JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 8, NUMBER 6 JUNE 1967

Branching Theorem for the Symplectic Groups

G. C. HEGERFELDT
Institut für Theoretische Physik, Universität Marburg, Germany

(Received 17 November 1966)

A theorem analogous to the Weyl branching law for the unimodular groups is derived for $Sp(2n)$.

IN the investigation of possible higher symmetry groups in elementary particle physics several authors have considered symplectic groups, as for instance $Sp(6)$ in Ref. 1 and $Sp(14)$ in Ref. 2. Beside this, these groups are of importance in nuclear physics.[3] It may therefore be useful to have an analogy of the well-known Weyl branching law for the unimodular group $SL(n)$,[4] that is, to have a formula for the splitting of an irreducible representation of $Sp(2n)$ under restriction to $Sp(2(n-1))$.

Two special cases, $Sp(6) \downarrow Sp(4)$ and $Sp(4) \downarrow Sp(2)$, have been treated in Ref. 5, a general formula for the symplectic groups, however, as far as the author knows, has not yet been published. In the following such a theorem is formulated and proved.

The irreducible representations D of $Sp(2n)$ can be classified by n nonnegative integers $\{m_1 \cdots m_n\}$ with $m_1 \geq \cdots \geq m_n \geq 0$ which determine a corresponding Young pattern (cf. Ref. 6). Write the space R_{2n}, in which the group $Sp(2n)$ of linear transformations acts, as a direct sum of a $2(n-1)$-dimensional and a 2-dimensional subspace $R_{(2n-1)}$ and R_2, respectively. Then the subgroup of $Sp(2n)$ which leaves $R_{2(n-1)}$ invariant and acts as identity transformation in R_2 is isomorphic to $Sp(2(n-1))$. We are now going to prove the following branching law for a restriction of $D_{\{m\}}$ to this subgroup:

Theorem: On restricting a representation $D_{\{m_1 \cdots m_n\}}$ of $Sp(2n)$ to the subgroup $Sp(2(n-1))$ one has the following splitting into irreducible representations $D_{\{m_1'' \cdots m_{n-1}''\}}$ of $Sp(2(n-1))$:

$$^{Sp(2n)}D_{\{m_1 \cdots m_n\}} \mid {}_{Sp(2(n-1))} = \sum_{m_1 \geq m_1' \geq \cdots \geq m_{n-1}' \geq m_n \geq m_n' \geq 0} \times \sum_{m_1' \geq m_1'' \geq \cdots \geq m_{n-1}'' \geq m_n'} {}^{Sp(2(n-1))}D_{\{m_1'' \cdots m_{n-1}''\}} \quad (1)$$

with m_i', m_j'' integers.

Proof. The characters $\chi_{\{m\}}$ of an irreducible representation $D_{\{m\}}$ of $Sp(2n)$ can be expressed as a function of the n independent characteristic roots $\epsilon_1, \cdots, \epsilon_n$ of an arbitrary matrix of $Sp(2n)$:

$$^{Sp(2n)}\chi_{\{m\}} = \chi_{\{m_1 \cdots m_n\}}^{(\epsilon_1 \cdots \epsilon_n)}.$$

For the subgroup $Sp(2(n-1))$ one has with a suitable numbering of the roots $\epsilon_n = 1$. Hence in order to prove (1) it suffices to establish the corresponding

[1] H. Bacry, J. Nuyts, and L. Van Hove, Nuovo Cimento **35**, 510 (1965).
[2] H. D. Doebner and G. C. Hegerfeldt, J. Math. Phys. **8**, 731 (1967).
[3] Cf., e.g., B. H. Flowers, Proc. Roy. Soc. (London) **A212**, 248 (1952).
[4] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications Inc., New York, 1963).
[5] M. L. Whippman, J. Math. Phys. **6**, 1534 (1965).
[6] H. Weyl, *Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946).

formula for the characters, namely,

$$
\begin{aligned}
&{}^{Sp(2n)}\chi^{(\epsilon_1\cdots\epsilon_n)}_{\{m_1\cdots m_n\}}\Big|_{\epsilon_n=1} \\
&\quad= \sum_{m_1\geq m_1'\geq\cdots\geq m_n\geq m_n'\geq 0} \\
&\qquad\times \sum_{m_1'\geq m_1''\geq\cdots\geq m_{n-1}''\geq m_n'} {}^{Sp(2(n-1))}\chi^{(\epsilon_1\cdots\epsilon_{n-1})}_{\{m_1''\cdots m_{n-1}''\}}\cdot
\end{aligned}
\tag{2}
$$

$\chi_{\{m\}}$ can be written as the quotient of two determinants (Ref. 6, p. 218):

$$
\chi^{(\epsilon_1\cdots\epsilon_n)}_{\{m_1\cdots m_n\}} = \frac{|\epsilon^{l_1}-\epsilon^{-l_1},\cdots,\epsilon^{l_n}-\epsilon^{-l_n}|}{|\epsilon^{n}-\epsilon^{-n},\cdots,\epsilon^{1}-\epsilon^{-1}|}, \tag{3}
$$

where the ith row of the determinants in numerator and denominator are obtained by attaching the suffix i to the ϵ's and where

$$
l_j = m_j + (n-j+1). \tag{4}
$$

For the denominator in (3) one has

$$
\begin{aligned}
&|\epsilon^{n}-\epsilon^{-n},\cdots,\epsilon^{1}-\epsilon^{-1}| \\
&\quad= \prod_{i=1}^{n}(\epsilon_i^{\frac{1}{2}}+\epsilon_i^{-\frac{1}{2}})\cdot|\epsilon^{n-\frac{1}{2}}-\epsilon^{-(n-\frac{1}{2})},\cdots,\epsilon^{\frac{1}{2}}-\epsilon^{-\frac{1}{2}}|,
\end{aligned}
\tag{5}
$$

which is checked by multiplying the ith row of the determinant on the right-hand side by $(\epsilon_i^{\frac{1}{2}}+\epsilon_i^{-\frac{1}{2}})$ and then, starting with the last column, subtract each column from the preceding one.

We use the following identity which can be derived from Ref. 7, Chap. VII, Eq. (12.2):

$$
\begin{aligned}
&\frac{|\epsilon^{l_1}-\epsilon^{-l_1},\cdots,\epsilon^{l_n}-\epsilon^{-l_n}|}{|\epsilon^{n-\frac{1}{2}}-\epsilon^{-(n-\frac{1}{2})},\cdots,\epsilon^{\frac{1}{2}}-\epsilon^{-\frac{1}{2}}|} \\
&= \frac{1}{|\epsilon^{n-1}+\epsilon^{-(n-1)},\cdots,\epsilon^{0}+\epsilon^{0}|} \\
&\quad\times \sum_{l_1>l_1'>\cdots>l_n>|l_n'|} (|\epsilon^{l_1'}+\epsilon^{-l_1'},\cdots,\epsilon^{l_n'}+\epsilon^{-l_n'}| \\
&\qquad + |\epsilon^{l_1'}-\epsilon^{-l_1'},\cdots,\epsilon^{l_n'}-\epsilon^{-l_n'}|),
\end{aligned}
\tag{6}
$$

the l_i' being half-integers.

Now let $\epsilon_n \to 1$ in (6). Then the last column and last row of the denominator on the right-hand side of (6) consist of 2's only, and for the sum one obtains (with $j = 1,\cdots,n-1$)

$$
\begin{aligned}
&\sum_{l_1>l_1'>\cdots>l_n>|l_n'|}\left(\begin{vmatrix} \epsilon_j^{l_1'}+\epsilon_j^{-l_1'},\cdots,\epsilon_j^{l_n'}+\epsilon_j^{-l_n'} \\ 2 \quad,\cdots,\quad 2 \end{vmatrix} + 0\right) \\
&= 2\cdot \sum_{l_1>l_1'>\cdots>l_n>l_n'>0} \begin{vmatrix} \epsilon_j^{l_1'}+\epsilon_j^{-l_1'},\cdots,\epsilon_j^{l_n'}+\epsilon_j^{-l_n'} \\ 2 \quad,\cdots,\quad 2 \end{vmatrix}
\end{aligned}
\tag{7}
$$

since l_n' runs from $-(l_n-\frac{1}{2})$ to $(l_n-\frac{1}{2})$ and the determinant is even under $l_n' \to -l_n'$.

From Eq. (12.5) of Chap. VII, Ref. 7, one can deduce the following relation (with $j = 1,\cdots,n-1$; l_i' half-integer, l_i'' integer):

$$
\begin{aligned}
&\frac{\begin{vmatrix} \epsilon_j^{l_1'}+\epsilon_j^{-l_1'},\cdots,\epsilon_j^{l_n'}+\epsilon_j^{-l_n'} \\ 2 \quad,\cdots,\quad 2 \end{vmatrix}}{\begin{vmatrix} \epsilon_j^{n-1}+\epsilon_j^{-(n-1)},\cdots,\epsilon_j^{0}+\epsilon_j^{0} \\ 2 \quad,\cdots,\quad 2 \end{vmatrix}} \\
&= \frac{1}{|\epsilon_j^{n-\frac{3}{2}}-\epsilon_j^{-(n-\frac{3}{2})},\cdots,\epsilon_j^{\frac{1}{2}}-\epsilon_j^{-\frac{1}{2}}|} \\
&\qquad \sum_{l_1'>l_1''>\cdots>l_{n-1}''>|l_n'|} \\
&\qquad\times |\epsilon_j^{l_1''}-\epsilon_j^{-l_1''},\cdots,\epsilon_j^{l_{n-1}''}-\epsilon_j^{-l_{n-1}''}|.
\end{aligned}
\tag{8}
$$

Combining Eqs. (3), (5), (7), and (8) one gets [with $l_j = m_j + (n-j+1)$]:

$$
\begin{aligned}
{}^{Sp(2n)}\chi^{(\epsilon_1,\cdots,1)}_{\{m\}} &= \frac{1}{2\prod_{i=1}^{n-1}(\epsilon_i^{\frac{1}{2}}+\epsilon_i^{-\frac{1}{2}})}\cdot \sum_{l_1>l_1'>\cdots>l_n>l_n'>0} \\
&\quad\times \sum_{l_1'>l_1''>\cdots>l_{n-1}''>l_n'} \\
&\quad\times \frac{|\epsilon^{l_1''}-\epsilon^{-l_1''},\cdots,\epsilon^{l_{n-1}''}-\epsilon^{-l_{n-1}''}|}{|\epsilon^{n-\frac{3}{2}}-\epsilon^{-(n-\frac{3}{2})},\cdots,\epsilon^{\frac{1}{2}}-\epsilon^{-\frac{1}{2}}|},
\end{aligned}
\tag{9}
$$

where l_i, l_j'' are integers, l_i' half-integers. Similar to (5) one has

$$
\begin{aligned}
&\prod_{i=1}^{n-1}(\epsilon_i^{\frac{1}{2}}+\epsilon_i^{-\frac{1}{2}})\,|\epsilon^{n-\frac{3}{2}}-\epsilon^{-(n-\frac{3}{2})},\cdots,\epsilon^{\frac{1}{2}}-\epsilon^{-\frac{1}{2}}| \\
&\qquad = |\epsilon^{n-1}-\epsilon^{-(n-1)},\cdots,\epsilon^{1}-\epsilon^{-1}|.
\end{aligned}
\tag{10}
$$

With (3) and (10) and $\bar{l}_i = l_i' + \frac{1}{2}$, Eq. (9) becomes

$$
\begin{aligned}
{}^{Sp(2n)}\chi^{(\epsilon_1,\cdots,\epsilon_{n-1},1)}_{\{m\}} &= \sum_{l_1\geq\bar{l}_1>l_2\geq\cdots\geq\bar{l}_n>0} \\
&\quad\times \sum_{l_1>l_1''\geq l_2>\cdots>l_{n-1}''\geq l_n} {}^{Sp(2(n-1))}\chi^{(\epsilon)}_{\{m_1''\cdots m_{n-1}''\}},
\end{aligned}
\tag{11}
$$

where

$$
m_i'' = l_i'' - (n-i), \quad i = 1,\cdots,n-1.
$$

Putting

$$
m_i' = \bar{l}_i - (n-i+1), \quad i = 1,\cdots,n-1,n, \tag{12}
$$

and summing over m_i', m_j'' instead of $\bar{l}_i$, l_j'' in (11) one arrives at Eq. (2), which is equivalent to (1).

ACKNOWLEDGMENTS

It is a pleasure to thank Professor L. C. Biedenharn, Dr. H. D. Doebner, and Dr. O. Melsheimer for a critical reading of the manuscript.

This work was supported in part by the Bundesministerium für Wissenschaftliche Forschung, Bad Godesberg.

[7] H. Boerner, *Darstellungen von Gruppen* (Springer-Verlag, Berlin, 1955); cf. also the English edition, Chap. VIII: H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Fresnel's Equation in General Relativity

HUGO F. KREMER
Instituto de Física, Universidade Federal do Paraná, Curitiba, Paraná, Brazil

(Received 10 May 1966)

The Fresnel equation is derived in general relativity using the classical method applied by Levi-Civita in the study of a nonrelativistic theory of electromagnetic induction. For the description of the anisotropic medium the theory proposed by Quan is adopted. The study of the Cauchy problem is also presented and the convergence of results assures us that the equation proposed is the good one.

I. INTRODUCTION

IN general relativity, we study the propagation of an electromagnetic wave in a medium with general electric and magnetic anisotropy. For the description of the anisotropic medium we adopt the theory proposed by Quan.[1] The Fresnel equation of wave normals is obtained by the application of the classical method, which consists of analyzing the discontinuity of the first derivative of the electromagnetic field $f_{\alpha\beta}$, once applied by Levi-Civita in the study of the same question in a nonrelativistic theory of electromagnetic induction.[2]

Essentially, we study a special type of singular regions of the 4-space of general relativity. These singular regions are assumed to be 3-spaces Σ_3 (called hypersurfaces of discontinuity), such that across them the electromagnetic field is continuous, but the first derivative may not be continuous. Across Σ_3 all the other physical quantities are supposed to be well behaved and in particular the metric tensor $g_{\alpha\beta}$ and its first derivatives are assumed to be continuous throughout. In the last section the study of the Cauchy problem for Quan's theory is presented.

The scheme followed in the paper generalizes the work on singular hypersurfaces made by Quan,[1] Saini,[3] and others, with the restriction that only the electromagnetic shock waves are studied.

As far as the author knows, a general covariant Fresnel's equation is not presented in the literature.[4] Recently, using semiclassical methods, the author studied the Fresnel equation for vacuum polarization[5] and the desire to confront this result with the equation for a phenomenological theory motivated his interest on the subject.

Throughout the paper, the Latin indices have a range 1, 2, 3 and the Greek indices a range 1, 2, 3, 4. The usual rules of the tensor algebra are used, (for example: $A^{[\alpha}B^{\beta]} \equiv \frac{1}{2}(A^\alpha B^\beta - A^\beta B^\alpha)$, and the partial and covariant derivatives are denoted by ∂_α and ∇_α, respectively. The permutation symbol is $\epsilon^{\alpha\beta\gamma\delta}$ and

$$\eta^{\alpha\beta\gamma\delta} \equiv -(|g|)^{\frac{1}{2}}\epsilon^{\alpha\beta\gamma\delta}, \quad \eta_{\alpha\beta\gamma\delta} \equiv (|g|)^{\frac{1}{2}}\epsilon_{\alpha\beta\gamma\delta}.$$

II. FUNDAMENTAL EQUATIONS

An electromagnetical induction is defined in a domain of the 4-space V_4 of general relativity when there are two antisymmetric tensor fields, $f_{\alpha\beta}$ and $p_{\alpha\beta}$, called electromagnetic and induction field and two nonsingular matrices $[\epsilon_\alpha^\beta]$, and $[\mu_\alpha^\beta]$, called induction matrices, for which

$$\begin{aligned} p_{\beta\alpha}U^\alpha &= \epsilon_\beta^\rho f_{\rho\alpha}U^\alpha, \\ f^*_{\beta\alpha}U^\alpha &= \mu_\beta^\rho p^*_{\rho\alpha}U^\alpha, \end{aligned} \tag{1}$$

where U^α is the 4-velocity of the charged medium defined in the domain.

The notation employed is

$$\begin{aligned} f_{\alpha\beta} &\equiv (\mathbf{E}, \mathbf{B}), \quad p_{\alpha\beta} \equiv (\mathbf{D}, \mathbf{H}), \\ f^*_{\alpha\beta} &= \tfrac{1}{2}\eta_{\alpha\beta\gamma\delta}f^{\gamma\delta}, \quad p^*_{\alpha\beta} = \tfrac{1}{2}\eta_{\alpha\beta\gamma\delta}p^{\gamma\delta}. \end{aligned} \tag{2}$$

If we introduce the 4-vectors, electric and magnetic fields and inductions, defined by

$$\begin{aligned} E_\alpha &\equiv f_{\alpha\rho}U^\rho, \quad B_\alpha \equiv f^*_{\alpha\rho}U^\rho, \\ D_\alpha &\equiv p_{\alpha\rho}U^\rho, \quad H_\alpha \equiv p^*_{\alpha\rho}U^\rho, \end{aligned} \tag{3}$$

the constitutive equations (1) may be written

$$D_\beta = \epsilon_\beta^\rho E_\rho, \quad B_\beta = \mu_\beta^\rho H_\rho. \tag{4}$$

It is seen that the matrices $[\epsilon_\alpha^\beta]$ and $[\mu_\alpha^\beta]$ represent two automorphisms of the vector space tangent to V_4.

[1] P. M. Quan, *Problèmes actuels en théorie de la relativité* (Revue d'Optique, Paris, 1959), p. 61.
[2] T. Levi-Civita, *Caracteristiche dei Sistemi differenziali e propagazione ondosa* (Zanichelli, Bologna, 1931).
[3] G. L. Saini, Proc. Roy. Soc. (London) **A260**, 61 (1961).
[4] Cf., for example, L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, London, 1960).
[5] H. F. Kremer, Phys. Rev. **139**, B254 (1965).

If we represent by $[\lambda_\alpha^\beta]$ and $[\tau_\alpha^\beta]$ the inverse transformations of $[\epsilon_\alpha^\beta]$ and $[\mu_\alpha^\beta]$,

$$\lambda_\alpha^\gamma \epsilon_\gamma^\beta = \delta_\alpha^\beta, \quad \tau_\alpha^\gamma \mu_\gamma^\beta = \delta_\alpha^\beta, \tag{5}$$

we get for the constitutive equations (1):

$$f_{\rho\alpha}U^\alpha = \lambda_\rho^\gamma p_{\gamma\alpha}U^\alpha, \quad p^*_{\rho\alpha}U^\alpha = \tau_\rho^\gamma f^*_{\gamma\alpha}U^\alpha. \tag{6}$$

The two fields $f_{\alpha\beta}$ and $p_{\alpha\beta}$ are supposed to satisfy the Maxwell equations

$$\nabla_\alpha f^{*\alpha\beta} = 0, \tag{7a}$$

$$\nabla_\alpha p^{\alpha\beta} = J^\beta, \tag{7b}$$

where the 4-current J^β is related to $f_{\alpha\beta}$ by the constitutive equation called Ohm's law

$$J^\beta = \sigma U_\alpha f^{\beta\alpha} + \delta U^\beta. \tag{8}$$

In the last equation the scalars σ and $\delta = U_\beta J^\beta$ are respectively the electrical conductivity and the excess charge.

In what follows, it is useful to consider the inverse relations of (3):

$$f^{\alpha\beta} = 2E^{[\alpha}U^{\beta]} - \eta^{\alpha\beta\mu\nu}B_\mu U_\nu, \tag{9a}$$

$$f^*_{\alpha\beta} = 2B_{[\alpha}U_{\beta]} + \eta_{\alpha\beta\mu\nu}E^\mu U^\nu, \tag{9b}$$

and

$$p^{\alpha\beta} = 2D^{[\alpha}U^{\beta]} - \eta^{\alpha\beta\mu\nu}H_\mu U_\nu, \tag{10a}$$

$$p^*_{\alpha\beta} = 2H_{[\alpha}U_{\beta]} + \eta_{\alpha\beta\mu\nu}D^\mu U^\nu. \tag{10b}$$

With the aid of Eqs. (10a), (3), (1), and (6) we obtain $p^{\alpha\beta}$ as a function of $f_{\alpha\beta}$ and $f^*_{\alpha\beta}$:

$$p^{\alpha\beta} = 2g^{\sigma[\alpha}U^{\beta]}\epsilon_\sigma^\rho f_{\rho\lambda}U^\lambda + \eta^{\alpha\beta\gamma\sigma}\tau_\sigma^\rho f^*_{\rho\mu}U^\mu U_\gamma. \tag{11}$$

The last equation is a generalization of the well-known equation valid for an isotropic medium ($\epsilon_\alpha^\beta = \delta_\alpha^\beta\epsilon$ and $\tau_\alpha^\beta = \mu^{-1}\delta_\alpha^\beta$):

$$p^{\alpha\beta} = (1/\mu)f^{\alpha\beta} + 2[(1 - \epsilon\mu)/\mu]U_\sigma f^{\sigma[\alpha}U^{\beta]}. \tag{12}$$

III. FRESNEL EQUATION

The equation of wave normals for the electromagnetic theory with inductions, may be obtained by the application of the method of Levi-Civita.[2]

In the domain of existence V_4 of $f_{\alpha\beta}$, we define a hypersurface of discontinuity Σ_3

$$\tau(x^\alpha) = C, \tag{13}$$

to which corresponds at a given moment a wave surface which separates V_4 into two distinct regions. The discontinuities $\Delta(\partial_\gamma f_{\alpha\beta})$ and $\Delta(\partial_\gamma p_{\alpha\beta})$ of the first derivatives of $f_{\alpha\beta}$ and $p_{\alpha\beta}$ across Σ_3 must satisfy the so-called "conditions of geometrical compatibility"

$$\Delta(\nabla_\gamma f_{\alpha\beta}) = \Delta(\partial_\gamma f_{\alpha\beta}) = \Lambda_{\alpha\beta}\partial_\gamma\tau, \tag{14a}$$

$$\Delta(\nabla_\gamma p_{\alpha\beta}) = \Delta(\partial_\gamma p_{\alpha\beta}) = K_{\alpha\beta}\partial_\gamma\tau, \tag{14b}$$

$\Lambda_{\alpha\beta}$ and $K_{\alpha\beta}$ representing the "discontinuity tensors."

The derivative of (11) gives

$$\nabla_\gamma p^{\alpha\beta} = 2g^{\sigma[\alpha}U^{\beta]}\epsilon_\sigma^\rho U^\lambda \nabla_\gamma f_{\rho\lambda} + \eta^{\alpha\beta\lambda\sigma}\tau_\sigma^\rho U^\mu U_\lambda \nabla_\gamma f^*_{\rho\mu} + \psi_\gamma^{\alpha\beta}, \tag{15}$$

where the quantities $\psi_\gamma^{\alpha\beta}$ are independent of the derivatives of $f_{\alpha\beta}$ and $f^*_{\alpha\beta}$.

If we calculate the discontinuity of $\nabla_\gamma p^{\alpha\beta}$ across Σ_3 with the use of (15), and taking into account that $f_{\alpha\beta}$ and $f^*_{\alpha\beta}$ are continuous across Σ_3 we have

$$\Delta(\partial_\gamma p^{\alpha\beta}) = 2g^{\sigma[\alpha}U^{\beta]}\epsilon_\sigma^\rho U^\lambda \Delta(\partial_\gamma f_{\rho\lambda}) + \eta^{\alpha\beta\lambda\sigma}\tau_\sigma^\rho U^\mu U_\lambda \Delta(\partial_\gamma f^*_{\rho\mu}). \tag{16}$$

Since not all the $\partial_\gamma\tau$ are zero, we get from (16), (14a), and (14b):

$$K^{\alpha\beta} = 2g^{\sigma[\alpha}U^{\beta]}\epsilon_\sigma^\rho U^\lambda \Lambda_{\rho\lambda} + \eta^{\alpha\beta\lambda\sigma}\tau_\sigma^\rho U^\mu U_\lambda \Lambda^*_{\rho\mu} \tag{17}$$

with

$$\Lambda^*_{\rho\mu} = \tfrac{1}{2}\eta_{\rho\mu\epsilon\xi}\Lambda^{\epsilon\xi} = \tfrac{1}{2}\eta_{\rho\mu\epsilon\xi}g^{\epsilon\lambda}g^{\xi\nu}\Lambda_{\lambda\nu}. \tag{18}$$

The relation (17) between $K^{\alpha\beta}$ and $\Lambda^{\alpha\beta}$ depends only on the relation between $f_{\alpha\beta}$ and $p_{\alpha\beta}$, that is, on the constitutive equations.

These tensors must also satisfy other conditions imposed by the field equations of the theory. These conditions, called "dynamic compatibility conditions," are obtained from (7a) and (7b):

$$\eta^{\beta\gamma\delta\epsilon}\Lambda_{\gamma\delta}\partial_\beta\tau = 0, \tag{19a}$$

$$K^{\beta\alpha}\partial_\beta\tau = 0. \tag{19b}$$

With the aid of (17) and (18), Eq. (19b) can be written in the form

$$(2g^{\lambda[\beta}U^{\alpha]}\epsilon_\lambda^\rho U^\sigma + \tfrac{1}{2}\eta^{\beta\alpha\lambda\nu}\eta_{\delta\mu\epsilon\xi}g^{\epsilon\rho}g^{\xi\sigma}U^\mu U_\lambda\tau_\nu^\delta)\Lambda_{\rho\sigma}\partial_\beta\tau = 0. \tag{20}$$

We note that (19a) admits the solution

$$\Lambda_{\rho\sigma} = Q_\sigma\partial_\rho\tau - Q_\rho\partial_\sigma\tau, \tag{21}$$

where Q_ρ is an arbitrary vector. Taking into account gauge invariance, we see that

$$\Lambda_{\rho\sigma} = q_\sigma\partial_\rho\tau - q_\rho\partial_\sigma\tau, \tag{22a}$$

$$q_\sigma = Q_\sigma + k\partial_\sigma\tau, \tag{22b}$$

with k, an arbitrary constant. With the condition $\partial_4\tau \neq 0$, we can always choose k so that $q_4 = 0$ and we adopt this gauge.

If we substitute (22a) into (20) we get

$$M^{\alpha\beta\rho\sigma}q_\sigma\partial_\beta\tau\partial_\rho\tau = 0, \tag{23}$$

where

$$M^{\alpha\beta\rho\sigma} \equiv \eta^{\beta\alpha\lambda\nu}\eta_{\delta\mu\epsilon\xi}U^\mu U_\lambda\tau_\nu^\delta g^{\epsilon\rho}g^{\xi\sigma} + 4g^{\lambda[\beta}U^{\alpha]}\epsilon_\lambda^{[\rho}U^{\sigma]}. \tag{24}$$

With the definition

$$H^{\alpha\sigma} \equiv M^{\alpha\beta\rho\sigma}\partial_\beta\tau\partial_\rho\tau, \tag{25}$$

Eq. (23) may be written

$$H^{\alpha\sigma}q_\sigma = 0. \tag{26}$$

These equations forming a linear and homogeneous system (26) are not independent because $H_{\alpha\sigma}$ satisfies the trivial identity

$$(H^{\alpha\sigma}q_\sigma)\partial_\alpha\tau \equiv 0. \tag{27}$$

Since $\partial_4\tau \neq 0$, we can express the fourth equation (26) as a linear combination of the three others and the system is reduced to

$$H^{\alpha s}q_s = 0. \tag{28}$$

The effective existence of discontinuities of $\partial_\gamma f_{\alpha\beta}$ across Σ_3 implies $q_s \neq 0$, that is,

$$\det|H^{\alpha s}| = 0 \tag{29}$$

or

$$\det|M^{\alpha\beta\rho s}\partial_\beta\tau\partial_\rho\tau| = 0. \tag{30}$$

This is the general covariant Fresnel equation for a medium with electric and magnetic anisotropy. If the medium presents electric anisotropy but is magnetically isotropic [$\tau_\alpha^\beta = (1/\mu)\delta_\alpha^\beta$], the last equation assumes the form

$$\begin{aligned}\det|[&(g^{\beta\rho}g^{\alpha s} - g^{\alpha\rho}g^{\beta s}) + (g^{\alpha\rho} - \mu g^{\lambda\alpha}\epsilon_\lambda^\rho)U^\beta U^s \\ &- (g^{\alpha s} - \mu g^{\lambda\alpha}\epsilon_\lambda^s)U^\beta U^\rho + (g^{\beta s} - \mu g^{\lambda\beta}\epsilon_\lambda^s)U^\rho U^\alpha \\ &- (g^{\beta\rho} - \mu g^{\lambda\beta}\epsilon^\rho)U^\alpha U^s]\partial_\beta\tau\partial_\rho\tau| = 0.\end{aligned} \tag{31}$$

For the complete isotropy (also $\epsilon_\alpha^\beta = \epsilon\delta_\alpha^\beta$), we have

$$\det|(\bar{g}^{\beta\rho}\bar{g}^{\alpha s} - \bar{g}^{\alpha\rho}\bar{g}^{\beta s})\partial_\beta\tau_\rho\partial\tau| = 0, \tag{32}$$

where

$$\bar{g}^{\alpha\rho} \equiv g^{\alpha\rho} - (1 - \epsilon\mu)U^\alpha U^\rho. \tag{33}$$

Developing Eq. (32) we get the simplest form

$$\bar{g}^{\alpha\beta}\partial_\alpha\tau\partial_\beta\tau = 0. \tag{34}$$

Finally, it is easy to show that the well-known equation[4]

$$\det|\eta^{as}\partial_m\tau\partial^m\tau - \partial^a\tau\partial^s\tau + \eta^{la}\mu\epsilon_l^s\partial_4\tau\partial_4\tau| = 0 \tag{35}$$

is obtained if we consider a Minkowski space ($g_{\alpha\beta} = \eta_{\alpha\beta}$) and a frame moving with the medium.

IV. CAUCHY PROBLEM

We now consider the full system Einstein–Maxwell equations, where the functions σ, $[\epsilon_\alpha^\beta]$, and $[\mu_\alpha^\beta]$ are supposed to be known. The field variables are now $g_{\alpha\beta}$, $f_{\alpha\rho}$, δ, and U^α and the Einstein equation determines $g_{\alpha\beta}$ and U^α.

Given on a three-dimensional manifold V_3 (whose local equation we take as $x^4 = 0$) the field $f_{\alpha\beta}$, we try to determine the values on V_3 of $\partial_4 f_{\alpha\beta}$.

According to the field equation (7a), we have on V_3:

$$\eta^{\lambda 4\nu\rho}\partial_4 f_{\nu\rho} + \varphi^\lambda(\text{C.d.}) = 0, \tag{36}$$

where ϕ^λ depends only on the Cauchy data (C.d.).

This equation gives, for $\lambda = k$, the values of $\partial_4 f_{ij}$ on V_3.

For $\lambda = 4$, (36) gives $\varphi^4(\text{C.d.}) = 0$, which represents a condition that the Cauchy data must satisfy on V_3.

The other group of the field equations (7b) may be decomposed in the system:

$$-M^{j44k}\partial_4 f_{4k} + \psi^j(\text{C.d.}) = \sigma U_\alpha f^{j\alpha} + \delta U^j, \tag{37a}$$

$$\psi^4(\text{C.d.}) = \sigma U_\alpha f^{4\alpha} + \delta U^4, \tag{37b}$$

where ψ_α is also known because of the Cauchy data.

If V_3 is not exceptional ($\det|M^{j44k}| \neq 0$), Eq. (37a) may be solved for $\partial_4 f_{4k}$. [We note incidentally that (37b) determines δ.]

The characteristics of the Maxwell system are the manifold solutions of

$$\det|M^{j44k}| = 0. \tag{38}$$

With the aid of the coordinate transformations defined by

$$x^{4'} = \tau(x^\alpha), \tag{39a}$$

$$x^{a'} = x^a \tag{39b}$$

for which

$$M^{a'4'4's'} = a_n^{a'}a_\beta^{4'}a_\rho^{4'}a_s^{s'}M^{a\beta\rho s} = \partial_\beta\tau\partial_\rho\tau M^{a\beta\rho s}, \tag{40}$$

we get the covariant form (30) obtained in the last section.

This convergence of results assures us that the equation obtained is the good one.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Derivation of the Principle of Compensation of Dangerous Diagrams

DONALD H. KOBE*
Quantum Chemistry Group, Uppsala University, Uppsala, Sweden

(Received 14 September 1966)

The principle of compensation of dangerous diagrams (PCDD) postulated by Bogoliubov to determine the coefficients in the canonical transformation to quasi-particles in superconducting systems is derived from four different criteria (1) the expected number of quasi-particles in the true ground state is a minimum, (2) the one-particle density matrix and the two-particle amplitude determined from the BCS ground state are equated to the true ones, (3) the expectation value of an arbitrary operator is simplified by diagonalizing its quadratic part, and (4) the starting point for the dressing of the quasi-particle is chosen in the most convenient way. The condition obtained for the PCDD is then expressed in terms of quasi-particle Green's functions. The ladder diagrams are eliminated by examining an integral equation for the Green's function describing the creation of two quasi-particles from the vacuum. Finally, the condition obtained here for the PCDD is compared with the condition obtained previously.

1. INTRODUCTION

THE *principle of compensation of dangerous diagrams* (PCDD) postulated by Bogoliubov[1] to determine the coefficients in the canonical transformation to quasi-particles (QP) in superconductivity theory has been of considerable interest since Henley and Wilets[2] showed that the second- and higher-order terms in it were very important for nuclear matter. The PCDD states that the coefficients in the canonical transformation should be determined by setting the sum of all the diagrams describing the creation of a pair of QP from the vacuum equal to zero.[1,3] It had previously been thought that the corrections to the lowest-order diagram were negligible,[4] but Henley and Wilets[2] showed that the energy gap equation sometimes did not even have solutions if the higher-order terms in the PCDD were included.

This result seems somewhat strange at first, because Bogoliubov *et al.*[5] have shown that the BCS model[6] with pairing forces is asymptotically exact in the limit of infinite volume. This problem has recently attracted a great deal of attention because it is one of the few nontrivial examples of an exactly solvable model in quantum field theory.[7] However, since Henley and Wilets[2] were investigating nonpairing forces in general it is no contradiction that they obtained the result that the second- and higher-order terms in the PCDD are important. Tolmachev and Tiablikov[4] calculated the second-order term in the PCDD for a pairing type interaction and pointed out that the term was of higher order in the coupling constant than the order of validity of the model Hamiltonian in the theory of superconductivity. The BCS theory has also been applied to finite nuclei[8] where it is not exact, and therefore it is important to investigate the foundations of the PCDD and the effect of higher-order terms.

The PCDD was originally postulated to remove some terms in the perturbation expansion of the ground state energy that could be divergent.[1,3] Since the expansion itself was not guaranteed to converge,[2] this argument was not very convincing. However, it was shown that the PCDD corresponded to

(I.1) maximizing the overlap between the true and the BCS ground state vector,[9]

(I.2) eliminating two QP states from the true ground state vector,[9]

* Present address: Department of Physics, Northeastern University, Boston, Massachusetts.

[1] N. N. Bogoliubov, Zh. Eksperim. i Teor. Fiz. **34**, 58 (1958) [English transl.: Soviet Phys.—JETP **7**, 41 (1958)]; Nuovo Cimento **7**, 794 (1958); Usp. Fiz. Nauk SSSR **67**, 549 (1959) [English transl.: Soviet Phys.—Usp. **2**, 236 (1959)].

[2] E. M. Henley and L. Wilets, Phys. Rev. **133**, B118 (1964); Phys. Rev. Letters **11**, 326 (1963).

[3] N. N. Bogoliubov, V. V. Tolmachev, and D. V. Shirkov, *A New Method in the Theory of Superconductivity* (Academy of Sciences of the USSR Press, Moscow, 1958) (English transl.: Consultants Bureau, New York, 1959); Fortschr. Physik **6**, 605 (1958).

[4] V. V. Tolmachev and S. V. Tiablikov, Zh. Eksperim. i Teor. Fiz. **34**, 66 (1958) [English transl.: Soviet Phys.—JETP **7**, 46 (1958)].

[5] N. N. Bogoliubov, D. N. Zubarev, and Yu. A. Tserkovnikov, Zh. Eksperim. i Teor. Fiz. **39**, 120 (1960) [English transl.: Soviet Phys.—JETP **12**, 88 (1961)]; N. N. Bogoliubov, Physica **26**, S1 (1960).

[6] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[7] R. Haag, Nuovo Cimento **25**, 287 (1962); H. Ezawa, J. Math. Phys. **5**, 1078 (1964); L. Leplae and H. Umezawa, Nuovo Cimento **33**, 372 (1964); C. T. Chen-Tsai, Chinese J. Phys. (Taiwan) **3**, 22 (1965).

[8] A. Bohr, B. R. Mottelson, and D. Pines, Phys. Rev. **110**, 936 (1958). S. T. Beliaev, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. **31**, No. 11 (1959); L. S. Kisslinger and R. A. Sorenson, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. **32**, No. 9 (1960); M. Baranger, Phys. Rev. **120**, 957 (1960); V. G. Soloviev, Kgl. Danske Videnskab. Selskab, Mat. Fys. Skrifter **1**, No. 11 (1961); C. J. Gallagher, Jr., and V. G. Soloviev, Kgl. Danske Videnskab. Selskab, Mat. Fys. Skrifter **2**, No. 2 (1962).

[9] D. H. Kobe, Phys. Rev. **140**, A825 (1965).

(I.3) diagonalizing the quadratic part of the reaction operator (Brillouin–Brueckner–Bogoliubov condition),[9] and

(I.4) generalizing Brueckner's extension of Hartree–Fock theory (the exact self-consistent-field theory) to QP.[10]

These criteria gave a condition for the PCDD which is called the PCDD(I). It could be expanded by perturbation theory,[4,10] but it was difficult to fit into the framework of QP Green's functions.[11] Using time-dependent perturbation theory it was even difficult to show that ladder diagrams do not contribute to the PCDD.[10] On the other hand, these criteria give some needed physical insight into a previously abstruse and mathematical principle.

It is the purpose of this paper to obtain another formulation of the PCDD which is also based on reasonable physical criteria, but which can be expressed conveniently in terms of QP Green's functions. The criteria which are used here are:

(II.1) the total expected number of QP in the true ground state is a minimum,

(II.2) the one particle density matrix and the two particle amplitude determined from the BCS ground state are equated to the true ones,

(II.3) the expectation value of an arbitrary operator is simplified by diagonalizing its quadratic part,[12] and

(II.4) the starting point for the dressing of the QP is chosen in the most convenient way.

These four criteria all lead to the same condition for the PCDD, which is called the PCDD(II). It is not exactly the same as the PCDD(I), but it is still compatible with the original statement of the PCDD that the sum of all dangerous diagrams vanish.[1,3]

Criterion (II.1) has an especially simple physical interpretation. If the number of QP in the true ground state is a minimum, the QP would be expected to behave more like an ideal gas. The QP interactions would not be as important and thus the free QP model would be expected to be a good approximation. The expansion of the true ground state in terms of the QP states would converge rapidly. Thus this criterion is just as physically appealing as the maximum overlap criterion (I.1).

Perhaps the most important aspect of the new condition for the PCDD is that it is easily formulated in terms of QP Green's functions which have previously been investigated.[11] The equations of motion for the Green's function describing the creation of two QP from the vacuum can be used to obtain a perturbation expansion of the PCDD(II), and even go beyond ordinary perturbation theory. It is also very easy to remove the ladder diagrams from the PCDD(II) by solving an integral equation.

In the next section the Bogoliubov QP is defined and the Hamiltonian transformed to QP operators. The criterion of minimum expected number of QP in the ground state is applied in Sec. 3 to obtain the PCDD(II). In Sec. 4 the criterion of best approximation to the true one- and two-particle density matrices is used to obtain the best coefficients in the canonical transformation. It is shown in Sec. 5 that the diagonalization of the quadratic part of the expectation value of an arbitrary operator also leads to the PCDD(II). The best starting point for the subsequent dressing of the QP is discussed in Sec. 6. Section 7 gives the Green's function formulation of the PCDD(II). The ladder diagrams can be eliminated by using the QP Green's functions as shown in Sec. 8. In Sec. 9 a comparison is made between the PCDD(I) and the PCDD(II). Finally the last section summarizes the different criteria and the advantages of the PCDD(II).

2. BOGOLIUBOV QUASI-PARTICLES

In this section the method of the canonical transformation is reviewed, because the equations are needed later. The necessity for a source term in the Hamiltonian and its implications are discussed. Finally the Hamiltonian is transformed to the QP operators.

The Hamiltonian for a system of particles interacting with two body forces is

$$H = \sum_1 (e_1 - \mu) a_1^\dagger a_1 - \tfrac{1}{4} \sum_{1234} \langle 12| V |34\rangle a_1^\dagger a_2^\dagger a_3 a_4 . \tag{2.1}$$

The operators $a_1^\dagger$ and a_1 are the creation and annihilation operators for fermions with momentum $\mathbf{k}_1$ and spin σ_1 ($\pm$ = up/down), and $(1) = (\mathbf{k}_1, \sigma_1)$, and $(2) = (\mathbf{k}_2, \sigma_2)$, etc. They satisfy the usual fermion anticommutation relations. The matrix element of the potential $\langle 12| V |34\rangle$ is antisymmetric and positive for attractive interactions, e_1 is the kinetic energy, and μ is the chemical potential.

Since we are interested in superconductivity, it is necessary to introduce a source term into the Hamiltonian in order to remove the degeneracy due to the

[10] D. H. Kobe, Quantum Chemistry Group, Uppsala, Sweden, Report No. 137, 1964 (unpublished); Ann. Phys. (N.Y.) **40**, 395 (1966).

[11] D. H. Kobe and W. B. Cheston, Ann. Phys. (N.Y.) **20**, 279 (1962).

[12] The quadratic part of an operator is that part containing a product of two creation or annihilation operators.

appearance of a condensate of bound pairs.[13] The source term is

$$H_s = -\nu \sum_1 (w_1 a_1^\dagger a_{-1}^\dagger + w_1^* a_{-1} a_1), \qquad (2.2)$$

where after the calculation the limit $\nu \to 0$ is taken. The complex function w_1 is arbitrary. The terms in Eq. (2.2) remove the conservation law for the number of particles. The new Hamiltonian

$$H' = H + H_s \qquad (2.3)$$

is now used. This Hamiltonian does not commute with the number operator, and so therefore the number of particles is not a good quantum number. If $|0\rangle$ is the true ground state vector of H', then it is not an eigenstate of the number operator.

In order to treat the problem of superconductivity, Bogoliubov[1,3] and Valatin[14] introduced the canonical transformation to quasi-particles (QP)[15]

$$\alpha_1 = u_1 a_1 + v_{-1} a_{-1}^\dagger . \qquad (2.4)$$

The QP is a particle with probability amplitude u_1 and a hole with probability amplitude v_1. The QP are also fermions if the transformation is to be canonical, so the QP creation and annihilation operators satisfy the usual fermion anticommutation relations. This condition implies that the coefficients satisfy

$$u_1^2 + v_1^2 = 1, \qquad (2.5a)$$

$$u_1 = u_{-1}, \qquad (2.5b)$$

$$v_1 = -v_{-1}. \qquad (2.5c)$$

Equation (2.4) and its Hermitian conjugate can be solved for the particle annihilation operator

$$a_1 = u_1 \alpha_1 - v_{-1} \alpha_{-1}^\dagger . \qquad (2.6)$$

Equation (2.6) expresses the particle annihilation operator in terms of the QP creation and annihilation operators.

The Hamiltonian H' in Eq. (2.3) obviously has a different form than the Hamiltonian H in Eq. (2.1). If Eq. (2.6) and its Hermitian conjugate are substituted into the Hamiltonian H' and the QP operators are put in normal order, the result will be of the same form as if only H had been used. Thus H and H' have the same form after the canonical transformation to QP has been made. In terms of QP the addition of the source term H_s to the original Hamiltonian H leaves it *form invariant*. The limit $\nu \to 0$ can be taken in H' and the original Hamiltonian is recovered. The only purpose of the source term is to give us a hunting license for nonzero two-particle amplitudes $\langle 0|a_1^\dagger a_{-1}^\dagger|0\rangle$. These would definitely be zero if the source terms were not present.[16]

The Hamiltonian H in Eq. (2.1) in terms of the QP operators can be written as[11]

$$H = \sum_{j,k} H_{jk}, \qquad (2.7)$$

where $j + k = 0, 2, 4$ and $(j, k) = (0, 1, 2, 3, 4)$. The term H_{jk} has j QP creation operators and k QP annihilation operators

$$H_{jk} = \sum_{1,2,\cdots,j+k} h_{jk}(1, 2, \cdots, j+k) \times \alpha_1^\dagger \alpha_2^\dagger \cdots \alpha_j^\dagger \alpha_{j+1} \cdots \alpha_{j+k}, \qquad (2.8)$$

where the coefficients h_{jk} can be found in Appendix A of Kobe and Cheston.[11]

The QP vacuum state is needed later as the unperturbed ground state in Sec. IV. The BCS ground state vector is[6]

$$|\mathrm{BCS}\rangle = \prod_j (u_j + v_j a_j^\dagger a_{-j}^\dagger)\,|\mathrm{vac}\rangle, \qquad (2.9)$$

where the product is only over half the total number of states, and $|\mathrm{vac}\rangle$ is the state of no particles. The BCS ground state can be shown to be the vacuum for the QP

$$a_k \alpha\,|\mathrm{BCS}\rangle = 0 \qquad (2.10)$$

for all k, by using Eq. (2.4). Because of this convenient property it is natural to choose the BCS state as the unperturbed ground state.

3. MINIMIZATION OF THE NUMBER OF QUASI-PARTICLES

The problem now is to determine the best choice of the coefficients u_1 and v_1 in Eq. (2.4), which then gives the best Bogoliubov QP. In previous papers[9,10] the criteria (I.1) through (I.4) have been used to obtain the PCDD(I). In this section another criterion is used to give a somewhat different condition for the best Bogoliubov QP which is called the PCDD(II). It is, however, completely compatible with the original statement of the PCDD given by Bogoliubov in terms of dangerous diagrams.[1,3]

The criterion (II.1) which is used now is that *the total expected number of QP in the true ground state is a minimum*. If the true ground state vector is given

[13] N. N. Bogoliubov, Physica **26**, S1 (1960). This procedure is analogous to the addition of an infinitesimal magnetic field in the theory of ferromagnetism. The limit as the field goes to zero is taken after the magnetization has been calculated, which gives a nonzero value below the Curie point. Bogoliubov has called all of these procedures *quasi-averages*. For superconductors the relation between the quasi-average method and the nonzero two-particle amplitudes (or anomalous propagators) was investigated by B. Johansson, Physica **32**, 2164 (1966).

[14] J. G. Valatin, Nuovo Cimento **7**, 843 (1958).

[15] The coefficients u_1 and v_1 are taken as real here. It is shown in the Appendix that this choice can be made without loss of generality. See also Ref. 13.

[16] Y. Nambu, Phys. Rev. **117**, 648 (1960).

by $|0\rangle$ the total expected number of QP in it is the expectation value of the QP number operator

$$n \equiv \sum_j \langle 0| \alpha_j^\dagger \alpha_j |0\rangle = \text{minimum} \tag{3.1}$$

subject to the constraints in Eq. (2.5).

This criterion is a very natural one, since if the number of QP in the true ground state were a minimum, the QP would be expected to behave almost ideally. Their number would be small and interactions would not be as important as if there were many QP present. Thus the free QP model would be expected to be a good approximation. This condition is analogous to reducing the pressure in a volume of gas, so that it behaves more ideally. By minimizing the number of QP, the expansion of the true ground state in terms of zero, two, four, etc. QP states would converge rapidly. This criterion would also be expected to be a good starting point for a treatment of QP interaction effects which is discussed in Sec. 6.

Because of the constraints in Eq. (2.5) it is necessary to introduce the Lagrangian multipliers Λ_j and minimize the function

$$n' = n + \sum_j \Lambda_j (u_j^2 + v_j^2 - 1). \tag{3.2}$$

However, the operator L_k can be defined as[10]

$$L_k = u_k(\partial/\partial v_k) - v_k(\partial/\partial u_k), \tag{3.3}$$

which has the property that

$$L_k \sum_j \Lambda_j (u_j^2 + v_j^2 - 1) = 0. \tag{3.4}$$

Therefore it does not make any difference if L_k is applied to Eq. (3.1) or (3.2).

If n is minimized (or extremized) with respect to the coefficients in the canonical transformation, the result is the condition

$$L_k n = 0 = \text{Re} \langle 0| \alpha_k^\dagger \alpha_{-k}^\dagger |0\rangle, \tag{3.5}$$

which is called the PCDD(II). It gives an equation from which the coefficients u_1 and v_1 can be determined to give the best Bogoliubov QP. In Sec. 9 it is compared with the PCDD(I), and, in Sec. 7 a means of expanding it in terms of QP Green's functions is given.

Equation (3.5) for the extremum is actually a minimum if the second derivatives n'_{uu}, n'_{vv}, and $n'_{uv} \cdot (-u/v)$ are all positive. The second derivatives n'_{uu} and n'_{vv} involve the Lagrangian multipliers Λ_j which must be eliminated by using the two equations obtained by setting the first derivatives equal to zero. When this procedure is carried out the conditions for a minimum are

$$\langle 0| a_k^\dagger a_k |0\rangle > \langle 0| \alpha_k^\dagger \alpha_k |0\rangle, \tag{3.6a}$$

$$\langle 0| a_k a_k^\dagger |0\rangle > \langle 0| \alpha_k^\dagger \alpha_k |0\rangle, \tag{3.6b}$$

$$2u_k v_k \,\text{Re} \langle 0| a_k^\dagger a_{-k}^\dagger |0\rangle > 0. \tag{3.6c}$$

These conditions should be satisfied since the number of QP in the true ground state is expected to be small. The number in the BCS state is, of course, zero. However, it is not yet known whether the extremum is a minimum or a maximum.

The conditions in Eq. (3.6) can be shown to be satisfied by solving Eq. (3.5) for the coefficients u_k and v_k, and then substituting them into Eq. (3.6). Equation (3.5) can be expressed in terms of the coefficients and particle amplitudes by using Eq. (2.4) which gives

$$(u_k^2 - v_k^2) A_k = 2u_k v_k B_k, \tag{3.7}$$

where A_k is defined to be the two-particle amplitude[17]

$$A_k = \text{Re} \langle 0| a_k^\dagger a_{-k}^\dagger |0\rangle \tag{3.8}$$

and B_k is defined as

$$2B_k = \langle 0| a_k a_k^\dagger |0\rangle - \langle 0| a_{-k}^\dagger a_{-k} |0\rangle, \tag{3.9}$$

which can be either positive or negative. Equation (3.7) can be solved with the help of Eq. (2.5) and the results are

$$2u_k v_k = A_k / C_k \tag{3.10}$$

and

$$u_k^2 - v_k^2 = B_k / C_k, \tag{3.11}$$

where

$$C_k = [A_k^2 + B_k^2]^{\frac{1}{2}}. \tag{3.12}$$

The expected number of QP in the true ground state can be expressed in terms of the particle amplitudes by using Eqs. (2.4), and (3.10) through (3.12) to give

$$\langle 0| \alpha_k^\dagger \alpha_k |0\rangle = \tfrac{1}{2} - C_k. \tag{3.13}$$

Since the expected number of QP in the BCS ground state is zero and $C_k = \frac{1}{2}$ for the BCS state, the positive solution has been taken in Eqs. (3.10) through (3.12). If Eqs. (3.13), (3.9), and (3.12) are used in Eq. (3.6) the conditions become

$$B_k < [A_k^2 + B_k^2]^{\frac{1}{2}}, \tag{3.14a}$$

$$-B_k < [A_k^2 + B_k^2]^{\frac{1}{2}}, \tag{3.14b}$$

$$A_k^2 / C_k > 0. \tag{3.14c}$$

These conditions are all satisfied for positive or negative B_k as long as A_k in Eq. (3.8) is not zero.

[17] A transformation on the single-particle orbitals as shown in Eq. (A10) can always be made to make $\langle 0| a_1^\dagger a_{-1}^\dagger |0\rangle$ real.

However, this condition is just the condition that the system is in the superconducting phase. Therefore for superconducting systems, the extremum condition given by the PCDD(II) in Eq. (3.5) makes the expected number of QP in the true ground state a true minimum.

4. BEST DENSITY

The criterion for the best Bogoliubov QP that is now discussed is that the one- and two-particle density matrices obtained with the BCS ground state are the best approximations to the true ones. The single-particle density matrix obtained from the BCS ground state can be set equal to the true single-particle density matrix. The two-particle density matrix obtained from the BCS ground state can be equated to a Gorkov[18] type factorization of the true two-particle density matrix. This criterion is essentially the same as the best density criterion in the independent particle model,[19] where the orbitals are Löwdin's natural spin orbitals.[20]

The matrix element of the single-particle density matrix determined from the BCS ground state is

$$\langle \mathrm{BCS}|\, a_k^\dagger a_k \,|\mathrm{BCS}\rangle = v_k^2 \tag{4.1}$$

if Eqs. (2.6) and (2.10) are used. This expression can be chosen such that it is equal to the true single-particle density matrix[21]

$$v_k^2 = \langle 0|\, a_k^\dagger a_k \,|0\rangle = \langle 0|\, a_{-k}^\dagger a_{-k} \,|0\rangle, \tag{4.2}$$

which can, for example, be calculated by the linearized equation of motion method.[22] This condition has in principle determined the coefficients because of Eq. (2.5). However, we investigate the two-particle density matrix to see if there are any auxiliary conditions which must be imposed.

The matrix element of the two-particle density matrix determined from the BCS ground state is

$$\begin{aligned}&\langle \mathrm{BCS}|\, a_1^\dagger a_2^\dagger a_3 a_4 \,|\mathrm{BCS}\rangle \\ &\quad = (u_1 v_1)(u_4 v_4)\delta_{3-4}\delta_{1-2} + v_1^2 v_2^2(\delta_{23}\delta_{14} - \delta_{13}\delta_{24})\end{aligned} \tag{4.3}$$

from Eqs. (2.6) and (2.10). However, the two-particle amplitude determined from the BCS ground state is

$$\langle \mathrm{BCS}|\, a_k^\dagger a_{-k}^\dagger \,|\mathrm{BCS}\rangle = u_k v_k . \tag{4.4}$$

The coefficients can be chosen such that $u_k v_k$ is equal to the true two-particle amplitude[17,23]

$$u_k v_k = \langle 0|\, a_k^\dagger a_{-k}^\dagger \,|0\rangle, \tag{4.5}$$

which is not zero in the superconducting state[13] because of the source term in Eq. (2.3). The two-particle amplitude and the density matrix can be shown to satisfy two coupled equations by the linearized equation of motion method.[22]

If Eqs. (4.2) and (4.5) are substituted into Eq. (4.3), we obtain a Gorkov[18] type factorization of the two-particle density matrix

$$\begin{aligned}&\langle 0|\, a_1^\dagger a_2^\dagger a_3 a_4 \,|0\rangle \\ &\quad \cong \langle \mathrm{BCS}|\, a_1^\dagger a_2^\dagger a_3 a_4 \,|\mathrm{BCS}\rangle \\ &\quad \cong \langle 0|\, a_1^\dagger a_{-1}^\dagger \,|0\rangle\langle 0|\, a_{-4} a_4 \,|0\rangle\, \delta_{3-4}\delta_{1-2} \\ &\qquad + \langle 0|\, a_1^\dagger a_1 \,|0\rangle\langle 0|\, a_2^\dagger a_2 \,|0\rangle\, (\delta_{23}\delta_{14} - \delta_{13}\delta_{24}).\end{aligned} \tag{4.6}$$

Equation (4.6) shows why it is not surprising that Gorkov[18] obtained the BCS theory[6] with this type of ansatz. The Green's function equations that he used are analogous to the linearized equation of motion method used by Pines.[22]

If the ansatz in Eq. (4.6) is used in the expectation value of the Hamiltonian, Eqs. (4.5) and (4.2) substituted into it, and the expression is minimized with respect to the coefficients, the BCS result[6] is obtained. This procedure corresponds to Valatin's[14] minimization of H_{00} in Eq. (2.7). All of these methods correspond to the PCDD in lowest order.

Equation (3.7) can easily be obtained by multiplying Eq. (4.5) by $(u_k^2 - v_k^2)$ and making use of Eqs. (4.2) and (2.5). If, however, Eq. (2.6) is substituted into Eq. (3.7) the condition

$$\mathrm{Re}\,\langle 0|\, \alpha_k^\dagger \alpha_{-k}^\dagger \,|0\rangle = 0 \tag{4.7}$$

is obtained which is the same as Eq. (3.5) for the PCDD(II). Thus the criterion of best approximation to the one- and two-particle density matrices gives the same condition as the minimization of the number of QP in the true ground state.[23a]

5. SIMPLIFICATION OF EXPECTATION VALUES

In his original paper on the canonical transformation applied to boson systems, Bogoliubov[24] neglected the QP interaction terms and merely diagonalized the quadratic part of the Hamiltonian which lead to only

[18] L. P. Gorkov, Zh. Eksperim. i Teor. Fiz. **34**, 735 (1958) [English transl.: Soviet Phys.—JETP **7**, 505 (1958)]; S. T. Beliaev, Physica **26**, S181 (1960); A. Zawadowszki and G. Pócsik, Phys. Letters **7**, 173 (1963); Nuovo Cimento **32**, 1110 (1964).

[19] W. Kutzelnigg and V. H. Smith, Jr., Quantum Chemistry Group, Uppsala, Sweden, Report No. 130, 1964 (unpublished); J. Chem. Phys. **41**, 896 (1964).

[20] P. O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[21] The number of particles in the state k is the same as the number in the state $-k$ by inversion symmetry.

[22] D. Pines, *The Many-Body Problem* (W. A. Benjamin, Inc., New York, 1961), pp. 94–96.

[23] If the coefficients are complex Eq. (A6) would be obtained from this criterion too.

[23a] V. H. Smith, Jr., Nuovo Cimento **48**, 443 (1967), has shown that the PCDD(I) is equivalent to setting

$$v_k^2 = \langle 0|\, a_k^\dagger a_k \,|\mathrm{BCS}\rangle,$$

which is the transition density matrix.

[24] N. N. Bogoliubov, J.Phys. (U.S.S.R.) **11**, 23 (1947).

the compensation of the lowest-order dangerous diagram. It was shown previously[9] that diagonalizing the quadratic part of the reaction operator (or t matrix) leads to the PCDD(I). No matter how important these two operators are for the energy, they should not be overemphasized, since there are many other important operators. One possible criterion for the best Bogoliubov QP would be to diagonalize the quadratic part of the operator that is most important to the particular problem. However, this criterion would make the coefficients dependent on which operator was chosen, which would not be a desirable feature. Since it is really expectation values which are of interest, the criterion of diagonalizing the quadratic part of the *expectation value* of an arbitrary operator can be used. In this way the expectation values of arbitrary operators can be simplified.

An arbitrary one-, two-, three-, or many-particle operator Q can easily be expressed in second quantization.[25] When it is transformed to the QP creation and annihilation operators in Eq. (2.4) and put in normal order, it can be written in a form similar to the Hamiltonian in Eq. (2.7). The true expectation value of the operator is

$$\langle 0|\, Q\, |0\rangle = Q_{00} + \langle 0|\, Q_{11}\, |0\rangle + 2\,\mathrm{Re}\,\langle 0|\, Q_{20}\, |0\rangle + \cdots, \tag{5.1}$$

where

$$\langle 0|\, Q_{11}\, |0\rangle = \sum_k q_{11}(k, k)\,\langle 0|\, \alpha_k^\dagger \alpha_k\, |0\rangle \tag{5.2}$$

and

$$\mathrm{Re}\,\langle 0|\, Q_{20}\, |0\rangle = \sum_k q_{20}(k, -k)\,\mathrm{Re}\,\langle 0|\, \alpha_k^\dagger \alpha_{-k}^\dagger\, |0\rangle. \tag{5.3}$$

The three dots in Eq. (5.1) represent the expectation values of the parts containing more than two QP creation and annihilation operators. The coefficients q_{ij} contain the matrix elements of the operator Q and the coefficients in the canonical transformation. It is assumed that q_{20} is real in Eq. (5.3).

Since the coefficients can be chosen arbitrarily, the form of Eq. (5.1) will be simplified if the third term on the right vanishes. From Eq. (5.3) it can be seent hat it will vanish if

$$\mathrm{Re}\,\langle 0|\, \alpha_k^\dagger \alpha_{-k}^\dagger\, |0\rangle = 0, \tag{5.4}$$

which is the same condition as Eq. (3.5) for the PCDD(II). It will also vanish if $q_{20}(k, -k)$ vanishes, but this condition gives coefficients that are dependent on the matrix elements of the operator, which is not desirable. Equation (5.4) for the coefficients is also advantageous in that Eq. (5.2) will be "small" since the number of QP in the state k in the exact ground state will be minimized by the condition.

Thus the condition of diagonalizing the quadratic part of the expectation value of an arbitrary operator also gives the PCDD(II). The criterion (I.3) which stipulates that the quadratic part of the reaction operator should be diagonal gives the PCDD(I). In this section it is the *expectation value* of the quadratic part of an arbitrary operator which is diagonalized, so that there is no contradiction between these two criteria.

6. BEST STARTING POINT FOR THE DRESSING OF THE QUASI-PARTICLE

Because of the QP interactions, the free QP becomes dressed with a cloud of virtual collective excitations.[26,27] The virtual emission and absorption of these collective excitations causes the energy of the bare QP to become dressed. This dressed energy is a better description of the QP since it takes into account the effect of QP interactions that are ignored in the free or bare QP model.

Another criterion that can be applied to determine the best choice of coefficients in the canonical transformation is that the starting point for the treatment of the QP dressing is chosen in the most convenient way.[28] This criterion is not unique, but it is shown that a reasonable choice will lead to the PCDD(II) obtained from the other criteria.

The concept of dressing can best be described in terms of Green's functions for the QP. In a previous paper[11] the set of coupled integral equations satisfied by the Fourier transform of the QP Green's functions

$$\mathcal{G}_{nm}(1, 2, \cdots, n, n+1, \cdots, n+m) = i\,\langle 0|\, T\{\alpha_1 \cdots \alpha_n \alpha_{n+1}^\dagger \cdots \alpha_{n+m}^\dagger\}\, |0\rangle \tag{6.1}$$

were investigated. The creation and annihilation operators in Eq. (6.1) are in the Heisenberg picture. The variables are defined as $(1) = (\mathbf{k}_1, \sigma_1, t_1)$, $(2) = (\mathbf{k}_2, \sigma_2, t_2)$, etc. The operator T is the time ordering operator that puts the largest times on the left and the smallest on the right with a plus sign for an even and a minus sign for an odd permutation of the original order.

The equation of motion for the single QP propagator can be obtained and is shown in Fig. 1. The QP self-energy was obtained in a previous paper[26] where the vertex in Fig.1(g), $h_{20}(k, -k)$, was set equal

[25] D. H. Kobe, Proc. Phys. Soc. London **88**, 9 (1966); E. R. Pike, *ibid.* **81**, 427 (1963); see also D. H. Kobe, Am. J. Phys. **34**, 1150 (1966).

[26] D. H. Kobe, Ann. Phys. (N.Y.) **28**, 400 (1964).

[27] J. R. Schrieffer, Nucl. Phys. **35**, 363 (1962).

[28] For the independent particle model the criterion of the best starting point for a subsequent treatment of correlation was proposed in Ref. 19.

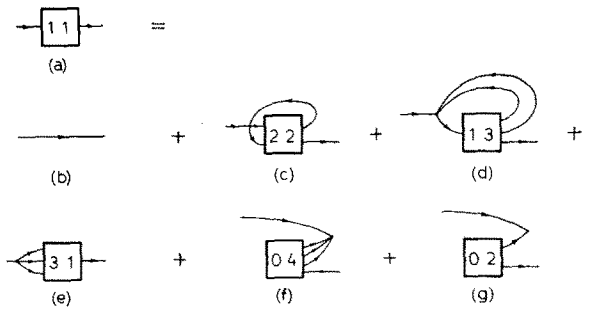


FIG. 1. The equation of motion for the single quasi-particle propagator.

to zero so that the term vanished. This procedure is just the PCDD in lowest order which is not adequate. The self energy was determined from the remaining terms in Fig. 1. However, higher-order dangerous diagrams would arise from some of the remaining terms in Fig. 1. For example, when Fig. 1(f) is expanded it results in a term that has the same form as Fig. 1(g) except that the vertex function is a second-order dangerous diagram and would not vanish if only $h_{20}(k, -k)$ were zero. It is thus desirable to find another condition which would result in the vanishing of the sum of all terms involving G_{02} or G_{20} in the expansion of the single quasi-particle propagator.

The concept of dressing the QP is really only an approximation which could be expected to be valid in the case of weak QP interaction. The true single QP propagator has poles at all the excitation energies of the system as the true spectral representation shows[29]

$$G_{11}(1, 2) = (-2\pi)^{-1}\delta(\omega_1 - \omega_2) \times \sum_s \left\{ \frac{\langle 0|\, \alpha_1\, |s\rangle\langle s|\, \alpha_2^\dagger\, |0\rangle}{\omega_1 - \omega_s + i0} + \frac{\langle 0|\, \alpha_2^\dagger\, |s\rangle\langle s|\, \alpha_1\, |0\rangle}{\omega_1 + \omega_s - i0} \right\}, \quad (6.2)$$

where $|s\rangle$ is a true eigenstate of the Hamiltonian H' in Eq. (2.3) and ω_s is its excitation energy. The arguments of $G_{11}(1, 2)$ contain the variables $(1) = (\mathbf{k}_1, \sigma_1, \omega_1)$, etc. The first term in Eq. (6.2) is the retarded part and the second term is the advanced part.

In the absence of QP interactions, the true ground state reduces to the BCS ground state $|\mathrm{BCS}\rangle$. The complete set of states reduce to the zero, one, two, etc. free QP states. Thus the only state that contributes to the sum over states in the first term in Eq. (6.1) is the state with one QP $|s\rangle = \alpha_1^\dagger |\mathrm{BCS}\rangle$ with an excitation energy $E_1 \equiv h_{11}(1, 1)$. The advanced part of Eq. (6.2) vanishes because of Eq. (2.10). Thus, in the absence of QP interactions, Eq. (6.2) reduces to the free QP propagator

$$G^0(1) = (-2\pi)^{-1}(\omega_1 - E_1 + i0)^{-1} \quad (6.3)$$

after integration (summation) over $\omega_2(\mathbf{k}_2, \sigma_2)$.

If the QP interaction is turned on very weakly, the form of Eq. (6.3) would still be expected to be a good approximation to the retarded part of the single QP propagator. The predominant contribution to the sum in Eq. (6.2) for the retarded part would come from the state with one QP created from the true ground state $|s\rangle = \alpha_1^\dagger |0\rangle$ with the excitation energy ξ_1. For the advanced part of Eq. (6.2) the predominant contribution would come from the state $|s\rangle = \alpha_{-1}^\dagger |0\rangle$ with the same excitation energy. Thus a good approximation to Eq. (6.2) would be[30]

$$G_{11}(1) = (-2\pi)^{-1}\left\{ \frac{\langle 0|\, \alpha_1\alpha_1^\dagger\, |0\rangle\langle 0|\, \alpha_1\alpha_1^\dagger\, |0\rangle}{\omega_1 - \xi_1 + i0} + \frac{\langle 0|\, \alpha_1^\dagger\alpha_{-1}^\dagger\, |0\rangle\langle 0|\, \alpha_{-1}\alpha_1\, |0\rangle}{\omega_1 + \xi_1 - i0} \right\} \quad (6.4)$$

after integration (summation) over $\omega_2(\mathbf{k}_2, \sigma_2)$. The energy ξ_1 is the dressed energy and can be determined in principle from the Dyson equation.[31] The advanced part of Eq. (6.4) is inconvenient, so it would be convenient to choose the coefficients such that it vanished. This choice is now shown to lead to the vanishing of G_{02} also.

After this discussion of the approximations used in G_{11} for the dressing of the QP, the same approximations can be made in Fig. 1(g) for the two QP creation propagator G_{02}. The functions G_{11} and G_{02} are coupled to each other and thus the same type of approximations should be used for both for the sake of consistency. The spectral representation of G_{02} is[29]

$$G_{02}(1, 2) = (-2\pi)^{-1}\delta(\omega_1 + \omega_2) \times \sum_s \left\{ \frac{\langle 0|\, \alpha_1^\dagger\, |s\rangle\langle s|\, \alpha_2^\dagger\, |0\rangle}{\omega_1 - \omega_s + i0} + \frac{\langle 0|\, \alpha_2^\dagger\, |s\rangle\langle s|\, \alpha_1^\dagger\, |0\rangle}{\omega_1 + \omega_s - i0} \right\}. \quad (6.5)$$

The state which would be expected to predominate for weak QP interactions in Eq. (6.5) is $|s\rangle = \alpha_{-1}^\dagger |0\rangle$ in the first term and $|s\rangle = \alpha_1^\dagger |0\rangle$ in the second. Both these states would have the same excitation energy ξ_1. The propagator G_{02} would then take the form

$$G_{02}(1) = (-2\pi)^{-1}\left\{ \frac{\langle 0|\, \alpha_1^\dagger\alpha_{-1}^\dagger\, |0\rangle\langle 0|\, \alpha_{-1}\alpha_{-1}^\dagger\, |0\rangle}{\omega_1 - \xi_1 + i0} + \frac{\langle 0|\, \alpha_{-1}^\dagger\alpha_1^\dagger\, |0\rangle\langle 0|\, \alpha_1\alpha_1^\dagger\, |0\rangle}{\omega_1 + \xi_1 - i0} \right\} \quad (6.6)$$

after integration (summation) over $\omega_2(\mathbf{k}_2, \sigma_2)$.

