \pi_\rho(A_l)\Omega_\rho).$$

Theorem A2.3[22]: If, for each $\rho \in E \cap \mathfrak{L}_G^{\perp}$, there exists a filter $\mathcal{F}$ on G such that for all A_1, $A_2 \in \mathfrak{A}$

$$\lim_{\mathcal{F}} \rho([A_1 , \tau_g A_2]) = 0,$$

then $\mathfrak{A}$ is G-Abelian, This is true:

(i) if $E \cap \mathfrak{L}_G^{\perp}$ is empty;

(ii) if $\mathfrak{A}$ is Abelian;

(iii) if $\mathfrak{A}$ is asymptotically Abelian,[23] i.e., G is locally compact noncompact and for all A_1, $A_2 \in \mathfrak{A}$

$$\lim_{g \rightarrow \infty} \| [A_1 , \tau_g A_2] \| = 0.$$

[22] A good characterization of G-Abelian algebras is given in Ref. 7.

[23] This terminology was introduced by Doplicher, Kastler, and Robinson (see Ref. 4).

Proposition A2.4: Let $\mathfrak{A}$ be G-Abelian and $\rho \in E \cap \mathfrak{L}_G^{\perp}$, then $\rho \in \mathfrak{E}(E \cap \mathfrak{L}_G^{\perp}) \Leftrightarrow P_\rho$ is one-dimensional.

(3) *Integral representations of G-invariant states.*

Proposition A3.1: Let $\mathfrak{A}$ be a B^* algebra with identity and $(\mathfrak{B}_\alpha)$ a countable family of self-adjoint subalgebras of $\mathfrak{A}$; define

$\mathcal{F} = \{\sigma \in E$: the restriction of σ to $\mathfrak{B}_\alpha$ has norm 1 for all $\alpha\}$.

If μ is a positive measure of norm 1 with resultant ρ on $\mathfrak{A}$, then

$$\rho \in \mathcal{F} \Leftrightarrow \mu \text{ is concentrated on } \mathcal{F}.$$

The proof is essentially that of Part 4 of the theorem in Ref. 3.

Theorem A3.2: Let $\mathfrak{A}$ be a B^* algebra and $(\mathfrak{A}_\alpha)$ a countable family of sub-B^* algebras of $\mathfrak{A}$ such that $\mathfrak{U}_\alpha \mathfrak{A}_\alpha$ is dense in $\mathfrak{A}$. Let $\mathfrak{I}_\alpha$ be a separable closed two-sided ideal for each α and define

$\mathcal{F} = \{\sigma \in E$: the restriction of σ to $\mathfrak{I}_\alpha$ has norm 1 for all $\alpha\}$.

(i) If $\rho \in \mathcal{F}$, the Hilbert space $\mathfrak{H}_\rho$ of the Gel'fand–Segal construction is separable.

(ii) If $E \cap \mathfrak{L}_G^{\perp}$ is a simplex (in particular, if $\mathfrak{A}$ has an identity and is G-Abelian) and if the positive measure μ of norm 1 on $E \cap \mathfrak{L}_G^{\perp}$ has resultant $\rho \in \mathcal{F}$, then

μ maximal on $E \cap \mathfrak{L}_G^{\perp}$ $\Leftrightarrow \mu$ concentrated on $\mathfrak{E}(E \cap \mathfrak{L}_G^{\perp})$.

The proof of (i) is essentially that of our Proposition 9.2 and the proof of (ii) is essentially that of Part 5 of the theorem in Ref. 3.

(4) *Groups with an invariant mean.*[24] Let G be a locally compact group and $\mathfrak{C}_B(G)$ be the Abelian B^* algebra of bounded continuous complex functions on G. If $f \in \mathfrak{C}_B(G)$, we denote by f_g the right translate of f by $g \in G$. A state $\mathcal{M}$ on $\mathfrak{C}_B(G)$ is called a right-invariant mean if, for all $g \in G$, $f \in \mathfrak{C}_B(G)$,

$$\mathcal{M}(f_g) = \mathcal{M}(f).$$

If there exists a right-invariant mean on G, there also exists a left-invariant mean and a two-sided invariant mean, one says then that G is a group with an invariant mean.

[24] Information about groups with an invariant mean is conveniently collected in J.-P. Pier, "Sur une classe de groupes localement compacts remarquables du point de vue de l'analyse harmonique," thesis, Nancy (1965) (unpublished) which we have used as a source for the indications given here.

G is a group with an invariant mean if it is Abelian, or compact, or admits a composition series consisting of such groups.

One can prove that G has an invariant mean if and only if for every $\epsilon > 0$ and compact $K \subset G$ there exists a function χ on G such that

(i) $\chi \geq 0$,

(ii) $\int dg\chi(g) = 1$,

(iii) $\int dg\, |\chi(gg_1) - \chi(g)| < \epsilon$, if $g_1 \in K$,

where the integrations are with respect to the right Haar measure. In that case let $\chi_{(K,\epsilon)}$ be such a function, the family $(\chi_{(K,\epsilon)})$ is a $\mathcal{M}$-directed set (see Sec. 10) for the order.

$$(K, \epsilon) \lesssim (K', \epsilon') \Leftrightarrow K \subset K', \quad \epsilon \geq \epsilon'$$

of the indices. If the topology of G has a countable basis, there is a subsequence of $(\chi_{(K,\epsilon)})$ which is a $\mathcal{M}$-directed set.

Theorem A4.1: Let (χ_α) be a $\mathcal{M}$-directed set on G, U a strongly continuous unitary representation of G in a complex Hilbert space $\mathfrak{H}$, and P the projection on the subspace of $\mathfrak{H}$ formed by the vectors invariant under U; then

$$\lim_{\alpha \to \infty} \int dg\chi_\alpha(g)U(g) = P, \quad \text{strongly.}$$

This is a mean ergodic theorem (see Ref. 5 for a proof in the case $G = R^\nu$).

(5) *States vanishing on positive elements.* Let $\mathfrak{A}$ be a B^* algebra with an identity, $\mathfrak{A}_1$ a sub-B^* algebra of $\mathfrak{A}$. A state ρ on $\mathfrak{A}$ vanishes on $\mathfrak{A}_1$, if and only if it vanishes on the positive elements of $\mathfrak{A}_1$. Let (A_λ) be a family of positive elements in $\mathfrak{A}$ and let

$$\mathfrak{V} = \{f \in \mathfrak{A}' : f(A_\lambda) = 0 \text{ for all } \lambda\}.$$

If $\rho \in E$ and μ is a positive measure on E such that $\mu \succ \delta_\rho$, then

$$\rho \in \mathfrak{V} \Leftrightarrow \operatorname{supp} \mu \subset \mathfrak{V}.$$

In particular, if $\mathfrak{A}$ is G-Abelian, $E \cap \mathfrak{L}_G^\perp \cap \mathfrak{V}$ is a simplex.

(6) *States of Quantum Statistical Mechanics.* For each Lebesgue-measurable set $\Lambda \subset R^\nu$, let $\mathcal{H}(\Lambda)$ be the Fock space of the canonical commutation relations constructed with the real square integrable functions on Λ as test functions. If $\Lambda_1 \cap \Lambda_2 = \phi$, one may, in a natural manner, identify $\mathcal{H}(\Lambda)$ to $\mathcal{H}(\Lambda_1) \otimes \mathcal{H}(\Lambda_2)$. It is natural[9,25] to identify the states occurring in quantum statistical mechanics to collections $(\rho(\Lambda))$ where $\rho(\Lambda)$ is a density matrix on $\mathcal{H}(\Lambda)$, and if $\Lambda_1 \cap \Lambda_2 = \phi$,

$$\rho(\Lambda_1) = \operatorname{Tr}_{\mathcal{H}(\Lambda_2)} \rho(\Lambda_1 \cup \Lambda_2).$$

We furthermore require the invariance of $\mathcal{H}(\Lambda)$ under the group G of translations (or the Euclidean group) of R^ν.

Let now Λ_n be the sphere of radius n around the origin, $\mathfrak{A}_n$ the algebra of all bounded operators on $\mathcal{H}(\Lambda_n)$ identified to a subalgebra of $\mathcal{H}(R^\nu)$, $\mathfrak{I}_n$ the ideal of $\mathfrak{A}_n$ consisting of the compact operators. Let $\mathfrak{A}$ be the C^* algebra on $\mathcal{H}(R^\nu)$ generated by the $\mathfrak{A}_n$. There is a one-to-one correspondence between families $(\rho(\Lambda))$ and the set $\mathcal{F}$ of states on $\mathfrak{A}$ with restrictions of norm 1 to each $\mathfrak{I}_n$. By Theorem A2.3 (iii), $\mathfrak{A}$ is G-Abelian and therefore by Theorem A3.2 the states in $\mathcal{F} \cap \mathfrak{L}_G^\perp$ have a unique integral representation on $\mathfrak{E}(E \cap \mathfrak{L}_G^\perp) \cap \mathcal{F}$. Furthermore the Hilbert space of their Gel'fand–Segal construction is separable and one can see that the corresponding unitary representation of G is strongly continuous.

[25] G. F. Dell'Antonio, S. Doplicher, and D. Ruelle, Commun. Math. Phys. 2, 223 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Application of the Theory of Padé Approximants to the Solution of the N/D Equations*

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

(Received 12 October 1966)

Approximate methods for the solution of the N/D equations are discussed. Two methods in particular are considered, one being the replacement of the unphysical singularities of the scattering amplitude by a series of poles and the other being a pole approximation to the spectral integral of the kinematical factor over the physical region. It is shown that Padé approximants may be used to define a sequence of the above approximate solutions, which converges to the exact solution of the N/D equations in those cases when the Fredholm method also gives a solution.

1. INTRODUCTION

PARTIAL wave dispersion relations play an important part in the study of the strong interactions of elementary particles. The usual problem is to solve these relations for the partial wave-scattering amplitude, given the unphysical singularities and inelasticities. The simplest way of doing this is to use the N/D method,[1] since the scattering amplitude will then satisfy automatically the unitary condition for physical energies. The dispersion relation is then equivalent to a pair of coupled integral equations in the functions N and D, from which either N or D can be eliminated to give a simple integral equation in the other function. It is in general not possible to solve this integral equation exactly and so approximate methods have to be used.

If the integral equation is not singular, then the methods of Fredholm[2] can be used. However, as discussed by Hamilton,[3] in many physically interesting problems the equation is singular and then, although the Fredholm method may still give a solution, other methods are often tried. The method usually followed, is to approximate the kernel of the integral equation by a degenerate kernel either by assuming that the unphysical singularities are poles[4] or by making a pole approximation to the spectral integral of the kinematical factor over the physical region.[5]

It is the purpose of this paper to show how, using the theory of Padé approximants,[6] a sequence of such degenerate kernels may be defined which converge uniformly to the given nondegenerate kernel, and then to investigate the convergence of the corresponding approximate solutions of the N/D equations to the exact solution. In fact, it is proved that the approximate solutions of the integral equation converge to the exact solution in those cases where the Fredholm method also gives a solution.

In Sec. 2 the N/D equations are introduced and the reduction to a single integral equation is given. The necessary and sufficient conditions for this equation to be nonsingular are discussed. The approximation of the kernel by degenerate kernels is described in Sec. 3 and the corresponding approximate solutions of the N/D equations are given. In Sec. 4, sequences of kernels which tend uniformly to the exact kernel of the integral equation in the two cases are obtained.

In Sec. 5 it is proved that the corresponding approximate solutions of the integral equation tends uniformly to the exact solution when the usual Fredholm method also gives a solution. This occurs when the Fredholm denominator is different from zero. Finally, in Sec. 6, the results of this paper are discussed.

2. THE INTEGRAL EQUATIONS FOR N AND D

Consider the partial wave amplitude $f_l(s)$ for the scattering of two equal scalar particles of mass μ, where s is the total centre of mass energy squared. If it is assumed that the scattering amplitude has no CDD poles, then it may be decomposed as follows,

$$f_l(s) = N_l(s)/D_l(s), \tag{2.1}$$

where $N_l(s)$ has the unphysical cut, $-\infty < s \leq s_1$ and $D_l(s)$ has only the physical cut, $4\mu^2 \leq s < \infty$ of $f_l(s)$. The unitarity condition on the right-hand cut gives

$$\text{Im } D_l(s) = -N_l(s)R_l(s)\rho_l(s), \tag{2.2}$$

where $R_l(s)$ is the inelasticity parameter and $\rho_l(s)$ is a

* Part of this work was done while the author was at Trinity College, Dublin, Ireland.

[1] G. F. Chew and S. Mandelstam, Phys. Rev. **119**, 467 (1960).

[2] See, for example, R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. 1.

[3] J. Hamilton, *Strong Interactions and High-Energy Physics*, R. G. Moorhouse, Ed. (Oliver and Boyd, London, 1964), p. 344.

[4] A. W. Martin, Phys. Rev. **135**, B967 (1964).

[5] H. Pagels, Phys. Rev. **140**, B1599 (1965).

[6] A review of the properties of Padé approximants and a list of references is given by G. A. Baker, J. Advan. Theoret. Phys. **1**, 1 (1965).

kinematical factor depending on the normalization of the scattering amplitude. In the following work it is assumed that the scattering is completely elastic, and then $R_l(s) = 1$. However, the results are unchanged if $R_l(s)$ is allowed to take on different values so long as it remains finite. The normalization of the scattering amplitude is chosen so that

$$\rho_l(s) = [(s - 4\mu^2)/s]^{\frac{1}{2}}. \tag{2.3}$$

The functions $N_l(s)$ and $D_l(s)$ then satisfy the following equations:

$$D_l(s) = 1 - \frac{(s - s_0)}{\pi}\int_R \frac{\rho_l(x)N_l(x)\,dx}{(x - s_0)(x - s)}, \tag{2.4}$$

$$N_l(s) = \frac{1}{\pi}\int_L \frac{\operatorname{Im} f_l(x)D_l(x)}{(x - s)}\,dx, \tag{2.5}$$

where $s = s_0$ is an arbitrary subtraction point and $D(s)$ is normalized to unity at this point. The exact solution of (2.4) and (2.5) for the scattering amplitude is independent of the value of s_0,[4] and so for convenience it is put equal to zero. It is now required to solve (2.4) and (2.5), given the function $\operatorname{Im} f_l(s)$ for s on the left-hand cut. Eliminating either $N_l(s)$ or $D_l(s)$, the following integral equations are obtained[4,5]:

$$\frac{D(s)}{s} = \frac{1}{s} + \frac{1}{\pi}\int_L \frac{x[F(x) - F(s)]\operatorname{Im} f(x)}{(x - s)}\left[\frac{D(x)}{x}\right]dx, \tag{2.6}$$

$$N(s) = B(s) + \frac{1}{\pi}\int_R \frac{\rho(x)N(x)}{x(x - s)}[xB(x) - sB(s)]\,dx, \tag{2.7}$$

where

$$F(s) = \frac{s}{\pi}\int_R \frac{\rho(x)}{x^2(x - s)}\,dx \tag{2.8}$$

$$B(s) = \frac{1}{\pi}\int_L \frac{\operatorname{Im} f(x)\,dx}{x - s} \tag{2.9}$$

and the suffix l has dropped from the quantities in the above equations, as it is now to be understood that a particular value of the angular momentum is being considered.

Consider first of all the integral equation (2.6). The kernel of this equation is

$$K(s, x) = [x/\pi(x - s)][F(x) - F(s)]\operatorname{Im} f(x) \tag{2.10}$$

and it is well behaved at $x = s$ as long as $dF(x)/dx$ exists. The full Fredholm theory may be applied to (2.6) only if

$$\int_L\int_L ds\,dx\,|K(s, x)|^2 < \infty, \tag{2.11}$$

and this condition necessitates that

$$\int_L |\operatorname{Im} f(x)|^2\,dx < \infty. \tag{2.12}$$

If $\operatorname{Im} f(x) = O(x^\delta)$ as $x \to -\infty$, then, for (2.12) to be true, $\delta < -\frac{1}{2}$ when $-\frac{1}{2} \leq \delta < 0$; the dispersion relations exist but (2.11) is not satisfied. The condition $\delta < -\frac{1}{2}$ is both necessary and sufficient for (2.11) to be true.

With the same asymptotic behavior of $\operatorname{Im} f(x)$, it may be proved that the symmetric form of (2.7) is Fredholm for all $\delta < 0$.[7] This equation may thus be solved using Fredholm's methods for all cases considered here. However, it may be more convenient in a particular situation to use an approximate degenerate kernel by replacing the unphysical singularities by poles. For completeness the behavior of the corresponding approximate solutions are also considered in this paper.

3. APPROXIMATE SOLUTIONS USING DEGENERATE KERNELS

To reduce the kernel of (2.6) to degenerate form, Pagels[5] approximated the function $F(s)$ for values of s on the left-hand cut by the expression

$$F(s) \simeq \sum_{r=1}^{N} \frac{c_r}{(s - a_r)}, \tag{3.1}$$

where c_r, a_r are constants and the poles of the approximate form lie on the right-hand cut. The kernel (2.10) is then given approximately by

$$K(s, x) \simeq \frac{x}{\pi(x - s)}\sum_{r=1}^{N} c_r\left[\frac{1}{x - a_r} - \frac{1}{s - a_r}\right]\operatorname{Im} f(x)$$
$$= \frac{-x}{\pi}\sum_{r=1}^{N}\frac{c_r}{(x - a_r)(s - a_r)}\operatorname{Im} f(x) \tag{3.2}$$

and so is degenerate. If $d(s)$ is the corresponding approximate solution for $D(s)$, then, from (2.6) for s on the left-hand cut,

$$d(s) = 1 - s\sum_{r=1}^{N}\left[\frac{c_r}{(s - a_r)}\right]n(a_r), \tag{3.3}$$

where $n(s)$ is the corresponding approximate solution for $N(s)$. Then substituting in (2.5),

$$n(s) = B(s) - \sum_{r=1}^{N}\frac{[sB(s) - a_rB(a_r)]c_rn(a_r)}{(s - a_r)}, \tag{3.4}$$

which is true for all s. Finally, from (2.6), for s on the right-hand cut,

$$d(s) = 1 + s\sum_{r=1}^{N}\frac{c_rn(a_r)}{(s - a_r)} + s\left[F(s) - \sum_{r=1}^{N}\frac{c_r}{(s - a_r)}\right], \tag{3.5}$$

[7] D. Atkinson, J. Math. Phys. 7, 1607 (1966).

where $F(s)$ is the exact function as given by (2.8). The values of $n(s)$ when $s = a_r$, occurring in the above equations, are obtained by setting $s = a_1, s = a_2, \cdots, s = a_N$ in turn, in (3.4), and then solving the resulting N linear equations.

The approximate solution of the N/D equations by replacing the unphysical singularities by a series of poles is now considered. The kernel of (2.7) is made degenerate by the approximation of $B(s)$ for s on the right-hand cut by[4]

$$B(s) \simeq \sum_{r=1}^{N} \frac{c_r}{(s - a_r)}. \tag{3.6}$$

Then substituting in Eq. (2.7),

$$n(s) = \sum_{r=1}^{N} \frac{c_r\, d(a_r)}{(s - a_r)} \tag{3.7}$$

for s on the right-hand cut, where once again $n(s)$ and $d(s)$ are the approximations to $N(s)$ and $D(s)$. Then for all s,

$$d(s) = 1 + s \sum_{r=1}^{N} \frac{c_r\, d(a_r)}{(s - a_r)} [F(a_r) - F(s)]. \tag{3.8}$$

Finally, for s on the left-hand cut,

$$n(s) = B(s)\, d(s) - \sum_{r=1}^{N} \frac{c_r}{(s - a_r)} [d(s) - d(a_r)]. \tag{3.9}$$

The constants $d(a_r)$ can be obtained as previously from (3.8).

4. UNIFORMLY CONVERGENT APPROXIMATIONS TO THE KERNELS

Consider the kernel of (2.6):

$$K(s, x) = \frac{x}{\pi(x - s)} [F(x) - F(s)] \operatorname{Im} f(x). \tag{2.10}$$

It has been shown in (3.2) that this kernel becomes degenerate if the approximation

$$F(s) \simeq \sum_{r=1}^{N} \frac{c_r}{(s - a_r)} \tag{3.1}$$

is made. In the following, it is shown that a sequence of approximations of the form (3.1) may be defined which tend uniformly to $F(s)$ for s on the left-hand cut. Now

$$F(s) = \frac{s}{\pi} \int_R \frac{\rho(x)}{x^2(x - s)}\, dx = \frac{-1}{\pi} \int_R \frac{\rho(x)}{x^2}\, dx + \frac{1}{\pi} \int_R \frac{\rho(x)}{(x - s)x}\, dx. \tag{2.8}$$

Make the substitutions

$$s = 4\mu^2 w/(1 + w), \quad x = 4\mu^2/(1 - y) \tag{4.1}$$

in the integral on the right-hand side of (2.8). Then

$$F(s) = \frac{(1 + w)}{4\mu^2\pi} \int_0^1 \frac{\rho[4\mu^2/(1 - y)]}{1 + wy}\, dy - \frac{1}{\pi} \int_R \frac{\rho(x)}{x^2}\, dx. \tag{4.2}$$

Now the function

$$G(w) = \frac{1}{4\mu^2\pi} \int_0^1 \frac{\rho[4\mu^2/(1 - y)]}{1 + wy}\, dy, \tag{4.3}$$

when expanded as a power series in w, is a series of Stieltjes,[8] since the numerator of the integrand in (4.3) is nonnegative. This power series expansion is

$$G(w) = \sum_{i=0}^{\infty} (-w)^i g_i,$$

$$\text{where} \quad g_i = \frac{1}{4\mu^2\pi} \int_0^1 \rho\left[\frac{4\mu^2}{1 - y}\right] y^i\, dy, \tag{4.4}$$

and it converges for $|w| < 1$.

The Padé approximant to this series is now defined. The $[N, M]$ Padé approximant is of the form of one polynomial $P(w)$ of degree M divided by another polynomial $Q(w)$ of degree N. The coefficients of the two polynomials are determined uniquely by equating like powers of w in the following equation:

$$G(w)Q(w) - P(w) = Aw^{M+N+1} + Bw^{M+N+2} + \cdots, \tag{4.5}$$

where

$$Q(0) = 1.0.$$

Baker has written a review article[6] on the properties of these approximants. The following theorems in that article are used here.

Theorem 1: If $f(w)$ is a series of Stieltjes with radius of convergence R, then the corresponding sequence of $[N, N + j]$ Padé approximates converge to $f(w)$ as $N \to \infty$, with constant $j \geq -1$ in the cut plane $(-\infty < w \leq -R)$. The convergence is uniform in any closed region of this cut plane.

Theorem 2: The poles of the $[N, N + j]$ Padé approximant to $f(w)$ lie on the cut $(-\infty < w \leq -R)$ and all the corresponding residues are positive.

In the case considered here $R = 1$, and so the $[N, N + j]$ Padé approximants converge to $G(w)$ in the cut plane and, in particular, for $-1 < w \leq 0$, which corresponds to s real and negative. The best approximant $[N, M]$ to $G(w)$ for a given odd value of $M + N$

[8] T. J. Stieltjes, Ann. Fac. Sci. Univ. Toulouse Sci. Math. Sci. Phys. 8, 9, 1 (1894).

for these values of w is of the form[9] $[N, N-1]$ and this one is used to obtain an approximate form for $F(s)$ on the left-hand cut in the s plane. Then

$$G(w) \simeq [N, N-1] = \sum_{r=1}^{N} \frac{\delta_r}{(\alpha_r + w)}, \quad (4.6)$$

where, from the second theorem quoted above, $\delta_r > 0$, $\alpha_r > 1$. Substituting for w in terms of s, the corresponding approximate form for $F(s)$ is

$$F(s) \simeq (1+w) \sum_{r=1}^{N} \frac{\delta_r}{(\alpha_r + w)} = \sum_{r=1}^{N} \frac{c_r}{(s - a_r)}, \quad (4.7)$$

where

$$c_r = 4\mu^2 \delta_r/(1-\alpha_r) < 0$$
$$\text{and} \quad a_r = 4\mu^2 \alpha_r/(\alpha_r - 1) > 0. \quad (4.8)$$

In (4.7) the constant term

$$\frac{1}{\pi} \int \frac{\rho(x)}{x^2}\, dx$$

has been omitted, but the kernel of (2.6) is unaltered by this as is obvious from (2.10).

From Theorems 1 and 2 the approximate form (4.7) converges uniformly to $F(s)$ as $N \to \infty$ for values of s on the left-hand cut in the s plane, and all its poles lie on the right-hand cut. It should be noted that δ_r and α_r in (4.6) depend on N, and so all the poles and residues of the approximate form (4.7) to $F(s)$ change as the number of poles is changed, and this applies also to the kernel (2.10).

The approximate for $B(s)$,

$$B(s) \simeq \sum_{r=1}^{N} \frac{c_r}{(s - a_r)}, \quad (3.10)$$

is obtained in almost the same way. Now

$$B(s) = \frac{1}{\pi} \int_{-\infty}^{-s_1} \frac{\operatorname{Im} f(x)}{(x - s)}\, dx. \quad (2.9)$$

Make the substitutions

$$s = -s_1 w/(1+w), \quad x = -s_1/(1-y). \quad (4.9)$$

Then (2.9) becomes

$$B(s) = -\frac{(1+w)}{\pi} \int_0^1 \frac{\operatorname{Im} f[-s_1/(1-y)]}{(1-y)(1+wy)}\, dy. \quad (4.10)$$

The integral on the right-hand side is not a series of Stieltjes since $\operatorname{Im} f(x)$ may take both positive and negative values on the left-hand cut in the s plane. However, writing (4.10) in the form

$$B(s) = \frac{(1+w)}{2\pi} \int_0^1 \frac{\{|\operatorname{Im} f| - \operatorname{Im} f\}}{(1-y)(1+wy)}\, dy - \frac{(1+w)}{2\pi} \int_0^1 \frac{\{|\operatorname{Im} f| + \operatorname{Im} f\}}{(1-y)(1+wy)}\, dy, \quad (4.11)$$

each of the integrals is the sum of a series of Stieltjes, and so they may be approximated as previously. The approximant converges uniformly for all values of s in any closed interval on the right-hand cut in the s plane, and the poles of the approximants lie on the left-hand cut.

5. CONVERGENCE OF APPROXIMATE SOLUTIONS OF THE N/D EQUATIONS TO THE EXACT SOLUTION

It follows from the last section that the approximate kernels of the integral equations (2.6), (2.7) tend uniformly to the exact kernels over finite ranges of the relevant variables. It is proved in this section that the corresponding, approximate solutions of (2.6), (2.7) converge to the exact solution. It is first of all shown that the approximate solutions tend to a limit, and then it is a simple matter to prove that this is the exact solution of the corresponding integral equation.

Now, as stated previously, the constants c_r and a_r appearing in the approximate solutions (3.3)–(3.5) and (3.7)–(3.9) depend on the order of the approximation. Therefore it would be difficult to prove the convergence of the approximate solutions directly. Instead the standard methods of Fredholm[10] are used to do this, it only being necessary to check that at each step the integrals occurring converge. As a preliminary, the approximate solutions are rederived using the standard procedures for solving an integral equation with degenerate kernel. Consider Eq. (2.6) and let

$$\phi_N(s) = \frac{d(s)}{s}; \quad g(s) = \frac{1}{s}, \quad \alpha_r(s) = \frac{c_r}{(s - a_r)}, \quad \beta_r(x) = \frac{-x \operatorname{Im} f(x)}{\pi(x - a_r)}. \quad (5.1)$$

The approximate form for the kernel is then

$$k_N(s, x) = \sum_{r=1}^{N} \alpha_r(s) \beta_r(x), \quad (5.2)$$

and the approximate form for (2.6) is

$$g(s) = \phi_N(s) - \lambda \int_L k_N(s, x) \phi_N(x)\, dx, \quad (5.3)$$

[9] This result follows in a simple way from Eqs. (2.20), (2.21), and (2.37) of Ref. 6. The discussion of the properties of Padé approximants for this part of the w-plane is anticipated in a following paper.

[10] See Ref. 2, pp. 142–145 for a presentation of these methods.

where $\lambda = 1.0$. Try to find a solution of (5.3) of the form

$$\phi_N(s) = g(s) + \sum_{r=1}^{N} \alpha_r(s) A_r, \tag{5.4}$$

where the A_r are constants. Then this is a solution if and only if

$$\sum_{r=1}^{N} \alpha_r(s)\left[A_r - \lambda g_r - \lambda \sum_{i=1}^{N} A_i B_{r,i}\right] = 0, \tag{5.5}$$

where

$$g_r = \int_L g(x)\beta_r(x)\,dx, \quad B_{r,i} = \int_L \beta_r(x)\alpha_i(x)\,dx. \tag{5.6}$$

Now (5.5) is true for all s on the left-hand cut if and only if

$$A_r - \lambda g_r - \lambda \sum_{r=1}^{N} A_i B_{r,i} = 0 \quad (r = 1, 2, \cdots, N). \tag{5.7}$$

Let $\Delta_N(\lambda)$ be the determinant of the N linear equations (5.7) in the unknowns A_i and let $\Delta_{i,r}$ be the complement of the element of the rth row and ith column. Then it follows that, if $\Delta_N(\lambda) \neq 0$, the solution of (5.7) is unique:

$$A_i = \lambda \sum_{r=1}^{N} \frac{\Delta_{i,r} g_r}{\Delta_N(\lambda)}, \tag{5.8}$$

and the corresponding form for the solution of the integral equation (5.3) is

$$\phi_N(s) = g(s) + \lambda \sum_{i=1}^{N} \frac{1}{\Delta_N(\lambda)} \left[\sum_{r=1}^{N} \Delta_{i,r} g_r\right] \alpha_i(s)$$

$$= g(s) + \lambda \int_L \frac{\Delta_N(s, x; \lambda)}{\Delta_N(\lambda)} g(x)\,dx, \tag{5.9}$$

where

$$\Delta_N(s, x; \lambda) = \sum_{i=1}^{N} \sum_{r=1}^{N} \Delta_{i,r} \alpha_i(s) \beta_r(x). \tag{5.10}$$

It is an easy matter to prove that the above solution for $\phi_N(s) = d(s)/s$ is exactly the same as that given by Eq. (3.3). The integrals occurring in Eqs. (5.3)–(5.10) exist when $\operatorname{Im} f(x) = O(x^\delta)$ as $x \to -\infty$ with $\delta < 0$ as was assumed earlier.

The resolvent of Eq. (5.3) may then be written as the ratio of two polynomials in λ as follows:

$$\Gamma_N(s, x; \lambda) = \Delta_N(s, x; \lambda)/\Delta_N(\lambda) \tag{5.11}$$

with

$$\Delta_N(\lambda) = 1 - \frac{\lambda}{1!}\Delta_1 + \frac{\lambda^2}{2!}\Delta_2 - \cdots + \frac{(-1)^N \lambda^N \Delta_N}{N!} \tag{5.12}$$

and

$$\Delta_N(s, x; \lambda) = \Delta_0(s, x) - \frac{\lambda}{1!}\Delta_1(s, x) + \cdots + \frac{(-1)^{(N-1)}\lambda^{(N-1)}}{(N-1)!}\Delta_{(N-1)}(s, x), \tag{5.13}$$

where

$$\Delta_h = \int_L \int_L \cdots \int_L \begin{vmatrix} k_N(x_1, x_1) & k_N(x_1, x_2) & \cdots & k_N(x_1, x_h) \\ k_N(x_2, x_1) & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ k_N(x_h, x_1) & & \cdots & k_N(x_h, x_h) \end{vmatrix} dx_1\,dx_2 \cdots dx_h, \tag{5.14}$$

$$\Delta_h(s, x) = \int_L \int_L \cdots \int_L \begin{vmatrix} k_N(s, x) & k_N(s, x_1) & \cdots & k_N(s, x_h) \\ k_N(x_1, x) & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ k_N(x_h, x) & & \cdots & k_N(x_h, x_h) \end{vmatrix} dx_1\,dx_2 \cdots dx_h. \tag{5.15}$$

These relations can be proved using Fredholm's methods for the case when the range of integrations is finite.[11] It has only to be checked that the integrals in (5.14) and (5.15) exist.

Now it may be shown that, when $-1 < w \leq 0$, the $[N, N-1]$ Padé approximant to $G(w)$ is of smaller magnitude than $G(w)$ itself.[9] Since from (4.8) $c_r < 0$, $a_r > 0$,

$$|k_N(s, x)| = \left| x \operatorname{Im} f(x) \sum_{r=1}^{N} \frac{c_r}{(x - a_r)(s - a_r)} \right| < \left| \frac{x}{s} \operatorname{Im} f(x) \sum_{r=1}^{N} \frac{c_r}{x - a_r} \right| < \left| \frac{x}{s} \operatorname{Im} f(x) F(x) \right| \tag{5.16}$$

[11] See, for example, W. Lovitt, *Linear Integral Equations* (McGraw-Hill Book Company, Inc., New York, 1924), pp. 34–42.

for x and s on the left-hand cut. Therefore, for all x and s on the left-hand cut, it follows from the asymptotic behavior of $\mathrm{Im} f(x)$ and $F(x)$ that

$$|k_N(s, x)| < M/|x^\epsilon s|, \tag{5.17}$$

where M and ϵ are constants independent of N and $-\delta > \epsilon > 0$.

Using (5.17) and Hadamard's inequality for determinants,[12] it is a simple matter to prove that all the integrals in (5.14) and (5.15) exist and that the series

$$\frac{1}{|sx^\epsilon|}\sum_{h=0}^{\infty}|(h+1)^{(h+1)}|^{\frac{1}{2}}M^{(h+1)}(s_1^{-\epsilon}\epsilon)^h\lambda^h/h!$$

and

$$1+\sum_{h=0}^{\infty}[h^h]^{\frac{1}{2}}(s_1^{-\epsilon}\epsilon)^h\lambda^h/h!,$$

which are convergent for all λ, are majorants of the series (5.12) and (5.13), respectively. The latter series therefore also converge for all values of λ and the convergence is uniform for all values of N. Now as N tends to infinity each term of these two series tends respectively to the corresponding terms of the Fredholm series,

$$\Delta(\lambda) = 1+\sum_{h=1}^{\infty}(-1)^h\Delta_h\lambda^h/h!, \tag{5.18}$$

$$\Delta(s, x; \lambda) = \sum_{h=0}^{\infty}(-1)^h\Delta_h(s, x)\lambda^h/h!, \tag{5.19}$$

where Δ_h and $\Delta_h(s, x)$ are defined by (5.14) and (5.15) but with k_N (s, x) replaced by $K(s, x)$. Since (5.17) is also satisfied when $k_N(s, x)$ is replaced by $K(s, x)$, it follows as above that (5.18) and (5.19) are both convergent series. It follows immediately that

$$\operatorname*{Lt}_{N\to\infty}\Delta_N(\lambda) = \Delta(\lambda) \quad \text{and} \quad \operatorname*{Lt}_{N\to\infty}\Delta_N(s, x; \lambda) = \Delta(s, x; \lambda), \tag{5.20}$$

where the limit is approached uniformly for all s and x on the left-hand cut in the latter case.

Therefore, if $\Delta(\lambda) \neq 0$,

$$\operatorname*{Lt}_{N\to\infty}\Gamma_N(s, x; \lambda) = \frac{\Delta(s, x; \lambda)}{\Delta(\lambda)}, \tag{5.21}$$

where $\Delta(s, x; \lambda)$ and $\Delta(\lambda)$ are given by (5.18) and (5.19), respectively, and the convergence is uniform for all s, x in any finite interval on the left-hand cut.

Therefore, from (5.9) and (5.11),

$$\operatorname*{Lt}_{N\to\infty}\phi_N(s) = g(s)+\lambda\int_L\left[\frac{\Delta(s, x; \lambda)}{\Delta(x)}\right]g(x)\,dx, \tag{5.22}$$

and the limit is approached uniformly for all values of s in a finite range on the left-hand cut. Now the right-hand side of (5.22) is just the form of Fredholm's solution to the exact integral equation which it satisfies if $\Delta(\lambda) \neq 0$. This is true even when the integral equation is singular, so long as the terms Δ_h and $\Delta_h(s, x)$ in the definitions (5.18) and (5.19) of $\Delta(\lambda)$ and $\Delta(s, x; \lambda)$ are finite, which is the case here.

Therefore, if $\Delta(1) \neq 0$, $\phi_N(s)$, the solution of the approximate form of (2.6) tends uniformly to a limiting function $\phi(s)$ as $N \to \infty$ for all values of s in a finite range on the left-hand cut, and this function $\phi(s)$ is a solution of the exact integral equation, i.e., $d(s)$, as given by (3.3), converges uniformly to $D(s)$ the solution of (2.6) for all values of s in a finite range on the left-hand cut. From (2.5) it follows that $n(s)$ given by (3.4) converges uniformly to $N(s)$ for s in any finite region of the s plane. Finally, for s in any finite closed region not containing the left-hand cut, $d(s)$ as given by (3.5) converges uniformly to $D(s)$.

Exactly the same procedure may be followed to prove that the approximate solutions of (2.7) converge to the exact solution. In this case

$$\begin{aligned}K(s, x) &= \frac{\rho(x)}{\pi x}\left[\frac{xB(x)-sB(s)}{x-s}\right]\\ &= \frac{\rho(x)}{\pi x}\left[\frac{xB_1(x)-sB_1(s)}{x-s}\right]\\ &\quad -\frac{\rho(x)}{\pi x}\left[\frac{xB_2(x)-sB_2(s)}{x-s}\right],\end{aligned} \tag{5.23}$$

where $B_1(x)$ and $B_2(x)$ are the two functions on the right-hand side of (4.13), and so

$$B(x) = B_1(x) - B_2(x).$$

The kernel (5.23) is now approximated by $k_N(s, x)$, which is obtained by taking the $[N, N-1]$ Padé approximants to the integrals occurring in $B_1(x)$ and $B_2(x)$. Therefore it is of the form

$$k_N(s, x) = \frac{\rho(x)}{\pi x}\left\{\sum_{r=1}^{N}\frac{c_r a_r}{(s-a_r)(x-a_r)} - \sum_{r=1}^{N}\frac{c_r' a_r'}{(s-a_r')(x-a_r')}\right\}, \tag{5.24}$$

where a_r, a_r', c_r, c_r' are negative constants. The approximate integral equation (2.7) may then be written in the standard form (5.3), where $k_N(s, x)$ is given above and

$$\phi_N(s) = n(s), \quad g(s) = B(s), \quad \text{and} \quad \lambda = 1.$$

The whole procedure from (5.3) to (5.22) may be carried through exactly as previously. There is one small difference since in this case[13]

$$|k_N(s, x)| < M'/|s^{\epsilon'}x| \tag{5.25}$$

[12] See Ref. 2, p. 36.

[13] The relation between the asymptotic behavior of $B(s)$ and $\mathrm{Im} f(s)$ has been discussed by L. Lanz and G. M. Prosperi, Nuovo Cimento 33, 201 (1964).

instead of (5.17), where $-\delta > \epsilon' > 0$ and $\epsilon' = 1$ if $-1 > \delta$ and where M' is a constant independent of N. But this does not alter the final results, which may be stated as follows.

If $\Delta(1) \neq 0$, then $n(s)$, as defined by (3.7), tends uniformly to $N(s)$, the solution of (2.7), for all s in any finite region on the right-hand cut. Then $d(s)$, as defined by (3.8), tends uniformly to $D(s)$, the solution of (2.4), for all s in a finite region of the s plane.

Finally, $n(s)$, as defined by (3.9), tends uniformly to $N(s)$ for all s in any finite closed region not containing the right-hand cut.

6. CONCLUSIONS

Approximate methods for solving the N/D equations have been discussed. These methods have to be used since in general the equivalent integral equations for N or D cannot be solved exactly. Although the integral equation for N is nonsingular[7] for the asymptotic behavior of the scattering amplitude considered here, it is often solved by taking an approximate degenerate[4] kernel instead of using the standard Fredholm's method for solution. The corresponding solutions were given in Sec. 3. In many physically interesting cases the integral equation for D is singular. The full Fredholm theory then will not apply to this equation although it may be possible to use Fredholm's methods to obtain a solution. However, approximate solutions may again be obtained by choosing an approximate degenerate kernel, and they have also been described in Sec. 3.

It has been the purpose of this paper to show that one can define sequences of the above approximate solutions which converge uniformly to the exact solution of the N/D equations. The first step was to obtain a sequence of degenerate kernels which converges to the exact kernel of the relevant integral equation. This was done in Sec. 4 using the method of Padé approximants.[14] Now for the partial wave dispersion relations to exist, it is sufficient that $\mathrm{Im}f(s) \to 0$ as $s \to -\infty$. The simple asymptotic behavior $\mathrm{Im}f(s) = O(x^{\delta})$ with $\delta < 0$ has been considered here. In Sec. 5, it was proved that, with this asymptotic behavior, the approximant solutions of the N/D equations converge uniformly to the exact solution when the above convergent sequences of kernels is used, so long as $\Delta(1) \neq 0$. [$\Delta(\lambda)$ is the denominator of the Fredholm solution of the relevant integral equation and is defined in (5.18).]

Now the full Fredholm theory may be applied to the integral equations (2.6) and (2.7) for D and N, respectively, so long as they are nonsingular. As stated previously, the integral equation for N is nonsingular. The integral equation for D is nonsingular if and only if $\delta < -\frac{1}{2}$. However, for $-\frac{1}{2} \leq \delta < 0$, the Fredholm series still converges, as was seen in Sec. 5. In both cases the Fredholm method gives a solution of the relevant integral equation and corresponding solutions of the N/D equations, as long as $\Delta(1) \neq 0$. Comparing with the previous paragraph, one sees that the approximate solutions of the N/D equations considered here converge to the exact solution when the Fredholm method also gives a solution.

When the integral equations are nonsingular, the zeros of $\Delta(\lambda)$ are denumerably infinite (or finite if the kernel is degenerate) and in any finite region of the λ plane there are only a finite number of them. One would then expect $\Delta(1) = 0$ for only certain special cases of the input function. However, the integral equation for D is singular for $\frac{1}{2} \leq \delta < 0$, and these properties of $\Delta(\lambda)$ are in general no longer true. In fact, singular integral equations can be found in which the zeros of $\Delta(\lambda)$ fill the real axis completely.[15] In the case considered here, it can be seen from (5.18) that $\Delta(\lambda) \neq 0$ for at least some small region about $\lambda = 0$. Whether this region includes the point $\lambda = 1$ or not depends on the value of M in (5.17), which in turn will depend on the form for $\mathrm{Im}f(s)$ on the left-hand cut.

In writing down Eqs. (2.4) and (2.5) for N and D it was assumed that no CDD poles were present. However, it is a simple matter to show that the introduction of such poles does not affect the convergence of the corresponding approximate solutions of the N/D equations. Finally, as mentioned in Sec. 4, the poles and residues of the approximate kernel $k_{(N+1)}(s, x)$ are all different from those of $k_N(s, x)$. Therefore an iteration procedure such as that derived by Bander[16] cannot be used here to obtain the approximate solution corresponding to $k_{(N+1)}(s, x)$ in terms of the approximate solution corresponding to $k_N(s, x)$.

ACKNOWLEDGMENTS

The author would like to thank Professor J. S. R. Chisholm for introducing him to the idea and properties of Padé approximants; also he thanks Professor D. Spearman for his comments.

[14] Padé approximants have also been applied to the theory of the solution of the N/D equations by A. P. Balachandran, Ann. Phys. (N.Y.) **30**, 476 (1964) and by D. Masson, J. Math, Phys. **8**, 512 (1967).

[15] F. G. Tricomi, *Integral Equations* (Interscience Publishers, Inc., New York, 1953).

[16] M. Bander, J. Math. Phys. **5**, 1427 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Near-Steady Transition Waves of a Collisionless Plasma*

RICHARD E. MEYER
The University of Wisconsin, Madison

(Received 29 September 1966)

Plane waves of small amplitude ϵ in a cold plasma which propagate into an equilibrium state and the head of which approaches steadiness are studied on the basis of a one-fluid model with transverse magnetic field. Their asymptotic behavior is shown to depend on whether the limits are taken in the order $\epsilon \to 0$, $t \to \infty$ or $t \to \infty$, $\epsilon \to 0$. The latter, nonclassical limit is the physically relevant one, and an approach is developed which yields uniqueness results for it. It is shown that the wave must ultimately become nonlinear, and a steady solution can be approached only in a conditional sense. If the wave lowers the magnetic pressure, it must do so monotonely. If not, then it must begin with a near-periodic wave train approaching steady solutions locally, but different ones in different places.

I. INTRODUCTION

"SHOCK" waves in a collisionless plasma which lead from one equilibrium state to another and settle down to a permanent form have attracted considerable interest over the last decade[1–11] in connection with plasma heating and the solar wind, and on account of their intrinsic physical interest. The model envisaged in most cases is that of a charge-neutral, zero-temperature plasma of singly-charged particles describable by a one-fluid theory.[1,4–6] It may not be realistic enough,[10] but a number of investigations[1,5,8,11] have shown that more elaborate models including temperature effects lead to quite similar results. This is due largely to the unimportance of collisions at sufficiently low density, which gives to any model the character of a process governed by conservation equations without dissipation. It is of interest, then, to understand the properties of the simple model clearly, especially since it has some subtle features with strong effect on what is observable. To reduce the complexity further, only a transverse magnetic field is considered.

The first question is whether there can be waves which are steady,[4,5,8,9] and thus readily observable, in the frame of some observer traveling with fixed velocity. The second question is whether these are actually approached asymptotically in time, and Morton's computations[8,9] confirm this, on the assumption that the plasma is governed by a Korteweg–de Vries equation—but only provided the amplitude is not too small. There is, in fact, a curious paradox: steady solutions leading from one equilibrium to a different one exist *only* for amplitudes above a critical value; for smaller amplitudes, the only steady solutions are either strictly periodic waves or solitary waves leading back to the initial state.[8] The following therefore concentrates on the case of arbitrarily small amplitude ϵ.

This introduces the double limit $t \to \infty$ and $\epsilon \to 0$. The classical procedure of mathematical physics is to let $\epsilon \to 0$, thereby linearizing the problem, and then to evaluate the solution asymptotically in t.[6,9] But that is nonphysical since observation always concerns a wave of given amplitude, studied over a sufficiently long time. Classically, one hopes that the limits commute, but this is disproved for the problem at hand. A direct attack on the correct double limit appears difficult, and the following indirect approach is used. The two double limits stand at the ends of a spectrum of single Kaplun limits[12] such that the real time is $t^* = t/\tau(\epsilon)$ and $\epsilon \to 0$ with t fixed and with $\tau(\epsilon) \to 0$ with ϵ. More precisely, τ is of class

$$\Omega = \{\sigma(\epsilon) \in C(0, 1) \mid \sigma > 0 \text{ and } \sigma \to 0 \text{ as } \epsilon \to 0\}.$$

Now for any $\mu \in \Omega$,

$$L_\mu = \{\sigma \in \Omega \mid (\sigma/\mu) \to 0 \text{ as } \epsilon \to 0\}$$

may be called the "left set of μ," and it has been

* This work was supported by the National Science Foundation under Grant GP-6143.

[1] C. S. Gardner, H. Goertzel, H. Grad, C. S. Morawetz, M. H. Rose, and H. Rubin, in *Proceedings of the Second United Nations International Conference on the Peaceful Uses of Atomic Energy* (United Nations, Geneva, 1958), Vol. 31, p. 230.

[2] J. H. Adlam and J. E. Allen, in Ref. 1, p. 221.

[3] L. Davis, R. Lust, and A. Schluter, Z. Naturforsch. **13a**, 916 (1958).

[4] J. H. Adlam and J. E. Allen, Phil. Mag. **3**, 448 (1958).

[5] A. Banos and A. R. Vernon, Nuovo Cimento **15**, 269 (1960).

[6] C. S. Gardner and G. K. Morikawa, New York University, Courant Inst. Math. Sciences Report TiD-6184 (1960).

[7] P. G. Saffman, J. Fluid Mech. **11**, 552 (1961).

[8] K. W. Morton, J. Fluid Mech. **14**, 369 (1962).

[9] K. W. Morton, Phys. Fluids **7**, 1800 (1964).

[10] P. J. Kellogg, Phys. Fluids **7**, 1555 (1964).

[11] C. S. Gardner and G. K. Morikawa, Commun. Pure Appl. Math. **18**, 35 (1965).

[12] S. Kaplun and P. A. Lagerstrom, J. Math. Mech. **6**, 585 (1957).

shown[13] that asymptotic approximation under

$$\lim_{\epsilon \to 0} \left(\lim_{t^* \to \infty} \right)$$

is equivalent to asymptotic approximation under a Kaplun limit, if the latter approximation is uniform[13] in a left set. The present investigation therefore starts by considering *all* $\tau \in \Omega$ for the plasma wave and proceeds to show that only a few approximations result and, finally, that only two among these can be uniform in any left set. The asymptotic solutions discovered before are not among these two and are therefore only "asymptotic transients" (Sec. V). The steady small-amplitude wave[4,5,8] is approached only in a conditional way, indicating a possible resolution of the paradox just mentioned.

Since collisionless plasma is dispersive, with highest phase velocity for the longest waves, the head of a wave propagating into an equilibrium state must be anticipated to have a large length scale L. But the particular relation $L = O(\epsilon^{-\frac{1}{2}})$ postulated by the earlier authors[3-6,9,11] turns out not to be the only relevant one, and it is desirable to assume only $L^{-1} = \delta(\epsilon) \in \Omega$ and to let the governing equations determine $\delta(\epsilon)$. Dispersion, moreover, tends to sort out the different wavelengths as time increases, and the time asymptotics of a dispersive process must therefore be anticipated to be thoroughly nonuniform in space. The present study is concerned only with the "head" of the wave, that is, with the wave front and as much of the wave behind the front as may be describable asymptotically in terms of a single length scale δ^{-1}. The analysis soon shows (Sec. III) that larger length scales may also be involved, but they are not considered here. It follows that the analysis cannot generally impose any "tail" boundary condition specifying the new equilibrium state to which the wave leads! Fortunately, the resulting indeterminacy turns out not to preclude uniqueness results.

It may be useful to look at the same difficulty also from another angle indicating that it need not depend on dispersion. The wave is governed by conservation equations formally expressible as

$$\partial N/\partial t + \partial M/\partial x = 0.$$

Steady solutions are obtained from $M = \text{const}$. More generally, M is formally given by

$$M = \int_x^\infty \left(\frac{\partial N}{\partial t} \right) dx,$$

if $M \to 0$ as $x \to +\infty$, because the wave travels in the direction of x increasing into an equilibrium state. Approach to steadiness means $|\partial N/\partial t| \to 0$, which is seen to imply, not directly steadiness, but rather that any appreciable effects of unsteadiness are postponed to large distances from the wave front. There is then no *a priori* assurance that the new equilibrium state will be approached by the wave's near-steady part for which $|\partial N/\partial t| \to 0$ implies $M \to 0$. For a sufficiently long transition, moreover, collisions, even if rare, might have a decisive influence on the tail of the wave. It appears all the more desirable to concentrate first on a clarification of the possible asymptotic forms of the head. In some instances (Secs. IV, V), specification of a steady tail boundary condition will turn out to be consistent with the asymptotic equations governing the head. But, generally, the only assumption is the implicit one that the tail boundary condition depends on time to a sufficiently small degree to permit an approach to steadiness at the head of the wave.

A further remark is necessary with regard to the meaning of time asymptotics. If a steady phenomenon is readily observed, it is implausible that the steady state could depend sensitively on the initial conditions from which it developed. This suggests either the use of a canonical (rather than strictly experimental) initial condition,[6,8,9] or, still more simply, a direct investigation of solutions that could emerge after a long time, without consideration of initial conditions. The latter approach is chosen here, and since time then enters into the problem only through the differential operator of the conservation equations, the only meaning assignable to the large time scale τ^{-1} of the Kaplun limit will be that $\tau(\epsilon)$ characterizes the smallness of the time derivatives at the head of the wave. Accordingly, the analysis is concerned with time asymptotics in the sense of asymptotics for a close approach to steadiness. The earlier remarks on the double limit $t^* \to \infty$, $\epsilon \to 0$ are equally relevant to the limit $\partial/\partial t^* \to 0$, $\epsilon \to 0$. The results obtained, moreover, encourage us to think of a progressively closer approach to steadiness as corresponding to progressively later stages in the development of the wave. But this interpretation is not implicit in the assumptions of the analysis.

A final difficulty arises from the fact (proved in Sec. VI) that a certain class of asymptotic solutions must indeed be governed by the Korteweg–de Vries equation.[14] It has the property that some of the steady solutions form a singular subset of the solution set in the sense that "neighboring" solutions differ radically from those steady ones, at least qualitatively. To

[13] R. E. Meyer, J. Inst. Math. Appl. (to be published).

[14] D. J. Korteweg and G. de Vries, Phil. Mag. (5) 39, 422 (1895).

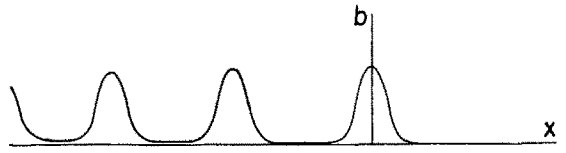

FIG. 1. Sketch of the head of the near-periodic wave train; the magnetic field perturbation is shown vs. distance at fixed time.

clarify the way in which steadiness can be approached therefore requires an unconventionally cautious analysis.

Conclusions. The main results may be summarized as follows. The asymptotic solutions discovered earlier[6,9] are only asymptotic transients (Sec. V) which can neither approach steadiness arbitrarily closely, nor continue indefinitely to approximate a solution. True asymptotics in either sense imply necessarily that the small nonlinear effects accumulate, and this leads to a sharp distinction between waves which lower the magnetic pressure and waves which raise it.

If the magnetic pressure ever falls below the initial equilibrium value within the head of a truly asymptotic wave form, then it must decrease *monotonely* at any fixed position as the head passes, and the head must spread linearly with time (Sec. IV). The only other possible wave approaching permanent shape (Sec. VI) must begin with a near-periodic train of almost cnoidal waves (Eq. 48), in which the magnetic pressure fluctuates, but always exceeds its initial value (Fig. 1). The details of the wave train depend on a parameter varying slowly in time and space, and the first crest of the train approaches the steady solitary-wave solution.[3-6]

II. FORMULATION

Consider plane waves of a charge-neutral plasma of singly ionized particles at zero temperature, without collisions,[1-3] and let n^* and u^* denote respectively the number density and velocity component of the ions (or electrons) in the direction of propagation (x^* increasing). Assume the magnetic field B^* to be purely transverse, take its initial direction to define that of z^* increasing, and let E^* and v^* denote, respectively, the y components of the electric field and of the difference between the ion and electron velocities. The only other nonzero field component is the charge-separation field $(m_+ - m_-)v^*B^*/(m_+ + m_-)$, where m_+ and m_- are the ion and electron masses, respectively. If the ion charge is e, the conservation of numbers and momentum is expressed by [1-3]

$$\partial n^*/\partial t^* + \partial(n^*u^*)/\partial x^* = 0, \tag{1}$$

$$\partial u^*/\partial t^* + u^*\partial u^*/\partial x^* = ev^*B^*/(m_+ + m_-), \tag{2}$$

$$\partial v^*/\partial t^* + u^*\partial v^*/\partial x^* = e(m_+ + m_-)(E^* - u^*B^*)/(m_+m_-), \tag{3}$$

and Maxwell's equations give

$$\partial B^*/\partial t^* + \partial E^*/\partial x^* = 0, \tag{4}$$

$$\partial B^*/\partial x^* + \mu e n^* v^* = 0, \tag{5}$$

where μ denotes the magnetic permeability.

Assume v^* to be twice differentiable; then by (3) and (4), $\partial v^*/\partial x^* + e(m_+ + m_-)B^*/(m_+m_-)$ satisfies the same conservation law as n^*. Since the wave is assumed to travel into plasma at rest and uniform number density n_0 and magnetic field B_0, it follows[6] that

$$\partial v^*/\partial x^* = e(m_+ + m_-)(B_0n^*/n_0 - B^*)/(m_+m_-), \tag{6}$$

and (1), (2), (5), (6) are the governing equations.

They are Galilean invariant, and we choose the frame of reference of an observer traveling in the direction of propagation with constant velocity

$$U = B_0[\mu n_0(m_+ + m_-)(1 - \lambda)]^{-\frac{1}{2}}$$

with respect to the plasma at rest. The boundary conditions ahead of the wave are then

$$u^* \to -U,\ v^* \to 0,\ n^* \to n_0,\ B^* \to B_0 \quad \text{as } x^* \to +\infty, \text{ all } t^*. \tag{7}$$

The constant λ determining U is to be chosen so that the wave can approach steadiness for our observer, and the physically natural interpretation of this notion is that the time scale is large compared with L/U, where L is the length scale. The most general stretching transformation reflecting this and the notion of small amplitude is

$$\begin{aligned} x &= \delta x^* e[\mu n_0(m_+ + m_-)]^{\frac{1}{2}}(m_+m_-)^{-\frac{1}{2}}, \\ t &= \delta\gamma U t^* e[\mu n_0(m_+ + m_-)]^{\frac{1}{2}}(m_+m_-)^{-\frac{1}{2}}, \\ u^*(x^*, t^*) &= -U + \epsilon U u(x, t; \epsilon), \\ v^*(x^*, t^*) &= \theta U v(x, t; \epsilon), \\ n^*(x^*, t^*) &= n_0(1 + \nu n(x, t; \epsilon)), \\ B^*(x^*, t^*) &= B_0(1 + \beta b(x, t; \epsilon)), \end{aligned} \tag{8}$$

where the Greek letters denote small parameters and, since Kaplun limits are to be studied, β, γ, δ, θ, ν all $\in \Omega$. The propagation constant λ may also depend on ϵ, and it is assumed that $\lambda(\epsilon) \in C(0, 1)$.