[29] D. H. Kobe, Ann. Phys. (N.Y.) 19, 448 (1962).

[30] See, e.g., P. Nozières, *Theory of Interacting Fermi Systems* (W. A. Benjamin, Inc., New York, 1964), Chap. 3, for a discussion of the concept of dressing.

[31] F. J. Dyson, Phys. Rev. 75, 1736 (1949).

The two QP creation propagator G_{02} and the advanced part of the single particle propagator in Eq. (6.4) will all vanish if the two QP amplitude in the numerator vanishes. Thus we can use the condition

$$\mathrm{Re}\,\langle 0|\,\alpha_k^\dagger \alpha_{-k}^\dagger\,|0\rangle = 0 \tag{6.7}$$

to determine the coefficients in the transformation. The imaginary part of the two QP amplitude is also zero because of Eq. (A8). Thus the term in Fig. 1(g) and other similar terms involving G_{02} and G_{20} are zero. From Eq. (6.4) the single-particle propagator is thus

$$G_{11}(1) = (-2\pi)^{-1}(\omega_1 - \xi_1 + i0)^{-1}Z_1, \tag{6.8}$$

where Z_1 is the numerator of the retarded part in Eq. (6.4). Equation (6.8) is in exactly the same form as the bare propagator in Eq. (6.3), and can easily be determined from the Dyson equation.[31] If the advanced part were not zero, the Dyson equation would be much more difficult to solve. Thus the criterion of the most convenient starting point for the dressing of the QP is also seen to lead to the PCDD(II) in Eq. (6.7).

7. QUASI-PARTICLE GREEN'S FUNCTION FORMULATION

In order to obtain an expansion of PCDD(II), it is convenient to express it in terms of the QP Green's functions. Then the QP Green's functions equations of motion[11] can be used to expand the PCDD(II) in terms of QP interactions. An ordinary perturbation expansion can be obtained,[10] but the Green's function method enables an infinite subset of graphs to be summed, and intermediate propagators to be dressed. Thus by using Green's functions one can go beyond the limitations imposed by ordinary perturbation theory. In the next section the Green's function method is used to eliminate the ladder diagrams.

From the spectral representation of the two QP creation propagator G_{02} in Eq. (6.5), it can be shown that the two QP amplitude is

$$g_{02}(1, 2) \equiv -\mathrm{Re}\, i \int_{-\infty}^{\infty} d\omega_1\, d\omega_2 G_{02}(1, 2)$$

$$= \mathrm{Re}\,\langle 0|\,\alpha_1^\dagger \alpha_2^\dagger\,|0\rangle. \tag{7.1}$$

The two QP annihilation propagator G_{20} has a spectral representation similar to Eq. (6.5).[29] It is related to the two QP amplitude by an equation similar to Eq. (7.1)

$$g_{20}(1, 2) \equiv -\mathrm{Re}\, i \int_{-\infty}^{\infty} d\omega_1\, d\omega_2 G_{20}(1, 2)$$

$$= \mathrm{Re}\,\langle 0|\,\alpha_1 \alpha_2\,|0\rangle. \tag{7.2}$$

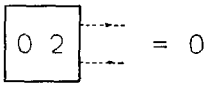


FIG. 2. The principle of compensation of dangerous diagrams in diagrammatic form. The dashed lines represent integration over the energies.

From Eqs. (7.1) and (7.2) it follows that

$$g_{02}(1, 2) = g_{20}(2, 1). \tag{7.3}$$

Because of the QP anticommutation relations, we also have the result

$$g_{02}(1, 2) = -g_{02}(2, 1). \tag{7.4}$$

From the PCDD(II) in Eqs. (3.5), (4.7), (5.4), and (6.7) the condition for the best QP is

$$g_{02}(1, -1) = 0, \tag{7.5}$$

which is shown graphically in Fig. 2.

The Green's function G_{02} satisfies the following equation of motion[11]:

$$G_{02}(1, 2) \tag{7.6a}$$

$$= \sum{}' 4\pi h'_{02}(11')G^0(1)G_{11}(1'2) \tag{7.6b}$$

$$+ \sum{}' 4\pi h'_{22}(1'2'13')G^0(1)G_{13}(3', 1'2') \tag{7.6c}$$

$$+ \sum{}' -2\pi h'_{31}(1'2'3'1)G^0(1)G_{04}(1'2'3'2) \tag{7.6d}$$

$$+ \sum{}' -6\pi h'_{13}(1'12'3')G^0(1)G_{22}(2'3', 1'2) \tag{7.6e}$$

$$+ \sum{}' 8\pi h'_{04}(13'2'1')G^0(1)G_{31}(3'2'1', 2). \tag{7.6f}$$

This equation is shown graphically in Fig. 3. It represents an "expansion" of the Green's function in terms of the last interaction, the outgoing QP encounters. The free QP propagator $G^0(1)$ is defined in Eq. (6.3). The prime on the h_{jk} functions defined in Eq. (2.8) means that there is also a delta function for conservation of energy at each vertex. The prime on the sum means to sum (integrate) over the primed momenta and spins (frequencies).

In order to obtain the PCDD(II), it is only necessary to integrate Eq. (7.6) as shown in Eq. (7.1) and use Eq. (7.5). The lowest-order term in the PCDD(II) can be obtained by neglecting all the QP interaction terms except (b) in Eq. (7.6) and Fig. 3. Then the PCDD(II) in Eq. (7.5) gives

$$h_{02}(k, -k) = 0, \tag{7.7}$$

which is just the lowest-order PCDD(II) obtained by diagonalizing the quadratic part of the Hamiltonian

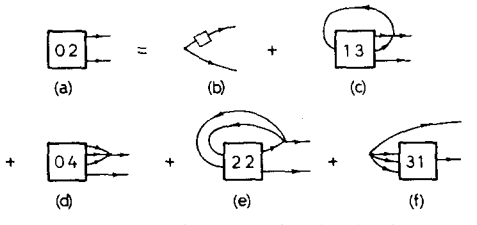


FIG. 3. The equation of motion for the Green's function describing the creation of two quasi-particles from the vacuum.

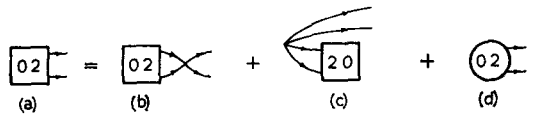


FIG. 4. The integral equation which sums the ladder diagrams. The outgoing lines are dressed.

in Eq. (2.7) and which is equivalent to the BCS theory. By substituting the equations of motion for the higher-order Green's functions in Eq. (7.6), the PCDD(II) can be obtained to any arbitrary order in perturbation theory. In particular, the second-order correction to the PCDD can be obtained from the term (f) in Eq. (7.6) and Fig. 3 if the three QP going into the box are annihilated and one is created.

8. ELIMINATION OF LADDER DIAGRAMS

By using the Green's function method of the last section, it is possible to show that ladder diagrams do not contribute to the PCDD(II). Tolmachev and Tiablikov[4] first pointed out that it is not necessary to consider the diagrams that have a ladder part on their outgoing lines. They used the PCDD(I), but it was shown that there is still some difficulty with the ladder diagrams.[10]

If the equations of motion for the higher-order Green's functions are substituted into Fig. 3 for G_{02}, then Fig. 4 is obtained. A similar equation can be obtained for G_{20}. The term in Fig. 4(d) is the sum of all nonladder type diagrams shown in Fig. 5. The iteration of the two equations for G_{02} and G_{20} results in the sum of all ladder diagrams being attached to the nonladder diagrams.[32]

In Fig. 4 the term (b) is obtained when the QP going into Fig. 3(c) passes through with only self-interactions. Likewise the term in Fig. 4(c) is obtained when one of the QP going into Fig. 3(f) passes through with only self-interactions. In Fig. 5 there would be some ladder diagrams with self-interactions on the outgoing lines, but these have been included in Fig. 4(b) and (c) by dressing the outgoing lines. Thus in Fig. 5 only the pure nonladder diagrams are considered.

Equation (7.6) can be written in the form of Fig. 4 by substituting the equations of motion for higher-order Green's functions[11] into it, which gives

$$G_{02}(1,2) = \sum{}' -4\pi i h'_{22}(1'2'12)G_1(1)G_1(2)G_{02}(1',2') + \sum{}' -24\pi i h'_{04}(122'1')G_1(1)G_1(2)G_{20}(2'1') + F_{02}(1,2) \quad (8.1)$$

The dressed single QP propagator $G_1(1)$ is defined in Eq. (6.4). If Eq. (8.1) is substituted into Eq. (7.2) and assuming the matrix elements of the potential and the two QP amplitude[17] are real, the result is

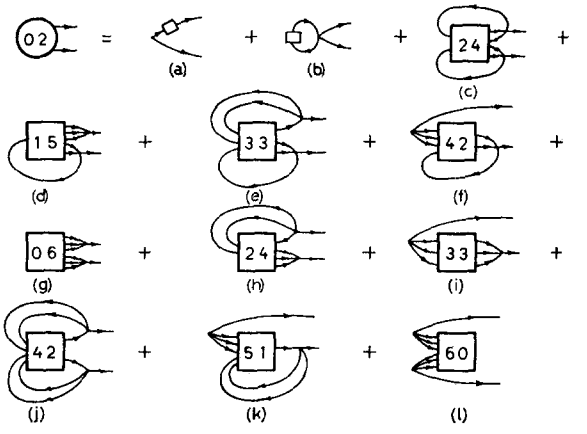


FIG. 5. The sum of all the nonladder diagrams. Ladder diagrams with self-energy interactions on the outgoing lines are not included.

$$g_{02}(1,2) = \sum{}' 4\pi h_{22}(1'2'12)g_1(12)g_{02}(1'2') + \sum{}' 24\pi h_{04}(121'2')g_1(12)g_{02}(2'1') + f_{02}(12). \quad (8.2)$$

Equations (7.3) and (7.4) have been used to obtain Eq. (8.2). The function g_1 is defined as

$$g_1(12) = (2\pi)^{-1}(\xi_1 + \xi_2)^{-1}[Z_1Z_2 + g_{02}^2(1,-1)\,g_{02}^2(2,-2)] \quad (8.3)$$

and the function f_{02} is defined as

$$f_{02}(1,2) = -\mathrm{Re}\, i \int_{-\infty}^{\infty}\!\!\int d\omega_1\, d\omega_2 F_{02}(1,2). \quad (8.4)$$

Equation (8.2) shows that if Eq. (7.5) for the PCDD(II) is satisfied, then $f_{02}(1,-1) = 0$ also. Conversely, if

$$f_{02}(1,-1) = 0 \quad (8.5)$$

then Eq. (7.5) for the PCDD(II) is also satisfied since the sum of the ladder graphs *after* the canonical transformation converges.[33] Equations (8.4) and (8.5) can be taken as a reformulation of the PCDD(II). In order to determine the coefficients in the canonical transformation the *sum of all the nonladder dangerous diagrams should be set equal to zero.*[34]

9. RELATION BETWEEN THE TWO CONDITIONS

In two previous papers[9,10] the criteria (I.1) through (I.4) in the Introduction were used to obtain the

[32] A typical ladder diagram is shown in Fig. 1 of D. H. Kobe, Ann. Phys. (N.Y.) **25**, 121 (1963).

[33] R. Balian and M. L. Mehta, Nucl. Phys. **31**, 587 (1962); B. Johansson (Ref. 13) has shown that the divergence of the ladder diagrams for the *particles* in the superconducting phase [D. J. Thouless, Ann. Phys. (N.Y.) **10**, 553 (1960)] is necessary when ν in the source term in Eq. (2.2) goes to zero to obtain a finite two-particle amplitude in Eq. (3.8).

[34] For finite systems it is necessary to re-examine the role of ladder diagrams in the PCDD according to E. M. Henley, R. C. Kennedy, and L. Wilets, Phys. Rev. **135**, A1172 (1964).

condition for the determination of the coefficients for the best Bogoliubov QP. These criteria gave what is called here the PCDD(I)

$$\text{I:}\quad \mathrm{Re}\,\langle 0|\,\alpha_k^\dagger\alpha_{-k}^\dagger\,|\mathrm{BCS}\rangle = 0, \tag{9.1}$$

where $|\mathrm{BCS}\rangle$ is the BCS ground state. Especially the maximum overlap criterion (I.1) gives some physical insight into the nature of the PCDD. It is this condition that was used by Tolmachev and Tiablikov[4] in calculating the second-order term in the PCDD.

The criteria (II.1) through (II.4) developed in this paper give the condition called the PCDD(II)

$$\text{II:}\quad \mathrm{Re}\,\langle 0|\,\alpha_k^\dagger\alpha_{-k}^\dagger\,|0\rangle = 0 \tag{9.2}$$

given in Secs. 3 through 6. These criteria are equally reasonable, and the PCDD(II) can be used as an alternative condition for determining the coefficients. It is especially convenient to use because QP Green's functions can be used to obtain an expansion for it.

The PCDD(II) reduces to the PCDD(I) if the true ground state $|0\rangle$ is replaced with the BCS ground state on the right side of Eq. (9.2). The difference between the PCDD(II) and the PCDD(I) can be seen by expanding the true ground state wave vector $|0\rangle$ in terms of zero, two, four, etc. QP states

$$\begin{aligned}|0\rangle = {} & c\,|\mathrm{BCS}\rangle \\ & + \sum_k c_k\alpha_k^\dagger\alpha_{-k}^\dagger\,|\mathrm{BCS}\rangle \\ & + \sum_{k>l} c_{kl}\alpha_k^\dagger\alpha_{-k}^\dagger\alpha_l^\dagger\alpha_{-l}^\dagger\,|\mathrm{BCS}\rangle \\ & + \sum_{k>l>m} c_{klm}\alpha_k^\dagger\alpha_{-k}^\dagger\alpha_l^\dagger\alpha_{-l}^\dagger\alpha_m^\dagger\alpha_{-m}^\dagger\,|\mathrm{BCS}\rangle + \cdots,\end{aligned} \tag{9.3}$$

where the sums are only over half the total number of states. The PCDD(I) given in Eq. (9.1) corresponds to

$$\begin{aligned} c &= \max, \\ c_k &= 0 \text{ for all } k,\end{aligned} \tag{9.4}$$

because of the maximum overlap criterion (I.1). However, the PCDD(II) corresponds to

$$\begin{aligned} n &= 2\sum_k |c_k|^2 + 4\sum_{k>l}|c_{kl}|^2 + 6\sum_{k>l>m}|c_{klm}|^2 + \cdots \\ &= \min\end{aligned} \tag{9.5}$$

from the minimization of the number of QP in Eq. (3.1). In order for Eq. (9.5) to be a minimum the coefficients corresponding to a large number of QP must be small. Thus the convergence of the expansion in Eq. (9.3) must be rapid.

From Eqs. (9.4) and (9.5) it can be seen that in general the two criteria are different, but that both are reasonable. Both the conditions in Eqs. (9.1) and (9.2) can be expanded by perturbation theory and the same general type of diagrams is obtained. However, in the perturbation expansion of Eq. (9.1) it is not possible to eliminate the ladder diagrams exactly.[10] For Eq. (9.2) Green's functions and their equations of motion[11] can be used. This method is very convenient and fits in very nicely with the self-energy of the QP.[26] In using Green's functions, it is possible to sum an infinite subset of graphs and easily go beyond ordinary perturbation theory. This advantage strongly recommends the PCDD(II).

It is not surprising that different criteria give somewhat different formulations of the PCDD. In the independent particle model[19] different criteria also give different orbitals. The condition obtained for orbitals from the maximum overlap principle is different from the best density criterion, and the two cannot in general be simultaneously satisfied.

10. CONCLUSION

Of the four criteria developed in this paper for determining the coefficients in the canonical transformation, the criterion (II.1) of minimizing the number of QP in the true ground state has the most appealing physical significance. It is analogous to reducing the pressure in a container of gas by removing molecules so that the gas will behave more ideally. If there are very few QP expected in the true ground state, the probability of finding many QP is small. Thus the expansion of the true ground state in terms of QP states would converge rapidly.

The criterion of best density (II.2) is a very reasonable one, since it would give the best approximation to the expectation values of one- and two-particle operators. It also sheds some light on the nature of the Gorkov[18] type of factorization of the two-particle density matrix.

The third criterion of the simplification of the expectation value of an arbitrary operator (II.3) is somewhat arbitrary. On the other hand, it is no more so than the original choice of coefficients as the ones diagonalizing the quadratic part of the Hamiltonian. The Hamiltonian is just one of many operators, so that it should not be overemphasized. It is significant, however, that the condition obtained from this criterion is also the one obtained by the other two criteria.

The fourth criterion, the best starting point for the dressing of the QP, is of interest in that all the terms involving a G_{02} or a G_{20} in the equation of motion for the single QP propagator are eliminated. In the original treatment of the QP self-energy,[26] only the compensation of the lowest-order dangerous diagram was used. In higher orders, there would be G_{02} or G_{20} terms with higher-order dangerous diagrams attached to them. With the PCDD(II) these terms would also be zero. The self-energy of the QP obtained previously does not have to be modified.

All of the criteria used here give the PCDD(II) instead of the PCDD(I) obtained previously. These two conditions for the PCDD were compared in the last section. The question as to which is best is probably a matter of taste. If ordinary perturbation theory is used, the simplest condition is the PCDD(I), since it involves the true ground state only once. However, there is some difficulty in removing the ladder diagrams exactly.

On the other hand, if the QP Green's functions are used then it is essential to use the PCDD(II). The ladder diagrams can easily be removed. A perturbation expansion can be obtained for the PCDD(II) in which dressed energies can easily be introduced. Then PCDD(II) can be used in connection with the QP self-energy to obtain equations coupling the energy gap to the single particle energy.[35] Thus the advantage of being able to work with the QP Green's function recommends the PCDD(II).

The eight criteria given in this and previous papers[9,10] have provided some needed physical insight into a previously abstruse and mathematical principle. The criterion of minimum number of QP in the true ground state stands alongside the maximum overlap criterion as giving an intuitive feeling for obtaining the best Bogoliubov quasi-particle.

ACKNOWLEDGMENTS

The author would like to thank Dr. R. Mattuck and Fil. lic. Börje Johansson for discussing the method of quasi-averages. This work was originally motivated by some discussions with Dr. V. H. Smith, Jr., on his work on the independent particle model with Dr. Kutzelnigg, and the author would like to thank him for many interesting conversations, especially about Sec. 4. His appreciation is also extended to Dr. F. Sasaki for many stimulating discussions.

Finally, the author would like to express sincere gratitude to Professor Per-Olov Löwdin for his encouragement and interest in this work, and his warm hospitality during the author's stay at the Quantum Chemistry Group in Uppsala.

The work reported in this paper has been sponsored in part by the King Gustaf VI Adlof's 70-Years Fund for Swedish Culture, Knut and Alice Wallenberg's Foundation, and in part by the Aerospace Research Laboratories, OAR, through the European Office of Aerospace Research (OAR), U.S. Air Force.

APPENDIX. COMPLEX COEFFICIENTS

Throughout this paper it has been assumed that the coefficients in the canonical transformation in Eq. (2.4) were real. However, this assumption can be made without loss of generality as is shown here.

The coefficients in Eq. (2.4) are now assumed to be arbitrary complex numbers

$$u_1 = |u_1|\, e^{i\chi_1}, \tag{A1}$$

$$v_1 = \epsilon_1 |v_1|\, e^{i\varphi_1}, \tag{A2}$$

where

$$\epsilon_1 = -\epsilon_{-1} = 1. \tag{A3}$$

The condition that the QP be fermions modifies only Eq. (2.5a) by replacing the coefficients with their absolute values

$$|u_1|^2 + |v_1|^2 = 1. \tag{A4}$$

The two-particle amplitude is in general complex and can be written as

$$A_1 = \langle 0|\, a_{-1}a_1\, |0\rangle = \epsilon_1 |A_1|\, e^{i\theta_1}. \tag{A5}$$

If the number of QP in Eq. (3.1) is minimized with respect to the coefficients in the transformation, the condition for the best QP is obtained. This means that Eq. (3.1) with the constraint in Eq. (A4) must be minimized with respect to $|u_1|$, $|v_1|$, χ_1, and φ_1. The minimization with respect to χ_1 and φ_1 gives the condition

$$\chi_1 - \varphi_1 + \theta_1 = 2\pi m, \tag{A6}$$

where $m = 0, \pm 1, \pm 2, \cdots$. This condition can also be shown to give a minimum by calculating the second derivatives.

The minimization with respect to the absolute value of the coefficients $|u_1|$ and $\epsilon_1 |v_1|$ gives the condition

$$s_1 = (|u_1|^2 - |v_1|^2)\epsilon_1 |A_1| - \epsilon_1 2\, |u_1 v_1|\, B_1 = 0. \tag{A7}$$

However, if the two QP amplitude is calculated and use is made of Eq. (A6) the result is

$$\langle 0|\, \alpha_1^\dagger \alpha_{-1}^\dagger\, |0\rangle = s_1 e^{-i(\varphi_1+\chi_1)}. \tag{A8}$$

Therefore the condition in Eq. (A7) can be expressed as

$$\langle 0|\, \alpha_1^\dagger \alpha_{-1}^\dagger\, |0\rangle = 0. \tag{A9}$$

Thus the phases φ_1 and χ_1 cannot be determined by the minimum principle.

A transformation on the single-particle orbitals,

$$a_1 = a_1' e^{i\theta_1/2}, \tag{A10}$$

can be made in Eq. (A5) and in the Hamiltonian, so that real two-particle amplitudes A_1 can be used throughout. The condition in Eq. (A6) is satisfied if both χ_1 and φ_1 are zero. Therefore real coefficients can be used in the canonical transformation without loss of generality.

[35] D. H. Kobe, Ann. Phys. (New York) 35, 42 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Deformation and Contraction of Lie Algebras

MONIQUE LEVY-NAHAS
Centre de Physique Théorique de l'Ecole Polytechnique, 17, rue Descartes, 75, Paris V, France

(Received 16 June 1966)

This paper deals with the theory of deformation of Lie algebras. A connection is established with the usual contraction theory, which leads to some "more singular" contractions. As a consequence it is shown that the only groups which can be contracted in the Poincaré group are $SO(4, 1)$ and $SO(3, 2)$.

INTRODUCTION

THIS paper is divided into two parts. The first part deals with the deformation of Lie algebras as it has been described in Gerstenhaber's paper. We follow this paper very closely, just establishing in a more precise way the geometrical interpretation of the cohomology groups $B^2(A, A)$ and $Z^2(A, A)$ which naturally appear in the study of the deformations of an algebra A. We thus recover simply a sufficient condition of "rigidity" for an algebra A, which is $H^2(A, A) = 0$. We illustrate these results by an application to the three-dimensional real Lie algebras, for which we determine all the possible deformations.

In a second part, we show the connection between the theory of deformation and the usual concept of contractions, defined in a general way by Segal and then by Inonü, Wigner, and Saletan. We are thus led to a slightly different notation of contraction. It is then possible to partially answer the question: "What are all the algebras which can be contracted into a given algebra A?", because these algebras have to be searched for among the deformations of A.

From the viewpoint of physical applications, we show that the only Lie algebras which can give the Poincaré algebra by contraction are the semisimple Lie algebras of the de Sitter groups $SO(4, 1)$ and $SO(3, 2)$.

I. DEFORMATION OF LIE ALGEBRAS

A. Some Remarks About the Deformation Theory

1. *Definitions*

We recall here briefly the main definitions and results given in a paper by Gerstenhaber.[1] Let A be a finite-dimensional Lie algebra over a field k, which we restrict to be R or C, and let V be the underlying vector space of A. The product law in A is a mapping $V \wedge V \to V$ denoted by [ ]. Let $V_K = V_k \otimes k((t))$ be the vector space obtained by extending the field k to the field $K = k((t))$, where K is the quotient field of the power series ring $k[[t]]$. A *deformation*[2] of A is a Lie algebra A_t over the underlying vector space V_K, which is given by a mapping $f_t: V_K \wedge V_K \to V_K$ expressible in the following form:

$$f_t(a, b) = [a, b] + tF_1(a, b) + t^2F_2(a, b) + \cdots. \quad (1.1)$$

The F_i are bilinear functions of $V \wedge V$ into V and are defined over k.

We also impose on the series (1.1) in t that it be convergent in the neighborhood of the origin. Conversely, a given set of functions $F_i(a, b)$ does not necessarily correspond to a deformation of A, because the mapping f_t thus defined does not necessarily obey the Jacobi conditions.

By writing that

$$f_t(f_t(a, b), c) + f_t(f_t(b, c), a) + f_t(f_t(c, a), b) = 0, \quad (1.2)$$

one obtains the "*integrability conditions*" connecting the F_i:

$$\sum_{\mathfrak{S}(a,b,c)} \sum_{\substack{\mu+\nu=n \\ n=0,1,2}} F_\mu(F_\nu(a, b), c) + F_\nu(F_\mu(a, b), c) = 0, \quad (1.3)$$

where we use the following notations: $F_0(a, b) = [a, b]$, i.e., the original Lie law on A, and $\mathfrak{S}(a, b, c)$ is the circular permutation of a, b, c. For $n = 0$, the condition (1.3) is trivially satisfied: It is the Jacobi condition for A. For $n = 1$, one gets

$$\sum_{\mathfrak{S}(a,b,c)} F_1([a, b], c) + [F_1(a, b), c] = 0 = \delta F_2(a, b, c). \quad (1.4)$$

This relation expresses that F_1 is a 2-cocycle for the Lie algebra A:

$$F_1 \in Z^2(A, A)$$

(A being an A module for the adjoint representation). Conversely, an element of $Z^2(A, A)$ will give rise to a deformation only if it is *integrable*, that is, if it can be the first element F_1 of a sequence $\{F_i\}$ which satisfies the set of conditions (1.3). For $n = 2$, the

[1] M. Gerstenhaber, Ann. Math. 79, No. 1 (1964).

[2] Words being defined are given in italic.

conditions (1.3) give

$$\sum_{\mathfrak{S}(a,b,c)} F_1(F_1(a, b), c) = -\delta F_2(a, b, c). \quad (1.5)$$

When $F_1 \in Z^2(A, A)$, the first member of (1.5) is an element of $Z^3(A, A)$. By the preceding condition, this element has to belong also to $B^3(A, A)$, which means that its 3-cohomology class must be the null class. So the integrability conditions involve $H^3(A, A)$. In fact, as is shown in Gerstenhaber's paper, if $H^3(A, A)$ is zero, each $F_1 \in Z^2(A, A)$ is integrable.

For the following, we adopt the viewpoint of Gerstenhaber (Chap. II) by taking as a parameter space for the deformation theory the algebraic manifold $\mathcal{C}$ formed by the set of the structure constants of the n-dimensional Lie algebra over K. In fact, by using the canonical isomorphism $\mathfrak{L}(A, B) \simeq A^* \otimes B$, we identify the structure constants to the corresponding elements of $(A \wedge A)^* \otimes A$.

A point of $\mathcal{C}$ represents an algebra with a fixed basis. Now the previous deformation, with the condition of convergence, corresponds exactly to an analytic curve on $\mathcal{C}$, starting for $t = 0$ from the representative point of the Lie algebra A. It is clear that the *infinitesimal deformation* given by the element $F_1(a, b)$ corresponds to a tangent vector at this curve. Let us call $Z^{2\text{int}}(A, A)$ the manifold formed by the integrable elements of $Z^2(A, A)$. On the manifold $\mathcal{C}$, each element of $Z^{2\text{int}}(A, A)$ is a tangent vector to $\mathcal{C}$ at the point p representative of the algebra A. Let T_p be the linear tangent manifold attached to $\mathcal{C}$ at the point p. When p is a simple point[3] of $\mathcal{C}$ (i.e., a non-singular point), $Z^{2\text{int}}(A, A)$ which corresponds to the tangent cone at p and the tangent space T_p are identical. Then, as raised by Gerstenhaber (p. 86), we are faced with the problem of comparing the vector space $Z^2(A, A)$ with the linear tangent manifold T_p.

2. *Geometrical Discussion*

(a) *The tangent manifold T_p.* A point of $\mathcal{C}$ is defined by $\frac{1}{2}n^2(n-1)$ parameters $\xi_1, \cdots, \xi_{[\frac{1}{2}n^2(n-1)]}$ which are the structure constants C_{ij}^k. The manifold $\mathcal{C}$ is defined by the Jacobi equation,

$$f_l(\xi_1, \cdots, \xi_{[\frac{1}{2}n^2(n-1)]}) = 0, \quad 1 \leq l \leq n, \quad (1.6)$$

where the f_l are homogeneous polynomials of degree two:

$$f_l(C_{ij}^k) \equiv \sum_{\mathfrak{S}(i,j,k);m} C_{ij}^m C_{mk}^l. \quad (1.6')$$

Algebraically, the linear tangent manifold to $\mathcal{C}$ at the point p is defined by the equation

$$\sum_{i=1}^{\frac{1}{2}n^2(n-1)} \xi_i \frac{\partial f_l}{\partial \xi_i}(p) = 0,$$

which leads to

$$\sum_{m;\mathfrak{S}(i,j,k)} C_{ij}^m C_{mk}^l + C_{ij}^m C_{mk}^l = 0, \quad (1.7)$$

where C_{ij}^k are the coordinates of p. But (1.7) is nothing other than Eq. (1.4) expressed in terms of the coordinates. The elements C_{ij}^k solution of (1.7) are then the components of the 2-cocycles $Z^2(A, A)$, where A is the Lie algebra defined by the structure constants C_{ij}^k.

But it must be noticed that this linear tangent manifold is not necessarily identical to the geometrical tangent manifold. It is larger in general. A sufficient condition for this being true is that the ideal generated by the polynomials be equal to its radical. This is a conjecture of Gerstenhaber for the Jacobi ideal,[1] and has not yet been proved. This equality of the ideal with its radical has been shown in simpler but similar cases such as that of idempotent matrices.[4]

(b) *The orbits of $GL(V)$.* The group $GL(V)$ of regular linear mapping from V onto V acts on $\mathcal{C}$. We choose the following way: The algebra $A: a \wedge b \to [a, b]$ is transformed in $A': a \wedge b \to \varphi(a, b)$ such that (for example, one can take the inverse definition)

$$\varphi(a, b) = g[g^{-1}a, g^{-1}b]. \quad (1.8)$$

The transformed structure constants are

$$C'^l_{mn} = (g^{-1})^i_m (g^{-1})^j_n C_{ij}^k (g)^l_k,$$

where the C_{ij}^k are the structure constants of A.

A' and A differ only by a change of basis and are isomorphic.

Thus the orbits of $GL(V)$ on $\mathcal{C}$ are the class of isomorphic algebras. When g has the form

$$g = 1 + tg_1 + t^2 g_2 + \cdots \Big| \begin{matrix} g \in GL(V), \\ g_i \in End V, \end{matrix}$$

(1.8) gives a deformation of A which stays in the orbit of A and so can be said "trivial".[1] Let us examine the infinitesimal part. To the first order one obtains

$$\begin{aligned} \varphi(a, b) &= [a, b] + t(\varphi[a, b] - [\varphi a, b] - [a, \varphi b] + \cdots) \\ &= [a, b] - t\delta\varphi(a, b), \end{aligned} \quad (1.9)$$

[3] W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry* (Cambridge University Press, New York, 1964), Vol. II.

[4] M. Raynaud, Compt. Rend. **258**, 2457 (1964); **260**, 4391 (1965).

which means that *each tangent vector to a curve in the orbit corresponds to an element of* $B^2(A, A)$. Conversely, a coboundary $\in B^2(A, A)$ is always integrable, and gives rise to a deformation of A starting in a tangent direction at the orbit of A. Therefore, the dimension of the orbit of A is equal to the dimension of $B^2(A, A)$. This result can be found in another way by noticing that the stabilizer of the point P is Aut A and that the orbit can be identified with $GL(V)/\text{Aut } A$. Now the infinitesimal automorphisms (derivations of A) are just the elements of $Z^1(A, A)$ and are always integrable. Then $\dim Z^1(A, A) = \dim (\text{Aut } A)$ and $\dim (\text{orbit}) = \dim GL(V) - \dim (\text{Aut } A) = n^2 - \dim Z^1(A, A) = \dim B^2(A, A)$.

Application. In general for a Lie algebra of dimension n, one has

$$\dim B^2(A, A) = n^2 - n + \dim (\text{center of } A) - \dim H^1(A, A),$$

which gives in the case of a semisimple Lie algebra,

$$\dim B^2(A, A) = n^2 - n.$$

But this is also the dimension of $Z^2(A, A)$, hence the dimension of the irreducible component $\mathcal{C}_i$ of $\mathcal{C}$ to which belongs the representative point of the semisimple algebra.

It is now known that the manifold $\mathcal{C}$ can have another irreducible component of higher dimension.[5]

(c) *A Rigidity Theorem.*[1,6] *Definition*: an algebra is *rigid* if it cannot be deformed into an inequivalent algebra.

As already remarked by Gerstenhaber, a sufficient condition of rigidity for A is that $H^2(A, A) = 0$. Geometrically this means that the tangent plane to the orbit at the point A, and the geometrical tangent plane to the manifold at the same point coincide.

In particular, a semisimple algebra is rigid. (We use this fact in the following.) But there can exist rigid algebras of an other type; for example, in dimension 2, the only non-Abelian Lie algebra is rigid.

Remark: It is clear that a necessary condition of rigidity for algebra A is

$$Z^{2\text{int}}(A, A) = B^2(A, A), \tag{1.10}$$

so that the tangent cone $Z^{2\text{int}}(A, A)$ is a linear space in that case. Now, if one can prove the conjecture of Gerstenhaber [see the paragraph following Eq. (1.7)], then, at a simple point, $Z^2 = Z^{2\text{int}}$ and (1.10) would give $H^2(A, A) = 0$ as a necessary condition of rigidity for A.

3. *Classification of the Deformations*

We do not know how to solve this problem under its more general form, but only for the infinitesimal deformations. First, in order to have a classification, we have to define an equivalence relation between the deformations.

We say that two deformations of the same algebra A (with a given basis) are equivalent, if the deformed algebras are for each value of t in the same orbit (i.e., they correspond to isomorphic algebras).

It must be observed that this definition does not imply that the deformations are in a well-determined orbit.

In a more explicit way, it means that, for each value of t, there exists an element of the group $GL(V)$, Φ_t, such that

$$g_t(a, b) = \Phi_t^{-1} f_t(\Phi_t a, \Phi_t b), \tag{1.11}$$

where g_t and f_t are two deformations of the same algebra A. Now, considering only the case where Φ_t admits a development of the following form,

$$\Phi_t = \Phi_0 + t\Phi_1 + t^2\Phi_2 + \cdots,$$

we see from (1.11) that Φ_0 must be an element of the stabilizer $\Gamma(A)$ of A (= an automorphism of A). (For $\Phi_0 = 1$, it is the equivalence definition given by Gerstenhaber.)

Through the relation (1.11), there are conditions at each order, i.e., on all the components F_i and G_i of f_t and g_t. At the first order, one gets, by noting

$$\begin{aligned} F_0(a, b) &= G_0(a, b) = [a, b], \\ \Phi_0 \circ G_1(a, b) - F_1(\Phi_0 a, \Phi_0 b) &= [\Phi_0 a, \Phi_1 b] + [\Phi_1 a, \Phi_0 b] - \Phi_1[a, b] \\ &= (\Phi_0 \circ G_1 - F_1 \circ \Phi_0)(a, b). \end{aligned} \tag{1.12}$$

Let ${}^{\Phi_0}A$ be the A module given by

$$\begin{cases} a \in A \\ m \in {}^{\Phi_0}A \rightsquigarrow a \cdot m = [\Phi_0 a, m] \end{cases}$$

[this is a structure of A module because $\Phi_0 \in \Gamma(A)$]. The relation (1.12) can also be written

$$\Phi_0 \circ G_1 - F_1 \circ \Phi_0 \in B^2(A, {}^{\Phi_0}A). \tag{1.12'}$$

In particular, for $\Phi_0 = 1$, G_1 and F_1 differ only by a coboundary $\in B^2(A, A)$.

Therefore, at the first order, the classification problem can be solved conveniently by the two

[5] Recently, Michèle Vergne has found an example of an irreducible set of dimension greater than $2n^3/27$. [Thèse de 3ème cycle, Paris, I.H.P. Mai 1966].

[6] A. Nijenhuis and R. W. Richardson, Bull. Am. Math. Soc. January (1966), p. 1.

following steps. (1) The equivalence relation (1.11) when using transformations $\Phi_t = 1 + t\Phi_1 + t^2\Phi_2 + \cdots$ implies that the infinitesimal deformations F_1 and G_1 belong to the same cohomology class. Let us call $H^{2\text{int}}(A, A)$ the manifold obtained by considering $Z^{2\text{int}}(A, A)$ modulo the coboundaries. (2) Then we have to determine in $H^{2\text{int}}(A, A)$ the equivalent classes due to the action of the stabilizer $\Gamma(A)$ of A, i.e., to determine the orbits in $H^{2\text{int}}(A, A)$ of $\Gamma(A)$ in the sense of Sec. 2B. To each of these orbits corresponds a class of deformations equivalent to the first order.

Note: One should stress that this *equivalence to the first order* is very poor since two nonequivalent deformations (F_i), (F_i') may nevertheless implement the same element $F_i' = F_1 \in Z^2(A, A)$ and therefore be equivalent to the first order.

B. An Example. The Three-Dimensional Real Lie Algebras

Although the classification of the three-dimensional Lie algebras is already known (for example, see Jacobson[7] for the complex case, or the thesis of Sharp[8] for the real case), we present here a description of the orbits which happens to be very convenient for the study of the deformations.

1. *The Orbits*[9]

The following description is due to L. Michel. The action of the group $GL(V) = GL(3, R)$ on the manifold $\mathcal{C}$ was described in Sec. I2. The structure constants are transformed by

$$C'^{l}_{mn} = (g^{-1})^i_m (g^{-1})^j_n C^k_{ij} g^l_k. \tag{1.13}$$

Now, in the case of dimension 3, it is possible to associate a matrix to the system of the structure constants by writing

$$C^l_{ij} = \rho^{lk}\epsilon_{ijk}, \quad 1 \leq i, j, k \leq 3.$$

ϵ_{ijk} is the usual completely antisymmetric tensor.

From (1.13) we obtain

$$\rho'^{lk} = \tfrac{1}{2}\epsilon^{ijk}C'^{l}_{ij} = \tfrac{1}{2}\epsilon^{ijk}(g^{-1})^m_i (g^{-1})^n_j \rho^{l'k'}\epsilon_{mnk'}(g)^l_{l'}. \tag{1.14}$$

It is easy to transform (1.14) by using the definition of det G, where G is the matrix $(g)^i_j$:

$$\det G^{-1}\epsilon^{mnl} = \epsilon^{ijk}(g^{-1})^m_i (g^{-1})^n_j (g^{-1})^l_k;$$

then one gets

$$\rho'^{lk} = \det G^{-1}\tfrac{1}{2}\epsilon^{mnp}\epsilon_{mnk'}g^k_p \rho^{l'k'} g^l_{l'}$$

or

$$\rho'^{lk} = \det G^{-1} g^l_{l'}\rho^{l'k'}g^k_{k'}. \tag{1.14a}$$

Then, denoting by R the matrix ρ^{ij}, one finally obtains

$$\boxed{R' = (\det G)^{-1} GRG^t.} \tag{1.15}$$

Now the Jacobi condition can be described in terms of the matrix R. First we decompose R in a sum of a symmetric and an antisymmetric part: $R = S + A$, i.e., $\rho^{ij} = \sigma^{ij} + \epsilon^{ijk}\hat{\alpha}_k$, where σ^{ij} is the (ij) element of S, and $\epsilon^{ijk}\hat{\alpha}_k$ is the (ij) element of A. Then the Jacobi condition

$$\sum_{\mathfrak{S}(ijk);1} C^l_{ij}C^a_{lk} = 0$$

can be written

$$\rho^{an}\rho^{lm}\epsilon_{nlm} = 0,$$

and by using the previous decomposition,

$$\boxed{\rho^{ab}\hat{\alpha}_b = 0}, \quad \text{or equivalently,} \quad \sigma^{ab}\hat{\alpha}_b = 0. \tag{1.16}$$

Now the rank of the matrix R is preserved by (1.15), and this rank is equal to the dimension of the derived algebra G'. We are thus led to a classification of the orbits following the dimension of G'.[7]

(a) dim $G' = 3$. Det $R \neq 0$, the condition (1.16) can only be satisfied with $\hat{\alpha} = 0$, which means that R is symmetric.

It is shown in Jacobson that condition (1.15) is equivalent to

$$R' = \rho GRG^t, \tag{1.15'}$$

in which ρ is an arbitrary real number.

Then, by using the decomposition $G = SO$ with O orthogonal and S symmetric, we see that it is always possible to diagonalize R by means of O, and by (1.15') to fix an eigenvalue to 1. Now, one easily sees that the action of S cannot change the sign of the other eigenvalues. We thus obtain two orbits, called A_9 and A_8, which can be respectively characterized by the elements

$$A_9 : R = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \quad \text{and} \quad A_8 : R = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 1 \end{pmatrix}.$$

A_9 and A_8 are the three-dimensional semisimple algebras. A_9 is the algebra of the three-dimensional rotation group and A_8 is its noncompact form. It is obvious that the dimension of these orbits is 6 (= dim of the symmetric 3×3 matrices space).

[7] N. Jacobson, *Lie Algebras* (Interscience Publishers Inc., New York, 1962), p. 11.

[8] W. T. Sharp, thesis, Princeton University (1960).

[9] Let us notice that the orbits are not connected; they are made of two connected pieces.

(b) dim $G' = 2$. Any matrix of rank 2 can be written by means of a suitable orthogonal transformation

$$R = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{pmatrix} + \begin{pmatrix} 0 & \alpha_3 & -\alpha_2 \\ -\alpha_3 & 0 & \alpha_1 \\ \alpha_2 & -\alpha_1 & 0 \end{pmatrix}.$$

When σ_1 and $\sigma_2 \neq 0$, condition (1.16) implies $\alpha_1 = \alpha_2 = 0$. Under the action of a symmetric matrix, which can be taken diagonal

$$S = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix},$$

the matrix R is transformed into

$$R' = \begin{pmatrix} \sigma_1\lambda_1^2 & \alpha_3\lambda_1\lambda_2 & 0 \\ -\alpha_3\lambda_1\lambda_2 & \sigma_2\lambda_2^2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and we see that the quantity $\sigma_1\sigma_2/\alpha_3^2$ is an invariant λ, which can be used to classify the orbits.

Finally, one obtains the following orbits, which we describe by giving a representative element:

$$A_7 : R = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix},$$

$$A_6 : R = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix},$$

$$\underset{\lambda \neq -1}{A_5(\lambda)} : R = \begin{pmatrix} 1 & 1 & 0 \\ -1 & \lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For the case $\sigma_1 = \sigma_2 = 0$,

$$A_4 : R = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The dimensions of these orbits can be computed in a straightforward way, by determining the stabilizer of the elements R. The results are

$$\begin{cases} \dim A_7, A_6, A_5(\lambda) = 5, \\ \dim A_4 = 3 \end{cases}$$

(antisymmetric 3 × 3 matrices).

(c) dim $G' = 1$. There are only two kinds of orbits:

$$A_3 \text{ corresponding to } R = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

[A_3 can be integrated in the family $A_5(\lambda)$.]

$$A_2 \text{ corresponding to } R = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix},$$

the dimensions of which are, respectively, 5 and 3.

(d) dim $G' = 0$. There is the orbit of the Abelian algebra A_1, which is reduced to a point on $\mathcal{C}$ (dimension zero).

2. *Deformations*

From I, we know that the dimensions of $B^2(A, A)$ and of $Z^2(A, A)$ are, respectively, the dimensions of the orbit of A and of the tangent linear manifold to $\mathcal{C}$ in A. A straightforward computation gives the results shown in Table I.

It can be seen directly from the equations of the manifold $\mathcal{C}$ that the singular points belong to A_1, A_2, A_6, and A_7. For the simple points the dimension of $\mathcal{C}$ is 6. For the other points, we have to determine the manifold $Z^{2\text{int}}(A, A)$ and see that its maximum dimension is also 6. We now discuss the deformations corresponding to each case.

(a) *Deformation of the Simple Points.* We know that A_8 and A_9 are rigid, which means that all their deformations are trivial. Then we have only to consider A_4 and the complete family $A_5(\lambda)$ (with $\lambda = -1$ included).

Now, all the nontrivial deformations of an algebra belonging to an $A_5(\lambda)$ or to A_4 are of the following form: they must cross over all the orbits $A_5(\lambda)$. For example, let us consider the following first-order

TABLE I.

A	dim $B^2(A, A)$	dim $Z^2(A, A)$
A_1	0	9
A_2	3	8
A_3	5	6
A_4	3	6
$\underset{\lambda \neq -1}{A_5(\lambda)}$	5	6
A_6	5	7
A_7	5	7
A_8	6	6
A_9	6	6

deformation of $A_5(\lambda)$.

$$A^t \begin{cases} [e_1, e_2] = 0, \\ [e_3, e_2] = -e_1 - e_2, \\ [e_3, e_1] = -e_1 + \lambda e_2 + te_2. \end{cases}$$

For each value of t, this algebra belongs to a different orbit $A_5(\lambda + t)$. (In fact, it is impossible to deform a nonsymmetric matrix into a symmetric one, so there are certainly no deformations of the preceding algebras into A_8 or A_9.)

Let us notice also that the dimension of $H^{2\text{int}}(A, A)$ [here $= H^2(A, A)$] measures the codimension of the orbit of the algebra which is deformed in the manifold formed by the deformed algebras. [The set of $A_5(\lambda)$ is a manifold of dimension 6.] However, we do not know if this result can be established generally.

(b) *Deformation of the Singular Algebras.* We now describe in more detail the case of A_6 and A_7, and especially of A_2. First we associate to the element $F_1(a, b)$, which characterizes the infinitesimal deformation (see Sec. A), a matrix R_1 as in Sec. B1.

We use the following notations:

$$R_1 = \begin{pmatrix} y & q & u \\ s & z & r \\ v & w & x \end{pmatrix}.$$

We note the deformation in term of matrices:

$$R_t = R_0 + tR_1 + \cdots,$$

where R_0 is the matrix associated to the nondeformed algebra.

Deformations of A_6 (A_7 can be treated in the same way). We choose in A_6 the algebra corresponding to

$$R = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}.$$

F_1 must be a cocycle: $F_1 \in Z^2(A_6, A_6)$. This imposes the conditions

$$\text{(a)} \quad \begin{cases} r = w, \\ u = v. \end{cases}$$

The coboundaries are defined by

$$\text{(b)} \quad \begin{cases} y = 0, \\ q = s, \end{cases}$$

and the integrability condition (1.5) is simply $x(q - s) = O(c)$. The manifold defined by (c) is the union of two components:

$$c_1: q = s,$$
$$c_2: x = 0,$$

which already yields the (maximum) dimension of $Z^{2\text{int}}$ as 6. Let us examine first the deformations which are linear in t: $R_t = R_0 + tR_1$. Then the preceding integrability condition involves two other equations:

$$\text{(c')} \quad v(q - s) = 0,$$
$$\text{(c'')} \quad w(q - s) = 0.$$

$F_1 \in C_1$: (c') and (c'') are satisfied. The matrix R_t is symmetric with a nonvanishing determinant for $x \neq 0$. Then there exists a deformation only into A_8. For A_7 two possibilities exist: into A_8 for $x < 0$, and into A_9 for $x > 0$.

$F_1 \in C_2$: In general in that case we have $q \neq s$. Therefore from (c') and (c'') we must have $v = w = 0$.

$$R_t = \begin{pmatrix} 1 + yt & tq & 0 \\ ts & -1 + tz & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

it is a deformation which crosses over the family $A_5(\lambda)$.

For $F_1 \in C_1 \cap C_2$, the linear transformations are trivial. One stays in the orbit of A_6. For deformations at higher orders, the conditions (c') and (c'') are no longer valid.

$F_1 \in C_1$: We always have a deformation into A_8 or A_9.

$F_1 \in C_2$: The deformation starts in a transverse direction to all the $A_5\,|\lambda|$ and then can (1) go by means of the higher-order terms completely inside one orbit $A_5(\lambda)$, or (2) stay always transverse to the family.

$F_1 \in C_1 \cap C_2$: All the cases are possible. The deformation starts tangent to the orbit, and then takes one of the previous ways.

Deformations of A_2: Let us consider the algebra

$$A_2 \begin{cases} [e_1, e_2] = e_3, \\ [e_2, e_3] = 0, \\ [e_3, e_1] = 0. \end{cases}$$

With the same notations, one easily finds that the spaces of the cocycles and of the coboundaries are determined by

$$Z^2(A_2, A_2): s = q,$$
$$B^2(A_2, A_2): s = q = 0, \quad y = z = 0; \quad u = v; \quad w = r.$$

The condition of integrability to the first order (1.4) is

$$\text{(d)} \quad \begin{cases} s(u - v) + y(w - r) = 0, \\ z(u - v) + s(w - r) = 0; \end{cases}$$
$$\text{(d')} \qquad -vr + wu = a,$$

where a is an arbitrary number (depending on F_2); if $F_2 = 0$, then $a = 0$.

The linear deformations. The manifold $Z^{2\text{int}}(A_2, A_2)$ is defined by the equations (d) and (d'). We distinguish the following submanifolds. [The distinction is based on the equation (d), in order to be able to consider the more general deformations.]

$V_1: \Delta = s^2 - yz \neq 0$, therefore $u = v$, $w = r$.
$V_2: \Delta = s^2 - yz = 0$ and in general $u \neq v$ and $w \neq r$.

One can verify that the manifolds are invariant under the action of the stabilizer G_0 of A_2.

$F_1 \in V_1$: The matrix R_t is symmetric, with a nonvanishing determinant in a neighborhood of $t = 0$ (because the lowest-order term is Δ). Thus there is a deformation into A_8 for $\Delta > 0$ and A_9 for $\Delta < 0$.

Example: deformation into A_9:

$$\begin{cases} [e_1, e_2] = e_3, \\ [e_2, e_3] = te_1, \\ [e_3, e_1] = te_2. \end{cases}$$

$F_1 \in V_2$: We have to consider the equation (d') with $a = 0$. Then one can see that the matrix R_t has a null determinant. Moreover, it is necessarily a nonsymmetric matrix, because $u \neq v$ and $r \neq w$; so it gives a deformation into the family $A_5(\lambda)$, which crosses over all the family. Example:

$$A_t \begin{cases} [e_1, e_2] = e_3 + te_1, \\ [e_2, e_3] = -te_3 + Kte_1, \\ [e_1, e_3] = 0. \end{cases}$$

For each value of t, A_t belongs to $A_5(Kt)$.

$F_1 \in V_2 \cap V_2$: ($\Delta = 0$, $u = v$; $w = r$). R_t is symmetric; its determinant is zero, so there is a deformation into A_6 or A_7. Example:

$$\begin{cases} [e_1, e_2] = e_3, \\ [e_2, e_3] = te_1, \\ [e_1, e_3] = 0. \end{cases}$$

One also has $\dim Z^{2\text{int}}(A_2, A_2) = \dim V_2 = \dim V_1 = 6$, $\dim V_3 = 5$.

The general deformations. $F_1 \in V_1$: Due to the rigidity of semisimple Lie algebras, there are always the two kinds of deformations into A_8 or A_9, and nothing more.

$F_1 \in V_2$: Exactly as for A_6, the deformation can stay always transverse to the family $A_5(\lambda)$, or can go more precisely in an orbit $A_5(\lambda)$ well determined (and then stay in it). Example:

$$\begin{cases} [e_1, e_2] = e_3 + te_1, \\ [e_2, e_3] = -te_3 + Kt^2e_1, \\ [e_1, e_3] = 0. \end{cases}$$

A_t belongs to the orbit $A_5(K)$.

$F_1 \in V_1 \cap V_2$: All the cases are possible. For example, to study the second-order deformations, one has to classify the deformations following the values of F_2 which is also a cocycle in that case [$F_2 \in V_1 \Rightarrow$ deformation into A_8 or A_9, etc.].

Finally we see that the study of the first-order (linear) deformations produces a rough classification of the equivalent deformations in well-determined families of orbits. Then the next-order deformations can go only in a more precise direction inside one of these families.

Remark: It is very easy to obtain the corresponding results in the complex case. Then A_9 and A_8 are in the same orbit, like A_9 and A_7 or $A_5(\lambda)$ and $A_5(-\lambda)$. The dimension over C of the cohomology groups are respectively the same, and the deformations can be immediately deduced from the previous ones.

II. CONTRACTION

The concept of contraction for Lie groups and algebras was introduced by Segal and by Inonü and Wigner in 1953.[10] It was later examined in a paper of Saletan[11] in which in particular a general condition for contraction is stated precisely. Our aim in this part, is to establish a connection between the process of contraction and deformation as it can be expected from a geometrical point of view. In doing this, we are led to add some slight refinements to Saletan's paper; it is therefore necessary to recall briefly the part of this paper that we need in the following.

A. The Saletan's Contraction

g is a Lie algebra over the underlaying vector space V, with the law: $a \wedge b \to [a, b]$. Let Φ_t be a linear mapping of $V_K = V_{k((t))}$ into itself (see Sec. IA), nonsingular for $t \neq 0$ and singular for $t = 0$, of the following form:

$$\Phi_t = u + tv, \tag{2.1}$$

where u and v are linear mappings from V into V defined over k and u is a singular mapping.

The contracted algebra g_1 is then defined by the limit law:

$$\lim_{t \to 0} \Phi_t^{-1}[\Phi_t a, \Phi_t b] = [a, b]^{(1)}. \tag{2.2}$$

Let us assume also that v is a regular mapping; then there is no loss of generality by taking for example $v = 1$.[12]

Now V is a finite-dimensional vector space, and the

[10] E. Inonü and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. 39, 510 (1953).
[11] E. J. Saletan, J. Math. Phys. 2, 1 (1961).
[12] This is not the convention of Saletan who chooses $v = 1 - u$, in order to have $\Phi_1 = 1$.

method of Saletan consists in using the fitting decomposition of[11] V under the form $V_R \oplus V_N$, where V_R and V_N are u-invariant subspaces defined in a canonical way with respect to u, such that u is surjective on V_R and nilpotent on V_N.

Then, if it exists, the law of the limit algebra g_1 is given by[13]

$$[a, b]^{(1)} = u^{-1}[ua, ub]_R + [ua, b]_N + [a, ub]_N - u[a, b]_N. \quad (2.3)$$

The indices N and R denote the components with respect to the decomposition $V = V_R \oplus V_N$. A necessary and sufficient condition for the existence of g_1 is

$$u([ua, b]_N + [a, ub]_N - u[a, b]_N) = [ua, ub]_N. \quad (2.4)$$

Let us notice that it is equivalent to write this condition (2.4) under the form

$$u[a, b]^{(1)} = [ua, ub]. \quad (2.4a)$$

And the condition (2.4a) simply expresses that u is an homomorphism of the algebra g_1 into the algebra g. But conversely, the mere existence of such homomorphism between two algebras g_1 and g is not a sufficient condition for g to be contracted in g_1 because in (2.4a) the law of g_1 depends on u.

B. First Connection with the Deformation Problem

Let us define $f_t(a, b) = \Phi_t^{-1}[\Phi_t a, \Phi_t b]$ with $\Phi_t = u + t$. Then we have on V_N by considering only the values of f_t belonging to V_N,

$$f_t^{(N)}(a, b) = [a, b]_N^{(1)} + t[a, b]_N;$$

and on V_R,

$$f_t^{(R)}(a, b) = [a, b]_R^{(1)} + t[a, b]_R + u^{-2}t(1 + tu^{-1})^{-1} \times [u\delta u(a, b) - [ua, ub]]_R,$$

where $(1 + tu^{-1})^{-1}$ is a formal series in t.

It is clear that, in a neighborhood of the origin, the preceding development is convergent. Therefore $f_t(a, b)$ is a deformation of the contracted algebra g_1 lying in the orbit of g. (For $t \neq 0$, one obtains an algebra isomorphic to g.) Geometrically, by considering the structure constants manifold C (defined in A), it means that the algebra g_1, and in fact all its orbit belongs to the edge of the orbit of g (see Fig. 1).

Remark: Instead of (1), we can consider the more general mapping of V_K into itself, defined in a neighborhood of the origin:

$$\Phi_t = u + tv + t^2w + t^3x + \cdots. \quad (2.5)$$

[13] Which can also be written $[a, b]^{(1)} = u^{-1}[ua, ub]_R + \delta u(a, b)_N$.

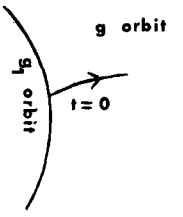


FIG. 1. Deformation of g_1 into the orbit of g.

Let us assume that $\psi_t = v + tw + t^2x + \cdots$ is a regular mapping for each t, and u is always a singular mapping. Then it is easy to see that this development does not give any new results.[14] It leads to a contracted algebra g_1 and to a deformation exactly as before.

Now, conversely, it would seem natural to require that any Lie algebra that can be deformed in a well-determined orbit can also be obtained by contraction of an element of this orbit. But with the previous definition of contraction, we obviously have a counterexample.

C. The Example of the Three-Dimensional Rotation Algebra

We have seen in Sec. IB that the three-dimensional Lie algebra A_2, defined by

$$\begin{cases} [e_1, e_2] = e_3, \\ [e_2, e_3] = 0, \\ [e_3, e_1] = 0, \end{cases}$$

can be deformed in the orbit of A_9, the rotation algebra. But it is easy to see that there is no homomorphism of A_2 into A_9, except the trivial one. The condition (2.4a) can only be satisfied with $u = 0$, but, by (2.3), it then leads to an Abelian Lie algebra, so that A_9 cannot be contracted into A_2. Besides, Saletan has proved that the only nontrivial contraction of A_9 defined by (2.3) and (2.4) leads to the Euclidean algebra A_7.

D. "More Singular" Contractions

Thus it is necessary to look at the case $u = 0$, that is, to take for the mapping Φ_t a more singular mapping at the limit $t = 0$ of the form

$$\Phi_t = tv + t^2w + t^3x + \cdots,$$

which can also be written

$$\Phi_t = t(v + t\Psi_t). \quad (2.6)$$

We suppose[15] that $v + t\psi_t$ is a nonsingular mapping

[14] Indeed by writing $\Phi_t = (u\psi_t^{-1} + 1)\psi_t$, we see that is sufficient to consider $w = uv^{-1}$ instead of u, and then the contracted algebra by Φ_t is isomorphic to the contracted algebra by $w + t$, the isomorphism being w.

[15] We suppose also ψ_t always nonsingular (even for $t = 0$).

when t is different from zero, but v alone is a singular mapping. Now, exactly as in Sec. IA, we do not restrict the generality by taking $\psi_t = 1$. Let us then study the limit (2.2) by using the same decomposition of the vector space V in $V_R \oplus V_N$ with respect to the mapping v. We obtain for

$$f_t(a, b) = \Phi_t^{-1}[\Phi_t a, \Phi_t b],$$

with the same notations,

$$\begin{aligned} f_t(a, b) &= f_t(a, b)_R + f_t(a, b)_N, \quad a, b \in V, \\ f_t(a, b)_R &= tv^{-1}(tv^{-1} + 1)^{-1}[t^2[a, b]_R \\ &\quad + t([ua, b]_R + [a, ub]_R) + [ua, ub]_R], \end{aligned} \tag{2.7}$$

and, by using the nilpotence of v, we have on V_N

$$(t + v)^{-1} = \frac{1}{t}\left(1 + \frac{v}{t}\right)^{-1} = \frac{1}{t}\sum_{n=0}^{q-1}\frac{(-v)^n}{t^n}.$$

q is the smallest number such that $v^q = 0$, so

$$\begin{aligned} f_t(a, b)_N = t\Big(&\frac{1}{t} - \frac{v}{t^2} + \frac{v^2}{t^3} + \cdots\Big) \\ &\times [v^2[a, b]_N - v([va, b]_N + [a, vb]_N) \\ &+ [va, vb]_N] + t(t[a, b]_N + [ua, b]_N \\ &+ [a, ub]_N - u[a, b]_N). \end{aligned} \tag{2.7a}$$

Therefore, if $f_t(a, b)$ has a limit for $t = 0$, it gives a Lie algebra $g^{(2)}$ corresponding to

$$\begin{aligned} [a, b]^{(2)} = v^2[a, b]_N - v([va, b]_N + [a, vb]_N) \\ + [va, vb]_N. \end{aligned} \tag{2.8}$$

Only the V_N component gives a contribution at the limit; the V_R part gives zero.

Equation (2.8) can also be written

$$[a, b]^{(2)} = [va, vb]_N - v(\delta v(a, b))_N \tag{2.9}$$

and a necessary and sufficient condition for the existence of this limit is, from (2.7a),

$$v[a, b]^{(2)} = 0. \tag{2.10}$$

Let us remark that, introducing

$$[a, b]^{(1)} = v^{-1}[va, vb]_R + \delta v(a, b)_N,$$

expression (2.9) takes the following form:

$$[a, b]^{(2)} = [va, vb] - v[a, b]^{(1)}. \tag{2.11}$$

By comparing this with (2.4a), we see that if the initial algebra g can be contracted by v (i.e., by taking $\Phi_t = v + t$), then the algebra $g^{(2)}$ obtained by contracting with $t\Phi_t$ is the Abelian algebra. Conversely, the existence of a non-Abelian Lie algebra $g^{(2)}$ "measures" the lack of verification of (2.4a).

Now, if it is impossible to obtain a contracted algebra by (2.6), the previous process obviously can be extended by taking for Φ_t more and more singular mappings at $t = 0$. For example, with $\Phi_t = t^2(w + t)$, one obtains an algebra $g^{(3)}$:

$$[a, b]^{(3)} = -w[a, b]^{(2)}, \tag{2.12}$$

[we define $[a, b]^{(2)}$ in function of w by means of the expression (2.9)], and the necessary and sufficient condition of existence for $g^{(3)}$ is still

$$w[a, b]^{(3)} = 0. \tag{2.13}$$

It follows from expression (2.7a) that it is possible to go on in this way: the contraction by $\Phi_t = t^n(u + t)$ gives an algebra g^{n+1} ($n \geq 1$) defined by

$$\begin{aligned} [a, b]^{(n+1)} &= -u[a, b]^{(n)} \\ &= (-)^{n-1}u^{n-1}[a, b]^{(2)}, \end{aligned} \tag{2.14}$$

where $[a, b]^{(2)}$ is defined by (2.9) with respect to u. The necessary and sufficient condition of contraction is

$$u[a, b]^{(n+1)} = 0. \tag{2.15}$$

If this condition is satisfied at order n, the next algebras (of higher order) are Abelian. Otherwise, we can continue until the algebra g^q, corresponding to $\Phi_t = t^{q-1}(u + t)$, is reached

$$\begin{aligned} [a, b]^{(q)} &= (-)^{q-1}u^{q-1}[a, b]^{(2)} \\ &= (-)^{q-1}u^{q-1}[ua, ub]_N, \end{aligned}$$

and there is no condition because $u[a, b]^{(q)}$ is always zero.

Remark: $[a, b]^{(q+1)} = 0$, so that it is always possible to contract into the Abelian algebra.

It is convenient to put these results into a diagram:

$$\begin{array}{ccc} (a, b) & \xrightarrow{u\times u} & (ua, ub) \\ \downarrow & \searrow^{\Delta} & \downarrow \\ [a, b]^{(1)} & \xrightarrow{u} [ua, ub] \xrightarrow{u} & [a, b]^{(3)} \xrightarrow{u} [a, b]^{(4)} \cdots, \end{array}$$

where $\Delta(a, b) = [a, b]^{(2)} = [ua, ub] - u[a, b]^{(1)}$. The condition (2.4a) of existence of g^1 is just the commutativity of the diagram. If $\Delta \neq 0$ and $u \circ \Delta = 0$, one can contract g into $g^{(2)}$, etc. In general, if $u^i \circ \Delta = 0$ and $u^{i-1} \circ \Delta \neq 0$, algebra g can be contracted into $g^{(i+1)}$ by using the mapping $\Phi_t = t^i(u + t)$.

E. Example

Let us return to the rotation case of Sec. IIC. The rotation algebra A_3 can be contracted into A_2 by using an application

$$\Phi_t = t(v + t(1 - v)).$$

Here we normalize differently[11] in order to have $\Phi_1 = 1$. We only have to take for v, in the basis e_1,

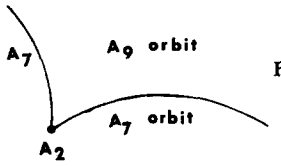


FIG. 2. Orbits of algebras A_2, A_7, and A_9.

e_2, e_3 used in Sec. IIC, the following application:

$$v = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix},$$

which gives

$$\Phi_t = \begin{pmatrix} t & & \\ & t & \\ & & t^2 \end{pmatrix}.$$

Then one can immediately verify that

$$[e_1, e_2]^{(2)} = e_3,$$
$$[e_2, e_3]^{(2)} = 0,$$
$$[e_3, e_1]^{(2)} = 0,$$

and that $v[e_i, e_j]^{(1)} = 0$.

Here one recognizes a contraction already pointed out by Inonü.[16]

From the geometric viewpoint, it corresponds to the following situation: The orbit of A_2 belongs to the edge of the "stratum" of the rotations algebra A_9, but also to the edge of the A_7's stratum. (Let us recall that the dimensions of the orbits are, respectively, 6 for A_9, 5 for A_7 and 3 for A_2.) It means that the situation of A_2 is more singular on C, for example, as it can be suggested by Fig. 2, where the orbit of A_2 is represented by a point. As is already known,[16] it is possible to reach A_2 from A_9 by means of two successive I–W contractions (and that is very clear geometrically). However, we do not know if this corresponds to a general situation; the solution of this problem seems to be related to a best knowledge of the structure constants manifold for an arbitrary dimension.

To sum up, through the more singular way of contracting, it is possible to obtain some Lie algebras, the orbits of which have themselves some more singular positions on the structure constants manifold. At the limit, it is always possible to reach the Abelian Lie algebra which belongs to the edge of all the orbits.

F. Deformation and Contraction

By (2.7) and (2.7a), we see that a deformation of the contracted algebra in a well-determined orbit corresponds exactly to a more singular contraction, as in Sec. IIB.. Now we are interested in the converse of this property.

Let us suppose that an algebra g_1 can be deformed in a well-determined orbit (an element of which being g, with the law: $[a, b]$). For each value of t ($t \neq 0$), the deformation $f_t(a, b)$ is isomorphic to $[a, b]$. Then there must exist an element of the group $GL(V)$ ψ_t such that, for each $t \neq 0$,

$$f_t(a, b) = \psi_t^{-1}[\psi_t a, \psi_t b]. \tag{2.16}$$

We know by assumption that the previous expression has a limit for $t = 0$, which corresponds to the algebra g_1. Now, to prove the reciprocity completely, the problem is to show that between the ψ_t (not uniquely) defined by (2.16), there is at least one which has an analytic expression in the neighborhood of $t = 0$ of the form (2.5), (or even with a first term in t^n,) the first element of this development being in general a nonregular application. Though this result seems extremely likely, we have not been able to prove it. Then it is not sure that the more general way of contracting is obtained by the previous Secs. IIA and IID contractions.

However, we have the partial result: all the algebras which can be contracted in a given one g_1 in the sense of Secs. IIA or IID have to be searched for among the deformations of g_1. The advantage is that the research of the deformations can be made in a more systematic way.

III. APPLICATION TO THE POINCARÉ'S GROUP

The knowledge of all the group or Lie algebras which can be contracted into the inhomogeneous Lorentz group $\mathfrak{I}$ may present some interest in physics [for example, from the viewpoint of general relativity,[17] or from the viewpoint of dynamical groups of Barut and Böhm.[18]] We already know that, between the semisimple Lie algebras, only the de Sitter's algebras can give $\mathfrak{I}$ by contraction, as it was proved by Sharp.[8] Now, by studying the deformations of the Poincaré Lie algebra, we can prove that they are the only possible algebras which can be contracted in $\mathfrak{I}$.

The Deformations of the Poincaré Lie Algebra

Let us consider the first-order deformations. Applying the results of Sec. I3, we have to compute $H^2(P, P)$ and $H^3(P, P)$. This is done in the Appendix. The result says that $H^2(P, P)$ is a one-dimensional space and that $H^3(P, P)$ is zero. It proves there are no

[16] E. Inonü, in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gürsey, Ed. (Gordon and Breach Science Publishers, Inc., New York, 1965), p. 391.

[17] F. Gürsey, in Ref. 16, p. 365.

[18] A. O. Barut and A. Böhm, Phys. Rev. **139**, B1107 (1965).