To find necessary conditions for solutions approaching steadiness at the head of the wave, it is now assumed that (1) to (7) possess solutions for which u, v, n, and b are $\in C^2$ and tend, as $\epsilon \to 0$, to functions of class C^2. More precisely it is assumed that $\{u, v, n, b\} \in \Gamma^2$ defined by

$$\Gamma^n = \left\{ f(x, t; \epsilon) \mid \lim_{\epsilon \to 0} f \in C^n(E) \right\},$$

where E is the set of all (x, t) and the limit is understood in the pointwise norm including the derivatives of order $\leq n$. Moreover, the solutions are assumed nontrivial in the sense that u, v, n, b, and

$$(|\partial u/\partial t| + |\partial v/\partial t| + |\partial n/\partial t| + |\partial b/\partial t|)$$

all belong to the class N of functions not possessing upper and lower bounds tending to zero with ϵ on every open, bounded x, t set.

All this expresses the physical assumption that ϵ, θ, ν, and β do represent the proper amplitude scales and $\tau^{-1} = (\gamma\delta)^{-1}$ and δ^{-1}, the smallest relevant time and length scales, respectively, at the head of the wave. The need to distinguish bounded sets from others in which $|x|$ or $|t|$ do not remain bounded as $\epsilon \to 0$ arises from the suspicion that scales larger than δ^{-1} and τ^{-1} may also be relevant to the wave. If the tail boundary condition can be consistently applied on the scale δ^{-1}, it is taken to be

$$b(x, t; 0) \to b_1 = \text{const} \quad \text{as} \quad x \to -\infty, \quad \text{all} \quad t. \tag{9}$$

If not, there remains a possibility of mistaking the outer skirt of a wave, near its very front, for a substantial part of the wave. The definition of amplitude and length scales is then completed by the assumption that $|b|$ remains bounded as $x \to -\infty$. (It emerges that choosing $|u|$ or $|n|$ here, in the place of $|b|$, would not change the results.) To simplify the language by giving adequate meaning to "boundedness" of $|t|$ in the absence of initial conditions, we add the convention that $t = 0$ represents a time at which the degree of unsteadiness at bounded $|x|$ is already so small that $(\gamma\delta)^{-1}$ is the locally relevant time scale.

III. ANALYSIS

The transformation (8) brings (1) into the form

$$\frac{\gamma\nu}{\epsilon}\frac{\partial n}{\partial t} + \frac{\partial q}{\partial x} = 0, \quad q \equiv u - \frac{\nu}{\epsilon}n + \nu nu. \tag{10}$$

Equations (1) and (2) may be combined into a momentum conservation law, but it is more convenient to work with the energy conservation law obtained directly from (2), by the help of (5) and (6), and which (8) transforms into

$$\gamma\frac{\partial u}{\partial t} + \frac{\partial h}{\partial x} = 0,$$
$$h \equiv -u + \frac{\beta}{\epsilon}(1 - \lambda)b + \frac{\epsilon}{2}(u^2 + \alpha^2 v^2), \tag{11}$$

where $\epsilon\alpha = \theta(m_+ m_-)^{\frac{1}{2}}(m_+ + m_-)^{-1}$. In addition, (8) transforms (5), (6), and (7), respectively, into

$$\delta\beta\partial b/\partial x + \epsilon\alpha v(1 + \nu n)(1 - \lambda)^{-\frac{1}{2}} = 0, \tag{12}$$

$$\epsilon\alpha\delta\partial v/\partial x = (\nu n - \beta b)(1 - \lambda)^{\frac{1}{2}}, \tag{13}$$

$$|u| + |v| + |n| + |b| \to 0 \quad \text{as} \quad x \to +\infty \quad \text{for all} \quad t, \epsilon. \tag{14}$$

For fixed $\epsilon > 0$, Eqs. (10)–(14) are equivalent to the exact equations governing the plasma model.

To begin the study of the limit $\epsilon \to 0$ defined in Sec. II, note that, if $\rho(\epsilon)$ and $\sigma(\epsilon)$ are any two-parameter functions defined by the stretching (8), and if it be shown that ρ/σ tends to a nonzero limit, then no generality is lost in taking $\lim \rho/\sigma = 1$, because (10) to (14) are homogeneous. Moreover, there is then no need to distinguish ρ from σ because information is sought only on the limit of solutions of Eqs. (10)–(14).

Next, note that

$$\{u, v, n, b\} \in \Gamma_0^2 = \{f \in \Gamma^2 \,|\, f(x, t; 0) \to 0 \text{ as } x \to \infty, \text{ all } t\}.$$

For if, e.g., $b(x, t; 0) \not\to 0$ as $x \to \infty$ for fixed t, then, given any $X > 0$, we could find $x_1 > X$ such that $|b(x_1, t; 0)| > 2c > 0$, and since $b \in \Gamma^2$, also $\epsilon_0 > 0$ such that $|b(x_1, t; \epsilon)| > c$ for $0 < \epsilon < 2\epsilon_0$, which which would contradict (14). It follows that

$$\partial\{u, v, n, b\}/\partial x \in N.$$

Indeed, consider, e.g., b on an arbitrary open-bounded set $S \subset E$. For every $\epsilon_1 > 0$, another such set R can be found such that $S \subset R$ and that $(x, t) \in S$ implies $(X, t) \in R$ for X sufficiently large to ensure $|b(X, t; \epsilon)| < \epsilon_1$ for $0 < \epsilon < \epsilon_0$, because $b \in \Gamma_0^0$. If $\partial b/\partial x \notin N$, then ϵ_0 could be chosen small enough to make $\underset{R}{\text{l.u.b.}}\, |b| < 2\epsilon_1$, for $0 < \epsilon < \epsilon_0$, contrary to the hypothesis $b \in N$.

Now consider (10); it shows $\nu/\epsilon \not\to 0$, for otherwise $\partial u/\partial x \notin N$. It may also be written

$$\partial n/\partial x - (\epsilon/\nu)\partial u/\partial x = \gamma\partial n/\partial t + \epsilon\partial(nu)/\partial x$$

for every $\epsilon > 0$, and since n and $u \in \Gamma^1$, the right-hand side $\to 0$ with ϵ on all bounded sets. But $\partial u/\partial x \in N$ implies the existence of a point (x, t) at which $\partial u/\partial x \to a \neq 0$ as $\epsilon \to 0$, and if ϵ/ν did not tend to a limit, then $\partial n/\partial x$ could not do so either at (x, t), contrary to the hypothesis $n \in \Gamma^1$. Moreover, $\epsilon/\nu \not\to 0$, for otherwise $\partial n/\partial x \notin N$. Without loss of generality, therefore, $\nu(\epsilon) = \epsilon$.

From (12), similarly, $\epsilon\alpha = \delta\beta(1 - \lambda)^{\frac{1}{2}}$ and $\partial b/\partial x \in \Gamma_0^2$, and then from (13), $\beta(\epsilon) = \nu(\epsilon)$ and $\delta^2\partial v/\partial x \in \Gamma_0^2$, and thus $\alpha(\epsilon) \in \Omega$. Moreover, (10) and (11) yield

$$\lambda\frac{\partial b}{\partial x} = \gamma\frac{\partial}{\partial t}(n + u) + \frac{\epsilon}{2}\frac{\partial}{\partial x}(u^2 + \alpha^2 v^2 + 2nu) - \delta^2\frac{\partial^2 v}{\partial x^2}, \tag{15}$$

so that

$$\alpha_0(\epsilon) \equiv \max(\epsilon, \delta^2, |\lambda|) \in \Omega.$$

It is useful to establish that $\lim_{\epsilon \to 0} \partial v/\partial x \to 0$ as $x \to \infty$ for any fixed t. Given $\eta > 0$, (14) implies the existence of $X_\eta(\epsilon)$ such that $\epsilon > 0$ and $x > X_\eta(\epsilon)$ together imply $|n - b| < \eta$. Let $\partial v/\partial x = f(x; \epsilon)$; then (13) has been reduced to $n - b = \delta^2 f$. Suppose $f(x; 0) \not\to 0$ as $x \to \infty$, then, given any X, we could find $x_1 > X$ such that $|f(x_1; 0)| = 2c$ for some $c > 0$. Since $f \in \Gamma^1$ by hypothesis, there must then be an $\epsilon_0 > 0$ such that $|f(x_1; \epsilon)| > c$ for $0 < \epsilon < 2\epsilon_0$, and if $\delta(\epsilon_0) = \delta_0$, also an $x_1 > X_\eta(\epsilon_0)$ such that

$$c < |f(x_1; \epsilon_0)| < \eta/\delta_0^2,$$

which is not possible for every $\eta > 0$. Hence

$$\partial v/\partial x \in \Gamma_0^1. \tag{16}$$

Since $q + h = (\epsilon/2)(u^2 + \alpha^2 v^2 + 2nu) - \lambda b - \delta^2 \partial v/\partial x$, it follows that

$$Q + H \equiv \alpha_0^{-1}(q + h) \in \Gamma_0^1. \tag{17}$$

On the other hand,

$$h - q = 2(b - u) - \lambda b + \delta^2 \frac{\partial v}{\partial x} + \frac{\epsilon}{2}(u^2 + \alpha^2 v^2 - 2nu) \in \Gamma_0^2, \tag{18}$$

whence, by (10), (11),

$$\left(\frac{\gamma}{2}\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)(h - q) = \frac{\gamma}{2}\frac{\partial}{\partial t}\left[\frac{\epsilon}{2}(u^2 + \alpha^2 v^2 - 2nu) - \lambda b - \delta^2 \frac{\partial v}{\partial x}\right] = \gamma \alpha_0 g_1, \quad g_1 \in \Gamma_0^0, \tag{19}$$

and

$$\frac{h - q}{\gamma \alpha_0} = \int_x^\infty g_1(\xi, \tau(\xi); \epsilon)\, d\xi, \quad \tau(\xi) = t - \frac{\gamma}{2}(\xi - x), \tag{20}$$

for every $\epsilon > 0$, by (14) and (18). For fixed x, t, then, $\eta > 0$ implies the existence of $X_\eta(\epsilon)$ such that

$$\left| \int_Y^Z g_1(\xi, \tau(\xi); \epsilon)\, d\xi \right| < \eta$$

whenever both $Z > Y > X_\eta(\epsilon)$ and $\epsilon > 0$. An argument similar to that used to prove (16) now shows that the integral in (20) converges also for $\epsilon = 0$, and it then depends continuously on x and t, by (19). Thus

$$\gamma^{-1}(H - Q) \equiv (\gamma \alpha_0)^{-1}(h - q) \in \Gamma_0^0, \tag{21}$$

and, from (16)–(18),

$$Q = \alpha_0^{-1} q \in \Gamma_0^0, \quad H = \alpha_0^{-1} h \in \Gamma_0^0, \quad \alpha_0^{-1}(u - b) \in \Gamma_0^0. \tag{22}$$

Since $\delta^2 \partial v/\partial x \in \Gamma^2$, moreover, (19) may be differentiated for $\epsilon \geq 0$, and the same argument then shows $\gamma^{-1}\partial(h - q)/\partial t$ and $\gamma^{-1}\partial(h - q)/\partial x \in \Gamma_0^0$, whence, by (18),

$$\alpha_1^{-1}\partial(b - u)/\partial t \in \Gamma_0^0, \quad \alpha_1^{-1}\partial(b - u)/\partial x \in \Gamma_0^0, \quad \alpha_1(\epsilon) = \max(\alpha_0, \gamma). \tag{23}$$

To sum up, it has been established that the perturbations of number density, magnetic field, and x velocity all have the same amplitude parameter ϵ. More remarkably, the corresponding mass-flow perturbation

$$q = (n^* u^* + n_0 U)/(\epsilon n_0 U),$$

energy perturbation

$$h = \frac{1}{\epsilon U^2}\left[\frac{u^{*2} - u^2}{2} + \frac{m_+ m_- v^{*2}}{2(m_+ + m_-)^2} + \frac{B_0(B^* - B_0)}{\mu(m_+ + m_-)n_0}\right],$$

and momentum perturbation

$$\frac{1}{\epsilon n_0 U^2}\left[\frac{\lambda - 3}{2} n_0 U^2 + n^* u^{*2} + \frac{B^{*2}}{2\mu(m_+ + m_-)}\right]$$

are all much smaller, when ϵ is sufficiently small, than the perturbations of number density, magnetic field, and x velocity. There is, moreover, a significant difference between $q + h$ and $q - h$. The result $q + h = O(\max(\delta^2, \epsilon, |\lambda|))$ is largely independent of the condition that the wave travels into equilibrium plasma; it reflects mainly the local amplitude, length, and time scales. The smallness of $q - h$, by contrast, is due to its very small rate of growth with distance from the wave front, and this makes $|q - h|$ even smaller than $|q + h|$ at the head of the wave.

The amplitude scale of the transverse velocity difference v^* has been shown to be

$$\theta = \epsilon\delta(m_+ + m_-)(1 - \lambda)^{\frac{1}{2}}(m_+ m_-)^{-\frac{1}{2}},$$

with $\lambda \in \Omega$ also, and the governing equations (10)–(13) have been reduced to

$$\gamma \frac{\partial b}{\partial t} + \frac{\partial q'}{\partial x} = 0, \quad q' = u - b + \epsilon u b + \delta^2 \omega = q + \gamma\delta^2 \partial v/\partial t = \alpha_0 Q', \tag{24}$$

$$\gamma \frac{\partial u}{\partial t} + \frac{\partial h}{\partial x} = 0, \quad h = (1 - \lambda)b - u + \epsilon r = \alpha_0 H, \tag{25}$$

$$v = -(1 + \epsilon n)^{-1} \partial b/\partial x, \tag{26}$$

$$n = b + \delta^2 \partial v/\partial x, \tag{27}$$

where H and $Q' \in \Gamma_0^0$, $\alpha_0(\epsilon) = \max(\epsilon, \delta^2, |\lambda|)$, and abbreviations

$$\omega = \gamma \partial v/\partial t - (1 - \epsilon u)\partial v/\partial x, \quad r = \tfrac{1}{2}(u^2 + (1 - \lambda)\delta^2 v^2) \tag{28}$$

are used for correction terms. To conclude this section, it is now shown that the head of the wave must be governed by limiting equations resulting from (24)–(27) as $\epsilon \to 0$ with either Ursell number[15] $\epsilon/\delta^2 \to 0$ or $\delta^2/\epsilon \to 0$ or $\delta^2 = \epsilon$.

To this end, note first that $\partial u/\partial t$, $\partial b/\partial t$ and $\partial n/\partial t$ are each $\in N$. For if one were not, the other two would also be $\notin N$, by (23) and (27). But by hypothesis (Sec. II), there would then be a bounded open set S with $\operatorname{lub}_S |\partial v/\partial t| \nrightarrow 0$ as $\epsilon \to 0$, and since v and $n \in \Gamma^1$, there would be a point (x, t) with neighborhood on which $glb \lim_{\epsilon \to 0} |\partial^2 b/\partial x \partial t| > 0$, by (26), which contradicts the supposition $\partial b/\partial t \notin N$.

Next,

$$\left(2\frac{\partial}{\partial t} - \frac{\lambda}{\gamma}\frac{\partial}{\partial x}\right) b = \frac{\partial}{\partial t}(b - u) - \frac{1}{\gamma}\frac{\partial}{\partial x}\left(\epsilon(r + ub) + \delta^2\omega\right), \quad (29)$$

obtained from (24) and (25) for $\epsilon > 0$, is used to show that

$$\epsilon/\gamma \to 0 \quad \text{and} \quad \delta^2/\gamma \to 0 \quad \text{as} \quad \epsilon \to 0 \quad (30)$$

is incompatible with our hypotheses. Suppose it was compatible; then, from (15),

$$(\lambda/\gamma)\partial b/\partial x - \partial(n + u)/\partial t \notin N,$$

and since $\partial b/\partial x \in N$ and $\partial(n + u)/\partial t \in \Gamma^1$, λ/γ must tend to a limit. But $\lambda/\gamma \nrightarrow 0$, for otherwise (30), (29), and (23) would imply $\partial b/\partial t \notin N$. Thus (30) would imply a nonzero limit for λ/γ and (29) would be an equation for b analogous to Eq. (19) for $h - q$ and would imply, similarly, that $b \notin N$, contrary to hypothesis.

The possibility that both $\epsilon/\lambda \to 0$ and $\delta^2/\lambda \to 0$ as $\epsilon \to 0$ may now also be ruled out because (15) would then imply $\partial b/\partial x - (\gamma/\lambda)\partial(n + u)/\partial t \notin N$, and since both $\partial b/\partial x$ and $\partial(n + u)/\partial t \in \Gamma^0 \cap N$ [because $\partial(n - u)/\partial t \notin N$ by (27), (23)], γ/λ would have to tend to a nonzero limit and (30) would follow.

For $\epsilon > 0$, (15) may be written

$$\frac{\partial^2 v}{\partial x^2} = \frac{\gamma}{\delta^2}\frac{\partial(n + u)}{\partial t} + \frac{\epsilon}{2\delta^2}\frac{\partial}{\partial x}(u^2 + \alpha^2 v^2 + 2nu) - \frac{\lambda}{\delta^2}\frac{\partial b}{\partial x}, \quad (31)$$

and to complete the argument requires confirmation that the four derivatives in (31) are each $\in N$. This has already been shown for $\partial b/\partial x$ and $\partial(n + u)/\partial t$, and if it were not true of $\partial^2 v/\partial x^2$, then $v(x, t; 0)$ would be linear in x, contrary to the hypothesis $v \in \Gamma_0^2 \cap N$. It has also been shown that $\alpha \in \Omega$ and u and $\partial u/\partial x \in N$, and since $n - b$, $\partial(n - b)\partial x$, $b - u$, and $\partial(b - u)/\partial x \notin N$, by (27), (22), and (23), the confirmation is complete. We recall that the four derivatives in (31) are also $\in \Gamma^0$, by hypothesis (Sec. II).

Now suppose that neither δ^2/ϵ nor ϵ/δ^2 tends to a limit. Then (31) implies either

(i) $\partial(u^2 + \alpha^2 v^2 + 2nu)/\partial x - (2\lambda/\epsilon)\partial b/\partial x \notin N$

or

(ii) $\partial^2 v/\partial x^2 + (\lambda/\delta^2)\partial b/\partial x \notin N.$

In case (i), again since both the derivatives $\in \Gamma^0 \cap N$, it would follow that λ/ϵ tends to a nonzero limit, and then, without loss of generality, $|\lambda/\epsilon| \to 1$, and (27), (22), (23) would imply $\partial(3u^2 \mp 2u)/\partial x \notin N$, contrary to the hypothesis $u \in \Gamma_0^0 \cap N$. In case (ii), similarly, $|\lambda/\delta^2| \to 1$, and by (26), $\partial^2 v/\partial x^2 \mp v \notin N$. Since $v \in \Gamma_0^2 \cap N$, it would follow that $v - ce^{-x} \notin N$ with constant $c \neq 0$, and (26) would imply $b - ce^{-x} \notin N$, contrary to the hypothesis that b remains bounded as $x \to -\infty$ (Sec. II).

IV. VERY LONG WAVES

Waves for which

$$\delta^2/\epsilon \to 0 \quad \text{as} \quad \epsilon \to 0 \quad (32)$$

are now studied, and we begin by showing that then λ/ϵ tends to a limit and $\gamma = \epsilon$. From (24) and (25),

$$\frac{\gamma}{\epsilon}\frac{\partial}{\partial t}(b + u) = \frac{\lambda}{\epsilon}\frac{\partial b}{\partial x} - \frac{\partial}{\partial x}\left(r + ub + \frac{\delta^2}{\epsilon}\omega\right)$$

for $\epsilon > 0$. Suppose now that λ/ϵ does not tend to a limit; then, since $r + ub \in \Gamma^1$ and $\partial b/\partial x \in \Gamma^0 \cap N$, it would follow that $\partial b/\partial x - (\gamma/\lambda)\partial(b + u)/\partial t \notin N$ and, since $b + u \in \Gamma^1$, λ/γ must tend to a limit, so that $\partial(b + u)/\partial t - (\lambda/\gamma)\partial b/\partial x \in \Gamma^1$. But $\epsilon/\gamma \to 0$ would imply (30), by (32); and since u, $\partial u/\partial x \in N$, we have $\partial(ru + b)/\partial x \in N$, by (28), (22), and (23), and it would follow that γ/ϵ must also tend to a limit. Thus $\lim(\lambda/\epsilon)$ does exist, and no generality is lost in taking $\alpha_0 = \epsilon$, and from (24), (25), and (28),

$$H + Q' = \tfrac{3}{2}u^2 - \frac{\lambda}{\epsilon}u + \frac{\delta^2}{\epsilon}\left(\omega + \frac{\epsilon}{2}(1 - \lambda)v^2\right) + \left(u - \frac{\lambda}{\epsilon}\right)(b - u),$$

$$\partial(H + Q')/\partial x = -(\gamma/\epsilon)\partial(b + u)/\partial t. \quad (33)$$

Since $\partial(b + u)/\partial t \in N$, by (23), it now follows from (17) that $\lim(\gamma/\epsilon)$ exists, and since (32), (22) and

[15] F. Ursell, Proc. Cambridge Phil. Soc. 49, 685 (1953).

$u \in N$ imply $H + Q' \in N$, (17) rules out $(\gamma/\epsilon) \to 0$, and no generality is lost in taking $\gamma = \epsilon$ and $\alpha_1 = \epsilon$.

By (23), (33) therefore implies

$$2\partial u/\partial t + (3u - \lambda/\epsilon)\partial u/\partial x \notin N$$

and, in the limit,

$$du/dt = 0 \quad \text{when} \quad dx/dt = \tfrac{1}{2}(3u - k), \qquad (34)$$

where $k = \lim \lambda/\epsilon$. By (32), moreover, this result is valid for $\tau = \gamma\delta$ in the left set of $\epsilon^{\frac{3}{2}}$ and, hence,[13] for the desired double limit (Sec. I). Now $3u - k$ increases with u and thus, if at some t the longitudinal velocity u (or n or b) fails to be a monotone, nondecreasing function of x, then it will fail to be $\in \Gamma^2$, and indeed even $\in \Gamma^0$, after a *finite* interval of t. A necessary condition for (32) to be asymptotically consistent with the governing equations is, therefore, that u, b and n are monotone, nondecreasing functions of x at any fixed t—for an observer who sees the plasma initially at rest, the wave must lower the longitudinal velocity, the magnetic pressure, and the number density, as it passes.

The conclusions arrived at are sufficiently explicit to permit remarks also on the real time asymptotics of this wave type, provided that either the tail boundary condition (9) applies with $b_1 < 0$ [so that (34) describes the whole wave] or the tail does not overtake the head. In that case, if the necessary monotoneity condition is satisfied at some t, (34) shows it to remain satisfied over any bounded time interval and indeed over any time interval Δt such that $\rho(\epsilon)\Delta t \to 0$ as $\epsilon \to 0$ for every $\rho \in \Omega$. Moreover, (34) shows the wave to spread by and by and thus to develop in a sense strengthening (32), so that real time asymptotics are equivalent to asymptotics for approach to steadiness. As far as the present analysis can carry, it therefore reveals no inconsistency in the existence hypotheses of Sec. II. The reader will have noticed the similarity between these waves and the simple waves of gas dynamics,[16] but there does not appear to be an exact solution of Eqs. (1)–(7) of such type.

V. LINEAR WAVES

Assume next that

$$\epsilon/\delta^2 \to 0 \quad \text{as} \quad \epsilon \to 0; \qquad (35)$$

then an argument analogous to that opening the preceding section shows that λ/δ^2 must tend to a limit,

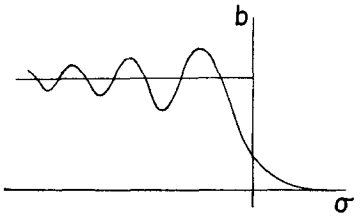

FIG. 2. Graph of (38) for $c > 0$; it represents the (transient) asymptotic form of the magnetic field perturbation vs. distance at fixed time for all transmission waves; scaling according to (35).

and $\alpha_0 = \delta^2$. From Eqs. (24)–(28), therefore,

$$H + Q' = \frac{1 - \epsilon u}{1 + \epsilon n}\frac{\partial^2 b}{\partial x^2} - \frac{\lambda}{\delta^2}b + \gamma\frac{\partial v}{\partial t} + \frac{\epsilon}{\delta^2}(r + ub) - \frac{\epsilon(1 - \epsilon u)}{(1 + \epsilon n)^2}\frac{\partial b}{\partial x}\frac{\partial n}{\partial x},$$

$$\partial(H + Q')/\partial x = -(\gamma/\delta^2)\partial(b + u)/\partial t, \qquad (36)$$

and since (23) implies $\partial(b + u)/\partial t \in N$, (17) shows $\lim(\gamma/\delta^2)$ to exist. If $\gamma/\delta^2 \to 0$, (17) and (35) imply $\partial^2 b/\partial x^2 - (\lambda/\delta^2)b \notin N$, so that b must be either trivial or unbounded, contrary to hypothesis (Sec. II). Thus $\gamma = \delta^2$, and by (23), (36) implies $2\partial b/\partial t + \partial^3 b/\partial x^3 - (\lambda/\delta^2)\partial b/\partial x \notin N$; and the last term in this expression may be absorbed into the time derivative by a transformation to the frame of an observer traveling with speed corresponding to $\lambda/\delta^2 = 0$, to obtain the limiting equation

$$2\partial b/\partial t + \partial^3 b/\partial x^3 = 0. \qquad (37)$$

It is readily verified to have a solution[6,17]

$$b(x, t) = c\int_\sigma^\infty Ai(\mu)\, d\mu, \quad \sigma = \left(\frac{2}{3t}\right)^{\frac{1}{3}}x,$$
$$Ai(\mu) = \frac{1}{\pi}\int_0^\infty \cos\left(\mu w + \frac{w^3}{3}\right) dw \qquad (38)$$

representing a wave of transition (Fig. 2) from the magnetic field B_0 to the magnetic field $B_0(1 + \epsilon c)$. Thus (38) satisfies (9), if $c = b_1$. Conversely, (37), (9), and (14) are invariant under the transformation $x = ax'$, $t = a^3 t'$, $b(x, t) = b'(x', t')$ for arbitrary constant $a \neq 0$, and the solution $b(x, t)$ for the initial condition $b(x, 0) = b_0(x)$ may thus be obtained from the solution $b'(x', t')$ for the initial condition

$$b'(x', 0) = b_0(ax').$$

For any monotone $b_0(x)$ satisfying (9) and (14), $b'(x', 0) \to b_1 H(-x')$ as $a^{-1} \to 0$ (where H denotes the unit step function), and that is the initial condition satisfied by (38) if $c = b_1$. Any transition wave governed by (37) and monotone in x at some t

[16] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience Publishers, Inc., New York, 1948).

[17] H. Jeffreys and B. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, Cambridge, England, 1946), pp. 508–518.

therefore approximates (38) asymptotically as $t \to \infty$. This holds, moreover, also for any such wave differing from a monotone one at some t by an integrable function [use the linearity of (37), a slightly different transformation, and the x derivative of (38)].

Gardner and Morikawa[6] conjectured that (38) is not truly asymptotic for Eqs. (1)–(7) because (37) does not admit the steady solitary-wave solution[2–5] of Eqs. (1)–(7). This is unconvincing, since Sec. IV indicates the existence of solutions approaching steadiness arbitrarily closely, but governed by a limiting equation that also fails to admit steady solutions. Rather than discredit the conjecture, however, the present result supplies proof that (37) can describe only a transient asymptotic stage in the approach to steadiness. This follows from (35), which has been seen to imply $\gamma = \delta^2$, so that $\tau = \delta^3$; if the approximation be uniform for τ in the left set of some $\sigma(\epsilon) \in C(0, 1)$, i.e., valid for every τ such that $\tau/\sigma \to 0$ with ϵ, then it must therefore be valid for $\delta^3 = \tau = \epsilon \min(\epsilon, \sigma)$, which implies (32), not (35). For the head of a wave characterized by (35), moreover, an increase of the time scale $(\gamma\delta)^{-1} = \delta^{-3}$ relative to the amplitude scale ϵ must weaken (35) so that ultimately $\rho\delta^2/\epsilon \to 0$ for every $\rho \in \Omega$—indicating that closer approach to steadiness requires boundedness of δ^2/ϵ.

The conclusions are again sufficiently explicit to permit relating near-steadiness to real time asymptotics and proving that (37) can represent no more than a transient also in the latter sense. Indeed, (38) shows b, u, and n to depend asymptotically (as $t \to \infty$) only on $t^{-\frac{1}{3}}x$, and the development of the wave therefore preserves the relation $\gamma = \delta^2$, while the rates of change of b, u, n with x and t—to which the definitions of δ and $\gamma\delta$ relate directly—decrease in magnitude as t increases.[15] Since the amplitude ϵb_1 is independent of time, (35) must weaken with time and must be anticipated to give way eventually to $\gamma = \delta^2 = \epsilon$. It should be stressed that (37) results[6] from the classical procedure of letting the amplitude tend to zero before letting the time tend to infinity, and our result therefore shows the order of these limit processes to be definitely noncommutative for the physical problem at hand.

VI. KORTEWEG–DE VRIES WAVES

Assume finally that

$$\delta^2 = \epsilon. \tag{39}$$

Then an argument similar to that opening Sec. IV shows that λ/ϵ tends to a limit, and $\alpha_0 = \epsilon$. From Eqs. (24)–(28), therefore,

$$H + Q' = \tfrac{3}{2}b^2 - \frac{\lambda}{\epsilon}b + \frac{1-\epsilon u}{1+\epsilon n}\frac{\partial^2 b}{\partial x^2} + \frac{3b+u}{2}(u-b) + \gamma\frac{\partial v}{\partial t} + \epsilon\left[\frac{1-\lambda}{2}v^2 - \frac{1-\epsilon u}{(1+\epsilon n)^2}\frac{\partial b}{\partial x}\frac{\partial n}{\partial x}\right], \tag{40}$$

$$\partial(H + Q')/\partial x = -(\gamma/\epsilon)\partial(b+u)/\partial t,$$

and by (23) and (17), $\lim(\gamma/\epsilon)$ exists. This leaves two possible cases.

(i) If $\gamma = \epsilon$, (23) implies

$$2\frac{\partial b}{\partial t} + \frac{\partial}{\partial x}\left[\frac{\partial^2 b}{\partial x^2} + \tfrac{3}{2}b^2 - \frac{\lambda}{\epsilon}b\right] \notin N, \tag{41}$$

so that the limiting equation is that of Korteweg and de Vries.[14] However, $\gamma = \delta^2 = \epsilon$ implies $\tau = \epsilon^{\frac{3}{2}}$, which does not represent a left set, and (41) can therefore relate only to a transient asymptotic approximation. Conversely, the numerical work of Morton[9] and Peregrine[18] indicates that (41) can describe the development of a transition wave, both from a shape with rather short length scale and also from one with very large length scale, to one of the last type remaining to be discussed.

(ii) If $\gamma/\epsilon \to 0$ as $\epsilon \to 0$, so that the time scale is large compared even with $\epsilon^{-\frac{3}{2}}$, (40) and (17) imply

$$\partial^2 b/\partial x^2 + \tfrac{3}{2}b^2 - kb \notin N, \quad k = \lim \lambda/\epsilon, \tag{42}$$

and since $b \in \Gamma_0^2$,

$$b \sim k\,\mathrm{sech}^2(\tfrac{1}{2}k^{\frac{1}{2}}x), \tag{43}$$

a solitary wave of amplitude parameter k. Since this result has been deduced for τ in the left set of $\epsilon^{\frac{3}{2}}$, it applies to the desired double limit (Sec. I). This wave type and the one discussed in Sec. IV are therefore the only possible, fully asymptotic waves consistent with Eqs. (1)–(7). Observe that (43) implies $k > 0$, if $b \in N$, so that the wave front must raise the magnetic pressure; conversely, it is the only fully asymptotic type of wave that can start with a rise in magnetic pressure, since the very long waves have been shown (Sec. IV) to require a monotone fall in magnetic pressure.

The result (43), however, has been deduced only for the head of the wave, and it may mislead, because (43) is an exceptional solution of (42). Under such circumstances, an asymptotic statement like (42) may not be precise enough, and we therefore return to (40). Assuming γ/ϵ bounded, it gives, by (22) and (23),

$$\partial^2 b/\partial x^2 + (3/2)b^2 - kb - \varepsilon f = H + Q', \quad f \epsilon \Gamma_0^1, \tag{44}$$

$$\partial(H + Q')/\partial x = -(\gamma/\epsilon)\partial(b+u)/\partial t, \tag{45}$$

[18] D. H. Peregrine, J. Fluid Mech. 25, 321 (1966).

and if $\gamma/\epsilon \to 0$, this pair is a *near-steady* form of an equation close to that of Korteweg and de Vries. Since the solution must approach (43) for bounded x, it follows that $k > 0$, and since $\partial b/\partial x \in \Gamma_0^2$ (Sec. III), integration of (44) gives

$$(\partial b'/\partial x') = b'^2 - b'^3 + 2Fb' - 2G \equiv C(b'; x', t; \epsilon), \tag{46}$$

where $b' = b/k$, $x' = k^{\frac{1}{2}}x$ and

$$k^2F = H + Q' + \epsilon f, \quad G = -\int_{x'}^{\infty} b' \frac{\partial F}{\partial x}\, dx. \tag{47}$$

Observe that $C(b'; x', t; \epsilon)$ is a near-cubic in b' with coefficients $2F$, $2G$ which, by (17) and (45), are small and *slowly varying* with x. The criterion for real solutions, bounded for bounded x, of (46) with constant F and G is[14]

$$\Delta = (1 + 9F - 27G)^2 - (1 + 6F)^3 \leq 0,$$

and the curve $\Delta = 0$ in the F, G plane is thus a critical curve for (46). By (17), (44), and (47), $F \to 0$ and $G \to 0$ (and thus also $\Delta \to 0$) as $x \to +\infty$ for any fixed t; and $F = G = 0$ corresponds to the solitary-wave solution (43) of (46). As x decreases, the representative point (F, G) must be anticipated to shift slightly from the origin of the F, G plane and, *however little* it shifts, the solution of (46) may then assume a quite different character.[19]

[19] T. B. Benjamin and M. J. Lighthill, Proc. Roy. Soc. (London) **A224**, 448 (1954).

If the representative point shifts into the region $\Delta > 0$, the solution of (46) satisfying (14) fails to be bounded even for bounded x, contrary to hypothesis. It cannot shift along the critical curve, as x decreases for fixed t, because $\Delta = 0$ implies $3\, dG/dF = 1 - (1 + 6F)^{\frac{1}{2}} \notin N$, while (47) implies $(\partial G/\partial x)/(\partial F/\partial x) = b' \in N$. For an arbitrarily close approach to steadiness at the head of the wave, with $\delta^2 = \epsilon$, it is therefore necessary that the representative point shift into the region $\Delta < 0$, and (46) is then the equation of a cnoidal wave[14]

$$b' = \rho + (1 + \nu - \rho)cn^2[\tfrac{1}{2}x'(1 + \nu - \sigma)^{\frac{1}{2}}] \tag{48}$$

of modulus $(1 + \nu - \rho)^{\frac{1}{2}}(1 + \nu - \sigma)^{-\frac{1}{2}}$, where $1 + \nu$, ρ and σ are the roots of the cubic $C(b'; x', t; \epsilon)$ in decreasing order. The wavelength is

$$\Lambda = 4(1 + \nu - \sigma)^{-\frac{1}{2}}K([(1 + \nu - \rho)/(1 + \nu - \sigma)]^{\frac{1}{2}}),$$

where K denotes the complete elliptic integral of the first kind and, since $|\nu| + |\rho| + |\sigma| \to 0$ as $\epsilon \to 0$ for bounded x', t, $\Lambda \sim -2 \log (\rho - \sigma)$. The wavelength is therefore logarithmically large compared with the scale that defines the head of the wave. For bounded x, (48) approaches (43) as $\epsilon \to 0$; but as $x \to -\infty$, the real asymptotic approximation is a near-periodic wave train (Fig. 1) dependent on the slowly varying parameter F.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Uniqueness of the Haag–Ruelle Scattering States*

R. F. STREATER
Physics Department, Imperial College, London, England

(Received 26 September 1966)

It is remarked that no complete proof exists in the literature that the asymptotic states of quantum field theory are independent of the spacelike surfaces chosen to define them. In this paper we present a proof which is valid for any plane surface; this follows the Haag–Ruelle method, supplemented by lemmas showing that the usual bounds on truncated Wightman functions and smooth solutions of the Klein–Gordon equation are uniform in certain spacelike regions. The same lemmas immediately show that asymptotic states may be defined using any spacelike surface possessing a normal, which lies inside a closed timelike cone at all points. The proof of the convergence of these states has hitherto been incomplete. It is then important to show that these states are independent of the surface; we sketch a proof of this.

INTRODUCTION

THE fundamental work of Haag[1] and Ruelle[2] on the asymptotic condition in axiomatic quantum field theory turns on two lemmas. The first concerns the asymptotic properties of the truncated Wightman functions in spacelike directions, and the second, the behavior of the smooth solutions of the Klein–Gordon equation, and its space integral, for large times. The first lemma had received some attention from previous authors[3] but the first proof of a theorem adequate for the problem was provided by Ruelle.[2] For the second lemma, Haag gave a rough argument in his original paper, which was made more mathematical by Ruelle. However, neither Ruelle's proof, nor the expositions of it,[4] go into complete details, and this must be considered unsatisfactory in view of the importance of the subject. As remarked by Segal,[5] it is not easy to supply all the details if one follows Ruelle's method.

Following a suggestion of Jost, Brodsky has given a complete proof of the required lemmas on the Klein–Gordon equation using operator methods,[5] and Jost himself has presented a straightforward proof.[6] These results, together with Ruelle's lemma on the truncated Wightman functions, enable one to follow the usual line of argument.[1,2,4] This demonstrates the convergence for large times of a sequence of time-dependent states $\Psi(t)$ formed from products of creation operators $B_\alpha(t)$. These are obtained from the field by integrating over a plane spacelike surface labeled by a parameter t. The limit states $\Psi^\pm$ obtained as $\pm t \to \infty$ are called asymptotic states of the theory. However, to show that the limit states are independent of the spacelike surface chosen requires a little more work. The published discussion on this point[2,4] ignores the problem of showing that the bounds are uniform in the parameters defining the normal to the surface. We fill this small gap in this paper, at the same time generalizing the existence proof to certain curved surfaces. The properties of the truncated Wightman functions beyond those given in Ref. 2 are easily obtained, using the method of Ruelle; the extension of the results of Jost and Ruelle is also quite elementary. This is virtually all that is needed to prove the uniqueness of the states obtained using curved surfaces.

1. PROOF OF CONVERGENCE

Let us recall the lemmas in more detail. Suppose $\phi(x)$ is a Wightman field with representation $U(a, \Lambda)$ of the Poincaré group. Suppose $U(a, 1) = e^{ip_\mu a_\mu}$, and the spectrum of P^μ exhibit the simple features of an ideal model, that is, a nondegenerate point vacuum Ψ_0 and a one-particle state Ψ_p of mass $m > 0$ separated from other states by a mass gap (see Figs. 1–3 in Ref. 7). Since polynomials in the smeared fields form a dense set when applied to Ψ_0, there exist operators (finite sum)

$$B_i = \sum_j \int g_{ij}(x_1, \cdots, x_j)\phi(x_1) \cdots \phi(x_j)\, dx_1 \cdots dx_j, \tag{1.1}$$

such that $B_i\Psi_0$ is not orthogonal to Ψ_p. Here we may

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research OAR through the European Office Aerospace Research, U.S. Air Force.

[1] R. Haag, Phys. Rev. **112**, 669 (1958).

[2] D. Ruelle, Helv. Phys. Acta **35**, 147 (1962).

[3] H. Araki, Ann. Phys. (N.Y.) **11**, 260 (1960); K. Hepp, R. Jost, D. Ruelle, and O. Steinman, Helv. Phys. Acta **34**, 542 (1961).

[4] A. Wightman, in *Theoretical Physics*, A. Salam, Ed. (International Atomic Energy Agency, Vienna, 1963); R. Jost, *General Theory of Quantized Fields* (American Mathematical Society, Providence, Rhode Island, 1965).

[5] I. E. Segal, in *Mathematical Theory of Particles and Fields*, (M.I.T. Press, Cambridge, Mass., 1966) A. R. Brodsky, Ph.D. thesis, Massachusetts Institute of Technology (1964). The details of Ruelle's proof have been given by H. Araki, Lectures at ETH, Zürich (Unpublished) 1962.

[6] R. Jost, Helv. Phys. Acta. **39**, 21 (1966).

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

take $g_{ij} \in \mathcal{S}(\mathbf{R}^{4j})$. The creation operator B_i can be made into a field by translations

$$B_i(x) = U(x)B_iU^{-1}(x), \tag{1.2}$$

where $U(x) = U(x, 1)$. Ruelle has proved that the truncated Wightman functions W_T formed from the fields B_i satisfy the following inequalities. Define

$$W_T^{\varphi}(a_1, \cdots, a_n) = \int (\Psi_0, B_1(x_1 + a_1) \cdots B_n(x_n + a_n)\Psi_0)_T \varphi(x_1, \cdots, x_n)\, dx_1 \cdots dx_n, \tag{1.3}$$

where $a_j = (a_j^0, \mathbf{a}_j)$. Then there exists a number A such that

$$|W_T^{\varphi}(a_1, \cdots, a_n)| < A/(1 + \lambda)^N, \tag{1.4}$$

where

$$\lambda = \max_{i,j} |\mathbf{a}_i - \mathbf{a}_j|$$

and all the a_j^0 are equal.

The arguments of Jost[6] or Ruelle lead to the following bounds for any smooth solution of the Klein–Gordon equation: Given $f(\mathbf{x}, t)$, there exists a number A (depending of course on f) such that

$$|f(\mathbf{x}, t)| < At^{-\frac{3}{2}}, \tag{1.5}$$

$$\int |f(\mathbf{x}, t)|\, d^3\mathbf{x} < A(1 + |t|)^{\frac{3}{2}}. \tag{1.6}$$

Suppose B, acting on the vacuum, creates a state whose one-particle component has wavefunction $h(\mathbf{p})$, that is,

$$(\Psi_{\mathbf{p}}, B\Psi_0) = h(\mathbf{p}); \tag{1.7}$$

then the time-dependent creation operator

$$B_{f\alpha}(t) = \int_{x^0=t} B_\alpha(\mathbf{x}, x^0) \overleftrightarrow{\frac{\partial}{\partial x^0}} f_\alpha(x, x_0)\, d^3x, \tag{1.8}$$

where $a\overleftrightarrow{\partial} b = a\partial b - (\partial a)b$, creates a state with wavefunction $\alpha(\mathbf{p})h(\mathbf{p})$, if the solution f_α of the Klein–Gordon equation has the form

$$f_\alpha(x) = \int e^{ipx}\theta(p_0)\delta(p^2 - m^2)\alpha(\mathbf{p})\, d^4p. \tag{1.9}$$

The method of Haag[1] now shows that the limit states

$$\lim_{t \to \pm\infty} B_{f_1}(t) \cdots B_{f_n}(t)\Psi_0 \tag{1.10}$$

exist in norm.

Instead of using the creation operators (1.8), we could define other time-dependent operators

$$B_\alpha^n(t) = \int \delta(n \cdot x - t)B(\mathbf{x}, x^0) \overleftrightarrow{\frac{\partial}{\partial x_\mu}} n_\mu f_\alpha(x)\, d^4x, \tag{1.11}$$

where n_μ is a timelike vector. We are thus integrating over a spacelike plane with normal given by the vector n_μ. A simple argument shows that, acting on the vacuum, $B_\alpha^n(t)$ creates the same one-particle state as $B_\alpha(t)$, which happens to be independent of time. What can we then say about the convergence of the states

$$\Psi(n, t) = B_{f_1}^n(t)B_{f_2}^n(t) \cdots B_{f_k}^n(t)\Psi_0 \tag{1.12}$$

as $t \to \pm\infty$? We can immediately assert the convergence of these states using the explicit Lorentz covariance of the theory. For there exists at least one Lorentz transformation Λ such that $\Lambda n = (1, 0, 0, 0)$, a purely timelike vector. Then we may write

$$B_\alpha^n(t) = \int_{x^0=t} d^3x B_\alpha(\Lambda x) \overleftrightarrow{\frac{\partial}{\partial x^0}} f(\Lambda x). \tag{1.13}$$

To put this in the form used above, we first note that $f_\Lambda(x) = f(\Lambda x)$ defines a new solution of the Klein–Gordon equation, which is therefore bounded by $At^{-\frac{3}{2}}$ if $t = x^0$ (A depending on λ). Secondly, in the proof of convergence, we need consider only the expectation values of $B_f^n(t)$ in the vacuum state, and this involves a study of

$$\langle B_1(\Lambda x_1) \cdots B_k(\Lambda x_k)\rangle_0^T = \langle V(\Lambda)^{-1}B_1(\Lambda x_1)V(\Lambda) \times V^{-1}(\Lambda)B_2(\Lambda x_2) \cdots B_k(\Lambda x_k)V(\Lambda)\rangle_0^T, \tag{1.14}$$

where $U(a, \Lambda) = U(a)V(\Lambda)$, say. Now we may rewrite this as

$$\langle B_1(\Lambda x_1) \cdots B_k(\Lambda x_k)\rangle_0^T = \langle B_1^\Lambda(x_1) \cdots B_k^\Lambda(x_k)\rangle_0^T, \tag{1.15}$$

where

$$B^\Lambda = V(\Lambda)BV^{-1}(\Lambda) \tag{1.16}$$

and

$$B^\Lambda(x) = U(x)B^\Lambda U^{-1}(x). \tag{1.17}$$

So far the argument does not use the form (1.1) for the operators B, and could apply also the algebraic formulation of Haag and Araki.[8] On using (1.1), we see that

$$B_\Lambda = \sum_j \int g_{ij\Lambda^{-1}}(x_1, \cdots, x_j) \times \phi(x_1) \cdots \phi(x_j)\, dx_1 \cdots dx_j, \tag{1.18}$$

which is again of the same form. Hence, the inequalities (1.4), (1.5), and (1.6) are sufficient to show the convergence of the states (1.12) to limit states $\Psi^{\pm}(n)$, which have as much right as $\Psi^{\pm}$ to be regarded as the asymptotic states of the theory, containing k particles with wavefunctions specified by the one-particle states produced by the creation operators. The argument given here is not sufficient to prove the convergence of similar products of creation operators

[8] See H. Araki, *Local Quantum Theory* (W. A. Benjamin, Inc., New York, to be published).

obtained from an arbitrary surface σ, which is not flat. Even more important, we do not know that $\Psi^{\pm}(n)$ is independent of n. As a vector it might not be even differentiable in the parameters defining n. The usual argument[2,4] considers $(d\Psi/dn)(n,t)$, showing that for $t \to \pm\infty$ this state converges to zero in norm. This would in fact be sufficient if the limit is uniform in the parameters η of n, that is, if for any ϵ, there exists a t_0 such that

$$\|(d\Psi/dt)(n,t)\| < \epsilon$$

for all $t > t_0$ and all n in some set. For then

$$\begin{aligned}\|\Psi(\eta,t) - \Psi(0,t)\| &= \left\| \int_0^{\eta} \frac{d\Psi(\eta,t)}{d\eta}\, d\eta \right\| \\ &\leq \int_0^{\eta} \left\| \frac{d\Psi(\eta,t)}{d\eta} \right\| dt \\ &\leq \epsilon\eta \quad \text{if} \quad t > t_0,\end{aligned}$$

showing that

$$\Psi(\eta,t) \to \Psi^{\pm} = \lim_{t\to\pm\infty} \Psi(0,t).$$

Here η is a parameter of the normal: $n = (\cosh\eta, \sinh\eta, 0, 0)$, say. In the next section we show how this uniform convergence follows from the corresponding uniform bounds similar to (1.4), (1.5), and (1.6); these bounds are derived in Sec. 3.

2. THE UNIFORM CONVERGENCE LEMMA

In order to prove that $d\Psi(\eta,t)/d\eta$ converges uniformly as $t \to \pm\infty$, we use the following strengthened forms of (1.14), (1.15), and (1.16).

Theorem 1: Let B_j, $j = 1, 2, \cdots, n$ be "local" quantities, that is, operators such that $B_j(x) = U(x)B_jU(-x)$ commutes with $B_k(y)$ if $(x-y)^2 < -k_{jk}^2$ [we use the timelike metric $x^2 = (x^0)^2 - \mathbf{x}^2$]. Then define, for $\varphi \in \mathcal{S}(\mathbf{R}^{4n})$,

$$W_T^{\varphi}(a_1, \cdots, a_n) = \int \varphi(x_1, \cdots, x_n) \times W_T(x_1 + a_1, \cdots, x_n + a_n)\, dx_1 \cdots dx_n, \quad (2.1)$$

where

$$W(x_1, \cdots, x_n) = (\Psi_0, B_1(x_1)B_2(x_2)\cdots B_n(x_n)\Psi_0) \quad (2.2)$$

and W_T is the usual truncated function defined in terms of W. Then for any $\alpha > 1$ and any N, there exists a number $C(\alpha)$ such that

$$[1 + \lambda]^N W_T^{\varphi}(a_1, \cdots, a_n) < C, \quad (2.3)$$

where

$$\lambda = \max_{j,k} |\mathbf{a}_j - \mathbf{a}_k|,$$

and the $a_1, \cdots, a_n$ are only restricted by

$$|\mathbf{a}_i - \mathbf{a}_j|^2 \geq \alpha^2 |a_i^0 - a_j^0|^2. \quad (2.4)$$

Remark: The "local" operators could come from a Haag field[8] just as well as being polynomials in a Wightman field.

Theorem 2: Let f be a smooth solution of the Klein–Gordon equation, and W a compact neighborhood of the identity in the Poincaré group. Then there exists a number A such that

$$|f_\Lambda(\mathbf{x},t)| = |f(\Lambda x)| \leq A(1 + |t|)^{-\frac{3}{2}}, \quad (2.5)$$

$$|\mathbf{x}|\,|f_\Lambda(\mathbf{x},t)| \leq A(1 + |t|)^{-\frac{1}{2}}, \quad (2.6)$$

for all $\mathbf{x}$ and t, and all $\Lambda \in W$.

Theorem 3: Given f and W as above, there exists a number A such that

$$\int |f(\Lambda x)|\, d^3x < A(1 + |t|)^{\frac{3}{2}} \quad (2.7)$$

for all $\Lambda \in W$ and all $x = (\mathbf{x}, t)$.

In discussing now the uniform convergence of $d\Psi(\eta,t)/d\eta$, we first remark that $\Psi(\eta,t)$ is a differentiable vector in η. This comes about because the creation operators in the form (1.13) involve the parameter Λ only in the test function, which is a smooth function. The differential coefficient is then given by the usual rules for the product:

$$\frac{d\Psi(\eta,t)}{d\eta} = \sum_{j=1}^{k} B_{f_1}^n(t) \cdots \frac{dB_{f_j}^n(t)}{dt} \cdots B_{f_k}^n(t)\Psi_0. \quad (2.8)$$

As remarked, $B_{f_j}^n(t)\Psi_0$ is a one-particle state, independent of both n and t; it follows that

$$[dB_{f_j}^n(t)/d\eta]\Psi_0 = 0. \quad (2.9)$$

Haag's method is to expand the norm of $d\Psi(\eta,t)/d\eta$, which is a $2k$-point Wightman function, in terms of truncated functions. It is clear that in any nonzero contribution, $dB_{f_j}^n(t)/dt$ cannot be next to the vacuum, and so must occur in at least three- or four-point function. The same goes for the conjugate $dB_{f_j}^n(t)^*/d\eta$ which occurs in the scalar product

$$\|d\Psi/d\eta\|^2 = (d\Psi/d\eta, d\Psi/d\eta). \quad (2.10)$$

In discussing the transformation Λ such that $\Lambda n = (1, 0, 0, 0)$, we may assume, without loss in generality, that

$$(\Lambda x)^0 = x^0 \cosh\eta + x^1 \sinh\eta, \quad (2.11)$$

$$(\Lambda x)^1 = x^0 \sinh\eta + x^1 \cosh\eta, \quad (2.12)$$

$$(\Lambda x)^2 = x^2, \quad (2.13)$$

$$(\Lambda x)^3 = x^3, \quad (2.14)$$

where $|\eta| \leq \eta_0$ determines the compact neighborhood W of the origin of the Poincaré group. Using the form (1.13) for $B_\alpha^n(t)$, it is clear that $dB_\alpha^n/d\eta(t)$ is a sum of operators of the form

$$B^{(\cdot)}(\Lambda x)x_i e^{\pm\eta} f^{(\cdot)}(\Lambda x), \tag{2.15}$$

where $i = 0, 1$ and the symbol $(\cdot)$ means possible differentiation with respect to x_0 or x_1. The differentiated functions satisfy the same properties of Theorems 1, 2, and 3 needed for the proof. We can show that the n point product of such objects can be uniformly bounded as follows. Only two of the terms in the product will involve $dB/d\eta$, which brings down the extra power of $|\mathbf{x}|$ or t. Then a typical truncated function is

$$\int (\Psi_0, B_1^{(\cdot)}(\Lambda x)^* \cdots B_k^{\cdot}(\Lambda x)\Psi_0)_T f_1^{(\cdot)}(\Lambda x_1) \cdots x_i f_i^{(\cdot)}(\Lambda x_i) \times e^{\pm\eta} f_k^{(\cdot)}(\Lambda x_k)\, dx_1 \cdots dx_k, \quad k \geq 3, \tag{2.16}$$

involving one $dB/d\eta$, or

$$\int (\Psi_0, B_1^{(\cdot)}(\Lambda x)^* \cdots B_k^{\cdot}(\Lambda x)\Psi_0)_T f_1^{(\cdot)}(\Lambda x_1) \cdots x_i \times f_i^{(\cdot)}(\Lambda x_i) \cdots x_j f_j^{(\cdot)}(\Lambda x_j) \cdots f_k^{(\cdot)}(\Lambda x_k), \quad k \geq 4, \tag{2.17}$$

involving both $dB/d\eta$ and $dB^*/d\eta$. These may be bounded uniformly in η, by using the theorems. Thus

$$|(\Psi_0, B_1(\Lambda x_1)^* \cdots B_k(\Lambda x_k))| < [A/(1 + \lambda_\eta)^N], \tag{2.18}$$

$$\lambda_\eta = \max_{i,j} |\Lambda \mathbf{x}_i - \Lambda \mathbf{x}_j| \geq \max_{i,j} |\mathbf{x}_i - \mathbf{x}_j|$$

since all the points $x_1, \cdots, x_k$ have the same time component. Putting

$$\lambda = \max_{i,j} |\mathbf{x}_i - \mathbf{x}_j|, \tag{2.19}$$

we can uniformly bound (2.18) by $A/(1 + \lambda)^N$, whose integral over $R^{3(k-1)}$ converges to a constant, say C. Thus (2.16) and (2.17) are bounded by

$$\int |f_1^{(\cdot)}(\Lambda x_1)|\, d^3\mathbf{x}_1 \int W_T(\Lambda(x_1 - x_2), \Lambda(x_2 - x_3), \cdots) \times (1 + t)^{-\frac{3}{2}(k-1)}(1 + |t|)\cosh \eta_0\, d(x_1 - x_2) \cdots, \quad k = 3, \tag{2.20}$$

and

$$\int |f_1^{(\cdot)}(\Lambda x_1)|\, d^3\mathbf{x}_1 \int W_T(\Lambda(x_1 - x_2), \Lambda(x_2 - x_3) \cdots) \times (1 + t)^{-\frac{3}{2}(k-1)}(1 + |t|)^2 \cosh \eta_0\, d(\mathbf{x}_1 - \mathbf{x}_2) \cdots, \quad k \geq 4, \tag{2.21}$$

giving the usual bounds, uniformly in Λ; that is, the functions with one $dB/d\eta$ are bounded by

$$A(1 + |t|)^{-\frac{3}{2}(k-1)+1},$$

and those with both $dB/d\eta$ and $dB^*/d\eta$ by

$$A(1 + |t|)^{-\frac{3}{2}(k-1)+2}.$$

Both of these tend to zero as $t \to \infty$.

3. PROOF OF THEOREMS

We are going to consider solutions of the form

$$f(x) = \int \theta(p_0)\delta(p^2 - m^2)\tilde{f}_0(\mathbf{p})e^{ipx}\, d^4p, \tag{3.1}$$

where $f_0 \in \mathcal{S}$. It is clear that $f(x)$ may be written as a convolution with the singular Green's function

$$G_0(\mathbf{x}, t) = \int \theta(p_0)e^{i(\omega t - \mathbf{p}\cdot\mathbf{x})}\delta(p^2 - m^2)\, d^4p, \tag{3.2}$$

namely

$$f(\mathbf{x}, t) = \int G_0(\mathbf{x} - \boldsymbol{\xi}, t) f_0(\boldsymbol{\xi})\, d^3\xi. \tag{3.3}$$

In order to be able to bound integrals of this type, Jost's idea was to avoid the singular Green's function, replacing it by the continuous function $G(x)$ whose Fourier transform is

$$\tilde{G}(p) = (1 + p_0)^{-3}\theta(p_0)\delta(p^2 - m^2). \tag{3.4}$$

Naturally $f_0(\xi)$ in (3.3) becomes replaced by $g(\xi)$, where $\tilde{g}(\mathbf{p}) = (1 + p_0)^3\tilde{f}_0(\mathbf{p})$. Because $f_0 \in \mathcal{S}$, we have $\tilde{g} \in \mathcal{S}$.