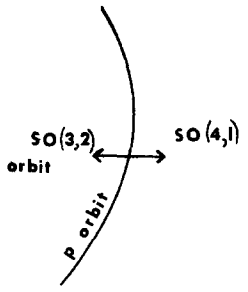


FIG. 3. Deformations of Poincare's algebra.

obstructions to the deformation (which means that every cocycle is integrable), and moreover there is only one "direction" of deformation (modulo the displacements in the $\mathfrak{I}$ orbit). This unique type of first-order deformations is easily obtained by taking, for example [see Ref. 17, p. 373],

$$-(i/\hbar)[J_{k\lambda}, J_{\mu\nu}] = \delta_{k\nu}J_{\lambda\mu} - \delta_{k\mu}J_{\lambda\nu} + \delta_{\lambda\mu}J_{k\nu} - \delta_{\lambda\nu}J_{k\mu},$$
$$-(i/\hbar)[\Pi_\lambda, J_{\mu\nu}] = \delta_{\lambda\mu}\Pi_\nu - \delta_{\lambda\nu}\Pi_\mu,$$
$$-(i/\hbar)[\Pi_\mu, \Pi_\lambda] = tJ_{\mu\nu}.$$

This corresponds to two classes of algebras following the sign of t (we have $t = -1/R^2$, where R is the radius of curvature of the universe[17]):

$t > 0$: it is the algebra of $S0(3, 2)$,

$t < 0$: it is the algebra of $S0(4, 1)$.

It comes from the rigidity of the semisimple Lie algebras that the higher-order deformations give nothing more.

Now this situation can be geometrically understood on the structure constants manifold corresponding to the 10-dimensional Lie algebras. Let us recall that the dimension of a semisimple Lie algebra's orbit is $n^2 - n$, which gives 90 for the de Sitter's orbits. The dimension of the Poincaré orbit is given by

$$\dim B^2(P, P) = n^2 - n - \dim H^1(P, P),$$

which is 89 because $H^1(P, P)$ is a one-dimensional space.[19] So the situation can be illustrated by Fig. 3.

ACKNOWLEDGMENTS

The author wishes to express her gratitude to Professor D. Fotiadi and Professor J. Lascoux for their constant interest and aid. She is also indebted to them for having conjectured most of the results. She also thanks Professor L. Michel for some stimulating discussions and Dr. P. Renouard for many interesting remarks.

[19] L. Michel, in *Lectures in Theoretical Physics* (University of Colorado Press, Boulder, Colorado, 1964), Vol. VIIa, p. 117.

APPENDIX. COMPUTATION OF $H^2(P, P)$ AND $H^3(P, P)$

We use the following formula of Serre and Hochschild[20]

$$H^n(g, M) = \sum_{i+j=n} H^i(g/K, F) \otimes H^j(K, M)^g, \quad \text{(A1)}$$

where M is a finite dimensional g module, K is an ideal of g such that g/K is semisimple, and F is the field.

The meaning of the symbol g on the right is defined later.

A. $H^2(P, P)$

By taking for K the translation ideal, P, we get directly from (A1):

$$H^2(P, P) = H^2(T, \mathbf{P})^{\mathbf{P}}$$

and the problem is reduced to the computation of $H^2(T, P)$.

1. *Coboundaries* $B^2(T, P)$

Let g be an element of $B^2(T, P)$; i.e., g is a bilinear function $T \wedge T \to P$ such that there exists a one-cochain $\Phi: T \to P$ verifying

$$g(x_i, x_j) = \delta\Phi(x_i, x_j) = [x_i, \Phi(x_j)] - [x_j, \Phi(x)]\ x_i \in T. \quad \text{(A2)}$$

Φ can be taken with its values in $L \simeq P/T$

$$\Phi(x_i) = \sum_{\mu\nu} d_i^{\mu\nu} M_{\mu\nu}, \quad \text{(A3)}$$

where $M_{\mu\nu}$ is the usual basis of L.

Now use the decomposition, $g = g_L + g_T$, where

$$g_L(x_i, x_j) = g(x_i, x_j)/_L$$

and

$$g_T(x_i, x_j) = g(x_i, x_j)/_T.$$

We see from (A2) that the component g_L is zero. So a coboundary is a function $g_T: T \wedge T \to T$,

$$g_T(x_i, x_j) = \alpha_{ij}^k x_k$$

satisfying (A2), which means, explicitly,

$$\alpha_{ij}^k = 2(d_j^{ik} g_{ii} - d_i^{jk} g_{jj}). \quad \text{(A4)}$$

But, given any $\alpha_{ij}^k = -\alpha_{ji}^k$, it is always possible to find a function Φ, such that (A2) [or (A4)] is verified. In fact this function is uniquely determined by

$$d_k^{ij} = \tfrac{1}{4}(\alpha_{kj}^i g^{jj} g^{ii} g_{kk} + \alpha_{kj}^i g^{jj} + \alpha_{ik}^j g^{ii})$$

and we have the result $B^2(T, \mathbf{P}) = C^2(T, T)$, where $C^2(T, T)$ is the space of all the bilinear functions: $T \wedge T \to T$.

2. *Cocycles* $Z^2(T, P)$

$f \in Z^2(T, P)$ if

$$\delta f(x_1, x_2, x_3) = \sum_{\mathfrak{I}(1,2,3)} [x_1, f(x_2, x_3)] = 0. \quad \text{(A5)}$$

[20] G. Hochschild and J. P. Serre, Ann. Math. 57, 603 (1953).

Using $f = f_T + f_L$ as before, we see that (A5) is only a condition on f_L.

3. $H^2(T, P)^P$

There is a canonical structure of P module on $C^2(T, P)$ which defined by

$$f \in C^2(T, \mathbf{P}) \rightsquigarrow g \cdot f \in C^2(T, \mathbf{P}); \quad g \in \mathbf{P},$$
$$(g \cdot f)(x_1, x_2) = g \cdot f(x_1, x_2) - f(g \cdot x_1, x_2) - f(x_1, g \cdot x_2).$$

$H^2(T, P)^P$ is the set of elements of $H^2(T, P)$ invariant by g, that is such that $g \cdot f$ is in the same cohomology class as f. This is always verified if $g \in T$, because in this case

$$(g \cdot f)(x_1, x_2) = [g, f(x_1, x_2)] \in T,$$
$$g \cdot f \in C^2(T, T) = B^2(T, \mathbf{P}).$$

A representative element in the class of f is given by an $f_L: T \wedge T \to L$. Now taking g in L, one sees that $g \cdot f_L$ is also a function $T \wedge T \to L$. Then the condition of invariance implies $g \cdot f_L = 0$, and is equivalent to the following proposition:

$$f_L \in B(T \wedge T, L)^L, \tag{A6}$$

where B is the space of functions $T \wedge T \to L$, and T, L being considered as L module (for the adjoint representation). As L modules,

$$\begin{cases} T \approx D^{\frac{1}{2},\frac{1}{2}}, \\ L \approx D^{1,0} + D^{0,1}. \end{cases}$$

So that we get from (A6), using canonical identifications,

$$f_L \in (D^{\frac{1}{2},\frac{1}{2}} \wedge D^{\frac{1}{2},\frac{1}{2}} \otimes D^{1,0} \otimes D^{0,1})^L. \tag{A7}$$

The only L invariant part is the trivial one: now $D^{0,0}$ appears twice in the decomposition, which means that $B(T \wedge T, L)^L$ is a *two-dimensional space*. It is easy to find two independent elements. We know the only invariant combinations of X_ρ, Y_σ, $M_{\mu\nu}$ are

$$(\alpha) \quad M_{\mu\nu} X^\rho Y^\sigma g^\mu_\rho g^\nu_\sigma,$$
$$(\beta) \quad M_{\mu\nu} X^\rho Y^\sigma \epsilon^{\mu\nu}_{\rho\sigma},$$

but only one of these elements (α) is a cocycle, so we have the final result that $H^2(T, P)^P = H^2(P, P)$ *is a one-dimensional space.*

B. $H^3(P, P)$

From $H^1(L, K) = H^2(L, K) = 0$ and also

$$H^0(T, P)^P = 0,$$

Eq. (A1) yields

$$H^3(P, P) = H^3(T, P)^P.$$

1. *Coboundaries* $B^3(T, P)$

$g = g_L + g_T$ [see (A1)] belongs to $B^2(T, P)$ if there exists a function $\Phi: T \wedge T \to P$ such that

$$g(x_1, x_2, x_3) = (\delta\Phi)(x_1, x_2, x_3) = \sum_{\mathfrak{S}(1,2,3)} [x_1, \Phi(x_2, x_3)], \tag{A8}$$

the component g_L is zero.

Let us define

$$\begin{cases} \Phi(x_i, x_j) = C^{\mu\nu}_{ij} M_{\mu\nu}, \\ g_T(x_i, x_j, x_k) = \alpha^l_{i,j,k} x_l. \end{cases}$$

The condition (A8) is equivalent to

$$\alpha^l_{ijk} = 2[C^{kl}_{ij} g_{kk} + C^{il}_{jk} g_{ii} + C^{jl}_{ki} g_{jj}]. \tag{A9}$$

Now it is always possible to find a set of coefficients C^{kl}_{ij} verifying (A9) (which is not uniquely determined in this case). For example,

$$C^{jl}_{ki} = \tfrac{1}{4}[\tfrac{4}{3}\alpha^l_{ijk} g^{jj} + \tfrac{1}{2}\alpha^k_{ijl} g^{ll} g^{jj} g_{kk} + \tfrac{1}{2}\alpha^i_{jkl} g^{ll} g_{ii} g^{jj} + \alpha^j_{ikl} g^{ll}],$$

so that $B^3(T, P)$ *is the set of all the 3-cochains* $C^3(T, T)$.

Cocycles $Z^3(T, P)$: The condition $f \in Z^3(T, P)$ only concerns the part f_L which characterizes the cohomology class.

2. $H^3(T, P)^P$

Exactly in the same way as for $H^2(T, P)^P$, there is only one condition of invariance: $(g \cdot f)(x_1, x_2, x_3) = 0$ with

$$\begin{cases} g \in L \\ f = f_L: T \wedge T \to L, \end{cases}$$

which means that

$$f_L \in \mathfrak{L}\left(\overset{A}{\underset{i=1}{\otimes}} T_i, L\right)^L,$$

or equivalently

$$f_L \in (\underbrace{D^{\frac{1}{2},\frac{1}{2}} \otimes D^{\frac{1}{2},\frac{1}{2}} \otimes D^{\frac{1}{2},\frac{1}{2}}}_{\text{Antisym}} \otimes D^{1,0} \otimes D^{0,1})^L$$

or

$$f_L \in (D^{\frac{1}{2},\frac{1}{2}} \otimes D^{1,0} \otimes D^{0,1})^L = 0,$$

because the trivial component $D^{0,0}$ is not contained in the previous product.

This proves that $H^3(P, P) = H^3(T, P)^P = 0$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Asymptotic Theory of Electromagnetic and Acoustic Diffraction by Smooth Convex Surfaces of Variable Curvature

SOONSUNG HONG
Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 5 October 1966)

A general method is presented for obtaining successive terms in short wavelength asymptotic expansions of the diffracted field produced by plane acoustic and electromagnetic waves incident on an arbitrary smooth convex surface. By introducing the geodesic coordinate system on arbitrary surfaces of non-constant curvature, both scalar and vector integral equations governing the surface fields are solved directly. The expressions for leading and second-order terms in the asymptotic expansion of the diffracted fields are obtained explicitly and the differences between acoustic and electromagnetic creeping waves are shown.

1. INTRODUCTION

A SHADOW is formed when a wave is incident upon a smooth convex body which is large compared to the incident wavelength. In the neighborhood of the shadow boundary the surface field does not abruptly vanish and a penumbra region exists. Some waves penetrate into the shadow region and account for the nonzero fields there. These phenomena are due to diffraction of the incident wave by the object.

The mathematical problem of analyzing diffraction of waves involves finding the short-wavelength asymptotic form of a solution of the wave equation satisfying an appropriate boundary condition of the diffracting surface and the radiation condition at infinity. Detailed studies of the surface field on a circular cylinder and a sphere,[1] for which the exact solutions are available, indicate that the incident wave is diffracted near the shadow boundary and the diffracted waves proceed along the geodesic into the shadow region, spilling off energy as they travel. Their phases are determined primarily by the distance traveled from the shadow boundary. The waves diffracted by a smooth convex surface are frequently called creeping waves.

In obtaining a description of the waves diffracted by an arbitrary smooth convex surface of variable curvature, two techniques can be used:

(1) finding the asymptotic form of an exact solution for a canonical body and generalizing the results;

(2) solving the boundary value problem directly by an asymptotic method for a general surface but in restricted regions.

The difficulty with the first method is that very few canonical problems can be solved exactly. Thus, in the well-known geometrical theory of diffraction,[2] a locally cylindrical body is chosen as the canonical body in analyzing diffraction of waves by arbitrary smooth surfaces. While this theory gives the correct leading term in the asymptotic expansion of the diffracted fields, it does not yield higher-order terms.

The purpose of this paper is to discuss an integral equation approach based on the second technique. It can yield not only the leading term but also higher-order terms in the asymptotic expansion of the fields diffracted by an arbitrary shape with a smooth convex surface. The method to be used is the following. The geodesic coordinate system is introduced to describe the geometry of the diffracting surface (Sec. 2). In terms of this coordinate system, the short-wavelength asymptotic form of the integral equation governing the surface fields is derived (Sec. 3.1) for the acoustic case, and its solutions are derived for the penumbra (Sec. 3.2) and shadow (Sec. 3.3) regions. The same procedure is repeated for the electromagnetic case (Sec. 4).

2. GEODESIC COORDINATE SYSTEM

From the analysis of the sphere solution,[1] it is observed that the creeping waves propagate along the geodesic. Thus, we propose to use the geodesic coordinate system to describe the diffracting surface. An important advantage of this coordinate system is that it can be defined on any smooth surface. For the sake of simplicity, it is assumed that the diffracting surface is symmetric with respect to the shadow boundary and that the torsion of the geodesic is zero. (See Fig. 1.)

Let us define the geodesic coordinate system as follows. The curve $u = 0$ is taken to represent the shadow boundary with v denoting arc length along it. At each point of $u = 0$ the incident wave is tangent in a given direction, and this defines a geodesic through each point of $u = 0$; these geodesics are taken as the

[1] W. Franz, Z. Naturforsch. **9A**, 705 (1954).
[2] B. R. Levy and J. B. Keller, Commun. Pure Appl. Math. **12**, 159 (1959).

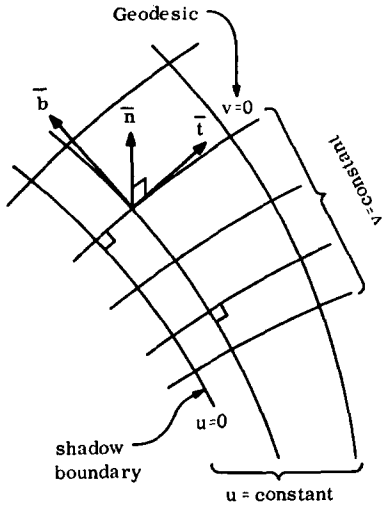


FIG. 1. Geodesic coordinate system.

coordinate curves $v = \text{const}$, with u taken as arc length along the geodesic measured positively from the shadow boundary. The geodesic coordinate system is orthogonal and the linear element is given as

$$ds^2 = du^2 + G\,dv^2, \quad \text{with} \quad G(u = 0) \equiv 1. \tag{2.1}$$

Because of the assumption that the geodesics are planar, G is independent of v.

The diffracting surface may be described by the Gauss–Weingarten equations[3]:

$$\begin{aligned} &\partial\mathbf{t}/\partial u = -\kappa_g\mathbf{n}, \\ &\partial\mathbf{t}/\partial v = \partial\mathbf{b}/\partial u = \kappa_{tt}\mathbf{b}, \\ &\partial\mathbf{b}/\partial v = -G[\kappa_{tt}\mathbf{t} + \kappa_{tn}\mathbf{n}] \text{ with } \kappa_t = (\kappa_{tt}^2 + \kappa_{tn}^2)^{\frac{1}{2}}, \\ &\partial\mathbf{n}/\partial u = \kappa_g\mathbf{t}, \quad \partial\mathbf{n}/\partial v = \kappa_{tn}\mathbf{b}, \end{aligned} \tag{2.2}$$

where

$$\partial\mathbf{r}/\partial u = \mathbf{t}, \quad \partial\mathbf{r}/\partial v = \mathbf{b} \quad \text{with } \mathbf{r} \text{ the position vector.} \tag{2.3}$$

Here $\mathbf{n}$, $\mathbf{t}$, and $\mathbf{b}/G^{\frac{1}{2}}$ are unit normal, tangent, and binormal vectors along the geodesic, respectively. κ_g is the curvature of the geodesic. κ_{tt} and κ_{tn} are, respectively, the tangential and the normal components of the curvature of the $u = \text{const}$ curves. Thus, the two principal curvatures are κ_g and κ_{tn} and their product is

$$\kappa_g\kappa_{tn} = -(1/G^{\frac{1}{2}})(\partial^2 G^{\frac{1}{2}}/\partial u^2), \tag{2.4}$$

while κ_{tt} is related to the function G by

$$\kappa_{tt} = (\partial G/\partial u)/2G. \tag{2.5}$$

In addition to Eq. (2.2), the Codazzi equation must be satisfied:

$$\begin{aligned} \partial\kappa_{tn}/\partial u &= \kappa_{tt}(\kappa_g - \kappa_{tn}), \\ \partial\kappa_g/\partial v &= 0. \end{aligned} \tag{2.6}$$

A more detailed analysis of the geodesic coordinate system can be found in most books on differential geometry.[3]

3. DIFFRACTION OF A PLANE ACOUSTIC (SCALAR) WAVE

The first problem to be considered is the diffraction of a plane acoustic (scalar) wave by an acoustically hard surface; i.e., a Neumann boundary condition is imposed.

3.1. Integral Equation Governing the Surface Field

We suppose that a plane acoustic wave is incident upon a smooth convex surface and that the normal derivative of the total field on the surface vanishes. Then the integral equation governing the surface field can be derived easily by Green's theorem[4]:

$$U(\mathbf{r}) = 2U_{\text{inc}}(\mathbf{r}) - \frac{1}{2\pi}\iint da' U(\mathbf{r}')\frac{1 - ikR}{R^3}\{\mathbf{n}(\mathbf{r}')\cdot\mathbf{R}\}e^{ikR}, \tag{3.1}$$

where $\mathbf{R} = \mathbf{r}' - \mathbf{r}$, and U_{inc} is the incident field. Without loss of generality, we consider the surface field on a geodesic which is called the curve $v = 0$. In terms of the geodesic coordinate system, the incident wave on the geodesic $v = 0$ is

$$U_{\text{inc}}(u, v = 0) = e^{ik\mathbf{t}(u=0,v=0)\cdot\mathbf{r}(u,v=0)}. \tag{3.2}$$

In the two equations above, the time dependence factor $e^{-i\omega t}$ is omitted. As observed in the study of a circular cylinder and a sphere,[1] the phase of the diffracted (creeping) wave is determined mainly by the distance traveled from the shadow boundary; thus we set

$$U(\mathbf{r}) = e^{iku}I(\mathbf{r}), \tag{3.3}$$

and for a large k ($= 2\pi/\lambda$, the wavenumber), $I(\mathbf{r})$ is assumed to be varying slowly in comparison with e^{iku}. Substitution of this expression into Eq. (3.1) gives

$$\begin{aligned} I(u, 0) = {} & 2\exp\{ik\mathbf{t}(0, 0)\cdot\mathbf{r}(u, 0) - iku\} \\ & - \frac{1}{2\pi}\iint G^{\frac{1}{2}}(u')\,du'\,dv' I(u', v')\frac{1 - ikR}{R^3} \\ & \times \{\mathbf{n}(u', v')\cdot\mathbf{R}\}\exp\{ikR - ik(u - u')\}. \end{aligned} \tag{3.4}$$

Since we are interested in the short-wavelength behavior of the solution, we replace the second term in Eq. (3.4) by its asymptotic form. For large k, the

[3] D. J. Struik, *Differential Geometry* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1950).

[4] H. Hönl, A. W. Maue, and K. Westpfahl, in *Handbuch der Physik* (Springer-Verlag, Berlin, 1961), Vol. XXV/1 pp. 218–544.

integrand has a saddle point where the derivative of the function $R - (u - u')$ vanishes. The Taylor series expansion of the vector $\mathbf{R}$ near $\mathbf{r} = \mathbf{r}'$ is derived easily by means of Eq. (2.2) and is given by

$$\begin{aligned}\mathbf{R} = \mathbf{r}'(u', v') - \mathbf{r}(u, 0) \simeq\ & (u' - u)\mathbf{t}(u) + v'\mathbf{b}(u) \\ & - \tfrac{1}{2}[(u' - u)^2\kappa_g(u)\mathbf{n}(u) - 2(u' - u)v'\kappa_{tt}(u)\mathbf{b}(u) \\ & + v'^2G(u)\{\kappa_{tt}(u)\mathbf{t}(u) + \kappa_{tn}(u)\mathbf{n}(u)\}] \\ & - \tfrac{1}{6}[(u' - u)^3\{\dot{\kappa}_g(u)\mathbf{n}(u) + \kappa_g^2(u)\mathbf{t}(u)\} \\ & + 3(u' - u)^2v'\kappa_g(u)\kappa_{tn}(u)\mathbf{b}(u) \\ & + 3(u' - u)v'^2G(u)\kappa_{tt}(u)\{\kappa_{tt}(u)\mathbf{t}(u) + \kappa_{tn}(u)\mathbf{n}(u)\} \\ & + v'^3G(u)k_t^2(u)\mathbf{b}(u)] + \cdots . \end{aligned} \tag{3.5}$$

Above, and in following pages, the curvatures (κ_g, etc.), $\mathbf{t}$, $\mathbf{b}$, and $\mathbf{n}$ without the argument for the v coordinate represents their values at $v = 0$. The dots denote the derivative with respect to the argument of the function. Using the above expression, the solution of the equation

$$\begin{aligned}&(\partial/\partial v')[(\mathbf{R}\cdot\mathbf{R})^{\frac{1}{2}} - (u - u')] \\ &\quad = (v'/R)G(u)[(u - u')\kappa_{tt}(u) + \cdots] = 0 \end{aligned} \tag{3.6}$$

yields the saddle point at $v' = 0$ for the v' integration. Applying the method of steepest descents[5] to the v' integration in Eq. (3.4), we obtain an asymptotic expression of the integral equation for large k.

$$\begin{aligned}I(u, 0) = 2\exp\{ik\mathbf{t}(0, 0)\cdot\mathbf{r}(u, 0) - iku\} - \frac{e^{(i\pi/4)}}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{u} du' e^{ikR_0 - ik(u-u')}\frac{[G(u')]^{\frac{1}{2}}}{[k(\partial^2R/\partial v'^2)(v' = 0)]^{\frac{1}{2}}} \\ \times\left[-ik\frac{\mathbf{n}(u', 0)\cdot\mathbf{R}_0}{R_0^2}I(u', 0) - \frac{\{\mathbf{n}(u', 0)\cdot\mathbf{R}_0\}(\partial^4R/\partial v'^4)(v' = 0)}{8R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}^2}I(u', 0)\right. \\ \left. + \frac{\kappa_{tn}(u')}{2R_0}I(u', 0) + \frac{\{\mathbf{n}(u', 0)\cdot\mathbf{R}_0\}}{2R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}}\frac{\partial^2I}{\partial v'^2}(u', 0)\right] + O(k^{-\frac{3}{2}}),\end{aligned} \tag{3.7}$$

where

$$\mathbf{R}_0 = \mathbf{r}'(u', 0) - \mathbf{r}(u, 0), \tag{3.8}$$

$$\begin{aligned}R_0 \simeq\ & (u - u') - \frac{\kappa_g^2(u)}{24}(u - u')^3 + \frac{\kappa_g(u)\dot{\kappa}_g(u)}{24}(u - u')^4 \\ & + \left[\frac{k_g^4(u)}{1920} - \frac{\dot{\kappa}_g^2(u)}{90} - \frac{\kappa_g(u)\ddot{\kappa}_g(u)}{80}\right](u - u')^5,\end{aligned} \tag{3.9}$$

$$\begin{aligned}&\frac{\partial^2R}{\partial v'^2}(v' = 0) \\ &\quad = \frac{G(u')}{R_0}[1 - \mathbf{R}_0\{\kappa_{tt}(u')\mathbf{t}(u') + \kappa_{tn}(u')\mathbf{n}(u')\}],\end{aligned} \tag{3.10}$$

and

$$\begin{aligned}\frac{\partial^4R}{\partial v'^4}(v' = 0) = -\frac{3G(u')}{R_0^2}\left[1 - \mathbf{R}_0\cdot\{\kappa_{tt}(u')\mathbf{t}(u')\right. \\ \left. + \kappa_{tn}(u')\mathbf{n}(u')\} + \frac{R_0^2\kappa_t^2(u')}{3}\right]\frac{\partial^2R}{\partial v'^2}(v' = 0).\end{aligned} \tag{3.11}$$

In Eq. (3.7), the contribution from the u' integration between u and infinity is neglected. The reason is the following. From Eq. (3.5) it can be shown that R_0 near $u = u'$ is

$$\begin{aligned}R_0 &\simeq (u - u') + \cdots, \quad \text{for} \quad u > u', \\ &\simeq (u' - u) + \cdots, \quad \text{for} \quad u < u'.\end{aligned}$$

Thus, $\exp[ik\{R_0 - (u - u')\}]$ in Eq. (3.7) has a saddle point at $u = u'$ only if $u \geqq u'$; therefore, by integrating by parts, one can show that the contribution from the region $u < u' \leq \infty$ is asymptotically negligible for large k.

Now the integral equation governing the surface field is thus reduced to a one-dimensional Volterra equation.

3.2. The Surface Field in the Penumbra Region

In this section, the asymptotic integral Eq. (3.7) governing the surface field is solved for the penumbra region. It is assumed that the curvatures are slowly varying and that ρ_g/ρ_{tn} is of order one or less. In order to obtain an appropriate form of Eq. (3.7) in the neighborhood of the shadow boundary, we set

$$\begin{aligned}M_0 &= [k\rho_g(0, 0)]^{\frac{1}{3}} \quad (\rho_g = 1/\kappa_g), \\ ku &= M_0^2\xi, \\ ku' &= M_0^2\tau,\end{aligned} \tag{3.12}$$

and further assume that $[k\rho_g(u, v)]^{\frac{1}{3}} \gg 1$. Near the shadow boundary ($u = 0$), the phase function, $\mathbf{t}(0, 0)\cdot\mathbf{r}(u, 0) - u$, of the incident wave term in Eq. (3.7) can be expanded in Taylor series by means of Eq. (2.2):

$$\begin{aligned}&\mathbf{t}(0, 0)\cdot\mathbf{r}(u, 0) - u \\ &\quad \simeq -\frac{u^3}{6}\kappa_g^2(0, 0) + \frac{u^5}{120}\kappa_g^4(0, 0)\{1 + 4\rho_g(0, 0)\ddot{\rho}_g(0, 0)\}\end{aligned} \tag{3.13}$$

[$\dot{\rho}_g(0, 0) = 0$ by assumption of symmetry of the diffracting surface with respect to the shadow boundary].

[5] L. M. Brekhovskikh, *Waves in Layered Media* (Academic Press Inc., New York, 1960), pp. 245–250.

Substitution of Eq. (3.12) into the above expression yields the asymptotic form of the incident wave:

$$e^{ik^t(0,0)\cdot r(u,0)-iku}$$

$$= e^{-i(\xi^3/6)}\left[1 + i\frac{\{1 + 4\rho_g(0,0)\ddot{\rho}_g(0,0)\}}{120M_0^2}\xi^5\right] + O(M_0^{-3}). \tag{3.14}$$

Similarly, an appropriate asymptotic form of the second term in the right-hand side of Eq. (3.7) can be derived easily by expanding the integrand near $u = u'$ and by substituting the relationships of (3.12) along with

$$\rho_g(u, 0) \simeq \rho_g(0, 0) + \tfrac{1}{2}u^2\ddot{\rho}_g(0, 0) \quad \text{near} \quad u = 0.$$

The high-frequency asymptotic form of the integral equation governing the surface field in the penumbra region is then

$$I(\xi, 0) = 2e^{-i(\xi^3/6)}\left[1 + i\frac{\{1 + 4\rho_g(0,0)\ddot{\rho}_g(0,0)\}}{120M_0^2}\xi^5\right] - \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\xi} d\tau I(\tau, 0)e^{-i[(\xi-\tau)^3/24]} \times \left[e^{-i(\pi/4)}\frac{(\xi-\tau)^{\frac{1}{2}}}{2} + \frac{K_2(\xi,\tau)}{M_0^2}\right] + O(M_0^{-3}), \tag{3.15}$$

where

$$K_2(\xi,\tau) = e^{-i(\pi/4)}(\xi-\tau)^{\frac{1}{2}}\left[-\frac{(\xi-\tau)^2}{96} + i\frac{(\xi-\tau)^5}{20} \times \left\{\frac{1}{192} + \frac{\rho_g(0,0)\ddot{\rho}_g(0,0)}{8}\right\} + \rho_g(0,0)\ddot{\rho}_g(0,0)\left\{\frac{\xi^2}{12} - \frac{\xi\tau}{3} - \frac{(\xi-\tau)^2}{8} + i\frac{\xi\tau(\xi-\tau)^3}{48}\right\} + \frac{\rho_g(0,0)}{8\rho_{tn}(0,0)}(\xi^2-\tau^2) + i\frac{\{\frac{3}{16} + [\rho_g(0,0)/2\rho_{tn}(0,0)]\}}{(\xi-\tau)}\right]. \tag{3.16}$$

Since there is no term of order M_0^{-1} in the above equation, we take the asymptotic expansion of I as

$$I(\xi, 0) = I_0(\xi, 0) + [I_1(\xi, 0)/M_0^2] + O(M_0^{-3}). \tag{3.17}$$

Substitution of this expression into Eq. (3.15) simplifies the integral equation and the following equations for I_0 and I_1 are obtained:

$$I_0(\xi, 0) = 2e^{-i(\xi^3/6)} - \frac{e^{-i(\pi/4)}}{4}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_{-\infty}^{\xi} d\tau I_0(\tau, 0)(\xi-\tau)^{\frac{1}{2}}e^{-i[(\xi-\tau)^3/24]} \tag{3.18}$$

and

$$I_1(\xi, 0) = i\frac{\{1 + 4\rho_g(0,0)\ddot{\rho}_g(0,0)\}}{60}\xi^5 e^{-i(\xi^3/6)} - \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\xi} d\tau I_0(\tau, 0)K_2(\xi,\tau)e^{-i[(\xi-\tau)^3/24]} - \frac{e^{-i(\pi/4)}}{4}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_{-\infty}^{\xi} d\tau I_1(\tau, 0)(\xi-\tau)^{\frac{1}{2}}e^{-i[(\xi-\tau)^3/24]}. \tag{3.19}$$

We observe, from the above two equations, that the kernel functions are the same and that substitution of the solution for the leading term I_0 yields the solution for the second-order term. Similarly, integral equations governing higher-order terms in the high-frequency expansion of the field can be derived by including further terms in the asymptotic expansion of Eqs. (3.7) and (3.15).

Since Eq. (3.18) is a Volterra type and its kernel is a function of $\xi - \tau$ only, the use of Fourier transform is suggested. We set

$$\tilde{I}_0(t) = \int_{-\infty}^{\infty} I_0(\xi, 0)e^{-i\xi t}\,dt. \tag{3.20}$$

Application of the Fourier transform to both sides of Eq. (3.18) and use of the convolution theorem yields

$$\tilde{I}_0(t) = 2\int_{-\infty}^{\infty} d\xi e^{-i\xi t - i(\xi^3/6)} \times \left[1 + \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{e^{-i(\pi/4)}}{4}\int_0^{\infty} d\xi\xi^{\frac{1}{2}}\cdot e^{-i\xi t - i(\xi^3/24)}\right]^{-1}. \tag{3.21}$$

The numerator of the above equation is an Airy function[6]

$$\int_{-\infty}^{\infty} d\xi e^{[-i\xi t - i(\xi^3/6)]} = 2^{\frac{4}{3}}\,\mathrm{Ai}\,(t2^{\frac{1}{3}}). \tag{3.22}$$

The denominator and other integrals for $\tilde{I}_1$ can be evaluated by means of the functions[7]

$$F_n(p) = \int_0^{\infty} dx x^{n-\frac{1}{2}}e^{-i(12)^{\frac{1}{3}}px - ix^3} \tag{3.23}$$

for various n. In particular,

$$F_0 = \pi^{\frac{3}{2}}2^{\frac{2}{3}}3^{-\frac{1}{6}}e^{i(\pi/4)}\,\mathrm{Ai}\,(p)[\mathrm{Ai}\,(p) - i\,\mathrm{Bi}\,(p)], \tag{3.24}$$

$$F_1 = \pi^{\frac{3}{2}}3^{-\frac{1}{2}}e^{3i(\pi/4)}[2\,\mathrm{Ai}\,(p)\{\dot{\mathrm{A}}\mathrm{i}\,(p) - i\,\dot{\mathrm{B}}\mathrm{i}\,(p)\} + (i/\pi)], \tag{3.25}$$

$$F_2 = 2\pi^{\frac{3}{2}}3^{-\frac{1}{2}}(12)^{-\frac{1}{3}}e^{5i(\pi/4)}[\{\dot{\mathrm{A}}\mathrm{i}\,(p)\}^2 + p\{\mathrm{Ai}\,(p)\}^2 - i\{p\,\mathrm{Ai}\,(p)\,\mathrm{Bi}\,(p) + \dot{\mathrm{A}}\mathrm{i}\,(p)\,\dot{\mathrm{B}}\mathrm{i}\,(p)\}]. \tag{3.26}$$

[6] J. C. P. Miller, *The Airy Integral* (Cambridge University Press, New York, 1946).

[7] V. H. Weston, "Pressure Pulse Received Due to an Explosion in the Atmosphere at an Arbitrary Altitude, Part 1," The University of Michigan Radiation Laboratory Technical Report 2886-1-T, C. 1-C. 4 (1960).

TABLE I. The function F_n.

n	C_n	D_n	E_n
3	2	$4p$	0
4	0	6	$4p$
5	$8p$	$16p^2$	10
6	28	$80p$	$16p^2$
7	$32p^2$	$108 + 64p^3$	$112p$
8	$288p$	$672p^2$	$220 + 64p^3$
9	$2912 + 512p^3$	$10048p + 1024p^4$	$3456p^2$

The remaining F_n may be expressed in terms of the above three functions by the relation

$$F_n(p) = C_n(12)^{-n/3}i^nF_0(p) + D_n(12)^{-(n-1)/3} \times i^{(n-1)}F_1(p) + E_n(12)^{-(n-2)/3}i^{(n-2)}F_2(p), \tag{3.27}$$

where C_n, D_n, and E_n are given in Table I.

Using the F_1 function, Eq. (3.21) reduces to

$$\tilde{I}_0(t) = 2^{\frac{1}{3}}\pi^{\frac{1}{2}}/\dot{w}_1(t2^{\frac{1}{3}}), \tag{3.28}$$

where

$$w_1 = i\pi^{\frac{1}{2}}[\mathrm{Ai} - i\,\mathrm{Bi}].$$

The inverse Fourier transform of Eq. (3.28) yields the solution for the leading term:

$$I_0(\xi, v = 0) = \frac{1}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp\,\frac{e^{ip\xi 2^{-\frac{1}{3}}}}{\dot{w}_1(p)} \quad (p = t2^{\frac{1}{3}}). \tag{3.29}$$

Similarly, the application of the Fourier transform to the integral equation governing the second-order term (3.19) gives

$$\tilde{I}_1(t) = \frac{\tilde{N}}{2\pi^{\frac{1}{2}}\,\mathrm{Ai}\,(p)\dot{w}_1(p)} \quad (p = t2^{\frac{1}{3}}), \tag{3.30}$$

where

$$\tilde{N} = \int_{-\infty}^{\infty} d\xi e^{-it\xi}\left[i\,\frac{\{1 + 4\rho_g(0,0)\ddot{\rho}_g(0,0)\}}{60}\,\xi^5 e^{-i(\xi^3/6)} - \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\xi} d\tau I_0(\tau,0)\mathrm{K}_2(\xi,\tau)e^{-i[(\xi-\tau)^3/24]}\right]. \tag{3.31}$$

Using the convolution theorem, the functions F_n of Table I and Eq. (3.28), one can show that

$$\frac{\hat{N}}{4\pi\,\mathrm{Ai}\,(p)} = p\left[-\frac{2}{15} + \rho_g(0,0)\ddot{\rho}_g(0,0)\frac{17}{15} + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right] + \frac{w_1(p)}{\dot{w}_1(p)}\left[-\frac{1}{5} + \frac{\beta^3}{30} + (\tfrac{8}{175} - \tfrac{6}{5}\beta^3)\rho_g(0,0)\ddot{\rho}_g(0,0)\right] - \left[\frac{pw_1(p)}{\dot{w}_1(p)}\right]^2\left[\tfrac{7}{3}\rho_g(0,0)\ddot{\rho}_g(0,0) + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right] + \rho_g(0,0)\ddot{\rho}_g(0,0)\frac{1}{3}\left[\frac{w_1(p)}{\dot{w}_1(p)}\right]^3 p^4. \tag{3.32}$$

Substitution of this expression into Eq. (3.30) and inverse Fourier transformation gives the solution for I_1. Combining I_1 with Eqs. (3.3) and (3.29), we obtain the desired expression for the surface field in the penumbra region:

$$U(u,0) \simeq e^{iku}\frac{1}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp\,\frac{e^{ip(u/d)}}{\dot{w}_1(p)} + \frac{e^{iku}}{[\frac{1}{2}k\rho_g(0,0)]^{\frac{2}{3}}\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp e^{ip(u/d)}\left[\frac{p}{\dot{w}_1(p)}\left\{-\frac{2}{15} + \rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\frac{17}{15} + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right\} + \frac{w_1(p)}{\{\dot{w}_1(p)\}^2}\left\{-\frac{1}{5} + \frac{p^3}{30} + \rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)(\tfrac{8}{175} - \tfrac{6}{5}p^3)\right\} - \left\{\frac{pw_1(p)}{\dot{w}_1(p)}\right\}^2 \times \left\{\tfrac{7}{3}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0) + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right\} + \frac{4}{3}\frac{p\{pw_1(p)\}^3}{\{\dot{w}_1(p)^4\}}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right], \tag{3.33}$$

where $d = [\lambda\rho_g^2(9,0)/\pi]^{\frac{1}{3}}$ and λ is the incident wavelength. When u/d is positive and sufficiently large (far away from the shadow boundary into the shadow region), Eq. (3.33) can be expressed as a rapidly convergent series in terms of the residues at the poles $\dot{w}_1(p) = 0$. This residue series represents the creeping waves. When u/d goes to negative infinity (illuminated region), Eq. (3.33) reduces to $2e^{ikt(0,0)\cdot r(u,0)}$, which is the geometrical optics term.[8] The width of the penumbra region is of order d.

33. The Surface Field in the Shadow Region

The incident plane wave cannot reach the shadow region directly (otherwise the shadow does not exist), and only the waves diffracted near the shadow boundary proceed into the shadow region. An expression for the surface field in this region may be obtained by following two steps[4]:

(1) Obtain the initial values of the diffracted (creeping) waves from the solution for the penumbra region at the shadow boundary.

(2) Solve the homogeneous integral equation [without the plane wave term in Eq. (3.7) and the limit of the integration only over the surface in the shadow region] and match the initial values at the shadow boundary.

The initial values of the diffracted waves may be obtained from Eq. (3.33). When u is positive, the integrals of this equation can be expressed in terms of the residues at the poles $\dot{w}_1(p) = 0$, and each

[8] N. A. Logan, "General Research in Diffraction Theory Vol. 1," Lockheed Missiles and Space Division Technical Report No. LMSD-288087 (1959).

residue represents a creeping wave.[9] The values of these residues at the shadow boundary ($u = 0$) yield the necessary initial values of the creeping waves. The residue series of Eq. (3.33) at the shadow boundary is

$$U(0,0) \sim 2\pi^{\frac{1}{2}} i \sum_{l=1}^{\infty} \frac{1}{p_l w_1(p_l)} \left[1 + \frac{1}{2\{\frac{1}{2}k\rho_g(0,0)\}^{\frac{1}{3}}} \times \left\{ p_l \left(-\frac{1}{15} + \frac{\rho_g(0,0)}{2\rho_{tn}(0,0)} + \tfrac{1}{90}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0) \right) + \frac{1}{p_l^2}\left(\frac{1}{5} + \frac{\rho_g(0,0)}{2\rho_{tn}(0,0)} - \tfrac{1}{30}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0) \right) \right\} \right], \tag{3.34}$$

where p_l is the lth root of $\dot{w}_1(p_l) = 0$. In the above expression, each term in the series represents the initial value (birth weight) of the lth mode of the creeping waves.

The short-wavelength asymptotic form of the homogeneous integral equation governing the surface field in the shadow region is the same as Eq. (3.7), except for the incident field term which now vanishes. Before attempting to solve the integral equation, we can observe that the common factor

$$[G(u')]^{\frac{1}{2}}[\partial^2 R/\partial v'^2]^{-\frac{1}{2}}$$

in the kernel of Eq. (3.7) behaves near the saddle point $u = u'$ as follows

$$[G(u')]^{\frac{1}{2}}\left[R_0 \frac{\partial^2 R}{\partial v'^2}(v' = 0)\right]^{-\frac{1}{2}} = [G(u')]^{\frac{1}{2}}[1 - \mathbf{R}_0 \cdot \{\kappa_{t\kappa}(u')\mathbf{t}(u') + \kappa_{tn}(u')\mathbf{n}(u')\}]^{-\frac{1}{2}} \simeq \left[\frac{G(u')}{G(u)}\right]^{\frac{1}{2}}. \tag{3.35}$$

[Refer to Eqs. (2.2), (2.4), (2.5), and (3.10).] The above relation indicates that the solution of the homogeneous integral equation has a factor $[G(u)]^{-\frac{1}{2}}$. In view of the phase factor for the solution in the penumbra region, Eq. (3.33), we set the solution of the homogeneous integral equation in the form

$$I(u,0) = A[G(u)]^{-\frac{1}{2}} \exp\left[ik \int_0^u ds \frac{2^{-\frac{1}{3}}}{M^2(s)} \times \left\{ \gamma_0(s) + \frac{\gamma_1(s)}{M(s)} + \frac{\gamma_2(s)}{M^2(s)} + \cdots \right\} \right], \tag{3.36}$$

where the constant A is the initial value, and

$$M(u) = [k\rho_g(u,0)]^{\frac{1}{3}}.$$

The propagation factors γ_0, γ_1, and γ_2 are yet to be determined. On setting

$$k(u - u') = M^2(u)\tau, \tag{3.37}$$

we obtain the following expression under the assumption that the curvatures are varying slowly:

$$k\int_{u'}^{u} \frac{ds}{M^2(s)} \left\{ \gamma_0(s) + \frac{\gamma_1(s)}{M(s)} + \frac{\gamma_2(s)}{M^2(s)} + \cdots \right\} \simeq \gamma_0(u)\tau + \frac{1}{M(u)}\left[\gamma_1(u)\tau - \frac{\tau^2}{2}\{\rho_g(u,0)\dot{\rho}_0(u) - \tfrac{2}{3}\dot{\rho}_g(u,0)\gamma_0(u)\} \right] + \frac{1}{M^2(u)}\left[\gamma_2(u)\tau + \frac{\tau^3}{6}\{\rho_g^2(u,0)\ddot{\gamma}_0(u) - \tfrac{2}{3}\rho_g(u,0)\ddot{\rho}_g(u,0)\gamma_0(u) - \tfrac{4}{3}\rho_g(u,0)\dot{\rho}_g(u,0)\dot{\gamma}_0(u) + \tfrac{10}{9}\dot{\rho}_g^2(u,0)\gamma_0(u)\} - \frac{\tau^2}{2}\{\rho_g(u,0)\dot{\gamma}_1(u) - \dot{\rho}_g(u,0)\gamma_1(u)\} \right] + O(M^{-3}). \tag{3.38}$$

Now combine Eq. (3.7), without the incident plane wave term, with Eq. (3.36) and expand the integrand near the saddle point $u = u'$ by Taylor series [using (Eq. 2.2)]. After these algebraic manipulations and making use of Eq. (3.37) and (3.38), we can obtain the following asymptotic homogeneous equation: for the propagation factors γ_i,

$$1 = -\frac{e^{-i(\pi/4)}}{2(2\pi)^{\frac{1}{2}}} \int_0^{\infty} d\tau \tau^{\frac{1}{2}} \left[1 - \frac{\dot{\rho}_g(u,0)}{M(u)}\left\{ -\tfrac{2}{3}\tau + i\frac{\tau^3}{24} \right\} + \frac{\tau^2}{M^2(u)}\left\{ -\frac{1}{48} + \frac{\dot{\rho}_g^2(u,0)}{2} - \frac{\rho_g(u,0)\ddot{\rho}_g(u,0)}{4} \right\} + i\frac{\tau^5}{M^2(u)}\left\{ \frac{1}{1920} - \tfrac{23}{360}\dot{\rho}_g^2(u,0) + \frac{\rho_g(u,0)\ddot{\rho}_g(u,0)}{80} \right\} - \frac{\tau^8}{1152M^2(u)}\dot{\rho}_g^2(u,0) + i\frac{\{\frac{3}{8} + [\rho_g(u,0)/\rho_{tn}(u,0)]\}}{M^2(u)\tau} \right] \times \left[1 + \frac{i2^{-\frac{1}{3}}}{M^2(u)}\left\{ \gamma_2(u)\tau + \frac{\tau^3}{6}(\rho_g^2(u,0)\ddot{\gamma}_0(u) - \tfrac{2}{3}\rho_g(u,0)\ddot{\rho}_g(u,0)\gamma_0(u) - \tfrac{4}{3}\rho_g(u,0)\dot{\rho}_g(u,0)\dot{\gamma}_0(u) + \tfrac{10}{9}\dot{\rho}_g^2(u,0)\gamma_0(u)) - \frac{\tau^2}{2}(\rho_g(u,0)\dot{\gamma}_1(u) - \dot{\rho}_g(u,0)\gamma_1(u)) \right\} \right] \exp\left[-i\frac{\tau^3}{24} + i\gamma_0(u)2^{-\frac{1}{3}}\tau + \frac{i2^{-\frac{1}{3}}}{M(u)} \times \left\{ \gamma_1(u)\tau - \frac{\tau^2}{2}(\rho_g(u,0)\dot{\gamma}_0(u,0) + 0(M^{-3}) - \tfrac{2}{3}\dot{\rho}_g(u,0)\gamma_0(u)) \right\} \right]. \tag{3.39}$$

[9] R. F. Goodrich, Trans. IRE-PGAP, AP-7, 528 (1959).

Upon comparing coefficients of the leading term in the expansion in powers of $1/M$, we obtain

$$1 = -\frac{e^{-i(\pi/4)}}{2(2\pi)^{\frac{1}{2}}}\int_0^\infty d\tau\tau^{\frac{1}{2}}\exp\left\{-i\frac{\tau^3}{24} - i\gamma_0(u)\tau 2^{-\frac{1}{3}}\right\}. \tag{3.40}$$

Comparing with Eq. (3.23), it can be shown that the right-hand side is related to the function F_1 of Eq. (3.25). Substitution of Eq. (3.23) and (3.25) into Eq. (3.40) yields the following:

$$\mathrm{Ai}\,(\gamma_0)\dot{w}_1(\gamma_0) = 0. \tag{3.41}$$

The solution of this equation determines γ_0. To be consistent with the initial values [Eq. (3.34)] of the creeping waves, the roots of $\dot{w}_1(\gamma_0) = 0$ must be chosen. In terms of the definition

$$\gamma_{0l} = e^{i(\pi/3)}\beta_l,$$

the various roots are given by Table II.

Since γ_0 is constant, comparison of the coefficients of $1/M$ in Eq. (3.39) yields

$$0 = -\frac{e^{-i(\pi/4)}}{2(2\pi)^{\frac{1}{2}}}\int_0^\infty d\tau\tau^{\frac{1}{2}}\left[\tfrac{2}{3}\dot{\rho}_g(u,0)\tau - i\frac{\dot{\rho}_g(u,0)}{24}\tau^4 - i\tau\left\{\gamma_1(u) + \frac{\gamma_0}{3}\dot{\rho}_g(u,0)\tau\right\}2^{-\frac{1}{3}}\right] \times \exp\left\{-i\gamma_0 2^{-\frac{1}{3}}\tau - i\frac{\tau^3}{24}\right\}. \tag{3.45}$$

Various integrals in Eq. (3.45) can be identified with Eq. (3.23). Thus, after substitution of F_n from Table II, we can evaluate γ_1 from Eq. (3.45), and the result is

$$\gamma_1(u) = i2^{\frac{1}{3}}[\tfrac{1}{6}\dot{\rho}_g(u,0)]. \tag{3.46}$$

TABLE II. The values of β_l and Ai $(-\beta_l)$.

l	β_l	Ai $(-\beta_l)$
1	1.01879	+0.53566
2	3.24820	−0.41902
3	4.82010	+0.38041

From the coefficients of M^{-2} in Eq. (3.39), one finds

$$\gamma_2(u)2^{-\frac{2}{3}}F_2(\gamma_0) = (12)^{-\frac{2}{3}}F_0(\gamma_0)\left[\frac{1}{5} + \frac{\rho_g(u,0)}{2\rho_{tn}(u,0)} - \frac{\rho_g(u,0)\ddot{\rho}_g(u,0)}{30} + \frac{\dot{\rho}_g^2(u,0)}{45}\right] + \gamma_0^2F_2(\gamma_0)[\tfrac{1}{60} - \tfrac{2}{45}\rho_g(u,0)\ddot{\rho}_g(u,0) + \tfrac{4}{135}\dot{\rho}_g^2(u.0)]. \tag{3.47}$$

Upon substituting the values of F_0 and F_2 given by Eqs. (3.24) and (3.26), we obtain γ_2, namely,

$$\gamma_2(u)2^{-\frac{2}{3}} = -\frac{1}{\gamma_0}\left[\frac{1}{10} + \frac{\rho_g(u,0)}{4\rho_{tn}(u,0)} - \frac{\rho_g(u,0)\ddot{\rho}_g(u,0)}{60} + \frac{\dot{\rho}_g^2(u,0)}{90}\right] + \gamma_0^2[\tfrac{1}{60} - \tfrac{2}{45}\rho_g(u,0)\ddot{\rho}_g(u,0) + \tfrac{4}{135}\dot{\rho}_g^2(u,0)]. \tag{3.48}$$

Combining Eqs. (3.36), (3.46), and (3.48) and matching the initial values given by Eq. (3.34) by letting $u = 0$, we obtain the desired solution for the surface field in the shadow region:

$$U(u,0) = \left[\frac{G(0)}{G(u)}\right]^{\frac{1}{4}}\left[\frac{\rho_g(0,0)}{\rho_g(u,0)}\right]^{\frac{1}{6}}e^{iku}\sum_{l=1}^{\infty}\frac{1}{\beta_l\mathrm{Ai}(-\beta_l)}\left[1 + \frac{e^{i(\pi/3)}}{2^{-\frac{2}{3}}M^2(0)}\left\{\beta_l\left(-\frac{1}{30} + \frac{\rho_g(0,0)}{4\rho_{tn}(0,0)} + \tfrac{1}{180}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right)\right.\right.$$
$$\left.\left. - \frac{1}{\beta_l^2}\left(\frac{1}{10} + \frac{\rho_g(0,0)}{4\rho_{tn}(0,0)} - \tfrac{1}{60}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right)\right\}\right]\exp\left[-e^{-i(\pi/6)}\beta_l\int_0^u\frac{ds}{\rho_g(s,0)}M(s)2^{-\frac{1}{3}}\right.$$
$$-e^{i(\pi/6)}\int_0^u\frac{ds}{\rho_g(s,0)}\frac{2^{\frac{1}{3}}}{M(s)}\left\{\frac{1}{\beta_l}\left(\frac{1}{10} + \frac{\rho_g(s,0)}{4\rho_{tn}(s,0)} - \frac{\rho_g(s,0)\ddot{\rho}_g(s,0)}{60} + \frac{\dot{\rho}_g^2(s,0)}{90}\right)\right.$$
$$\left.\left.\left. + \beta_l^2(\tfrac{1}{60} - \tfrac{2}{45}\rho_g(s,0)\ddot{\rho}_g(s,0) + \tfrac{4}{135}\dot{\rho}_g^2(s,0))\right\}\right], \tag{3.49}$$

where $M(u) = [k\rho_g(u,0)]^{\frac{1}{3}}$ and various values of β_l and $A_i(-\beta_l)$ are given in Table II. In deriving Eq. (3.49), the following relationships are used:

$$\gamma_{0l} = e^{i(\pi/3)}\beta_l \quad \text{and} \quad w_1(p) = e^{i(\pi/6)}2\pi^{\frac{1}{2}}\,\mathrm{Ai}\,(pe^{i(2\pi/3)}). \tag{3.50}$$

4. DIFFRACTION OF A PLANE ELECTROMAGNETIC (VECTOR) WAVE

The second problem investigated is the diffraction of a plane electromagnetic wave by a perfectly conducting smooth convex surface of nonconstant curvature. Since much of the analysis is similar to that which we have already discussed for the acoustic case, the details are omitted wherever possible.

4.1. Integral Equation Governing the Induced Currents on the Conducting Surface

If a plane electromagnetic wave is incident upon a smooth convex conducting surface, the integral equation governing the induced currents on the conductor

is[4]

$$\mathbf{J}(\mathbf{r}) = 2\mathbf{n}(\mathbf{r}) \times \mathbf{H}^{\text{inc}}(\mathbf{r}) - \frac{1}{2\pi}\mathbf{n}(\mathbf{r}) \times \iint da' \frac{1 - ikR}{R^3}\{\mathbf{J}(\mathbf{r}) \times \mathbf{R}e^{ikR}\}. \quad (4.1)$$

Here $\mathbf{H}^{\text{inc}}$ is the incident field and the time dependence factor $e^{-i\omega t}$ is omitted. Again, without loss of generality, we consider the induced current along the geodesic $v = 0$. The expression for the incident field is

$$\mathbf{H}^{\text{inc}}(u, 0) = [-\cos\theta_0\mathbf{b}(0, 0) + \sin\theta_0\mathbf{n}(0, 0)] \times e^{ik\mathbf{t}(0,0)\cdot\mathbf{r}(u,0)}, \quad (4.2)$$

where $\theta_0 = \sin^{-1}\{\mathbf{n}(0, 0) \times \mathbf{H}_0/|\mathbf{H}_0|\}$ is the polarization angle of the incident wave. With the substitution

$$\mathbf{J}(\mathbf{r}) = [\mathbf{t}(\mathbf{r})I_t(\mathbf{r}) + \mathbf{b}(\mathbf{r})I_b(\mathbf{r})]e^{iku}, \quad (4.3)$$

the vector integral Eq. (4.1) is reduced to two coupled scalar equations:

$$I_t(u, 0) = 2I_t^{\text{inc}}(u, 0) - \frac{1}{2\pi}\iint du'\,dv'[G(u')]^{\frac{1}{2}}\frac{1 - ikR}{R^3} \times \mathbf{t}(u)\cdot[\mathbf{n}(u) \times \{I_t(u', v')\mathbf{t}(u', v') \times \mathbf{R} + I_b(u', v')\mathbf{b}(u', v') \times \mathbf{R}\}]e^{ikR - ik(u-u')} \quad (4.4)$$

and

$$I_b(u, 0) = 2I_b^{\text{inc}}(u, 0) - \frac{1}{2\pi}\iint du'\,dv'[G(u')]^{\frac{1}{2}}\frac{1 - ikR}{R^3} \times \mathbf{b}(u)\cdot[\mathbf{n}(u) \times \{I_t(u', v')\mathbf{t}(u', v') \times \mathbf{R} + I_b(u', v')\mathbf{b}(u', v') \times \mathbf{R}\}]e^{ikR - ik(u-u')}. \quad (4.5)$$

The two equations above are similar to that of acoustic case, Eq. (3.4). They also have saddle points at $v' = 0$ for the v' integration and at $u = u'$ for the u' integration along the $v' = 0$ curve. Performing the v' integration by the method of steepest descents, we obtain

$$I_t(u, 0) = 2I_t^{\text{inc}}(u, 0) - \frac{1}{2\pi}\int_{-\infty}^{u} du'[G(u')]^{\frac{1}{2}}\left(\frac{2\pi i}{k}\right)^{\frac{1}{2}}\frac{1}{[(\partial^2R/\partial v'^2)(v' = 0)]^{\frac{1}{2}}}\left[\frac{-ikI_t(u', 0)}{R_0^2}\mathbf{t}(u)\cdot\{\mathbf{n}(u) \times (\mathbf{t}(u') \times \mathbf{R}_0)\}\right.$$
$$- \frac{(\partial^4R/\partial v'^4)(v' = 0)}{8R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}^2}\mathbf{t}(u)\cdot\{\mathbf{n}(u) \times (\mathbf{t}(u') \times \mathbf{R}_0)\}I_t(u', 0) - \frac{I_t(u', 0)}{2R_0^2}\frac{G(u')}{\{(\partial^2R/\partial v'^2)(v' = 0)\}}\mathbf{t}(u)$$
$$\cdot\{(\mathbf{n}(u)\cdot\mathbf{R}_0\kappa_{tt}(u') - \mathbf{n}(u)\cdot\mathbf{t}(u'))\kappa_t(u')\mathbf{N}(u') + (\mathbf{t}(u') - \kappa_{tt}(u')\mathbf{R}_0)\kappa_t(u')\mathbf{n}(u)\cdot\mathbf{N}(u')\}$$
$$+ \frac{(\partial^2I_t/\partial v'^2)(u', 0)}{2R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}}\mathbf{t}(u)\cdot\{\mathbf{n}(u) \times (\mathbf{t}(u') \times \mathbf{R}_0)\} + \frac{(\partial I_b/\partial v')(u', 0)}{\{(\partial^2R/\partial v'^2)(v' = 0)\}R_0^2}G(u')\mathbf{t}(u)$$
$$\left.\cdot\{\mathbf{R}_0(\mathbf{n}(u)\cdot\mathbf{N}(u'))\kappa_t(u') - \kappa_t(u')\mathbf{N}(u')(\mathbf{n}(u)\cdot\mathbf{R}_0)\}\right] + O(k^{-\frac{3}{2}}) \quad (4.6)$$

and

$$I_b(u, 0) = 2I_b^{\text{inc}}(u, v = 0) - \frac{1}{2\pi}\int_{-\infty}^{u} du'[G(u')]^{\frac{1}{2}}\left(\frac{2\pi i}{k}\right)^{\frac{1}{2}}\frac{1}{[(\partial^2R/\partial v'^2)(v' = 0)]^{\frac{1}{2}}}$$
$$\times\left[-ik\frac{\{\mathbf{n}(u)\cdot\mathbf{R}_0\}\{\mathbf{b}(u)\cdot\mathbf{b}(u')\}}{R_0^2} - \frac{\{(\partial^4R/\partial v'^4)(v' = 0)\}\{\mathbf{n}(u)\cdot\mathbf{R}_0\}\{\mathbf{b}(u)\cdot\mathbf{b}(u')\}}{8R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}^2}\right.$$
$$\times I_b(u', 0) - \frac{G(u')I_b(u', 0)}{\{(\partial^2R/\partial v'^2)(v' = 0)\}2R_0^2}\{\mathbf{b}(u)\cdot\mathbf{b}(u')\}\{\kappa_t^2(u')\mathbf{n}(u)\cdot\mathbf{R}_0 - \kappa_t(u')\mathbf{n}(u)\cdot\mathbf{N}(u')\}$$
$$+ \frac{(\partial^2I_b/\partial v'^2)(u', 0)}{2R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}}\{\mathbf{b}(u)\cdot\mathbf{b}(u')\}\{\mathbf{n}(u)\cdot\mathbf{R}_0\} + \frac{\partial I_t}{\partial v'}(v' = 0, u')$$
$$\left.\times\frac{\{\mathbf{b}(u)\cdot\mathbf{b}(u')\}}{R_0^2\{(\partial^2R/\partial v'^2)(v' = 0)\}}\{\kappa_{tt}(u')\mathbf{n}(u)\cdot\mathbf{R}_0 - \mathbf{n}(u)\cdot\mathbf{t}(u')\}\right] + O(k^{-\frac{3}{2}}), \quad (4.7)$$

where

$$\kappa_t(u)\mathbf{N}(u) \equiv \kappa_{tt}(u)\mathbf{t}(u) + \kappa_{tn}(u)\mathbf{n}(u).$$

4.2. Induced Currents in the Penumbra and the Shadow Regions

In this section, Eqs. (4.6) and (4.7) are solved by the same technique used in the acoustic case. In the penumbra region, substitution of Eq. (3.12) into Eq. (4.2) gives the asymptotic form of the incident field:

$$I_t^{\text{inc}}(u, v = 0) = \mathbf{t}(u)\cdot[\mathbf{n}(u) \times \mathbf{r}'_{\text{inc}}]e^{-iku}$$
$$= \cos\theta_0 e^{-i(\xi^3/6)}\left[1 + i\frac{\{1 + 4\rho_g(0)\ddot{\rho}_g(0)\}}{120M_0^2}\xi^5\right] + O(M_0^{-3}) \quad (4.8)$$

and

$$I_b^{\text{inc}}(u, v = 0) = \mathbf{b}(u)\cdot[\mathbf{n}(u) \times \mathbf{H}_{\text{inc}}]e^{-iku} = -\sin\theta_0\frac{\xi}{M_0}e^{-i(\xi^3/6)} + O(M_0^{-3}). \quad (4.9)$$

Combining Eqs. (3.12), (4.6), and (4.8), we have

$$I_t(\xi, 0) = 2\cos\theta_0 e^{-i(\xi^3/6)}\left[1 + i\frac{\{1 + 4\rho_g(0)\ddot{\rho}_g(0)\}}{120M_0^2}\xi^5\right] - \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\xi} d\tau I_t(\xi, 0)e^{-i[(\xi-\tau)^3/24]} \times\left\{e^{-i(\pi/4)}\frac{(\xi - \tau)^2}{2} + \frac{K_{2t}(\xi, \tau)}{M_0^2}\right\} + O(M_0^{-3}), \quad (4.10)$$

where

$$K_{2t}(\xi,\tau) = e^{-i(\pi/4)}(\xi-\tau)^{\frac{1}{2}} \times \left[-\frac{(\xi-\tau)^2}{96} + i\frac{(\xi-\tau)^5}{20}\left\{\frac{1}{192} + \frac{\rho_g(u=0)\ddot{\rho}_g(0)}{8}\right\} + \rho_g(0)\ddot{\rho}_g(0)\left\{\frac{\xi^2}{12} - \frac{\xi\tau}{3} - \frac{(\xi-\tau)^2}{8} + \frac{i}{48}(\xi-\tau)^3\xi\tau\right\} + \frac{1}{8}\frac{\rho_g(0)}{\rho_{tn}(0)}(\xi^2-\tau^2) + i\frac{(\frac{3}{16} - \rho_g(0)/2\rho_{tn}(0))}{(\xi-\tau)}\right]. \quad (4.11)$$

Now, the above asymptotic form of the integral equation governing I_t is independent of I_b, and thus the original coupled vector integral Eq. (4.1) is decoupled in the asymptotic sense for large k.

Upon comparing Eq. (4.10) with acoustic Eq. (3.15), we can easily observe that the only difference between the two equations is the sign of the term $\rho_g(0,0)/2\rho_{tn}(0,0)(\xi-\tau)$ in K_{2t} and K_2. Thus, we can obtain immediately the solution for I_t from the acoustic solution given by Eq. (3.33).

The asymptotic form of the integral equation governing I_b in the penumbra region is

$$I_b(\xi,0) = -2\sin\theta_0\frac{\xi}{M_0}e^{-i(\xi^3/6)} + \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\xi} d\tau I_b(\tau,0)e^{-i[(\xi-\tau)^3/24]} \times \left[\frac{e^{-i(\pi/4)}}{2}(\xi-\tau)^{\frac{1}{2}} + O(M_0^{-2})\right] + O(M_0^{-3}). \quad (4.12)$$

On setting

$$I_b(\xi,0) = [I_{0b}(\xi,0)/M_0] + O(M_0^{-3}), \quad (4.13)$$

application of the Fourier transform and the F_n function [Eq. (3.23)] yields:

$$\tilde{I}_{b0} = -i\sin\theta_0 4 2^{\frac{2}{3}}\pi\,\dot{\mathrm{Ai}}(p) + \tilde{I}_{b0} \times [-1 + i2\pi\,\mathrm{Ai}(p)\{\dot{\mathrm{Ai}}(p) - i\,\dot{\mathrm{Bi}}(p)\}]. \quad (4.14)$$

Substituting the Wronskian relation

$$\mathrm{Ai}\,\dot{\mathrm{Bi}} - \mathrm{Bi}\,\dot{\mathrm{Ai}} = 1/\pi$$

into Eq. (4.14), the inverse Fourier transform gives the solution for I_{b0}:

$$I_b(\xi,0) = \frac{i\sin\theta_0}{2^{-\frac{1}{3}}M_0\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp\,\frac{e^{i\xi 2^{-\frac{1}{3}}p}}{w_1(p)}. \quad (4.15)$$

Thus, combining solutions for I_t and I_b, we obtain the expression for the induced currents in the penumbra region as

$$\mathbf{J}(u,0) = \mathbf{t}(u,0)\cos\theta_0 e^{iku}\frac{1}{\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp\,\frac{e^{ip(u/d)}}{\dot{w}_1(p)} + \mathbf{t}(u,0)\cos\theta_0 e^{iku}\frac{1}{\pi^{\frac{1}{2}}2^{-\frac{2}{3}}M_0^2}\int_{-\infty}^{\infty} dp e^{ip(u/d)} \times \left[\frac{p}{\dot{w}_1(p)}\left\{-\frac{2}{15} + \rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\frac{17}{15} + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right\} + \frac{w_1(p)}{\{\dot{w}_1(p)\}^2}\left\{-\frac{1}{5} + \frac{p^3}{30}\frac{\rho_g(0,0)}{\rho_{tn}(0,0)} + \rho_g(0,0) \times \frac{d^2\rho_g}{du^2}(0,0)(\tfrac{8}{15} - \tfrac{6}{5}p^3)\right\} - \frac{\{pw_1(p)\}^2}{\{\dot{w}_1(p)\}^3} \times \left\{\tfrac{7}{3}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0) + \frac{\rho_g(0,0)}{\rho_{tn}(0,0)}\right\} + \frac{4}{3}\frac{p\{pw_1(p)\}^3}{\{\dot{w}_1(p)\}^4}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right] + \mathbf{b}(u,0)\frac{i\sin\theta_0}{2^{-\frac{1}{3}}M_0\pi^{\frac{1}{2}}}\int_{-\infty}^{\infty} dp\,\frac{e^{ip(u/d)}}{w_1(p)} + O(M_0^{-3}), \quad (4.16)$$

where

$$d = [\lambda\rho_g^2(0,0)/\pi]^{\frac{1}{3}}.$$

Due to the similarity of asymptotic forms of electromagnetic and acoustic integral equations, solutions for the shadow region can be obtained by the same method used in the acoustic case. The induced current in the shadow region is

$$\mathbf{J}(u,0) = \mathbf{t}(u,0)\cos\theta_0\left[\frac{G(0)}{G(u)}\right]^{\frac{1}{4}}\left[\frac{\rho_g(0,0)}{\rho_g(u,0)}\right]^{\frac{1}{6}}e^{iku}\sum_{l=1}^{\infty}\frac{1}{\beta_l\,\mathrm{Ai}(-\beta_l)} \times \left[1 + \frac{e^{i(\pi/3)}}{2^{-\frac{2}{3}}M^2(0)}\left\{\beta_l\left(-\frac{1}{30} + \frac{\rho_g(0,0)}{4\rho_{tn}(0,0)} + \tfrac{1}{180}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right) - \frac{1}{\beta_l^2}\left(\frac{1}{10} - \frac{\rho_g(0,0)}{4\rho_{tn}(0,0)} - \tfrac{1}{60}\rho_g(0,0)\frac{d^2\rho_g}{du^2}(0,0)\right)\right\}\right] \exp\left[-e^{-i(\pi/6)}\beta_l\int_0^u\frac{ds}{\rho_g(s,0)}2^{-\frac{1}{3}}M(s) - e^{i(\pi/6)}\int_0^u\frac{ds}{\rho_g(s,0)2^{-\frac{1}{3}}M(s,0)}\left\{\frac{1}{\beta_l}\left(\frac{1}{10} - \frac{\rho_g(s,0)}{4\rho_{tn}(s,0)} - \frac{\rho_g(s,0)(d^2\rho_g/ds^2)(s,0)}{60} + \frac{[(d\rho_g/ds)(s,0)]^2}{90}\right) + \beta_l^2\left(\frac{1}{60} - \tfrac{2}{45}\rho_g(s,0)\frac{d^2\rho_g}{ds^2}(s,0) + \frac{4}{135}\left(\frac{d\rho_g}{ds}(s,0)\right)^2\right)\right\}\right] + \mathbf{b}(u,0)\sin\theta_0\left[\frac{G(0)}{G(u)}\right]^{\frac{1}{4}}\left[\frac{\rho_g(0,0)}{\rho_g(u,0)}\right]^{\frac{1}{6}} \times \frac{e^{iku+i(\pi/6)}}{2^{-\frac{1}{3}}M(0)}\sum_{l=1}^{\infty}\frac{1}{\dot{\mathrm{Ai}}(-\alpha_l)}\exp\left[-e^{-i(\pi/6)}\alpha_l\int_0^u\frac{ds}{\rho_g(s,0)}2^{-\frac{1}{3}}M(s)\right] + O(M^{-3}). \quad (4.17)$$

Here β_l and Ai $(-\beta_l)$ are given in Table II, and α_l is the lth root of Ai $(-\alpha_l) = 0$ (see Table III).