Instead of $(1 + p_0)^{-3}$, any other smooth decreasing function of $\mathbf{p}$ would do; Jost's choice is convenient because of the identity

$$\frac{2}{(p_0 + 1)^3} = \int_0^\infty \alpha^2 e^{-\alpha} e^{-\alpha p_0}\, d\alpha, \tag{3.5}$$

which means that

$$G(\mathbf{x}, t) = c\int_0^\infty \alpha^2 e^{-\alpha} F(\alpha, \mathbf{x}, t)\, d\alpha, \tag{3.6}$$

where

$$F(\alpha, \mathbf{x}, t) = \int e^{-\alpha p_0} e^{i(wt - \mathbf{p}\cdot\mathbf{x})}\theta(p_0)\delta(p^2 - m^2)\, dp^4. \tag{3.7}$$

Using the method and some results of [6], we first put uniform bounds on G. We put $r = |\mathbf{x}|$.

Lemma 1: For any $\sigma > 1$ and $N > 0$ there exists a number A such that

$$|G(\mathbf{x}, t)| < A(\sigma, N)/(1 + r)^N \tag{3.8}$$

for all x with $r > \sigma t$.

Proof. In the course of Ref. 6 it is shown that there is a constant B such that

$$F(\alpha, \mathbf{x}, t) < B\frac{1 + \alpha^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}|\zeta|^{-\frac{3}{2}}e^{-\mathrm{Re}\,\zeta}, \tag{3.9}$$

where

$$\xi = (r^2 - t^2 + \alpha^2 - 2i\alpha t)^{\frac{1}{2}}. \tag{3.10}$$

Putting $u = r^2 - t^2 + \alpha^2$, $v = -2\alpha t$, one obtains for $r^2 \geq \sigma^2 t^2$

$$|\xi|^2 = (u^2 + v^2)^{\frac{1}{2}} \geq u \geq (1 - 1/\sigma^2)r^2 \tag{3.11}$$

and

$$\operatorname{Re} \zeta = \tfrac{1}{2}\big(u + (u^2 + v^2)^{\frac{1}{2}}\big) \geq u \geq (1 - 1/\sigma^2)r^2. \tag{3.12}$$

Therefore, from (3.9),

$$|F(\alpha, \mathbf{x}, t)| < B[(1 + \alpha^{\frac{1}{2}})/\alpha^{\frac{1}{2}}](1 - 1/\sigma^2)^{-\frac{3}{4}} r^{-\frac{3}{2}} \times e^{-(1-1/\sigma^2)^{\frac{1}{2}} r}. \tag{3.13}$$

Thus we obtain

$$|G(\mathbf{x}, t)| \leq C \frac{e^{-(1-\sigma^{-2})^{\frac{1}{2}} r}}{r^{\frac{3}{2}}} \int_0^\infty \frac{1 + \alpha^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}} \alpha^2 e^{-\alpha}\, d\alpha, \tag{3.14}$$

from which the result follows, since G is a bounded function at $r = 0$ and the integral in (3.14) converges. We have in fact shown rather more than stated in the lemma since exponential decrease is faster than "rapid decrease" in general. But we lose this information in the next lemma.

Lemma 2: Let $f(\mathbf{x}, t)$ be a solution of the Klein–Gordon equation of the form (3.1). Then, for any $\sigma > 1$ and any N, there exists a number $A(\sigma, N)$ such that

$$|f(\mathbf{x}, t)| < A(\sigma, N)/(1 + r)^N \tag{3.15}$$

for all x such that $r \geq \sigma t$.

Proof. We have

$$f(x) = \int g(\xi) G(\mathbf{x} - \xi, t)\, d^3\xi \tag{3.16}$$

with $g \in \mathcal{S}$. The integral can be split into two parts I_1 and I_2; in the first, ξ is such that $(1 < \sigma' < \sigma)$

$$\text{(I)} \qquad |\mathbf{x} - \xi| \geq \sigma' t, \tag{3.17}$$

and in the second

$$\text{(II)} \qquad |\mathbf{x} - \xi| \leq \sigma' t.$$

Thus

$$|I_1| = \int_1 g(\xi) G(\mathbf{x} - \xi, t)\, d^3\xi \leq \int_{R^3} |g(\xi)| \frac{A(\sigma', N + 3)\, d^3\xi}{(1 + |\mathbf{x} - \xi|)^{N+3}} \tag{3.18}$$

by (3.8). The integral in (3.18) as a function of $\mathbf{x}$ is bounded by an expression of the form $A(1 + r)^{-N}$. In the second term of the integral, I_2, we have the condition

$$|\mathbf{x} - \xi| \leq \sigma' t \quad \text{and} \quad |\mathbf{x}| \geq \sigma t.$$

Thus we get the lower bound for $|\xi|$:

$$\mathbf{x}^2 - 2\mathbf{x} \cdot \xi + \xi^2 \leq \sigma'^2 t^2,$$

giving

$$-2\mathbf{x} \cdot \xi + \xi^2 \leq \sigma'^2 t^2 - r^2 \leq 0.$$

Thus

$$-\xi^2 + 2\mathbf{x} \cdot \xi \geq r^2 - \sigma'^2 t^2 \geq r^2(1 - \sigma'^2/\sigma^2) \geq 0.$$

Therefore

$$2\mathbf{x} \cdot \xi \geq r^2[(\sigma^2 - \sigma'^2)/\sigma^2],$$

giving

$$|\xi| \geq \tfrac{1}{2} r[(\sigma^2 - \sigma'^2)/\sigma^2] = \gamma r,$$

say, by using

$$2\mathbf{x} \cdot \xi \leq 2\,|\mathbf{x}|\,|\xi|.$$

But we know $g \in \mathcal{S}$, so that for N given, there exists a C such that

$$|g(\xi)| < C/(1 + |\xi|)^{N+3}.$$

Hence

$$|I_2| = \int_2 g(\xi) G(\mathbf{x} - \xi, t)\, d^3\xi \leq \int_{|\xi| \geq \gamma r} \frac{CM}{(1 + |\xi|)^{N+3}}\, d^3\xi,$$

where M bounds $G(\mathbf{x}, t)$. The last integral is obviously bounded by an expression of the form $C/(1 + r)^N$. This proves Lemma 2.

Proof of Theorem 2. By rotation invariance of the inequalities, it is sufficient to prove the theorem for Λx given by (2.11)–(2.14), with $|\eta| \leq \eta_0$. We prove it for $\eta > 0$ and $t > 0$. The other cases are proved similarly.

The inequality (1.5), proved in Ref. 6, shows that there exists a constant C such that

$$|f(\Lambda x)| < \frac{C}{1 + |t \cosh \eta + x^1 \sinh \eta|^{\frac{3}{2}}}. \tag{3.19}$$

Divide space into two regions; in the first

$$t \cosh \eta + x^1 \sinh \eta > \rho_0 t, \quad 0 < \rho_0 < 1, \tag{3.20}$$

where ρ_0, independent of $\mathbf{x}$, t, and η, is chosen later. Equation (3.20) with (3.19) gives

$$|f(\Lambda x)| < C/[1 + (\rho_0 t)^{\frac{3}{2}}]$$

in the first region, which can clearly be put in the required form. In the second region

$$|t \cosh \eta + x^1 \sinh \eta| < \rho_0 t, \tag{3.21}$$

giving, certainly,

$$t \cosh \eta + x^1 \sinh \eta < \rho_0 t,$$

so that

$$0 < t(\cosh \eta - \rho_0) < -x^1 \sinh \eta,$$

showing $x^1 < 0$ as $\eta > 0$, and

$$-x^1 > t[(\cosh\eta - \rho_0)/\sinh\eta].$$

Hence

$$\begin{aligned}-x_1\cosh\eta - t\sinh\eta &> (t(\cosh\eta-\rho_0)\cosh\eta - t\sinh^2\eta)(\sinh\eta)^{-1}\\ &= t\left(\frac{1-\rho_0\cosh\eta}{\sinh\eta}\right) > \frac{1-\rho_0\cosh\eta_0}{\sinh\eta_0}\,t,\\ &> 0, \quad (3.22)\end{aligned}$$

provided we choose $\rho_0 < 1/\cosh\eta_0$. But, by (3.21),

$$t > |t\cosh\eta + x_1\sinh\eta|/\rho_0,$$

and therefore, from (3.22),

$$\begin{aligned}x_2^2 + x_3^2 + |x_1\cosh\eta + t\sinh\eta|^2 &\geq |x_1\cosh\eta + t\sinh\eta|^2 \geq \left(\frac{1-\rho_0\cosh\eta_0}{\sinh\eta_0}\right)^2 t\\ &\geq \left(\frac{1-\rho_0\cosh\eta_0}{\rho_0\sinh\eta_0}\right)^2 |t\cosh\eta + x_1\sinh\eta|.\end{aligned}$$

Choose ρ_0 such that $(1-\rho_0\cosh\eta_0)/\rho_0\sinh\eta_0 > 1$. This is clearly possible in the range $0 < \rho_0 < 1$ since the expression diverges as $\rho_0 \to 0$.

Then we can apply Lemma 2 to show that, in the second region,

$$\begin{aligned}|f(\Lambda x)| &< \frac{A((1-\rho_0\cosh\eta_0)/\rho_0\sinh\eta_0, 2)}{(1 + x_2^2 + x_3^2 + |x_1\cosh\eta + t\sinh\eta|^2)^2}\\ &< \frac{A_2}{\{1 + [(1-\rho_0\cosh\eta_0)/\sinh\eta_0]^2 t^2\}^2}.\end{aligned}$$

This clearly implies the inequality (2.5). To prove (2.6), if $|\mathbf{x}| < \rho_1 t$ for some ρ_1 independent of $\mathbf{x}$, t, and Λ, then by the result just proved,

$$|\mathbf{x}|\,|f(\Lambda x)| < \rho_1 t A(1 + |t|)^{-\frac{3}{2}},$$

as required. We need consider only the points with $|\mathbf{x}| \geq \rho_1 t$. This can happen in two ways:

$$|x_1| \geq (\rho_1/\nu)t \quad (3.23)$$

or

$$|x_1| \leq \rho_1/\nu \quad \text{and} \quad (x_2^2 + x_3^2)^{\frac{1}{2}} > [(\nu-1)/\nu]\rho_1 t. \quad (3.24)$$

We choose ν, independent of Λ, $\mathbf{x}$, and t, later on.

If $|x_1| > (\rho_1 t/\nu)$, then

$$\begin{aligned}|x_1\cosh\eta + t\sinh\eta| &\geq |x_1\cosh\eta| - t\sinh\eta\\ &\geq |x_1\cosh\eta| - |x_1\sinh\eta| - t\sinh\eta - t\cosh\eta + (t\cosh\eta + |x_1\sinh\eta|)\\ &\geq |t\cosh\eta + x_1\sinh\eta|\end{aligned}$$

provided

$$|x_1|(\cosh\eta - \sinh\eta) \geq t(\cosh\eta + \sinh\eta),$$

that is, provided $|x_1| \geq e^{2\eta}t$ for all $|\eta| < \eta_0$ and $\mathbf{x}$, t of the region. This holds, in view of (2.23), if we choose $\rho_1/\nu = e^{2\eta_0}$. In this region, by Lemma 2, Eq. (3.15), we have

$$\begin{aligned}|\mathbf{x}|\,|f(\Lambda x)| &< \frac{A|\mathbf{x}|}{(1 + x_2^2 + x_3^2 + |x_1\cosh\eta + t\sinh\eta|^2)^N}\\ &< \frac{B}{1 + (x_2^2 + x_3^2 + x_1^2)^{2N-1}} < \frac{B}{1 + |x_1|^{2N-1}}\\ &< \frac{B}{1 + [(\rho_1/\nu)t]^{2N-1}},\end{aligned}$$

which shows the result for this region.

In the region (3.24) we have

$$\begin{aligned}x_2^2 + x_3^2 + (x_1\cosh\eta + t\sinh\eta)^2 &\geq [(\nu-1)/\nu]^2\rho_1^2 t^2 + (x_1\cosh\eta + t\sinh\eta)^2\\ &\geq [(\nu-1)/\nu]^2\rho_1^2 t^2 \geq \rho_2(t\cosh\eta_0 + |x_1|\sinh\eta_0)^2 \quad (3.25)\end{aligned}$$

provided ρ_1, ρ_2 are chosen such that $\rho_2 > 1$ and

$$[(\nu-1)/\nu]^2\rho_1^2 \geq \rho_2^2[\cosh\eta_0 + \sinh\eta_0(\rho_1/\nu)]^2,$$

consistent with $\rho_1 = \nu e^{2\eta_0}$, since $|x_1| < \rho_1/\nu$ in the region (3.24). This is clearly possible if we make ν large enough. Then, from (3.25),

$$\begin{aligned}x_2^2 + x_3^2 + (x_1\cosh\eta + t\sinh\eta)^2 &\geq \rho_2^2(\cosh\eta_0 t + \sinh\eta_0|x_1|)^2\\ &\geq \rho_2^2|t\cosh\eta + x_1\sinh\eta|^2.\end{aligned}$$

Thus, again by Lemma 2, we can bound $r\,|f(\Lambda x)|$ by

$$\begin{aligned}&\left|(x_2^2 + x_3^2) + \frac{\rho_1^2 t^2}{\nu^2}\right|^{\frac{1}{2}} \max|f(\Lambda x)|\\ &\leq \frac{[(x_2^2 + x_3^2) + (\rho_1^2/\nu^2)t^2]^{\frac{1}{2}}A}{[1 + x_2^2 + x_3^2 + (x_1\cosh\eta + t\sinh\eta)^2|^{N+1}}\\ &\leq C\frac{\left[1 + \dfrac{\rho_1^2}{\nu^2\{[(\nu-1)/\nu]\rho_1\}^2}\right]^{\frac{1}{2}}}{[1 + \{[(\nu-1)/\nu]\rho_1\}^2 t^2]^N} \quad \text{for some } C,\end{aligned}$$

which proves Theorem 2.

Proof of Theorem 3. The idea of this proof is the classical one, that the volume of space inside which $f(\mathbf{x}, t)$ is not very small, increases at worst like $|t|^3$ as $t \to \pm\infty$. We first prove the result for $\eta > 0$, $t > 0$; the other cases are proved similarly.

We split the integral into two main parts, each with a finite number of subdivisions. Then each integral is shown to be bounded uniformly in Λ by a function of the form $A(1 + |t|)^{\frac{3}{2}}$. In the part (a) we have

$$\text{(a)} \quad x_2^2 + x_3^2 + (x_1\cosh\eta + t\sinh\eta)^2 \geq \sigma^2(x_1\sinh\eta + t\cosh\eta)^2, \quad (3.26)$$

where $\sigma > 1$ is chosen later. In part (b) we have

$$\text{(b)}\quad x_2^2 + x_3^2 + (x_1 \cosh\eta + t \sinh\eta)^2 \leq \sigma^2(x_1 \sinh\eta + t\cosh\eta)^2. \quad (3.27)$$

The region (a) is again split up into the regions (a1) and (a2), given by

$$\text{(a1)}\qquad |x_1 \cosh\eta + t\sinh\eta| > \alpha |x_1|, \qquad (3.28)$$

$$\text{(a2)}\qquad |x_1 \cosh\eta + t\sinh\eta| < \alpha |x_1|, \qquad (3.28a)$$

α to be chosen later. In the region (a1), clearly, by Lemma 2,

$$|f(\Lambda x)| < \frac{A(\sigma, N)}{1 + [x_2^2 + x_3^2 + |x_1\cosh\eta + t\sinh\eta|^2]^N} < \frac{A(\sigma, N)}{1 + [x_2^2 + x_3^2 + \alpha^2 x_1^2]^N} \qquad (3.29)$$

by (a1). Therefore

$$\int |f(\Lambda x)|\, d^3x < C,$$

some number, from which (2.7) follows. Expanding the inequality (a2), we arrive at four cases:

$$\text{(a2i)}\quad 0 \leq x_1\cosh\eta + t\sinh\eta \leq \alpha x_1;$$

$$\text{(a2ii)}\quad 0 \leq x_1\cosh\eta + t\sinh\eta \leq -\alpha x_1;$$

$$\text{(a2iii)}\quad 0 \leq -x_1\cosh\eta - t\sinh\eta \leq \alpha x_1;$$

$$\text{(a2iv)}\quad 0 \leq -x_1\cosh\eta - t\sinh\eta \leq -\alpha x_1.$$

The first and third are impossible, since $\alpha < 1$ and $t > 0$, $\eta > 0$. From (a2ii) we see

$$t\sinh\eta \geq -x_1\cosh\eta = |x_1|\cosh\eta \geq |x_1|,$$

so

$$|x_1| \leq t\sinh\eta_0. \qquad (3.30)$$

In (a2iv) $x_1 < 0$ and

$$|x_1|(1-\alpha) \leq |x_1|(\cosh\eta - \alpha) = -x_1(\cosh\eta - \alpha) \leq t\sinh\eta \leq t\sinh\eta_0 \qquad (3.31)$$

showing that, in (a2),

$$|x_1| \leq [\sinh\eta_0/(1-\alpha)]t = \beta t,$$

say by combining (3.30) and (3.31).
It follows that

$$\int |f(\Lambda x)|\, d^3x \leq \int_0^{\beta t} dx_1 \int \frac{A(\sigma, N)\, dx_2\, dx_3}{1 + [x_2^2 + x_3^2 + |x_1\cosh\eta + t\sinh\eta|^2]^N} \leq \int_0^{\beta t} dx_1 \int \frac{A(\sigma, N)\, dx_2\, dx_3}{1 + [x_2^2 + x_3^2]^N} \leq C\beta t \quad \text{for some } C.$$

This deals with the region (a). In region (b) we wish to show that $|\mathbf{x}|$ is bounded by a fixed multiple of t. Because of (3.27), $|x_1| < \delta |t|$ is sufficient for this for some δ independent of η. For then

$$|x_2| \leq \sigma |x_1 \sinh\eta + t\cosh\eta| \leq \sigma |x_1| \sinh\eta_0 + \sigma\cosh\eta_0 |t| \leq (\sigma\alpha \sinh\eta_0 + \sigma\cosh\eta_0)|t| = \epsilon |t|,$$

say, and the same for $|x_3|$. Then

$$\int |f(\Lambda x)|\, d^3x < \frac{C}{(1+t)^{\frac{3}{2}}} \int_{-\delta t}^{\delta t} dx_1 \int_{-\epsilon t}^{\epsilon t} dx_2 \int_{\epsilon t}^{\epsilon t} dx_3 < B(1 + |t|)^{\frac{3}{2}} \quad \text{for some } B.$$

Equation (3.27), valid in region (b), clearly implies

$$|x_1\cosh\eta + t\sinh\eta| \leq \sigma |x_1\sinh\eta + t\cosh\eta|.$$

Again there are four cases:

(bi)
$$0 < x_1\cosh\eta + t\sinh\eta < \sigma(x_1\sinh\eta + t\cosh\eta);$$

(bii)
$$0 < -x_1\cosh\eta - t\sinh\eta < \sigma(x_1\sinh\eta + t\cosh\eta);$$

(biii)
$$0 < x_1\cosh\eta + t\sinh\eta < \sigma(-x_1\sinh\eta - t\cosh\eta);$$

(biv)
$$0 < -x_1\cosh\eta - t\sinh\eta < \sigma(-x_1\sinh\eta - t\cosh\eta).$$

In (bi), if $x_1 < 0$, then $-x_1\cosh\eta < t\sinh\eta$ implies $|x_1| < t\sinh\eta_0$, as desired. So we may assume $x_1 > 0$. Then

$$x_1(\cosh\eta - \sigma\sinh\eta) < \sigma t\cosh\eta - t\sinh\eta < t\sigma\cosh\eta_0).$$

Let us choose $\sigma < \cosh\eta_0/\sinh\eta_0$. Then $\cosh\eta - \sigma\sinh\eta > 0$ for all $\eta < \eta_0$, and we have

$$|x_1| < t\,\frac{\sigma\cosh\eta_0}{\cosh\eta - \sigma\sinh\eta} < \frac{t\sigma\cosh\eta_0}{\cosh\eta_0 - \sigma\sinh\eta_0},$$

the desired result. In (bii), $x_1 > 0$ is not possible (we have $t > 0$, $\eta > 0$). Thus (bii) gives

$$|x_1| < |x_1|(\cosh\eta + \sigma\sinh\eta) = -x_1(\cosh\eta + \sigma\sinh\eta) < t(\sigma\cosh\eta + \sinh\eta),$$

so

$$|x_1| < t(\sigma\cosh\eta_0 + \sinh\eta_0),$$

proving the result for (bii). In (biii) x_1 must be negative, and then $-x_1\cosh\eta < t\sinh\eta$ leads to

$$|x_1| < t\sinh\eta_0.$$

In (biv) x_1 must be negative since

$$-x_1\sinh\eta - t\cosh\eta > 0.$$

But then

$$-x_1(\cosh\eta - \sigma\sinh\eta) < t(\sinh\eta - \sigma\cosh\eta)$$

is not possible since the left-hand side is positive and the right-hand side negative. This proves Theorem 3.

Proof of Theorem 1. The proof follows that of Ref. 2. Suppose we are considering a sequence of configurations for the $a_1, \cdots, a_n$ in which

$$\lambda = \max_{j,k} |\mathbf{a}_j - \mathbf{a}_k|$$

becomes large, and such that the maximum is attained for a fixed $j = i_0$, $k = i_0'$. Consider also the family $F(i_0, i_0')$ of all partitions of the set $(1, \cdots, n)$ into subsets X, X' such that $i_0 \in X$, $i' \in X'$. Define

$$\mu^2 = \max_{X \in F(i_0, i_0')} \left[\min_{i \in X, i' \in X'} |\mathbf{a}_i - \mathbf{a}_i'|^2 \right].$$

Suppose also that the configurations also are all such that this maximum is achieved for a fixed partition $X = X_0$, $X' = X_0'$ and for fixed $i = j_0$, $i' = j_0'$. Then $n\mu \geq \lambda$. As in Ref. 2, we set

$$X_0 = (i_1 < i_2 < \cdots < i_k)$$

and

$$X' = (i_{k'+1} < \cdots < i_{n'}),$$

and define a permutation

$$I = (1, 2, \cdots, n)$$

$$J = (i_1, i_2, \cdots, i_k, i_{k'+1}, \cdots, i_{n'})$$

with the permuted expectation values W_φ^I and W_ψ^J defined by permuting the fields appropriately. We then note that

$$F(x_1, \cdots, x_n) = W^I(x_1, \cdots, x_n) - W^J(x_1, \cdots, x_n) \quad (3.32)$$

vanishes whenever the differences $\xi_{ii'} = x_i - x_{i'}$ are sufficiently spacelike. Therefore $\varphi(x)$ does not contribute to

$$F_\varphi(a_1, \cdots, a_n) = \int \varphi(x_1, \cdots, x_n) F(x_1 + a_1, \cdots, x_n + a_n)\, dx$$

for any x such that

$$(x_i + a_i - x_{i'} - a_{i'})^2 < -K^2, \quad K = \max K_{ij}.$$

This is ensured for sufficiently large μ, provided that

$$\|x_i\|^2 = r^2 + x_0^2 < A^2\mu^2$$

for some A, independent of μ, chosen later. For then

$$\begin{aligned}(x_i - x_{i'})^2 &+ (a_i - a_{i'})^2 + 2(x_i - x_{i'})(a_i - a_{i'}) \\ &= (x_i^0 - x_{i'}^0)^2 - (\mathbf{x}_i - \mathbf{x}_i')^2 + (a_i^0 - a_{i'}^0) - (\mathbf{a}_i - \mathbf{a}_{i'})^2 \\ &\quad + 2(x_i^0 - x_{i'}^0)(a_i^0 - a_{i'}^0) - 2(\mathbf{x}_i - \mathbf{x}_i') \cdot (\mathbf{a}_i - \mathbf{a}_{i'}) \\ &\leq 4A^2\mu^2 + \mu^2(-1 + 1/\alpha^2) + 4A\mu(n\mu/\alpha) + 4A\mu \cdot n\mu \\ &< -K^2\end{aligned}$$

if A is small enough (depending only on α and K). Thus for large enough μ, i.e., large enough λ, greater than λ_0 say, depending only on α and K, this is negative. The rest of the proof goes as in Ref. 2.

4. CONCLUDING REMARKS

Because of the uniform nature of our bounds, the lemmas can be immediately applied to prove that certain smooth curved surfaces may be used to define creation operators. Let $\sigma(t)$ be a set of spacelike surfaces, with a normal at every point, which lies everywhere inside a compact timelike cone; then if $n = (\cosh\eta, \sinh\eta, 0, 0)$, we know there exists an η_0 such that $|\eta| \leq \eta_0$ on σ. The operators

$$B_\sigma^f(t) = \int_{\sigma(t)} d\sigma^\mu B(x) \frac{\overleftrightarrow{\partial}}{\partial x^\mu} n_\mu f_\alpha(x)$$

exist as creation operators, and

$$B_\sigma^f(t)\Psi_0$$

is a one-particle state independent of σ and t, as is well known. Now suppose the time dependence of $\sigma(t)$ is smooth enough so that $B_\sigma^f(t)$ is differentiable in t, and is such that the fixed point

$$(x_0(t, \sigma), 0, 0, 0) \in \sigma(t)$$

goes to infinity as $t \to \infty$. Then our method proves that

$$\lim_{t \to \pm\infty} B_\sigma^{f_1}(t) \cdots B_\sigma^{f_n}(t)\Psi_0$$

exist in norm. These states have a right to be called asymptotic states; it is therefore an important problem to prove that they are independent of the series of spacelike surfaces used to define the limit states. To prove this in general involves more work, though it is easy to show, as above, that it is independent of the orientation. Note that the surfaces $\sigma(t)$ need not be all the same shape, though the dependence on shape with time must be smooth.

To show the independence of the limit states on the sequence, we remark that the limit states may be obtained by following the sequence $t_n : x_0(t, \sigma_1) = n$ an integer; as $t \to \infty$, so $n \to \infty$, where

$$(x_0(t, \sigma_1)0, 0, 0) \in \sigma_2.$$

Both sequences converge. But there exists a set of other surfaces, smoothly interpolating between these two surfaces, i.e., a surface $\sigma(t)$ such that $\sigma(t_n) = \sigma_1(t_n)$ and $\sigma(t_{n+\frac{1}{2}}) = \sigma_2(t_{n+\frac{1}{2}})$ as surfaces. We know that the surfaces $\sigma(t)$ converge, showing the other limits must coincide.

The result of this paper, the independence of the states on the Lorentz frame, was obtained without

using the covariance of the theory under the Lorentz group. It is interesting that if the theory is covariant under the Lorentz group, then our result is essential to show that the asymptotic states transform covariantly. It would indicate that the Lorentz invariance of the S matrix is not independent of almost locality.

One might like to extend the results of scattering theory to states of the form

$$\Psi_{\sigma_1 \cdots \sigma_n}(t) = B^{f_1}_{\sigma_1}(t) B^{f_2}_{\sigma_1}(t) \cdots B^{f_n}_{\sigma_n}(t) \Psi_0 .$$

However, such states do not seem to converge; the reason is that if $\sigma_1 \neq \sigma_2$, there are always some points on σ_1 and σ_2 with a timelike difference. We cannot obtain suitable bounds in this case, unless σ_1 and σ_2 are very close together at ∞. We do not pursue this question.

ACKNOWLEDGMENTS

The author thanks W. F. Anderson for discussions and I. E. Segal for a prepublication copy of his arguments in Ref. 5, and for drawing his attention to some difficulties in Ref. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Unitary Representations of $U(2, 2)$ and Massless Fields*

BEHRAM KURŞUNOĞLU
Center for Theoretical Studies, University of Miami, Coral Gables, Florida

(Received 19 December 1966)

This paper contains a discussion of unitary irreducible representations of the group $U(2, 2)$ in terms of the noncompact algebra of creation and annihilation operators and some applications to massless fields. In particular, the $U(2, 2)$ algebra yields discrete values for p_4 (energy), one of its generators. The little group and wave equations of massless fields are also derived from the Lie algebra of $U(2, 2)$.

1. INTRODUCTION

THIS paper is a contribution to the explosion of group theoretical publications pertaining to elementary particle concepts. The present state of theoretical research on elementary particles seems to indicate that there exist ever increasing possibilities for the so-called "classification" of particles. Recent attempts[1] for the unification of internal and space–time symmetries into a single group theoretical structure aiming at an hypothesis of simultaneous charge, hypercharge, and spin independence of strong interactions (at high energy) have led to further discussions of the subject by others.[2] These authors have shown that there are some basic difficulties in the models proposed earlier.[1] In particular, if one adheres to the existing interpretations of the isotopic spin, then spin and isotopic spin assignments to various generators of the group $SU(3, 1)$ lead to noncommuting operators for the respective observables. Therefore, what remains as acceptable is the product of two commuting groups, i.e., the cover group is just the direct product of the Poincaré group with an internal symmetry group.

In the light of these investigations the fundamental issue appears to be the possible existence of a "noncompact symmetry group" whose unitary representations together with some reasonable physical assumptions on the nature of interactions of fields may provide a good beginning for particle physics.

We use here, as in the previous paper,[1] the techniques of creation and annihilation operators for the representation of the group $SU(2, 2)$, which is locally isomorphic to $SO(4, 2)$. Our discussion is confined only to unitary, irreducible representations.

2. REPRESENTATION OF $U(2, 2)$

In order to establish the method, we consider a special set of ten Hermitian operators satisfying the commutation relations for the inhomogeneous Lorentz group. These are given by p_μ (four translation operators), and by the relativistic definition of angular momenta,[3]

$$R_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu = i \langle x| FM_{\mu\nu} |p\rangle, \qquad (2.1)$$

where

$$F = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad |x\rangle = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad |p\rangle = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} \qquad (2.2)$$

and x_μ, p_ν ($\mu, \nu = 1, 2, 3, 4$) are subject to commutation relations

$$[x_\mu, p_\nu] = -i\hbar g_{\mu\nu}, \quad [x_\mu, x_\nu] = [p_\mu, p_\nu] = 0, \qquad (2.3)$$

with $g_{\mu\nu}$ being the elements of F. Every Lorentz matrix L satisfies the condition

$$\tilde{L}FL = F, \qquad (2.4)$$

where $\tilde{L}$ is the transposed form of L.

The operators x_μ and p_μ under a Lorentz transformation transform according to

$$|\acute{x}\rangle = L|x\rangle, \quad |\acute{p}\rangle = L|p\rangle. \qquad (2.5)$$

In a way similar to (2.1), we introduce complex creation and annihilation operators. For example, the

* This research is supported by the U.S. Air Force Office of Scientific Research, Washington D.C., U.S. Air Force Contract No. 49 (638)-1260.

[1] *Proceedings of the First Coral Gables Conference on Symmetry Principles at High Energy* (W. H. Freeman and Company, San Francisco, 1964). See also B. Kurşunoğlu, Phys. Rev. **135**, B761 (1964).

[2] W. D. McGlinn, Phys. Rev. Letters **12**, 467 (1964); F. Coester, M. Hamermesh, and W. D. McGlinn, Phys. Rev. **135**, B451 (1964); H. Bacry and J. Nuyts, Phys. Letters **12**, 2, 156 (1964); M. E. Mayer, H. S. Schnitzer, E. C. G. Sudarshan, R. Acharya, and M. Y. Han, Phys. Rev. **136**, B888 (1964); A. Beskow and U. Ottoson, Nuovo Cimento **34**, 248 (1964).

[3] B. Kurşunoğlu, *Modern Quantum Theory* (W. H. Freeman and Company, San Francisco, 1962). See p. 254, Eq. (VIII.8.3) also p. 50 for the definition of the 4×4 matrices $M_{\mu\nu}$, which are generators of rotations and Lorentz transformations. The matrices $M_{\mu\nu}$ constitute a nonunitary representation of the homogeneous group. This book is hereafter referred to as MQT.

Hermitian generators of the homogeneous Lorentz group can be represented by [MQT, p. 257, Eq. (VIII.8.21)]

$$J_{\mu\nu} = \tfrac{1}{2}\langle a|\, \beta\sigma_{\mu\nu}\, |a\rangle, \tag{2.6}$$

where

$$\beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad |a\rangle = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix},$$

$$\langle a| = [a_1^\dagger, a_2^\dagger, a_3^\dagger, a_4^\dagger], \tag{2.7}$$

and the operators a_α, $a_\rho^\dagger$ ($\alpha, \rho = 1, 2, 3, 4$) satisfy the commutation relations

$$[a_\alpha, a_\rho^\dagger] = \beta_{\alpha\rho}, \quad [a_\alpha, a_\rho] = [a_\alpha^\dagger, a_\rho^\dagger] = 0 \tag{2.8}$$

with β being taken as the "metric" of the 4-dimensional complex space.

We could, if we wished, use two-component representations for the a's. For example, the commutation relations (2.8) can be replaced by the equivalent set

$$\begin{aligned} &[a_\lambda, a_\omega^\dagger] = [b_\lambda, b_\omega^\dagger] = \delta_{\lambda\omega}, \\ &[a_\lambda, b_\omega] = [a_\lambda^\dagger, b_\omega^\dagger] = 0, \\ &[a_\lambda, a_\omega] = [b_\lambda, b_\omega] = 0, \\ &\quad\vdots \end{aligned} \tag{2.8'}$$

where we put

$$a_3^\dagger = b_1, \quad a_4^\dagger = b_2$$

and the subscripts $\lambda, \omega = 1, 2$.

We are using a representation of γ's given by

$$\begin{aligned} [\gamma_\mu, \gamma_\nu]_+ &= -2g_{\mu\nu}, \quad \gamma_4 = i\beta, \\ \gamma_5 &= \gamma_1\gamma_2\gamma_3\gamma_4, \\ \Lambda_\pm &= \tfrac{1}{2}(1 \pm i\gamma_5), \\ \sigma_{\mu\nu} &= -\tfrac{1}{2}i[\gamma_\mu, \gamma_\nu], \\ \gamma_5\sigma_{\mu\nu} &= \tfrac{1}{2}\epsilon_{\mu\nu\alpha\beta}\sigma^{\alpha\beta}, \end{aligned}$$

and

$$\begin{aligned} g_{11} = g_{22} = g_{33} = -g_{44} = -1, \\ g_{j4} = g_{4j} = 0, \quad g_{kj} = 0, \quad k \neq l, \end{aligned}$$

where γ_j ($j = 1, 2, 3$) are Hermitian and γ_4 is anti-Hermitian.

The corresponding commutation and anticommutation relations are

$$[\tfrac{1}{2}\sigma_{\mu\nu}, \tfrac{1}{2}\sigma_{\alpha\beta}] = \tfrac{1}{2}i(g_{\alpha\nu}\sigma_{\mu\beta} + g_{\beta\nu}\sigma_{\alpha\mu} - g_{\alpha\mu}\sigma_{\nu\beta} - g_{\mu\beta}\sigma_{\alpha\nu}), \tag{2.9}$$

$$\tfrac{1}{2}[\sigma_{\mu\nu}, \sigma_{\alpha\beta}]_+ = -\gamma_5\epsilon_{\mu\nu\alpha\beta} + g_{\alpha\mu}g_{\beta\nu} - g_{\mu\beta}g_{\alpha\nu}, \tag{2.10}$$

$$[\sigma_{\mu\nu}, \gamma_5] = 0, \tag{2.11}$$

$$\tfrac{1}{2}[\sigma_{\mu\nu}, \gamma_\rho] = i(g_{\rho\nu}\gamma_\mu - g_{\rho\mu}\gamma_\nu), \tag{2.12}$$

$$\tfrac{1}{2}[\sigma_{\mu\nu}, \gamma_\rho]_+ = -i\epsilon_{\mu\nu\rho\beta}\gamma^\beta\gamma_5. \tag{2.13}$$

From the two representations (2.1) and (2.6), it follows that we may define a transformation operator S, in analogy to a Lorentz transformation L, which satisfies the condition

$$S^\dagger\beta S = \beta \tag{2.14}$$

in complex 4-dimensional space. However, the relation (2.14) implies that the 4×4 complex matrix S depends only on 16 free real parameters, so that the group in question is a 16-parameter group. All nonunitary 4-dimensional representations of $U(2, 2)$ must satisfy the matrix condition (2.14). This means that the S-transformations on 4-dimensional complex objects Z_α must leave the statement

$$|Z_1|^2 + |Z_2|^2 - |Z_3|^2 - |Z_4|^2$$

unchanged. The condition (2.14) is valid only for proper $U(2, 2)$ transformation. For improper transformations the right side of (2.14) should be replaced by $-\beta$. In this paper we are not concerned with the latter case.[4] Under a Lorentz transformation of the generators $J_{\mu\nu}$, the operator column vector $|a\rangle$ transforms according to

$$|\acute{a}\rangle = S\,|a\rangle, \tag{2.15}$$

but the correspondence here is in the sense of a homomorphism. The commutation relations (2.8) are invariant under S-transformations satisfying the condition (2.14).

A special type of S-transformations are gauge transformations of the type $\exp(i\phi)$. Furthermore, from (2.14) it follows that the determinant of an S-transformation is defined up to a phase factor. Hence the group of S-transformations can be decomposed according to $U = U_1 \times S_0 \times Z$, where U_1 is the one-dimensional unitary group and S_0 is the group of S-transformations with determinant $+1$. The factor Z is of the form $\exp(\tfrac{1}{2}i\pi n)$, $n = 1, 2, \cdots$, representing an invariant S-transformation subgroup of fourth order whose members consist of ± 1 and $\pm i$. This means that there are four types of vector operators a_α pertaining to the representations of the group $U(2, 2)$.[5]

In terms of the operators a_α and $a_\alpha^\dagger$, the Hermitian generators of $U(2, 2)$ for the positive energies are

---

[4] See Eq. (VIII.5.55) on p. 240, and Eqs. (VIII.8.21) and (VIII.8.22) on p. 257 of MQT. Equations (VIII.5.56) and (VIII.5.57) on p. 241 of MQT are examples of S-transformations. The operators γ_5, $i\gamma_\mu$, $\gamma_5\gamma_\mu$ are also generators of S-transformations.

[5] This is a special case of an arbitrary phase factor discussed in R. E. Marshak and E. C. G. Sudarshan, *Introduction to Elementary Particle Physics* (Interscience Publishers, Inc., New York, 1960), p. 59.

given by

$$J_{\mu\nu} = \tfrac{1}{2}\langle a|\, \beta\sigma_{\mu\nu}\, |a\rangle, \tag{2.16}$$

$$p_\mu^+ = -\langle a|\, \gamma_4\Lambda_+\gamma_\mu\, |a\rangle, \tag{2.17}$$

$$p_\mu^- = -\langle a|\, \gamma_4\Lambda_-\gamma_\mu\, |a\rangle, \tag{2.18}$$

$$\zeta = \tfrac{1}{2}\langle a|\, \beta\gamma_5\, |a\rangle, \tag{2.19}$$

$$\Gamma = \tfrac{1}{2}\langle a|\, \beta\, |a\rangle. \tag{2.20}$$

The 16 Hermitian operators as defined by (2.16)–(2.20) provide an irreducible unitary representation[6] of $U(2, 2)$. The commutation rules of $U(2, 2)$ are given by

$$[J_{\mu\nu}, J_{\alpha\beta}] = i(g_{\alpha\nu}J_{\mu\beta} + g_{\beta\nu}J_{\alpha\mu} - g_{\alpha\mu}J_{\nu\beta} - g_{\mu\beta}J_{\alpha\nu}), \tag{2.21}$$

$$[J_{\mu\nu}, p_\rho^+] = i(g_{\rho\nu}p_\mu^+ - g_{\rho\mu}p_\nu^+), \tag{2.22}$$

$$[p_\mu^+, p_\nu^+] = 0, \tag{2.23}$$

$$[J_{\mu\nu}, p_\rho^-] = i(g_{\rho\nu}p_\mu^- - g_{\rho\mu}p_\nu^-), \tag{2.24}$$

$$[p_\mu^-, p_\nu^-] = 0, \tag{2.25}$$

$$[p_\mu^+, p_\nu^-] = 2i(g_{\mu\nu}\zeta - \tfrac{1}{2}J_{\mu\nu}), \tag{2.26}$$

$$[p_\mu^+, \zeta] = ip_\mu^+, \tag{2.27}$$

$$[p_\mu^-, \zeta] = -ip_\mu^-, \tag{2.28}$$

$$[J_{\mu\nu}, \zeta] = 0. \tag{2.29}$$

These are satisfied by (2.16)–(2.20).

The operator Γ commutes with all the rest of the generators. From the above commutation rules it is seen that the group $U(2, 2)$ contains the Poincaré group as its subgroup. The special representation (2.16)–(2.20) refers to a massless case.

An invariant of $U(2, 2)$ is given by

$$\mathcal{M}_I = \tfrac{1}{2}(p_\mu^+p_-^\mu + p_\mu^-p_+^\mu + J_{\mu\nu}J^{\mu\nu}) - \zeta^2. \tag{2.30}$$

The invariants, for any of the p's either p_+ or p_-, $I_1 = p_\mu p^\mu$, $I_2 = \tfrac{1}{2}J^{\mu\nu}J_{\mu\nu}p_\rho p^\rho - J_{\mu\rho}J^{\mu\rho}p_\nu p^\mu$ of the subgroup vanish, as can easily be shown via (2.16)–(2.20).

Now, from the definition (2.16) of $J_{\mu\nu}$ we obtain

$$J_l = J_{1l} + J_{2l} \quad (l = 1, 2, 3), \tag{2.31}$$
$$J_l = \tfrac{1}{2}\epsilon_{lsk}J_{sk},$$

where

$$J_{1l} = \tfrac{1}{2}\langle A|\, \sigma_l\, |A\rangle, \quad J_{2l} = -\tfrac{1}{2}\langle B|\, \sigma_l\, |B\rangle \tag{2.32}$$

and where σ_l ($l = 1, 2, 3$) are the usual Pauli matrices, and

$$|A\rangle = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad |B\rangle = \begin{bmatrix} a_3 \\ a_4 \end{bmatrix}.$$

They satisfy the commutation rules for the commuting angular momenta,[7]

$$[J_{1i}, J_{2j}] = 0, \quad [J_{1i}, J_{1j}] = i\epsilon_{ijl}J_{1l}, \quad [J_{2i}, J_{2j}] = i\epsilon_{ijl}J_{2l}, \tag{2.33}$$

where

$$J^2 = j(j + 1), \quad J_1^2 = j_1(j_1 + 1), \quad J_2^2 = j_2(j_2 + 1), \tag{2.34}$$

$$j_1 = \tfrac{1}{2}(a_1^\dagger a_1 + a_2^\dagger a_2), \quad j_2 = \tfrac{1}{2}(a_3a_3^\dagger + a_4a_4^\dagger). \tag{2.35}$$

Hence we see that the space part of $J_{\mu\nu}$ is decomposable into a direct product of two 3-dimensional rotation groups. The resultant angular momentum j is associated with angular momenta

$$j = |j_1 - j_2|, |j_1 - j_2| + 1, \cdots, j_1 + j_2, \tag{2.36}$$

where $|j_1 - j_2|$ ($= s$) is the minimum value of j, it is the spin quantum number of the representation assuming the values $0, \frac{1}{2}, 1, \cdots$.

From (2.17) (dropping superscript +) and definitions of γ's (see p. 235 of MQT), the translation operator p_μ can be written as

$$\begin{aligned} p_1 &= J_{1x} - J_{2x} + \tfrac{1}{2}(a_1^\dagger a_4 + a_4^\dagger a_1 + a_3^\dagger a_2 + a_2^\dagger a_3), \\ p_2 &= J_{1y} - J_{2y} - \tfrac{1}{2}i(a_1^\dagger a_4 - a_4^\dagger a_1 + a_3^\dagger a_2 - a_2^\dagger a_3), \\ p_3 &= J_{1z} - J_{2z} + \tfrac{1}{2}(a_3^\dagger a_1 + a_1^\dagger a_3 - a_4^\dagger a_2 - a_2^\dagger a_4), \\ p_4 &= j_1 + j_2 + 1 + \tfrac{1}{2}(a_3^\dagger a_1 + a_1^\dagger a_3 + a_4^\dagger a_2 + a_2^\dagger a_4). \end{aligned} \tag{2.37}$$

Using these definitions we can construct the helicity operator of massless particles in the form

$$\begin{aligned} \Gamma_0 = \mathbf{J}\cdot\hat{\mathbf{p}} &= 1 + \tfrac{1}{2}\langle a|\, \beta\, |a\rangle \\ &= \tfrac{1}{2}(a_1^\dagger a_1 + a_2^\dagger a_2 - a_3a_3^\dagger - a_4a_4^\dagger), \end{aligned} \tag{2.38}$$

where

$$\hat{\mathbf{p}} = \frac{\mathbf{p}}{p_4}, \tag{2.39}$$

and where $\frac{1}{2}\langle a|\, \beta\, |a\rangle$ commutes with the ten generators of the group and is therefore a group invariant. We consider only positive energy representations where the helicity operator Γ_0 together with p_4, J^2, and J_3 form a complete commuting set. A set of simultaneous eigenstates of these commuting operators is designated by $|n, \xi\rangle$. The requirement of nonnegativity for j_1 and j_2 also assures positive sign for the energy and the former is obtained only by defining the vacuum state by the conditions

$$\begin{aligned} a_1\, |0\rangle = 0, \quad a_3^\dagger\, |0\rangle = 0, \\ a_2\, |0\rangle = 0, \quad a_4^\dagger\, |0\rangle = 0. \end{aligned} \tag{2.40}$$

[6] B. Kurşunoğlu, in *Proceedings of Second Coral Gables Conference on Symmetry Principles at High Energy* (W. H. Freeman and Company, San Francisco, 1965), p. 163.

[7] The commutation rules (2.33) are the same as the commutation relations corresponding to the Lie algebra of the 4-dimensional Euclidean group, namely the group O_4.

These definitions in terms of two-component operators as defined in (2.8) can be written as $a_\lambda|0\rangle = 0$, $b_\lambda|0\rangle = 0$, respectively. The vacuum state as defined here is invariant under $SL(2, C)$ subgroup of $U(2, 2)$.

In complete analogy with Fock representation of harmonic oscillator (see Chap. 7 of MQT), we find that occupation number operators are given by

$$N_1 = a_1^\dagger a_1, \quad N_2 = a_2^\dagger a_2, \quad N_3 = a_3 a_3^\dagger, \quad N_4 = a_4 a_4^\dagger, \tag{2.41}$$

which satisfy the eigenvalue equations

$$N_\alpha|n_\alpha\rangle = n_\alpha|n_\alpha\rangle \quad (\alpha = 1, 2, 3, 4), \tag{2.42}$$

where

$$n_\alpha = 0, 1, 2, 3, \cdots.$$

The normalized eigenstates are defined by

$$|n_1\rangle = (n_1!)^{-\frac{1}{2}}(a_1^\dagger)^{n_1}|0\rangle, \quad |n_2\rangle = (n_2!)^{-\frac{1}{2}}(a_2^\dagger)^{n_2}|0\rangle,$$
$$|n_3\rangle = (n_3!)^{-\frac{1}{2}}(a_3)^{n_3}|0\rangle, \quad |n_4\rangle = (n_4!)^{-\frac{1}{2}}(a_4)^{n_4}|0\rangle, \tag{2.43}$$

so that the simultaneous eigenstates $|n, \xi\rangle$ of the complete commuting set p_4, ζ_0, J^2, and J_3 are products of these eigenstates.

From (2.38) it follows that the helicity operator can be expressed in the form

$$\mathbf{J}\cdot\hat{\mathbf{p}} = \tfrac{1}{2}(N_1 + N_2 - N_3 - N_4) = \tfrac{1}{2}N, \tag{2.44}$$

and it acts on the state $|n, \xi\rangle$ according to

$$\mathbf{J}\cdot\hat{\mathbf{p}}|n, \xi\rangle = \tfrac{1}{2}n|n, \xi\rangle,$$

where

$$\tfrac{1}{2}n = \tfrac{1}{2}(n_1 + n_2 - n_3 - n_4) = j_1 - j_2 = \pm s$$

assumes both positive and negative half odd-integral and integral values including zero. Hence we can write

$$\mathbf{J}\cdot\hat{\mathbf{p}}|n, \xi\rangle = \pm s|n, \xi\rangle. \tag{2.45}$$

The eigenvalue equation can further be simplified by noting that it is equivalent to

$$\mathbf{J}\cdot\mathbf{p}|n_\tau, \xi\rangle = \tau_3 s p_4|n_\tau, \xi\rangle, \tag{2.46}$$

where

$$|n_\tau, \xi\rangle = \begin{bmatrix} |n_+, \xi\rangle \\ |n_-, \xi\rangle \end{bmatrix}, \quad \tau_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\mathbf{J}\cdot\mathbf{p}|n_+, \xi\rangle = sp_4|n_+, \xi\rangle,$$
$$|n_+, \xi\rangle = \tfrac{1}{2}(1 + \tau_3)|n_\tau, \xi\rangle,$$
$$\mathbf{J}\cdot\mathbf{p}|n_-, \xi\rangle = -sp_4|n_-, \xi\rangle,$$
$$|n_-, \xi\rangle = \tfrac{1}{2}(1 - \tau_3)|n_\tau, \xi\rangle.$$

Hence, the most general state is a superposition of two orthogonal states,

$$|n_\tau, \xi\rangle = |n_+, \xi\rangle + |n_-, \xi\rangle, \tag{2.47}$$

referring either to two different states of polarization or to two different particles. It depends on the reflection symmetries of various spin states whether one has just a different state of polarization or a different particle. Two states of polarizations, whether they refer to identical particles (e.g., zeron $s = 0$, photons $s = 1$) or two different particle states (e.g., ν, $\tilde{\nu}$ with $s = \frac{1}{2}$) as eigenstates of τ_3, span a 2-dimensional space.

Finally we note from (2.37) and (2.43) that the diagonal element of the operator p_4 with respect to the state $|n, \xi\rangle$ is given by

$$\langle n, \xi|p_4|n, \xi\rangle = \tfrac{1}{2}(n_1 + n_2 + n_3 + n_4) + 1$$
$$= \tfrac{1}{2}n + 1, \tag{2.48}$$

where $n = 0, 1, 2, \cdots$ so that zero-point oscillations are also included in the algebra of $U(2, 2)$.

3. WAVE EQUATIONS

As is well known the group of translations, being an Abelian subgroup of the Poincaré group, has only 1-dimensional irreducible unitary representations. For a translation of states by a real vector $\mathbf{b}$, the unitary operator is $\exp[-ib^\mu p_\mu]$. This group of translations also contains the representations $\exp[-ib^\mu p'_\mu]$ provided p'_μ is obtained from p_μ by a proper Lorentz transformation

$$p'_\mu = L_\mu^\nu p_\nu.$$

An infinitesimal translation of a function of coordinates by an amount ϵb_μ is represented by

$$\exp[i\epsilon b^\mu p_\mu]\psi(x)\exp[-i\epsilon b^\nu p_\mu]$$
$$= (1 + i\epsilon b^\mu p_\mu)\psi(x)(1 - i\epsilon b^\nu p_\nu)$$
$$= \psi(x) + i\epsilon b^\mu[p_\mu, \psi(x)].$$

Hence in the limit of $\epsilon \to 0$ we obtain

$$b^\mu(\partial\psi/\partial x^\mu) = ib^\mu[p_\mu, \psi(x)]$$

or, since this is valid for all b_μ, we have

$$-i(\partial\psi/\partial x^\mu) = [p_\mu, \psi]. \tag{3.1}$$

Now consider the eigenstates $|r, t\rangle$ of the complete commuting set $\mathbf{q}$

$$\mathbf{q}|\mathbf{r}, t\rangle = \mathbf{r}|\mathbf{r}, t\rangle. \tag{3.2}$$

The translation operator acts according to

$$\exp[-i\epsilon b^\mu p_\mu]|x\rangle = |x + \epsilon b\rangle = (1 - i\epsilon b^\mu p_\mu)|x\rangle.$$

Hence, this being valid for every b, we get

$$p_\mu|x\rangle = -i(\partial/\partial x^\mu)|x\rangle. \tag{3.3}$$

A way of obtaining a wave equation may proceed

by representing the state $|\mathbf{r}, t\rangle$ in a Hilbert space spanned by $|n, \xi\rangle$. Thus, writing

$$\langle \mathbf{r}, t \,|\, n, \xi\rangle = |\mathbf{r}, t, s\rangle$$

and regarding it as $2s + 1$ component wavefunction we can derive a wave equation.

From (2.45) we obtain

$$\langle \mathbf{r}, t|\, \mathbf{J}\cdot\mathbf{p}\,|n, \xi\rangle = \pm s\,\langle \mathbf{r}, t|\, p_4\,|n, \xi\rangle$$

or introducing the unit operator

$$\int |\mathbf{r}, t\rangle\langle \mathbf{r}, t|\, d^3\mathbf{r},$$

using (3.3) and performing the obvious steps we get the wave equations

$$H\,|\mathbf{r}, t, s\rangle = \pm i\hbar(\partial/\partial t)\,|\mathbf{r}, t, s\rangle, \tag{3.4}$$

where the Hamiltonian H is given by

$$H = (c/\hbar s)\mathbf{J}\cdot\mathbf{p}. \tag{3.5}$$

For spin $\frac{1}{2}$ particle ($s = \frac{1}{2}$) we have $\mathbf{J} = \frac{1}{2}\hbar\boldsymbol{\sigma}$. The corresponding wave equations are

$$H\,|\nu\rangle = i\hbar\frac{\partial}{\partial t}\,|\nu\rangle, \tag{3.6}$$

$$H\,|\nu\rangle = -i\hbar\frac{\partial}{\partial t}\,|\nu\rangle, \tag{3.7}$$

where

$$H = c\boldsymbol{\sigma}\cdot\mathbf{p}, \quad |\nu\rangle = |r, t, \tfrac{1}{2}\rangle = \text{2-component spinor.}$$

If we call $|\nu\rangle$ the neutrino state then the anti-neutrino state can be defined by

$$|\bar{\nu}\rangle = T\,|\nu\rangle, \tag{3.8}$$

where $T = i\sigma_2\bar{C}$ is the time reversal operator for a 2-component spinor state and $\bar{C}$ is just complex conjugation operation. The operator T acts on σ_i according to (see p. 221 of MQT)

$$T^{-1}\sigma_i T = -\sigma_i. \tag{3.9}$$

Hence the wave equation (3.7) can be written as

$$H\,|\bar{\nu}\rangle = i\hbar\frac{\partial}{\partial t}\,|\bar{\nu}\rangle, \tag{3.10}$$

which is of the same form as (3.6) but refers to anti-neutrino. Reflection symmetry here consists of time reversal operation alone, since space parity is not valid in this case.

As a second example, we take $s = 1$ with J represented by $J_i = \hbar K_i$, $(i = 1, 2, 3)$, where K_i are the generators of 3-dimensional rotations. Thus (3.4) yields the wave equations

$$H\,|\eta\rangle = i\hbar\frac{\partial}{\partial t}\,|\eta\rangle, \tag{3.11}$$

$$H\,|\eta\rangle = -i\hbar\frac{\partial}{\partial t}\,|\eta\rangle, \tag{3.12}$$

where $|\eta\rangle$ is a 3-component complex vector, and $H = c\mathbf{K}\cdot\mathbf{p}$ is the Hamiltonian of a single photon. Now, defining $|p\rangle$ as a 3-dimensional column vector in terms of p_i, $(i = 1, 2, 3)$ and operating on H from the left we obtain

$$\langle p|\, H = 0,$$

which is due to H being a 3×3 anti-symmetric matrix operator in p's. Hence the Eq. (3.11) yields

$$\boldsymbol{\nabla}\cdot\boldsymbol{\eta} = 0, \tag{3.13}$$

which is the transversality condition of the photon wave (see Chap. II of MQT).

The wave equation (3.12) refers to a state of polarization opposite to the one described by (3.11). This can be seen by performing a parity operation on $|\boldsymbol{\eta}\rangle$.[8] Thus if we take

$$|\bar{\eta}\rangle = \bar{C}\,|\eta\rangle \tag{3.14}$$

and noting the transformation

$$\bar{C}H\bar{C} = -H,$$

the wave equation (3.12) becomes

$$H\,|\bar{\eta}\rangle = i\hbar\frac{\partial}{\partial t}\,|\bar{\eta}\rangle, \tag{3.15}$$

which is of the same form as (3.11) but refers to a state of polarization opposite to the one contained in (3.11). The corresponding transversality condition is obtained as $\boldsymbol{\nabla}\cdot\bar{\eta} = 0$.

A third example is the wave equation for zeron. We first observe that

$$\langle n, \xi|\, \hat{\mathbf{S}}\cdot\hat{\mathbf{p}}\,|n, \xi\rangle = \pm 1, \quad \langle n, \xi \,|\, n, \xi\rangle = 1$$

for every n, where $\hat{\mathbf{S}} = s^{-1}\mathbf{J}$. Thus for zero spin we must have $\hat{\mathbf{S}} = \pm\hat{\mathbf{p}}$. Hence

$$p^2\,|0, \xi\rangle = p_4^2\,|0, \xi\rangle,$$

which, using the same methods, yields the scalar wave equation

$$\Delta^2\phi - (c^2)^{-1}(\partial^2\phi/\partial t^2) = 0. \tag{3.16}$$

[8] All of this analysis could also be carried out by using the other alternative of helicity operator which involves $p_{\bar{\mu}}$ in place of p_{μ}^{+}. The results are the same and the two helicity operators commute. At this point it is tempting to assume that the electron and muon neutrinos belong to the massless representations of $U(2, 2)$ instead of that of Poincaré group. However, this interpretation entails doubling of integral spin massless particles, also on the basis of the discussion of $U(2, 2)$ in this paper alone it is not possible to speculate on the existence of two kinds of photons.