TABLE III. The values of α_l and $\dot{\text{A}}\text{i}\,(-\alpha_l)$.

l	α_l	$\dot{\text{A}}\text{i}\,(-\alpha_l)$
1	2.33811	+0.70121
2	4.08795	−0.80311
3	5.52056	+0.86520

5. DISCUSSION

In both the acoustic and the electromagnetic diffraction problems considered, the short-wavelength asymptotic expressions for the surface fields have been obtained for the penumbra and the shadow regions. The second-order terms in the asymptotic expansion of the surface fields are new results. The leading terms are the same as those of Fock[10] and Levy and Keller.[2]

In the solutions for the shadow region, the factor $[G(0)/G(u)]^{\frac{1}{4}}$ is of interest. By definition of the function G [Eq. (2.1)], $G^{\frac{1}{2}}\,dv$ represents the width between the two adjacent geodesics. Thus, referring to the geometrical theory of diffraction,[2] $[G(0)/G(u)]^{\frac{1}{4}}$ represents the so-called ray convergence factor for the creeping waves. In the geometrical theory of diffraction, this factor was obtained by physical reasoning (conservation of energy), and in the present paper, this factor is justified mathematically. The leading term for the acoustic and electromagnetic creeping waves is the same as that predicted by the geometrical theory of diffraction. This leading term, except the factor $[G(0)/G(u)]^{\frac{1}{4}}$, is independent of curvature in the direction transverse to the geodesic.

In the solution of electromagnetic diffraction problems, it is shown that up to the terms of order $[k\rho_g]^{-\frac{3}{2}}$ in the asymptotic expansion, there is no coupling between tangential and binormal components of the creeping waves. However, identity between the acoustic creeping waves under Neumann boundary condition and the tangential component of the electromagnetic creeping waves is true only in the leading term. The transverse curvature appears in the second-order term. The effect of transverse curvature on the electromagnetic creeping waves differs from that on the acoustic creeping waves. This is one of the new results of the present investigation.

When the radius of curvature ρ_t in the transverse direction is infinite, the diffracting surface becomes cylindrical. In this case, the propagation factors of the creeping waves in the shadow region agree with those obtained by Franz and Klante,[11] and by Keller and Levy.[12] When the principal radii of curvature (ρ_g and ρ_{tn}) are the same and constant, the diffracting surface is spherical. In this case, the solutions of the creeping waves reduce to the results of Senior,[13] who obtained the creeping wave solution (including the second-order terms) for the sphere by means of a Watson transformation of the Mie series (exact) solution.

The solutions for the shadow regions are not valid near a caustic where the radius of curvature (ρ_t) in the direction transverse to the geodesic is no longer large compared to the incident wavelength. The author feels that the integral equation method used here will still be applicable in investigating the surface fields near the caustic, provided that the saddle point integration for the v' coordinate (Secs. 3.1 and 4.1) is modified by some suitable means.

ACKNOWLEDGMENTS

The author is grateful to Dr. V. H. Weston and Dr. R. F. Goodrich for their many valuable suggestions, and to L. P. Zukowski for checking the algebra. He is also indebted to many members of The University of Michigan Radiation Laboratory for comments which were of value in preparing the manuscript.

The research reported in this paper was sponsored by the USAF Ballistic Systems Division, Norton AFB, California, under Contract AF 04(694)-834. This work was done while the author was with the University of Michigan Radiation Laboratory.

[10] V. Fock, J. Phys. **10**, 130 (1949).

[11] W. Franz and K. Klante, Trans. IRE-PGAP **AP-7**, 568 (1959).

[12] J. B. Keller and B. R. Levy, IRE Trans. Antennas Propagation **AP-7**, 552 (1959).

[13] T. B. A. Senior, The University of Michigan Radiation Laboratory Note (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Lowering and Raising Operators for the Orthogonal Group in the Chain $O(n) \supset O(n-1) \supset \cdots$, and their Graphs*

SING CHIN PANG AND K. T. HECHT
Physics Department, University of Michigan, Ann Arbor, Michigan

(Received 4 August 1966)

Normalized lowering and raising operators are constructed for the orthogonal group in the canonical group chain $O(n) \supset O(n-1) \supset \cdots \supset O(2)$ with the aid of graphs which simplify their construction. By successive application of such lowering operators for $O(n)$, $O(n-1)$, $\cdots$ on the highest weight states for each step of the chain, an explicit construction is given for the normalized basis vectors. To illustrate the usefulness of the construction, a derivation is given of the Gel'fand–Zetlin matrix elements of the infinitesimal generators of $O(n)$.

1. INTRODUCTION

THE semisimple Lie groups have recently found many new applications in physics. The unitary groups in particular have received wide attention as a result of this renewed importance, and the irreducible representations of $U(n)$ (arbitrary n), have been studied in considerable detail.[1,2] Although the orthogonal group $O(n)$ has received less attention, it recently also found some new applications to physical problems. In particular, the groups $O(5)$ and $O(8)$ have become of interest in nuclear spectroscopy in connection with the quasi-spin formalism for neutron and proton configurations.[3,4] The group chain $O(n) \supset O(n-1) \supset \cdots$ has also been found of interest in general many-body theory in the construction of n-body states of definite permutational symmetry.[5]

The basis vectors of an arbitrary irreducible representation of $O(n)$ are completely characterized by the chain of canonical subgroups $O(n-1) \supset O(n-2) \supset \cdots O(2)$. This canonical group chain has been studied many years ago by Gel'fand and Zetlin,[6] who give the matrix elements of the infinitesimal operators of $O(n)$, for arbitrary n, in this basis.[7] Since the mathematically natural chain of subgroups, such as $O(n) \supset O(n-1) \supset \cdots$, often does not include the subgroups of actual physical interest,[3,4] the application to physical problems, in general, involves a transformation from the mathematically natural to a physically relevant scheme. To effect such a transformation, it becomes important to have an explicit construction of the basis vectors of an arbitrary irreducible representation of the group.

It is the purpose of this paper to give an explicit construction of the basis vectors of the irreducible representations of $O(n)$ in the Gel'fand scheme through the successive application of lowering operators acting on the highest weight state. The concept of lowering (or raising) operators was employed by Nagel and Moshinsky[1] to construct the full set of basis vectors of $U(n)$ in the canonical group chain $U(n) \supset U(n-1) \supset \cdots$. Although the present work has set itself the analogous task for the group chain $O(n) \supset O(n-1) \cdots$ and thus forms a parallel to the work of Nagel and Moshinsky, the techniques employed are somewhat different. In particular, since the lowering (or raising) operators for $O(n)$ are complicated polynomial functions of the infinitesimal generators of the group, a graphical technique has been found useful in the construction of these operators.

In Sec. 2 a review is given of some of the properties of the group $O(n)$ and the canonical chain of subgroups employed in the Gel'fand basis. In Sec. 3 the raising and lowering operators are constructed with the aid of graphs. Section 4 presents the calculation of the normalization coefficients of the lowering operators. These are the fundamental numbers of the construction since the successive application of lowering operators must yield a normalized basis vector for easy application in actual problems. Finally, in Sec. 5, a brief derivation is given of the Gel'fand and Zetlin results for the matrix elements of the infinitesimal operators to illustrate the usefulness of the present construction.

* Supported by the U.S. Office of Naval Research, Contract NONR 1224(59).

[1] J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).

[2] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963); G. E. Baird and L. C. Biedenharn, *ibid.* **4**, 1449 (1963). For earlier references consult these references.

[3] B. H. Flowers and S. Szpikowski, Proc. Phys. Soc. (London) **84**, 193 (1964); J. C. Parikh, Nucl. Phys. **63**, 214 (1965). J. N. Ginocchio, *ibid.* **74**, 321 (1965); M. Ichimura, Progr. Theoret. Phys. (Kyoto) **32**, 757 (1964); **33**, 215 (1965). K. T. Hecht, Phys. Rev. **139**, B794 (1965).

[4] B. H. Flowers and S. Szpikowski, Proc. Phys. Soc. (London) **84**, 673 (1964).

[5] P. Kramer and M. Moshinsky, Nucl. Phys. **82**, 241 (1966).

[6] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk. USSR **71**,, 1017 (1950). I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro *Representations of the Rotation and Lorentz Groups and Their Application* (The Macmillan Company, New York, 1963), p. 353.

[7] The Gel'fand–Zetlin result has also been derived by algebraic techniques by J. D. Louck, Los Alamos Scientific Laboratory Reports LA 2451 (1960).

2. SOME PROPERTIES OF $O(n)$

A. Generators of $O(n)$

The natural infinitesimal generators of $O(n)$ are the set of skew-symmetric, Hermitian operators J_{ij} with the commutation relations

$$[J_{mj}, J_{kl}] = i(\delta_{mk}J_{jl} + \delta_{jl}J_{mk} - \delta_{jk}J_{ml} - \delta_{ml}J_{jk}), \quad (2.1)$$

where m, j, k, and l run from 1 to n. The number of independent generators of $O(n)$ is therefore $\frac{1}{2}n(n-1)$.

The infinitesimal generators of a Lie group are best expressed in standard form[8] in which they are organized into one set of k commuting operators (H type), where k is the rank of the group, and a set of raising and lowering generators[9] (E type). In $O(3)$, for example, H, E_1, E_{-1} correspond to J_{12}, $J_{13} + iJ_{23}$, $J_{13} - iJ_{23}$, respectively. For both $O(2k+1)$ and $O(2k)$ it is convenient to choose the k commuting operators as $J_{12}, J_{34}, J_{56}, \cdots, J_{2k-1,2k}$. It is useful to further classify the raising and lowering generators into two types, those which connect the group $O(n)$ to its subgroups, to be denoted by Q, and those which operate within the space of the subgroups only, to be denoted by ρ, so that there are three types of operators in all. In $O(7)$, for example, operators of type Q are linear combinations of the J_{i7}, while operators of type ρ involve only J_{ij} with *both* $i, j < 7$.

The three types of operators are defined as follows:

(a) $O(2k+1)$

Type (1) $H_\alpha = J_{2\alpha-1,2\alpha}, \quad \alpha = 1, 2, \cdots, k,$

(2) $Q_{2k+1,\pm\alpha} = J_{2\alpha-1,2k+1} \pm iJ_{2\alpha,2k+1}, \quad \alpha = 1, 2, \cdots, k, \quad (2.2)$

(3) $\rho_{\alpha\beta} = [Q_{2k+1,\alpha}, Q_{2k+1,\beta}],$
$\alpha, \beta = \pm 1, \cdots, \pm k, \quad \beta \neq -\alpha;$

(b) $O(2k)$

Type (1) $H_\alpha = J_{2\alpha-1,2\alpha}, \quad \alpha = 1, 2, \cdots, k-1,$

(2) $Q_{2k,k} = J_{2k-1,2k} \quad (=H_k),$
$Q_{2k,\pm\alpha} = J_{2\alpha-1,2k} \pm iJ_{2\alpha,2k}, \quad \alpha = 1, 2, \cdots, k-1, \quad (2.3)$

(3) $\rho_{\alpha\beta} = [Q_{2k,\alpha}, Q_{2k,\beta}],$
$\alpha, \beta = \pm 1, \pm 2, \cdots, \pm(k-1), k, \quad \beta \neq -\alpha,$

(Note that H_k is now included among the type 2 operators, and that $\rho_{\alpha,-\alpha}$ is not of type 3 but, from its definition, is merely equal to $2J_{2\alpha-1,2\alpha}$.) The basic commutators of these operators are then

$$[J_{2\alpha-1,2\alpha}, J_{2\beta-1,2\beta}] = 0, \quad (2.4)$$

$$[J_{2\alpha-1,2\alpha}, \rho_{\beta\gamma}] = (\delta_{\alpha\beta} + \delta_{\alpha\gamma} - \delta_{\alpha,-\beta} - \delta_{\alpha,-\gamma})\rho_{\beta\gamma}, \quad (2.5)$$

$$[J_{2\alpha-1,2\alpha}, Q_{n\beta}] = (\delta_{\alpha\beta} - \delta_{\alpha,-\beta})Q_{n\beta}, \quad (2.6)$$

$$[\rho_{\alpha\beta}, Q_{n\gamma}] = 2(\delta_{\beta,-\gamma}Q_{n\alpha} - \delta_{\alpha,-\gamma}Q_{n\beta}), \quad (2.7)$$

$$[\rho_{\alpha\beta}, \rho_{\gamma\delta}] = 2(\delta_{\alpha,-\delta}\rho_{\beta\gamma} + \delta_{\beta,-\gamma}\rho_{\alpha\delta} - \delta_{\alpha,-\gamma}\rho_{\beta\delta} - \delta_{\beta,-\delta}\rho_{\alpha\gamma}). \quad (2.8)$$

The $\rho_{\alpha\beta}$ can also be represented as Q-type operators of the subgroups of $O(n)$

$$\begin{aligned} \rho_{\alpha\beta} &= i[Q_{2\beta-1,\alpha} + iQ_{2\beta,\alpha}], \\ \rho_{\alpha,-\beta} &= i[Q_{2\beta-1,\alpha} - iQ_{2\beta,\alpha}], \quad 0 < \alpha < \beta, \\ \rho_{-\alpha,\beta} &= i[Q_{2\beta-1,-\alpha} - iQ_{2\beta,-\alpha}], \quad 2\beta < n, \\ \rho_{-\alpha,-\beta} &= i[Q_{2\beta-1,-\alpha} + iQ_{2\beta,-\alpha}]. \end{aligned} \quad (2.9)$$

B. The Gel'fand Basis

Gel'fand and Zetlin[6] have provided a way to completely specify the basis vectors of the irreducible representations of $O(n)$ according to the canonical chain of subgroups $O(n) \supset O(n-1) \supset \cdots \supset O(2)$. For the case $n = 2k+1$

$$|M_{n\mu}\rangle = \begin{bmatrix} m_{2k+1,1} & m_{2k+1,2} & \cdots & m_{2k+1,k-1} & m_{2k+1,k} \\ m_{2k,1} & m_{2k,2} & \cdots & m_{2k,k-1} & m_{2k,k} \\ m_{2k-1,1} & m_{2k-1,2} & \cdots & m_{2k-1,k-1} & \\ \cdot & \cdot & \cdots & m_{2k-2,k-1} & \\ \cdot & \cdot & \cdots & & \\ \cdot & \cdot & \cdots & & \\ m_{51} & m_{52} & & & \\ m_{41} & m_{42} & & & \\ m_{31} & & & & \\ m_{21} & & & & \end{bmatrix}. \quad (2.10)$$

[8] G. Racah, CERN reprint 61-8 (1961).

[9] The raising and lowering *generators* are not to be confused with the raising and lowering *operators* which are the subject of this paper. Except for $O(3)$ the lowering and raising *operators* are complicated polynomial functions of the lowering and raising *generators*.

For the case $n = 2k$

$$|\mathcal{M}_{n\mu}\rangle = \begin{bmatrix} m_{2k,1} & m_{2k,2} & \cdots & m_{2k,k-1} & m_{2k,2} \\ m_{2k-1,1} & m_{2k-1,2} & \cdots & m_{2k-1,k-1} & \\ m_{2k-2,1} & m_{2k-2,2} & \cdots & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ m_{41} & m_{42} & & & \\ m_{31} & & & & \\ m_{21} & & & & \end{bmatrix}. \tag{2.11}$$

The k numbers in the top row characterize the irreducible representations of $O(n)$. The numbers in the next row characterize one of the possible irreducible representations of $O(n-1)$ contained in the specific irreducible representation of $O(n)$, and so forth for successive subgroups of the chain. The numbers in each row thus characterize one of the possible irreducible representations of a specific subgroup. The numbers m_{61}, m_{62}, m_{63}, for example, characterize one of the irreducible representations of $O(6)$.[10]

The Gel'fand basis vectors are not eigenvectors of the k commuting operators $J_{2\alpha-1,2\alpha}$. The basis differs in this respect from the corresponding Gel'fand–Zetlin basis for the unitary groups.[1] Although the full set of m_{ij} are thus not simply related to the components of the weights, they are nevertheless related to the highest weights of the irreducible representations, since the highest weight state of $O(n)$ is an eigenvector of the set of $J_{2\alpha-1,2\alpha}$. The significance of the $m_{n,i}$ is therefore the following:

(a) For $n = 2k + 1$,

$m_{2k+1,1}$ is the maximum possible eigenvalue of J_{12} in $O(2k+1)$,

$m_{2k+1,2}$ is the maximum possible eigenvalue of J_{34} when the eigenvalue of J_{12} is $m_{2k+1,1}$ in $O(2k+1)$,

.
.
.

$m_{2k+1,i}$ is the maximum possible eigenvalue of $J_{2i-1,2i}$ when the eigenvalues of $J_{2\alpha-1,2\alpha}$ are equal to $m_{2k+1,\alpha}$ for all $\alpha < i$ in $O(2k+1)$,

.
.
.

$m_{2k+1,k}$ is the maximum possible eigenvalue of $J_{2k-1,2k}$ when the eigenvalues of $J_{2\alpha-1,2\alpha}$ are equal to $m_{2k+1,\alpha}$ for all $\alpha < k$ in $O(2k+1)$;

(b) For $n = 2k$,

$m_{2k,1}$ is the maximum possible eigenvalue of J_{12} in $O(2k)$,

.
.
.

$m_{2k,i}$ is the maximum possible eigenvalue of $J_{2i-1,2i}$ when the eigenvalues of $J_{2\alpha-1,2\alpha}$ are equal to $m_{2k,\alpha}$ for all $\alpha < i$ in $O(2k)$,

.
.
.

$m_{2k,k-1}$ is the maximum possible eigenvalue of $J_{2k-3,2k-2}$ when the eigenvalues of $J_{2\alpha-1,2\alpha}$ are equal to $m_{2k,\alpha}$ for all $\alpha < k-1$ in $O(2k)$,

$m_{2k,k}$ is the eigenvalue of $J_{2k-1,2k}$ when the eigenvalues of $J_{2\alpha-1,2\alpha}$ are equal to $m_{2k,\alpha}$ for all $\alpha < k-1$ in $O(2k)$.

The irreducible representations of the subgroups in the chain are characterized in the same way.

The numbers m_{ij} are simultaneously either integral or half integral with restrictions which have been given by Gel'fand and Zetlin[6]:

$$m_{2p+1,i} \geq m_{2p,i} \geq m_{2p+1,i+1} \quad (i = 1, 2, \cdots, p),$$

$$m_{2p,i} \geq m_{2p-1,i} \geq m_{2p,i+1} \quad (i = 1, 2, \cdots, p-1),$$

$$m_{2p+1,p} \geq |m_{2pp}|. \tag{2.12}$$

These properties are clear once the lowering and raising operators are derived in this paper.

Since the type-1 operators $J_{2\alpha-1,2\alpha}$ are not diagonal in the general Gel'fand basis, it is convenient to define a whole hierarchy of subbases of decreasing

[10] A slight change has been made in the Gel'fand–Zetlin notation. The first index has been shifted up by one unit so that $m_{\nu 1}, m_{\nu 2}, \cdots$ characterize the irreducible representation of $O(\nu)$. The chain of numbers thus ends with m_{21} [irreducible representation of $O(2)$], rather than with m_{11}.

complexity:

$$
\begin{aligned}
&[\mathcal{M}^{(2)}_{n\mu}] \equiv [\mathcal{M}_{n\mu}],\\
&[\mathcal{M}^{(3)}_{n\mu}] = [\mathcal{M}_{n\mu};\ m_{31} = m_{21}],\\
&[\mathcal{M}^{(4)}_{n\mu}] = [\mathcal{M}_{n\mu};\ m_{41} = m_{31} = m_{21}],\\
&[\mathcal{M}^{(5)}_{n\mu}] = \left[\mathcal{M}_{n\mu};\ \begin{matrix} m_{51} = m_{41} = m_{31} = m_{21} \\ m_{52} = m_{42} \end{matrix}\right],\\
&\qquad\vdots\\
&[\mathcal{M}^{(2i)}_{n\mu}] = \left[\mathcal{M}_{n\mu};\ m_{2i,\alpha} = m_{\beta\alpha} \quad \begin{matrix} \alpha = 1, 2, \cdots, i \\ \beta = 2\alpha, 2\alpha + 1, \cdots, 2i \end{matrix}\right],\\
&[\mathcal{M}^{(2i+1)}_{n\mu}] = \left[\mathcal{M}_{n\mu};\ m_{2i+1,\alpha} = m_{\beta\alpha} \quad \begin{matrix} \alpha = 1, 2, \cdots, i \\ \beta = 2\alpha, 2\alpha + 1, \cdots, 2i \end{matrix}\right],
\end{aligned}
\tag{2.13}
$$

The base vectors of $[\mathcal{M}^{(q)}_{n\mu}]$ with $q = 2i$ or $q = 2i + 1$ have the special property that they are eigenvectors of the set of commuting operators $J_{2\alpha-1,2\alpha}$ with $\alpha = 1, \cdots, i$. Any vector of $[\mathcal{M}^{(q)}_{n\mu}]$ is specified by $(n - q + 1)$ rows of numbers.

The particular subbasis $[\mathcal{M}^{(n-1)}_{n\mu}]$, made up of the base vectors of highest weight in the immediate subgroup $O(n - 1)$ of $O(n)$, is of greatest importance in the present discussion. Its states are specified by only two rows of numbers and it has the following special properties.

(1) *All* of the type-1 operators, $J_{2\alpha-1,2\alpha}$ (with $\alpha = 1, \cdots, k$ for $n = 2k + 1$, and $\alpha = 1, \cdots, k - 1$ for $n = 2k$), are diagonal in this basis.

(2) All of the type-3 raising generators $\rho_{\alpha\beta}$, $\rho_{\alpha,-\beta}$ $(0 < \alpha < \beta)$ give zero when operating on any vector of the basis $[\mathcal{M}^{(n-1)}_{n\mu}]$. This condition is necessary and sufficient to define the basis $[\mathcal{M}^{(n-1)}_{n\mu}]$. (Note that the generators ρ_{12}, $\rho_{1,-2}$, ρ_{13}, $\rho_{1,-3}$, ρ_{23}, $\rho_{2,-3}$, $\cdots$ are naturally considered as raising generators, whereas $\rho_{-1,2}$, $\rho_{-1,-2}$, $\rho_{-1,3}$, $\rho_{-1,-3}$, $\rho_{-2,3}$ $\rho_{-2,-3}$, $\cdots$ are lowering generators.)

The raising and lowering operators which are the subject of this paper are best defined in terms of the subbasis $[\mathcal{M}^{(n-1)}_{n\mu}]$. They are the operators which raise or lower by one integer one of the quantum numbers $m_{n-1,i}$ of the second row without leaving the subbasis $[\mathcal{M}^{(n-1)}_{n\mu}]$, that is the space of base vectors of highest weight in the immediate subgroup. In particular, the full set of states of $[\mathcal{M}^{(n-1)}_{n\mu}]$ can be constructed by repeated operation with the various lowering operators of $O(n)$ on the highest weight state of a specific irreducible representation, namely $[\mathcal{M}^{(n)}_{n\mu}]$. The set of states of $[\mathcal{M}^{(n-2)}_{n\mu}]$ can then be constructed by successive operation with lowering operators of $O(n - 1)$ on the states of $[\mathcal{M}^{(n-1)}_{n\mu}]$ which are highest weight states of irreducible representations of $O(n - 1)$, and so forth, until the full set of Gel'fand states has been reached by successive stepdown operations with the lowering operators of $O(n)$, $O(n - 1), \cdots, O(3)$.

3. THE RAISING AND LOWERING OPERATORS AND THEIR GRAPHS

In $O(3)$ the raising (and lowering) *generators* $J_{13} \pm iJ_{23} \equiv Q_{3,\pm1}$ are themselves raising (and lowering) *operators*; that is, $Q_{3,+1}$ operating on a state $|l, m\rangle$ converts it into a state $|l, m + 1\rangle$. In $O(n)$, with $n > 3$, the raising and lowering generators $Q_{n,i}$ have matrix elements connecting very many different states of the general Gel'fand basis, and when operating on a state of $[\mathcal{M}^{(n-1)}_{n\mu}]$ do not give states belonging solely to $[\mathcal{M}^{(n-1)}_{n\mu}]$.

By forming polynomial functions of the raising and lowering generators, it is possible to construct raising and lowering operators, to be denoted by $O_{n,\pm i}$, which have the simple property that they raise (or lower) by one integer one of the quantum numbers $m_{n-1,i}$ of the subbasis $[\mathcal{M}^{(n-1)}_{n\mu}]$ without leaving this subbasis, that is, the space of base vectors of highest weight in the immediate subgroup $O(n - 1)$.[11] Specifically $O_{n\pm i}$ is defined by

$$
\begin{aligned}
O_{n,\pm i} &\left|\begin{matrix} m_{n1} & m_{n2} & \cdots & m_{ni} & \cdots & m_{nk} \\ m_{n-1,1} & m_{n-1,2} & \cdots & m_{n-1,i} & \cdots & \cdot \end{matrix}\right\rangle\\
&= N' \left|\begin{matrix} m_{n1} & m_{n2} & \cdots & m_{ni} & \cdots & m_{nk} \\ m_{n-1,1} & m_{n-1,2} & \cdots & m_{n-1,i} \pm 1 & \cdots & \cdot \end{matrix}\right\rangle,
\end{aligned}
\tag{3.1}
$$

where N' is a normalization factor and $|\ \rangle$ denotes a normalized state. To save writing, only the column that suffers change is indicated:

$$
\begin{aligned}
&O_{n,\pm i} \left|\begin{matrix} m_{ni} \\ m_{n-1,i} \end{matrix}\right\rangle = N' \left|\begin{matrix} m_{ni} \\ m_{n-1,i} \pm 1 \end{matrix}\right\rangle,\\
&\qquad \begin{cases} i = 1, 2, \cdots, k, & n = 2k + 1, \\ i = 1, 2, \cdots, k - 1, & n = 2k. \end{cases}
\end{aligned}
\tag{3.2}
$$

For $n = 2k$ it is also convenient to introduce the zero-step operator, $O_{2k,k}$

$$
\begin{aligned}
O_{2k,k} &\left|\begin{matrix} m_{2k,1} & \cdots & m_{2k,i} & \cdots & m_{2k,k-1} & m_{2k,k} \\ m_{2k-1,1} & \cdots & m_{2k-1,i} & \cdots & m_{2k-1,k-1} & \end{matrix}\right\rangle\\
&= N' \left|\begin{matrix} m_{2k,1} & \cdots & m_{2k,i} & \cdots & m_{2k,k-1} & m_{2k,k} \\ m_{2k-1,1} & \cdots & m_{2k-1,l} & \cdots & m_{2k-1,k-1} & \end{matrix}\right\rangle
\end{aligned}
\tag{3.3}
$$

[11] For the specific cases $n = 5$ and 6 explicit expressions for raising and lowering operators have been given previously. J. Flores, E. Chacon, P. A. Mello, and M. de Llano, Nucl. Phys. **72**, 352 (1965), and ($n = 5$) K. T. Hecht, *ibid.* **63**, 177 (1965).

or

$$O_{2k,k}\begin{vmatrix} m_{2k,i} \\ m_{2k-1,i} \end{vmatrix} = N' \begin{vmatrix} m_{2k,i} \\ m_{2k-1,i} \end{vmatrix}. \tag{3.4}$$

Since

$$\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} m_{ni} \\ m_{n-1,i} \pm 1 \end{vmatrix} \in [\mathcal{M}_{n\mu}^{(n-1)}],$$

$O_{n,\pm i}$ and $O_{2k,k}$ must satisfy

$$\begin{aligned} &(1)\quad [J_{2\alpha-1,2\alpha}, O_{ni}] = \delta_{\alpha i} O_{ni}, \\ &\qquad [J_{2\alpha-1,2\alpha}, O_{n-i}] = -\delta_{\alpha i} O_{n,-i}, \\ &\qquad [J_{2\alpha-1,2\alpha}, O_{2k,k}] = 0, \quad 0 < \alpha < k, \end{aligned} \tag{3.5}$$

$$\begin{aligned} &(2)\quad [\rho_{\alpha\beta}, O_{n,\pm i}]\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0, \\ &\qquad [\rho_{\alpha\beta}, O_{2k,k}]\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0, \\ &\qquad [\rho_{\alpha,-\beta}, O_{n,\pm i}]\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0, \\ &\qquad [\rho_{\alpha,-\beta}, O_{2k,k}]\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0, \quad 0 < \alpha < |\beta|. \end{aligned} \tag{3.6}$$

Equations (3.5) and (3.6) are necessary and sufficient conditions that $O_{n,\pm i}$ be raising (lowering) operators. Equations (3.6) apply to all of the raising generators of the subgroup $O(n-1)$ and ensure that the state $O_{n,\pm i}|m_{ni}, m_{n-1,i}\rangle$ is a highest weight state of the subgroup $O(n-1)$ since the state $|m_{ni}, m_{n-1,i}\rangle$ has this property. Since the raising and lowering operators are complicated functions of the generators they are best described in terms of graphs, and manipulations involving these operators are also best performed with the aid of these graphs.

A. Raising Operators and Their Associated Graphs

Contents of R_i graphs

Graphs associated with the raising operator $O_{n,i}$ are to be denoted by R_i; these graphs consist of the following (see Table I).

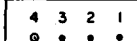
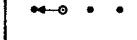
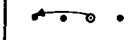
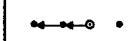
(1) A single row of i ordered points numbered from 1 to i with order increasing from right to left.

(2) A connected chain of arrows always pointing from right to left, with (a) any point $1 \leq j \leq i$ as starting point, to be indicated by a circle, (b) end point always at i, (c) the arrows which form the links of the connected chain may connect some (possibly all) of the points between the starting point j and the end point i but may skip around others (possibly none).

Operator Representations of the R_i Graphs

Each of the many possible graphs of type R_i represents one of the terms of the raising operator $O_{n,i}$.

TABLE I. The graphs of R_{ni} for any $n > 8$.

GRAPHS	OPERATOR REPRESENTATION OF THE GRAPHS
4 3 2 1 	$Q_{n4}\ a_{3\text{-}4}\ a_{2\text{-}4}\ a_{1\text{-}4}$
	$(-\rho_{4\text{-}3})\ Q_{n3}\ a_{2\text{-}4}\ a_{1\text{-}4}$
	$(-\rho_{4\text{-}2})\ Q_{n2}\ a_{3\text{-}4}\ a_{1\text{-}4}$
	$(-\rho_{4\text{-}3})(-\rho_{3\text{-}2})\ Q_{n2}\ a_{1\text{-}4}$
	$(-\rho_{4\text{-}1})\ Q_{n1}\ a_{3\text{-}4}\ a_{2\text{-}4}$
	$(-\rho_{4\text{-}3})(-\rho_{3\text{-}1})\ Q_{n1}\ a_{2\text{-}4}$
	$(-\rho_{4\text{-}2})(-\rho_{2\text{-}1})\ Q_{n1}\ a_{3\text{-}4}$
	$(-\rho_{4\text{-}3})(-\rho_{3\text{-}2})(-\rho_{2\text{-}1})\ Q_{n1}$
$a_{i\text{-}j} = a_i - a_j$	$a_s = 2(J_{2\alpha-1,2\alpha} + k - \alpha)$ $n = 2k$ or $2k+1$

(1) The circle around the starting point j represents the operator $Q_{n,j}$.

(2) An arrow link of the chain connecting points α and β, with $\alpha < \beta$, represents the operator $(-\rho_{\beta,-\alpha})$. [Note that the operator $(-\rho_{\beta,-\alpha}) = \rho_{-\alpha,\beta}$ with $\alpha < \beta$ is a lowering generator of one of the subgroups of $O(n)$.]

(3) A free point, not connected by one of the arrow links of the chain, is associated in the operator representation of the graph by $a_{j,-i} = a_j - a_i$, $a_\alpha = 2(J_{2\alpha-1,2\alpha} + k - \alpha)$ for $n = 2k$ or $n = 2k+1$. (Note that $a_{j,-i} = a_{-i,j}$, and the vectors of $[\mathcal{M}_{n\mu}^{(n-1)}]$ are eigenvectors of a_α.)

(4) The full operator represented by one of the R_i graphs is the product of all the factors of type $(-\rho_{\beta,-\alpha})$ and $Q_{n,j}$ implied by the various links of the graph. The order of the Q and ρ operators in the product reading from right to left is the same as the order of the links of the chain again reading from right to left, with $-\rho_{i,-\alpha}$ on the extreme left and $Q_{n,j}$ on the right followed on the right by all the commuting operator functions $a_{-i,j}$.

The Raising Operators

Theorem: O_{ni} is equal to the sum of the operators represented by all possible graphs R_{ni}.

Proof: Since all raising generators $\rho_{\alpha,\beta}$, $\rho_{\alpha,-\beta}$ $(0 < \alpha < |\beta|)$ can be expressed in terms of commutators of generators of the type $\rho_{j,j+1}$ and $\rho_{j,-(j+1)}$, Eqs. (3.5) and (3.6) follow from

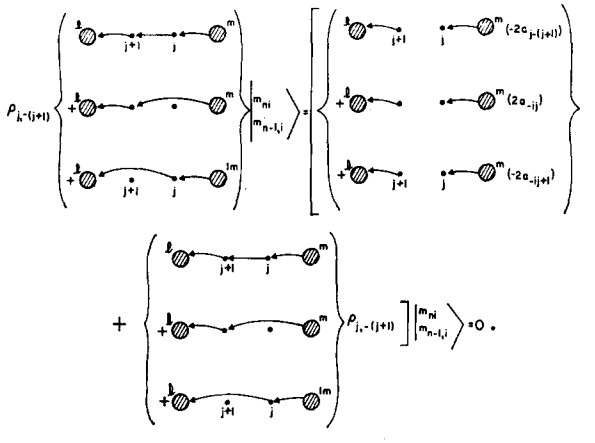
$$[J_{2\alpha-1,2\alpha}, O_{ni}] = \delta_{\alpha i} O_{ni}, \tag{3.7}$$

$$\rho_{j,j+1} O_{ni}\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0, \tag{3.8}$$

$$j = 1, 2, \cdots, (k-1),$$

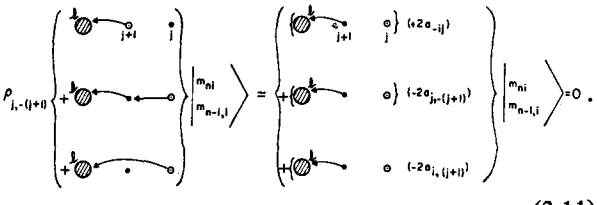
$$\rho_{j,-(j+1)} O_{ni}\begin{vmatrix} m_{ni} \\ m_{n-1,i} \end{vmatrix} = 0. \tag{3.9}$$

Equation (3.7) follows at once from the commutators of $J_{2\alpha-1,2\alpha}$ with the ρ's and Q's [Eqs (2.4)–(2.6)]. Equation (3.8) follows from the fact that a ρ with two positive indices, when commuted through to the right of all factors $\rho_{\beta,-\alpha}$ and Q_j of O_{ni}, leaves one ρ with two positive indices (which in turn has one such surviving term when commuted to the right), and a ρ with two positive indices gives zero when operating on a state of $[\mathcal{M}^{(n-1)}_{n,i}]$. Equation (3.9) follows since all the $\rho_{j,-(j+1)}$ commute through to the right side of all factors ρ and Q of $O_{n,i}$ except for the types which involve the indices j and $j+1$, and $\rho_{j,-(j+1)}$ operating on terms including these satisfy the relations written in terms of graphs:



(3.10)

And also



(3.11)

Special examples

$O(6)$

$$O_{61} = Q_{6,1}$$

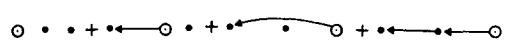
$$O_{62} = Q_{6,2}a_{1-2} + (-\rho_{2-1})Q_{6,1}$$

$O(7)$

$$O_{7,1} = Q_{7,1}$$

$$O_{7,2} = Q_{7,2}a_{1-2} + (-\rho_{2-1})Q_{7,1}$$

$$O_{7,3} = Q_{7,3}a_{1-3}a_{2-3} + (-\rho_{3-2})Q_{7,2}a_{1-3} + (-\rho_{3-1})Q_{7,1}a_{2-3} + (-\rho_{3-2})(-\rho_{2-1})Q_{7,1}$$



B. Lowering Operators and Their Associated Graphs

The graphs for the raising operators $O_{n,1}$ are identical for all $k > i$. The graphs for the lowering operators $O_{n,-1}$, however, are not only dependent on the specific value of n but have a slightly different character for the odd- and even-dimensional rotation groups, $n = 2k+1$ and $n = 2k$, so that the two cases must be discussed separately. Graphs associated with the lowering operators $O_{2k+1,-i}$ are to be denoted by $\mathfrak{L}_{2k+1,i}$

(1) $\mathfrak{L}_{2k+1,i}$ *Graphs*

The $\mathfrak{L}_{2k+1,i}$ graphs consist of the following (see Table II).

(a) Two rows of ordered points, k points in the bottom row numbered from 1 to k with order increasing from right to left, and $k - i + 1$ points in the top row with order decreasing from left to right starting with k at the left and ending with i so that the point, j $(i \leq j \leq k)$, in the top row sits above the point j of the bottom row.

TABLE II. The graphs of $\mathfrak{L}_{7,2}$.

GRAPHS	OPERATOR REPRESENTATION OF THE GRAPHS
•3 ⊙2 •3 •2 •1	$Q_{7-2}\, a_{2-3}\, a_{23}\, a_{22}\, a_{21}$
⊙—• • •	$(-\rho_{-23})\, Q_{7-3}\, a_{23}\, a_{22}\, a_{21}$
⊙—• •	$(-\rho_{-2-3})\, Q_{73}\, a_{2-3}\, a_{22}\, a_{21}$
⊙—• •	$(-\rho_{-23})(-\rho_{-3-2})\, Q_{72}\, a_{23}\, a_{21}$
⊙—• •	$(-\rho_{-2-3})(-\rho_{3-2})\, Q_{72}\, a_{2-3}\, a_{21}$
• • ⊙	$(-\rho_{-2-1})\, Q_{71}\, a_{2-3}\, a_{23}\, a_{22}$
• • ⊙	$(-\rho_{-23})(-\rho_{-3-1})\, Q_{71}\, a_{23}\, a_{22}$
• • ⊙	$(-\rho_{-2-3})(-\rho_{3-1})\, Q_{71}\, a_{2-3}\, a_{22}$
• • ⊙	$(-\rho_{-23})(-\rho_{-3-2})(-\rho_{2-1})\, Q_{71}\, a_{23}$
• • ⊙	$(-\rho_{-2-3})(-\rho_{3-2})(-\rho_{2-1})\, Q_{71}\, a_{2-3}$
$a_{ij} = a_i + a_j$ $\quad$ $a_{i-j} = a_i - a_j$ $a_\alpha = 2(J_{2\alpha-1,2\alpha} + 3 - \alpha)$	

(b) A connected chain of arrows forming a clockwise path, the arrows always pointing from right to left in the bottom row and from left to right in the top row, with (i) any point of either the top or bottom row as starting point, to be indicated by a circle, (ii) end point always at i of the top row, (iii) no vertical arrows (that is, no connections from point l in the bottom row to point l in the top row), (iv) no arrows pointing downward [that is, no arrows with starting points (tails) in the top row and end points (arrowheads) in the bottom row], (v) the arrows which form the links of the connected chain may then be directed from point α in the botton row to point β in the

bottom row with ($\beta > \alpha$), from point μ in the top row to point ν in the top row with ($\nu < \mu$), or from point α in the bottom row to any point σ in the top row $\sigma \lesssim \alpha$, but $\sigma \neq \alpha$.

(2) *Operator Representations of the* $\mathfrak{L}_{2k+1,i}$ *Graphs*

(a) The circle around the starting point, say j, represents the operator $Q_{2k+1,j}$ when it is in the bottom row and $Q_{2k+1,-j}$ when in the top row.

(b) An arrow link of the chain connecting point α to point β represents the operator

(i) $(-\rho_{\beta,-\alpha})$ when $\alpha < \beta$, both points in bottom row,

(ii) $(-\rho_{-\beta,\alpha})$ when $\alpha > \beta$, both points in top row,

(iii) $(-\rho_{\beta,-\alpha})$ with α in the bottom row, β in top row.

(c) A free point not connected by one of the arrow links of the chain is represented by the operator function $a_{i,-\alpha} = a_i - a_\alpha$ when α is in the top row and $a_{i\alpha} = a_i + a_\alpha$ when α is in the bottom row, where $a_\alpha = 2(J_{2\alpha-1,2\alpha} + k - \alpha)$, as before.

(d) The full operator, represented by one of the graphs $\mathfrak{L}_{2k+1,i}$, is again the product of all factors of type ρ and Q implied by the links of the graph. The order of the Q and ρ operators in the product reading from right to left is the same as the order of the links of the chain starting with the encircled point and ending at point i of the top row, with Q followed on the right by all the factors $a_{i\pm\alpha}$ implied by the free points of the graph.

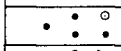
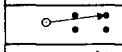
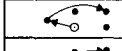
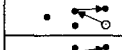
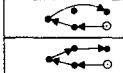
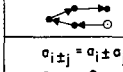
(3) $\mathfrak{L}_{2k,i}$ *Graphs*

Graphs associated with the lowering operators $O_{2k,-i}$ of the orthogonal group in an even number of dimensions are to be denoted by $\mathfrak{L}_{2k,i}$ (see Table III). The graphs $\mathfrak{L}_{2k,i}$ have the same structure as the graphs $\mathfrak{L}_{2k+1,i}$ with the exception that the two points k are replaced by a single point to be placed halfway between the top and bottom rows but to the left of the two points $(k-1)$. The rules for the construction of the operators represented by the graphs $\mathfrak{L}_{2k,i}$ are the same as those for the graphs $\mathfrak{L}_{2k+1,i}$ except for the following.

(a) A free point, not connected by one of the arrow links of the chains and if placed in the αth position of the bottom row, is to be denoted by $b_{i\alpha} = a_{i\alpha} - 2$. If the kth point is a free point it is to be denoted by $c_i = \frac{1}{2}(a_i - 2)$. (Free points of the top row are associated with $a_{i,-\alpha}$, as for $\mathfrak{L}_{2k+1,i}$.)

(b) For the special case $i = k$, required for the zero-step operator, the free points of the bottom row (say in position α) are now to be denoted merely by a_α. (The points of the top row play no role whatsoever in this special case.)

TABLE III. The graphs of $\mathfrak{L}_{6,1}$.

GRAPHS	OPERATOR REPRESENTATION OF THE GRAPHS
	$Q_{6\text{-}1}\ a_{1\text{-}2}\ c_1\ b_{12}\ b_{11}$
	$(-\rho_{\text{-}12})\ Q_{6\text{-}2}\ c_1\ b_{12}\ b_{11}$
	$(-\rho_{\text{-}13})\ Q_{6\text{-}3}\ a_{1\text{-}2}\ b_{12}\ b_{11}$
	$(-\rho_{\text{-}12})(-\rho_{\text{-}23})\ Q_{6\text{-}3}\ b_{12}\ b_{11}$
	$(-\rho_{\text{-}1\text{-}2})\ Q_{62}\ a_{1\text{-}2}\ c_1\ b_{11}$
	$(-\rho_{\text{-}13})(-\rho_{\text{-}3\text{-}2})\ Q_{62}\ a_{1\text{-}2}\ b_{11}$
	$(-\rho_{\text{-}12})(-\rho_{\text{-}23})(-\rho_{\text{-}3\text{-}2})\ Q_{62}\ b_{11}$
	$(-\rho_{\text{-}12})(-\rho_{\text{-}2\text{-}1})\ Q_{61}\ c_1\ b_{12}$
	$(-\rho_{\text{-}13})(-\rho_{\text{-}3\text{-}1})\ Q_{61}\ a_{1\text{-}2}\ b_{12}$
	$(-\rho_{\text{-}12})(-\rho_{\text{-}23})(-\rho_{\text{-}3\text{-}1})\ Q_{61}\ b_{12}$
	$(-\rho_{\text{-}1\text{-}2})(-\rho_{2\text{-}1})\ Q_{61}\ a_{1\text{-}2}\ c_1$
	$(-\rho_{\text{-}13})(-\rho_{\text{-}3\text{-}2})(-\rho_{2\text{-}1})\ Q_{61}\ a_{1\text{-}2}$
	$(-\rho_{\text{-}12})(-\rho_{\text{-}23})(-\rho_{\text{-}3\text{-}2})(-\rho_{2\text{-}1})\ Q_{61}$
$a_{i\pm j} = a_i \pm a_j$	$b_{ij} = a_{ij} - 2$	$c_i = 1/2(a_i - 2)$
$Q_{6,3} = Q_{6,\text{-}3}$	$\rho_{\alpha,3} = \rho_{\alpha,\text{-}3}$	$a_\alpha = 2(J_{2\alpha-1,2\alpha} + 3 - \alpha)$

Theorem: $O_{2k+1,-i}$, $O_{2k,-i}$ is equal to the sum of the operators represented by all the possible graphs of $\mathfrak{L}_{2k+1,i}$ and $\mathfrak{L}_{2k,i}$ respectively.

Proof: (a) $[J_{2\alpha-1,2\alpha}, O_{n,-i}] = -O_{n,-i}\delta_{\alpha i}$

This again follows at once from the commutators of Eqs. (2.4)–(2.6).

(b) For the relation

$$\rho_{j,-(j+1)} O_{n,-i} \left| \begin{matrix} m_{ni} \\ m_{n-1,i} \end{matrix} \right\rangle = 0,$$

the proof is essentially the same as that for the raising operator except that there are two sets of terms like those of Eqs. (3.10)–(3.11). One set arises when j and $j+1$ are both in the bottom row, the other when j and $j+1$ are both in the top row. Both sets of terms sum to zero independently of each other. [The points m and l of Eqs. (3.10)–(3.11) can now be in either top or bottom row.]

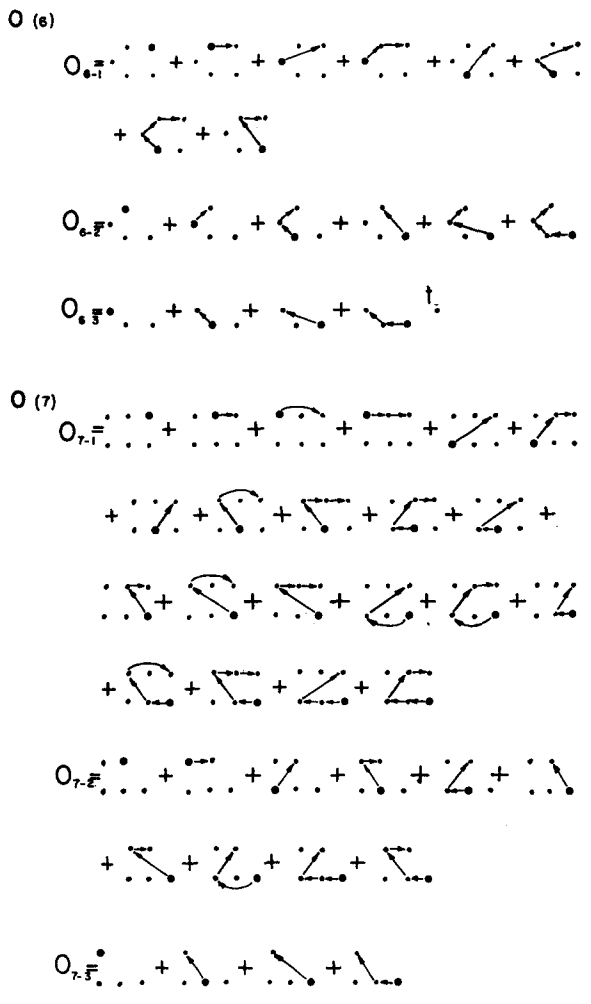
The proof of the relation

$$\rho_{j,j+1} O_{n,-i} \left| \begin{matrix} m_{ni} \\ m_{n-1,i} \end{matrix} \right\rangle = 0$$

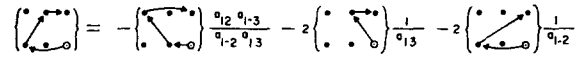
is much more complicated since more summations of graphs are involved. However, the method is identical[12] to that illustrated by Eqs. (3.10)–(3.11).

[12] S. C. Pang, University of Michigan dissertation (to be published).

(4) *Special Examples of Lowering Operators*[13]



Not all the graphs give independent operators. In $O_{7,-1}$, for example, only 15 out of the 21 graphs are independent. The remaining six give operators which can be written as linear combinations of the 15 independent ones. Terms 16 and 18, for example, are related by



It is most convenient, however, to treat all graphs on an equal footing to preserve both the over-all symmetry of the expression for $O_{n,-i}$ and the uniform and simple factoring of the operators associated with an individual graph. The operator representations of the various graphs all have the same structure, differing only in the number of factors of type ρ and a. The total number of operator factors for each graph of $\mathfrak{L}_{ni}$ is equal to $n - i$.

[13†] Note that O_{63} is an example of a neutral or zero-step operator of type $O_{2k,k}$.

C. Some Properties of the Raising and Lowering Operators

The raising and lowering operators which have been constructed have meaning only when they operate on the basis $[\mathcal{M}_{n\mu}^{(n-1)}]$. It is interesting to note that the operators $O_{n,\pm\alpha}$ together with the $J_{2\alpha-1,2\alpha}$ form a Lie algebra with respect to $[\mathcal{M}_{n\mu}^{(n-1)}]$. The raising and lowering operators have not yet been normalized. However, the unnormalized operators $O_{n,\alpha}$ have the simple property

$$[O_{ni}, O_{nj}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle = 0, \quad i \neq -j. \tag{3.12}$$

With respect to the basis $[\mathcal{M}_{n\mu}^{(n-1)}]$, the set of operators $O_{n\alpha}$, $O_{n,-\alpha}$, $J_{2\alpha-1,2\alpha}$ thus commutes with any other set $O_{n\beta}$, $O_{n,-\beta}$, $J_{2\beta-1,2\beta}$ $(\beta \neq \alpha)$, so that the Lie algebras mentioned above breaks up into a set of k, $(k-1)$, commuting algebras of order three for dimension $n = 2k + 1$ $(n = 2k)$, respectively. Equation (3.12) can be verified by direct computation or obtained from the following considerations.

From the uniqueness of the base vector

$$\left|\begin{matrix} m_{ni} & m_{nj} \\ m_{n-1,i} \pm 1 & m_{n-1,j} \pm 1 \end{matrix}\right\rangle$$

the states $O_{ni}O_{nj}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle$, $O_{nj}O_{ni}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle$ can differ by at most a constant:

$$O_{ni}O_{nj}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle = c_{ij}O_{nj}O_{ni}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle. \tag{3.13}$$

The constant c_{ij} can be shown to be unity by comparing the coefficients of the terms with the largest number of factors of type ρ on each side of Eq. (3.13). The term with the largest number of factors ρ for a single operator O_{ni} arises from a single graph and has the coefficient unity in all cases except $i < 0$, $n = 2k + 1$. In the latter case it arises from two graphs (e.g., graphs Nos. 9 and 10 of Table II) whose summed coefficient (on the right) is equal to a_{ii}. This has the same eigenvalue when operating on the state $|\mathcal{M}_{n\mu}^{(n-1)}\rangle$ or on $O_{nj}\,|\mathcal{M}_{n\mu}^{(n-1)}\rangle$. Thus $c_{ij} = 1$.

4. THE NORMALIZATION

The raising and lowering operators $O_{n,i}$ do not yield normalized basis vectors. It is therefore important to define normalized raising and lowering operators, to be denoted by $U_{n,i}$, which differ from the O_{ni} merely by a normalization factor. The calculation of these normalization factors is presented in this section.

The results for the even- and odd-dimensional orthogonal group are somewhat different. For $n = 2k + 1$

the normalized raising and lowering operators are[14]

$$U_{2k+1,i} = \left| \left(\prod_{\alpha=1}^{i-1} \frac{a_{i\alpha}}{(a_{-i,\alpha}+2)(a_{-i\alpha^0}^{(2k+1)}+2)a_{i\alpha^0}} \right) \left(\prod_{\beta=i+1}^{k} \frac{a_{i,-\beta}a_{i\beta}}{a_{i\beta^0}^{(2k+1)}(a_{i,-\beta^0}^{(2k+1)}-2)} \right) \frac{2}{a_{ii^0}^{(2k+1)}} \frac{2}{(a_{i^0,-i}^{(2k+1)}+2)} \right|^{\frac{1}{2}} O_{2k+1,i}, \tag{4.1}$$

$$U_{2k+1,-i} = O_{2k+1,-i} \frac{1}{a_{ii}} \left| \frac{2}{a_{ii^0}^{(2k+1)}} \frac{2}{(a_{i^0,-i}^{(2k+1)}+2)} \prod_{\alpha=1}^{i-1} \frac{(a_{-i\alpha}+2)}{a_{i\alpha}(a_{-i,\alpha^0}^{(2k+1)}+2)a_{i\alpha^0}^{(2k+1)}} \prod_{\beta=i+1}^{k} \frac{1}{a_{i,-\beta}a_{i\beta}a_{i\beta^0}^{(2k+1)}(a_{i,-\beta^0}^{(2k+1)}-2)} \right|^{\frac{1}{2}}, \tag{4.2}$$

where

$$a_{i\pm\alpha} = 2(J_{2i-1,2i} + k - i) \pm 2(J_{2\alpha-1,2\alpha} + k - \alpha) \tag{4.3}$$

with eigenvalue $2(m_{2ki} + k - i) \pm 2(m_{2k\alpha} + k - \alpha)$ in the restricted basis $[\mathcal{M}_{2k+1,\mu}^{(2k)}]$. The superscript zero on a subscript of $a_{i\alpha}$ has the following meaning: the eigenvalue of the corresponding $J_{2\alpha-1,2\alpha}$ is to take its highest possible value in $O(2k+1)$. For example, the eigenvalue of $a_{i\alpha^0}^{(2k+1)}$ is

$$2(m_{2ki} + k - i) + 2(m_{2k+1,\alpha} + k - \alpha).$$

For $n = 2k$ the normalized raising and lowering operators are

$$U_{2ki} = \frac{|b_{ii}c_i|^{\frac{1}{2}}}{2} \left| \left(\prod_{\alpha=1}^{i-1} \frac{b_{i\alpha}}{(a_{-i,\alpha^0}^{(2k)}+2)(a_{-i,\alpha}+2)b_{i\alpha^0}^{(2k)}} \right) \left(\prod_{\beta=i+1}^{k-1} \frac{a_{i-\beta}b_{i\beta}}{(a_{i,-\beta^0}^{(2k)}-2)b_{i\beta^0}^{(2k)}} \right) \frac{2}{b_{ii^0}^{(2k)}} \frac{2}{(a_{i^0,-i}^{(2k)}+2)} \frac{2}{(a_{i,-k^0}^{(2k)}-2)} \frac{2}{b_{ik^0}} \right|^{\frac{1}{2}} O_{2ki}, \tag{4.4}$$

$$U_{2k,-i} = O_{2k,-i} \frac{1}{2} \left| \frac{1}{b_{ii}c_i} \frac{2}{b_{ii^0}^{(2k)}} \frac{2}{(a_{i,-k^0}^{(2k)}-2)} \frac{2}{(a_{i^0,-i}^{(2k)}+2)} \frac{2}{b_{ik^0}^{(2k)}} \prod_{\alpha=1}^{i-1} \frac{(a_{-i\alpha}+2)}{(a_{-i\alpha^0}^{(2k)}+2)b_{i\alpha}b_{i\alpha^0}^{(2k)}} \prod_{\beta=i+1}^{k-1} \frac{1}{a_{i,-\beta}b_{i\beta}(a_{i,-\beta^0}^{(2k)}-2)b_{i\beta^0}^{(2k)}} \right|^{\frac{1}{2}}, \tag{4.5}$$

where $a_{i\alpha}$, $a_{i\alpha^0}^{(2k)}$ are defined as before, and $b_{i\alpha} = a_{i\alpha} - 2$, $c_i = \frac{1}{2}(a_i - 2)$.

The general basis vector for the orthogonal group can then be generated by successive applications of these operators. Taking the case $n = 2k+1$ as an example,

$$\begin{aligned}
|\mathcal{M}_{2k+1\mu}^{(2k)}\rangle &= U_{2k+1-1}^{m_{2k+1,1}-m_{2k,1}} U_{2k+1,-2}^{m_{2k+1,2}-m_{2k,2}} \cdots U_{2k+1,-k}^{m_{2k+1,k}-m_{2k,k}} |\mathcal{M}_{2k+1,\mu}^{(2k+1)}\rangle = \overrightarrow{\prod_{\alpha=1}^{k}} U_{2k+1,-\alpha}^{m_{2k+1,\alpha}-m_{2k^\alpha}} |\mathcal{M}_{2k+1,\mu}^{(2k+1)}\rangle, \\
|\mathcal{M}_{2k+1,\mu}^{(2k-1)}\rangle &= U_{2k,-1}^{m_{2k,1}-m_{2k-1,1}} U_{2k,-2}^{m_{2k,2}-m_{2k-1,2}} \cdots U_{2k,-(k-1)}^{m_{2k,k-1}-m_{2k-1,k-1}} |\mathcal{M}_{2k+1,\mu}^{(2k)}\rangle = \overrightarrow{\prod_{\beta=1}^{k-1}} U_{2k,-\beta}^{m_{2k,\beta}-m_{2k-1,\beta}} |\mathcal{M}_{2k+1,\mu}^{(2k)}\rangle, \\
&\;\;\vdots \\
|\mathcal{M}_{2k+1,\mu}^{(3)}\rangle &= U_{3-1}^{m_{31}-m_{21}} |\mathcal{M}_{2k+1,\mu}^{(3)}\rangle.
\end{aligned} \tag{4.6}$$

Therefore

$$|\mathcal{M}_{2k+1,\mu}\rangle = \overrightarrow{\prod_{i=3}^{2k+1}} \overrightarrow{\prod_{j=1}^{[i-1]}} U_{i-j}^{m_{i\alpha}-m_{i-1,\alpha}} |\mathcal{M}_{2k+1,\mu}^{(2k+1)}\rangle, \tag{4.7}$$

where

$$[n] = \begin{cases} \frac{1}{2}(n-1), & n \text{ odd}, \\ \frac{1}{2}n, & n \text{ even}. \end{cases}$$

The symbol $\overrightarrow{\prod}$ with an arrow means that terms are to be arranged in increasing order from left to right. Note that the eigenvalues of the $a_{i\alpha}$ depend upon the exact position of these factors in the ordered product.

A. Normalization Factor for the Case $n = 2k+1$

Since the lowering operators $O_{2k+1,-\alpha}$ form a commuting set of operators in the restricted basis $[\mathcal{M}_{2k+1,\mu}^{(2k)}]$, it is sufficient to consider the special vector $|i\rangle$ in the calculation of the normalization factor associated with $O_{2k+1,-i}$, where $|i\rangle$ is defined by

$$|i\rangle = |[\mathcal{M}_{2k+1,\mu}^{(2k)}];\ m_{2k,\alpha} = m_{2k+1,\alpha} \quad \text{for} \quad \alpha \neq i\rangle. \tag{4.8}$$

Before calculating the normalization factors, a number of preparatory steps are taken.

(1) *The Quadratic Casimir Invariant*

It is well known that $\sum_{i<j}^{2k+1} J_{ij}^2$ is a quadratic invariant of $O(2k+1)$,

$$\sum_{i<j}^{2k+1} J_{ij}^2 |\mathcal{M}_{2k+1,\mu}^{(\alpha)}\rangle = C_{2k+1} |\mathcal{M}_{2k+1,\mu}^{(\alpha)}\rangle \quad \text{for all } \alpha. \tag{4.9}$$

Expressing J_{ij} in terms of the Q operators of both

[14] The superscript $(2k+1)$ will be omitted whenever it is obvious.

Eqs. (2.2) and (2.3) the invariant takes the form

$$\sum_{i<j}^{2k+1} J_{ij}^2 = \sum_{i=1}^{k} Q_{2k+1,-i}Q_{2k+1,i} + \sum_{0<i<j}^{k} Q_{2j,-i}Q_{2j,i} + \sum_{0<i<j}^{k} Q_{2j-1,-i}Q_{2j-1,i} + \sum_{i=1}^{k} J_{2i-1,2i}^2 + \sum_{i=1}^{k}(2k-2i+1)J_{2i-1,2i}. \quad (4.10)$$

By applying (4.10) to

$$|\mathcal{M}^{(2k+1)}_{2k+1,\mu}\rangle$$

and using the fact that the raising generators Q_i give zero when operating on the highest-weight state, the invariant can be evaluated:

$$C_{2k+1} = \sum_{\alpha=1}^{k} m_{2k+1,\alpha}^2 + \sum_{\alpha=1}^{k}(2k-2\alpha+1)m_{2k+1,\alpha}. \quad (4.11)$$

(2) *Some Preparatory Lemmas*

Lemma 1:

$$Q_{n,j}\,|i\rangle = 0 \quad \text{for} \quad 0<j<i, \quad (4.12)$$

where the vector $|i\rangle$ is defined by Eq. (4.8).

Proof:

$$O_{n\alpha} = \sum_{\beta=1}^{\alpha} g_{\alpha\beta}(\rho)Q_{n\beta},$$

$$Q_{n\beta} = \sum_{\beta=1}^{\alpha} h_{\alpha\beta}(\rho)O_{n\beta},$$

since

$$O_{n\beta}\,|i\rangle = 0 \quad \beta \neq i \quad (m_{2k\beta} = m_{2k+1\beta} \quad \beta \neq i),$$

it follows that

$$Q_{n,j}\,|i\rangle = 0, \quad 0<j<i.$$

Lemma 2:

$$\langle i|\sum_{j\geq i}^{k} Q_{2k+1,-j}Q_{2k+1,j}\,|i\rangle = \langle i|\,(m_{2k+1,i} - J_{2i-1,2i}) \times (m_{2k+1,i} + J_{2i-1,2i} + 2k - 2i + 1)\,|i\rangle. \quad (4.13)$$

This is a consequence of Lemma 1 and Eqs. (4.10) and (4.11).

Lemma 3:

$$\langle i|\,Q_{n,-i}Q_{ni}\,|i\rangle = \langle i|\,\frac{2\sum_{m\geq i}^{l-1} Q_{n,-l}Q_{nl}}{a_{i-l}}\,|i\rangle = \langle i|\,\frac{2}{a_{i-l}}\prod_{\alpha=i+1}^{l-1}\frac{a_{i-\alpha}+2}{a_{i-\alpha}}\,Q_{n,-i}Q_{ni}\,|i\rangle, \quad l>i. \quad (4.14)$$

This follows from the relations

(i) $\langle i|\,Q_{n,-\beta}O_{n\beta}\,|i\rangle = 0, \quad i<\beta,$

(ii) $\langle i|\,\rho_{-\epsilon\lambda} = 0, \quad 0<\epsilon<|\lambda|,$

and a process of mathematical induction. Note that (ii) follows from $\langle i|\,\rho_{-\epsilon\lambda} = (-\rho_{\epsilon,-\lambda}\,|i\rangle)^{\dagger}$ and the fact that $-\rho_{\epsilon,-\lambda}$ with $0<\epsilon<|\lambda|$ is a raising generator of a subgroup of $O(n)$. Set $\beta = i+1$ in relation (i). As a consequence of Lemma 1 only two terms of $O_{n\beta}$ (corresponding to the first two graphs of Table I) survive. Commuting $Q_{n,-(i+1)}$ through the factor $\rho_{i+1,-i}$ and using relation (ii), the term arising through the second graph reduces to

$$-2\langle i|\;Q_{n,-i}Q_{ni}\prod_{\alpha=1}^{i-1} a_{\alpha,-(i+1)}\,|i\rangle.$$

Together with the first term this leads to the special case of Eq. (4.14), with $l = i+1$. By similar techniques the case with arbitrary l can be related to that with $l-1$.

Lemma 4:

$$\langle i|\,Q_{2k+1,-i}Q_{2k+1,i}\,|i\rangle = \langle i|\prod_{l=i+1}^{k}\frac{a_{i,-l}}{(a_{i-l}+2)}(m_{2k+1,i} - J_{2i-1,2i}) \times (m_{2k+1,i} + J_{2i-1,2i} + 2k - 2i + 1)\,|i\rangle. \quad (4.15)$$

This is a direct consequence of Lemma 3 and Lemma 2.

(3) *Evaluation of* $\langle i|\,O_{2k+1,i}O_{2k+1,-i}\,|i\rangle$

All terms in the raising operator $O_{2k+1,i}$, except the one term containing $Q_{2k+1,i}$, have at least one factor $\rho_{-\epsilon\lambda}$ $(0<\epsilon<|\lambda|)$ on the left-hand side. Since $\langle i|\,\rho_{-\epsilon\lambda} = 0$, the basic matrix element reduces to

$$\langle i|\,O_{2k+1,i}O_{2k+1,-i}\,|i\rangle = \langle i|\,Q_{2k+1,i}O_{2k+1,-i}\,|i\rangle\prod_{\alpha=1}^{i-1}\langle i|\,(a_{\alpha-i}+2)\,|i\rangle. \quad (4.16)$$

The matrix element $\langle i|\,Q_{2k+1,i}O_{2k+1,-i}\,|i\rangle$ is evaluated by commuting all of the factors $\rho_{-\epsilon\lambda}$ of $O_{2k+1,-i}$ through to the left-hand side where they give zero when operating on $\langle i|$. After this process only matrix elements of the type $\langle i|\,Q_{2k+1,-j}Q_{2k+1,j}\,|i\rangle$, $(j \geq i)$, survive. Their coefficients are evaluated in Appendix A by a process of summing of graphs. The matrix elements themselves are given by Lemmas 3 and 4. Combining these results (Appendix A), the basic matrix element is

$$\langle i|\,O_{2k+1,i}O_{2k+1,-i}\,|i\rangle = \langle i|\left(\prod_{\alpha=1}^{k} a_{i\alpha}\prod_{\gamma=i+1}^{k}(a_{i,-\gamma}-2)\prod_{\beta=1}^{i-1}(a_{-i\beta}+2)\right) \times (m_{2k+1,i} + J_{2i-1,2i} + 2k - 2i)\,|i\rangle. \quad (m_{2k+1,i} - J_{2i-1,2i} + 1) \quad (4.17)$$

In the state $|i\rangle$ all $J_{2\alpha-1,2\alpha}$ except that with $\alpha = i$ yield their highest-weight value

$$\begin{aligned}\langle i|\, J_{2\alpha-1,2\alpha}\,|i\rangle &= m_{2k+1,\alpha}\,,\\ \langle i|\, J_{2i-1,2i}\,|i\rangle &= m_{2k,i}\,.\end{aligned}\qquad(\alpha \neq i)\qquad(4.18)$$

Thus

$$\begin{aligned}&\langle i|\, O_{2k+1,i}O_{2k+1,-i}\,|i\rangle\\ &= \langle \mathcal{M}^{(2k)}_{2k+1,\mu}|\,(m_{2k+1,i} - J_{2i-1,2i} + 2k - 2i)\,|\mathcal{M}^{(2k)}_{2k+1,\mu}\rangle\\ &\quad\times \tfrac{1}{2}\langle \mathcal{M}^{(2k)}_{2k+1,\mu}|\prod_{\alpha=1}^{k} a_{i\alpha^0}^{(2k+1)}\prod_{\gamma=i+1}^{k}\\ &\quad\times (a_{i,-\gamma^0}^{(2k+1)} - 2)\prod_{\beta=1}^{i}(a_{-i\beta^0}^{(2k+1)} + 2)\,|\mathcal{M}^{(2k)}_{2k+1,\mu}\rangle.\end{aligned}\qquad(4.19)$$

The superscript zero on a subscript of $a_{i\alpha}$ has been defined in connection with Eq. (4.19). For example,

$$\langle \mathcal{M}^{(2k)}_{2k+1,\mu}|\, a_{i,-\gamma^0}^{(2k+1)}\,|\mathcal{M}^{(2k)}_{2k+1,\mu}\rangle = 2(m_{2k,i} - m_{2k+1,\gamma} + \gamma - i).\qquad(4.20)$$

B. Normalization Factor for the Case $n = 2k$

(1) *The Quadratic Casimir Invariant*

$$\sum_{i<j}^{2k} J_{ij}^2\,|\mathcal{M}^{(\alpha)}_{2k,\mu}\rangle = C_{2k}\,|\mathcal{M}^{(\alpha)}_{2k,\mu}\rangle \quad\text{for all } \alpha.\qquad(4.21)$$

Expressing the J_{ij} in terms of Q operators as before

$$\begin{aligned}\sum_{i<j}^{2k} J_{ij}^2 =& \sum_{i=1}^{k-1} Q_{2k,-i}Q_{2ki} + \sum_{0<i<j}^{k-1} Q_{2j,-i}Q_{2ji}\\ &+ \sum_{0<i<j}^{k} Q_{2j-1,-i}Q_{2j-1,i} + \sum_{i=1}^{k} J_{2i-1,2i}^2\\ &+ \sum_{i=1}^{k-1}(2k - 2i)J_{2i-1,2i}.\end{aligned}\qquad(4.22)$$

$$C_{2k} = \sum_{i=1}^{k} m_{2k,-i}^2 + \sum_{i=1}^{k-1}(2k - 2i)m_{2k,i}.\qquad(4.23)$$

As before, it is convenient to define the special vector

$$|i\rangle = |\mathcal{M}^{(2k-1)}_{2k,\mu};\, m_{2k,\alpha} = m_{2k-1,\alpha}\quad \alpha \neq i\rangle.\qquad(4.24)$$

Since the raising operators for $O(2k)$ and $O(2k+1)$ have the same form, Eqs. (4.12) and (4.14) hold, and

$$\langle i|\sum_{j\geq i}^{k-1} Q_{2k,-j}Q_{2kj}\,|i\rangle = \langle i|\prod_{\alpha=i+1}^{k-1}\frac{(a_{i-\alpha}+2)}{a_{i\alpha}}\,Q_{2k,-i}Q_{2k,i}\,|i\rangle.\qquad(4.25)$$

Putting this relation back into the expression for the quadratic Casimir invariant gives

$$\begin{aligned}\langle i|\Big[\prod_{\alpha=i+1}^{k-1}\frac{(a_{i-\alpha}+2)}{a_{i-\alpha}}\Big]Q_{2k,-i}Q_{2k,i} + J_{2k-1,2k}^2\\ + J_{2i-1,2i}^2 + (2k-2i)J_{2i-1,2i}\,|i\rangle\\ = m_{2ki}^2 + (2k-2i)m_{2k,i} + m_{2k,k}^2.\end{aligned}\qquad(4.26)$$

Unlike the corresponding equation for the case of the odd-dimensional orthogonal group, this relation is not sufficient to evaluate the matrix element $\langle i|\, Q_{2k,-i}Q_{2k,i}\,|i\rangle$, since the matrix element

$$\langle i|\, J_{2k-1,2k}J_{2k-1,2k}\,|i\rangle$$

is not known. However, there is now one more invariant at our disposal.