4. THE LITTLE GROUP

The group of Lorentz transformations which leave a null vector invariant is isomorphic to the two-dimensional Euclidean group. This is a known result.[9] However, here we derive it in a direct way.

Under an S-transformation, the requirement of invariance of p_μ is contained in the statements

$$p'_\mu = -\langle a|\, S^\dagger \gamma_4 \Lambda_+ \gamma_\mu S\, |a\rangle = -\langle a|\, \gamma_4 \Lambda_+ \gamma_\mu\, |a\rangle = p \tag{4.1}$$

This must hold for every a_α and $a_\alpha^\dagger$, which is possible only if the S-transformations in question commute with p_μ. The operator $\mathbf{J} \cdot \hat{\mathbf{p}} = \frac{1}{2}N$ is the only nontrivial invariant of the group, and therefore, a given S-transformation must be a function of $\frac{1}{2}N$ and also must satisfy (2.14). Such an operator is uniquely defined to be

$$S = e^{\frac{1}{2}iN\theta}, \tag{4.2}$$

where θ can be regarded as an angle of rotation in the xy plane. For an electromagnetic wave, θ is the angle of rotation of the electric vector in the plane perpendicular to its momentum.

The result (4.2) proves the required isomorphism between the group of Lorentz transformations which leave a null vector invariant and the 2-dimensional Euclidean group. Thus the representation of the little group for massless particles is 1-dimensional. The representatives of S are of the form

$$\langle n', \xi|\, S\, |n, \xi\rangle = \delta_{nn'} e^{\pm is\theta}, \tag{4.3}$$

where

$$-s \leq \tfrac{1}{2}n \leq s, \quad -s \leq \tfrac{1}{2}n' \leq s$$

and the dimension of the representation is $2s + 1$.

[9] E. P. Wigner, in Theoretical Physics. A. Salem, Ed. (International Atomic Energy Agency, Vienna, 1963), pp. 59–82.

Reduction of Some Representations of the Generalized Bondi–Metzner Group*

VITTORIO CANTONI†
Department of Mathematics, King's College, London, England

(Received 29 November 1966)

The infinitesimal operators of a class of representations of the GBM group described in a previous paper are given in explicit form. All unitary representations of this type with positive "rest mass" are shown to be completely reducible; each irreducible component is characterized by the eigenvalue m^2 of the rest-mass operator, by the lowest eigenvalue $s(s+1)$ of the "spin" operator and by the sign of an eigenvalue of a suitably defined operator $R_4^{\pm}$.

I. INTRODUCTION

It has been shown[1] that from any linear representation $\mathfrak{T} \equiv \{T(l)\}$ of the inhomogeneous Lorentz group L, acting on a linear space $\mathfrak{K}$, it is possible to construct a linear representation $\tilde{\mathfrak{T}} \equiv \{\tilde{T}(g)\}$ of the GBM group, acting on a linear space $\tilde{\mathfrak{K}}$, $\tilde{\mathfrak{T}}$ being unitary and faithful whenever $\mathfrak{T}$ is such. The restriction of $\tilde{\mathfrak{T}}$ to a specific subgroup of the GBM group isomorphic with L will be denoted by $\tilde{\mathfrak{T}}_L \equiv \{\tilde{T}(l)\}$ and called the induced representation of L. Whenever $\mathfrak{T}$ is irreducible, so that its rest-mass operator P^2 is a multiple of the identity with eigenvalue m^2, the rest-mass operator $\tilde{P}^2$ of the (reducible) induced representation $\tilde{\mathfrak{T}}_L$ is also a multiple of the identity, with the same eigenvalue.

In Sec. II of this paper the representation of the GBM Lie algebra associated with $\tilde{\mathfrak{T}}$ is derived in explicit form, and a basis is introduced in the representation space $\tilde{\mathfrak{K}}$. In Sec. III, $\mathfrak{T}$ is assumed to be unitary and irreducible, with positive rest mass: the irreducible components of the induced representation $\tilde{\mathfrak{T}}_L$ of L are determined. It is shown that the GBM representations $\tilde{\mathfrak{T}}$ of this type are themselves completely reducible, and their irreducible components are characterized.

II. INFINITESIMAL OPERATORS AND REPRESENTATION SPACE

A. Infinitesimal Four-Rotations

The notation being the same as in Ref. 1, consider any one-parameter subgroup $\{\Lambda_t\}$ of the homogeneous Lorentz group ($\Lambda_0 = 1$; $\Lambda_{t'}\Lambda_{t'} = \Lambda_{t'+t'}$).

The element $\tilde{T}_{\Lambda_t} \in \tilde{\mathfrak{T}}$, corresponding to $\Lambda_t \in \{\Lambda_t\}$, is defined by

$$\tilde{T}_{\Lambda_t}(\theta, \varphi) = T(\Lambda_t), \quad [T(\Lambda_t) \in \mathfrak{T}];$$

$\tilde{T}_{\Lambda_t}$ transforms the generic element $\tilde{\Phi} \in \tilde{\mathfrak{K}}$ into the element $\tilde{\Phi}_{\Lambda_t} \in \tilde{\mathfrak{K}}$ defined by

$$\tilde{\Phi}_{\Lambda_t}(\theta, \varphi) = K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)T(\Lambda_t)\tilde{\Phi}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi),$$
$$[\tilde{\Phi}_{\Lambda_t}(\theta, \varphi), \tilde{\Phi}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi) \in \mathfrak{K}].$$

Throughout this paper the mappings $\tilde{\Phi} \in \tilde{\mathfrak{K}}$ are assumed to be twice differentiable on the two-sphere. $K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)$, $\Lambda_t^{-1}\theta$, and $\Lambda_t^{-1}\varphi$ are differentiable with respect to t (see the Appendix). Therefore one can write

$$\begin{aligned}
&\tilde{\Phi}_{\Lambda_t}(\theta, \varphi) \\
&= \left[1 + \left(\frac{\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)}{\partial t}\right)_{t=0} t + O(t^2)\right] \\
&\quad \times [1 + Mt + O(t^2)]\left[\tilde{\Phi}(\theta, \varphi) + \left(\frac{\partial \tilde{\Phi}(\theta, \varphi)}{\partial \theta}\frac{\partial \Lambda_t^{-1}\theta}{\partial t}\right.\right. \\
&\quad \left.\left. + \frac{\partial \tilde{\Phi}(\theta, \varphi)}{\partial \varphi}\frac{\partial \Lambda_t^{-1}\varphi}{\partial t}\right)_{t=0} t + O(t^2)\right] \\
&= \left\{1 + \left[M + \left(\frac{\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)}{\partial t}\right)_{t=0}\right.\right. \\
&\quad + \left(\frac{\partial \Lambda_t^{-1}\theta}{\partial t}\right)_{t=0}\frac{\partial}{\partial \theta} + \left.\left.\left(\frac{\partial \Lambda_t^{-1}\varphi}{\partial t}\right)_{t=0}\frac{\partial}{\partial \varphi}\right] t + O(t^2)\right\} \\
&\quad \times \tilde{\Phi}(\theta, \varphi),
\end{aligned}$$

where M is the infinitesimal operator of $\mathfrak{T}$ corresponding to the one-parameter subgroup $\{\Lambda_t\}$.

Hence one gets the expression for the infinitesimal operator $\tilde{M}$ corresponding to $\{\Lambda_t\}$ in the representation $\tilde{\mathfrak{T}}$:

$$\begin{aligned}
\tilde{M}(\theta, \varphi) = M &+ \left(\frac{\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)}{\partial t}\right)_{t=0} \\
&+ \left(\frac{\partial \Lambda_t^{-1}\theta}{\partial t}\right)_{t=0}\frac{\partial}{\partial \theta} + \left(\frac{\partial \Lambda_t^{-1}\varphi}{\partial t}\right)_{t=0}\frac{\partial}{\partial \varphi} \quad (1)
\end{aligned}$$

* Sponsored in part by the Aerospace Research Laboratories under Contract AF 61 (052)-877, through the European Office of Aerospace Research (OAR), U.S. Air Force.

† NATO research scholar, on leave from Gruppo di ricerca No. 36, C.N.R., Roma.

[1] V. Cantoni, J. Math. Phys. 7, 1361 (1966).

which transforms any element $\tilde{\Phi} \in \tilde{\mathcal{H}}$ into the element $\tilde{\Psi}_M \in \tilde{\mathcal{H}}$ such that $\tilde{\Phi}_M(\theta, \varphi) = \tilde{M}(\theta, \varphi)\tilde{\Phi}(\theta, \varphi)$. The expressions for

$$\left(\frac{\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)}{\partial t}\right)_{t=0}, \quad \left(\frac{\partial \Lambda_t^{-1}\theta}{\partial t}\right)_{t=0},$$

and

$$\left(\frac{\partial \Lambda_t^{-1}\varphi}{\partial t}\right)_{t=0}$$

are derived in the Appendix for the six one-parameter subgroups of four-rotations in the coordinate planes (x^i, x^j), $(i, j = 0, 1, 2, 3; i \neq j)$. One obtains for the corresponding infinitesimal operators $\tilde{M}_{ij}$ in the representation $\tilde{\mathcal{C}}$:

$$\begin{aligned}
\tilde{M}_{12}(\theta, \varphi) &= M_{12} - \frac{\partial}{\partial \varphi} \equiv M_{12} + \mathfrak{M}_{12}, \\
\tilde{M}_{23}(\theta, \varphi) &= M_{23} - \sin\varphi \frac{\partial}{\partial \theta} \\
&\quad - \operatorname{cotg}\theta \cos\varphi \frac{\partial}{\partial \varphi} \equiv M_{23} + \mathfrak{M}_{23}, \\
\tilde{M}_{31}(\theta, \varphi) &= M_{31} + \cos\varphi \frac{\partial}{\partial \theta} \\
&\quad - \operatorname{cotg}\theta \sin\varphi \frac{\partial}{\partial \varphi} \equiv M_{31} + \mathfrak{M}_{31}, \qquad (2) \\
\tilde{M}_{01}(\theta, \varphi) &= M_{01} - \sin\theta\cos\varphi \\
&\quad + \cos\theta\cos\varphi \frac{\partial}{\partial \theta} - \frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial \varphi}, \\
\tilde{M}_{02}(\theta, \varphi) &= M_{02} - \sin\theta\sin\varphi \\
&\quad + \cos\theta\sin\varphi \frac{\partial}{\partial \theta} + \frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial \varphi}, \\
\tilde{M}_{03}(\theta, \varphi) &= M_{03} - \cos\theta - \sin\theta \frac{\partial}{\partial \theta},
\end{aligned}$$

where the $\mathfrak{M}_{ij}$'s $(i, j = 1, 2, 3; i \neq j)$ are the operators corresponding to the infinitesimal rotations in the representations of the three-dimensional rotation group acting on spherical functions.

As usual, it is convenient to replace the operators $\tilde{M}_{ij}$ by their linear combinations

$$\begin{aligned}
&\tilde{H}_+ = i\tilde{M}_{23} - \tilde{M}_{31}, \quad \tilde{H}_- = i\tilde{M}_{23} + \tilde{M}_{31}, \quad \tilde{H}_3 = i\tilde{M}_{12}, \\
&\tilde{F}_+ = i\tilde{M}_{01} - \tilde{M}_{02}, \quad \tilde{F}_- = i\tilde{M}_{01} + \tilde{M}_{02}, \quad \tilde{F}_3 = i\tilde{M}_{03}.
\end{aligned} \qquad (3)$$

The "angular momentum" operator

$$\tilde{H}^2 = \tfrac{1}{2}(\tilde{H}_+\tilde{H}_- + \tilde{H}_-\tilde{H}_+) + \tilde{H}_3^2$$

can be written in the form

$$\tilde{H}^2 = H^2 + \mathfrak{H}^2 + 2H_3\mathfrak{H}_3 + H_+\mathfrak{H}_- + H_-\mathfrak{H}_+, \qquad (4)$$

where

$$H^2 = \tfrac{1}{2}(H_+H_- + H_-H_+) + H_3^2$$

and

$$\mathfrak{H}^2 = \tfrac{1}{2}(\mathfrak{H}_+\mathfrak{H}_- + \mathfrak{H}_-\mathfrak{H}_+) + \mathfrak{H}_3^2.$$

B. Infinitesimal Supertranslations

Let $\mathbf{a} = (a^0, a^1, a^2, a^3)$ be the vector of a 4-translation of L, represented by $T(\mathbf{a})$ in the representation $\mathcal{C}$. It is well known that $T(\mathbf{a})$ has the form

$$T(\mathbf{a}) = \exp(-ia^kP_k),$$

where P_k is the kth component of the linear momentum operator in $\mathcal{C}$.

Consider now a supertranslation $(1, \alpha)$, characterized by the function $\alpha(\theta, \varphi)$, and denote by $\mathbf{a}_{\theta\varphi}$ the vector of its asymptotically tangent translation in the ray direction (θ, φ). In the representation $\tilde{\mathcal{C}}$, the operator $\tilde{T}_\alpha$ corresponding to $(1, \alpha)$ is defined by

$$\tilde{T}_\alpha(\theta, \varphi) = \exp(-ia^k_{\theta\varphi}P_k)$$

and transforms any element $\tilde{\Phi} \in \tilde{\mathcal{H}}$ into the element $\tilde{\Phi}_\alpha \in \tilde{\mathcal{H}}$ defined by

$$\tilde{\Phi}_\alpha(\theta, \varphi) = \tilde{T}_\alpha(\theta, \varphi)\tilde{\Phi}(\theta, \varphi)$$

(see Ref. 1, Sec. V).

The supertranslation $(1, \alpha)$ belongs to the one-parameter subgroup $\{t\alpha\}$ of supertranslations whose generic element is characterized by the function $t\alpha(\theta, \varphi)$, t being the parameter; the corresponding infinitesimal operator $\tilde{S}_\alpha$ in $\tilde{\mathcal{C}}$ is defined by

$$\tilde{S}_\alpha(\theta, \varphi) = -ia^k_{\theta\varphi}P_k,$$

and transforms any element $\tilde{\Phi} \in \tilde{\mathcal{H}}$ into the element $\tilde{\Phi}_s \in \tilde{\mathcal{H}}$ defined by

$$\tilde{\Phi}_s(\theta, \varphi) = -ia^k_{\theta\varphi}P_k\tilde{\Phi}(\theta, \varphi).$$

Following Sachs,[2] one can expand $\alpha(\theta, \varphi)$ in (normalized) spherical harmonics $Y_{lm}(\theta, \varphi)$,

$$\alpha(\theta, \varphi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l}(z_{lm}Y_{lm} + (-1)^m\bar{z}_{lm}Y_{l-m})$$

(the bar denotes complex conjugation), and consider the complete set of one-parameter subgroups

$$\begin{aligned}
t\alpha_{lm} &= t(Y_{lm} + (-1)^m Y_{l-m}), \\
t\beta_{lm} &= it(Y_{lm} - (-1)^m Y_{l-m}).
\end{aligned}$$

The operators $\tilde{A}_{lm}$ and $\tilde{B}_{lm}$ of $\tilde{\mathcal{C}}$ which represent the

[2] R. Sachs, Phys. Rev. **128**, 2851 (1962).

corresponding infinitesimal elements may conveniently be replaced by their linear combinations:

$$\tilde{S}_{lm} = \tfrac{1}{2}(\tilde{A}_{lm} - i\tilde{B}_{lm}),$$
$$\tilde{S}_{l-m} = \tfrac{1}{2}(-1)^m(\tilde{A}_{lm} + i\tilde{B}_{lm}),$$
$$(l = 0, 1, \cdots; m = 0, 1, \cdots, l).$$

Making use of the expressions for $a^i_{\theta\varphi}$ $(i = 0, 1, 2, 3)$ given in Ref. 1, one obtains

$$\begin{aligned}\tilde{S}_{lm} = {} & -i(1 - \tfrac{1}{2})l(l+1)Y_{lm}P_0 \\ & + i\Big[\tfrac{1}{2}l(l+1)\sin\theta\cos\varphi Y_{lm} \\ & - \frac{\sin\varphi}{\sin\theta}\frac{\partial Y_{lm}}{\partial\varphi} + \cos\theta\cos\varphi\frac{\partial Y_{lm}}{\partial\theta}\Big]P_1 \\ & + i\Big[\tfrac{1}{2}l(l+1)\sin\theta\sin\varphi Y_{lm} \\ & + \frac{\cos\varphi}{\sin\theta}\frac{\partial Y_{lm}}{\partial\varphi} + \cos\theta\sin\varphi\frac{\partial Y_{lm}}{\partial\theta}\Big]P_2 \\ & + i\Big[\tfrac{1}{2}l(l+1)\cos\theta Y_{lm} - \sin\theta\frac{\partial Y_{lm}}{\partial\theta}\Big]P_3, \\ & (l = 0, 1, \cdots; m = -l, -l+1, \cdots, l). \qquad (5)\end{aligned}$$

The operators $\tilde{P}_i$ $(i = 1, 2, 3)$ defined by $\tilde{P}_i(\theta, \varphi) = P_i$ and the operator $\tilde{P}^2 = -\tilde{P}_0^2 + \tilde{P}_1^2 + \tilde{P}_2^2 + \tilde{P}_3^2$ may be expressed in terms of the $\tilde{S}_{lm}$'s:

$$\begin{aligned}\tilde{P}_0 &= i(4\pi)^{\frac{1}{2}}\tilde{S}_{00} \\ \tilde{P}_1 &= -i(\tfrac{2}{3}\pi)^{\frac{1}{2}}(\tilde{S}_{11} - \tilde{S}_{1-1}) \\ \tilde{P}_2 &= -(\tfrac{2}{3}\pi)^{\frac{1}{2}}(\tilde{S}_{11} + \tilde{S}_{1-1}) \qquad (6) \\ \tilde{P}_3 &= -i(\tfrac{4}{3}\pi)^{\frac{1}{2}}\tilde{S}_{10} \\ \tilde{P}^2 &= \tfrac{4}{3}\pi(2\tilde{S}_{11}\tilde{S}_{1-1} - \tilde{S}_{10}^2) + 4\tilde{S}_{00}^2.\end{aligned}$$

According to Sec. VII, Ref. 1, $\tilde{P}_i$ and $\tilde{P}^2$ are, respectively, the linear momentum and the rest-mass operators of the induced representation $\tilde{\mathcal{C}}_L$ of L.

C. Introduction of a Basis in $\tilde{\mathcal{K}}$

It is well known[3,4] that the representation space $\mathcal{K}$ of the representation $\mathcal{C}$ of L admits a basis whose elements are simultaneous eigenvectors of the momentum operators P_0, P_1, P_2, P_3, the spin operator and the operator $W_3 = \frac{1}{2}\epsilon_{3ijk}M^{ij}P^k$ (ϵ denoting the permutation tensor and indices being raised with the Lorentz metric). Whenever $\mathcal{C}$ is irreducible (as it is here assumed to be), P_iP^i and the spin operator are multiples of the identity, and the elements of the basis may only be labeled by the corresponding eigenvalues of P_1, P_2, P_3, and W_3. The generic element of such a basis is denoted by $\boldsymbol{\varphi}_{p\zeta}$, where p stands for the set of eigenvalues (p_1, p_2, p_3) of P_1, P_2, and P_3, and ζ is a discrete index corresponding to the eigenvalue of W_3. The generic element of $\mathcal{K}$ has the form

$$\Phi = \sum_\zeta \int f(p, \zeta)\varphi_{p\zeta}\, dp.$$

According to the definition of the representation space $\tilde{\mathcal{K}}$ of $\tilde{\mathcal{C}}$, the generic element $\tilde{\Phi} \in \tilde{\mathcal{K}}$ maps each point (θ, φ) of the 2-sphere on an element $\tilde{\Phi}(\theta, \varphi) \in \mathcal{K}$, so that one can write

$$\tilde{\Phi}(\theta, \varphi) = \sum_\zeta \int f(p, \zeta; \theta, \varphi)\boldsymbol{\varphi}_{p\zeta}\, dp,$$

and the functions $f(p, \zeta; \theta, \varphi)$ are continuous and twice differentiable on the 2-sphere for all values of p and ζ.

If f is expanded in spherical harmonics, $\tilde{\Phi}(\theta, \varphi)$ takes the form

$$\tilde{\Phi}(\theta, \varphi) = \sum_\zeta \sum_{l=1}^{\infty} \sum_{m=-l}^{l} \int f_{lm}(p, \zeta)Y_{lm}(\theta, \varphi)\boldsymbol{\varphi}_{p\zeta}\, dp$$

from which it is clear that the elements $\tilde{\varphi}_{p\zeta lm} \in \tilde{\mathcal{K}}$ defined by

$$\tilde{\varphi}_{p\zeta lm}(\theta, \varphi) = Y_{lm}(\theta, \varphi)\boldsymbol{\varphi}_{p\zeta} \qquad (7)$$

constitute a basis for $\tilde{\mathcal{K}}$.

Note that $\tilde{\mathcal{K}}$ can be regarded as a linear space over the ring of twice-differentiable complex functions on the 2-sphere, the "product" of a function $f(\theta, \varphi)$ and an element $\tilde{\Phi} \in \tilde{\mathcal{K}}$ being the element $\tilde{\Psi} \in \tilde{\mathcal{K}}$ defined by

$$\tilde{\Psi}(\theta, \varphi) = f(\theta, \varphi)\tilde{\Phi}(\theta, \varphi).$$

In particular the element $\tilde{\varphi}_{p\zeta lm}$ of the basis can be denoted by $Y_{lm}\tilde{\varphi}_{p\zeta}$ and regarded as the product of $Y_{lm}(\theta, \varphi)$ and $\tilde{\varphi}_{p\zeta}$, where $\tilde{\varphi}_{p\zeta} \in \tilde{\mathcal{K}}$ is defined by $\tilde{\varphi}_{p\zeta}(\theta, \varphi) = \varphi_{p\zeta}$.

III. REPRESENTATIONS WITH POSITIVE REST MASS

A. Analysis of the Induced Representation $\tilde{\mathcal{C}}_L$

Throughout this section it is assumed that $\mathcal{C}$ is unitary and irreducible, and that

$$P^2\Phi = m^2\Phi \quad (\Phi \in \mathcal{K}) \qquad (8)$$

with $m > 0$.

The induced representation $\tilde{\mathcal{C}}_L$ of L is unitary, and therefore fully reducible.[3] Since (8) implies $\tilde{P}^2\tilde{\Phi} = m^2\tilde{\Phi}$, $(\tilde{\Phi} \in \tilde{\mathcal{K}})$, the irreducible components of $\tilde{\mathcal{C}}_L$ all correspond to the same eigenvalue m^2 of the rest-mass operator $\tilde{P}^2$, but not, in general, to the same eigenvalue of the spin operator.

[3] E. Wigner, Ann. Math. **40**, 149 (1939).
[4] S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row & Peterson, Evanston, Illinois, 1961).

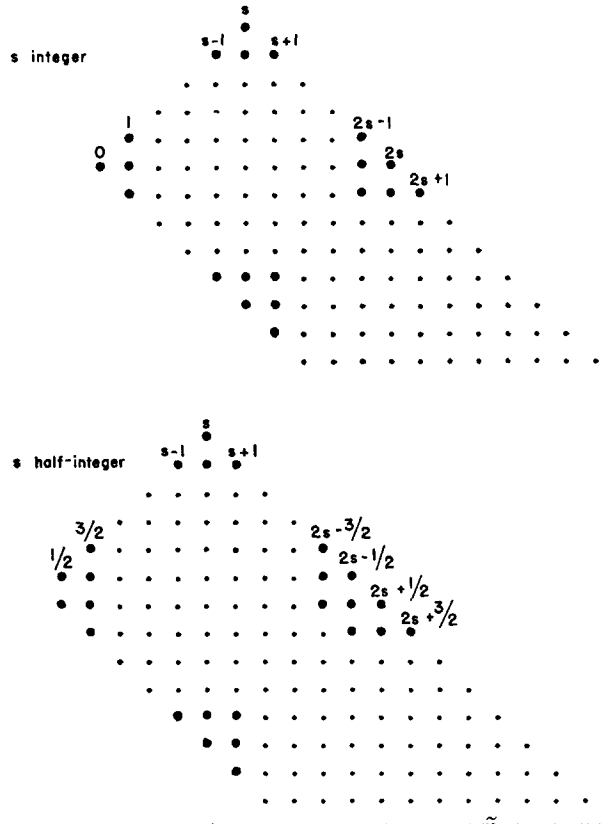

FIG. 1. Here each dot represents a subspace of $\tilde{\mathcal{H}}_0$ irreducible under O^3 (or the corresponding irreducible representation). The $(l+1)$th row represents the subspace of $\tilde{\mathcal{H}}_0$ spanned by the elements $\{Y_{lm}\tilde{\varphi}_{0\zeta}\}$. Elements of the same column correspond to the same weight, which labels the column.

The irreducible components of $\tilde{\mathcal{C}}_L$ can be found[4] by considering the subspace $\tilde{\mathcal{H}}_0$ of $\tilde{\mathcal{H}}$ spanned by the elements $\{\tilde{\varphi}_{0\zeta lm}\} \equiv \{Y_{lm}\tilde{\varphi}_{0\zeta}\}$ of the basis, and by determining the subspaces of $\tilde{\mathcal{H}}_0$ which are irreducible under the three-dimensional rotation group O^3 [the operators which correspond to the infinitesimal three-rotations being the $\tilde{M}_{ij}$'s $(i,j=1,2,3;\ i\neq j)$ given by (2), up to a common factor m].

It is clear from (2) that for every fixed value of l the subspace of $\tilde{\mathcal{H}}_0$ spanned by the elements $\{Y_{lm}\tilde{\varphi}_{0\zeta}\}$ of the basis may be identified with the representation space of the product of two irreducible representations of the rotation group, namely the representation with weight l acting on spherical harmonics of order l, and the representation with weight s acting on the subspace $\mathcal{H}_0$ of $\mathcal{H}$ spanned by the elements $\{\varphi_{0\zeta}\}$, s denoting the spin of the representation. Such a product is reducible,[5] and can be decomposed into irreducible representations of O^3 with weights $l+s,\ l+s-1,\cdots,\ |l-s|$, each of these representations appearing once in the decomposition.

The number l takes all possible values $0, 1, 2, \cdots$ in $\tilde{\mathcal{H}}_0$; consequently if s is an integer (half-integer), $\tilde{\mathcal{H}}_0$ is the direct sum of subspaces invariant under O^3 and corresponding to irreducible representations of O^3 with integral (half-integral) weights. It is clear from the diagrams (Fig. 1) that the number of such subspaces corresponding to any given integral (half-integral) weight σ is $2\sigma+1$ for $\sigma \leq s$ and $2s+1$ for $\sigma \geq s$.

Hence one is led to the following conclusion: *Whenever the representation $\mathcal{C}$ of L is irreducible and unitary with positive rest mass and integral (half-integral) spin s, the induced representation $\tilde{\mathcal{C}}_L$ can be decomposed into irreducible representations of L with mass m; each integral (half-integral) value σ of the spin appears N_σ times in the decomposition with $N_\sigma = 2\sigma+1$ for $\sigma \leq s$ and $N_\sigma = 2s+1$ for $\sigma \geq s$.*

B. Properties of the Representations $\tilde{\mathcal{C}}$

(1) *The whole representation space $\tilde{\mathcal{H}}$ can be generated from the element $\tilde{\varphi}_{0s00}$*; i.e., every element $\tilde{\Phi} \in \tilde{\mathcal{H}}$ is of the type $\tilde{\Phi} = \tilde{T}_\Phi \tilde{\varphi}_{0s00}$, where $\tilde{T}_\Phi$ is a linear combination of products of operators belonging to the representation of the GBM Lie algebra.

Proof. Since every element $\tilde{\varphi}_{p\zeta lm}$ of the basis which does not belong to $\tilde{\mathcal{H}}_0$ can be obtained from some element $\psi \in \tilde{\mathcal{H}}_0$ by acting upon it with the operator $\tilde{T}_\Lambda$ of a suitable homogeneous Lorentz transformation Λ, it is sufficient to show that $\tilde{\mathcal{H}}_0$ can be generated from the element $\tilde{\varphi}_{0s00}$.