(2) *The Quadratic Invariant in the Restricted Basis* $[\mathcal{M}^{(2k-1)}_{2k\mu}]$

Since the (zero-step) neutral operator $O_{2k,k}$ commutes with all raising and lowering operators when applied to the basis $[\mathcal{M}^{(2k-1)}_{2k,\mu}]$, it is an invariant in this restricted basis. To get a relation between the matrix elements of the quadratic factors $J_{2k-1,2k}^2$ and $Q_{2k,-i}Q_{2k,i}$ consider $\langle i|\, O_{2k,k}O_{2k,k}\,|i\rangle$, where

$$\begin{aligned}\langle i|\, O_{2k,k}O_{2k,k}\,|i\rangle &= \langle i|\prod_{\alpha=1}^{k-1} a_\alpha J_{2k-1,2k}O_{2k,k}\,|i\rangle\\ &= \langle i|\prod_{\alpha=1}^{k-1} a_\alpha\,|i\rangle\langle i|\, J_{2k-1,2k}O_{2k,k}\,|i\rangle\end{aligned}\qquad(4.27)$$

through the relation $\langle i|\,\rho_{-j\epsilon} = 0,\ 0 < j < |\epsilon|$. Summing up of the matrix elements from all the possible graphs in $O_{2k,k}$ with techniques similar to those illustrated in Appendix A leads to

$$\begin{aligned}&\langle i|\, O_{2k,k}O_{2k,k}\,|i\rangle\\ &= \langle i|\prod_{\alpha=1}^{k-1} a_\alpha^2\,|i\rangle\\ &\quad\times\langle i|\, J_{2k-1,2k}^2 - \prod_{\alpha=i+1}^{k-1}\frac{(a_{i-\alpha}+2)}{a_{i-\alpha}a_i}\,Q_{2k,-i}Q_{2k,i}\,|i\rangle.\end{aligned}\qquad(4.28)$$

On the other hand, since $O_{2k,k}$ is an invariant

$$\langle i|\, O_{2k,k}O_{2k,k}\,|i\rangle = \langle \mathcal{M}^{(2k)}_{2k,\mu}|\, O_{2k,k}O_{2k,k}\,|\mathcal{M}^{(2k)}_{2k,\mu}\rangle.\qquad(4.29)$$

Also

$$\langle i|\, a_\alpha^2\,|i\rangle = \langle \mathcal{M}^{(2k)}_{2k,\mu}|\, a_\alpha^2\,|\mathcal{M}^{(2k)}_{2k,\mu}\rangle \quad\text{for } \alpha \neq i.\qquad(4.30)$$

By applying (4.28), (4.29), and (4.30), the quartic invariant leads to the relation

$$\begin{aligned}\langle i|\, J_{2k-1,2k}^2 a_i^2 - a_i\prod_{\alpha=i+1}^{k-1}\frac{(a_{i-\alpha}+2)}{a_{i-\alpha}}\,Q_{2k,-i}Q_{2k,i}\,|i\rangle\\ = 4m_{2k,k}^2(m_{2k,i} + k - i)^2.\end{aligned}\qquad(4.31)$$

(3) *Evaluation of* $\langle i|\, O_iO_{-i}\,|i\rangle$

Since we have two equations and two unknowns we can determine both $\langle i|\, Q_{2k,-i}Q_{2k,i}|i\rangle$ and $\langle i|J_{2k-1,2k}^2|i\rangle$. The technique for the summing up of the graphs is similar to the case of $O(2k+1)$ illustrated in Appendix

A (and Sec. 4A3.) and leads to

$$\langle i|\, O_{2k,i}O_{2k,-i}\,|i\rangle$$
$$= \tfrac{1}{4}\langle \mathcal{M}^{(2k-1)}_{2k,\mu}|\prod_{\alpha=i+1}^{k}(a^{(2k)}_{i-\alpha^0}-2)\prod_{\beta=1}^{i}(a^{(2k)}_{-i,\beta^0}+2)$$
$$\times\prod_{\gamma=1}^{k} b^{(2k)}_{i\gamma^0}\,|\mathcal{M}^{(2k-1)}_{2k,\mu}\rangle\langle\mathcal{M}^{(2k-1)}_{2k,\mu}|\,\frac{b^{(2k)}_{ii^0}}{b_{ii}}\,|\mathcal{M}^{(2k-1)}_{2k,\mu}\rangle. \quad (4.32)$$

The superscript zero on a subscript of $a_{i\alpha}$ has the same meaning as before. For example,

$$\langle \mathcal{M}^{(2k-1)}_{2k,\mu}|\, b^{(2k)}_{i\gamma^0}\,|\mathcal{M}^{(2k-1)}_{2k,\mu}\rangle$$
$$= 2(m_{2k-1,i}+m_{2k,\gamma}+2k-i-\gamma-1). \quad (4.33)$$

C. The Normalization Coefficients

Let the normalized lowering (raising) operators be denoted by $U_{n,\pm i}$. If the state $|\mathcal{M}^{(n-1)}_{n\mu}\rangle$ is normalized

$$\langle \mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}|\, U_{ni}U_{n,-i}\,|\mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}\rangle$$
$$= \langle \mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}-1\,|\,\mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}-1\rangle = 1. \quad (4.34)$$

But

$$\langle \mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}-1| = \langle \mathcal{M}^{(n-1)}_{n\mu};\, m_{n-1,i}|\,(U_{n,-i})^{\dagger}. \quad (4.35)$$

The normalized lowering (raising) operators should thus have the property

$$U_{ni} = (U_{n,-i})^{\dagger}. \quad (4.36)$$

The lowering (raising) operators of type $O_{n,\pm i}$ do not satisfy this relation. However, if O_{ni} is a lowering (raising) operator of $[\mathcal{M}^{(n-1)}_{n\mu}]$, so is

$$f_{ni}(J_{12}, J_{34}, \cdots)O_{ni},$$

where f_{ni} is a function of $J_{2\alpha-1,2\alpha}$ only ($\alpha = 1, \cdots, k$ for $n = 2k+1$, $\alpha = 1, \cdots, k-1$ for $n = 2k$), and where $f_{n,\pm i}$ can be chosen such that

$$f_{ni}O_{ni} = (O_{n,-i}f_{n,-i})^{\dagger}. \quad (4.37)$$

Since any arbitrary function $g(Q, \rho, J)\rho_{\alpha\beta}$, with $0 < \alpha < 1\,|\beta|$, is a null operator when acting on $[\mathcal{M}^{(n-1)}_{n\mu}]$ and can be added to a raising or lowering operator without changing its raising or lowering property, the functions $f_{ni}, f_{n,-i}$ must be evaluated by comparing the ρ independent terms on each side of Eq. (4.37). This leads to

$$f_{2k+1,i} = \prod_{\alpha=i+1}^{k} a_{i,-\alpha}\prod_{\beta=1}^{k} a_{i\beta}, \quad (4.38)$$

$$f_{2k+1,-i} = \prod_{\gamma=1}^{i-1}(a_{-i,\gamma}+2), \quad (4.39)$$

$$f_{2k,-i} = \prod_{\alpha=1}^{i-1}(a_{-i,\alpha}+2), \quad (4.40)$$

$$f_{2k,i} = c_i b_{ii}\prod_{\beta=1}^{i-1} b_{i\beta}\prod_{\alpha=i+1}^{k-1} a_{i,-\alpha}b_{i\alpha}. \quad (4.41)$$

Thus

$$U_{ni} = (f_{ni}/N_{ni})O_{ni}, \quad (4.42)$$

$$U_{n,-i} = O_{n,-1}(f_{n,-i}/N_{ni}), \quad (4.43)$$

where N_{ni} is a factor which is defined to be real. With

$$\langle \mathcal{M}^{(n-1)}_{n\mu}|\, U_{ni}U_{n,-i}\,|\mathcal{M}^{(n-1)}_{n\mu}\rangle = 1,$$
$$\langle \mathcal{M}^{(n-1)}_{n\mu}|\, f_{ni}O_{ni}O_{n,-i}f_{n,-i}\,|\mathcal{M}^{(n-1)}_{n\mu}\rangle = N^2_{ni}. \quad (4.44)$$

Note that the U_{ni}, unlike the O_{ni}, do not form a commuting set of lowering (raising) operators, $[U_{n\alpha}, U_{n\beta}] \neq 0$, since $[f_{n\alpha}, O_{n\beta}] \neq 0$. However,

$$[O_{ni}O_{n,-i}, f_{n\alpha}] = 0, \quad \text{for any } \alpha. \quad (4.45)$$

Therefore,

$$N^2_{ni} = \langle \mathcal{M}^{(n-1)}_{n\mu}|\, f_{ni}f_{n,-i}\,|\mathcal{M}^{(n-1)}_{n\mu}\rangle\langle i|\, O_{ni}O_{n,-i}\,|i\rangle, \quad (4.46)$$

With Eqs. (4.19), (4.32), and (4.38)–(4.41)

$$N_{2k+1,i} = a^{(2k+1)}_{ii}\left|\prod_{\alpha=i+1}^{k} a_{i,-\alpha}a_{i\alpha}a^{(2k+1)}_{i\alpha^0}(a^{(2k+1)}_{i,-\alpha^0}-2)\right.$$
$$\left.\times\prod_{\beta=1}^{i-1} a_{i\beta}(a_{-i,\beta}+2)(a^{(2k+1)}_{-i,\beta^0}+2)a^{(2k+1)}_{i\beta^0}\right|^{\frac{1}{2}}, \quad (4.47)$$

$$N_{2k,i} = \left|\frac{b_{ii}c_i}{4}\,b^{(2k)}_{ii^0}(a^{(2k)}_{i^0,-i}+2)(a^{(2k)}_{i,-k^0}-2)b^{(2k)}_{ik^0}\right|^{\frac{1}{2}}$$
$$\times\left|\prod_{\alpha=1}^{i-1}(a_{-i,\alpha}+2)b_{i\alpha}b^{(2k)}_{i\alpha^0}(a^{(2k)}_{-i,\alpha^0}+2)\right.$$
$$\left.\times\prod_{\beta=i+1}^{n-1} a_{i,-\beta}b_{i\beta}(a^{(2k)}_{i,-\beta^0}-2)b^{(2k)}_{i\beta^0}\right|^{\frac{1}{2}}. \quad (4.48)$$

5. MATRIX ELEMENTS OF $J_{n-1,n}$

In the evaluation of the matrix elements of the infinitesimal generators, the matrix elements of $J_{n-1,n}$ play the fundamental role since the matrix elements of all other J_{ij} can be simply related to these. Matrix elements of $J_{n-1,n}$ have been given by Gel'fand and Zetlin.[6,7] A derivation of the Gel'fand–Zetlin result is given here to illustrate the usefulness of the lowering (raising) operators.

Since $J_{n-1,n}$ commutes with all J_{ij} with both $i, j < n-1$, $J_{n-1,n}$ is a scalar operator with respect to $O(n-2)$. The matrix elements of $J_{n-1,n}$ are thus diagonal in $m_{n-2,\alpha}$ and independent of $m_{\nu,\alpha}$, $\nu \leq n-3$. With respect to $O(n)$, $J_{n-1,n}$ transforms according to the regular representation $[11000\cdots]$. With respect to $O(n-1)$ its irreducible tensor character is that of the vector representation $[1000\cdots]$. It thus connects states in which any one of the $m_{n-1,\alpha}$ differ by ± 1 only. (For $n-1$ odd, it also has a diagonal matrix element.)

$$\langle \mathcal{M}'_{n\mu}|\, J_{n-1,n}\,|\mathcal{M}_{n\mu}\rangle = \langle \mathcal{M}'^{(n-2)}_{n\mu}|\, J_{n-1,n}\,|\mathcal{M}^{(n-2)}_{n\mu}\rangle$$
$$= \left\langle \begin{matrix} m_{ni} \\ m'_{n-1,i} \\ m_{n-2,i} \end{matrix}\right| J_{n-1n} \left|\begin{matrix} m_{ni} \\ m_{n-1,i} \\ m_{n-2,i} \end{matrix}\right\rangle. \quad (5.1)$$

For convenience, only the relevant m_{vi} in the one column subject to change are written out. The matrix elements in the $[\mathcal{M}_{n\mu}^{(n-2)}]$ basis could be evaluated through a construction involving successive application of lowering operators of type $U_{n,i}$ followed by $U_{n-1,i}$. It is more convenient to factor the matrix element of $J_{n-1,n}$ into two parts by using the Wigner-Eckart theorem. The reduced matrix element, independent of the $m_{n-2,\alpha}$, can be chosen as the matrix element of $J_{n-1,n}$ in the restricted basis $[\mathcal{M}_{n\mu}^{(n-1)}]$, while the $m_{n-2,\alpha}$-dependent factor can be expressed as the matrix element of a vector operator in $(n-1)$-dimensional space

$$\langle \mathcal{M}'_{n\mu}| J_{n-1,n} |\mathcal{M}_{n\mu}\rangle = \left\langle \begin{matrix} m_{ni} \\ m'_{n-1,i} \\ m_{n-1,i} \end{matrix} \right| J_{n-1,n} \left| \begin{matrix} m_{ni} \\ m_{n-1,i} \\ m_{n-1,i} \end{matrix} \right\rangle \times \left\langle \begin{matrix} m'_{n-1,i} \\ m_{n-2,i} \end{matrix} \right| V \left| \begin{matrix} m_{n-1,i} \\ m_{n-2,i} \end{matrix} \right\rangle, \quad (5.2)$$

where V has irreducible tensor character $[1000\cdots]$ with respect to $O(n-1)$ and $[000\cdots]$ with respect to $O(n-2)$, and its matrix element is normalized to unity when $m_{n-2,i} = m_{n-1,i}$ (all i). The first factor imposes the restriction $m'_{n-1,i} \geq m_{n-1,i}$. However, the matrix element with $m'_{n-1,i} = m_{n-1,i} - 1$ can be obtained from that with $m'_{n-1,i} = m_{n-1,i} + 1$ through the Hermiticity of $J_{n-1,n}$.

A. Evaluation of $\left\langle \begin{matrix} m_{ni} \\ m'_{n-1,i} \\ m_{n-1,i} \end{matrix} \right| J_{n-1i,n} \left| \begin{matrix} m_{ni} \\ m_{n-1,i} \\ m_{n-1,i} \end{matrix} \right\rangle$

(1) $n = 2k$

$O_{2k,k}$ is a linear combination of $J_{2k-1,2k}$ and $Q_{2k,\alpha}$. Re-expressing $O_{2k,k}$ instead as a linear combination of $J_{2k-1,2k}$ and $O_{2k,i}$

$$\left\{ O_{2k,k} = \left[J_{2k-1,2k} \prod_{\alpha=1}^{k-1} a_\alpha^{(2k)} + \sum_{\alpha=1}^{k-1} O_{2k-1,-\alpha} O_{2k,\alpha} h_i \right] \right\} \left| \begin{matrix} m_{2k,i} \\ m_{2k-1,i} \end{matrix} \right\rangle = 0, \quad (5.3)$$

where h_i are functions of $J_{2i-1,2i}$, $(i < k)$, which are to be determined from the conditions required for $O_{2k,k}$

$$\text{(a)} \qquad [\rho_{jl}, O_{2k,k}] \left| \begin{matrix} m_{n,j} \\ m_{n-1,j} \end{matrix} \right\rangle = 0, \qquad 0 < j < |l| \leq k-1, \quad (5.4)$$

$$\text{(b)} \qquad [Q_{2k-1,l}, O_{2k,k}] \left| \begin{matrix} m_{n,j} \\ m_{n-1,j} \end{matrix} \right\rangle = 0. \quad (5.5)$$

Condition (a) is automatically satisfied. In order to satisfy condition (b):

$$\left\{ [Q_{2k-1,j}, J_{2k-1,2k}] \prod_{\alpha=1}^{k-1} a_\alpha^{(2k)} + \sum_{\alpha=1}^{k-1} [Q_{2k-1,j}, O_{2k-1,-\alpha}] O_{2k\alpha} h_\alpha \right\} \left| \begin{matrix} m_{n,j} \\ m_{n-1,j} \end{matrix} \right\rangle = 0. \quad (5.6)$$

From the coefficients of $Q_{2k,i}$, however, the h_j follow directly

$$h_j = i \left(\prod_{\alpha=1}^{j-1} \frac{a_\alpha^{(2k)}}{(a_{j\alpha}^{(2k)} - 2) a_{-j,\alpha}^{(2k)}} \prod_{\beta=j+1}^{k-1} \frac{a_\beta^{(2k)}}{a_{j,-\beta}^{(2k)} (a_{j\beta}^{(2k)} - 2)} \right) \frac{1}{a_{jj}^{(2k)}}. \quad (5.7)$$

Also

$$O_{2k,k} \left| \begin{matrix} m_{nj} \\ m_{n-1,j} \end{matrix} \right\rangle = \frac{1}{2} \prod_{i=1}^{k} a_{j^0}^{(2k)} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle. \quad (5.8)$$

Re-expressing the $O_{2k-1,-\alpha}$ and $O_{2k,\alpha}$ of Eq. (5.3) in terms of $U_{2k-1,-\alpha}$ and $U_{2k,\alpha}$, the matrix element of $J_{2k-1,2k}$ can be read off from Eq. (5.3):

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right| J_{2k-1,2k} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = \left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right| \frac{a_{k^0}^{(2k)}}{2} \prod_{\alpha=1}^{k} \frac{a_{\alpha^0}^{(2k)}}{a_\alpha^{(2k)}} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle, \quad (5.9)$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} + 1 \\ m_{2k-1,j} \end{matrix} \right| J_{2k-1,2k} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = \left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right| \frac{-2i}{a_j^{(2k)} (a_j + 1)^{\frac{1}{2}}} \left| \left(\prod_{\alpha=1}^{j-1} \frac{a_{-j\alpha^0}^{(2k)} a_{j\alpha}^{(2k)}}{a_{-j\alpha}^{(2k)} a_{j\alpha}^{(2k)}} \prod_{\beta=j+1}^{k-1} \frac{a_{j-\beta^0}^{(2k)} a_{j\beta^0}^{(2k)}}{a_{j-\beta}^{(2k)} a_{j\beta}^{(2k)}} \right) \frac{a_{jj^0}^{(2k)}}{2} \frac{a_{j^0-j}^{(2k)}}{2} \frac{a_{j-k^0}^{(2k)}}{2} \frac{a_{jk^0}^{(2k)}}{2} \right|^{\frac{1}{2}} \left| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle \quad (j = 1, 2, \cdots, k-1) \quad (5.10)$$

(2) *The $n = 2k + 1$ Case*

The procedure is similar to that for $n = 2k$. First, since $J_{2k,2k+1}$ has no diagonal matrix elements, in place of the neutral operator there is now the relation

$$\left\{ J_{2k,2k+1} - \left[\sum_{\alpha=1}^{k-1} O_{2k,-\alpha} O_{2k+1,\alpha} h_\alpha + O_{2k+1,k} h_k + O_{2k+1,-k} h_{-k} \right] \right\} \left| \begin{matrix} m_{2k+1,i} \\ m_{2k,i} \end{matrix} \right\rangle = 0, \quad (5.11)$$

where the h_α are evaluated as before through

$$\left[Q_{2k,j}, J_{2k,2k+1} - \sum_{\alpha=1}^{k-1} O_{2k,-\alpha} O_{2k+1,\alpha} h_\alpha - O_{2k+1,k} h_k - O_{2k+1,-k} h_{-k}\right] \left|\begin{matrix} m_{2k+1,i} \\ m_{2k,i} \end{matrix}\right\rangle = 0. \quad (5.12)$$

From the coefficients of $Q_{2k+1,j}$,

$$h_j = i\left[2(b_{jj}^{(2k)} + 4)\left(\frac{a_{j-k}^{(2k)}}{2}\frac{a_{jk}^{(2k)}}{2}\right) \times \prod_{\alpha=1}^{j-1} a_{j\alpha}^{(2k)} a_{-j,\alpha}^{(2k)} \prod_{\beta=j+1}^{k-1} a_{j,-\beta}^{(2k)} a_{j\beta}^{(2k)}\right)^{-1}. \quad (5.13)$$

Similarly, re-expressing the O operators in terms of U operators, the matrix element of $J_{2k,2k+1}$ can be read off from Eq. (5.11).

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j}+1 \\ m_{2k,j} \end{matrix} \middle| J_{2k,2k+1} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k,j} \end{matrix} \right\rangle = \left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \end{matrix} \middle| \frac{-i}{2} \middle| \frac{a_{-jj^0}^{(2k+1)}(a_{jj^0}^{(2k+1)}+2)a_j^{(2k+1)}}{a_{jk}^{(2k+1)}(a_{j-k}^{(2k+1)}+2)} \middle|^{\frac{1}{2}} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \end{matrix} \right\rangle$$
$$\times \left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \end{matrix} \middle| \prod_{\alpha=1}^{j-1} \middle| \frac{a_{-j\alpha^0}^{(2k+1)}(a_{j\alpha^0}^{(2k+1)}+2)}{a_{-j\alpha}^{(2k+1)}(a_{j\alpha}^{(2k+1)}+2)} \middle|^{\frac{1}{2}} \prod_{\beta=j+1}^{k} \middle| \frac{a_{j-\beta^0}^{(2k+1)}(a_{j\beta^0}^{(2k+1)}+2)}{a_{j-\beta}^{(2k+1)}(a_{j\beta}^{(2k+1)}+2)} \middle|^{\frac{1}{2}} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \end{matrix} \right\rangle, \quad j = 1, 2, \cdots, k. \quad (5.14)$$

B. Evaluation of $\left\langle \begin{matrix} m'_{n-1,i} \\ m_{n-2,i} \end{matrix} \middle| V^{(n-1)} \middle| \begin{matrix} m_{n-1,i} \\ m_{n-2,i} \end{matrix} \right\rangle$

$V^{(n-1)}$ has the transformation properties of

$$\left|\begin{matrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \end{matrix}\right\rangle$$

and is to be normalized such that

$$\left\langle \begin{matrix} m_{n-1,1} & m_{n-1,2} & m_{n-1,3} & \cdots & m'_{n-1,i} & m_{n-1,i+1} & \cdots \\ m_{n-1,1} & m_{n-1,2} & m_{n-1,3} & \cdots & m_{n-1,i} & m_{n-1,i+1} & \cdots \end{matrix} \middle| V^{(n-1)} \middle| \begin{matrix} m_{n-1,1} & m_{n-1,2} & \cdots & m_{n-1,i} & \cdots \\ m_{n-1,1} & m_{n-1,2} & \cdots & m_{n-1,i} & \cdots \end{matrix} \right\rangle = 1. \quad (5.15)$$

It is convenient to introduce the following shorthand notation. Change $m_{ni} \to \alpha_i$ $m_{n-1,i} \to \beta_i$ $m_{n-2,i} \to \gamma_i \leq \beta_i$ and define

$$\left|\begin{matrix} \beta_i \\ \gamma_i \end{matrix}\right\rangle = \left|\begin{matrix} \beta_1 & \beta_2 & \cdots & \beta_{i-1} & \beta_i & \beta_{i+1} & \beta_{i+2} & \cdots \\ \beta_1 & \beta_2 & \cdots & \beta_{i-1} & \gamma_i & \gamma_{i+1} & \gamma_{i+2} & \cdots \end{matrix}\right\rangle, \quad (5.16)$$

$$\left|\begin{matrix} \beta_j+1 & \beta_i \\ \beta_j & \gamma_i \end{matrix}\right\rangle = \left|\begin{matrix} \beta_1 & \beta_2 & \cdots & \beta_j+1 & \beta_{j+1} & \beta_{j+2} & \cdots & \beta_{i-1} & \beta_i & \beta_{i+1} & \beta_{i+2} & \cdots \\ \beta_1 & \beta_2 & \cdots & \beta_j & \beta_{j+1} & \beta_{j+2} & \cdots & \beta_{i-1} & \gamma_i & \gamma_{i+1} & \gamma_{i+2} & \cdots \end{matrix}\right\rangle, \quad (5.17)$$

$$\left|\begin{matrix} \beta_i & \beta_j+1 \\ \gamma_i & \gamma_j \end{matrix}\right\rangle = \left|\begin{matrix} \beta_1 & \beta_2 & \cdots & \beta_{i-1} & \beta_i & \beta_{i+1} & \cdots & \beta_{j-1} & \beta_j+1 & \beta_{j+1} & \cdots \\ \beta_1 & \beta_2 & \cdots & \beta_{i-1} & \gamma_i & \gamma_{i+1} & \cdots & \gamma_{j-1} & \gamma_j & \gamma_{j+1} & \cdots \end{matrix}\right\rangle, \quad (5.18)$$

$$\left|\begin{matrix} 1 \\ 0 \end{matrix}\right\rangle = \left|\begin{matrix} 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \end{matrix}\right\rangle. \quad (5.19)$$

(1) *The $n = 2k$ Case*

Define coefficients Γ_{ij} by the relation

$$\left|\begin{matrix} 1 \\ 0 \end{matrix}\right\rangle \left|\begin{matrix} \beta_i \\ \gamma_i \end{matrix}\right\rangle = \sum_{j=1}^{i-1} \Gamma_{ij} \binom{\beta_j \ \ \beta_i}{\beta_j \ \ \gamma_i} \left|\begin{matrix} \beta_j+1 & \beta_i \\ \beta_j & \gamma_i \end{matrix}\right\rangle + \sum_{j=i}^{k-1} \Gamma_{ij} \binom{\beta_i \ \ \beta_j}{\gamma_i \ \ \gamma_j} \left|\begin{matrix} \beta_i & \beta_j+1 \\ \gamma_i & \gamma_j \end{matrix}\right\rangle + \sum_{j=i}^{k-1} \Gamma_{i-j} \binom{\beta_i \ \ \beta_j-1}{\gamma_i \ \ \gamma_j} \left|\begin{matrix} \beta_i & \beta_j-1 \\ \gamma_i & \gamma_j \end{matrix}\right\rangle + \Gamma_{i0} \binom{\beta_i}{\gamma_i} \left|\begin{matrix} \beta_i \\ \gamma_i \end{matrix}\right\rangle, \quad (5.20)$$

where the Γ_{ij} are generalized Wigner coefficients for the Kronecker product $[100\cdots] \times [\beta_1\beta_2\beta_3\cdots]$ of $O(2k-1)$. Note that Γ_{1j}, the coefficient with *all* $\gamma_i < \beta_i$ starting with γ_1, is equal to the matrix element of $V^{(2k-1)}$ provided $\Gamma_{k-1,j}$, the coefficient with *all* $\gamma_i = \beta_i$, satisfies the normalization condition $\Gamma_{k-1,j} = 1$ required by Eq. (5.15). The coefficients Γ_{ij} can be related to the coefficients $\Gamma_{i+1,j}$ by recursion techniques, leading after repeated recursion to a relation between Γ_{1j} and $\Gamma_{k-1,j}$. Since the recursion is to be established through the raising generator Q_i, it is necessary to define further coefficients, Ω_{ij}, by the

relation

$$\left\{Q_{2k-1,i}\begin{vmatrix}1\\0\end{vmatrix}\right\}\begin{vmatrix}\overline{\beta_i}\\\gamma_i\end{vmatrix}=\sum_{j=1}^{i-1}\Omega_{ij}\begin{pmatrix}\beta_j & \beta_i\\ \beta_j & \gamma_i\end{pmatrix}\begin{vmatrix}\overline{\beta_j+1 \quad \beta_i}\\ \beta_j \quad \gamma_i+1\end{vmatrix}$$
$$+\sum_{j=i}^{k-1}\Omega_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}\begin{vmatrix}\overline{\beta_i \quad \beta_j+1}\\ \gamma_i+1 \quad \gamma_j\end{vmatrix}$$
$$+\sum_{j=i}^{k-1}\Omega_{i,-j}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}\begin{vmatrix}\overline{\beta_i \quad \beta_j-1}\\ \gamma_i+1 \quad \gamma_j\end{vmatrix}$$
$$+\Omega_{i0}\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\begin{vmatrix}\overline{\beta_i}\\ \gamma_i+1\end{vmatrix}$$
$$+\sum_{j=1}^{i-1}\Lambda(\rho)\begin{vmatrix}\overline{\beta_j+1 \quad \beta_i}\\ \beta_j+1 \quad \gamma_i\end{vmatrix}. \tag{5.21}$$

The operators $\Lambda(\rho)$ of the last term of the equation, when acting on states (5.16), create states outside the basis $[\mathcal{M}^{(2k-2)}_{2k-1,\mu}]$. These are orthogonal to the states of present interest so that the last term of Eq. (5.21) plays no further role in the discussion.

Applying $Q_{2k-1,i}$ again to Eq. (5.21), a set of recursion relations is established for the coefficients Ω_{ij}

$$\frac{\Omega_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}}{\Omega_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i+1 & \gamma_j\end{pmatrix}}=\frac{q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}}{q_i\begin{pmatrix}\beta_i & \beta_j+1\\ \gamma_i+1 & \gamma_j\end{pmatrix}}, \tag{5.22}$$

$$\frac{\Omega_{i0}\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}}{\Omega_{i0}\begin{pmatrix}\beta_i\\ \gamma_i+1\end{pmatrix}}=\frac{q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}}{q_i\begin{pmatrix}\beta_i\\ \gamma_i+1\end{pmatrix}}, \tag{5.23}$$

where

$$\left\langle\begin{vmatrix}\overline{\beta_i}\\ \gamma_i+1\end{vmatrix}\right|Q_{2k-1,i}\left|\begin{vmatrix}\overline{\beta_i}\\ \gamma_i\end{vmatrix}\right\rangle=q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}. \tag{5.24}$$

Applying $Q_{2k-1,i}$ also to Eq. (5.20), another set of recursion relations is established,

$$\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}q_i\begin{pmatrix}\beta_i & \beta_j+1\\ \gamma_i & \gamma_j\end{pmatrix}$$
$$=q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i+1 & \gamma_j\end{pmatrix}+\Omega_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}, \tag{5.25}$$

$$\Gamma_{i0}\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}=q_i\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\Gamma_{i0}\begin{pmatrix}\beta_i\\ \gamma_i+1\end{pmatrix}+\Omega_{i0}\begin{pmatrix}\beta_i\\ \gamma_j\end{pmatrix}. \tag{5.26}$$

The recursion process for Γ_{ij} can be started if the coefficients Γ_{ij}, Ω_{ij} can be related for a particular value of γ_i. The cases Γ_{ij} and Γ_{i0} are somewhat different.

a. *The Case* Γ_{ij}. From Eqs. (4.1)

$$q_i\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}=\Big|(\beta_i-\gamma_i)(\beta_i+\gamma_i+2k-2i-1)$$
$$\times\prod_{\lambda=i+1}^{k-1}\frac{a_{i,-\lambda^0}(a_{i,\lambda^0}+2)}{(a_{i,\lambda}+2)(a_{i,-\lambda}+2)}\Big|^{\frac{1}{2}}. \tag{5.27}$$

For $m_i=\beta_i-j+i+1$

$$0=q_i\begin{pmatrix}\beta_i & \beta_j+1\\ m_i & \gamma_j\end{pmatrix}\neq q_i\begin{pmatrix}\beta_i\\ m_i\end{pmatrix} \tag{5.28}$$

and Eq. (5.25) reduces to

$$q_i\begin{pmatrix}\beta_i\\ m_i\end{pmatrix}\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ m_i+1 & \gamma_j\end{pmatrix}+\Omega_{ij}\begin{pmatrix}\beta_i & \beta_j\\ m_i & \gamma_j\end{pmatrix}=0. \tag{5.29}$$

With this starting relation and the recursion relations (5.22) and (5.25),

$$\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \gamma_i & \gamma_j\end{pmatrix}=\frac{\gamma_i-\beta_j+j-i-1}{\beta_i-\beta_j+j-i-1}$$
$$\times\prod_{\lambda_i=0}^{\beta_i-\gamma_i-1}\frac{q_i\begin{pmatrix}\beta_i\\ \gamma_i+\lambda_i\end{pmatrix}}{q_i\begin{pmatrix}\beta_i & \beta_j+1\\ \gamma_i+\lambda_i & \gamma_j\end{pmatrix}}\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \beta_i & \gamma_j\end{pmatrix}, \tag{5.30}$$

where Γ_{ij} has been related to

$$\Gamma_{i+1,j}\left[\equiv\Gamma_{ij}\begin{pmatrix}\beta_i & \beta_j\\ \beta_i & \gamma_j\end{pmatrix}\right].$$

In the same way the relationship $\Gamma_{1j}\to\Gamma_{2j}\to\Gamma_{3j}\to\cdots\to\Gamma_{jj}$ can be established, leading to

$$\Gamma_{1j}\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_j & \cdots & \beta_{k-1}\\ \gamma_1 & \gamma_2 & \cdots & \gamma_j & \cdots & \gamma_{k-1}\end{pmatrix}=\Gamma_{jj}\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_j & \beta_{j+1} & \beta_{j+2} & \cdots\\ \beta_1 & \beta_2 & \cdots & \beta_j & \gamma_{j+1} & \gamma_{j+2} & \cdots\end{pmatrix}$$
$$\times\left|\prod_{i=1}^{j}\frac{(\gamma_i-\beta_j+j-i-1)(\gamma_i+\beta_j+2k-i-j-1)}{(\beta_i-\beta_j+j-i-1)(\beta_i+\beta_j+2k-i-j-1)}\right|^{\frac{1}{2}}. \tag{5.31}$$

So far the recursive chain stops at $i=j$ since Eqs. (5.22) and (5.25) are valid only if $i\leq j$. To complete the recursive chain, the relationship $\Gamma_{jj}\to\Gamma_{j+1,j}\to\cdots\to\Gamma_{k-1,j}$ must be established. For this

purpose consider

$$\text{(a)}\quad \left|\begin{matrix}1\\0\end{matrix}\right\rangle \left|\begin{matrix}\beta_j & \beta_{j+1}\\ \beta_j & \gamma_{j+1}\end{matrix}\right\rangle = \gamma_{jj}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}} \left|\begin{matrix}\beta_j+1 & \beta_{j+1}\\ \beta_j & \gamma_{j+1}\end{matrix}\right\rangle + \cdots, \tag{5.32}$$

$$\text{(b)}\quad \left|\begin{matrix}1\\0\end{matrix}\right\rangle \left|\begin{matrix}\beta_j & \beta_{j+1}\\ \beta_j & \gamma_{j+1}+1\end{matrix}\right\rangle = \gamma_{jj}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}+1} \left|\begin{matrix}\beta_j+1 & \beta_{j+1}\\ \beta_j & \gamma_{j+1}+1\end{matrix}\right\rangle + \cdots. \tag{5.33}$$

[The omitted states are similar to those of Eq. (5.21). They are orthogonal to the states of present interest.] Operating with $Q_{2k-1,i}$ on (b), and with $Q_{2k-1,j+1}Q_{2k-1,j}$ on (a) and comparing the two equations

$$\frac{\Gamma_{jj}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}}}{\Gamma_{jj}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}+1}} = \frac{q_{j+1}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}} q_j\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}+1}}{q_j\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}} q_{j+1}\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j+1 \quad \gamma_{j+1}}}. \tag{5.34}$$

Then

$$\Gamma_{j+1,j}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}} = \Gamma_{j+2,j}\binom{\beta_{j+1} \quad \beta_{j+2}}{\beta_{j+1} \quad \gamma_{j+2}} \frac{q_j\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j \quad \beta_{j+1}}}{q_j\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}}} \times \prod_{\lambda_{j+1}=0}^{\beta_{j+1}-\gamma_{j+1}-1} \frac{q_{j+1}\binom{\beta_j \quad \beta_{j+1}}{\beta_j \quad \gamma_{j+1}+\lambda_{j+1}}}{q_{j+1}\binom{\beta_j+1 \quad \beta_{j+1}}{\beta_j+1 \quad \gamma_{j+1}+\lambda_{j+1}}}. \tag{5.35}$$

By repeated application of Eq. (5.35), finally

$$\Gamma_{jj}\binom{\beta_j \quad \beta_{j+1} \quad \cdots}{\beta_j \quad \gamma_{j+1} \quad \cdots} = \Gamma_{k-1,j}\binom{\beta_{k-1}}{\beta_{k-1}} \times \prod_{i=j+1}^{k-1} \frac{q_j\binom{\beta_j+1 \quad \beta_i}{\beta_j \quad \beta_i}}{q_j\binom{\beta_j+1 \quad \beta_i}{\beta_j \quad \gamma_i}} \prod_{\lambda_i=0}^{\beta_i-\gamma_i-1} \frac{q_i\binom{\beta_j \quad \beta_i}{\beta_j \quad \gamma_i+\lambda_i}}{q_i\binom{\beta_j+1 \quad \beta_i}{\beta_j+1 \quad \gamma_i+\lambda_i}}. \tag{5.36}$$

Combining Eqs. (5.31) and (5.36) with the restriction $\Gamma_{k-1,j}=1$,

$$\left\langle\begin{matrix}\beta_1 & \beta_2 & \cdots & \beta_j+1 & \beta_{j+1} & \beta_{j+2} & \cdots & \beta_{k-1}\\ \gamma_1 & \gamma_2 & \cdots & \gamma_j & \gamma_{j+1} & \gamma_{j+2} & \cdots & \gamma_{k-1}\end{matrix}\right| V^{(2k-1)} \left|\begin{matrix}\beta_1 & \beta_2 & \cdots & \beta_j & \beta_{j+1} & \cdots & \beta_{k-1}\\ \gamma_1 & \gamma_2 & & \gamma_j & \gamma_{j+1} & \cdots & \gamma_{k-1}\end{matrix}\right\rangle = \left|\prod_{i=1}^{k-1} \frac{(\beta_j-\gamma_i+i-j+1)(\beta_j+\gamma_i+2k-i-j-1)}{(\beta_j-\beta_i+i-j+1)(\beta_j+\beta_i+2k-i-j-1)}\right|^{\frac{1}{2}}. \tag{5.37}$$

b. *The Case* Γ_{i0}. From (5.23) and (5.26)

$$\Gamma_{i0}\binom{\beta_i}{\gamma_i} = \Gamma_{i0}\binom{\beta_i}{\beta_i} + (\beta_i-\gamma_i)\frac{\Omega_{i0}\binom{\beta_i}{\beta_i-1}}{q_i\binom{\beta_i}{\beta_i-1}}. \tag{5.38}$$

In order to start the recursion process, the relation between $\Gamma_{i0}\binom{\beta_i}{\beta_i}$ and $\Omega_{i0}\binom{\beta_i}{\beta_i-1}\Big/q_i\binom{\beta_i}{\beta_i-1}$ must be known. The technique used for the case Γ_{ij} cannot be applied here. However, by applying the quadratic invariant to both sides of

$$\left\{Q_{2k-1,i}\left|\begin{matrix}1\\0\end{matrix}\right\rangle\right\} \left|\begin{matrix}\beta_i\\ \beta_i-1\end{matrix}\right\rangle = \Omega_{i0}\binom{\beta_i}{\beta_i-1} \left|\begin{matrix}\beta_i\\ \beta_i\end{matrix}\right\rangle + \cdots, \tag{5.39}$$

the desired relation is obtained (details are given in Appendix B) as

$$\Gamma_{i0}\binom{\beta_i}{\beta_i} = -(\beta_i+k-i-1) \times \Omega_{i0}\binom{\beta_i}{\beta_i-1}\Big/q_i\binom{\beta_i}{\beta_i-1}, \tag{5.40}$$

$$\Gamma_{10}\binom{\beta_1}{\gamma_1} = \prod_{i=1}^{k-1}\frac{\gamma_i+k-i-1}{\beta_i+k-i-1}\Gamma_{k-1,0}\binom{\beta_{k-1}}{\beta_{k-1}}, \tag{5.41}$$

so that

$$\left\langle\begin{matrix}\beta_1 & \beta_2 & \cdots & \beta_i & \cdots & \beta_{k-1}\\ \gamma_1 & \gamma_2 & & \gamma_i & & \gamma_{k-1}\end{matrix}\right| \times V^{(2k-1)} \left|\begin{matrix}\beta_1 & \beta_2 & \cdots & \beta_i & \cdots & \beta_{k-1}\\ \gamma_1 & \gamma_2 & \cdots & \gamma_i & \cdots & \gamma_{k-1}\end{matrix}\right\rangle = \prod_{i=1}^{k-1}\frac{\gamma_i+k-i-1}{\beta+k-i-1}. \tag{5.42}$$

(2) *The Case $n = 2k + 1$*

The procedure is exactly the same, except that the term Γ_{i0} does not exist ($V^{(2k)}$ has no diagonal matrix element). The result is

$$\left\langle \begin{matrix} \beta_1 & \beta_2 & \cdots & \beta_{j-1} & \beta_j+1 & \beta_{j+1} & \cdots & \beta_{k-1} & \beta_k \\ \gamma_1 & \gamma_2 & \cdots & \gamma_{j-1} & \gamma_j & \gamma_{j+1} & & \gamma_{k-1} & \end{matrix} \middle| V^{(2k)} \middle| \begin{matrix} \beta_1 & \beta_2 & \cdots & \beta_{j-1} & \beta_j+1 & \beta_{j+1} & \cdots & \beta_{k-1} & \beta_k \\ \gamma_1 & \gamma_2 & & \gamma_{j-1} & \gamma_j & \gamma_{j+1} & \cdots & \gamma_{k-1} & \end{matrix} \right\rangle = \left| \prod_{i=1}^{k-1} \frac{(\beta_j - \gamma_i + i - j + 1)(\beta_j + \gamma_i + 2k - i - j)}{(\beta_j - \beta_i + i - j + 1)(\beta_j + \beta_i + 2k - i - j)} \right|^{\frac{1}{2}}. \tag{5.43}$$

C. Evaluation of $\left\langle \begin{matrix} m_{nj} \\ m'_{n-1,n} \\ m_{n-2,j} \end{matrix} \middle| J_{n-1,n} \middle| \begin{matrix} m_{nj} \\ m_{n-1,j} \\ m_{n-2,j} \end{matrix} \right\rangle$

Combining the results of subsections A and B above, Eqs. (5.2), (5.9), (5.10), (5.14), (5.37), (5.42), and (5.43), the Gel'fand–Zetlin matrix elements are obtained. With

$$l_{2k,\alpha} = m_{2k,\alpha} + k - \alpha,$$
$$l_{2k-1,\alpha} = m_{2k-1,\alpha} + k - \alpha,$$

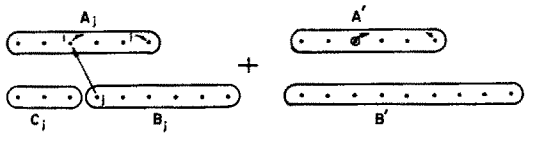
$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j}+1 \\ m_{2k-2,j} \end{matrix} \middle| J_{2k-1,2k} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \right\rangle = -i \left| \frac{\prod_{\alpha=1}^{k-1} (l_{2k-2,\alpha}^2 - l_{2k-1,j}^2) \prod_{\beta=1}^{k} (l_{2k,\beta}^2 - l_{2k-1,j}^2)}{l_{2k-1,j}^2 (4l_{2k-1,j}^2 - 1) \prod_{\alpha \neq j}^{k-1} (l_{2k-1,\alpha}^2 - l_{2k-1,j}^2)[(l_{2k-1,\alpha} - 1)^2 - l_{2k-1,j}^2]} \right|^{\frac{1}{2}}, \tag{5.44}$$

$$\left\langle \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \middle| J_{2k-1,2k} \middle| \begin{matrix} m_{2k,j} \\ m_{2k-1,j} \\ m_{2k-2,j} \end{matrix} \right\rangle = \frac{\prod_{\alpha=1}^{k-1} l_{2k-2,\alpha} \prod_{\beta=1}^{k} l_{2k,\beta}}{\prod_{\alpha=1}^{k-1} l_{2k-1,\alpha}(l_{2k-1,\alpha} - 1)}, \tag{5.45}$$

$$\left\langle \begin{matrix} m_{2k+1,j} \\ m_{2k,j}+1 \\ m_{2k-1,j} \end{matrix} \middle| J_{2k,2k+1} \middle| \begin{matrix} m_{2k+1,j} \\ m_{2k,j} \\ m_{2k-1,j} \end{matrix} \right\rangle = \frac{-i}{2} \left| \frac{\prod_{\alpha=1}^{k-1} (l_{2k-1,\alpha} - l_{2k,j} - 1)(l_{2k-1,\alpha} + l_{2k,j}) \prod_{\beta=1}^{k} (l_{2k+1,\beta} - l_{2k,j} - 1)(l_{2k+1,\beta} + l_{2k,j})}{\prod_{\alpha \neq j}^{k} (l_{2k\alpha}^2 - l_{2kj}^2)[l_{2k\alpha}^2 - (l_{2kj} + 1)^2]} \right|^{\frac{1}{2}}. \tag{5.46}$$

APPENDIX A. EVALUATION OF $\langle i \mid Q_{2k+1,i} O_{2k+1,-i} \mid i \rangle$

There are many graphs in $O_{2k+1,-i}$. For some types of calculation certain ways of grouping them are more convenient than others. The following example demonstrates one way.



$$O_{2k+1,-i} = \sum_{j=1}^{k} \sum_{p=1}^{j} \sum_{l=i+1}^{k} \{-i, l\}(-\rho_{-l,-j})\{j, -p\} Q_{2k+1,p} \times \prod_{\gamma=j+1}^{k} a_{i,\gamma} \prod_{\gamma=l+1}^{k} a_{i,-\gamma} \prod_{\gamma=1}^{p-1} a_{i,\gamma} + \sum_{j=i}^{k} \{-i, j\} Q_{2k+1,-j} \prod_{\gamma=1}^{k} a_{i\gamma} \prod_{\gamma=j+1}^{k} a_{i,-\gamma}, \tag{A1}$$

$$\left.\begin{aligned} A_j &= \sum_{l=i+1}^{k} \{-i, l\}(-\rho_{-l,-j}) \prod_{\gamma=l+1}^{k} a_{i,-\gamma}, \qquad \{-i, i\} = \{j, -j\} = 1 \\ B_j &= \sum_{p=1}^{j} \{j, -p\} Q_{2k+1,p} \prod_{\gamma=1}^{p-1} a_{i\gamma}, \\ C_j &= \prod_{\gamma=j+1}^{k} a_{i\gamma}, \\ A' &= \sum_{j=i}^{k} \{-i, j\} Q_{2k+1,-j} \prod_{\gamma=j+1}^{k} a_{i,-\gamma}, \\ B' &= \prod_{\gamma=1}^{k} a_{i\gamma}, \end{aligned}\right\} \tag{A2}$$

$$O_{2k+1,-i} = \sum_{j=1}^{k} A_j B_j C_j + A' B', \tag{A3}$$

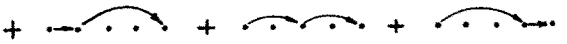
where $\{-\alpha, \beta\}$ means summation of all possible graphs, which have a chain away from the $|\beta|$th point

on the top when β is positive and on the bottom when β is negative; and a chain ending at the $|\alpha|$th point on the top when α is positive and on the bottom when α is negative.

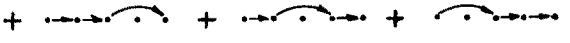
Example:

$\{-1, 5\}$ includes the following graphs in the top row:

(With one arrow link of the chain and three points.)



(All the possible distinct graphs generated by removing any one free point from the first graph, with the appropriate chains.)



(All the possible distinct graphs generated by removing any two free points from the first graph, with the appropriate chains.)

(All the possible distinct graphs generated by removing three free points from the first graph.)

Since the distribution of free points uniquely determines the graph, it is sometimes more convenient to define the graph by its free points.

With

$$\langle i|\, Q_{2k+1,\alpha}(-\rho_{-\alpha\beta}) = \langle i|\, (-2)Q_{2k+1,\beta} \quad 0 < \alpha < |\beta| \tag{A4}$$

each chain contraction gives a factor (-2), and

$$\begin{aligned}\langle i|\, Q_{2k+1,i}\{-i, l\} &= -2\langle i|\, Q_{2k+1,l} \prod_{\alpha=i+1}^{l-1} a_{i,-\alpha}\Bigg[1 + (-2)\sum_{\beta=i+1}^{l-1}\frac{1}{a_{i,-\beta}} \\ &\quad + (-2)^2 \sum_{\substack{\beta,\gamma=i+1\\ \beta\neq\gamma}}^{l-1} \frac{1}{a_{i,-\beta}a_{i,-\gamma}} + \cdots (-2)^{l-i-1}\frac{1}{\prod_{\alpha=i+1}^{l-1} a_{i,-\alpha}}\Bigg].\end{aligned} \tag{A5}$$

The second term in the parenthesis comes from the removal of one free point from the first graph which is

$$(-\rho_{-il}) \prod_{\alpha=i+1}^{l-1} a_{i,-\alpha},$$

and the last term in the parenthesis comes from the removal of all free points from the first graph. A similar removal of free points gives the intermediate terms. By summing up all the terms,

$$\langle i|\, Q_{2k+1,i}\{-i, l\} = -2\langle i|\, Q_{2k+1,l} \prod_{\alpha=i+1}^{l-1} (a_{i,-\alpha} - 2). \tag{A6}$$

Similarly

$$\begin{aligned}\langle i|\, Q_{2k+1,i}A_j = \langle i|\, (-2)Q_{2k+1,-j}\Bigg[&\prod_{\gamma=i+1}^{k}(a_{i,-\gamma} - 2) \\ &+ 2\prod_{\gamma=i+1}^{j-1}(a_{i,-\gamma} - 2)\prod_{\gamma=j+1}^{k} a_{i,-\gamma}\Bigg].\end{aligned} \tag{A7}$$

The first term is the summation of the contribution of all possible graphs with any number of free points from $i + 1$ to k in the A_j; but it has included the graphs

$$\{-i, j\}(-\rho_{-j,-j}) \prod_{\alpha=j+1}^{k} a_{i,-\alpha},$$

which should be zero, since $\rho_{-j-j} = 0$. The second term is therefore needed to take away the improper contribution of

$$\{-i, j\}(-\rho_{-j,-j}) \prod_{\alpha=j+1}^{k} a_{i,-\alpha}.$$

Similarly

$$\begin{aligned}\langle i|\, Q_{2k+1,-j}B_j = \langle i|\, &\sum_{p=1}^{j-1}(-2)Q_{2k+1,-p}Q_{2k+1,p} \\ &\times \prod_{\alpha=p+1}^{j-1}(a_{i\alpha} - 2)\prod_{\alpha=1}^{p-1} a_{i\alpha} + Q_{2k+1,-j}Q_{2k+1,j}\prod_{\alpha=1}^{j-1} a_{i\alpha}.\end{aligned} \tag{A8}$$

With Eq. (4.12) and Eq. (A2)

$$\langle i|\, Q_{2k+1,-j}B_j\,|i\rangle = \langle i|\, Q_{2k+1,-i}Q_{2k+1,i}\frac{2\prod_{\alpha=1}^{i-1} a_{i\alpha}\prod_{\alpha=i+1}^{j} a_{i\alpha}\prod_{\alpha=i+1}^{j-1}(a_{i,-\alpha} + 2)}{\prod_{\alpha=i+1}^{j} a_{i,-\alpha}}|i\rangle. \tag{A9}$$

By combining with (A7), (A2), and (A9)

$$\langle i|\, Q_{2k+1,i}\sum_{j=i}^{k} A_jB_jC_j\,|i\rangle = \sum_{j=i}^{k} - 4\langle i|\, \frac{Q_{2k+1,-i}Q_{2k+1,i}\prod_{\alpha=i+1}^{j-1}(a_{i-\alpha}^2 - 4)\prod_{\alpha=1}^{k} a_{i\alpha}\Big[\prod_{\alpha=j}^{k}(a_{i-\alpha} - 2) + 2\prod_{\alpha=j+1}^{k} a_{i,-\alpha}\Big]}{a_{ii}\prod_{\alpha=i+1}^{j} a_{i,-\alpha}}|i\rangle. \tag{A10}$$

Similarly, with Eq. (4.12), Eq. (A2), and Eq. (A6)

$$\langle i|\, Q_{2k+1,i}A'B'\,|i\rangle = \sum_{j=i}^{k} -4\langle i|\, J_{2j-1,2j}\prod_{\alpha=i+1}^{j-1}(a_{i,-\alpha}-2)\prod_{\alpha=j+1}^{k}a_{i,-\alpha}\prod_{\alpha=1}^{k}a_{i\alpha}\,|i\rangle$$
$$+\sum_{j=i}^{k} -4\langle i|\, Q_{2k+1,-i}Q_{2k+1,i}\frac{\prod_{\alpha=i+1}^{j-1}(a_{i,-\alpha}^2-4)\prod_{\alpha=j+1}^{k}a_{i,-\alpha}\prod_{\alpha=1}^{k}a_{i\alpha}}{\prod_{\alpha=i+1}^{j}a_{i,-\alpha}}\,|i\rangle. \tag{A11}$$

By summing up Eq. (A10) and (A11), finally

$$\langle i|\, Q_{2k+1,i}O_{2k+1,-i}\,|i\rangle = \langle i|\left[\prod_{\alpha=1}^{k}a_{i\alpha}\prod_{\alpha=i+1}^{k}(a_{i,-\alpha}-2)\right](m_{2k+1,i}-J_{2i-1,2i}+1)(m_{2k+1,i}+J_{2i-1,2i}+2k-2i)\,|i\rangle. \tag{A12}$$

A similar process works for $O(2k)$, and gives

$$\langle i|\, Q_{2k,i}O_{2k,-i}\,|i\rangle = \langle i|\,(m_{2k,i}+J_{2i-1,2i}+2k-2i-1)(m_{2k,i}-J_{2i-1,2i}+1)\,|i\rangle$$
$$\times\langle i|\prod_{\alpha=i+1}^{k}(a_{i,-\alpha}-2)\prod_{\alpha=1}^{k}b_{i\alpha}\,|i\rangle\langle i|\frac{1}{b_{ii}}\,|i\rangle. \tag{A13}$$

APPENDIX B. DERIVATION OF EQ. (5.40)

To derive Eq. (5.40),

$$\Gamma_{i0}\begin{pmatrix}\beta_i\\ \beta_i\end{pmatrix} = -(\beta_i+k-i-1)\times\Omega_{i0}\begin{pmatrix}\beta_i\\ \beta_i-1\end{pmatrix}\Big/q_i\begin{pmatrix}\beta_i\\ \beta_i-1\end{pmatrix}, \tag{5.40}$$

the quadratic Casimir invariant C_{2k-1} is applied to Eq. (5.39). In order to simplify the evaluation of these terms, the following points are useful:

(1) $$Q_{\alpha i}\left|\begin{matrix}\beta_i\\ \gamma_i\end{matrix}\right\rangle = 0 \quad i>0,\quad \alpha<2k-1 \tag{B1}$$

since $\left|\begin{matrix}\beta_i\\ \gamma_i\end{matrix}\right\rangle$ belongs to $[\mathcal{M}^{(2k-2)}_{2k-1\mu}]$

(2)a. $$Q_{2k-1,j}Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle = 0,\quad i\neq -j. \tag{B2}$$

The net result of the two Q operations in succession would either have to change one of the $m_{2k-2,\alpha}$ by two or two of the $m_{2k-2,\alpha}$ by one each. Both cases are impossible since $m_{2k-1,\alpha}$ is $[1000\cdots]$.

b. $$Q_{2k-1,-i}Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle = 2\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle, \tag{B3}$$

a direct consequence of Eq. (4.14) and Eq. (4.15).

(3) $$C_{2k-1} = \sum_{i<j}^{2k-1}J_{ij}^2$$
$$= \sum_{j=1}^{k-1}Q_{2k-1,-j}Q_{2k-1,j} + \sum_{j<\alpha\le k-1}Q_{2\alpha,-j}Q_{2\alpha,j}$$
$$+\sum_{j<\alpha\le k-1}Q_{2\alpha-1,-j}Q_{2\alpha-1,j} + \sum_{\alpha=1}^{k-1}J_{2\alpha-1,2\alpha}^2$$
$$+\sum_{\alpha=1}^{k-1}(2k-2\alpha-1)J_{2\alpha-1,2\alpha}. \tag{B4}$$

Proof: Operating on Eq. (5.39):

$$\left\{Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle\right\}\left|\begin{matrix}\beta_i\\ \beta_i-1\end{matrix}\right\rangle = \Omega_{i0}\begin{pmatrix}\beta_i\\ \beta_i-1\end{pmatrix}\left|\begin{matrix}\beta_i\\ \beta_i\end{matrix}\right\rangle + \cdots \tag{5.39}$$

with C_{2k-1} in the form of Eq. (B4), bearing in mind (B1), (B2),

$$\left\{Q_{2k-1,-i}Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle\right\}\left\{Q_{2k-1,i}\left|\begin{matrix}\beta_i\\ \beta_i-1\end{matrix}\right\rangle\right\}$$
$$+\left\{Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle\right\}\left[\sum_{j=1}^{k-1}Q_{2k-1,-j}Q_{2k-1,j}+\sum_{\alpha=1}^{k-1}J_{2\alpha-1,2\alpha}^2\right.$$
$$\left.+\sum_{\alpha=1}^{k-1}(2k-2\alpha-1)J_{2\alpha-1,2\alpha}\right]\left|\begin{matrix}\beta_i\\ \beta_i-1\end{matrix}\right\rangle$$
$$+(2k-2i+2\beta_i-2)\left\{Q_{2k-1,i}\left|\begin{matrix}1\\ 0\end{matrix}\right\rangle\right\}\left|\begin{matrix}\beta_i\\ \beta_i-1\end{matrix}\right\rangle$$
$$=\Omega_{i0}\begin{pmatrix}\beta_i\\ \beta_i-1\end{pmatrix}C_{2k-1}\left|\begin{matrix}\beta_i\\ \beta_i\end{matrix}\right\rangle. \tag{B5}$$

The second term on the left-hand side cancels the term on the right-hand side, and with Eqs. (B3), (5.20), and (5.24) the derivation of Eq. (5.40) is attained.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Decomposition of the Unitary Irreducible Representations of the Group $SL(2C)$ Restricted to the Subgroup $SU(1, 1)$

A. SCIARRINO
Istituto di Fisica e Scuola di Perfezionamento in Fisica, Università di Roma, Rome, Italy

M. TOLLER
Istituto di Fisica—Università di Roma, Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy

(Received 31 October 1966)

The unitary irreducible representations of the group $SL(2C)$ belonging to the principal series restricted to the subgroup $SU(1, 1)$ are decomposed into a direct integral of unitary irreducible representations of $SU(1, 1)$. The matrix elements of the unitary operator which performs the decomposition are given explicitly and used to obtain a relation between the matrix elements of the unitary irreducible representations of the groups $SL(2C)$ and $SU(1, 1)$. Similar identities between the matrix elements of nonunitary representations of these groups are obtained by means of analytic continuation. The relevance of these results to the theory of complex angular momentum and of high energy nearly forwardscattering is pointed out.

1. INTRODUCTION

RECENTLY various authors[1,2] have suggested and investigated a connection between complex angular momentum and the (not necessarily unitary) irreducible representations of the Poincaré group corresponding to a spacelike four-momentum (the momentum transfer). These representations are strictly connected[3] to the representations of the little group corresponding to a spacelike four-momentum, i.e., the three-dimensional Lorentz group, or the corresponding spinor group, the group $SU(1, 1)$ which is homomorphic to it. The scattering amplitude at fixed momentum transfer can be expanded in terms of the matrix elements of these representations. This expansion is strictly connected with the expansion obtained by means of the Sommerfeld–Watson transform.[4] It can be considered as a generalization of the partial wave analysis, which is an expansion of the scattering amplitude at fixed energy in terms of the matrix elements of the representations of the group $SU2$.

It has also been suggested [5–7] that, when the momentum transfer vanishes, it is more natural to expand the scattering amplitude in terms of the matrix elements of the representations of the little group corresponding to a vanishing four-momentum, i.e., the homogeneous Lorentz group or the corresponding spinor group $SL(2C)$ which is homomorphic to it. In other words, many considerations lead us to think that, at vanishing momentum transfer, the expansion in terms of the matrix elements of the representations of $SL(2C)$ permits a simpler description of the high-energy scattering amplitude.

As the scattering amplitude is an analytic, and therefore continuous, function of the momentum transfer, a connection must exist between the $SL(2C)$ expansion at vanishing momentum transfer and the $SU(1, 1)$ expansion for very small momentum transfer. This connection takes a more suggestive form if we make the assumption (supported by recent research[8]) that the scattering amplitude is dominated in the very-high-energy region by Regge pole contributions. As shown in the above-mentioned papers, each Regge pole contribution can be described in terms of the matrix elements of an irreducible nonunitary representation of $SU(1, 1)$. Gribov and Volkov[9,10] have pointed out that, when the momentum transfer vanishes, the various Regge pole contributions can no longer be independent from each other, but must be arranged in families of poles which are displaced from one another by integral numbers. However, there is a large ambiguity in determining the structure of these families, so further hypotheses are needed. According to the ideas sketched above, we suggest[11] that at vanishing momentum transfer the sum of the Regge pole contributions belonging to a family gives rise to a contribution which can be described in terms of an

[1] H. Joos, *Lectures in Theoretical Physics*, W. E. Brittin and A. O. Barut, Eds. (University of Colorado Press, Boulder, Colo., 1965), Vol. VIIA, p. 132; L. Sertorio and M. Toller, Nuovo Cimento **33**, 413 (1964); F. T. Hadjoannou, Nuovo Cimento **44**, 185 (1966).
[2] J. F. Boyce, J. Math. Phys. **8**, 675 (1967).
[3] E. P. Wigner, Ann. Math. **40**, 149 (1939).
[4] E. J. Squires, *Complex Angular Momentum and Particle Physics* (W. A. Benjamin, Inc., New York, 1963). This book contains the references to the original papers.
[5] M. Toller, Nuovo Cimento **37**, 631 (1965).
[6] M. Toller, "The Laplace Transform on the Lorentz Group and a Generalization of the Regge Pole Hypothesis," Istituto di Fisica dell'Università di Roma, Report No. 76 (1965).
[7] M. Toller, "Some Consequences of a Generalization of the Regge–Pole Hypothesis," Istituto di Fisica dell' Università di Roma, Report No. 84 (1965).
[8] R. J. N. Phillips and W. Rarita, Phys. Rev. **139**, B1336 (1965).
[9] D. V. Volkov and V. N. Gribov, Zh. Eksperim. i Teor. Fiz. **44**, 1068 (1963) [English transl.: Soviet Phys.—JETP **17**, 720 (1963)].
[10] V. N. Gribov, Zh. Eksperim. i Teor. Fiz. **43**, 1529 (1962) [English transl.: Soviet Phys.—JETP **16**, 1080 (1963)].
[11] Volkov and Gribov use a different assumption.

irreducible representation of $SL(2C)$[12]; we call it a "Lorentz pole contribution."

In order to investigate the above-mentioned concepts in more detail, it is necessary to study the connection between the representations of the groups $SL(2C)$ and $SU(1,1)$. This is the aim of the present paper.

In Secs. 2 and 3 we describe the irreducible unitary representations of $SL(2C)$[13] and of the subgroup $SU(1,1)$.[14] In Sec. 4 we perform the decomposition of the irreducible unitary representations of $SL(2C)$ restricted to the subgroup $SU(1,1)$ into a direct integral of unitary irreducible representations of $SU(1,1)$. In Sec. 5 we investigate the properties of the matrix elements of the unitary operator which performs the decomposition. By means of these matrix elements we can write a formula which gives the matrix elements of the representations of $SL(2C)$ in terms of the matrix elements of the representations of $SU(1,1)$. In Sec. 6 we modify and extend this formula, introducing by means of analytic continuation the matrix elements of nonunitary irreducible representations of the groups considered. In Sec. 7 we show that the identities so obtained are the very mathematical instruments needed in order to clarify and exploit the ideas sketched at the beginning of this Introduction.

The arguments treated in this paper are also useful in connection with the problem of decomposing an irreducible representation of the complex inhomogeneous Lorentz group restricted to the real inhomogeneous Lorentz group (Poincarè group). The importance of this problem in the relativistic theory of scattering has been pointed out by Roffman.[15,16]

2. UNITARY IRREDUCIBLE REPRESENTATIONS OF THE GROUP $SL(2C)$

The elements of the group $SL(2C)$ are the unimodular complex 2×2 matrices of the form

$$a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad a_{11}a_{22} - a_{12}a_{21} = 1. \quad (2.1)$$

The unitary (and also the non-unitary) irreducible representations of this group have been investigated and classified by Gel'fand and Naimark.[13] The method used by these authors for the construction of these representations has been generalized by Mackey[17,18] in his theory of induced representations.

In order to introduce suitable notations, we give in this section a short description of the irreducible unitary representations of $SL(2C)$ belonging to the principal series. We do not consider in this paper the representations of the complementary series. As in the following, we use Mackey's general techniques, and emphasize the fact that we are dealing with induced representations.

It is useful to introduce the following notation for the elements of some one-parameter subgroups of $SL(2C)$:

$$u_x(\theta) = \begin{pmatrix} \cos\frac{1}{2}\theta & -i\sin\frac{1}{2}\theta \\ -i\sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix}, \quad a_x(\zeta) = \begin{pmatrix} \cosh\frac{1}{2}\zeta & \sinh\frac{1}{2}\zeta \\ \sinh\frac{1}{2}\zeta & \cosh\frac{1}{2}\zeta \end{pmatrix},$$

$$u_y(\theta) = \begin{pmatrix} \cos\frac{1}{2}\theta & -\sin\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix}, \quad a_y(\zeta) = \begin{pmatrix} \cosh\frac{1}{2}\zeta & -i\sinh\frac{1}{2}\zeta \\ i\sinh\frac{1}{2}\zeta & \cosh\frac{1}{2}\zeta \end{pmatrix},$$

$$u_z(\theta) = \begin{pmatrix} \exp(-i\tfrac{1}{2}\theta) & 0 \\ 0 & \exp(i\tfrac{1}{2}\theta) \end{pmatrix}, \quad a_z(\zeta) = \begin{pmatrix} \exp\frac{1}{2}\zeta & 0 \\ 0 & \exp(-\frac{1}{2}\zeta) \end{pmatrix}. \quad (2.2)$$

In the homomorphism between $SL(2C)$ and the Lorentz group, the first three elements give rise to rotations of an angle θ around the axes of the coordinate system, and the other three elements give rise to pure Lorentz transformations along the coordinate axes with relative velocities $\beta = \tanh\zeta$.

$SU2$ is the unitary subgroup of $SL(2C)$ (homomorphic to the rotation group). As well known, every element $u \in SU2$ can be written in the form

$$u = u_z(\mu)u_y(\theta)u_z(\nu), \quad 0 \le \mu < 4\pi, \quad 0 \le \theta \le \pi, \quad 0 \le \nu < 2\pi. \quad (2.3)$$

By means of this parametrization, the invariant measure on $SU2$ takes the form

$$d\mu = (4\pi)^{-2}\sin\theta\, d\mu\, d\nu\, d\theta. \quad (2.4)$$

The induced representations we consider are constructed by means of the subgroup K whose elements are complex matrices of the form

$$k = \begin{pmatrix} p^{-1} & q \\ 0 & p \end{pmatrix}, \quad p \neq 0, \quad (2.5)$$

[12] A similar classification of the Regge trajectories has been proposed by G. Domokos and P. Suranyi, Nucl. Phys. **54**, 529 (1964). These authors use as a starting point a Bethe–Salpeter equation with the integration path of the energy variable rotated in the complex plane [G. C. Wick, Phys. Rev. **96**, 1124 (1954)]. In consequence, they consider the four-dimensional rotation group instead of the Lorentz group.

[13] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon Press, Inc., London, 1964). This book contains the references to the original papers by Gel'fand and Naimark.

[14] V. Bargmann, Ann. Math. **48**, 568 (1947).

[15] E. H. Roffman, Phys. Rev. Letters **16**, 210 (1966).

[16] E. H. Roffman, Commun. Math. Phys. **4**, 237 (1967).

[17] G. Mackey, "The Theory of Group Representations," Lecture notes, University of Chicago (1955).

[18] G. Mackey, Bull. Am. Math. Soc. **69**, 628 (1963). This paper contains references to previous work on induced representations.

and, by means of the following one-dimensional unitary representations of K,

$$L^{M\lambda}(k) = |p(k)|^{2(\lambda+M)}[(p(k)]^{-2M}, \tag{2.6}$$

where $p(k)$ is an element of the matrix k as shown in Eq. (2.5); M is an integral or half-integral and λ is a pure imaginary parameter.

We mention some properties of the right cosets of $SL(2C)$ with respect to the subgroup K[13]:

(a) Every coset (one excepted) contains one and only one element of the form

$$z = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix}, \tag{2.7}$$

where z represents both a complex number and the corresponding matrix. This means that every matrix $a \in SL(2C)$ (with $a_{22} \neq 0$) can be written in the form $a = kz$ with $k \in K$.

(b) Every coset contains infinite unitary matrices. Two unitary matrices u and u' belong to the same coset if and only if

$$u' = u_z(\mu)u \tag{2.8}$$

for some value of μ. This means that every matrix $a \in SL(2C)$ can be written (in an infinite number of ways) in the form $a = ku$ with $k \in K$ and $u \in SU2$.

A measure in the space of the right cosets can be considered as a measure $d\mu(z)$ in the complex plane of z. We choose this measure in the following way:

$$\int \psi(z)\, d\mu(z) = \int_{SU2} \psi(u)\, d\mu, \tag{2.9}$$

where $\psi(a)$ is an arbitrary function defined over $SL(2C)$ with the property

$$\psi(ka) = \psi(a), \quad k \in K. \tag{2.10}$$

A simple calculation shows that Eq. (2.9) requires that

$$d\mu(z) = d\,\mathrm{Re}\, z\, d\,\mathrm{Im}\, z/\pi(1 + |z|^2)^2. \tag{2.11}$$

An element $a \in SL(2C)$ given by Eq. (2.1) operates a transformation in the space of the right cosets or in the space of the representative elements z. More explicitly, this transformation is

$$z' \to z'' = kz'a, \quad k \in K, \tag{2.12}$$

or, performing the calculations

$$z'' = (z'a_{11} + a_{21})/(z'a_{12} + a_{22}). \tag{2.13}$$

From Eq. (2.11) we know that the measure $d\mu(z)$ is affected in the following way by the transformation (2.12):

$$\frac{d\mu(z'')}{d\mu(z')} = \left(\frac{1 + |z'|^2}{|z'a_{12} + a_{22}|^2 + |z'a_{11} + a_{21}|^2}\right)^2. \tag{2.14}$$

Given an element a of $SL(2C)$, we indicate by $(a)_0$ an arbitrary unitary matrix which belongs to the same coset. Clearly $a(a)_0^{-1}$ belongs to the group K. If a is given by Eq. (2.1), we have

$$|p[a(a)_0^{-1}]|^2 = |a_{21}|^2 + |a_{22}|^2. \tag{2.15}$$

Therefore Eq. (2.14) can be written in the form

$$\frac{d\mu(z'')}{d\mu(z')} = \left|\frac{p[z'(z')_0^{-1}]}{p[z'a(z'a)_0^{-1}]}\right|^4 = \left|\frac{p[a'(a')_0^{-1}]}{p[a'a(a'a)_0^{-1}]}\right|^4. \tag{2.16}$$

In the last expression of this equation, a' is an arbitrary element belonging to the same coset as z' [note that $(a')_0 = (kz')_0 = (z')_0$ and that $p(k)$ is a representation of K].

The Hilbert space $\mathcal{H}$, where the induced representation $\mathfrak{D}^{M\lambda} = UL^{M\lambda}$ operates, is formed by functions $f(a)$ defined on $SL(2C)$ which satisfy the covariance condition

$$f(ka) = L^{M\lambda}(k)f(a), \quad k \in K. \tag{2.17}$$

The scalar product is defined by

$$(f, f') = \int_{SU2} \overline{f(u)} f'(u)\, d\mu \tag{2.18}$$

and the representation operators operate in the following way:

$$[\mathfrak{D}^{M\lambda}(a)f](a') = \left|\frac{p[a'(a')_0^{-1}]}{p[a'a(a'a)_0^{-1}]}\right|^2 f(a'a). \tag{2.19}$$

Equations (2.17)–(2.19) define the induced representation completely.

Equation (2.19) can also be written in terms of functions defined on $SU2$ alone:

$$\begin{aligned}[\mathfrak{D}^{M\lambda}(a)f](u) &= |p[ua(ua)_0^{-1}]|^{-2} f(ua) \\ &= |p[ua(ua)_0^{-1}]|^{-2} L^{M\lambda}[ua(ua)_0^{-1}] f[(ua)_0] \\ &= |p[ua(ua)_0^{-1}]|^{2(\lambda+M-1)} (p[ua(ua)_0^{-1}])^{-2M} f[(ua)_0].\end{aligned} \tag{2.20}$$

The covariance condition in this case takes the simpler form

$$f(u_z(\mu)u) = \exp(-iM\mu) f(u). \tag{2.21}$$

In order to write the representation operators in matrix form, we have to choose a basis in the Hilbert space of the functions $f(u)$ which satisfies Eq. (2.21). A very convenient choice is the following:

$$\Phi^M_{jm}(u) = (2j + 1)^{\frac{1}{2}} R^j_{Mm}(u). \tag{2.22}$$

$R^j_{Mm}(u)$ are the matrix elements of the unitary irreducible representations of $SU2$, which can be written in the form[19]

$$R^j_{mm'}(u_z(\mu)u_y(\theta)u_z(\nu)) = \exp(-im\mu - im'\nu) r^j_{mm'}(\theta), \tag{2.23}$$

where the functions $r^j_{mm'}(\theta)$ are given in Eq. (A6) of Appendix A.

It is known[13] that every element of $SL(2C)$ can be

[19] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).

decomposed in the following way:

$$a = u_1 a_z(\zeta) u_2, \quad u_1, u_2 \in SU2, \quad 0 \leq \zeta < \infty. \tag{2.24}$$

From Eq. (2.20) we have

$$[\mathfrak{D}^{M\lambda}(u)\Phi^M_{jm}](u') = \Phi^M_{jm}(u'u) = \sum_{m'} \Phi^M_{jm'}(u') R^j_{m'm}(u), \tag{2.25}$$

and therefore

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(u) = (\phi^M_{jm}, \mathfrak{D}^{M\lambda}(u)\Phi^M_{j'm'}) = \delta_{jj'} R^j_{mm'}(u). \tag{2.26}$$

We show in Appendix A that

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(a_z(\zeta)) = \delta_{mm'}\, d^{M\lambda}_{mjj'}(\zeta). \tag{2.27}$$

From the last two equations we can obtain the general form of a matrix element of the representation operator

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(u_1 a_z(\zeta) u_2) = \sum_{m''} R^j_{mm''}(u_1)\, d^{M\lambda}_{m''jj'}(\zeta) R^{j'}_{m''m'}(u_2). \tag{2.28}$$

Many properties of the functions $d^{M\lambda}_{mjj'}(\zeta)$ are given in Appendix A.

In particular, we know that representations $\mathfrak{D}^{M\lambda}$ and $\mathfrak{D}^{-M,-\lambda}$ are equivalent and therefore the following identity must hold:

$$\mathfrak{D}^{-M,-\lambda}_{jmj'm'}(a) = \mathfrak{U}^{-M,-\lambda}_j \mathfrak{D}^{M\lambda}_{jmj'm'}(a) \mathfrak{U}^{M\lambda}_{j'}, \tag{2.29}$$

where the functions $\mathfrak{U}^{M\lambda}_j$ are given by Eq. (A18) of Appendix A. In the following we use the decomposition

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(a) = \mathcal{A}^{M\lambda}_{jmj'm'}(a) + \mathfrak{U}^{M\lambda}_j \mathcal{A}^{-M,-\lambda}_{jmj'm'}(a) \mathfrak{U}^{-M,-\lambda}_{j'}, \tag{2.30}$$

where

$$\mathcal{A}^{M\lambda}_{jmj'm'}(u_1 a_z(\zeta) u_2) = \sum_{m''} R^j_{mm''}(u_1) a^{M\lambda}_{m''jj'}(\zeta) R^{j'}_{m''m'}(u_2). \tag{2.31}$$

As shown in Appendix A, the functions $a^{M\lambda}_{mjj'}(\zeta)$ have a simple asymptotic behavior for $\zeta \to \infty$.

3. UNITARY IRREDUCIBLE REPRESENTATIONS OF THE GROUP $SU(1, 1)$

The group $SU(1, 1)$ is a subgroup of $SL(2C)$ formed by all the complex matrices of the form

$$v = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad |\alpha|^2 - |\beta|^2 = 1. \tag{3.1}$$

Every element of this group can be written in a unique way as follows:

$$v = u_z(\mu) a_x(\zeta) u_z(\nu), \quad 0 \leq \mu < 4\pi, \quad 0 \leq \nu < 2\pi, \quad 0 \leq \zeta < \infty. \tag{3.2}$$

With this parametrization, the invariant measure on the group takes the form

$$dv = (4\pi)^{-2} \sinh \zeta \, d\mu \, d\nu \, d\zeta. \tag{3.3}$$

The irreducible representations of $SU(1, 1)$ have been described by Bargmann.[14] The representation matrices have the form

$$D^{\Lambda}_{mm'}(u_z(\mu) a_x(\zeta) u_z(\nu)) = \exp(-im\mu - im'\nu)\, d^{\Lambda}_{mm'}(\zeta). \tag{3.4}$$

The superscript Λ is a shorthand notation for the parameters which label the equivalence classes of irreducible representations. In order to obtain more explicit formulas, it has to be replaced in the following way:

(a) Representations of the continuous classes: Λ stands for the two parameters ϵ, l. The parameter ϵ takes the value 0 for the representations of the integral type and $\frac{1}{2}$ for the representations of the half-integral type. The unitary representations of the principal series correspond to complex values of the parameter l with $\mathrm{Re}\, l = -\frac{1}{2}$. For other values of l we have non-unitary representations, some of which are equivalent to the unitary representations of the complementary series. The subscripts m and m' take the values ϵ, $\epsilon \pm 1$, $\epsilon \pm 2, \cdots$.

(b) Representations of the discrete classes: Λ stands for $k\pm$ with $k = \frac{1}{2}, 1, \frac{3}{2}, \cdots$. The subscripts m and m' take the values $k, k+1, k+2, \cdots$ for the representations D^{k+} and $-k, -(k+1), -(k+2), \cdots$ for the representations D^{k-}.

The Plancherel formula can be written in the form

$$\int_{SU(1,1)} |f(v)|^2 \, dv = \int \sum_{mm'} |F^{\Lambda}_{mm'}|^2 \, d\Lambda, \tag{3.5}$$

where

$$F^{\Lambda}_{mm'} = \int_{SU(1,1)} f(v) D^{\Lambda}_{mm'}(v) \, dv. \tag{3.6}$$

The Plancherel measure $d\Lambda$ which appears in Eq. (3.5) is defined by

$$\int \psi(\Lambda)\, d\Lambda = \sum_{\epsilon} \int_0^{\infty} \psi(\epsilon, is - \tfrac{1}{2}) \eta(\epsilon, is - \tfrac{1}{2})\, ds + \sum_k (2k-1)[\psi(k+) + \psi(k-)], \tag{3.7}$$

where

$$\eta(0, l) = (2l+1)\cot \pi l, \qquad \eta(\tfrac{1}{2}, l) = -(2l+1)\tan \pi l. \tag{3.8}$$

For the functions $d^{\epsilon l}_{mm'}(\zeta) = d^l_{mm'}(\zeta)$ we use a phase convention different from that used by Bargmann.[14] In such a way we obtain functions which are analytic in the whole complex l plane. For $m \geq m'$ they can be expressed in terms of the hypergeometric function in the following way[5]:

$$d^l_{mm'}(\zeta) = \frac{1}{(m-m')!} \frac{\Gamma(l+m+1)}{\Gamma(l+m'+1)} \left(\cosh \frac{\zeta}{2}\right)^{m+m'} \times \left(\sinh \frac{\zeta}{2}\right)^{m-m'} F_{21}\left(m-l, m+l+1; m-m'+1; -\left(\sinh \frac{\zeta}{2}\right)^2\right), \tag{3.9}$$

and they satisfy the symmetry conditions

$$\overline{d^l_{mm'}(\zeta)} = d^l_{-m,-m'}(\zeta), \qquad d^l_{mm'}(\zeta) = d^l_{m'm}(\zeta). \tag{3.10}$$

From Eq. (3.9) we can derive the identity

$$d^{-l-1}_{mm'}(\zeta) = U^{-l-1}_m d^l_{mm'}(\zeta) U^l_{m'}, \tag{3.11}$$

where

$$U^l_m = \frac{\Gamma(l+m+1)}{\Gamma(m-l)}, \quad U^{-l-1}_m U^l_m = 1. \tag{3.12}$$

These equations show that the representations $D^{\epsilon l}$ and $D^{\epsilon,-l-1}$ are equivalent.

We use in the following the decomposition

$$d^l_{mm'}(\zeta) = a^l_{mm'}(\zeta) + U^l_m a^{-l-1}_{mm'}(\zeta) U^{-l-1}_{m'}, \tag{3.13}$$

where

$$a^l_{mm'}(\zeta) = \frac{(-1)^{m-m'}\Gamma(-2l-1)}{\Gamma(-l-m)\Gamma(-l+m)}\left(\sinh\frac{\zeta}{2}\right)^{-2l-2} \times \left(\coth\frac{\zeta}{2}\right)^{m+m'} F_{21}\left(l+m'+1, l+m+1; 2l+2; -\left(\sinh\frac{\zeta}{2}\right)^{-2}\right). \tag{3.14}$$

Functions (3.14) have a simple asymptotic behavior for large values of ζ. They are analytic in the whole complex plane, apart from poles for half-integral values of $l-m$.

The functions $d^{k+}_{mm'}(\zeta)$ have been computed by Bargmann.[14] They can be given in terms of the functions defined above by means of the formulas

$$d^{k+}_{mm'}(\zeta) = \left[\frac{(m-k)!\,(m'+k-1)!}{(m'-k)!\,(m+k-1)!}\right]^{\frac{1}{2}} d^{(k-1)}_{mm'}(\zeta) \tag{3.15}$$

or

$$d^{k+}_{mm'}(\zeta) = 2\left[\frac{(m-k)!\,(m'+k-1)!}{(m'-k)!\,(m+k-1)!}\right]^{\frac{1}{2}} a^{(k-1)}_{mm'}(\zeta). \tag{3.16}$$

The functions $d^{k-}_{mm'}(\zeta)$ are given by

$$d^{k-}_{mm'}(\zeta) = d^{k+}_{-m',-m}(\zeta) = (-1)^{m-m'} d^{k+}_{-m,-m'}(\zeta). \tag{3.17}$$

It is useful to define the matrices also:

$$A^l_{mm'}(u_z(\mu)a_x(\zeta)u_z(\nu)) = \exp(-im\mu - im'\nu)\alpha^l_{mm'}(\zeta). \tag{3.18}$$

It is clear that Eqs. (3.11) and (3.13) can also be written in terms of the matrices $A^l_{mm'}(v)$ and $D^{\epsilon l}_{mm'}(v)$.

4. DECOMPOSITION PROCESS

In this section we decompose the irreducible unitary representations $\mathfrak{D}^{M\lambda}$ of $SL(2C)$ restricted to $SU(1,1)$ into a direct integral containing the irreducible unitary representations D^{Λ} of $SU(1,1)$ described in Sec. 3. Our proceeding can be divided in two stages. In the first stage, following a general method developed by Mackey,[17] we decompose the restrictions of the representations $\mathfrak{D}^{M\lambda}$, which have been described in Sec. 2 as induced representations, in pairs of reducible representations of $SU(1,1)$, which we indicate by O^{-M} and O^{+M}. In the second stage we decompose these representations by means of the Plancherel formula in a direct integral of irreducible unitary representations. At last we find an explicit expression for the matrix elements $\mathfrak{D}^{M\lambda}_{jmj'm'}(v)$ in function of the matrix elements $D^{\Lambda}_{mm'}(v)$.

In order to perform the first step, we must determine the double cosets of $SL(2C)$ with respect to its subgroups K and $SU(1,1)$. We recall that these double cosets are the sets composed by the elements of $SL(2C)$ that may be written in the form

$$kav, \quad k \in K, \quad v \in SU(1,1),$$

where a is a fixed element of $SL(2C)$ which can be chosen as representative element of the double coset. Clearly, the double cosets are pairwise disjoint, and their union is $SL(2C)$.

From the properties given in Sec. 2, we know that the following decomposition is unique:

$$\begin{gathered} a = ku_y(\theta)u_z(\nu), \quad a \in SL(2C), \\ k \in K, \quad 0 \le \theta \le \pi, \quad 0 \le \nu < 2\pi. \end{gathered} \tag{4.1}$$

It is easily obtained that

$$\begin{aligned} &u_y(\theta) = kb^+a_x(\zeta), \quad \tanh\tfrac{1}{2}\zeta = \tan\tfrac{1}{2}\theta, \\ &\qquad\qquad \text{for} \quad 0 \le \theta < \tfrac{1}{2}\pi; \\ &u_y(\theta) = kb^-a_x(\zeta), \quad \tanh\tfrac{1}{2}\zeta = \cot\tfrac{1}{2}\theta, \\ &\qquad\qquad \text{for} \quad \tfrac{1}{2}\pi < \theta \le \pi, \end{aligned} \tag{4.2}$$

where

$$b^+ = e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad b^- \begin{pmatrix} 10 & - \\ 1 & 0 \end{pmatrix}. \tag{4.3}$$

In both cases we have

$$p(k) = (\cosh\zeta)^{-\frac{1}{2}}. \tag{4.4}$$

From Eqs. (4.1) and (4.2) it follows that each element $a \in SL(2C)$ can be written in one of the following three forms:

$$\begin{aligned} a &= kb^+a_x(\zeta)u_z(\nu), \\ a &= kb^-a_x(\zeta)u_z(\nu), \\ a &= ku_y(\tfrac{1}{2}\pi)u_z(\nu). \end{aligned} \tag{4.5}$$

If we note that $a_x(\zeta)u_z(\nu) \in SU(1,1)$, we see that Eqs. (4.5) represent three double cosets. It can be shown that they are distinct. In the following we disregard the third coset which has vanishing invariant measure.

For every function $f(a)$, belonging to the space where the representations described in Sec. 2 operate, we introduce the following two functions defined over $SU(1,1)$:

$$f^{\pm}(v) = f(b^{\pm}v). \tag{4.6}$$

From Eqs. (4.6) and (2.17) it follows that these functions satisfy the covariance relation

$$f^{\pm}(hv) = L^{M\lambda}(b^{\pm}h(b^{\pm})^{-1})f^{\pm}(v), \quad h \in H^{\pm}, \tag{4.7}$$

where $H^{\pm}$ are subgroups of $SU(1,1)$, defined by

$$H^{\pm} = SU(1,1) \cap (b^{\pm})^{-1} K b^{\pm},$$

which coincide with the group H of the elements $u_z(\mu)$.

More explicitly, the relation (4.7) can be written as

$$f^{\pm}(u_z(\mu)v) = \exp(\mp iM\mu) f^{\pm}(v). \tag{4.8}$$

From a straightforward calculation using the change of variables (4.2), we have

$$\begin{aligned}\int_{SU2} |f(u)|^2 \, d\mu &= \frac{1}{4\pi} \int_0^{2\pi} d\nu \int_0^{\pi} |f(u_y(\theta)u_z(\nu))|^2 \sin\theta \, d\theta \\ &= \frac{1}{4\pi} \sum_{\rho} \int_0^{2\pi} d\nu \int_0^{\infty} |f^{\rho}(a_x(\zeta)u_z(\nu))|^2 \frac{\sinh\zeta}{(\cosh\zeta)^2} \, d\zeta \\ &= \sum_{\rho} \int_{SU(1,1)} |F^{\rho}(v)|^2 \, dv, \end{aligned} \tag{4.9}$$

where ρ is an index which assumes the two values $-$ and $+$, and $F^{\pm}(v)$ is defined by

$$F^{\pm}(v) = (\cosh\zeta)^{-1} f^{\pm}(v), \tag{4.10}$$

where v and ζ are connected by Eq. (3.2). The functions $F^{\pm}(\zeta)$ satisfy the covariance relation (4.8).

If we consider two Hilbert spaces $\mathcal{H}^{+M}$ and $\mathcal{H}^{-M}$ composed by the functions $F^{\pm}(v)$ with the norm

$$\|F^{\pm}\|^2 = \int_{SU(1,1)} |F^{\pm}(v)|^2 \, dv, \tag{4.11}$$

we have obtained, by means of Eqs. (4.7), (4.9), and (4.10), an isomorphism between the Hilbert space $\mathcal{H}$, where $\mathfrak{D}^{M\lambda}$ acts, and the space

$$\mathcal{H}^{+M} \oplus \mathcal{H}^{-M}.$$

In this way we have split the representation $\mathfrak{D}^{M\lambda}$ in the direct sum of two representations $O^{\pm M}$ which act in $\mathcal{H}^{\pm M}$. More explicitly, from Eq. (2.19) we have

$$[O^{\pm M}(v)f^{\pm}](v') = \left| \frac{p[b^{\pm}v'(b^{\pm}v')_0^{-1}]}{p[b^{\pm}v'v(b^{\pm}v'v)_0^{-1}]} \right|^2 f^{\pm}(vv'). \tag{4.12}$$

Writing v in the form (3.2), we have from Eq. (2.15)

$$|p(b^{\pm}v(b^{\pm}v)_0^{-1})|^2 = \cosh\zeta, \tag{4.13}$$

and, using Eq. (4.10), we can write Eq. (4.12) in the form

$$[O^{\pm}(v)F^{\pm}](v') = F^{\pm}(v'v). \tag{4.14}$$

From Eqs. (4.8), (4.11), and (4.14), we see that the representations $O^{\pm M}$ are representations induced by the following representations of the subgroup H.

$$u_z(\mu) \to \exp(\mp iM\mu). \tag{4.15}$$

If we compare the representations $O^{\pm M}$ with the regular representation of $SU(1,1)$, we see that the only difference consists in the condition (4.8) which defines a subspace of the space $L^2(SU(1,1))$ where the regular representation operates. This means that the representations $O^{\pm M}$ are contained in the regular representation, which admits a decomposition in a direct integral of irreducible representations by means of the Plancherel formula (3.5), (3.6). The representations we are considering can be decomposed in a similar way. We introduce the function $F_m^{\Lambda\pm}$ defined by

$$\int_{SU(1,1)} D^{\Lambda}_{mm'}(v^{-1}) F^{\pm}(v) \, dv = \delta_{\pm M, m'} F_m^{\Lambda\pm}. \tag{4.16}$$

We introduce the subsets $\Omega_{\pm M}$ of the set of the equivalence classes of irreducible unitary representations of $SU(1,1)$. An element Λ of this set belongs to $\Omega_{\pm M}$ if the representation $D^{\Lambda}_{mm'}$ contains the representation (4.15) of the subgroup H—that is, if $\pm M$ is one of the values that can be assumed by the matrix indices m and m' (see Sec. 3). Note that if $\Lambda \notin \Omega_{\pm M}$ the integral (4.16) vanishes.

The Plancherel formula (3.5) takes the form

$$\int_{SU(1,1)} |F^{\pm}(v)|^2 \, dv = \int_{\Omega_{\pm M}} \sum_m |F_m^{\Lambda\pm}|^2 \, d\Lambda. \tag{4.17}$$

The integral can be extended to the subset $\Omega_{\pm M}$ because, outside, the integrand vanishes. If we introduce the Hilbert spaces $\mathcal{H}^{\Lambda\pm}$ formed by the functions $F_m^{\Lambda\pm}$ with the norm

$$\|F^{\Lambda\pm}\|^2 = \sum_m |F_m^{\Lambda\pm}|^2, \tag{4.18}$$

from Eq. (4.17) we see that Eq. (4.16) defined the isomorphism between the spaces $\mathcal{H}^{\pm M}$ and the spaces

$$\int_{\Omega_{\pm M}}^{\oplus} \mathcal{H}^{\Lambda\pm} \, d\Lambda, \tag{4.19}$$

respectively.

We determine how the functions $F_m^{\Lambda\pm}$ are affected by the transformation (4.14) of the function which generates them.

$$\begin{aligned} F_m^{\Lambda\pm} &\to \int_{SU(1,1)} D^{\Lambda}_{m,\pm M}(v'^{-1}) F^{\pm}(v'v) \, dv' \\ &= \int_{SU(1,1)} D^{\Lambda}_{m,\pm M}(vv'^{-1}) F^{\pm}(v') \, dv' = \sum_{m'} D^{\Lambda}_{mm'}(v) F_{m'}^{\Lambda\pm}. \end{aligned} \tag{4.20}$$

This means that the representations which operate on the spaces $\mathcal{H}^{\Lambda\pm}$ are exactly the representations D^{Λ}. Therefore the result of the decomposition process is

$$\begin{aligned} [\mathfrak{D}^{M\lambda}]^{SU(1,1)} &\simeq O^{+M} \oplus O^{-M} \\ &\simeq \int_{\Omega_{+M}}^{\oplus} D^{\Lambda} \, d\Lambda \oplus \int_{\Omega_{-M}}^{\oplus} D^{\Lambda} \, d\Lambda. \end{aligned} \tag{4.21}$$

Equations (4.6), (4.10), and (4.16) give a linear relation between the function $f(a)$ and the functions $F_m^{\Lambda\pm}$. If, in particular, we take

$$f(u) = \Phi^M_{jm}(u) \tag{4.22}$$

[see Eq. (2.22)], we have

$$\begin{aligned} &F^{\pm}(u_z(\mu)a_x(\zeta)u_z(\nu)) \\ &= (\cosh\zeta)^{\lambda-1} \Phi^M_{jm}(u_z(\pm\mu)u_y(\theta^{\pm})u_z(\nu)) \\ &= (2j+1)^{\frac{1}{2}} (\cosh\zeta)^{\lambda-1} \exp(\mp iM\mu - im\nu) r^j_{Mm}(\theta^{\pm}), \\ &\qquad 0 \le \theta^+ < \tfrac{1}{2}\pi, \quad \tfrac{1}{2}\pi < \theta^- \le \pi, \end{aligned} \tag{4.23}$$

where the relation between $\theta^{\pm}$ and ζ is given by Eq. (4.2), and

$$F^{\Lambda\pm}_{m'} = (2j+1)^{\frac{1}{2}} \int_0^{4\pi} d\mu \int_0^{2\pi} dv \int_0^{\infty} (\cosh\zeta)^{\lambda-1} r^j_{Mm}(\theta^{\pm}) \times d^{\Lambda}_{m',\pm M}(\zeta)(-1)^{m'\mp M} \exp[i(m'-m)v](4\pi)^{-2} \times \sinh\zeta\, d\zeta = \delta_{mm'} K^{M\lambda}_m(\Lambda, \pm; j), \quad (4.24)$$

where

$$K^{M\lambda}_m(\Lambda, \pm; j) = (2j+1)^{\frac{1}{2}}(-1)^{m\mp M} \int_0^{\infty} (\cosh\zeta)^{\lambda-1} \times r^j_{Mm}(\theta^{\pm})\, d^{\Lambda}_{m,\pm M}(\zeta)^{\frac{1}{2}} \sinh\zeta\, d\zeta. \quad (4.25)$$

Note that this function is defined only for $\Lambda \in \Omega_{\pm M} \cap \Omega_m$. In general, from Eqs. (4.9) and (4.17), we have

$$\int_{SU2} \overline{f_{(1)}(u)} f_{(2)}(u)\, du = \sum_{\rho} \int_{\Omega_{\rho M}} \sum_{m} \overline{F^{\Lambda\rho}_{(1)m}}\, F^{\Lambda\rho}_{(2)m}\, d\Lambda, \quad (4.26)$$

and if we take

$$f_{(1)}(u) = \Phi^M_{jm}(u), \quad f_2(u) = [\mathfrak{D}^{M\lambda}(v)\Phi^M_{j'm'}](u), \quad v \in SU(1,1),$$

we obtain from (4.24) and (4.26)

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) = \sum_{\rho} \int_{\Omega} \overline{K^{M\lambda}_m(\Lambda, \rho; j)}\, D^{\Lambda}_{mm'}(v) K^{M\lambda}_{m'}(\Lambda, \rho; j')\, d\Lambda, \quad (4.27)$$

where

$$\Omega = \Omega_{\rho M} \cap \Omega_m \cap \Omega_{m'}.$$

This is the basic formula which connects the representations of $SL(2C)$ and $SU(1,1)$.

If we consider the special case $v = e$, we obtain the othonormality condition

$$\sum_{\rho} \int_{\Omega_{\rho M} \cap \Omega m} \overline{K^{M\lambda}_m(\Lambda, \rho; j)}\, K^{M\lambda}_m(\Lambda, \rho; j')\, d\Lambda = \delta_{jj'}. \quad (4.28)$$

As the functions $\Phi^M_{jm}(u)$ form a complete set in the space of the functions $f(u)$ which satisfy Eq. (2.21) so also the functions $K^{M\lambda}_m(\Lambda, \rho; j)$ form a complete set in the space of the functions $F^{\Lambda\rho}_m$ defined for $\Lambda \in \Omega_{\rho M} \cap \Omega_m$. Therefore we can derive in the standard way the completeness condition which can be written formally as

$$\sum_j K^{M\lambda}_m(\Lambda, \rho; j) \overline{K^{M\lambda}_m(\Lambda', \rho'; j)} = \delta_{\rho\rho'}\delta(\Lambda, \Lambda'), \quad (4.29)$$

where the improper function $\delta(\Lambda, \Lambda')$ is defined by

$$\int \psi(\Lambda)\delta(\Lambda, \Lambda')\, d\Lambda = \psi(\Lambda'). \quad (4.30)$$

Equations (4.28) and (4.29) can be considered as unitarity conditions if we consider the functions $K^{M\lambda}_m(\Lambda, \rho; j)$ as matrix elements of a unitary operator.

By means of Eq. (4.29) we can invert Eq. (4.27) and we obtain

$$\sum_{jj'} K^{M\lambda}_m(\Lambda, \rho; j) \mathfrak{D}^{M\lambda}_{jmj'm'}(v) \overline{K^{M\lambda}_{m'}(\Lambda', \rho'; j')} = \sum_{\rho''} \int \delta_{\rho\rho''}\delta(\Lambda, \Lambda'') D^{\Lambda''}_{mm'}(v) \delta_{\rho''\rho'}\delta(\Lambda'', \Lambda')\, d\Lambda'' = \delta_{\rho\rho'}\delta(\Lambda, \Lambda') D^{\Lambda}_{mm'}(v). \quad (4.31)$$

5. PROPERTIES OF THE FUNCTIONS $K^{M\lambda}_m(\Lambda, \rho; j)$

In this section we find some properties of the functions defined by Eq. (4.25).

If we take into account the identity[19]

$$r^j_{mm'}(\theta) = (-1)^{m-m'} r^j_{-m,-m'}(\theta) \quad (5.1)$$

and Eqs. (3.10) and (3.17), we obtain easily the following symmetry properties:

$$K^{-M,\lambda}_{-m}(\epsilon, l, \rho; j) = (-1)^{M-m} K^{M\lambda}_m(\epsilon, l, \rho; j), \quad (5.2)$$

$$K^{-M,\lambda}_{-m}(k\pm, \rho; j) = \rho^{2M} K^{M\lambda}_m(k\mp, \rho; j). \quad (5.3)$$

From the reality of the functions $r^j_{mm'}(\theta)$ and from the last of Eqs. (3.10) we obtain

$$\overline{K^{M\lambda}_m(\epsilon, l, \rho; j)} = K^{M\lambda}_m(\epsilon, \bar{l}, \rho; j), \quad \overline{K^{M\lambda}_m(k\pm, \rho; j)} = K^{M\bar{\lambda}}_m(k\pm, \rho; j), \quad (5.4)$$

and from Eq. (3.11) we have

$$K^{M\lambda}_m(\epsilon, -l-1, \rho; j) = U^{-l-1}_m U^l_{\rho M} K^{M\lambda}_m(\epsilon, l, \rho; j). \quad (5.5)$$

From Eqs. (4.2) we know that the variables $\theta^{\pm}$ which appear in Eq. (4.25) are connected by the relation

$$\theta^- = \pi - \theta^+. \quad (5.6)$$

If we take into account the identity[19]

$$r^j_{mm'}(\pi - \theta) = (-1)^{j-m'} r^j_{-m,m'}(\theta), \quad (5.7)$$

we obtain at once the relation

$$K^{M\lambda}_m(\Lambda, -; j) = (-1)^{j-m} K^{-M,\lambda}_m(\Lambda, +; j). \quad (5.8)$$

On account of this equation, in the following we may consider only the functions $K^{M\lambda}_m(\Lambda, +; j)$.

It is useful to introduce the following functions:

$$E^{M\lambda}_m(l, j) = (2j+1)^{\frac{1}{2}}(-1)^{m-M} \int_0^{\infty} (\cosh\zeta)^{\lambda-1} \times r^j_{Mm}(\theta^+) a^l_{mM}(\zeta)^{\frac{1}{2}} \sinh\zeta\, d\zeta. \quad (5.9)$$

This integral converges for $\mathrm{Re}\, l > \mathrm{Re}\, \lambda - 1$, whereas the integral which defines $K^{M\lambda}_m(\epsilon, l, +; j)$ converges only for $\mathrm{Re}\, \lambda - 1 < \mathrm{Re}\, l < -\mathrm{Re}\, \lambda$. If we substitute in Eq. (4.25) the Eqs. (3.13) and (3.16) we obtain

$$K^{M\lambda}_m(\epsilon, l, +; j) = E^{M\lambda}_m(l, j) + U^l{}_m U^{-l-1}_M E^{M\lambda}_m(-l-1, j), \quad (5.10)$$

$$K^{M\lambda}_m(k+, +; j) \quad 2\left[\frac{(m-k)!\,(M+k-1)!}{(M-k)!\,(m+k-1)!}\right]^{\frac{1}{2}} E^{M\lambda}_m(k-1, j). \quad (5.11)$$

Therefore all the functions we are considering can be expressed in terms of the functions (5.9) by means of Eqs. (5.10), (5.11), (5.8), and (5.3).

In Appendix B we show that the integral (4.25) can be evaluated in terms of a finite sum of Meijer G functions[20] and the result is the following, for $m \geq M$[21]:

$$K_m^{M\lambda}(\epsilon, l, +; j) = \Delta_{Mm}^{j}[\Gamma(l+M+1)\Gamma(m-l)\Gamma(j-\lambda+1)]^{-1} \times \sum_{\alpha\beta}{}' 2^{\beta-j-1}(-1)^{\alpha}[(j-M-\alpha)!\,(\alpha+M-m)! \times (m+j-\alpha-\beta)!\,\alpha!\,\beta!]^{-1} G_{33}^{23}\left(\frac{1}{2}\middle| \begin{matrix} \beta-i, & -l, & l+1 \\ \beta-\lambda, & m, & M \end{matrix}\right), \tag{5.12}$$

where

$$\Delta_{Mm}^{j} = [(2j+1)(j+M)!\,(j-M)! \times (j+m)!\,(j-m)!]^{\frac{1}{2}}. \tag{5.13}$$

For $M > m$ we have to use Eq. (5.2).

Equation (5.12) permits the analytic continuation of the function $K_m^{M\lambda}(\epsilon, l, +; j)$ to complex values of l and λ for which the integral which defines this function does not converge.

We show in Appendix B that this function is meromorphic in the whole complex planes of λ and l; it has poles only for

$$\begin{aligned} l &= \lambda - n - 1, \\ l &= -\lambda + n, \end{aligned} \tag{5.14}$$

where n is a nonnegative integer.

Also the functions $E_m^{M\lambda}(l, j)$ defined by Eq. (5.9) can be written explicitly by means of a finite sum of Meijer G functions and continued analytically in λ and l. The result is[21]

$$\begin{aligned} E_m^{M\lambda}(l, j) = \Delta_{Mm}^{j}&[-2\cos\pi(l-M)]^{-1} \times [\Gamma(l+M+1)\Gamma(m-l)\Gamma(j-\lambda+1)]^{-1} \\ &\times \sum_{\alpha\beta}{}' 2^{\beta-j-1}(-1)^{\alpha} \times [(j-M-\alpha)!\,(\alpha+M-m)! \\ &\times (m+j-\alpha-\beta)!\,\alpha!\,\beta!]^{-1} \\ &\times G_{33}^{32}\left(\frac{1}{2}\middle| \begin{matrix} \beta-j, & -l, & l+1 \\ \beta-\lambda, & m, & M \end{matrix}\right). \end{aligned} \tag{5.15}$$

These functions are meromorphic for all the values of λ and l and may have poles for

$$l = \lambda - n - 1 \tag{5.16}$$

and for half-integral values of $l - M$. The poles of this last kind, which arise from the factor $[\cos\pi(l-M)]^{-1}$, cancel each other in the right-hand side of Eq. (5.10).

If we use the identity

$$G_{33}^{32}\left(X\middle| \begin{matrix} a_1, & a+n, & a \\ b_1, & b_2, & b_3 \end{matrix}\right) = (-1)^n G_{33}^{32}\left(X\middle| \begin{matrix} a_1, & a, & a+n \\ b_1, & b_2, & b_3 \end{matrix}\right), \tag{5.17}$$

which holds for integral n and follows immediately from the definition of G function [see Ref. 20, Eq. (5.3.1)], we obtain from Eq. (5.15) for integral values of $2k$

$$E_m^{M\lambda}(k-1, j) = U_m^{k-1}U_M^{-k}E_m^{M\lambda}(-k, j), \tag{5.18}$$

and from Eqs. (5.10), (5.11), (5.8), (5.2), and (5.3) we have

$$K_m^{M\lambda}(k+, \rho; j) = \left[\frac{(m-k)!\,(\rho M+k-1)!}{(\rho M-k)!\,(m+k-1)!}\right]^{\frac{1}{2}} K_m^{M\lambda}(\epsilon, k-1, \rho; j), \tag{5.19}$$

$$K_m^{M\lambda}(k-, \rho; j) = (-1)^{m-\rho M}\left[\frac{(-m-k)!\,(-\rho M+k-1)!}{(-\rho M-k)!\,(-m+k-1)!}\right]^{\frac{1}{2}} \times K_m^{M\lambda}(\epsilon, k-1, \rho; j).$$

Note that the right-hand sides of these equations have not been obtained directly from the integral (4.25), but are the analytic continuation of the function $K_m^{M\lambda}(\epsilon, l, \rho; j)$ defined by the integral (4.25) for $\operatorname{Re} l = -\frac{1}{2}$.

We use the following notation for the residues of the poles of the function $K_m^{M\lambda}(\epsilon, l, +; j)$:

$$\lim_{l\to\lambda-n-1}(l-\lambda+n+1)K_m^{M\lambda}(\epsilon, l, +; j) = W_{jm}^{M\lambda n}. \tag{5.20}$$

From Eq. (5.5) we also have

$$\lim_{l\to-\lambda+n}(l+\lambda-n)K_m^{M\lambda}(\epsilon, l, +; j) = -U_m^{-\lambda+n}U_M^{\lambda-n-1}W_{jm}^{M\lambda n}. \tag{5.21}$$

The explicit form of the residues $W_{jm}^{M\lambda n}$ has been calculated in Appendix B and for $m \geq M$ is

$$\begin{aligned} W_{jm}^{M\lambda n} = \Delta_{Mm}^{j}&\frac{2^{\lambda-j-1}\Gamma(n-2\lambda+1)(-1)^{M-m}}{\Gamma(n-m-\lambda+1)\Gamma(n+m-\lambda+1)} \\ &\times \sum_{\alpha\beta\gamma}{}' (-1)^{\alpha+\gamma}2^{-\gamma}[(j-M-\alpha)! \\ &\times (\alpha+M-m)!\,(m+j-\alpha-\beta)! \\ &\times (n-\beta-\gamma)!\,\alpha!\,\beta!\,\gamma!]^{-1}[m+\lambda-n]_{n-\beta-\gamma} \\ &\times [M+\lambda-n]_{n-\beta-\gamma}[j-\lambda+1]_{\gamma}[n-2\lambda+1]_{\beta+\gamma}, \end{aligned} \tag{5.22}$$

where

$$[a]_n = a(a+1)\cdots(a+n-1).$$

From Eq. (5.2) we obtain the symmetry property

$$W_{j,-m}^{-M,\lambda,n} = (-1)^{M-m}W_{jm}^{M\lambda n}, \tag{5.23}$$

[20] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. I.

[21] We denote by Σ_{α}' the sum over all the values of α for which the arguments in the factorials which appear in the denominator are nonnegative integers.

which can be used to calculate these functions for $M > m$.

A remarkable and perhaps unexpected property of the residues is the following:

$$(-1)^{j-m}W^{-M,\lambda,n}_{jm} = (-1)^n W^{M\lambda n}_{jm}. \quad (5.24)$$

This formula has been proved by direct algebraic calculation for $n = 0, 1$ and has been tested numerically by means of an electronic computer for a very large number of choices of the parameters. We have not found a general proof for it.

In order to prove the convergence of certain integrals it is necessary to know the behavior of the function $K^{M\lambda}_m(\epsilon, l_0 + il', \rho; j)$ when $l' \to \pm\infty$ for fixed values of the other parameters. In Appendix B we show that this function decreases faster than any negative power of $|l'|$.

6. RELATIONS BETWEEN THE MATRIX ELEMENTS OF NONUNITARY REPRESENTATIONS

If we use Eqs. (5.4) and (3.7) and note that for unitary representations it is $\bar{\lambda} = -\lambda$ and $\bar{l} = -l - 1$, we can write the basic formula (4.27) in the form

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) = -i \sum_\rho \int_{-\frac{1}{2}}^{-\frac{1}{2}+i\infty} K^{M,-\lambda}_m(\epsilon, -l-1, \rho; j) \times D^{\epsilon l}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, l, \rho; j')\eta(\epsilon, l)\, dl + \mathfrak{S}(v), \quad (6.1)$$

where $\epsilon - M$ is an integer (this convention holds also for the following formulas) and $\mathfrak{S}(v)$ is the contribution of the representations of the discrete classes, given by

$$\mathfrak{S}(v) = \sum_\rho \sum_{k\pm} (2k-1) K^{M,-\lambda}_m(k\pm, \rho; j) \times D^{k\pm}_{mm'}(v) K^{M\lambda}_{m'}(k\pm, \rho; j'), \quad (6.2)$$

where the sum over $k\pm$ has to be extended to all the values of this index for which the quantities which appear in the equation have a sense (see Secs. 2 and 4).

From Eqs. (3.8), (3.11), and (5.5) we see that the integrand in Eq. (6.1) is a function symmetric for $l \to -l - 1$. It follows that Eq. (6.1) can also be written in the form

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) = \frac{1}{2i} \sum_\rho \int_{C-\frac{1}{2}} K^{M,-\lambda}_m(\epsilon, -l-1, \rho; j) \times D^{\epsilon l}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, l, \rho; j')\eta(\epsilon, l)\, dl + \mathfrak{S}(v). \quad (6.3)$$

Here and in the following we indicate by C_L a straight path from $L - i\infty$ to $L + i\infty$.

Equation (6.3) can be continued analytically in the complex plane of the parameter λ. The singularities of the integrand are poles which may appear only at the points

$$\begin{aligned} &l = -\lambda + n, \quad l = \lambda - n - 1, \\ &l = \lambda + n, \quad l = -\lambda - n - 1, \\ &l = \epsilon + n, \quad l = -\epsilon - n - 1, \end{aligned} \quad (6.4)$$

where n indicates an arbitrary nonnegative integer. Note that some of the poles move when λ varies, and it is possible that some of them cross the integration path. In order to avoid this, the integration path has to be deformed as shown in Fig. 1, where the new path has been indicated by C.

In the following we do not consider integral values of 2λ. Then the poles of the integrand given by Eq. (6.4) can never coincide and the deformation of the integration path is feasible. Moreover, with this restriction the functions appearing in Eq. (6.2) can be continued analytically in the complex λ plane by means of Eq. (5.19).

If Re λ becomes larger than $\frac{1}{2}$, the poles at $l = \lambda - 1 - n$ and $l = -\lambda + n$ with $0 \le n \le N(\lambda - \frac{1}{2})$ cross the path $C_{-\frac{1}{2}}$. Here and in the following we indicate by $N(a)$ the largest integer smaller than Re a. If we want to obtain the analytic continuation of the integral in Eq. (6.3), we have to use the deformed integration path C shown in Fig. 1. Alternatively we can use the integration path $C_{-\frac{1}{2}}$ and take into account the contributions of the poles.

Owing to the symmetry property of the integrand, the contributions of the poles at $l = \lambda - 1 - n$ and at $l = -\lambda + n$ are the same. The residues can be obtained from Eq. (5.20) and we have, taking Eq. (5.8) into account

$$\begin{aligned} \lim_{l \to \lambda - n - 1} (l - \lambda + n + 1) \sum_\rho K^{M,-\lambda}_m(\epsilon, -l-1, \rho; j) &\times K^{M\lambda}_{m'}(\epsilon, l, \rho; j') \\ = K^{M,-\lambda}_m(\epsilon, -\lambda + n, +; j) W^{M\lambda n}_{j'm'} &+ (-1)^{j-m+j'-m'} \\ \times K^{-M,-\lambda}_m(\epsilon, -\lambda + n, +; j) W^{-M,\lambda,n}_{j'm'} \\ = V^{M\lambda n}_{jm} W^{M\lambda n}_{j'm'}, \end{aligned} \quad (6.5)$$

where

$$V^{M\lambda n}_{jm} = K^{M,-\lambda}_m(\epsilon, -\lambda + n, +; j) + (-1)^{j-m+n} \times K^{-M,-\lambda}_m(\epsilon, -\lambda + n, +; j). \quad (6.6)$$

The factorization (6.5) of the residues has been obtained by means of the identity (5.24).

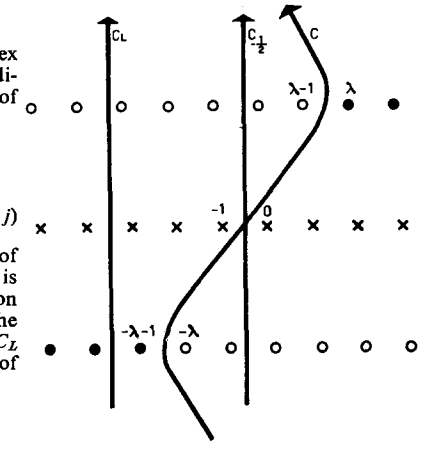


FIG. 1. The complex l plane. We have indicated by ○ the poles of $K^{M\lambda}_{m'}(\epsilon, l, \rho; j')$, by ● the poles of $K^{M,-\lambda}_m(\epsilon, -l-1, \rho; j)$ and by × the poles of $\eta(\epsilon, l)$ ($\epsilon = 0$). $C_{-\frac{1}{2}}$ is the original integration path Eq. (6.3), C is the deformed path and C_L is the shifted path of Eq. (6.14).

By means of this procedure, the analytic continuation of Eq. (6.3) can be written as

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) = \frac{1}{2i}\sum_{\rho}\int_{C-\frac{1}{2}} K^{M,-\lambda}_{m}(\epsilon, -l-1, \rho; j) \times D^{\epsilon l}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, l, \rho; j')\eta(\epsilon, l)\, dl + 2\pi \sum_{n=0}^{N(\lambda-\frac{1}{2})} \eta(\epsilon, \lambda-n-1) V^{M\lambda n}_{jm} W^{M\lambda n}_{j'm'} \times D^{\epsilon,\lambda-n-1}_{mm'}(v) + \mathfrak{G}(v). \quad (6.7)$$

If $\operatorname{Re}\lambda < -\frac{1}{2}$, the poles at $l = \lambda + n$ and $l = -\lambda - n - 1$ with $n = 0, 1, \cdots, N(-\lambda - \frac{1}{2})$ cross the path $C_{-\frac{1}{2}}$. From Eqs. (5.5) and (6.5) we have, after some calculations,

$$\lim_{l\to -\lambda-n-1} (l+\lambda+n+1)\sum_{\rho} K^{M,-\lambda}_{m}(\epsilon, -l-1, \rho; j) \times K^{M\lambda}_{m'}(\epsilon, l, \rho; j') = U^{\lambda+n}_{m} U^{-\lambda-n-1}_{m'} W^{M,-\lambda,n}_{jm} V^{M,-\lambda,n}_{j'm'} \quad (6.8)$$

and the contribution of the poles takes the form

$$2\pi \sum_{n=0}^{N(-\lambda-\frac{1}{2})} \eta(\epsilon, \lambda+n) W^{M,-\lambda,n}_{jm} V^{M,-\lambda,n}_{j'm'} D^{\epsilon,\lambda+n}_{mm'}(v), \quad (6.9)$$

where we have used Eq. (3.11).

From Eqs. (5.2) and (6.6) we have

$$V^{-M,\lambda,n}_{j,-m} = (-1)^{M-m} V^{M\lambda n}_{jm}, \quad (6.10)$$

and from Eq. (6.6) we obtain directly

$$V^{-M,\lambda,n}_{jm} = (-1)^{j-m+n} V^{M\lambda n}_{jm}. \quad (6.11)$$

Note the analogy between Eqs. (6.10), (6.11), and Eqs. (5.23), (5.24).

The sum of the pole contributions in Eqs. (6.7) and (6.9) can be considered as an asymptotic approximation of the function $\mathfrak{D}^{M\lambda}_{jmj'm'}(v)$ when $\zeta \to \infty$ [here and in the following we assume that the elements v of $SU(1, 1)$ are parametrized by means of Eq. (3.2)]. In fact, both the integral along $C_{-\frac{1}{2}}$ and the contribution $\mathfrak{G}(v)$ of the representations of the discrete classes decrease for $\zeta \to \infty$ as $\exp(-\frac{1}{2}\zeta)$ or faster.

In order to obtain a complete asymptotic expansion, we have to modify Eq. (6.3) taking into account the identity

$$K^{M,-\lambda}_{m}(\epsilon, -l-1, \rho; j) D^{\epsilon l}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, l, \rho; j') = K^{M,-\lambda}_{m}(\epsilon, -l-1, \rho; j) A^{l}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, l, \rho; j') + K^{M,-\lambda}_{m}(\epsilon, l, \rho; j) A^{-l-1}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, -l-1, \rho; j'), \quad (6.12)$$

which easily can be obtained from Eqs. (3.13), (3.18), and (5.5). The two terms in the right-hand side of Eq. (6.12) give the same contribution to the integral of Eq. (6.3), and therefore we have

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) = -i\sum_{\rho}\int_{C} K^{M,-\lambda}_{m}(\epsilon, l, \rho; j) A^{-l-1}_{mm'}(v) \times K^{M\lambda}_{m'}(\epsilon, -l-1, \rho; j')\eta(\epsilon, l)\, dl + \mathfrak{G}(v) = -i\sum_{\rho}\int_{C} K^{M,-\lambda}_{m}(\epsilon, -l-1, \rho; j) \times U^{l}_{m} A^{-l-1}_{mm'}(v) U^{-l-1}_{m'} K^{M\lambda}_{m'}(\epsilon, l, \rho; j') \times \eta(\epsilon, l)\, dl + \mathfrak{G}(v). \quad (6.13)$$

Now we shift the integration path C on the left and take into account the contribution of the poles crossed by the path. As it is (see Appendix C)

$$\int_{C_L} K^{M,-\lambda}_{m}(\epsilon, l, \rho; j) A^{-l-1}_{mm'}(v) K^{M\lambda}_{m'}(\epsilon, -l-1, \rho; j') \times \eta(\epsilon, l)\, dl = O(\exp L\zeta). \quad (6.14)$$

If we let $L \to -\infty$, the contribution of the poles gives an asymptotic series for the left-hand side of Eq. (6.13). We have to consider the poles at $l = \lambda - n - 1$, at $l = -\lambda - n - 1$, and at $l = -\epsilon - n - 1$. We call $\mathfrak{G}'(v)$ the contribution of the poles of the last kind. The residues of the poles of the first two kinds are given by Eqs. (6.5) and (6.8), and we have

$$\mathfrak{D}^{M\lambda}_{jmj'm'}(v) \sim 2\pi\sum_{n=0}^{\infty} \eta(\epsilon, \lambda-n-1) V^{M\lambda n}_{jm} \times W^{M\lambda n}_{j'm'} U^{\lambda-n-1}_{m} A^{-\lambda+n}_{mm'}(v) U^{-\lambda+n}_{m'} + 2\pi\sum_{n=0}^{\infty} \eta(\epsilon, \lambda+n) W^{M,-\lambda,n}_{jm} V^{M,-\lambda,n}_{j'm'} \times A^{\lambda+n}_{mm'}(v) + \mathfrak{G}(v) + \mathfrak{G}'(v). \quad (6.15)$$

The left-hand side of this equation can be decomposed by means of Eq. (2.30). From Eqs. (3.14) and (A12), we can see, after some calculations, that the functions $\mathcal{A}^{M\lambda}_{jmj'm'}(v)$, $A^{\lambda+n}_{mm'}(v)$, $\mathcal{A}^{-M,-\lambda}_{jmj'm'}(v)$, $A^{-\lambda+n}_{mm'}(v)$ can be expanded for $v = a_x(\zeta)$ in asymptotic series of the type

$$\sum_{i=1}^{\infty} \beta_i \exp(\alpha_i \zeta), \quad (6.16)$$

where for the first two functions the coefficients α_i are of the type $-\lambda + \nu$ (ν is an integer), and for the last two functions the coefficients α_i are of the type $\lambda + \nu$. In a similar way, using Eq. (3.16), we can expand the functions $\mathfrak{G}(v)$ and $\mathfrak{G}'(v)$ in an asymptotic series of the type (6.16) where the coefficients α_i are of the type $M + \nu$.

As we have assumed that 2λ is not integral, the coefficients α_i of the three forms described above can never coincide. In an expansion of the type (6.16), the coefficients are unequivocally determined, and therefore we can split Eq. (6.15) into three different equations, collecting the terms which have asymptotic

expansions with coefficients α_i of the same type. In such a way we obtain

$$\mathfrak{S}(v) + \mathfrak{S}'(v) = 0, \tag{6.17}$$

$$\mathcal{A}^{M\lambda}_{jmj'm'}(v) \sim 2\pi \sum_{n=0}^{\infty} \eta(\epsilon, \lambda + n) \times W^{M,-\lambda,n}_{jm} V^{M,-\lambda,n}_{j'm'} A^{\lambda+n}_{mm'}(v), \tag{6.18}$$

$$\mathfrak{U}^{M\lambda}_{j} \mathcal{A}^{-M,-\lambda}_{jmj'm'}(v) \mathfrak{U}^{-M,-\lambda}_{j'} \sim 2\pi \sum_{n=0}^{\infty} \eta(\epsilon, \lambda - n - 1) V^{M\lambda n}_{jm} W^{M\lambda n}_{j'm'} \times U^{\lambda-n-1}_{m} A^{-\lambda+n}_{mm'}(v) U^{-\lambda+n}_{m'}. \tag{6.19}$$

Of course, Eq. (6.17) could be proved directly by means of Eqs. (3.16) and (5.19). We have not tried to prove the convergence of the series appearing in Eqs. (6.18) and (6.19). For the physical applications we have in mind it is only necessary to know that these equations represent asymptotic expansions, as we have proved above.

7. DECOMPOSITION OF A "LORENTZ POLE CONTRIBUTION"

It has been shown in Ref. 5 that the scattering amplitude at fixed momentum transfer can be considered as a function defined over a subgroup of $SL(2C)$. This subgroup is homomorphic to the group of the Lorentz transformations which do not change the four-momentum transfer. If the momentum transfer is a spacelike four-vector, this group is $SU(1,1)$; and if the four-momentum transfer vanishes, this group is $SL(2C)$ itself.

It has also been shown[5] that the contribution of a Regge pole with factorizable residue can be written in the form

$$f(v) \sim \sum_{mm'} \rho_m \rho'_{m'} A^{-l-1}_{mm'}(v), \quad v \in SU(1,1). \tag{7.1}$$

Following the ideas suggested in the Introduction, when the four-momentum transfer vanishes, we define a "Lorentz pole contribution" to the scattering amplitude [defined over $SL(2C)$] in the following way[6,7]:

$$f(a) \sim \sum_{jmj'm'} \rho_{jm} \rho'_{j'm'} \mathcal{A}^{-M,-\lambda}_{jmj'm'}(a), \quad a \in SL(2C). \tag{7.2}$$

A contribution of this kind can be expanded into Regge pole contributions of the form (7.1) by means of Eq. (6.18) in the following way:

$$f(v) \sim \sum_{n=0}^{\infty} \sum_{mm'} \rho^{(n)}_m \rho'^{(n)}_{m'} A^{-\lambda+n}_{mm'}(v), \tag{7.3}$$

where

$$\rho^{(n)}_m = 2\pi\eta(\epsilon, -\lambda + n) \sum_j \rho_{jm} W^{-M,\lambda,n}_{jm}, \qquad \rho'^{(n)}_{m'} = \sum_{j'} \rho'_{j'm'} V^{-M,\lambda,n}_{j'm'}. \tag{7.4}$$

In conclusion, we see that a Lorentz pole contribution with factorizable residue can be decomposed into a series of Regge pole contributions, each with factorizable residue and with $l_n = \lambda - 1 - n$, $n = 0, 1, 2 \cdots$.

The factorizability of the residues of the generated Regge poles is not a trivial result and is a consequence of the identity (5.24).

Of course, one of the sums in Eq. (7.4) could give a vanishing result and the corresponding pole contribution could disappear.

Note added in proof: Results similar to those obtained in the present paper have been obtained independently by S. Ström [Arkiv Fysik **34**, 215 (1967)] by means of the infinitesimal method.

APPENDIX A. PROPERTIES OF THE FUNCTIONS $d^{M\lambda}_{mjj'}(\zeta)$

The matrix elements of the irreducible representations of $SL(2C)$ have been calculated by Ström.[22,23] Here we give a short treatment consistent with the notations and the phase conventions used in the text.

In order to calculate the matrix element (2.27), we have to find the matrix $[ua_z(\zeta)]_0$. It is easy to verify directly that, if u is given by Eq. (2.3), we can write

$$[ua_z(\zeta)]_0 = u_z(\mu) u_y(\theta') u_z(\nu), \tag{A1}$$

where

$$\tan \tfrac{1}{2}\theta' = \exp \zeta \tan \tfrac{1}{2}\theta. \tag{A2}$$

Moreover it is the case that

$$p[ua_z(\zeta)(ua_z(\zeta))_0^{-1}] = \exp(-\tfrac{1}{2}\zeta) \cos \tfrac{1}{2}\theta (\cos \tfrac{1}{2}\theta')^{-1}, \tag{A3}$$

and from Eqs. (2.18), (2.20), and (2.22) we have

$$\begin{aligned} &\mathfrak{D}^{M\lambda}_{jmj'm'}(a_z(\zeta)) \\ &= \int_{SU2} \overline{\Phi^{M}_{jm}(u)} [\mathfrak{D}^{M\lambda}(a_z(\zeta)) \Phi^{M}_{j'm'}](u)\, du \\ &= (2j+1)^{\frac{1}{2}} (2j'+1)^{\frac{1}{2}} (4\pi)^{-2} \\ &\quad \times \int_0^{4\pi} d\mu \int_0^{2\pi} d\nu \exp[i(m - m')\nu] \\ &\quad \times \exp[(1-\lambda)\zeta] (\cos \tfrac{1}{2}\theta)^{2(\lambda-1)} (\cos \tfrac{1}{2}\theta')^{2(1-\lambda)} \overline{r^{j}_{Mm}(\theta)} \\ &\quad \times r^{j'}_{Mm'}(\theta') \sin\theta\, d\theta = \delta_{mm'}\, d^{M\lambda}_{mjj'}(\zeta), \end{aligned} \tag{A4}$$

where

$$d^{M\lambda}_{mjj'}(\zeta) = (2j+1)^{\frac{1}{2}} (2j'+1)^{\frac{1}{2}} \times \int_0^{\pi} \exp[(1-\lambda)\zeta] \left(\cos\frac{\theta}{2}\right)^{2(\lambda-1)} \times \left(\cos\frac{\theta'}{2}\right)^{2(1-\lambda)} \overline{r^{j}_{Mm}(\theta)}\, r^{j'}_{Mm}(\theta') \tfrac{1}{2} \sin\theta\, d\theta. \tag{A5}$$

[22] S. Ström, Arkiv Fysik **29**, 467 (1965).
[23] S. Ström, Arkiv Fysik **33**, 465 (1967).

We use for the matrix elements of the representations of $SU(2)$ the expression[19,21]

$$r^j_{mm'}(\theta) = [(j+m)!\,(j-m)!\,(j+m')!\,(j-m')!]^{\frac{1}{2}} \times \sum_\alpha{}' (-1)^{\alpha+m-m'}[(j-m-\alpha)! \times (j+m'-\alpha)!\,(\alpha+m-m')!\,\alpha!]^{-1} \times \left(\cos\frac{\theta}{2}\right)^{2j+m'-m-2\alpha}\left(\sin\frac{\theta}{2}\right)^{m-m'+2\alpha} \quad \text{(A6)}$$

These functions have the symmetry properties

$$r^j_{mm'}(\theta) = \overline{r^j_{mm'}(\theta)} = r^j_{-m',-m}(\theta) = (-1)^{m-m'} r^j_{m'm}(\theta). \quad \text{(A7)}$$

Inserting these relations in Eq. (A5), we obtain the following properties of the functions $d^M_{mj'}(\zeta)$:

$$\begin{aligned} \overline{d^{M\lambda}_{mj'}(\zeta)} &= d^{M\bar\lambda}_{mj'}(\zeta), \\ d^{M\lambda}_{mj'}(\zeta) &= d^{-M,\lambda}_{-m,j,j'}(\zeta) \\ &= d^{m\lambda}_{Mj'}(\zeta). \end{aligned} \quad \text{(A8)}$$

If we use the expression (A7) for the functions $r^j_{mm'}(\theta)$, the integral (A5) can be evaluated in terms of elementary functions by means of the following change of variable:

$$\begin{aligned} x &= \exp(-\zeta)(\cos\tfrac{1}{2}\theta)^2(\cos\tfrac{1}{2}\theta')^{-2}, \\ (\cos\tfrac{1}{2}\theta')^2 &= [x^{-1} - \exp(-\zeta)][\exp\zeta - \exp(-\zeta)]^{-1}, \\ (\cos\tfrac{1}{2}\theta)^2 &= [\exp\zeta - x][\exp\zeta - \exp(-\zeta)]^{-1}, \\ \tfrac{1}{2}\sin\theta\,d\theta &= [\exp\zeta - \exp(-\zeta)]^{-1}\,dx. \end{aligned} \quad \text{(A9)}$$

In this we may obtain

$$\begin{aligned} d^{M\lambda}_{mj'}(\zeta) &= \Delta^j_{Mm}\Delta^{j'}_{Mm}[\exp\zeta - \exp(-\zeta)]^{-j-j'-1} \\ &\times \sum_{\alpha\alpha'}{}'(-1)^{\alpha+\alpha'}[(j-M-\alpha)! \\ &\times (j+m-\alpha)!\,(\alpha+M-m)!\,\alpha! \\ &\times (j'-M-\alpha')!\,(j'+m-\alpha')!\,(\alpha'+M-m)!\,\alpha'!]^{-1} \\ &\times \exp[\zeta(M-m-j'+2\alpha')] \\ &\times \int_{\exp(-\zeta)}^{\exp\zeta} x^{\lambda-j'-1}(\exp\zeta - x)^{j+j'+m-M-\alpha-\alpha'} \\ &\times [x - \exp(-\zeta)]^{M-m+\alpha+\alpha'}\,dx, \end{aligned} \quad \text{(A10)}$$

where Δ^j_{Mm} is given by Eq. (5.13). We can expand the integrand in Eq. (A10) by means of the binomial formula and perform the integration term by term. If we separate the contributions of the upper and lower integration limits, we obtain

$$d^{M\lambda}_{mj'}(\zeta) = a^{M\lambda}_{mj'}(\zeta) + (-1)^{j-j'}a^{-M,-\lambda}_{mj'j}(\zeta), \quad \text{(A11)}$$

where

$$\begin{aligned} a^{M\lambda}_{mj'}(\zeta) &= \Delta^j_{Mm}\Delta^{j'}_{Mm}[\exp\zeta - \exp(-\zeta)]^{-j-j'-1} \\ &\times \sum_{\alpha\alpha' rs}{}'(-1)^{j'-m+s}(j+j'+m-M-\alpha-\alpha')! \\ &\times (M-m+\alpha+\alpha')! \\ &\times [(j-M-\alpha)!\,(j+m-\alpha)!\,(\alpha+M-m)!\,\alpha! \\ &\times (j'-M-\alpha')!\,(j'+m-\alpha')! \\ &\times (\alpha'+M-m)!\,\alpha'!\,(r-\alpha)!\,(\alpha+j'-M-s-r)! \\ &\times (j+j'+m-M-\alpha'-r)! \\ &\times (2M-m-j'+\alpha'+r+s)!]^{-1}(M+s-\lambda)^{-1} \\ &\times \exp[\zeta(-\lambda+j+j'-M+m-2r)]. \end{aligned} \quad \text{(A12)}$$

If we substitute Eq. (A11) into the equation

$$d^{-M,-\lambda}_{mj'}(\zeta) = \mathfrak{U}^{-M,-\lambda}_j\, d^{M\lambda}_{mj'}(\zeta)\mathfrak{U}^{M\lambda}_{j'}, \quad \text{(A13)}$$

which is a simplified form of Eq. (2.29), and identify terms which have a similar dependence on ζ, we obtain

$$(-1)^{j-j'}a^{M\lambda}_{mj'j}(\zeta) = \mathfrak{U}^{-M,-\lambda}_j a^{M\lambda}_{mj'j}(\zeta)\mathfrak{U}^{M\lambda}_{j'}. \quad \text{(A14)}$$

Therefore Eq. (A11) can also be written in the form

$$d^{M\lambda}_{mj'}(\zeta) = a^{M\lambda}_{mj'}(\zeta) + \mathfrak{U}^{M\lambda}_j a^{-M,-\lambda}_{mj'j}(\zeta)\mathfrak{U}^{-M,-\lambda}_{j'}, \quad \text{(A15)}$$

from which Eq. (2.30) follows immediately. In a similar way, if we substitute Eq. (A11) into Eqs. (A8) we obtain

$$\begin{aligned} \overline{a^{M\lambda}_{mj'}(\zeta)} &= a^{M\bar\lambda}_{mj'}(\zeta), \\ a^{M\lambda}_{mj'}(\zeta) &= a^{-M,\lambda}_{-m,j,j'}(\zeta) = a^{m\lambda}_{Mj'}(\zeta). \end{aligned} \quad \text{(A16)}$$

From Eq. (A12) we can derive, after some calculations, an asymptotic expression for the functions $a^{M\lambda}_{mj'}(\zeta)$.

If $M \geq m$,

$$\begin{aligned} a^{M\lambda}_{mj'}(\zeta) &= (2j+1)^{\frac{1}{2}}(2j'+1)^{\frac{1}{2}} \\ &\times \left[\frac{(j-m)!\,(j+M)!\,(j'-m)!\,(j'+M)!}{(j+m)!\,(j-M)!\,(j'+m)!\,(j'-M)!}\right]^{\frac{1}{2}} \\ &\times \frac{(-1)^{j'+m}[\lambda-m+1]_{j'+m}}{(M-m)!\,[-\lambda-M]_{j'+M+1}} \\ &\times \exp[\zeta(-\lambda-1-M+m)][1+O(\exp(-2\zeta))]. \end{aligned} \quad \text{(A17)}$$

If we perform the limit $\zeta \to \infty$ in Eq. (A14) by means of Eq. (A17), we can compute the coefficients $\mathfrak{U}^{M\lambda}_j$. The result is

$$\mathfrak{U}^{M\lambda}_j = \prod_{s=|M|}^{j} \frac{s-\lambda}{s+\lambda}. \quad \text{(A18)}$$

APPENDIX B. CALCULATION OF THE FUNCTIONS $K^{M\lambda}_m(\epsilon, l, +; j)$

In order to evaluate the integrals (4.25) and (5.9) we perform the following change of variable:

$$(\sinh\tfrac{1}{2}\zeta)^2 = z, \quad \tfrac{1}{2}\sinh\zeta\,d\zeta = dz,$$
$$\left(\cos\frac{\theta^+}{2}\right)^2 = \frac{1+z}{1+2z}, \quad \left(\sin\frac{\theta^+}{2}\right)^2 = \frac{z}{1+2z}. \quad \text{(B1)}$$

The functions $d^l_{mm'}(\zeta)$ and $a^l_{mm'}(\zeta)$ can be written by

means of the Meijer G functions using the Eqs. (5.6.24) and (5.3.9) of Ref. 20. The result is

$$d^l_{mm'}(\zeta) = [\Gamma(m-l)\Gamma(m'+l+1)]^{-1} \times \left(\frac{1+z}{z}\right)^{(m+m')/2} G^{12}_{22}\left(z \,\middle|\, \begin{matrix} -l, & l+1 \\ m, & m' \end{matrix}\right) \quad (m \geq m'), \tag{B2}$$

$$a^l_{mm'}(\zeta) = [-2\cos\pi(l-m')\Gamma(m-l)\Gamma(m'+l+1)]^{-1} \times \left(\frac{1+z}{z}\right)^{(m+m')/2} G^{21}_{22}\left(z \,\middle|\, \begin{matrix} -l, & l+1 \\ m, & m' \end{matrix}\right). \tag{B3}$$

If we use for the functions $r^j_{mm'}(\theta)$ the expression (A6), the integral (4.25) becomes, for $m \geq M$[21],

$$K^{M\lambda}_m(\epsilon, l, +; j) = \Delta^j_{Mm}[\Gamma(m-l)\Gamma(M+l+1)]^{-1} \times \sum_\alpha{}' (-1)^\alpha [(j-M-\alpha)!\,(j+m-\alpha)! \times (\alpha+M-m)!\,\alpha!]^{-1} \int_0^\infty (1+2z)^{\lambda-j-1} z^{\alpha-m} \times (1+z)^{m+j-\alpha} G^{12}_{22}\left(z \,\middle|\, \begin{matrix} -l, & l+1 \\ m, & M \end{matrix}\right) dz, \tag{B4}$$

where Δ^j_{Mm} is given by Eq. (5.13). If we expand the term $(1+z)^{m+j-\alpha}$ by means of the binomial formula, we obtain

$$K^{M\lambda}_m(\epsilon, l, +; j) = \Delta^j_{Mm}[\Gamma(m-l)\Gamma(M+l+1)]^{-1} \sum_{\alpha\beta}{}' (-1)^\alpha 2^{\lambda-j-1} \times [(j-M-\alpha)!\,(\alpha+M-m)!\,\alpha!\,\beta! \times (m+j-\alpha-\beta)!]^{-1} \int_0^\infty (z+\tfrac{1}{2})^{\lambda-j-1} z^{\alpha+\beta-m} \times G^{12}_{22}\left(z \,\middle|\, \begin{matrix} -l, & l+1 \\ m, & M \end{matrix}\right) dz. \tag{B5}$$

This integration can be performed by means of the formula (20.5.4) of Ref. 24, and the result, after the change $\beta \to m+j-\alpha-\beta$ in the summation index, is given in Eq. (5.12). The result (5.15) is obtained in a completely similar way.