According to (7) one has for the $2s+1$ elements $\tilde{\varphi}_{0\zeta 00}$:

$$\tilde{\varphi}_{0\zeta 00}(\theta,\varphi) = Y_{00}(\theta,\varphi)\varphi_{0\zeta} = (4\pi)^{-\frac{1}{2}}\varphi_{0\zeta} \qquad (\zeta = -s, -s+1, \cdots, s),$$

and it is clear from (2) that by repeated application of multiples of $\tilde{H}_+$ and $\tilde{H}_-$ any of these elements can be transformed into any other of them. On the other hand, in the representation $\mathcal{C}$ of L one has

$$\begin{aligned} P_i\boldsymbol{\varphi}_{0\zeta} &= 0 \quad (i=1,2,3), \\ P_0\boldsymbol{\varphi}_{0\zeta} &= m\varphi_{0\zeta} \end{aligned} \tag{9}$$

so that in $\tilde{\mathcal{C}}$ one gets from (5) and (7) for $l \neq 1$:

$$\begin{aligned} \tilde{Q}_{lm}\tilde{\varphi}_{0\zeta 00} &= (4\pi)^{\frac{1}{2}}\big(1 - \tfrac{1}{2}l(l+1)\big)^{-1}\tilde{S}_{lm}\tilde{\varphi}_{0\zeta 00} \\ &= Y_{lm}\tilde{\varphi}_{0\zeta 00} \equiv \tilde{\varphi}_{0\zeta lm}, \end{aligned} \tag{10}$$

where

$$\tilde{Q}_{lm} \equiv (4\pi)^{\frac{1}{2}}\big(1 - \tfrac{1}{2}l(l+1)\big)^{-1}\tilde{S}_{lm} \quad (l \neq 1).$$

This shows that every element $\tilde{\varphi}_{0\zeta lm}$ of the basis of $\tilde{\mathcal{H}}_0$ with $l \neq 1$ can be generated from any of the elements $\tilde{\varphi}_{0\zeta 00}$.

To complete the proof, it must be shown that the elements $\tilde{\varphi}_{0\zeta 1m}$ can also be generated from an element $\tilde{\varphi}_{0\zeta 00}$. By repeated use of the relation[5]

$$(l+1)[(2l+1)(2l+3)]^{-\frac{1}{2}}Y_{l+1\,0} + l[(2l+1)(2l-1)]^{-\frac{1}{2}}Y_{l-1\,0} = (4\pi/3)^{\frac{1}{2}}Y_{10}Y_{10}$$

satisfied by normalized Legendre polynomials, one can derive the following expression for Y_{10}:

$$Y_{10} = \tfrac{1}{3}(140\pi/3)^{\frac{1}{2}}Y_{20}Y_{30} - 50/9.33^{-\frac{1}{2}}Y_{50} - 28/9.21^{-\frac{1}{2}}Y_{30}. \quad (11)$$

From (11), taking (5), (7), (10), and (2) into account, it is seen that the operators

$$\begin{aligned}\tilde{Q}_{10} &\equiv \tfrac{1}{3}(140\pi/3)^{\frac{1}{2}}\tilde{Q}_{20}\tilde{Q}_{30} - 50/9.33^{-\frac{1}{2}}\tilde{Q}_{50} - 28/9.21^{-\frac{1}{2}}\tilde{Q}_{30},\\ \tilde{Q}_{11} &= 2^{-\frac{1}{2}}(\tilde{H}_+\tilde{Q}_{10} - \tilde{Q}_{10}\tilde{H}_+),\\ \tilde{Q}_{1-1} &= 2^{-\frac{1}{2}}(\tilde{H}_-\tilde{Q}_{10} - \tilde{Q}_{10}\tilde{H}_-),\end{aligned} \quad (12)$$

transform the element $\tilde{\varphi}_{0\zeta 00}$ into the elements

$$\tilde{Q}_{1m}\tilde{\varphi}_{0\zeta 00} = \tilde{\varphi}_{0\zeta 1m} \quad (m = -1, 0, 1),$$

which is the desired result.

Note that each operator $\tilde{Q}_{lm}$ transforms any element of $\tilde{\mathcal{H}}_0$ into its product (in the sense of Sec. IIC) with the function Y_{lm} with the same indices.

(2) The operators $\tilde{H}_+$, $\tilde{H}_-$, $\tilde{H}_3$, and $\tilde{Q}_{lm}$ satisfy the commutation relations

$$\begin{aligned}[\tilde{H}_+, \tilde{Q}_{lm}] &= [(l+m+1)(l-m)]^{\frac{1}{2}}\tilde{Q}_{lm+1},\\ [\tilde{H}_-, \tilde{Q}_{lm}] &= [(l+m)(l-m+1)]^{\frac{1}{2}}\tilde{Q}_{lm-1},\\ [\tilde{H}_3, \tilde{Q}_{lm}] &= m\tilde{Q}_{lm},\end{aligned} \quad (13)$$

$$\begin{aligned}[\tilde{H}^2, \tilde{Q}_{lm}] &= l(l+1)\tilde{Q}_{lm} + 2m\tilde{Q}_{lm}\tilde{H}_3\\ &\quad + [(l+m)(l-m+1)]^{\frac{1}{2}}\tilde{Q}_{lm-1}\tilde{H}_+\\ &\quad + [(l+m+1)(l-m)]^{\frac{1}{2}}\tilde{Q}_{lm+1}\tilde{H}_-,\end{aligned} \quad (14)$$

$$\begin{aligned}[\tilde{H}_-, \tilde{Q}_{11}^{\lambda}] &= 2^{\frac{1}{2}}\tilde{Q}_{10}\tilde{Q}_{11}^{\lambda-1}\\ [\tilde{H}_-^2, \tilde{Q}_{11}^{\lambda}] &= 2^{\frac{3}{2}}\lambda\tilde{Q}_{10}\tilde{Q}_{11}^{\lambda-1}\tilde{H}_-\\ &\quad + 2\lambda(\lambda-1)\tilde{Q}_{10}^2\tilde{Q}_{11}^{\lambda-2} + 2\lambda\tilde{Q}_{1-1}\tilde{Q}_{11}^{\lambda-1},\end{aligned} \quad (15)$$

which can be obtained from (2), (3), (5), (10), and (12) by straightforward calculation (see also Ref. 2).

(3) Whenever an element $\tilde{\Phi} \in \tilde{\mathcal{H}}_0$ is a simultaneous eigenvector of $\tilde{H}^2$ and $\tilde{H}_3$ with eigenvalues $\lambda(\lambda+1)$ and μ, respectively, it is denoted by $\tilde{\Phi}_{\lambda\mu}$ and called an element of type (λ, μ).

[5] I. M. Gel'fand and Z. Ya. Šapiro, Am. Math. Soc. Transl., Ser. 2, Vol. 2 (1956).

Consider the operators

$$\begin{aligned}\tilde{R}^+ &\equiv \tilde{Q}_{11},\\ \tilde{R}_\lambda^= &\equiv 2^{\frac{1}{2}}\lambda\tilde{Q}_{10} - \tilde{Q}_{11}\tilde{H}_-,\\ \tilde{R}_\lambda^- &\equiv 2\lambda(2\lambda-1)\tilde{Q}_{1-1} - 2^{\frac{1}{2}}(2\lambda-1)\tilde{Q}_{10}\tilde{H}_- + \tilde{Q}_{11}\tilde{H}_-^2 \quad (\lambda = 0, \tfrac{1}{2}, 1, \cdots).\end{aligned} \quad (16)$$

Making use of (13) and (14) one can verify the relations

$$\begin{aligned}\tilde{H}^2\tilde{R}^+\tilde{\Phi}_{\lambda\lambda} &= (\lambda+1)(\lambda+2)\tilde{R}^+\tilde{\Phi}_{\lambda\lambda},\\ \tilde{H}_3\tilde{R}^+\tilde{\Phi}_{\lambda\lambda} &= (\lambda+1)\tilde{R}^+\tilde{\Phi}_{\lambda\lambda},\\ \tilde{H}^2\tilde{R}_\lambda^=\tilde{\Phi}_{\lambda\lambda} &= \lambda(\lambda+1)\tilde{R}_\lambda^=\tilde{\Phi}_{\lambda\lambda},\\ \tilde{H}_3\tilde{R}_\lambda^=\tilde{\Phi}_{\lambda\lambda} &= \lambda\tilde{R}_\lambda^=\tilde{\Phi}_{\lambda\lambda},\\ \tilde{H}^2\tilde{R}_\lambda^-\tilde{\Phi}_{\lambda\lambda} &= (\lambda-1)\lambda\tilde{R}_\lambda^-\tilde{\Phi}_{\lambda\lambda},\\ \tilde{H}_3\tilde{R}_\lambda^-\tilde{\Phi}_{\lambda\lambda} &= (\lambda-1)\tilde{R}_\lambda^-\tilde{\Phi}_{\lambda\lambda},\end{aligned}$$

which show that $\tilde{R}^+$, $\tilde{R}_\lambda^=$, and $\tilde{R}_\lambda^-$ transform any vector of type (λ, λ) into vectors of type $(\lambda+1, \lambda+1)$, (λ, λ), and $(\lambda-1, \lambda-1)$, respectively.

The operators $\tilde{R}^+$, $\tilde{R}^=$, and $\tilde{R}^-$ satisfy the relations

$$\tilde{R}_\lambda^{=2} = \frac{3}{2\pi}\lambda^2 + \tilde{R}^+\tilde{R}_\lambda^-,$$

$$\tilde{R}_{\lambda+1}^-\tilde{R}^+ = -\frac{3}{2\pi}(2\lambda+1) + \tilde{R}^+\tilde{R}_\lambda^-.$$

C. Reduction of the Representations $\tilde{\mathfrak{C}}$

Whenever it is necessary to specify the particular value s of the spin of the representation $\mathfrak{C}$ from which $\tilde{\mathfrak{C}}$ is constructed, the latter is denoted by $\tilde{\mathfrak{C}}_s$.

1. *Irreducibility of* $\tilde{\mathfrak{C}}_0$

In the case $s = 0$ it is already known from Sec. IIIA that the induced representation $\tilde{\mathfrak{C}}_L$ contains, exactly once, each irreducible representation of L with mass m and integral spin. Here the index ζ takes the unique value 0, and it is clear from (4) that the elements $\tilde{\varphi}_{00lm}$ (for any fixed value of l, and $m = -l, -l+1, \cdots, l$) span the intersection of $\tilde{\mathcal{H}}_0$ with the subspace of $\tilde{\mathcal{H}}$ which is invariant under $\tilde{\mathfrak{C}}_L$ and corresponds to the value l of the spin.

Suppose that $\tilde{\mathcal{H}}$ has a proper subspace $\tilde{h}$ which is invariant under the GBM group: *a fortiori* $\tilde{h}$ is invariant under L, and must contain at least one of the elements $\tilde{\varphi}_{00\lambda\lambda}$ for some λ. Hence to prove that $\tilde{\mathcal{H}}$ has no proper invariant subspace under the GBM group, it is sufficient to show that $\tilde{\mathcal{H}}$ can be generated from any of the elements $\tilde{\varphi}_{00\lambda\lambda}$, or on account of the result of Sec. IIIB1, that any element $\tilde{\varphi}_{00\lambda\lambda}$ can be transformed into the element $\tilde{\varphi}_{0000}$.

The element $\tilde{R}^{+\lambda}\tilde{\varphi}_{00}$ is obviously of type (λ, λ) and equal to $\tilde{\varphi}_{00\lambda\lambda}$ up to a scalar factor. By acting upon it with the operator $\tilde{R}_\lambda^-$, taking (15), (16), and the condition $\tilde{H}_-\tilde{\varphi}_{00} = 0$ into account, one gets

$$\tilde{R}_\lambda^-\tilde{R}^{+\lambda}\tilde{\varphi}_{00} = -\frac{3}{2\pi}\lambda^2\tilde{R}^{+\lambda-1}\tilde{\varphi}_{00};$$

the element obtained is equal to $\tilde{\varphi}_{00\lambda-1\lambda-1}$ up to a scalar factor. Thus by successive application of the operators $\tilde{R}_\lambda^-$, $\tilde{R}_{\lambda-1}^-, \cdots, \tilde{R}_1^-$, the element $\tilde{\varphi}_{00\lambda\lambda}$ is transformed into a multiple of $\tilde{\varphi}_{0000}$, and the irreducibility of $\tilde{\mathfrak{S}}_0$ is proved.

2. *Reduction of* $\tilde{\mathfrak{S}}_{\frac{1}{2}}$

In the case $s = \frac{1}{2}$, a decomposition of $\tilde{\mathfrak{S}}_L$ into irreducible representations of L gives, exactly twice, each representation with mass m and half-integral spin, and the index ζ takes the values $-\frac{1}{2}$ and $\frac{1}{2}$.

The operator $\tilde{R}_{\frac{1}{2}}^=$ has two independent eigenvectors of type $(\frac{1}{2}, \frac{1}{2})$, namely the elements $\tilde{\psi}_+ = \tilde{\varphi}_{0\frac{1}{2}} + (8\pi/3)^{\frac{1}{2}}\tilde{R}_{\frac{1}{2}}^=\tilde{\varphi}_{0\frac{1}{2}}$ and $\tilde{\psi}_- = \tilde{\varphi}_{0\frac{1}{2}} - (8\pi/3)^{\frac{1}{2}}\tilde{R}_{\frac{1}{2}}^=\tilde{\varphi}_{0\frac{1}{2}}$; the corresponding eigenvalues are $(3/8\pi)^{\frac{1}{2}}$ and $-(3/8\pi)^{\frac{1}{2}}$, respectively. It is shown that the two subspaces of $\tilde{\mathfrak{K}}$ generated from $\tilde{\psi}_+$ and $\tilde{\psi}_-$ by acting upon them with all possible sums of products of operators of the GBM Lie algebra are disjoint, invariant, and irreducible, and that $\tilde{\mathfrak{S}}_{\frac{1}{2}}$ is the direct sum of the corresponding irreducible representations of the GBM group.

Consider the two sequences

$$\begin{aligned} S_+ &\equiv \{\tilde{\psi}_+, \tilde{R}^+\tilde{\psi}_+, \tilde{R}^{+2}\tilde{\psi}_+, \cdots, \tilde{R}^{+\lambda}\psi_+, \cdots\}, \\ S_- &\equiv \{\tilde{\psi}_-, {}^+\tilde{\psi}_-, \tilde{R}^{+2}\tilde{\psi}_-, \cdots, \tilde{R}^{+\lambda}\tilde{\psi}_-, \cdots\} \end{aligned} \tag{17}$$

of elements of $\tilde{\mathfrak{K}}$: the kth element of each sequence is of type $(\frac{1}{2}(k-1), \frac{1}{2}(k-1))$ and belongs to a subspace of $\tilde{\mathfrak{K}}$ irreducible under $\tilde{\mathfrak{S}}_L$, corresponding to a representation of L with spin $\frac{1}{2}(k-1)$. Denote by $\tilde{\mathfrak{K}}_+$ the sum of all the subspaces of $\tilde{\mathfrak{K}}$ which are irreducible under $\tilde{\mathfrak{S}}_L$ and contain one of the elements of the sequence ζ_+, and by $\tilde{\mathfrak{K}}_-$ the subspace of $\tilde{\mathfrak{K}}$ defined in an analogous way in terms of the sequence ζ_-. It is easy to verify that the linear independence of $\tilde{\psi}_+$ and $\tilde{\psi}_-$ implies the linear independence of $\tilde{R}^{+\lambda}\tilde{\psi}_+$ and $\tilde{R}^{+\lambda}\tilde{\psi}_-$ for all values of λ, so that, for each half-integer $\lambda + \frac{1}{2}$, $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ contain distinct subspaces which are irreducible under $\tilde{\mathfrak{S}}_L$ and correspond to a representation of L with spin $\lambda + \frac{1}{2}$, $(\lambda = 0, 1, \cdots)$. Hence one is immediately led to the conclusion that $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ are disjoint, and that $\tilde{\mathfrak{K}} = \tilde{\mathfrak{K}}_+ + \tilde{\mathfrak{K}}_-$.

It must be shown that $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ are invariant and irreducible under the operators of $\tilde{\mathfrak{S}}_{\frac{1}{2}}$.

The irreducibility follows, exactly as in the case of the representation $\tilde{\mathfrak{S}}_0$, from the fact that any vector of each sequence (17) can be transformed into any other vector of the same sequence by acting upon it with a suitable product of the operators $\tilde{R}^+$ and $\tilde{R}_\lambda^-$. To prove that $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ are invariant subspaces it is sufficient to show that $\tilde{\psi}_+$ cannot be transformed into $\tilde{\psi}_-$, and vice versa. Consider any element $\tilde{\psi}$ of type $(\frac{1}{2}, \frac{1}{2})$: the subspace of $\tilde{\mathfrak{K}}_0$ generated from $\tilde{\psi}$ is easily seen to be constituted of elements of the form

$$\tilde{\psi}' = \sum_{l,m} a_{lm}\tilde{Q}_{lm}\tilde{\psi} + \sum_{l,m} b_{lm}\tilde{Q}_{lm}\tilde{H}_-\psi,$$

where the a_{lm}'s and the b_{lm}'s are arbitrary constants. $\tilde{\psi}'$ is also of type $(\frac{1}{2}, \frac{1}{2})$ if and only if $\tilde{H}_3\psi' = \frac{1}{2}\tilde{\psi}'$ and $\tilde{H}_+\tilde{\psi}' = 0$: from (2), (3) and the concluding remark of Sec. IIIB1, one sees that this is only possible if $a_{10} + 2^{-\frac{1}{2}}b_{11} = 0$ and all the other constants a_{lm} and b_{lm} vanish, except a_{00} which remains arbitrary. In other words, any operator which transforms an element of type $(\frac{1}{2}, \frac{1}{2})$ into another element of the same type is a linear combination of the operator $\tilde{R}_{\frac{1}{2}}^=$ and the identity operator. Since $\tilde{\psi}_+$ and $\tilde{\psi}_-$ are independent eigenvectors of the operator $\tilde{R}_{\frac{1}{2}}^=$, both of type $(\frac{1}{2}, \frac{1}{2})$, there exists no operator which transforms one of them into the other, and the invariance of $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ is proved.

Thus $\tilde{\mathfrak{S}}_{\frac{1}{2}}$ is the direct sum of two irreducible representations $t_{\frac{1}{2}}^+$ and $t_{\frac{1}{2}}^-$ of the GBM group, with representation spaces $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$, respectively.

3. *Reduction of* $\tilde{\mathfrak{S}}_s$ $(s > \frac{1}{2})$.

In the general case the operator $\tilde{R}_s^=$ has $2s + 1$ independent eigenvectors of type (s, s). Two of these eigenvectors, denoted by $\tilde{\psi}_+$ and $\tilde{\psi}_-$, are characterized by the additional condition

$$\tilde{R}_s^-\tilde{\psi}_+ = 0, \quad \tilde{R}_s^-\tilde{\psi}_- = 0;$$

their eigenvalues are $(3/2\pi)^{\frac{1}{2}}s$ and $-(3/2\pi)^{\frac{1}{2}}s$, respectively.

The subspaces $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ of $\tilde{\mathfrak{K}}$ which can be generated from $\tilde{\psi}_+$ and $\tilde{\psi}_-$ are disjoint, invariant, and irreducible, and the restrictions of $\tilde{\mathfrak{S}}_s$ to $\tilde{\mathfrak{K}}_+$ and $\tilde{\mathfrak{K}}_-$ are irreducible representations of the GBM group (denoted by t_s^+ and t_s^-, respectively). The reduction of the corresponding induced representations of L gives in both cases, exactly once, each irreducible representation of L with mass m and spin s, $s + 1$, $s + 2, \cdots$. The restriction of $\tilde{\mathfrak{S}}_s$ to the orthogonal complement of $\tilde{\mathfrak{K}}_+ + \tilde{\mathfrak{K}}_-$ in $\tilde{\mathfrak{K}}$ is equivalent to the representation $\tilde{\mathfrak{S}}_{s-1}$. The proofs of these statements

are straightforward generalizations of the analogous proofs given in the special case $s = \frac{1}{2}$.

One can conclude that *the irreducible components of the representations considered are characterized by the eigenvalue m^2 of the rest-mass operator, by the lowest eigenvalue $s(s+1)$ of the spin operator, and by the sign of the eigenvalue belonging to the eigenvector of type (s, s) of the operator $\tilde{R}_s^{=}$.*

ACKNOWLEDGMENT

The author is deeply indebted to Dr. F. A. E. Pirani for his constant encouragement and help.

APPENDIX

Denote by (Λ_j^i) the matrix of the homogeneous Lorentz transformation Λ in the Cartesian frame $\{y^i\}$ associated with the polar coordinates (u, r, θ, φ). One has[1]

$$r' = K_\Lambda(\theta, \varphi)r + J + O(r^{-1}),$$
$$u = y^0 - r, \quad u' = y^{0'} - r',$$

so that the transformation

$$u' = K_\Lambda^{-1}(\theta, \varphi)u + O(r^{-1})$$

can be written in the form

$$\Lambda_0^0 y^0 + \Lambda_1^0 r \sin\theta\cos\varphi + \Lambda_2^0 r \sin\theta\sin\varphi + \Lambda_3^0 r\cos\theta - K_\Lambda r - J + O(r^{-1}) = K_\Lambda^{-1}u + O(r^{-1}).$$

Dividing by r and passing to the limit for $r \to \infty$, $u = \text{const}$, one gets the following expression for K_Λ:

$$K_\Lambda(\theta, \varphi) = \Lambda_0^0 + \Lambda_1^0 \sin\theta\cos\varphi + \Lambda_2^0 \sin\theta\sin\varphi + \Lambda_3^0\cos\theta.$$

Hence, denoting by $A^{\alpha\beta}$ $(\alpha, \beta = 0, 1, 2, 3)$ the value of $[\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)/\partial t]_{t=0}$ corresponding to the one-parameter subgroup of rotations in the coordinate plane (y^α, y^β), one obtains

$$A^{ij} = 0 \quad (i, j = 1, 2, 3),$$
$$A^{01} = [(\partial/\partial t)(\cosh t - \sinh t \sin\Lambda_t^{-1}\theta\cos\Lambda_t^{-1}\varphi)]_{t=0} = -\sin\theta\cos\varphi,$$
$$A^{02} = -\sin\theta\sin\varphi,$$
$$A^{03} = -\cos\theta.$$

Denote by $B^{\alpha\beta}$ and $C^{\alpha\beta}$ the values of $(\partial\Lambda_t^{-1}\theta/\partial t)_{t=0}$ and $(\partial\Lambda_t^{-1}\varphi/\partial t)_{t=0}$, respectively, for the one-parameter subgroup of four-rotations in the coordinate plane (y^α, y^β). By differentiation with respect to t of the equations

$$\cos\Lambda_t^{-1}\theta = \frac{\Lambda^3_{-ti}y^i}{r'}, \quad \left(r' = \left[\sum_\alpha (\Lambda^i_{-t\alpha}y^\alpha)^2\right]^{\frac{1}{2}}\right),$$

and

$$\tan\Lambda_t^{-1}\varphi = \frac{\Lambda^2_{-ti}y^i}{\Lambda^1_{-tj}y^j},$$

one obtains, passing to the limit for $r \to \infty$, $u = \text{const}$:

$$B^{12} = -1, \quad B^{23} = -\sin\varphi, \quad B^{31} = \cos\varphi,$$
$$B^{01} = \cos\theta\cos\varphi, \quad B^{02} = \cos\theta\sin\varphi,$$
$$B^{03} = -\sin\theta,$$
$$C^{12} = 0, \quad C^{23} = -\cot\theta\cos\varphi,$$
$$C^{31} = -\cot\theta\sin\varphi,$$
$$C^{01} = -\frac{\sin\varphi}{\sin\theta}, \quad C^{02} = \frac{\cos\varphi}{\sin\theta}, \quad C^{03} = 0.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Erratum: Nonrelativistic Quantum Theory of an Electron in an Arbitrarily Intense Laser Field

H. H. NICKLE

[J. Math. Phys. 7, 1497 (1966)]

(Received 6 December 1965)

THE term $(e^2\mathbf{A}_\omega \cdot \mathbf{A}_\omega^*/mc^2)a_n$ was omitted from the right-hand side of Eq. (2.14). This term can be combined with the first term on the right-hand side of Eq. (2.14) if we replace $\hbar\omega_n$ by

$$\hbar\tilde{\omega}_n \equiv \hbar\omega_n + (e^2\mathbf{A}_\omega \cdot \mathbf{A}_\omega^*/mc^2),$$

which corresponds to the addition of the electron's zero-point energy due to the presence of the electromagnetic field described by the vector potential (2.3). Thus, in the subsequent equations it is necessary to replace $\hbar\omega_n$ by $\hbar\tilde{\omega}_n$ in *every* equation referring to the solution (or properties of the electron state) in the presence of the laser field.

are straightforward generalizations of the analogous proofs given in the special case $s = \frac{1}{2}$.

One can conclude that *the irreducible components of the representations considered are characterized by the eigenvalue m^2 of the rest-mass operator, by the lowest eigenvalue $s(s+1)$ of the spin operator, and by the sign of the eigenvalue belonging to the eigenvector of type (s, s) of the operator $\tilde{R}_s^{=}$.*

ACKNOWLEDGMENT

The author is deeply indebted to Dr. F. A. E. Pirani for his constant encouragement and help.

APPENDIX

Denote by (Λ_j^i) the matrix of the homogeneous Lorentz transformation Λ in the Cartesian frame $\{y^i\}$ associated with the polar coordinates (u, r, θ, φ). One has[1]

$$r' = K_\Lambda(\theta, \varphi)r + J + O(r^{-1}),$$
$$u = y^0 - r, \quad u' = y^{0'} - r',$$

so that the transformation

$$u' = K_\Lambda^{-1}(\theta, \varphi)u + O(r^{-1})$$

can be written in the form

$$\Lambda_0^0 y^0 + \Lambda_1^0 r \sin\theta\cos\varphi + \Lambda_2^0 r \sin\theta\sin\varphi + \Lambda_3^0 r\cos\theta - K_\Lambda r - J + O(r^{-1}) = K_\Lambda^{-1}u + O(r^{-1}).$$

Dividing by r and passing to the limit for $r \to \infty$, $u = \text{const}$, one gets the following expression for K_Λ:

$$K_\Lambda(\theta, \varphi) = \Lambda_0^0 + \Lambda_1^0 \sin\theta\cos\varphi + \Lambda_2^0 \sin\theta\sin\varphi + \Lambda_3^0\cos\theta.$$

Hence, denoting by $A^{\alpha\beta}$ $(\alpha, \beta = 0, 1, 2, 3)$ the value of $[\partial K_{\Lambda_t}(\Lambda_t^{-1}\theta, \Lambda_t^{-1}\varphi)/\partial t]_{t=0}$ corresponding to the one-parameter subgroup of rotations in the coordinate plane (y^α, y^β), one obtains

$$A^{ij} = 0 \quad (i, j = 1, 2, 3),$$
$$A^{01} = [(\partial/\partial t)(\cosh t - \sinh t \sin\Lambda_t^{-1}\theta\cos\Lambda_t^{-1}\varphi)]_{t=0} = -\sin\theta\cos\varphi,$$
$$A^{02} = -\sin\theta\sin\varphi,$$
$$A^{03} = -\cos\theta.$$

Denote by $B^{\alpha\beta}$ and $C^{\alpha\beta}$ the values of $(\partial\Lambda_t^{-1}\theta/\partial t)_{t=0}$ and $(\partial\Lambda_t^{-1}\varphi/\partial t)_{t=0}$, respectively, for the one-parameter subgroup of four-rotations in the coordinate plane (y^α, y^β). By differentiation with respect to t of the equations

$$\cos\Lambda_t^{-1}\theta = \frac{\Lambda^3_{-ti}y^i}{r'}, \quad \left(r' = \left[\sum_\alpha (\Lambda^i_{-t\alpha}y^\alpha)^2\right]^{\frac{1}{2}}\right),$$

and

$$\tan\Lambda_t^{-1}\varphi = \frac{\Lambda^2_{-ti}y^i}{\Lambda^1_{-tj}y^j},$$

one obtains, passing to the limit for $r \to \infty$, $u = \text{const}$:

$$B^{12} = -1, \quad B^{23} = -\sin\varphi, \quad B^{31} = \cos\varphi,$$
$$B^{01} = \cos\theta\cos\varphi, \quad B^{02} = \cos\theta\sin\varphi,$$
$$B^{03} = -\sin\theta,$$
$$C^{12} = 0, \quad C^{23} = -\cot\theta\cos\varphi,$$
$$C^{31} = -\cot\theta\sin\varphi,$$
$$C^{01} = -\frac{\sin\varphi}{\sin\theta}, \quad C^{02} = \frac{\cos\varphi}{\sin\theta}, \quad C^{03} = 0.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 8 AUGUST 1967

Erratum: Nonrelativistic Quantum Theory of an Electron in an Arbitrarily Intense Laser Field

H. H. NICKLE

[J. Math. Phys. 7, 1497 (1966)]

(Received 6 December 1965)

THE term $(e^2\mathbf{A}_\omega \cdot \mathbf{A}_\omega^*/mc^2)a_n$ was omitted from the right-hand side of Eq. (2.14). This term can be combined with the first term on the right-hand side of Eq. (2.14) if we replace $\hbar\omega_n$ by

$$\hbar\tilde{\omega}_n \equiv \hbar\omega_n + (e^2\mathbf{A}_\omega \cdot \mathbf{A}_\omega^*/mc^2),$$

which corresponds to the addition of the electron's zero-point energy due to the presence of the electromagnetic field described by the vector potential (2.3). Thus, in the subsequent equations it is necessary to replace $\hbar\omega_n$ by $\hbar\tilde{\omega}_n$ in *every* equation referring to the solution (or properties of the electron state) in the presence of the laser field.