From the general definition of the G function given by Eq. (5.3.1) of Ref. 20, we have

$$G^{23}_{33}\left(\tfrac{1}{2} \,\middle|\, \begin{matrix} \beta-j, & -l, & l+1 \\ \beta-\lambda, & m, & M \end{matrix}\right) = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} \Gamma(\beta-\lambda-s)\Gamma(m-s)\Gamma(1-\beta+j+s) \times \Gamma(1+l+s)\Gamma(-l+s)[\Gamma(1-M+s)]^{-1} 2^{-s}\, ds, \tag{B6}$$

where the integration path leaves the poles of

$$\Gamma(\beta-\lambda-s)\Gamma(m-s) \tag{B7}$$

on the right and the poles of

$$\Gamma(1+l+s)\Gamma(-l+s) \tag{B8}$$

on the left. The poles of $\Gamma(1-\beta+j+s)$ are canceled by the poles of $\Gamma(1-M+s)$.

The singularities in the G function arise when the integration path is pinched by two of the above-mentioned poles. This happens for

$$\begin{aligned} m+1+l &= -n, \\ m-l &= -n, \\ \beta+l-\lambda+1 &= -n, \\ \beta-\lambda-l &= -n, \end{aligned} \tag{B9}$$

where n is a nonnegative integer. The first two conditions give rise to poles which are canceled by the poles of the Γ functions appearing in Eq. (5.12), and the last two conditions give rise to the poles described by Eq. (5.14).

The discussion of the singularities of the function (5.15) can be performed in a similar way.

In order to determine the residues $W^{M\lambda n}_{jm}$ defined by Eq. (5.20), we use Eq. (5.3.5) of Ref. 20 and write Eq. (5.12) in the form

$$K^{M\lambda}_m(\epsilon, l, +; j) = \Delta^j_{Mm} \sum_{\alpha\beta}{}' (-1)^\alpha 2^{\beta-j-1} [(j-M-\alpha)!\,(\alpha+M-m)!\,(j+m-\alpha-\beta)!\,\alpha!\,\beta!]^{-1} \times \left\{ \frac{2^{\lambda-\beta}\Gamma(\lambda+m-\beta)\Gamma(-\lambda+l+\beta+1)\Gamma(-\lambda-l+\beta)}{\Gamma(l+M+1)\Gamma(m-l)\Gamma(-\lambda-M+\beta+1)} \times F_{32}\left[\begin{matrix} -\lambda+j+1, & -\lambda+l+\beta+1, & -\lambda+l+\beta; \\ -\lambda-m+\beta+1, & -\lambda-M+\beta+1; & \tfrac{1}{2} \end{matrix}\right] + \frac{2^{-m}\Gamma(-\lambda-m+\beta)\Gamma(m+j-\beta+1)\Gamma(m+l+1)}{\Gamma(l+M+1)\Gamma(m-M+1)\Gamma(-\lambda+j+1)} \times F_{32}\left[\begin{matrix} m+j-\beta+1, & m+l+1, & m-l; \\ \lambda+m-\beta+1, & m-M+1; & \tfrac{1}{2} \end{matrix}\right] \right\}. \tag{B10}$$

[24] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Tables of Integral Transforms* (McGraw-Hill Book Company, New York, 1954), Vol. 2.

We can easily see that the only singularity for $l \to \lambda - n - 1$ is due to the function $\Gamma(-\lambda + l + \beta + 1)$. Therefore, using the formula

$$\lim_{z \to -n} (z + n)\Gamma(z) = \frac{(-1)^n}{n!}, \tag{B11}$$

we have

$$W_{jm}^{M\lambda n} = \Delta_{Mm}^{j} \sum_{\alpha\beta}{}' (-1)^{\alpha+\beta+n} 2^{\lambda-j-1} [(j - M - \alpha)!\,(\alpha + M - m)!\,(j + m - \alpha - \beta)!\,\alpha!\,\beta!\,(n - \beta)!]^{-1}$$
$$\times \frac{\Gamma(\lambda + m - \beta)\Gamma(-2\lambda + n + \beta + 1)}{\Gamma(\lambda - n + M)\Gamma(m - \lambda + n + 1)\Gamma(-\lambda - M + \beta + 1)}$$
$$\times F_{32}\begin{bmatrix} -\lambda + j + 1, & \beta - n, & -2\lambda + \beta + n + 1; \\ -\lambda - m + \beta + 1, & -\lambda - M + \beta + 1; & \tfrac{1}{2} \end{bmatrix}. \tag{B12}$$

As $\beta - n$ is a nonpositive integer, the generalized hypergeometric series is a finite sum. If we write it explicitly, after some calculations using well-known properties of the Γ functions, we obtain Eq. (5.22).

At last we want to study the behavior of the function $K_m^{M\lambda}(\epsilon, l_0 + il', +; j)$ when $l' \to \pm\infty$ for fixed values of the other parameters. We use for the G functions which appear in Eq. (5.12) the integral representation (B6), and we shift the integration path on the left taking into account the contributions of the poles which are crossed. The new integration path is along the line $\operatorname{Re} s = s_0$ where $s_0 < \beta - \lambda$, $s_0 < m$.

From the Stirling formula [Ref. 20, Eq. (1.18.6)], we see that for real a and for $x \to \pm\infty$

$$|\Gamma(a + ix)| \sim (2\pi)^{\frac{1}{2}} |x|^{a-\frac{1}{2}} \exp(-\tfrac{1}{2}\pi |x|). \tag{B13}$$

It follows that a function $\varphi(a)$ exists with the property

$$|\Gamma(a + ix)| \leq \varphi(a)(1 + |x|)^{a-\frac{1}{2}} \exp(-\tfrac{1}{2}\pi |x|). \tag{B14}$$

From this equation, if $s_0 < -|l_0 + \tfrac{1}{2}|$, after some calculation we have

$$|\Gamma(1 + s_0 + l_0 + i(s' + l'))\Gamma(s_0 - l_0 + i(s' - l'))|$$
$$\leq \varphi(s_0 + l_0)\varphi(s_0 - l_0)\, |2l'|^{|l_0+\frac{1}{2}|+s_0} \exp(-\pi |l'|). \tag{B15}$$

This inequality can be used to find an upper bound for the integral along the shifted path. We obtain easily that this integral is of the order

$$O[|l'|^{|l_0+\frac{1}{2}|+s_0} \exp(-\pi |l'|)].$$

If we explicitly evaluate the contributions of the poles, we see by means of Eq. (B13) that they decrease more rapidly than $\exp(-2\pi |l'|)$, and are therefore dominated by the contribution of the integral along the shifted path. As s_0 can be chosen arbitrarily negative, we see that the G function is of the order $O(|l'|^{\alpha} \exp[-\pi |l'|])$, where α is arbitrary.

If we insert this result into Eq. (5.12) and again use Eq. (B13), we see that the function $K_m^{M\lambda}(\epsilon, l_0 + il', +; j)$ decreases faster than any negative power of $|l'|$.

APPENDIX C. ASYMPTOTIC PROPERTIES OF AN INTEGRAL

In order to prove Eq. (6.14) we need the following inequality:

$$|A_{mm'}^{l}(v)| = |a_{mm'}^{l}(\zeta)|$$
$$\leq |\tan \pi(l + m) \cot \pi(\operatorname{Re} l + m) a_{mm'}^{\operatorname{Re} l}(\zeta)| \qquad (\operatorname{Re} l > 1 - m). \tag{C1}$$

This inequality is an immediate consequence of the following integral representation, which holds for $\operatorname{Re} l > 1 - m$ and is a consequence of Eq. (2.1.10) of Ref. 20.

$$a_{mm'}^{l}(\zeta) = (2\pi)^{-1}(-1)^{m+m'} \tan \pi(l + m)(\sinh \tfrac{1}{2}\zeta)^{-2l-2}$$
$$\times (\tanh \tfrac{1}{2}\zeta)^{m+m'} \int_0^1 t^{l-m}(1 - t)^{l+m}$$
$$\times [1 + t(\sinh \tfrac{1}{2}\zeta)^{-2}]^{-l-1+m'}\, dt. \tag{C2}$$

From Eq. (C1) we see that the integral (6.14) is smaller in modulus, then

$$|a_{mm'}^{-L-1}(\zeta) \cot \pi(L + m)| \int_{-\infty}^{+\infty} |K_m^{M,-\lambda}(\epsilon, L + il', \rho; j)|$$
$$\times |K_{m'}^{M\lambda}(\epsilon, -L - 1 - il', \rho; j')| \cdot |2L + 2il' + 1|\, dl'. \tag{C3}$$

As from Eq. (3.14), we have that

$$a_{mm'}^{-L-1}(\zeta) = O(\exp(-L\zeta)).$$

Eq. (6.14) is proved if we show that the integral in Eq. (C3) converges. But this follows from the fact proved in the Appendix B that

$$|K_m^{M\lambda}(\epsilon, l_0 + il', \rho; j)|$$

decreases faster than any power of $|l'|$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Mathematical Methods for Evaluating Second-Order Three-Body Interactions between Atoms or Ions with Gaussian Wavefunctions

SHIMSHON ZIMERING
Battelle Institute, Advanced Studies Center, Geneva, Switzerland

(Received 7 July 1966)

It is shown that the integrals occurring in the expression for the interaction energy between three atoms or ions in second-order perturbation theory with Gaussian wavefunctions can be reduced to single integrals of three different types. The first two types are erf x and erf ix functions, whereas the third type is a single integral of the error function which is easily evaluated by electronic computation.

INTRODUCTION

IT is well known that Gaussian-type electron wavefunctions are particularly simple for the evaluation of multicenter integrals occurring in the energy expressions for atomic and molecular systems via a variational method. Equivalently, first-order perturbation energies can be equally and readily evaluated on the basis of such Gaussian functions.

In a previous publication,[1] we have shown that Gaussian-type wavefunctions can also be utilized in second-order perturbation theory for interacting systems of atoms or ions, in the evaluation of both direct and exchange integrals. However, the methods developed were essentially based on asymptotic expansions, which are sufficiently accurate only if the distances between the atoms are relatively large. This condition is fulfilled in the problem of evaluating the three-atom energy in rare-gas crystals; on the other hand, they cannot be applied to ionic solids, since in this case the distances are much smaller because of electrostatic compression of the crystals (Madelung energy).

For the evaluation of three-ion crystal energies it is necessary, therefore, to avoid asymptotic expansions of the integrals. In this paper we present general methods for evaluating three-atom and three-ion interaction energies in second-order perturbation theory, valid for all distances between the atoms or ions. It appears that all expressions can be written as linear combinations of three types of integrals; two of these are the well-known erf x and erf ix functions, whereas the third type can be readily evaluated by electronic computation. In this analysis we will consider the general case of atoms or ions of two different sizes, i.e., the inverse widths of their Gaussian wavefunctions are generally different.

[1] S. Zimering, J. Math. Phys. 6, 336 (1965).

TYPES OF INTEGRALS OCCURRING IN THE EXPRESSION FOR THE SECOND-ORDER INTERACTION ENERGY

We consider a triplet of atoms or ions (abc) and three electrons, denoted by 1, 2, 3. The distances between the atoms are denoted by R_{ab}, that between the electrons by r_{12}, and those between nuclei and electrons by r_{a1}, etc.

We introduce the following notations: φ_a, φ_b, and φ_c are the ground-state wavefunctions for atoms a, b, and c, respectively,

$$\varphi_a(1) = (\beta_a/\Pi^{\frac{1}{2}})^{\frac{3}{2}} \exp(-\beta_a^2 r_{a1}^2/2) \text{ etc.},$$

where β_a is the Gaussian parameter for atom a, etc.

H'_{ab} is the electrostatic interaction (perturbation Hamiltonian) between atoms a and b.

$\Delta_{ab} = \int \varphi_a \varphi_b \, d\tau$ is the overlap integral between a and b,

$$\Delta_{abc}^2 = \Delta_{ab}^2 + \Delta_{ac}^2 + \Delta_{bc}^2 - 2\Delta_{ab}\Delta_{ac}\Delta_{bc},$$

Ψ is the zero-order total wavefunction (Slater determinant),

$$\Psi = [3!\,(1 - \Delta_{abc}^2)]^{-\frac{1}{2}} \det\{\varphi_a(1)\varphi_b(2)\varphi_c(3)\} \quad (1)$$

and

$$\Psi_{(ab)} = [2!\,(1 - \Delta_{ab}^2)]^{-\frac{1}{2}} \det\{\varphi_a(1)\varphi_b(2)\}.$$

We have, for the first-order perturbation energy,

$$\langle H'_{abc}\rangle = \langle H'_{ab}\rangle + \langle H'_{ac}\rangle + \langle H'_{bc}\rangle,$$

where

$$\langle H'_{ab}\rangle = \iiint \Psi^* H'_{ab} \Psi \, d\tau_1 \, d\tau_2 \, d\tau_3.$$

On the other hand, the sum of *pair interactions* for the triplet (abc) is, in first order, given by

$$\langle H'_{abc}\rangle_0 = \langle H'_{ab}\rangle_0 + \langle H'_{ac}\rangle_0 + \langle H'_{bc}\rangle_0,$$

with

$$\langle H'_{ab}\rangle_0 = \iint \Psi^*_{(ab)} H'_{ab} \Psi_{(ab)} \, d\tau_1 \, d\tau_2.$$

The expression for the second-order energy is

$$E_2 = \sum_{\chi \neq 0} (H'_{abc})_{0\chi}(H'_{abc})_{\chi 0}/(E_0 - E_\chi)$$
$$= -\{\langle H'^2_{abc}\rangle - \langle H'_{abc}\rangle^2\}/E_{\mathrm{av}},$$

where χ numbers the excited states of the system (energy E_χ) and E_{av} is defined by the averaging procedure. On the other hand, the second-order total *pair energy* is given by

$$E_2^{(0)} = -(1/E_{\mathrm{av}})\{\langle H'^2_{ab}\rangle_0 - \langle H'_{ab}\rangle_0^2 + [(ac), (bc)]\},$$

where $[(ac), (bc)]$ signifies that the corresponding expressions for the pairs (ac) and (bc) are to be added. Since $\langle H'_{abc}\rangle$ and $\langle H'_{abc}\rangle_0$ can be obtained readily from first-order calculations,[1] the only difficulty lies in the evaluation of $\langle H'^2_{abc}\rangle$, defined by

$$\langle H'^2_{abc}\rangle = \iiint \Psi^* H'^2_{abc} \Psi \, d\tau_1 \, d\tau_2 \, d\tau_3 .$$

We consider the general case where the interacting atoms or ions are represented by two different Gaussian parameters β and β'.

By substituting the expression for H'_{abc} into $\langle H'^2_{abc}\rangle$, we find, using (1), that this quantity is a linear combination of the following seven types of volume integrals (some of them multiplied by overlap integrals):

$$J_1 = \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^3 \int \frac{\exp(-\alpha^2 r_{A1}^2)}{(\alpha r_{B1})^2} \, d\tau_1 , \tag{2}$$

$$J_2 = \zeta^{\frac{3}{2}} \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^6 \iint \frac{\exp -[\alpha^2(r_{A1}^2 + \zeta r_{B2}^2)]}{(\alpha r_{12})^2} \, d\tau_1 \, d\tau_2 , \tag{3}$$

$$J_3 = \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^3 \int \frac{\exp(-\alpha^2 r_{A1}^2)}{\alpha r_{B1}} \, d\tau_1 , \tag{4}$$

$$J_4 = \zeta^{\frac{3}{2}} \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^6 \iint \frac{\exp[-\alpha^2(r_{A1}^2 + \zeta r_{B2}^2)]}{\alpha r_{12}} \, d\tau_1 \, d\tau_2 , \tag{5}$$

$$J_5 = \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^3 \int \frac{\exp(-\alpha^2 r_{C1}^2)}{\alpha r_{A1} \alpha r_{B1}} \, d\tau_1 , \tag{6}$$

$$J_6 = \zeta^{\frac{3}{2}} \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^6 \iint \frac{\exp[-\alpha^2(r_{C1}^2 + \zeta r_{A2}^2)]}{\alpha r_{B1} \alpha r_{12}} \, d\tau_1 \, d\tau_2 , \tag{7}$$

and

$$J_7 = \zeta^{\frac{3}{2}} \eta^{\frac{3}{2}} \left(\frac{\alpha}{\Pi^{\frac{1}{2}}}\right)^9$$
$$\times \iiint \frac{\exp[-\alpha^2(r_{C1}^2 + \zeta r_{A2}^2 + \eta r_{B3}^2)]}{\alpha r_{12} \alpha r_{13}} \, d\tau_1 \, d\tau_2 \, d\tau_3 . \tag{8}$$

In Eqs. (2)–(8), $d\tau$ denotes the element of volume, r_{A1} the distance between a fixed point A on the triangle (abc) and electron 1 [for example, A may be the middle of the side (bc)]; α takes the values β, β', and $\beta[\frac{1}{2}(1 + \beta'^2/\beta^2)]^{\frac{1}{2}}$ and ζ and η are different ratios between these last parameters. These seven integrals are generalizations of Eqs. (41)–(50) given in Ref. 1 for $\alpha = \beta = \beta'$.

Upon inspection of the volume integrals J_1 to J_7, it appears possible to reduce them to the three following basic integrals:

$$A(X) = \frac{2}{X} e^{-X^2} \int_0^X e^{u^2} \, du,$$

$$B(X) = \frac{2}{X\Pi^{\frac{1}{2}}} \int_0^X e^{-u^2} \, du = \frac{\operatorname{erf} X}{X},$$

and

$$C\big((X_1, \eta), (X_2, \zeta), X_3\big) =$$
$$\frac{2}{\Pi^{\frac{1}{2}}} \int_0^{(\zeta/\zeta+1)^{\frac{1}{2}}} \frac{e^{-X_2^2 z^2} \operatorname{erf}([P(z)]^{\frac{1}{2}}[\eta/1 + \eta(1 - z^2)]^{\frac{1}{2}})}{[P(z)]^{\frac{1}{2}}} \, dz, \tag{9}$$

where

$$P(z) = X_1^2 + (X_3^2 - X_2^2 - X_1^2)z^2 + X_2^2 z^4,$$

$\zeta, \eta > 0$ and X_1, X_2, X_3 are dimensionless quantities proportional to the lengths of the three sides of the triangle specified by X_1, X_2, X_3.

We note that $P(z) > 0$ $(1 \geq z \geq 0)$ for every triangular configuration except when $\sphericalangle(X_1, X_2) = 0$. In particular, when $X_1 = X_2$ and $X_3 = 0$, we have $P(1) = 0$ and

$$C\big((X, \infty), (X, \infty), 0\big) = A(X).$$

We note also that in many cases the function $C\big((X_1, \eta), (X_2, \zeta), X_3\big)$ can be simplified[1]; for example,

$$C\big((X, \infty), (0, 1), X\big) = 1 - X^2 B^2(X) + (2/\Pi^{\frac{1}{2}})$$
$$\cdot [B(2^{\frac{1}{2}}X) - e^{-X^2} B(X)].$$

Finally we remark that $A(0) = 2$, $B(0) = 2\Pi^{-\frac{1}{2}}$.

The relations between the volume integrals $J_1 - J_7$ and the basic integrals A to C are as follows:

$$J_1 = A(\alpha R_{AB}), \tag{10}$$

$$J_2 = \frac{\zeta}{\zeta + 1} A\left(\left(\frac{\zeta}{\zeta + 1}\right)^{\frac{1}{2}} \alpha R_{AB}\right), \tag{11}$$

$$J_3 = B(\alpha R_{AB}), \tag{12}$$

$$J_4 = \left(\frac{\zeta}{\zeta + 1}\right)^{\frac{1}{2}} B\left(\left(\frac{\zeta}{\zeta + 1}\right)^{\frac{1}{2}} \alpha R_{AB}\right), \tag{13}$$

$$J_5 = C\big((\alpha R_{BC}, \infty), (\alpha R_{AC}, \infty), \alpha R_{AB}\big), \tag{14}$$

$$J_6 = C\big((\alpha R_{BC}, \infty), (\alpha R_{AC}, \zeta), \alpha R_{AB}\big), \tag{15}$$

$$J_7 = C\big((\alpha R_{BC}, \eta), (\alpha R_{AC}, \zeta), \alpha R_{AB}\big). \tag{16}$$

DERIVATION OF EQS. (10)–(16)

Equations (10)–(13) follow by using the formulas

$$\frac{1}{r^2} = \int_0^\infty e^{-r^2 t}\, dt \tag{17}$$

and

$$\frac{1}{r} = 2\Pi^{-\frac{1}{2}} \int_0^\infty e^{-r^2 t^2}\, dt, \tag{18}$$

and by integrating with respect to the coordinates of electrons 1 and 2. We sketch the derivation of J_2, which is the most complicated one. Taking Cartesian coordinates (x, y, z) for electron 1 and (u, v, w) for electron 2 and using (17), we obtain

$$\begin{aligned} J_2 = \zeta^{\frac{3}{2}}\alpha^6\Pi^{-3} \int_0^\infty \Big\{ & \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \\ & \times \exp\left[-\alpha^2(x^2 + y^2 + z^2)\right] \\ & \times \exp\left[-\zeta\alpha^2\{u^2 + (v - R_{AB})^2 + w^2\}\right] \\ & \times \exp\left[-t\alpha^2\{(x-u)^2 + (y-v)^2 + (z-w)^2\}\right] \\ & \times dx\, dy\, dz\, du\, dv\, dw \Big\}\, dt. \end{aligned}$$

Integration with respect to x, y, z, u, v, and w gives

$$\begin{aligned} J_2 = \zeta^{\frac{3}{2}} \exp\left[-\frac{R^{*2}}{(1+\zeta)}\right] \\ \times \int_0^\infty \frac{\exp\left[\zeta R^{*2}/(1+\zeta)(\zeta + t + t\zeta)\right]}{(\zeta + t + t\zeta)^{\frac{3}{2}}}\, dt, \end{aligned} \tag{19}$$

where $R^* = \zeta^{\frac{1}{2}}\alpha R_{AB}$. Equation (11) is obtained by putting in (19),

$$u = R^*\{\zeta/(1+\zeta)(\zeta + t + t\zeta)\}^{\frac{1}{2}}.$$

For the derivation of Eqs. (14)–(16), we sketch the first and the last ones, since the derivation of J_6 is very similar to that of J_5.

Let the Cartesian coordinates of the vertices A, B, C be $(b_1, b_2, 0)$, $(a, 0, 0)$, and $(0, 0, 0)$, respectively, and that of electron 1 be (x, y, z). Then, using (18), we obtain

$$\begin{aligned} J_5 = 4\alpha^3\Pi^{-\frac{5}{2}} \\ \times \int_0^\infty \int_0^\infty \Big\{ \Big(\int_{-\infty}^\infty & \exp\left\{-\alpha^2[x^2 + (x-a)^2u^2 \right. \\ & \left. + (x - b_1)^2t^2]\right\} dx \Big) \\ \times \Big(\int_{-\infty}^\infty & \exp\left\{-\alpha^2[y^2(1+u^2) + (y-b_2)^2t^2]\right\} dy \Big) \\ \times \Big(\int_{-\infty}^\infty & \exp\left[-\alpha^2 z^2(1 + t^2 + u^2)\right] dz \Big) \Big\}\, du\, dt. \end{aligned}$$

By integration with respect to x, y, and z we find

$$J_5 = \frac{4}{\Pi} \int_0^\infty \int_0^\infty f(t, u)\, dt\, du, \tag{20}$$

where

$$f(t, u) = \frac{\exp\left[-(a^2t^2 + b^2u^2 + c^2t^2u^2)/(1 + t^2 + u^2)\right]}{(1 + t^2 + u^2)^{\frac{3}{2}}} \tag{21}$$

and

$$a = \alpha R_{BC}, \quad b = \alpha R_{AC}, \quad c = \alpha R_{AB}.$$

By using the substitution

$$w = \left\{\frac{P(u)}{u^2 + 1}\, \frac{t^2}{1 + t^2 + u^2}\right\}^{\frac{1}{2}}, \tag{22}$$

where

$$P(u) = a^2(u^2 + 1) - b^2u^2 + c^2u^2(u^2 + 1),$$

we obtain

$$(1 + t^2 + u^2)^{-\frac{3}{2}}\, dt = \{(u^2 + 1)P(u)\}^{-\frac{1}{2}}\, dw;$$

$$\begin{aligned} \exp\left[-(a^2t^2 + b^2u^2 + c^2t^2u^2)/(1 + t^2 + u^2)\right] \\ = \exp\left\{-b^2u^2/(1 + u^2)\right\} \exp\left(-w^2\right) \end{aligned}$$

and, finally,

$$\begin{aligned} J_5 = 2\Pi^{-\frac{1}{2}} \int_0^\infty \\ \times \frac{\exp\left\{-b^2[u^2/(1+u^2)]\right\} \operatorname{erf}\left\{[P(u)/(1+u^2)]^{\frac{1}{2}}\right\}}{[(u^2+1)P(u)]^{\frac{1}{2}}}\, du. \end{aligned} \tag{23}$$

Equation (14) is obtained from (23) by the substitution

$$z = \{u^2/(1 + u^2)\}^{\frac{1}{2}}. \tag{24}$$

For the derivation of J_7 we use the formula

$$\frac{\operatorname{erf} r}{r} = 2\Pi^{-\frac{1}{2}} \int_0^1 e^{-r^2t^2}\, dt. \tag{25}$$

Let, again, the Cartesian coordinates of the vertices A, B, C be $(b_1, b_2, 0)$, $(a, 0, 0)$, $(0, 0, 0)$, respectively, and that of electron 1 be (x, y, z). Selecting for electrons 2 and 3 polar coordinates $u = r_{A2}$, $\alpha = \sphericalangle 1A2$ and $w = r_{B3}$, $\beta = \sphericalangle 1B3$, respectively, we obtain

$$\begin{aligned} J_7 = J_7\big((\alpha R_{BC}, \eta), (\alpha R_{AC}, \zeta), \alpha R_{AB}\big) = 4\alpha^7\Pi^{-\frac{5}{2}}\zeta^{\frac{3}{2}}\eta^{\frac{3}{2}} \\ \times \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \Big\{ & \exp\left[-\alpha^2(x^2 + y^2 + z^2)\right] \\ & \times \left[\int_0^\infty \int_{-1}^1 \frac{u^2 e^{-\zeta\alpha^2u^2}\, dm\, du}{(r_{A1}^2 + u^2 - 2r_{A1}um)^{\frac{1}{2}}}\right] \\ & \times \left[\int_0^\infty \int_{-1}^1 \frac{w^2 e^{-n\alpha^2w^2}\, dn\, dw}{(r_{B1}^2 + w^2 - 2r_{B1}wn)^{\frac{1}{2}}}\right] \Big\}\, dx\, dy\, dz, \end{aligned}$$

where $m = \cos\alpha$ and $n = \cos\beta$.

By putting $x = \zeta^{\frac{1}{2}}\alpha u$ and using the formula[1]

$$\int_0^\infty \int_{-1}^1 \frac{x^2 e^{-x^2}\, dm\, dx}{(R^2 + x^2 - 2Rxm)^{\frac{1}{2}}} = \frac{\Pi^{\frac{1}{2}} \operatorname{erf} R}{2R},$$

we obtain

$$J_7 = \alpha^3 \zeta^{\frac{1}{2}} \eta^{\frac{1}{2}} \Pi^{-\frac{3}{2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\exp(-\alpha^2 r_{C1}^2) \times \frac{\operatorname{erf}(\zeta^{\frac{1}{2}}\alpha r_{A1})}{\zeta^{\frac{1}{2}}\alpha r_{A1}} \frac{\operatorname{erf}(\eta^{\frac{1}{2}}\alpha r_{B1})}{\eta^{\frac{1}{2}}\alpha r_{B1}}\right] dx\, dy\, dz, \quad (26)$$

where

$$r_{A1} = \{(x - b_1)^2 + (y - b_2)^2 + z^2\}^{\frac{1}{2}},$$
$$r_{B1} = \{(x - a)^2 + y^2 + z^2\}^{\frac{1}{2}},$$

and

$$r_{C1} = \{x^2 + y^2 + z^2\}^{\frac{1}{2}}.$$

Further, using (25) and integrating in (26) with respect to x, y, and z, it is found that

$$J_7 = \frac{4}{\Pi} \int_{u=0}^{\zeta^{\frac{1}{2}}} \int_{t=0}^{\eta^{\frac{1}{2}}} f(t, u)\, dt\, du, \quad (27)$$

with $f(t, u)$ defined by (21).

The formula (16) follows now from (27), by using the substitutions (22) and (24).

ASYMPTOTIC EXPANSIONS FOR INTEGRALS A, B, AND C

In the previous sections we have given analytical expressions for the integrals A, B, and C occurring in the evaluation of three-atom or three-ion interactions. Values for the integrals A and B are available in the form of tables, whereas C must be evaluated by electronic computation. The computation of C can be simplified considerably for large values of X_1, X_2, X_3 (≥ 2.5) by using asymptotic series expansions. We here give the expansions used, including for completeness also the corresponding series for the integrals A and B, and discuss briefly their derivation. The asymptotic series expansions are

$$A(X) = \frac{1}{X^2} + \frac{1}{2X^4} + \frac{1\cdot 3}{4X^6} + \frac{1\cdot 3\cdot 5}{8X^8} + \cdots = \sum_{n=1}^{N} \frac{(2n-3)!!}{2^{n-1}X^{2n}} + O(X^{-2N-2}),$$

$$\frac{1}{X} - B(X) = \frac{e^{-X^2}}{\Pi^{\frac{1}{2}}X^2}\left(1 - \frac{1}{2X^2} + \frac{1\cdot 3}{(2X^2)^2} - \cdots\right) = \frac{e^{-X^2}}{\Pi^{\frac{1}{2}}X^2}\left(\sum_{n=1}^{N} \frac{(-1)^{n-1}(2n-3)!!}{(2X^2)^{n-1}} + O(X^{-2N})\right)$$

[where $(-1)!! = 1$], and

$$C((X_1, \eta), (X_2, \zeta), X_3) = \frac{1}{X_1X_2} + \frac{\cos\Theta}{2X_1^2X_2^2} + \frac{3(3\cos^2\Theta - 1)}{8X_1^3X_2^3} + \frac{15\cos\Theta(5\cos^2\Theta - 3)}{16X_1^4X_2^4} + \frac{105(35\cos^4\Theta - 30\cos^2\Theta + 3)}{128X_1^5X_2^5} + O(X_1^{-6}X_2^{-6}), \quad (28)$$

for $X_1 \geq X_2 \to \infty$ and $\Theta = \sphericalangle(X_1, X_2) \geq \frac{1}{3}\Pi$.

The first two expansions are easily established by applying l'Hospital's rule. For C, we consider a triangular configuration for which

$$X_1 \geq X_2 \quad \text{and} \quad \Theta \geq \Pi/3. \quad (29)$$

We assume, moreover, that X_1 and $[\zeta/(1+\zeta)]^{\frac{1}{2}}X_2$ are sufficiently large {i.e., X_1 and $[\zeta/(1+\zeta)]^{\frac{1}{2}}X_2 \geq 2, 5$} and that η is not too small ($\eta > 0, 8$).

Putting

$$k = X_1^{-2}(3X_1^2 + X_2^2 - X_3^2) = 2(1 + X_2\cos\Theta/X_1),$$
$$\mu = 1 - (X_2/X_1)^2 \quad (30)$$

and

$$g(z) = \{z/(1+z^2)\}^2(k + \mu z^2),$$

we have, by definition,

$$\{P(z)\}^{\frac{1}{2}} = X_1(1+z^2)\{1 - g(z)\}$$

and obtain

$$C((X_1, \eta), (X_2, \zeta), X_3) = 2\Pi^{-\frac{1}{2}}X_1^{-1} \int_0^{[\zeta/(1+\zeta)]^{\frac{1}{2}}} \frac{e^{-X_2^2z^2} \operatorname{erf}\left\{X_1(1+z^2)\left[(1 - g(z))\frac{\eta}{1+\eta(1-z^2)}\right]^{\frac{1}{2}}\right\}}{(1+z^2)[1-g(z)]^{\frac{1}{2}}}\, dz. \quad (31)$$

Since

$$0 \leq \{z/(1+z^2)\}^2 < \tfrac{1}{4}, \quad (1 > z \geq 0),$$

it follows from (29) and (30) that $1 > g(z) \geq 0$ ($1 > z \geq 0$). Equation (28) is obtained from (31) by developing the function $\{(1+z^2)[1-g(z)]^{\frac{1}{2}}\}^{-1}$ in a power series for $1 > z \geq 0$, approximating

$$\operatorname{erf}\left\{X_1(1+z^2)\left[(1 - g(z))\frac{\eta}{1+\eta(1-z^2)}\right]^{\frac{1}{2}}\right\}$$

by 1 and approximating

$$\int_0^{[\zeta(\zeta+1)]^{\frac{1}{2}}-\epsilon} z^{2n} e^{-X_2^2z^2}\, dz$$

by

$$\int_0^\infty z^{2n} e^{-X_2^2z^2}\, dz = \frac{\Pi^{\frac{1}{2}}(2n-1)!!}{2^{n+1}X_2^{2n+1}}, \quad (n = 0, 1, 2, \cdots),$$

where ϵ is of small positive value.

To indicate the accuracy of the third asymptotic expansion, we mention that, for $X_1 \geq X_2 = 3$ and $\Theta \geq \frac{1}{3}\Pi$, Eq. (28) gives an accuracy of 5×10^{-5} with four terms.

EVALUATION OF THE INTEGRALS A TO C

In this paper we have shown that the evaluation of the second-order three-atom or three-ion energy of interaction can be reduced to that of calculating a number of volume integrals [(2)–(8)]. These integrals can all be expressed in terms of three basic integrals A to C [(10)–(16)]. Two of these last three integrals, A and B, can be obtained from tabulated values of the error function[2] and of Dawson's integral[3,4]

$$e^{-x^2}\int_0^x e^{t^2}\,dt,$$

respectively. On the other hand, values of the integral C [Eq. (9)] are not available, but these can be readily obtained by electronic computation.

It should be noted that, compared with our previous analysis of second-order three-body interactions,[1,5] the present results constitute a considerable simplification. In particular, Eqs. (10)–(16) are valid for the whole range of triangular dimensions, whereas the asymptotic expansions were only applicable in the range of large dimensions (in terms of the Gaussian parameter β and the nearest-neighbor distance R in a solid, this range of validity is $\beta R \geq 2$). The latter limiting case applies for rare-gas crystals. In ionic crystals, however, βR is considerably smaller because of electrostatic compression.[6]

ACKNOWLEDGMENT

The author is indebted to Professor L. Jansen for his help and encouragement in this work.

[2] National Bureau of Standards, *Tables of the Error Function and its Derivative*, Applied Math. Series 41 (U.S. Government Printing Office, Washington, D.C., 1954), 2nd ed.

[3] B. Lohmander and S. Rittsten, Kungl. Fysiogr. Sällsk. i Lund Förh. **28**, 45 (1958).

[4] W. Lash Miller and A. R. Gordon, J. Phys. Chem. **35**, 2874 (1931).

[5] L. Jansen, Phys. Rev. **135**, A1292 (1964).

[6] After this manuscript had been submitted for publication, an article by P. J. Roberts, Proc. Phys. Soc. (London) **89**, 67 (1966) came to the author's attention. This article deals with the same subject, specifying to the case $\beta = \beta^1$ for the Gaussian parameters.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Asymptotic Behavior of Stieltjes Transforms. I

W. S. WOOLCOCK
Department of Physics, University College, London, England

(Received 21 June 1966)

Several theorems are proved concerning the asymptotic behavior of Stieltjes transforms as $|z|$ approaches infinity, in a sector of the complex z plane which does not include the cut in the transform. The asymptotic behavior of the transform is related to the asymptotic behavior, for large values of the argument, of the function whose transform is taken.

I. INTRODUCTION

STIELTJES transforms have become familiar to physicists because of the extensive use of dispersion relations in calculations concerning elementary particle interactions. Their asymptotic properties are important in discussing questions such as the number of subtractions required in dispersion relations. Some results concerning these properties are given in a paper of Lanz and Prosperi,[1] which includes a list of earlier papers containing some discussion of such asymptotic properties. Lanz and Prosperi concentrate on results in which a bound on the transform is derived from a bound on the original; results in which a precise asymptotic behavior of the transform can be established are mentioned only in passing.

In this paper we prove a number of results on the behavior of Stieltjes transforms as $|z|$ approaches infinity in a direction away from the cut in the transform. These results are, in fact, shown to hold uniformly in a sector in the complex plane which does not include the cut. The proof of theorems on the asymptotic behavior of principal value integrals and their extension to hold uniformly in a sector which includes the cut requires further conditions on the

[1] L. Lanz and G. M. Prosperi, Nuovo Cimento **33**, 201 (1964).

To indicate the accuracy of the third asymptotic expansion, we mention that, for $X_1 \geq X_2 = 3$ and $\Theta \geq \frac{1}{3}\Pi$, Eq. (28) gives an accuracy of 5×10^{-5} with four terms.

EVALUATION OF THE INTEGRALS A TO C

In this paper we have shown that the evaluation of the second-order three-atom or three-ion energy of interaction can be reduced to that of calculating a number of volume integrals [(2)–(8)]. These integrals can all be expressed in terms of three basic integrals A to C [(10)–(16)]. Two of these last three integrals, A and B, can be obtained from tabulated values of the error function[2] and of Dawson's integral[3,4]

$$e^{-x^2}\int_0^x e^{t^2}\,dt,$$

respectively. On the other hand, values of the integral C [Eq. (9)] are not available, but these can be readily obtained by electronic computation.

It should be noted that, compared with our previous analysis of second-order three-body interactions,[1,5] the present results constitute a considerable simplification. In particular, Eqs. (10)–(16) are valid for the whole range of triangular dimensions, whereas the asymptotic expansions were only applicable in the range of large dimensions (in terms of the Gaussian parameter β and the nearest-neighbor distance R in a solid, this range of validity is $\beta R \geq 2$). The latter limiting case applies for rare-gas crystals. In ionic crystals, however, βR is considerably smaller because of electrostatic compression.[6]

ACKNOWLEDGMENT

The author is indebted to Professor L. Jansen for his help and encouragement in this work.

[2] National Bureau of Standards, *Tables of the Error Function and its Derivative*, Applied Math. Series 41 (U.S. Government Printing Office, Washington, D.C., 1954), 2nd ed.

[3] B. Lohmander and S. Rittsten, Kungl. Fysiogr. Sällsk. i Lund Förh. **28**, 45 (1958).

[4] W. Lash Miller and A. R. Gordon, J. Phys. Chem. **35**, 2874 (1931).

[5] L. Jansen, Phys. Rev. **135**, A1292 (1964).

[6] After this manuscript had been submitted for publication, an article by P. J. Roberts, Proc. Phys. Soc. (London) **89**, 67 (1966) came to the author's attention. This article deals with the same subject, specifying to the case $\beta = \beta^1$ for the Gaussian parameters.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Asymptotic Behavior of Stieltjes Transforms. I

W. S. WOOLCOCK
Department of Physics, University College, London, England

(Received 21 June 1966)

Several theorems are proved concerning the asymptotic behavior of Stieltjes transforms as $|z|$ approaches infinity, in a sector of the complex z plane which does not include the cut in the transform. The asymptotic behavior of the transform is related to the asymptotic behavior, for large values of the argument, of the function whose transform is taken.

I. INTRODUCTION

STIELTJES transforms have become familiar to physicists because of the extensive use of dispersion relations in calculations concerning elementary particle interactions. Their asymptotic properties are important in discussing questions such as the number of subtractions required in dispersion relations. Some results concerning these properties are given in a paper of Lanz and Prosperi,[1] which includes a list of earlier papers containing some discussion of such asymptotic properties. Lanz and Prosperi concentrate on results in which a bound on the transform is derived from a bound on the original; results in which a precise asymptotic behavior of the transform can be established are mentioned only in passing.

In this paper we prove a number of results on the behavior of Stieltjes transforms as $|z|$ approaches infinity in a direction away from the cut in the transform. These results are, in fact, shown to hold uniformly in a sector in the complex plane which does not include the cut. The proof of theorems on the asymptotic behavior of principal value integrals and their extension to hold uniformly in a sector which includes the cut requires further conditions on the

[1] L. Lanz and G. M. Prosperi, Nuovo Cimento **33**, 201 (1964).

original function and involves quite different techniques from those used in this paper. It is appropriate, therefore, to discuss these latter results in a separate paper.

In Sec. II we establish our notation and write down conditions on the original function which will be assumed to hold for all the theorems and corollaries to be proved. We also make some preliminary transformations and state two theorems which are used in Sec. III. Section III itself contains the proof of several results on the precise asymptotic behavior of Stieltjes transforms. Section IV repeats a result of Lanz and Prosperi; it is included here for completeness.

The aim has been to put down all the results on the asymptotic behavior of Stieltjes transforms that we have been able to discover, in the hope that it may be of value to have these collected together. Of the results in Sec. III, the first theorem is an extension of a well-known result which is given, for example, by Widder.[2] The main preliminary transformation in Sec. II follows Widder's proof of this result. We have not been able to find the other theorems and corollaries in Sec. III in the mathematical literature.

II. PRELIMINARIES

Let $g(x)$ be a real valued function defined for $x \geq 0$, and let $g(x)$ belong to $L([a, b])$ for any choice of a, b with $0 < a < b$. Let the limits

$$\int_{\to 0}^{\to \infty} \frac{g(x)\,dx}{x} \quad \text{and} \quad \int_{\to 0} g(x)\,dx$$

exist. (The notation is that of Titchmarsh.[3])

The *Stieltjes transform* $f(z)$ of the function $g(x)$ is defined by

$$f(z) = \int_{\to 0}^{\to \infty} \frac{g(t)\,dt}{t + z}$$

for $z \neq 0$, $\operatorname{Arg} z \neq \pi$. For each such z, the manipulations carried out in this section show that the integral defining $f(z)$ does indeed converge at both limits. (For every $z \neq 0$, $\operatorname{Arg} z$ is unique, and $-\pi < \operatorname{Arg} z \leq \pi$.) The function $f(z)$ is an analytic function, regular in the whole complex plane cut along the negative real axis. We are going to study the behavior of $f(z)$ as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$. Uniform convergence properties are established in a sector $|\operatorname{Arg} z| \leq \pi - \delta$, where $0 < \delta < \pi$. This sector is henceforth denoted by S_δ.

We dispose of the complication at the lower end of the range of integration by writing

$$f(z) = f_1(z) + \int_{\to 0}^{1} \frac{g(t)\,dt}{t + z},$$

where

$$f_1(z) = \int_{1}^{\to \infty} \frac{g(t)\,dt}{t + z}.$$

Now

$$z\int_{\to 0}^{1} \frac{g(t)\,dt}{t + z} - \int_{\to 0}^{1} g(t)\,dt = -\int_{\to 0}^{1} \frac{tg(t)\,dt}{t + z}.$$

Defining the continuous function $\phi(x)$ by

$$\phi(x) = \int_{\to 0}^{x} g(t)\,dt \quad (x > 0)$$

and applying the formula for integration by parts[4] to the interval $[\alpha, 1]$, with $0 < \alpha < 1$, we have

$$\int_{\alpha}^{1} \frac{tg(t)\,dt}{t + z} = \frac{t\phi(t)}{t + z}\Bigg|_{\alpha}^{1} - z\int_{\alpha}^{1} \frac{\phi(t)\,dt}{(t + z)^2}.$$

Letting $\alpha \to 0$,

$$\int_{\to 0}^{1} \frac{tg(t)\,dt}{t + z} = \frac{\phi(1)}{1 + z} - z\int_{0}^{1} \frac{\phi(t)\,dt}{(t + z)^2}.$$

The integral on the right side exists in the ordinary Riemann sense. Moreover, for $z \in S_\delta$ and fixed $r \neq 0$,

$$\left| z\int_{0}^{1} \frac{\phi(t)\,dt}{(t + z)^2} \right| \leq r\int_{0}^{1} \frac{|\phi(t)|\,dt}{(t - r\cos\delta)^2 + r^2\sin^2\delta} \leq \frac{1}{r\sin^2\delta}\int_{0}^{1} |\phi(t)|\,dt.$$

It follows that

$$\int_{\to 0}^{1} \frac{g(t)\,dt}{t + z} = \int_{\to 0}^{1} g(t)\,dt/z + o(1/|z|)$$

as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$. This result means that the theorems of Sec. III need be proved only for $f_1(z)$.

Next we write $f_1(z)$ in a different form. Defining the continuous function $\psi(x)$ by

$$\psi(x) = \int_{x}^{\to \infty} \frac{g(t)\,dt}{t} \quad (x \geq 1)$$

and applying the formula for integration by parts to the interval $[1, X]$, with $X > 1$, we have

$$\int_{1}^{X} \frac{g(t)\,dt}{t + z} = -\frac{t\psi(t)}{t + z}\Bigg|_{1}^{X} + z\int_{1}^{X} \frac{\psi(t)\,dt}{(t + z)^2}.$$

[2] D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, N.J., 1941), Chap. VIII, Theorem 3d, p. 333. In the notation of Sec. II, Widder's result states that $f(x) \to 0$ as $x \to \infty$ through real values.

[3] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford University Press, New York, 1948), 2nd ed., pp. 9–10.

[4] This holds for absolutely continuous functions and Lebesgue integrals. See, for example, Theorem (18.19) of E. Hewitt and K. Stromberg, *Real and Abstract Analysis* (Springer-Verlag, Berlin, 1965).

Now let $X \to \infty$. Since $\psi(x) \to 0$ as $x \to \infty$, the integral

$$\int_1^\infty \frac{\psi(t)\,dt}{(t+z)^2}$$

is absolutely convergent and

$$f_1(z) = \frac{\psi(1)}{1+z} + z\int_1^\infty \frac{\psi(t)\,dt}{(t+z)^2}.$$

This form for $f_1(z)$ is very convenient for proving the theorems of the next section.

We now give statements of two theorems which we use in Sec. III.

Theorem A: Let the complex valued function $h(t, r)$ of the real variables t, r, defined for all $t \geq t_0$, $r \geq r_0$, belong to $L([t_0, T])$ for each $T > t_0$ and each $r \geq r_0$, and let the limit

$$\int_{t_0}^{\to\infty} h(t, r)\,dt$$

exist for each $r \geq r_0$.

Suppose that, as $r \to \infty$,

$$\lim_{r\to\infty} h(t, r)$$

exists for each $t \geq t_0$; denote this limit function by $h_1(t)$. Let the convergence of $h(t, r)$ to $h_1(t)$ be uniform in each interval $[t_0, T]$ with $T > t_0$.

Then the limits

$$\int_{t_0}^{\to\infty} h_1(t)\,dt$$

and

$$\lim_{r\to\infty} \int_{t_0}^{\to\infty} h(t, r)\,dt$$

exist and are equal if and only if, given $\epsilon > 0$, there exists T (depending on ϵ) such that, for each $t > T$, there is a number R (depending on t) for which

$$\left|\int_t^{\to\infty} h(u, r)\,du\right| < \epsilon \quad \text{for all} \quad r > R.$$

This theorem is a simple application of a theorem on repeated limits given in Sec. 305 of Vol. I of Hobson.[5] We do not use the full power of this theorem; in particular we are always able to find a number R which is independent of t.

Theorem B: Let $h(u, r)$ be a complex valued function of the real variables u, r, defined for all $u \geq u_0$, $r \geq r_0$ and belonging to $L([u_0, \infty])$ for each $r \geq r_0$. Let

$$\lim_{r\to\infty} h(u, r)$$

exist almost everywhere on $[u_0, \infty]$. Suppose that there exists a (real valued) function $M(u)$ belonging to $L([u_0, \infty])$ such that $|h(u, r)| \leq M(u)$ almost everywhere on $[u_0, \infty]$ for each $r \geq r_0$. Then

$$\lim_{r\to\infty} h(u, r)$$

belongs to $L([u_0, \infty])$ and

$$\lim_{r\to\infty} \int_{u_0}^\infty h(u, r)\,du = \int_{u_0}^\infty \lim_{r\to\infty} h(u, r)\,du.$$

This is just Lebesgue's dominated convergence theorem,[6] written in a way which can be applied immediately in Sec. III.

III. PRECISE ASYMPTOTIC BEHAVIOR

We turn now to the main results of this paper and give a sequence of three theorems. It is assumed that the conditions in the first paragraph of Sec. II hold throughout.

Theorem 1: The transform $f(z) \to 0$ as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$.

Proof. It suffices to prove that

$$z\int_1^\infty \frac{\psi(t)\,dt}{(t+z)^2} \to 0$$

under the conditions just stated. Now, for $z \in S_\delta$ and fixed $r \neq 0$, we have

$$\left|z\int_1^\infty \frac{\psi(t)\,dt}{(t+z)^2}\right| \leq r\int_1^\infty \frac{|\psi(t)|\,dt}{(t - r\cos\delta)^2 + r^2\sin^2\delta}.$$

Theorem A may be applied immediately, with

$$h(t, r) = \frac{r\,|\psi(t)|}{(t - r\cos\delta)^2 + r^2\sin^2\delta} \qquad (t \geq 1, r \geq 1 \text{ say}).$$

As $r \to \infty$, $h(t, r) \to 0$ for each $t \geq 1$. Moreover, this convergence is uniform in each interval $[1, T]$ with $T > 1$, since, for $t \in [1, T]$,

$$h(t, r) \leq m(T)/r\sin^2\delta,$$

where $m(T) = \max\{|\psi(t)| : t \in [1, T]\}$.

Finally, given $\epsilon > 0$, we can find T such that $|\psi(t)| < \epsilon \sin\delta/(\pi - \delta)$ for all $t \geq T$, since $\psi(t) \to 0$ as $t \to \infty$. Hence, for $t \geq T$, we have

$$\int_t^\infty h(u, r)\,du < \frac{\epsilon\sin\delta}{\pi - \delta} r\int_0^\infty \frac{du}{(u - r\cos\delta)^2 + r^2\sin^2\delta} = \epsilon,$$

[5] E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* (Cambridge University Press, New York, 1927), 3rd ed., Vol. I.

[6] See, for example, Theorem (12.30) of Hewitt and Stromberg Ref. 4.

independently of r. The conditions of Theorem A are therefore satisfied and the result follows.

An alternative proof[7] using Theorem B requires the transformation $t = ru$. Then

$$r\int_1^\infty \frac{|\psi(t)|\,dt}{(t - r\cos\delta)^2 + r^2\sin^2\delta} = \int_0^\infty \frac{\chi_{[1/r,\infty]}(u)\,|\psi(ru)|\,du}{(u - \cos\delta)^2 + \sin^2\delta},$$

where $\chi_{[1/r,\infty]}(u)$ is the characteristic function of the interval $[1/r, \infty]$. Now with

$$h(u, r) = \frac{\chi_{[1/r,\infty]}(u)\,|\psi(ru)|}{(u - \cos\delta)^2 + \sin^2\delta} \quad (u \geq 0, r \geq 1 \text{ say}),$$

$$M(u) = M/[(u - \cos\delta)^2 + \sin^2\delta],$$

where $M = \max\{|\psi(t)| : t \geq 1\}$, the conditions of Theorem B are satisfied, since, for each $u > 0$, $\psi(ru)$ [and therefore $h(u, r)$] $\to 0$ as $r \to \infty$, and clearly $M(u)$ belongs to $L([0, \infty])$. Hence the result.

Corollary: Let the limit $\int^{\to\infty} g(t)\,dt$ exist. Then

$$zf(z) \to \int_{\to 0}^{\to\infty} g(t)\,dt,$$

as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$.

Proof.

$$zf(z) - \int_{\to 0}^{\to\infty} g(t)\,dt = -\int_{\to 0}^{\to\infty} \frac{tg(t)\,dt}{t + z}.$$

Apply Theorem 1 to the Stieltjes transform on the right side.

Theorem 2: Let *either* of the following conditions hold:

(i) the limit $\int^{\to\infty} \frac{g(x)\,dx}{x^{1-\alpha}}$ exists;

(ii) $x^\alpha g(x) \to 0$ as $x \to \infty$,

where $0 < \alpha < 1$. Then $z^\alpha f(z) \to 0$, as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$.

[*Notes:* (1) The alternative conditions above are not comparable. For, if $x^\alpha g(x) = \sin x$, then

$$\int^{\to\infty} \frac{g(x)\,dx}{x^{1-\alpha}}$$

exists, but $x^\alpha g(x) \not\to 0$ as $x \to \infty$. On the other hand, if $x^\alpha g(x) = (\ln x)^{-1}$, then

$$\int^{\to\infty} \frac{g(x)\,dx}{x^{1-\alpha}}$$

does not exist, but $x^\alpha g(x) \to 0$ as $x \to \infty$.

(2) Theorem 1 extends Theorem 2 under condition (i) to the case $\alpha = 0$, while the corollary to Theorem 1 extends it to the case $\alpha = 1$.]

Proof. The proof amounts to showing that

$$z^{\alpha+1}\int_1^\infty \frac{\psi(t)\,dt}{(t + z)^2} \to 0$$

under the conditions stated above. We first prove that $x^\alpha\psi(x) \to 0$ as $x \to \infty$ under each of the conditions (i), (ii) in turn.

To use condition (i), write

$$\psi(x) = \int_x^{\to\infty} \frac{g(t)}{t^{1-\alpha}}\frac{1}{t^\alpha}\,dt$$

and apply the second mean value theorem for integrals. (See, for example, Sec. 422 of Hobson.[5]) This gives

$$\int_x^X \frac{g(t)}{t^{1-\alpha}}\frac{1}{t^\alpha}\,dt = \frac{1}{x^\alpha}\int_x^\xi \frac{g(t)\,dt}{t^{1-\alpha}},$$

where ξ is some number, depending on X, such that $x \leq \xi \leq X$. Letting $X \to \infty$ and then $x \to \infty$, it follows from the existence of the limit

$$\int^{\to\infty} \frac{g(t)\,dt}{t^{1-\alpha}}$$

that $x^\alpha\psi(x) \to 0$ as $x \to \infty$. It is obvious that this result holds under condition (ii).

For $z \in S_\delta$ and fixed $r \neq 0$, we have

$$\left|z^{\alpha+1}\int_1^\infty \frac{\psi(t)\,dt}{(t + z)^2}\right| \leq r^{\alpha+1}\int_1^\infty \frac{|\psi(t)|\,dt}{(t - r\cos\delta)^2 + r^2\sin^2\delta}.$$

Again we may apply Theorem A with

$$h(t, r) = \frac{r^{\alpha+1}\,|\psi(t)|}{(t - r\cos\delta)^2 + r^2\sin^2\delta} \quad (t \geq 1, r \geq 1 \text{ say}).$$

Since $\alpha < 1$, $h(t, r) \to 0$ as $r \to \infty$, for each $t \geq 1$. This convergence is uniform in each interval $[1, T]$ with $T > 1$ since, for $t \in [1, T]$,

$$h(t, r) \leq m(T)/r^{1-\alpha}\sin^2\delta.$$

Finally, given $\epsilon > 0$, we can find T such that

$$|t^\alpha\psi(t)| < \frac{\epsilon\sin\delta\sin\alpha\pi}{\pi\sin[\alpha(\pi - \delta)]}$$

[7] The author is indebted to H. Kestleman for pointing out this alternative method of proof.

for $t \geq T$. Thus for $t \geq T$, we have

$$\int_t^\infty h(u, r)\, du < \frac{\epsilon \sin \delta \sin \alpha\pi}{\pi \sin [\alpha(\pi - \delta)]}\, r^{\alpha+1} \times \int_0^\infty \frac{du}{u^\alpha[(u - r\cos\delta)^2 + r^2 \sin^2 \delta]} = \epsilon,$$

independently of r. The result follows.

To apply Theorem B, put $t = ru$ as before. Then

$$r^{\alpha+1}\int_1^\infty \frac{\psi(t)\, dt}{(t - r\cos\delta)^2 + r^2 \sin^2 \delta} = \int_0^\infty \frac{\chi_{[1/r,\infty]}(u)(ru)^\alpha\, |\psi(ru)|\, du}{u^\alpha[(u - \cos\delta)^2 + \sin^2 \delta]}.$$

With

$$h(u, r) = \frac{\chi_{[1/r,\infty]}(u)(ru)^\alpha\, |\psi(ru)|}{u^\alpha[(u - \cos\delta)^2 + \sin^2 \delta]} \qquad (u \geq 0, r \geq 1 \text{ say}),$$

$$M(u) = \frac{M}{u^\alpha[(u - \cos\delta)^2 + \sin^2 \delta]},$$

where $M = \max\{t^\alpha\, |\psi(t)| : t \geq 1\}$, the conditions of Theorem B are satisfied, since, for each $u > 0$, $(ru)^\alpha\psi(ru)$ [and therefore $h(u, r)$] $\to 0$ as $r \to \infty$, and $M(u)$ belongs to $L([0, \infty])$. This proves the theorem.

Corollary: Let either of the following conditions hold:

(i) the limit $\int^{\to\infty} \frac{(x^\alpha g(x) - A)\, dx}{x}$ exists;

(ii) for large x, $g(x) \sim Ax^{-\alpha}$,

where A, α are constants and $0 < \alpha < 1$. Then, for large $|z|$,

$$f(z) \sim A\pi \csc(\pi\alpha) z^{-\alpha},$$

for fixed $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ. (The function z^α has its principal value in the complex plane cut along the negative real axis.)

Proof. Put $(x^\alpha g(x) - A) = x^\alpha h(x)$. Then the conditions (i), (ii) above are equivalent to the conditions (i), (ii) of Theorem 2 on the function $h(x)$. Now

$$f(z) = A\int_{\to 0}^{\to\infty} \frac{dt}{t^\alpha(t + z)} + \int_{\to 0}^{\to\infty} \frac{h(t)\, dt}{t + z}.$$

The first term on the right side is $A\pi \csc(\pi\alpha) z^{-\alpha}$. Applying Theorem 2, we see that

$$z^\alpha \int_{\to 0}^{\to\infty} \frac{h(t)\, dt}{t + z} \to 0,$$

as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$. Under these conditions, therefore, $z^\alpha f(z) \to A\pi \csc \pi\alpha$, which is the required result.

Finally we extend Theorem 2 under condition (ii) to $\alpha = 1$.

Theorem 3: Let $xg(x) \to 0$ as $x \to \infty$. Then $zf(z)/\log z \to 0$, as $|z| \to \infty$ in any direction for which $\operatorname{Arg} z \neq \pi$, and uniformly for z in S_δ as $|z| \to \infty$. (The function $\log z$ denotes the principal value of $\log z$ in the complex plane cut along the negative real axis.)

Proof: It is enough to show that

$$\frac{z^2}{\log z}\int_1^\infty \frac{\psi(t)\, dt}{(t + z)^2} \to 0$$

under the conditions stated. In exactly the same way as in Theorem 2 it follows from the condition on $g(x)$ that $x\psi(x) \to 0$ as $x \to \infty$.

For $z \in S_\delta$ and fixed $r > 1$, we have

$$\left|\frac{z^2}{\log z}\int_1^\infty \frac{\psi(t)\, dt}{(t + z)^2}\right| \leq \frac{r^2}{\log r}\int_1^\infty \frac{|\psi(t)|\, dt}{(t - r\cos\delta)^2 + r^2 \sin^2 \delta}.$$

We apply Theorem A once more, with

$$h(t, r) = \frac{r^2\, |\psi(t)|}{\log r[(t - r\cos\delta)^2 + r^2 \sin^2 \delta]} \qquad (t \geq 1, r \geq 2 \text{ say}).$$

Again $h(t, r) \to 0$ as $r \to \infty$, for each $t \geq 1$. The convergence is uniform in each interval $[1, T]$ with $T > 1$ since, for $t \in [1, T]$,

$$h(t, r) \leq m(T)/\log r \sin^2 \delta.$$

Also, given $\epsilon > 0$, we can find T such that $|t\psi(t)| < \frac{1}{2}\epsilon$ for $t \geq T$. Hence, for $t \geq T$, we have

$$\int_t^\infty h(u, r)\, du < \frac{\epsilon}{2}\frac{r^2}{\log r}\int_1^\infty \frac{dt}{t[(t - r\cos\delta)^2 + r^2 \sin^2 \delta]} \leq \epsilon\left[\frac{1}{2} + \frac{1}{2\log r}\left\{\log\left(1 + \frac{1}{r}\right) + \frac{\pi}{\sin\delta}\right\}\right] < \epsilon \text{ for all sufficiently large } r.$$

The conditions of Theorem A are satisfied and the result follows.

Theorem B cannot be applied in this case, but, by making the transformation $t = ru$ once more, an alternative proof can be found. We need to show that

$$\frac{1}{\log r}\int_{1/r}^\infty \frac{(ru)\, |\psi(ru)|\, du}{u[(u - \cos\delta)^2 + \sin^2 \delta]} \to 0 \quad \text{as} \quad r \to \infty.$$

Let $M = \max\{t\, |\psi(t)| : t \geq 1\}$. Since

$$\int_1^\infty \frac{(ru)\, |\psi(ru)|\, du}{u[(u - \cos\delta)^2 + \sin^2 \delta]} \leq M\int_1^\infty \frac{du}{u[(u - \cos\delta)^2 + \sin^2 \delta]},$$

a constant, it is sufficient to show that

$$\frac{1}{\log r}\int_{1/r}^{1}\frac{(ru)\,|\psi(ru)|\,du}{u[(u-\cos\delta)^2+\sin^2\delta]}\to 0 \quad \text{as} \quad r\to\infty.$$

Now, given $\epsilon>0$, there exists T (≥ 1) such that $t\,|\psi(t)|<\frac{1}{2}\epsilon\sin^2\delta$ for $t\geq T$. Thus, for $r>T$,

$$\begin{aligned}\left|\frac{1}{\log r}\int_{1/r}^{1}\frac{(ru)\,|\psi(ru)|\,du}{u[(u-\cos\delta)^2+\sin^2\delta]}\right| &\leq \frac{1}{\log r\sin^2\delta}\left[M\int_{1/r}^{T/r}\frac{du}{u}+\tfrac{1}{2}\epsilon\sin^2\delta\int_{T/r}^{1}\frac{du}{u}\right]\\ &= \frac{1}{\log r\sin^2\delta}\left(M\log T+\tfrac{1}{2}\epsilon\sin^2\delta\log\frac{r}{T}\right)\\ &\leq \epsilon\left(\frac{1}{2}+\frac{M\log T}{\epsilon\sin^2\delta\log r}\right)\\ &< \epsilon \text{ for all sufficiently large } r.\end{aligned}$$

Corollary: Suppose that, for large x, $g(x)\sim A/x$, where A is a constant. Then, for large $|z|$,

$$f(z)\sim A\log z/z,$$

for fixed $\operatorname{Arg} z\neq\pi$, and uniformly for z in S_δ.

Proof: If we put $(xg(x)-A)=xh(x)$, the function $h(x)$ satisfies the condition of Theorem 3. Now

$$\begin{aligned}f_1(z) &= A\int_1^{\to\infty}\frac{dt}{t(t+z)}+\int_1^{\to\infty}\frac{h(t)\,dt}{t+z}\\ &= \frac{A}{z}\log(1+z)+\int_1^{\to\infty}\frac{h(t)\,dt}{t+z}.\end{aligned}$$

Thus

$$\frac{zf_1(z)}{\log z}=A\frac{\log(1+z)}{\log z}+\frac{z}{\log z}\int_1^{\to\infty}\frac{h(t)\,dt}{t+z}.$$

Since, from Sec. II,

$$\frac{z}{\log z}\int_{\to 0}^{1}\frac{g(t)\,dt}{t+z}\to 0,$$

it follows from Theorem 3 that $zf(z)/\log z\to A$, as $|z|\to\infty$ in any direction for which $\operatorname{Arg} z\neq\pi$, and uniformly for z in S_δ as $|z|\to\infty$.

We conclude this section by noting that, if $g(x)=h(x)x^{-p}$, where p is a positive integer and $h(x)$ satisfies the conditions of any of the theorems (or their corollaries) given above, one can apply these results by writing

$$\begin{aligned}f_1(z)=\int_1^{\to\infty}\frac{h(t)\,dt}{t^p(t+z)} &= \sum_{m=1}^{p}(-1)^{m-1}z^{-m}\int_1^{\to\infty}\frac{h(t)\,dt}{t^{p-m+1}}\\ &\quad+(-1)^pz^{-p}\int_1^{\to\infty}\frac{h(t)\,dt}{t+z}.\end{aligned}$$

IV. BOUNDS ON TRANSFORMS

Our final theorem is due to Lanz and Prosperi[1]; in it a bound on $f(z)$ for large $|z|$ is derived from a bound on $g(x)$ for large x.

Theorem 4: Let $|g(x)|<A/x^\alpha$ for all $x\geq X$ (>0), where $0<\alpha\leq 1$. Then

$$\begin{aligned}|f(z)| &< C/|z|^\alpha \qquad (0<\alpha<1),\\ f(z) &< C'\log|z|/|z| \quad (\alpha=1),\end{aligned}$$

for all sufficiently large $|z|$, for fixed $\operatorname{Arg} z\neq\pi$ and uniformly for z in S_δ.

[*Note:* For $\alpha>1$, $\int^{\to\infty}g(x)\,dx$ converges and the corollary to Theorem 1 applies.]

Proof: In each case, write

$$f(z)=\int_{\to 0}^{X}\frac{g(t)\,dt}{t+z}+\int_X^{\to\infty}\frac{g(t)\,dt}{t+z}.$$

Just as in Sec. II,

$$\int_{\to 0}^{X}\frac{g(t)\,dt}{t+z}\sim\int_{\to 0}^{X}\frac{g(t)\,dt}{z}$$

for large $|z|$. We need deal, therefore, only with the second integral on the right side.

For $z\in S_\delta$ and fixed $r\neq 0$, we have

$$\begin{aligned}\left|\int_X^{\to\infty}\frac{g(t)\,dt}{t+z}\right| &\leq A\int_X^{\infty}\frac{dt}{t^\alpha[(t-r\cos\delta)^2+r^2\sin^2\delta]^{\frac{1}{2}}}\\ &= \frac{A}{r^\alpha}\int_{X/r}^{\infty}\frac{du}{u^\alpha[(u-\cos\delta)^2+\sin^2\delta]^{\frac{1}{2}}}.\end{aligned}$$

For $0<\alpha<1$,

$$\left|\int_X^{\to\infty}\frac{g(t)\,dt}{t+z}\right|<\frac{A}{r^\alpha}\int_0^{\infty}\frac{du}{u^\alpha[(u-\cos\delta)^2+\sin^2\delta]^{\frac{1}{2}}},$$

which is the required result.

For $\alpha=1$,

$$\begin{aligned}&\left|\int_X^{\to\infty}\frac{g(t)\,dt}{t+z}\right|\\ &\leq\frac{A}{r}\int_{X/r}^{\infty}\frac{du}{u[u^2-2u\cos\delta+1]^{\frac{1}{2}}}\\ &=\frac{A}{r}\log\left[\frac{X(1+\cos\delta)}{X\cos\delta-r+(X^2-2Xr\cos\delta+r^2)^{\frac{1}{2}}}\right]\\ &=\frac{A\log r}{r}+O\left(\frac{1}{r}\right)\quad\text{for } r\to\infty \text{ as required.}\end{aligned}$$

ACKNOWLEDGMENTS

The author is greatly indebted to H. Kestleman for several discussions, criticisms and helpful suggestions.

This work was supported in part by a grant from the Office of Aerospace Research (European Office), U.S. Air Force.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Probability Distribution of the Radius of Gyration of a Flexible Polymer

S. R. CORIELL
National Bureau of Standards, Washington, D.C.

J. L. JACKSON*
Department of Physics, Howard University, Washington, D.C.

(Received 10 November 1966)

Various aspects of the mathematics of the probability distribution $P_N(S_x)$ of one component of the square of the radius of gyration of an ideal Brownian chain with N units are presented. A rigorous expression for $P_N(S_x)$ in the form of a contour integral is obtained. The resulting integral is written in terms of Tchebichef polynomials. Several rigorous and approximate results are obtained for both the limiting distribution (N infinite) and for finite N.

I. INTRODUCTION

A SIMPLE and widely used model for the configuration of a large flexible polymer molecule in solution is the Brownian motion model. According to this model the position of the jth unit of the molecule relative to one of the ends is given by the sum of j steps in a Brownian motion, i.e.,

$$\mathbf{R}_j = \mathbf{r}_1 + \mathbf{r}_2 + \cdots \mathbf{r}_j, \tag{1}$$

with the $\mathbf{r}_j$ independent, identically distributed random variables. We confine our attention to a chain wherein the probability of the individual step is given by a spherically symmetric Gaussian distribution

$$P(x, y, z) = (2\pi\sigma^2)^{-\frac{3}{2}} \exp\left[-(x^2 + y^2 + z^2)/2\sigma^2\right]. \tag{2}$$

Although this is not as realistic as a model in which the steps are confined to the surface of a sphere or restricted to allowed angles, for many purposes the Gaussian model is adequate and it makes the mathematics more tractable.

An important measure of the size of the polymer is its radius of gyration S. The x component is given by[1,2]

$$\begin{aligned} S_x &= N^{-1}\sum_{j=1}^{N}\left(x_j - N^{-1}\sum_{k=1}^{N}x_k\right)^2 \\ &= N^{-1}\sum_{j=1}^{N}x_j^2 - N^{-2}\left(\sum_{j=1}^{N}x_j\right)^2, \end{aligned} \tag{3}$$

where x_j is the x component of $\mathbf{R}_j$, the vector from one end of the polymer to the jth unit, and N is the total number of elements in the chain. The probability distribution of S_x, S_y, and S_z is of importance in the light scattering, adsorption, expansion, and viscosity of polymers,[2] and there has been a great deal of current work on the calculation of the probability distribution.[3–8]

In this paper we present an extensive discussion of the mathematical properties of the probability distribution function of the radius of gyration $P_N(S_x)$. This includes its properties as a function of N for large but finite N as well as calculations relevant to the limit function as N goes to infinity. In the limit the distribution becomes a function of $S_x/N\sigma^2$ alone. The fact that this is true has the physical significance that the radius of gyration never behaves like a "macroscopic" variable no matter how large N becomes. Its distribution does not become more sharply peaked as the chain grows longer.

The mathematical problem is a quite complicated one. Even for the case $N \to \infty$, the result, which is obtained as a contour integral, is difficult to evaluate. For N finite, not only does the corresponding contour integral become more complicated, but there is the additional complexity attendant to the presence of yet another variable, N. The first paper to treat the problem was by Fixman.[3] Fixman obtained the limiting distribution for the three-dimensional radius of gyration ($S_x + S_y + S_z$), calculated its moments, and obtained asymptotic approximations. In a series of papers on the problem for finite but large N, Forsman and Hughes[5,6] and Forsman[7,8] presented the results of extensive investigations of the distribution function with particular emphasis on the function for very

* Consultant to the National Bureau of Standards.

[1] The variable S_x defined here is, of course, the x part of the square of what is customarily called the radius of gyration. To avoid the awkwardness of writing our variable as the square of another one or of introducing a new or longer name for S_x, we simply refer to S_x as the "radius of gyration."

[2] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, New York, 1953).

[3] M. Fixman, J. Chem. Phys. **36**, 306 (1962).

[4] M. Fixman, J. Chem. Phys. **36**, 3123 (1962).

[5] W. C. Forsman and R. E. Hughes, J. Chem. Phys. **38**, 2118 (1963).

[6] W. C. Forsman and R. E. Hughes, J. Chem. Phys. **38**, 2123 (1963).

[7] W. C. Forsman, J. Chem. Phys. **42**, 2829 (1965).

[8] W. C. Forsman, J. Chem. Phys. **44**, 1716 (1966).

small values of the radius of gyration, as this is related to the increase in entropy of a polymer when it is deposited on a surface.

We present a wide variety of results, both exact and approximate, relating to the function $P_N(S_x)$ and to the limiting distribution function for the radius of gyration divided by N in the limit as N approaches infinity. We obtain an integral expression for $P_N(S_x)$ which is exact for all N. At no point do we assume N to be large in the derivation. Working from the new form for $P_N(S_x)$, we present a broad survey of the mathematical properties of the function. Many of the results we obtain have been presented previously in the aforementioned papers.[3-8] Much, however, is new.

In Sec. II, we derive the integral expression for $P_N(S_x)$. The derivation involves the evaluation of a characteristic determinant, which is performed by a generating function technique. It is shown that the characteristic function of $P_N(S_x)$ may be written in terms of the Tchebichef polynomial of order $N-1$.

We have placed all of our exact results on the properties of $P_N(S_x)$ for finite N in Sec. III. This includes a derivation of an expression for $P_N(S_x)$ as a sum of real integrals (obtained by deforming the contour to go around the branch cuts of the characteristic function), an exact evaluation of the moments, and a derivation of the form of P_N in the limit as $N \to \infty$.

In Sec. IV, we use the method of saddle points on the finite N integral. We show that a single equation for the saddle point may be used to obtain simple approximations to P_N over various ranges of the values of the variable, S_x.

In Sec. V we obtain a form for P_∞ which is suitable for evaluation by quadrature. The results of the numerical integration are presented and compared with the asymptotic results for P_∞.

II. DERIVATION OF CHARACTERISTIC FUNCTION

According to Eq. (2), the probability of an x displacement, $x_j - x_{j-1}$, at the jth step is

$$P(x_j - x_{j-1}) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-(x_j - x_{j-1})^2/2\sigma^2\right]. \quad (4)$$

The distribution of S_x is then

$$P_N(S_x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_N \left\{ \delta\left[S_x - \frac{1}{N}\Sigma x_j^2 + \frac{1}{N^2}(\Sigma x_j)^2 \right] \times \frac{\exp\left[-x_1^2/2\sigma^2\right] \exp\left[-(x_2 - x_1)^2/2\sigma^2\right] \cdots \exp\left[-(x_N - x_{N-1})^2/2\sigma^2\right]}{(2\pi\sigma^2)^{N/2}} \right\}, \quad (5)$$

where δ is the Dirac delta function. Introducing the Fourier representation of the delta function,

$$P_N(S_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda' \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{dx_1 \cdots dx_N}{(2\pi\sigma^2)^{N/2}} \times \exp\left\{ i\lambda' \left[S_x - \frac{1}{N}\Sigma x_j^2 + \frac{1}{N^2}(\Sigma x_j)^2 \right] - \left[x_1^2/2\sigma^2\right] - \cdots - \left[(x_N - x_{N-1})^2/2\sigma^2\right] \right\}. \quad (6)$$

The x integrations yield

$$P_N(S_x) = \frac{1}{2\pi(2\pi\sigma^2)^{N/2}} \int_{-\infty}^{\infty} d\lambda' e^{i\lambda' x} \frac{\pi^{N/2}\sigma^N}{|D|^{\frac{1}{2}}} = \frac{1}{2\pi(2)^{N/2}} \int_{-\infty}^{\infty} d\lambda' \frac{e^{i\lambda' x}}{|D|^{\frac{1}{2}}}, \quad (7)$$

where $|D|$ is the determinant of the matrix D defined by

$$D_{kl} = -(i\lambda/N^2) + [1 + (i\lambda/N)]\delta_{kl} - \tfrac{1}{2}\delta_{kN}\delta_{lN} - \tfrac{1}{2}(\delta_{k,l+1} + \delta_{k,l-1}), \quad (8)$$

where $\lambda = \sigma^2\lambda'$ and δ_{kl} is the Kronecker delta.

We evaluate $|D|$ by a generating function method, and obtain the characteristic polynomial of D. In the Appendix we present an alternative method of evaluating $|D|$ by transforming the matrix D to a known form. Let $[C_1, C_2, \cdots, C_N]$ be an eigenvector of D with eigenvalue μ, i.e.,

$$\sum_{l=1}^{N} D_{kl}C_l = \mu C_k \quad (k = 1, \cdots, N). \quad (9)$$

We define the generating function $f(x)$ by

$$f(x) = \sum_{j=1}^{N} C_j x^j. \quad (10)$$

Multiplying Eq. (9) by x^k and summing over k, one obtains with Eq. (8),

$$f(x) = \frac{-\frac{i\lambda 2x}{N^2}\left(\sum_{j=1}^{N} C_j\right)\left(\sum_{j=1}^{N} x^j\right) - C_N x^{N+1} + C_N x^{N+2} + C_1 x}{x^2 - 2x[1 + (i\lambda/N) - \mu] + 1}. \quad (11)$$

For $x = 1$, the above equation yields

$$C_1 = 2\mu \sum_{j=1}^{N} C_j. \tag{12}$$

As the normalization of the eigenvectors is arbitrary, for simplicity we choose $\sum_{j=1}^{N} C_j = 1$. Then $C_1 = 2\mu$, and Eq. (11) becomes

$$f(x) = \frac{-\frac{2i\lambda}{N^2} x \sum_{j=1}^{N} x^j - C_N x^{N+1} + C_N x^{N+2} + 2\mu x}{x^2 - 2xt + 1}, \tag{13}$$

where $t = 1 + (i\lambda/N) - \mu$. We now make use of the fact that $(x^2 - 2xt + 1)^{-1}$ is the generating function for Tchebichef polynomials of the second kind, $U_j(t)$,[9] i.e.,

$$(x^2 - 2xt + 1)^{-1} = \sum_{j=0}^{\infty} U_j(t) x^j. \tag{14}$$

The Tchebichef polynomials are given by

$$U_n(t) = \sum_{m=0}^{[n/2]} \frac{(-1)^m (n-m)!\,(2t)^{n-2m}}{m!\,(n-2m)!}, \tag{15}$$

where $[n/2]$ is $n/2$ for n even and $(n-1)/2$ for n odd. Using Eq. (14), we write Eq. (13) as

$$\begin{aligned} f(x) &= \left\{ -(2i\lambda/N^2) \sum_{j=1}^{N} x^{j+1} - C_N x^{N+1} \right. \\ &\qquad \left. + C_N x^{N+2} + 2\mu x \right\} \sum_{j=0}^{\infty} U_j(t) x^j \\ &= \sum_{j=1}^{N} C_j x^j. \end{aligned} \tag{16}$$

Equating coefficients of x^m in Eq. (16), we obtain for $1 \leq m \leq N$

$$-(2i\lambda/N^2) \sum_{j=0}^{m-2} U_j(t) + 2\mu U_{m-1}(t) = C_m. \tag{17}$$

Summing the above equation over m from $m = 1$ to N and using the normalization condition for the C_m, $\sum_{m=1}^{N} C_m = 1$, yields

$$\begin{aligned} -1 - (2i\lambda/N^2) \sum_{m=0}^{N-2} (N - 1 - m) U_m(t) \\ + 2\mu \sum_{m=1}^{N} U_{m-1}(t) = 0. \end{aligned} \tag{18}$$

Recalling that $t = 1 + (i\lambda/N) - \mu$, we see that Eq. (18) is an Nth degree polynomial in μ for the N eigenvalues of the matrix D and is proportional to the characteristic polynomial of D. If Eq. (18) is then written in the form

$$\prod_{i=1}^{N} (d_i \mu - b_i) = 0, \tag{19}$$

then the eigenvalues of the matrix D are $\mu_i = b_i/d_i$. The determinant of D is

$$|D| = \prod_{i=1}^{N} \mu_i = \prod_{i=1}^{N} (b_i/d_i). \tag{20}$$

The product $\prod_{i=1}^{N} d_i$ is the coefficient of μ^N in Eq. (18). Thus, using Eq. (15), we get

$$\prod_{i=1}^{N} d_i = 2(-2)^{N-1} = (-1)^{N-1} 2^N. \tag{21}$$

The product $\prod_{i=1}^{N} (-b_i)$ is the constant in Eq. (18), and is given by

$$\begin{aligned} \prod_{i=1}^{N} (-b_i) &= -1 - (2i\lambda/N^2) \\ &\quad \times \sum_{m=0}^{N-2} (N - m - 1) U_m\left(1 + \frac{i\lambda}{N}\right). \end{aligned} \tag{22}$$

We use the identity

$$\sum_{m=0}^{N-2} (N - m - 1) U_m(t) = [N - U_{N-1}(t)]/[2(1 - t)]. \tag{23}$$

One may verify Eq. (23) with the help of the equation[9]

$$U_k(t) = \sin[(k+1) \arccos t]/\sin[\arccos t]. \tag{24}$$

If one then uses the formulas[10] for $\sum_{k=1}^{N} \sin k\theta$ and $\sum_{k=1}^{N} k \sin k\theta$, one obtains Eq. (23). Substituting from Eq. (23) into Eq. (22), we obtain

$$\prod_{i=1}^{N} (-b_i) = -\left(\frac{1}{N}\right) U_{N-1}\left(1 + \frac{i\lambda}{N}\right). \tag{25}$$

Thus, we find that

$$|D| = (N2^N)^{-1} U_{N-1}(1 + i\lambda/N). \tag{26}$$

Thus the distribution of the x component of the radius of gyration is

$$P_N(S_x) = N^{\frac{1}{2}} (2\pi)^{-1} \int_{-\infty}^{\infty} d\lambda' e^{i\lambda' S_x} \left[U_{N-1}\left(1 + \frac{i\lambda'\sigma^2}{N}\right) \right]^{-\frac{1}{2}}. \tag{27}$$

Using Eq. (24), this may also be written

$$P_N(S_x) = N^{\frac{1}{2}} (2\pi)^{-1} \int_{-\infty}^{\infty} d\lambda' e^{i\lambda' S_x} \left[\frac{(\sin\theta)}{(\sin N\theta)} \right]^{\frac{1}{2}}, \tag{28}$$

where

$$\theta = \arccos(1 + i\lambda'\sigma^2/N). \tag{29}$$

III. PROPERTIES OF THE DISTRIBUTION FUNCTION, $P_N(S_x)$

The distribution function $P_N(S_x)$ of the x component of the radius of gyration of an ideal Brownian chain of N steps is given by a complex integral in Eq. (27)

[9] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II.

[10] L. B. W. Jolley, *Summation of Series* (Dover Publications, Inc., New York, 1961).

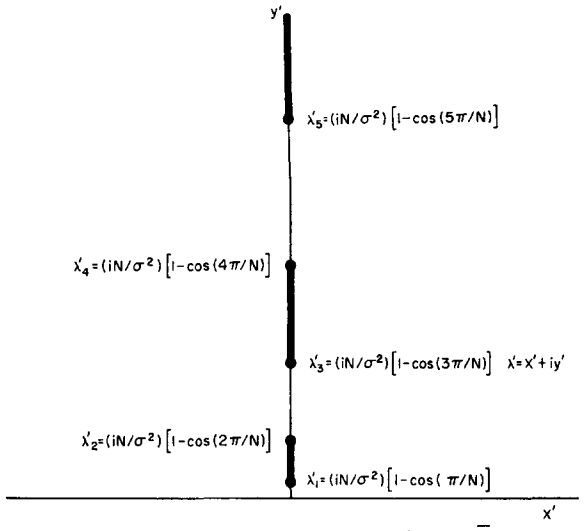


FIG. 1. Branch cuts for $[U_{N-1}(1 + i\lambda'\sigma^2/N)]^{-\frac{1}{2}}$.

or (28). This section is devoted to a variety of the properties of the finite N solution. In particular we first locate the branch points, specify branch cuts, and obtain a real integral for $P_N(S_x)$ by integrating along the branch cuts. We also use the integral to calculate exactly the moments of the distribution. Finally we show how one may obtain an integral for the radius of gyration in the limit as N becomes infinite, i.e.,

$$\lim_{N\to\infty} P_N\left(\frac{S_x}{N}\right).$$

A. Analytic Properties

The integral of Eq. (28) has a branch point at each of the zeros of $\sin N\theta$ except those for which $\sin\theta$, too, is zero. That is, there are branch points at

$$\theta_k = k\pi/N \quad (k = 1, \cdots N-1). \tag{30}$$

Therefore, from Eq. (29), the branch points in the λ' plane are

$$\lambda'_k = (iN/\sigma^2)(1 - \cos k\pi/N). \tag{31}$$

We see from this equation that values of k other than those specified to the right of Eq. (30) would not yield any additional value of the λ'_k. As we see all of the branch points are on the positive imaginary axis. We choose to make branch cuts between λ'_{2j-1} and λ'_{2j} with $j = 1, 2, \cdots \frac{1}{2}(N-1)$, for N odd or $j = 1, 2, \cdots \frac{1}{2}(N-2)$, for N even. For N even, there is an additional branch cut along the imaginary axis from λ'_{N-1} to infinity. The branch cuts are illustrated in Fig. 1.

The fact that there are no singularities in the lower half-plane means that $P_N(S_x)$ is identically zero[11] for S_x negative—as, of course, it must be since S_x by definition is a positive variable.

B. Integral Over Branch Cuts

We now obtain an expression for $P_N(S_x)$ in terms of definite integrals of real functions. This is done by showing that the integral over the real axis [Eq. (27)] is equal to integrals around the branch cuts in Fig. 1. As illustrated in Fig. 2, to the integral over the real axis we add: (1) the quarter circle of radius $R(R\to\infty)$ from $(R, 0)$ to $(|\epsilon|, [R^2 - \epsilon^2]^{\frac{1}{2}})$ $(|\epsilon|\to 0$ and $\lambda' = x' + iy')$; (2) the line $x' = |\epsilon|$ from R to 0; (3) the line $y' = 0$ from $|\epsilon|$ to $-|\epsilon|$; (4) the line $x' = -|\epsilon|$ from 0 to R; (5) the quarter circle of radius R from $(-|\epsilon|, [R^2 - \epsilon^2]^{\frac{1}{2}})$ to $(-R, 0)$. This contour does not enclose any singularities and hence the integral over the total contour vanishes. The integrals over the quarter circles (paths 1 and 5) vanish as $R\to\infty$, since the real part of the exponent is negative in the upper half plane.[11]

As $\epsilon\to 0$, the integrals on either side of the positive imaginary axis cancel each other everywhere expect on opposite sides of a branch cut. Recalling that the first branch cut went from branch point λ'_1 to λ'_2, the second from λ'_3 to λ'_4, etc., and introducing the definition $\lambda'_N = i\infty$ for N even, we have

$$P_N(S_x) = \frac{N^{\frac{1}{2}}}{2\pi}\lim_{\epsilon\to 0}\sum_{j=1}^{[N/2]}\left\{\int_{|\epsilon|+\lambda'_{2j-1}}^{|\epsilon|+\lambda'_{2j}} d\lambda' e^{i\lambda' S_x}\left[\frac{(\sin\theta)}{(\sin N\theta)}\right]^{\frac{1}{2}} - \int_{-|\epsilon|+\lambda'_{2j-1}}^{-|\epsilon|+\lambda'_{2j}} d\lambda' e^{i\lambda' S_x}\left[\frac{(\sin\theta)}{(\sin N\theta)}\right]^{\frac{1}{2}}\right\}, \tag{32}$$

where $[N/2]$ is the greatest integer $\leq (N/2)$.

In order to understand the behavior of the integrand near the branch points, it is convenient to write

$$\left[\frac{\sin\theta}{\sin N\theta}\right]^{\frac{1}{2}} = \left[U_{N-1}\left(1 + \frac{i\lambda'\sigma^2}{N}\right)\right]^{-\frac{1}{2}} = \left\{\left(\frac{2\sigma^2}{N}\right)^{N-1}\prod_{k=1}^{N-1}[i(\lambda' - \lambda'_k)]\right\}^{-\frac{1}{2}}. \tag{33}$$

Here, the last equality is obtained from the fact that

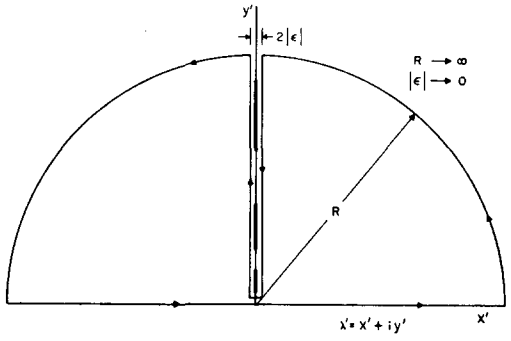


FIG. 2. Contour used in evaluating the integral of Eq. (27).

[11] L. V. Ahlfors, *Complex Analysis* (McGraw-Hill Book Company, Inc., New York, 1953).

the λ_k''s are the roots of the Tchebichef polynomial, and that the coefficient of $\lambda'^{(N-1)}$ in U_{N-1} is $(2i\sigma^2/N)^{N-1}$. We write $(\lambda' - \lambda_k') \equiv \rho_k \exp i(\phi_k')$, where, on the visible Riemann sheet, $-\frac{3}{2}\pi < \phi_k' < \frac{1}{2}\pi$. We further define

$$\phi_k = \frac{\pi}{2} + \phi_k', \quad \bar{\rho} = \prod_{k=1}^{N-1} \rho_k,$$

and

$$K_1 = (N/2\sigma^2)^{\frac{1}{2}(N-1)}.$$

Thus

$$\left[\frac{\sin\theta}{\sin N\theta}\right]^{\frac{1}{2}} = K_1\left\{\bar{\rho}\exp\left[i\sum_{k=1}^{N-1}\phi_k\right]\right\}^{-\frac{1}{2}}. \tag{34}$$

This expression is convenient for examining the behavior of $[\sin\theta/\sin N\theta]^{\frac{1}{2}}$. The positive square root must be taken, as, at $\lambda' = 0$, the result must go into that obtained from Eq. (6) with $\lambda' = 0$, and this, of course, is positive. Thus,

$$\left[\frac{\sin\theta}{\sin N\theta}\right]^{\frac{1}{2}} = K_1\,|\bar{\rho}^{-\frac{1}{2}}|\exp\left[-\frac{i}{2}\sum_{k=1}^{N-1}\phi_k\right]. \tag{35}$$

As we move upward along the imaginary axis with $\mathrm{Re}\,\lambda' > 0$, $\phi_k \to 0$ or π depending on whether we are below or above λ_k', respectively. For $\mathrm{Re}\,\lambda' < 0$, $\phi_k \to 0$ or $(-\pi)$ depending on whether $\mathrm{Im}\,(\lambda')$ is less than or greater than $\mathrm{Im}\,(\lambda_k')$, respectively. Since the integrals in Eq. (32) are evaluated between λ_{2j-1}' and λ_{2j}', we see that the argument of $[\sin\theta/\sin N\theta]^{\frac{1}{2}}$ is $\exp[-\frac{1}{2}i\pi(2j-1)] = i(-1)^j$ for $\mathrm{Re}\,\lambda' > 0$, and is $\exp[+\frac{1}{2}i\pi(2j-1)] = -i(-1)^j$ for $\mathrm{Re}\,\lambda' < 0$. Thus the integrands are of different sign on the two sides of the branch cut. We further see that because of the $(-1)^j$ factor the integrands alternate in sign on neighboring branch cuts. Since $[\sin\theta/\sin N\theta]$ is real along the imaginary axis, we directly obtain

$$P_N(S_x) = N^{\frac{1}{2}}\pi^{-1}\sum_{j=1}^{[N/2]} i(-1)^j \int_{\lambda'_{2j-1}}^{\lambda'_{2j}} d\lambda' e^{i\lambda' S_x}\left|\left[\frac{\sin\theta}{\sin N\theta}\right]^{\frac{1}{2}}\right|. \tag{36}$$

Upon making the substitution $\gamma = -(2j-1)\pi + N \times \arccos(1 + i\lambda'\sigma^2/N)$, the above integral can be written as

$$\begin{aligned} P_N(S_x) = {} & N^{\frac{1}{2}}\pi^{-1}\sigma^{-2} \\ & \times \sum_{j=1}^{[N/2]} (-1)^{j+1}\int_0^{\pi} d\gamma\{\exp[-(NS_x/\sigma^2) \\ & \times \{1 - \cos[(\gamma/N) + (2j-1)(\pi/N)]\}] \\ & \times [\sin\{(\gamma/N) + (2j-1)(\pi/N)\}]^{\frac{3}{2}}[\sin\gamma]^{-\frac{1}{2}}\}. \end{aligned} \tag{37}$$

Thus, the problem of finding $P_N(S_x)$ for all values of N is reduced to that of evaluating a sum of integrals with real integrands, which is amenable to numerical computation.

In addition, Eq. (37) can be used to find an approximation to $P_N(S_x)$ for large S_x. For S_x large, the $j = 1$ term in the sum dominates the series. Assuming $(2\pi/N) \ll 1$ we have

$$\begin{aligned} P_N(S_x) = {} & N^{-1}\pi^{-1}\sigma^{-2} \\ & \times \int_0^{\pi} d\gamma \exp[(-\tfrac{1}{2}S_x/N\sigma^2)(\gamma+\pi)^2](\gamma+\pi)^{\frac{3}{2}}(\sin\gamma)^{-\frac{1}{2}}, \end{aligned} \tag{38}$$

or, in terms of $y = S_x/N\sigma^2$,

$$P_N(y) = \pi^{-1}\int_0^{\pi} d\gamma \exp[-\tfrac{1}{2}y(\pi+\gamma)^2](\pi+\gamma)^{\frac{3}{2}}(\sin\gamma)^{-\frac{1}{2}}. \tag{39}$$

For large y, the only significant contribution to the integral occurs near $\gamma = 0$. The above integral can then be written

$$P_N(y) = \pi^{\frac{1}{2}}\exp(-\tfrac{1}{2}\pi^2 y)\int_0^{\infty} d\gamma\gamma^{-\frac{1}{2}}\exp(-y\pi\gamma), \tag{40}$$

where we have replaced the upper limit by infinity. Evaluation of the above integral yields

$$P_N(y) = (\pi/y)^{\frac{1}{2}}\exp(-\tfrac{1}{2}\pi^2 y). \tag{41}$$

This is in agreement with the result stated by Forsman.[12]

C. Exact Moments

We can obtain the moments of $P_N(S_x)$ from Eq. (27). Defining the mth moment,

$$\alpha_m = \int_{-\infty}^{\infty} S_x^m P_N(S_x)\, dS_x, \tag{42}$$

we find

$$\alpha_m = (-1)^m N^{\frac{1}{2}}\left(\frac{\sigma^2}{N}\right)^m \frac{d^m}{dy^m}\{[U_{N-1}(y)]^{-\frac{1}{2}}\}\Big|_{y=1}. \tag{43}$$

Using the fact that the derivatives of the Tchebichef polynomials are Gegenbauer polynomials one obtains[13]

$$\frac{d^m U_{N-1}(y)}{dy^m}\Big|_{y=1} = \frac{2^m(N+m)!\,m!}{[N-(m+1)]!\,(2m+1)!}. \tag{44}$$

As $U_{N-1}(1) = N$, one may readily evaluate the moments. The first four are seen to be

$$\begin{aligned} \alpha_1 &= \frac{1}{6}N\sigma^2\left[1 - \frac{1}{N^2}\right], \\ \alpha_2 &= \frac{1}{20}N^2\sigma^4\left[1 - \frac{1}{N^4}\right], \\ \alpha_3 &= \frac{1}{2520}N^3\sigma^6\left[61 + \frac{63}{N^2} - \frac{21}{N^4} - \frac{103}{N^6}\right], \\ \alpha_4 &= \frac{1}{75600}N^4\sigma^8\left[1261 + \frac{2440}{N^2} \right. \\ &\qquad \left. + \frac{1638}{N^4} - \frac{1240}{N^6} - \frac{4099}{N^8}\right]. \end{aligned} \tag{45}$$

[12] See Eq. (1) of Ref. 7.
[13] Reference 9, pp. 186 and 174.

D. Limiting Distribution as $N \to \infty$

We now consider the distribution function in the limit $N \to \infty$. For this case it is not difficult to verify that the part of the integral in Eq. (27) which does not go to zero as $N \to \infty$ corresponds to $\lambda'\sigma^2/N \ll 1$. We then have $\theta = i[2i\lambda'\sigma^2/N]^{\frac{1}{2}}$ and $\sin\theta = \theta$. Letting $s = N\sigma^2\lambda'$, we obtain from Eq. (28)

$$P_\infty(S_x) = (2\pi N\sigma^2)^{-1}\int_{-\infty}^{\infty} ds \exp\left[\frac{iS_x s}{N\sigma^2}\right] \times \left\{\frac{[\sinh(2is)^{\frac{1}{2}}]}{(2is)^{\frac{1}{2}}}\right\}^{-\frac{1}{2}}. \quad (46)$$

Defining $y = S_x/N\sigma^2$, we write Eq. (46) as

$$P_\infty(y) = (2\pi)^{-1}\int_{-\infty}^{\infty} ds e^{ivs}\left\{\frac{[\sinh(2is)^{\frac{1}{2}}]}{(2is)^{\frac{1}{2}}}\right\}^{-\frac{1}{2}}. \quad (47)$$

We note that the limiting distribution is only a function of $S_x/N\sigma^2$.[14] Equation (47) is the one-dimensional analog of the result obtained by Fixman.[3] In Sec. V, we evaluate this integral numerically to obtain $P_\infty(y)$.

IV. SADDLE POINT METHOD

In this section we show how the integral for $P_N(S_x)$, as given in Eq. (27), may be evaluated approximately by the method of steepest descent. We see that the saddle point a_0 is always on the imaginary axis, running from $a_0 = -i\infty$ for $S_x = 0$ to the bottom of the first branch point, $a_0 = i(N/\sigma^2)(1 - \cos\pi/N)$ for $S_x = \infty$. That is, for all values of the radius of gyration S_x, there is a saddle point on the imaginary axis below the lowest branch point. Thus the contour of integration along the real axis may always be displaced to go through the saddle point, without the contour crossing any singularities. The approximation yields

$$P_N(S_x) = \exp[g(a_0)][-2\pi/g''(a_0)]^{\frac{1}{2}}, \quad (48)$$

where

$$g(\lambda) = \ln(N^{\frac{1}{2}}/2\pi) + i\lambda' S_x + \tfrac{1}{2}\ln(\sin\theta) - \tfrac{1}{2}\ln(\sin N\theta), \quad (49)$$

and the equation for the saddle point, $\partial g/\partial\lambda' = 0$, is

$$y = \tfrac{1}{2}N^{-2}\csc\theta_a[\cot\theta_a - N\cot N\theta_a], \quad (50)$$

where $\theta_a = \arccos(1 + ia_0\sigma^2/N)$. Differentiating, we obtain

$$g''(a_0) = (\sigma^4/2N^2)\csc^2\theta_a[\csc^2\theta_a + \cot^2\theta_a - N\cot\theta_a\cot N\theta_a - N^2\csc^2 N\theta_a]. \quad (51)$$

[14] It is interesting to note that here we have an example of a quantity which does not have a "macroscopic limit" as the number of elements N approaches infinity. By this we mean that the fluctuation of the observable quantity (the radius of gyration) does not become relatively smaller as N increases.

TABLE I. Position of saddle point for $y = (S_x/N\sigma^2) = \infty, \frac{1}{6}$, and 0.

θ_a	y	λ'
(π/N)	∞	$\frac{iN}{\sigma}\left(1-\cos\frac{\pi}{N}\right)$ [a]
0	$\frac{1}{6}$ $(=\bar{y})$	0
$i\infty$	0	$-i\infty$

[a] First branch point in λ' plane.

It is also convenient to define $\beta_a = -i\theta_a$, in terms of which Eq. (50) becomes

$$y = -\tfrac{1}{2}N^{-2}\operatorname{csch}\beta_a[\coth\beta_a - N\coth N\beta_a]. \quad (52)$$

A study of Eqs. (50) and (52) shows that there is a solution such that a_0 lies on the imaginary axis beneath the first branch point for any value of y. In the limit $y \to \infty$, a_0 approaches the first branch point, i.e., $a_0 \to (iN/\sigma^2)[1 - \cos(\pi/N)]$ and θ_a tends to (π/N). As y decreases, θ_a decreases and a_0 moves down the imaginary axis. At $y = \frac{1}{6}$ (the mean value of y), $\theta_a = a_0 = 0$. For $y < \frac{1}{6}$, a_0 lies on the negative imaginary axis while θ_a is now on the positive imaginary axis. As y becomes smaller, a_0 and θ_a move further from the origin. In the limit $y \to 0$, a_0 approaches $(-i\infty)$ and θ_a approaches $(i\infty)$. The values of y for $\theta_a = (\pi/N)$, 0, and $(i\infty)$ are given in Table I along with the corresponding values of a_0. We can obtain simple approximate expressions for the solutions of Eq. (50) and (52) for three ranges of y, viz., (i) $2\pi y \gg 1$; (ii) $y \cong \frac{1}{6}$; and (iii) $y \ll \frac{1}{2}$. Whenever $\theta_a \ll 1$, which holds for case (ii) and also for case (i) when $N \gg 1$, then, as discussed in Sec. III, the expression for $P_N(y)$ reduces to $P_\infty(y)$, and our treatment is simply the one dimensional analog of Fixman's three-dimensional results. However, for case (iii) (small y), we obtain a more general result than has previously been derived.

Case (i) $2\pi y \gg 1$

We have

$$\begin{aligned} \theta_a &\cong (\pi/N) - (\sigma^2/2\pi S_x), \\ a_0 &\cong i(\pi^2/2N\sigma^2)[1 - 1/\pi^2 y], \\ g''(a_0) &\cong -2(yN\sigma^2)^2, \\ P(y) &\cong (\pi e/2y)^{\frac{1}{2}}\exp(-\tfrac{1}{2}\pi^2 y). \end{aligned} \quad (53)$$

This differs only by the factor $(e/2)^{\frac{1}{2}}$ from the result obtained in Sec. III [Eq. (41)].

Case (ii) $y \cong \frac{1}{6}$

In this region we obtain an approximate expression for the distribution near its mean. For the approximation to the solution of Eq. (50) [or Eq. (52)] to be

valid, we must have $N\theta_a \ll 1$ ($\theta_a = 0$ corresponds to $y = \frac{1}{6}$). In terms of the region of validity for y, this inequality is

$$[90(y - \tfrac{1}{6})]^{\frac{1}{2}} \ll 1.$$

In this region we have

$$a_0 \cong (iN\theta_a^2/2\sigma^2),$$
$$g''(a_0) \cong -(N^2\sigma^4/45)[1 + (20/7)(6y - 1)], \quad (54)$$

and

$$P(y) \cong (45/2\pi)^{\frac{1}{2}}[1 - (60/7)(y - \tfrac{1}{6})].$$

It is seen in the last section, where we present the results of a numerical integration for $P(y)$, that Eq. (54) is indeed in excellent agreement with the computed function.

Case (iii) $y \ll \frac{1}{6}$

Here we obtain an approximation to $P(y)$ in the region where y is much less than its mean value. In terms of the equation relating the value of y to the saddle point, Eq. (52), we can obtain a simple result as long as

$$\coth \beta \ll N. \quad (55)$$

Then Eq. (52) becomes

$$Ny \cong \tfrac{1}{2} \operatorname{csch} \beta_a. \quad (56)$$

Here we have used the fact that Eq. (55) implies $N\beta \gg 1$. Inserting Eq. (56) into Eq. (55) we see that the range of y over which this approximate form holds is $y \ll \frac{1}{2}$.

We then have the results

$$a_0 \cong -i(N/\sigma^2)\{[1 + (\sigma^4/4S_x^2)]^{\frac{1}{2}} - 1\},$$
$$g''(a_0) \cong -(2/N)(yN\sigma^2)^2[1 + 4y^2N^2]^{\frac{1}{2}}, \quad (57)$$

$$P(y) \cong \frac{\exp(N^2y\{-1 + [1 + (2Ny)^{-2}]^{\frac{1}{2}}\})}{4\pi^{\frac{1}{2}}y^2\{1 + (2Ny)^{-2}\}\{(2Ny)^{-1} + [1 + (2Ny)^{-2}]^{\frac{1}{2}}\}}.$$

If $1/2N \ll y \ll \frac{1}{2}$, then Eq. (57) reduces to

$$P(y) \cong \tfrac{1}{4}\pi^{-\frac{1}{2}}y^{-2} \exp[-(8y)^{-1}]. \quad (58)$$

If $y \ll 1/2N$, Eq. (57) becomes

$$P(y) \cong (N/8\pi)^{\frac{1}{2}}y^{-\frac{3}{2}}e^{N/2}(yN)^{N/2}. \quad (59)$$

From this we see that the single expression for $P(y)$ given in Eq. (57) covers both of the small y cases discussed in Ref. 7 by Forsman.

V. NUMERICAL EVALUATION OF THE DISTRIBUTION FUNCTION AS $N \to \infty$

An integral representation of the distribution function $P_N(y)$ in the limit $N \to \infty$ is given by Eq. (47). In this section, we obtain an integral representation of $P_\infty(y)$ with a real integrand and evaluate $P_\infty(y)$ numerically. To accomplish this we need to express the quantity $\{[\sinh(2is)^{\frac{1}{2}}]/(2is)^{\frac{1}{2}}\}^{-\frac{1}{2}}$ in terms of its real and imaginary parts. Defining

$$E(s) = \{[\sinh(2is)^{\frac{1}{2}}]/(2is)^{\frac{1}{2}}\}^{\frac{1}{2}}$$
$$= \{A(s) + iB(s)\}^{\frac{1}{2}}, \quad (60)$$

where $A(s)$ and $B(s)$ are real, we have

$$A(s) = \tfrac{1}{2}s^{-\frac{1}{2}}[\cos s^{\frac{1}{2}} \sinh s^{\frac{1}{2}} + \sin s^{\frac{1}{2}} \cosh s^{\frac{1}{2}}], \quad (61a)$$

$$B(s) = \tfrac{1}{2}s^{-\frac{1}{2}}[-\cos s^{\frac{1}{2}} \sinh s^{\frac{1}{2}} + \sin s^{\frac{1}{2}} \cosh s^{\frac{1}{2}}]. \quad (61b)$$

It can be seen that $A(s)$ is even while $B(s)$ is odd. We can then write

$$E(s) = E_R(s) + iE_I(s), \quad (62a)$$

where $E_R(s)$ and $E_I(s)$ are real and may be written

$$E_R(s) = 2^{-\frac{1}{2}}S_R(s)\{A(s) + [A^2(s) + B^2(s)]^{\frac{1}{2}}\}^{\frac{1}{2}}, \quad (62b)$$

$$E_I(s) = 2^{-\frac{1}{2}}S_I(s)\{-A(s) + [A^2(s) + B^2(s)]^{\frac{1}{2}}\}^{\frac{1}{2}}, \quad (62c)$$

where $S_R(s)$ and $S_I(s)$ are functions of s and are equal to plus or minus unity. These functions are discussed in the next paragraph. It can be seen that, since $A(s)$ and $B(s)$ are even and odd functions of s, respectively, the functions $E_R(s)$ and $E_I(s)$ are even and odd, respectively. We can now write $P_\infty(y)$ as

$$P_\infty(y) = (2\pi)^{-1}\int_{-\infty}^{\infty} ds e^{isy}\{E_R(s) + iE_I(s)\}^{-1}$$
$$= (2\pi)^{-1}\int_{-\infty}^{\infty} ds[E_R^2 + E_I^2]^{-1}[E_R \cos(sy) - E_I \sin(sy) + iE_I \cos(sy) + iE_R \sin(sy)]. \quad (63)$$

Because of the symmetries of $E_R(s)$, $E_I(s)$, $\sin(sy)$, and $\cos(sy)$, the integrals containing $E_I \cos(sy)$ and $E_R \sin(sy)$ vanish. Similarly we can change the limits of integration to zero to infinity, obtaining

$$P_\infty(y) = \pi^{-1}\int_0^{\infty} ds[E_R^2 + E_I^2)^{-1} \times [E_R \cos(sy) - E_I \sin(sy)]. \quad (64)$$

Making use of the fact that $P_\infty(-|y|) = 0$, we have

$$P_\infty(y) = 2\pi^{-1}\int_0^{\infty} ds[E_R^2 + E_I^2]^{-1}E_R \cos(sy). \quad (65)$$

It is necessary to determine the function $S_R(s)$ in order to evaluate $P_\infty(y)$ as given in Eq. (65). To do this we recall the discussion of Sec. III concerning the branch cuts. Using the nomenclature of Sec. III, we

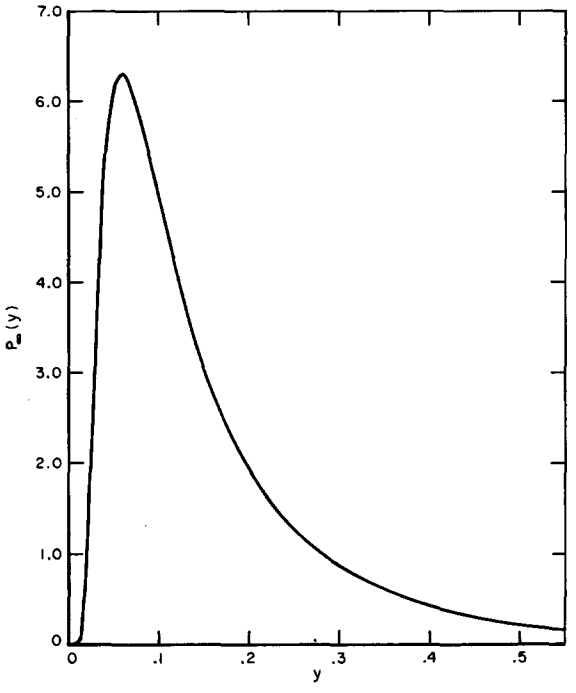


FIG. 3. The probability distribution $P_\infty(y)$ of one component of the square of the radius of gyration.

write $E(s)$ as[15]

$$E(s) = N^{-\frac{1}{2}}\left[\frac{\sin\theta}{\sin N\theta}\right]^{-\frac{1}{2}} = N^{-\frac{1}{2}}K_1^{-1}\bar{\rho}^{\frac{1}{2}}\exp\left[\frac{i}{2}\sum_{k=1}^{N-1}\phi_k\right]. \tag{66}$$

Hence we have

$$E_R = \left(\frac{\bar{\rho}}{NK_1^2}\right)^{\frac{1}{2}}\cos\left[\tfrac{1}{2}\sum_{k=1}^{N-1}\phi_k\right], \tag{67}$$

$$E_I = \left(\frac{\bar{\rho}}{NK_1^2}\right)^{\frac{1}{2}}\sin\left[\tfrac{1}{2}\sum_{k=1}^{N-1}\phi_k\right]. \tag{68}$$

When s is near the origin, $(\frac{1}{2}\sum_{k=1}^{N-1}\phi_k) \ll 1$ and E_R is positive. It is clear that E_R changes sign whenever $(\frac{1}{2}\sum_{k=1}^{N-1}\phi_k) = (2n+1)\frac{1}{2}\pi$; $n = 0, \pm1, \pm2, \cdots$. At these values $E_R = 0$ and we can find the values of s at which $(\frac{1}{2}\sum_{k=1}^{N-1}\phi_k) = (2n+1)\frac{1}{2}\pi$ by setting $E_R(s) = 0$. This is given by the solution of the equation $\cos(s_n^{\frac{1}{2}})\sinh(s_n^{\frac{1}{2}}) = \sin(s_n^{\frac{1}{2}})\cosh(s_n^{\frac{1}{2}})$ with $\cos(s_n^{\frac{1}{2}}) < 0$ and $\sin(s_n^{\frac{1}{2}}) < 0$. The approximate solution is $|s_n| \cong \{2\pi n + 5\pi/4 - \exp[-2(2\pi n + 5\pi/4)]\}^2$, $n = 0, 1, 2, \cdots$. Thus $E_R(s)$ is positive for $0 < s < s_1$, negative for $s_1 < s < s_2$, positive for $s_2 < s < s_3$, and so on, which specifies the function $S_R(s)$ introduced in Eq. (62b).

The integral for $P_\infty(y)$ as given in Eq. (65) has been evaluated numerically. The results are shown in Fig. 3. The maximum in $P_\infty(y)$ occurs at $y = 0.060 \pm 0.001$. The zero, first, and second moments were computed from $P_\infty(y)$ and agree within 1% with the exact moments. Numerically, we obtain $P_\infty(\frac{1}{6}) \cong 2.61$ and

$$\left.\frac{dP_\infty}{dy}\right|_{y=\frac{1}{6}} \cong 23.8,$$

while the saddle point method of Sec. IV gives $P_\infty(\frac{1}{6}) = 2.68$ and

$$\left.\frac{dP}{dy}\right|_{y=\frac{1}{6}} = 23.0.$$

Forsman and Hughes[5] have obtained an approximation to $P_\infty(y)$ by an iterative procedure. Their results are in reasonable agreement with the results obtained here, e.g., the maximum in $P_\infty(y)$ as obtained by Forsman and Hughes occurs at $y = 0.066$ and $P_\infty(0.066) = 5.5$ (estimated from Fig. 1 of their paper) while our maximum is at $y = 0.060 \pm 0.001$ with $P_\infty(0.060) = 6.30$. Equation (58), which was obtained by the saddle point method and which is valid for $y \ll \frac{1}{2}$, predicts a maximum in $P_\infty(y)$ at $y = \frac{1}{16}$ with $P_\infty(\frac{1}{16}) = 4.9$.

ACKNOWLEDGMENTS

The authors would like to thank Dr. R. J. Rubin and Dr. J. A. Simmons for helpful discussions and suggestions.

We also wish to thank the Advanced Research Projects Agency for partial support of this work. One of us (JLJ) would also like to thank the National Science Foundation and the Office of Naval Research for their support.

APPENDIX

In this appendix, we present another method of finding the determinant of the $(N \times N)$ matrix D,

$$D_{kl} = -(i\lambda/N^2) + [1 + (i\lambda/N)]\delta_{kl} - \tfrac{1}{2}\delta_{kN}\delta_{lN} - \tfrac{1}{2}[\delta_{k,l+1} + \delta_{k,l-1}]. \tag{A1}$$

We define the $(N \times N)$ matrices L and B by

$$L_{kl} = \delta_{kl} - \delta_{k,l+1}, \tag{A2}$$

$$B_{kl} = \tfrac{1}{2}\delta_{kl} - (i\lambda/N^2)(k-1)(l-1) + (i\lambda/N) \times [(k-1)S(l-k) + (l-1)S(k-l)], \tag{A3}$$

where

$$S(x) = \begin{cases} 1; & x > 0, \\ \frac{1}{2}; & x = 0, \\ 0; & x < 0. \end{cases} \tag{A4}$$

By matrix multiplication, it can be shown that $D = L^T B L$. Since $|L| = |L^T| = 1$, we have $|D| = |B|$. The

[15] Rather than rewrite the corresponding functions in the limit as $N \to \infty$ we use the finite N forms from Sec. III [i.e., Eq. (34)]. The explanation is not affected by leaving N finite.

matrix B has the form

$$B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & A \end{pmatrix}, \tag{A5}$$

where the $(N-1)$ by $(N-1)$ matrix A is defined by

$$A_{kl} = B_{k+1,l+1};$$

i.e.,

$$A_{kl} = \tfrac{1}{2}\delta_{kl} - (i\lambda/N^2)kl + (i\lambda/N) \times [kS(l-k) + lS(k-l)], \tag{A6}$$

and we have $|B| = \frac{1}{2}|A|$.

We next define the $(N-1)$ by $(N-1)$ matrix C by the equation

$$A = \tfrac{1}{2}I + (i\lambda/N^2)C, \tag{A7}$$

where I is the identity matrix. Thus,

$$C_{kl} = -kl + N[kS(l-k) + lS(k-l)]. \tag{A8}$$

The inverse of C is given by

$$(C^{-1})_{kl} = (1/N)[2\delta_{kl} - \delta_{k,l+1} - \delta_{k+1,l}], \tag{A9}$$

which can be proved by matrix multiplication. Noting that $A = C[\frac{1}{2}C^{-1} + (i\lambda/N^2)I]$, we have

$$|D| = \tfrac{1}{2}|C|\left|\tfrac{1}{2}C^{-1} + \frac{i\lambda}{N^2}I\right|$$

or

$$|D| = \frac{|NC^{-1} + (2i\lambda/N)I|}{2^N |NC^{-1}|}. \tag{A10}$$

Let $P_{N-1}(x)$ be the $(N-1)$ by $(N-1)$ matrix defined by $[P_{N-1}(x)]_{kl} = x\delta_{kl} - \delta_{k,l+1} - \delta_{k+1,l}$. It has been shown by Wolstenholme[16] (see Rutherford[17]) that

$$|P_{N-1}(x)| = \sin N\theta/\sin\theta = U_{N-1}(\tfrac{1}{2}x), \tag{A11}$$

where $\theta = \arccos(\frac{1}{2}x)$. Combining this result with Eq. (A10), we obtain

$$|D| = \frac{U_{N-1}(1 + i\lambda/N)}{N2^N}, \tag{A12}$$

which is the same as Eq. (26).

[16] J. Wolstenholme, Educ. Times **27**, 67 (1874).

[17] D. E. Rutherford, Proc. Roy. Soc. Edinburgh **62A**, 229 (1946–47).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Criterion for Bose–Einstein Condensation and Representation of Canonical Commutation Relations

HIROSHI EZAWA*
Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

MARSHALL LUBAN†
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania

(Received 14 September 1966)

The existence of a Bose–Einstein condensation in an interacting many-boson system at $T = 0°K$ is proved under certain conditions on the particle density and the interparticle potential. Starting with the tentative assumption that the condensation exists, we study the fluctuation in the occupation number of the condensate with due regard to its interactions (1) with particles outside the condensate as well as (2) with the fluctuation itself. If the condensate fluctuation has a normalizable ground state, then the assumed existence of the condensation is tenable. For the case of the pair-Hamiltonian model satisfying the conditions for condensation, the interactions of the second category are of no importance. In the limit of infinite volume, this Hamiltonian can be diagonalized in an irreducible representation of a Bose-field operator $\phi(\mathbf{x})$, where $\phi(\mathbf{x})$ has nonvanishing ground state expectation value, in accordance with the usual c-number replacement of creation and destruction operators for the condensate particles. The full Hamiltonian for a system of pairwise interacting bosons is studied only in a low-density limit. Bose–Einstein condensation exists when the over-all space integral of the interparticle potential is positive. In this case the interactions of the second category play an important role in ensuring a normalizable ground state for the condensate fluctuation. There is an indication that in the limit of infinite volume the Hamiltonian cannot be diagonalized in any irreducible representation of the field operator ϕ. Yet the c-number replacement of the condensate operators is legitimate as far as states of particles outside the condensate are concerned. Some speculations are made as to what may happen for systems of moderate density.

I. INTRODUCTION

IT is universally believed that at sufficiently low temperatures a Bose–Einstein (B.E.) condensation occurs in a system of $N \to \infty$ interacting bosons.[1] To date no proof has been given to support this belief except for several simplified model systems.[2] In a classic paper on interacting bosons, Bogoliubov[3] proposed that one simplify the second-quantized Hamiltonian, so as to take advantage of the assumed B.E. condensation, by replacing $a_0^\dagger$ and a_0, the creation and destruction operators for the zero-momentum single-particle state, by a c number $N_0^{\frac{1}{2}}$. The quantity N_0 is the average number of particles occupying this state, and it is supposed to be $O(N)$, of course. Bogoliubov's procedure has been very useful for the purpose of providing a first-principles explanation of some of the low-temperature properties of liquid ^{4}He. Nevertheless, it is necessary to ascertain the domain of applicability of this otherwise *ad hoc* procedure. In this regard, an *a posteriori* verification[4–6] that $N_0 = O(N)$ *does not* establish its legitimacy.

Based on a result[7,8] from the theory of the representation of canonical commutation relations, one of us (H. E.)[8] proposed the operator replacement

$$a_0 = N_0^{\frac{1}{2}} + c, \tag{1.1}$$

where the operator c and its adjoint $c^\dagger$ satisfy the usual boson commutation relations,

$$\begin{gathered}[c, c^\dagger] = 1, \quad [c, c] = [c^\dagger, c^\dagger] = 0, \\ [c^{(\dagger)}, a_{\mathbf{k}}^{(\dagger)}] = 0 \quad (\mathbf{k} \neq 0).\end{gathered} \tag{1.2}$$

It is the main subject of this paper to study the usefulness of the replacement (1.1) for the purpose of obtaining an existence criterion for the B.E. condensation.

Two remarks are in order about this operator replacement. First, if we adhered to the usual Fock

---

* Present address: Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo, Japan.

† Present address: Department of Physics, Bar-Ilan University, Ramat Gan, Israel.

[1] The limit we are referring to is that for which both the number N of particles and the volume V enclosing the system become infinitely large, with the particle density $\rho = N/V$ held fixed. Such a limit is called the volume limit.

[2] W. H. Bassichis and L. L. Foldy, Phys. Rev. 133, A935 (1964). Their study concerns a model system which, in the absence of interactions, consists of only three energy states. The existence of a B.E. condensation can also be demonstrated for the Hartree–Fock model of interacting bosons.

[3] N. N. Bogoliubov, J. Phys. (USSR) 11, 23 (1947).

[4] T. D. Lee, K. Huang, and C. N. Yang, Phys. Rev. 106, 1135 (1957).

[5] N. M. Hugenholtz and D. Pines, Phys. Rev. 116, 489 (1959).

[6] S. Beliaev, Soviet Phys.—JETP 7, 289 (1958).

[7] H. Araki and E. J. Woods, J. Math. Phys. 4, 637 (1963).

[8] H. Ezawa, J. Math. Phys. 6, 380 (1965).

representation[9] for the operators $a_{\mathbf{k}}^{(\dagger)}$, then any state vector describing B.E. condensation would keep rotating in the Fock space and therefore no convergence could be expected in the volume limit when $N_0 \to \infty$. It should be recognized here that the B.E. condensation can be defined only in the volume limit, $N,\ V \to \infty,\ N/V = \rho$ fixed.[10] Thus we need some replacement for the quantum mechanical variables $a_0^{(\dagger)}$. The idea of the replacement (1.1) is to shift the "origin" of $a_0^{(\dagger)}$ in accordance with the size N_0 of the condensate. The new operators $c^{(\dagger)}$ describe the fluctuations of the $\mathbf{k} = 0$-state occupation number from its average. Underlying (1.1) is the hope that $c^{(\dagger)}$ remain "finite" in the volume limit when $N_0 \to \infty$. If unfortunately the condensate of a given system is such that its fluctuations[11] cannot be described by any finite $c^{(\dagger)}$, then a further change of variables will be needed. The proper choice of the variables will be determined by the Hamiltonian of the system.[12] Parenthetically we note here that a finite description of the condensate may be achieved in some other ways too. For instance, we may try the variables P, Q defined by $a_0 = 2^{-\frac{1}{2}}(N_0^{\frac{1}{2}}P - iN_0^{-\frac{1}{2}}Q)$ and

$$a_0^\dagger = 2^{-\frac{1}{2}}(N_0^{\frac{1}{2}}P + iN_0^{-\frac{1}{2}}Q);$$

they satisfy the commutation relation $[P, Q] = -i$. The number of condensate particles, $a_0^\dagger a_0 = \frac{1}{2}N_0 \times (P^2 + N_0^{-2}Q^2 - N_0^{-1})$, can be $O(N)$ for finite operators P, Q.

The second remark concerns the particular features of our choice of variables (1.1). For one thing, the Bogoliubov replacement $a_0^{(\dagger)} \to N_0^{\frac{1}{2}}$ is closely related to (1.1). If it can be shown in fact that $c^{(\dagger)} = o(N_0^{\frac{1}{2}})$ for the condensate of a given system, then the Bogoliubov replacement is justified for the treatment of the system. The notation $c^{(\dagger)} = o(N_0^{\frac{1}{2}})$ should be understood to mean that the operators $c^{(\dagger)}$ can be treated as if they are quantities of $o(N_0^{\frac{1}{2}})$. The precise meaning can be given only after we find a Hilbert space appropriate for describing $c^{(\dagger)}$. Another interesting feature of the replacement (1.1) is that it leads to an irreducible representation of boson-field operators if $c^{(\dagger)} = o(N_0^{\frac{1}{2}})$. If we used the representation in terms of the finite operators P, Q defined above, the term $N_0^{-\frac{1}{2}}Q$ becomes, in effect, negligible in the volume limit and only the operator P remains relevant, so that P can be treated as a c number. The representation of the boson-*field* operators becomes a direct sum of representations, each corresponding to a particular value of the c number P. The representation is thus reducible in the volume limit.

We now formulate within the context of the operator replacement (1.1) the existence criterion for the B.E. condensation for the special case of $T = 0°\text{K}$. Let $\overline{\mathcal{H}}$ denote the form of the Hamiltonian which obtains when the operators a_0 and $a_0^\dagger$ are replaced in accordance with (1.1). Since we have to use a representation in which the total number of particles is not sharp, we introduce $\mathcal{K} = \overline{\mathcal{H}} - \mu\mathcal{N}$, where $\mathcal{N} = \sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$ and μ are the total number operator and the chemical potential, respectively. We call $\mathcal{K}$, rather than $\overline{\mathcal{H}}$, the Hamiltonian of the system. Since the Hamiltonian $\mathcal{K}$ and therefore its ground state $|\Omega\rangle$ involve two parameters N_0 and μ, we can impose two subsidiary conditions on the ground state,

$$\langle\Omega|\, c\,|\Omega\rangle = \langle\Omega|\, c^\dagger\,|\Omega\rangle = 0, \tag{1.3}$$

$$\langle\Omega|\, \mathcal{N}\,|\Omega\rangle = N_0 + \langle\Omega|\, c^\dagger c + \sum_{\mathbf{k}}{}' a_{\mathbf{k}}^\dagger a_{\mathbf{k}}\,|\Omega\rangle = N, \tag{1.4}$$

besides the usual normalization requirement

$$\langle\Omega\,|\,\Omega\rangle = 1. \tag{1.5}$$

Here and in the following a prime on a summation symbol for momentum states $\mathbf{k}$ means that the $\mathbf{k} = 0$ term is to be omitted. Note that if we understand the canonical transformation as providing an analog of the method of small oscillations, then (1.3) should be compared to the equilibrium condition with which we determine the center of oscillation. A B.E. condensation occurs if the number N_0 determined from (1.3)–(1.5) is in fact macroscopic, $N_0 = O(N)$. For the purpose of practical calculations, it is actually more convenient to start by assuming that N_0 is macroscopic. Then one must determine whether or not the Hamiltonian has a bona fide ground state satisfying the restrictions (1.3)–(1.5).

To facilitate such calculations, in Sec. II of this paper we present a modified version of the Born–Oppenheimer method,[13] which, as is well known, was originally designed for the treatment of molecular vibrations. Just as Born and Oppenheimer eliminated the electron–nucleus interactions to obtain an effective

[9] For the classification of the representation of canonical commutation relations, see L. Gårding and A. S. Wightman, Proc. Natl. Acad. Sci. U.S. **40**, 622 (1954); A. S. Wightman and S. S. Schweber, Phys. Rev. **98**, 812 (1955); R. Haag, *Lectures in Theoretical Physics* (Interscience Publishers, Inc., New York, 1961), Vol. III.

[10] For the general definitions of B.E. condensation, see O. Penrose and L. Onsager, Phys. Rev. **104**, 576 (1956); C. N. Yang, Rev. Mod. Phys. **34**, 694 (1962).

[11] The size of the condensate fluctuation is given by the formula

$$\sigma(a_0^\dagger a_0) = N_0\sigma(c + c^\dagger) + 2N_0^{\frac{1}{2}}\sigma(c + c^\dagger, c^\dagger c) + \sigma(c^\dagger c),$$

where

$$\sigma(A, B) = \tfrac{1}{2}\langle AB + BA\rangle - \langle A\rangle\langle B\rangle, \quad \sigma(A) = \sigma(A, A),$$

and the angular bracket means expectation value.

[12] H. Araki, J. Math. Phys. **1**, 492 (1960).

[13] M. Born and R. Oppenheimer, Ann. Physik **84**, 457 (1927); R. de Kronig, Z. Physik **50**, 347 (1928).

nuclear potential, so we construct an effective Hamiltonian Λ_n for the condensate fluctuation by solving an "eigenvalue problem" for the particles outside the condensate ($\mathbf{k} \neq 0$). To be specific, let $\mathfrak{H}_B$ and $\mathfrak{H}_C$ denote the Hilbert spaces appropriate to operators $a_{\mathbf{k}}^{(\dagger)}$, ($\mathbf{k} \neq 0$), and $c^{(\dagger)}$, respectively. Then Λ_n is an operator on $\mathfrak{H}_C$ such that

$$\mathcal{H}\psi_n = \psi_n \Lambda_n, \tag{1.6}$$

where ψ_n is a vector of $\mathfrak{H}_B$ describing the state of the $\mathbf{k} \neq 0$ particles and, at the same time, an operator defined on $\mathfrak{H}_C$. In other words, ψ_n is a linear combination of vectors of $\mathfrak{H}_B$ whose coefficients are functions of the operators $c^{(\dagger)}$. The reason why we call Λ_n an effective condensate Hamiltonian is that, given the "state" ψ_n of the $\mathbf{k} \neq 0$ particles, the spectrum of the associated condensate fluctuation is obtained by solving the eigenvalue problem

$$\Lambda_n |\phi_{n\alpha}) = E_{n\alpha} |\phi_{n\alpha}), \tag{1.7}$$

where the $|\phi_n) \in \mathfrak{H}_C$ are the state vectors of the condensate. The physical meaning of the eigenvalue $E_{n\alpha}$ can be found by combining (1.6) and (1.7):

$$\mathcal{H} |n, \alpha\rangle = E_{n\alpha} |n, \alpha\rangle, \tag{1.8}$$

where

$$|n, \alpha\rangle = \psi_n |\phi_{n\alpha}) \in \mathfrak{H}_B \otimes \mathfrak{H}_C.$$

Thus $E_{n\alpha}$ is the energy eigenvalue of the total system. The differences $E_{n\alpha} - E_{n0}$ give the excitation energies of the condensate fluctuation. Let us denote the ground state of the total Hamiltonian by $n = \alpha = 0$: $|0, 0\rangle = |\Omega\rangle$. As we see in the following sections the condition (1.3) serves to determine the chemical potential[14] by requiring that Λ_0 has no term linear in $c^{(\dagger)}$. The number N_0 is then determined by (1.4). If the above program yields a normalizable ground state and if the number N_0 turns out to be macroscopic as desired, then we can conclude that the boson system undergoes a B.E. condensation.

In Secs. III–VI we study the eigenvalue problem (1.6) and (1.7) for the case of the pair Hamiltonian model, whose thermodynamical properties have been studied by one of us (M. L.).[15] Then in Sec. VII we try to discuss some challenging problems one meets when one examines the case of the full boson Hamiltonian. Note that the pair Hamiltonian includes only a small subclass of interaction terms included in the latter. Despite several unphysical properties, the pair Hamiltonian is interesting because it admits of a complete mathematical analysis.

In Sec. III we begin our study of the pair Hamiltonian by solving (1.6), the first part of our eigenvalue problem. Reasonable solutions can be obtained only when the interparticle potential and the particle density $\rho = N/V$ satisfy certain conditions ($k \neq 0$ stability conditions), e.g., $v(0) > 0$, where $v(\mathbf{k})$ is the potential in the momentum representation.[16] For the moment we assume that these conditions are satisfied.

In Sec. IV we calculate the effective condensate Hamiltonian Λ_0 by making use of the perturbation formulas presented in Appendix A. The Hamiltonian consists of two parts, a finite part Λ_0^0 and an infinitesimal part Λ_0'; both are power series in $c^{(\dagger)}$, and the coefficients are $O(1)$ in the former and $o(1)$ in the latter as $N \to \infty$. In the second-order approximation we obtain

$$\Lambda_0^0 = W_0 + f_0 c^\dagger c + \tfrac{1}{2} h_0 (cc + c^\dagger c^\dagger), \tag{1.9}$$

where $W_0/N, f_0$, and h_0 are c numbers of $O(1)$.[17] For the present case of the pair Hamiltonian, the use of Green functions shows that there is good reason to believe that this expression for Λ_0^0 is asymptotically correct to *all* orders as $N \to \infty$. The Hamiltonian (1.9) can be studied more conveniently after the canonical transformation,

$$p = 2^{-\frac{1}{2}}(c + c^\dagger), \quad q = i2^{-\frac{1}{2}}(c - c^\dagger); \tag{1.10}$$

the new variables satisfy the usual momentum-coordinate commutation relation,

$$[p, q] = -i. \tag{1.11}$$

Then (1.9) becomes

$$\Lambda_0^0 = W_0 - \tfrac{1}{2} f_0 + \tfrac{1}{2}(f_0 + h_0)(p^2 + \lambda q^2), \tag{1.12}$$

where

$$\lambda = (f_0 - h_0)/(f_0 + h_0). \tag{1.13}$$

It is now clear that Λ_0^0 has a normalizable ground state $|\phi_{00}\rangle$ if and only if

$$f_0 \pm h_0 > 0. \tag{1.14}$$

The inequalities (1.14) constitute what we call the $\mathbf{k} = 0$ stability conditions. When (1.14) is satisfied, the requirements (1.3) and (1.5) can be satisfied by $|\Omega\rangle = \psi_0 |\phi_{00}\rangle$. The number N_0 we get from (1.4) is macroscopic in accord with our presupposition of the B.E. condensation. Further, as desired, the operators

[14] It may be seen that the chemical potentials μ_n determined, respectively, by the condition that Λ_n has no terms linear in $N_0^{\frac{1}{2}} c^{(\dagger)}$, all coincide, $\mu_n = \mu_0$, as long as the states ψ_n differ from ψ_0 by a finite number of $k \neq 0$ excitations.

[15] M. Luban, Phys. Rev. **128**, 965 (1962).

[16] The condition $v(0) > 0$ means, therefore, that the *space* integral of the interparticle potential must be positive.

[17] We can calculate the effective condensate Hamiltonian Λ_n corresponding to the state ψ_n of the $\mathbf{k} \neq 0$ particles. It may be seen that $\Lambda_n = \Lambda_0$ as long as the state ψ_n differs from ψ_0 by a finite number of $\mathbf{k} \neq 0$ excitations.

$c^{(\dagger)}$ are finite in the sense that all matrix elements[18] $\langle\Omega|\,(c^\dagger)^i(c)^j\,|\Omega\rangle$, $i, j = 0, 1, 2, \cdots$, are $O(1)$. Thus the Bogoliubov replacement $a_0^{(\dagger)} \to N_0^{\frac{1}{2}}$ is valid.

In Secs. V and VI we study the implications of the $\mathbf{k} = 0$- and $\mathbf{k} \neq 0$-stability conditions on the interparticle potential and the particle density ρ. In particular, the condition $f_0 - h_0 > 0$ is found to imply that $\tilde{v}$ must be partially attractive.[19,20] Thus a potential that is purely repulsive, $\tilde{v}(\mathbf{x}) > 0$, in coordinate space will lead to a violation of (1.14). For a many-boson system with such a potential, the discussion in Appendix B shows that the operator part $c^{(\dagger)}$ of (1.1) must be comparable in order of magnitude with the c number part $N_0^{\frac{1}{2}}$. Two possibilities are therefore open. Either (1) a B.E. condensation occurs in such a system but the operator replacement (1.1) is not useful, or (2) a B.E. condensation, at least of the usual variety, fails to occur in the system. Regarding the latter possibility, we can also show that the assumption of *no* condensation for such a system is untenable (see Sec. V).

When we turn our attention to the case of the full Hamiltonian in Sec. VII, the situation becomes less clear because we are unable to calculate the effective condensate Hamiltonian in an asymptotically exact fashion. In the low-density limit, where a second-order perturbation calculation is believed to be meaningful, the finite part Λ_0^0 of the condensate Hamiltonian turns out to have the form of (1.9) with the coefficients[21] $f_0 + h_0 = 2\rho_0 v(0)$, and $f_0 - h_0 = 0$; the restoring force term in (1.12) vanishes and therefore the operators $c^{(\dagger)}$ cannot remain $O(1)$ as $N \to \infty$. Vanishing of the restoring force is closely related to the Hugenholtz–Pines theorem[5,22] and the corresponding result of Gavoret and Nozières,[22,23] which assert, respectively, that the one- and two-particle excitation energies vanish in the low-momentum limit. Since $c^{(\dagger)}$ may become large, the infinitesimal part Λ_0' of the condensate Hamiltonian will play a significant role. Inclusion of Λ_0' leads to a Hamiltonian Λ_0 of the form a constant times the following:

$$(c + c^\dagger)^2 + 2N_0^{-\frac{1}{2}}(c^\dagger cc + c^\dagger c^\dagger c) + N_0^{-1}c^\dagger c^\dagger cc = [(a_0^\dagger a_0 - N_0 - \tfrac{1}{2})^2 - \tfrac{1}{4}]/N_0, \quad (1.15)$$

which is diagonal in the occupation number representation of $a_0^\dagger a_0$. The energy is minimum when $a_0^\dagger a_0 = N_0$ or $N_0 + 1$ so that the system undergoes a B.E. condensation. It is remarkable that Λ_0 is diagonal only in the occupation number representation of $a_0^\dagger a_0$, a representation which is *reducible* in the volume limit.[7] The reducibility is a reflection of the nonconvergence of the state vector, a situation mentioned at the beginning of this section.

Finally, in Sec. VIII we discuss the relevance of our present results to the theory of representations of canonical commutation relations.

II. MODIFIED BORN–OPPENHEIMER APPROXIMATION

In order to discuss the condensate fluctuation we use a perturbation method which can be regarded as a natural extension of the Born–Oppenheimer approximation.[13,24] Deferring the presentation of the mathematical details until Appendix A, in this section we discuss the pertinent features of our method.

Suppose we are given a dynamical system which consists of two subsystems B and C. The Hamiltonian consists of three parts,

$$\mathcal{H} = \mathcal{H}_B + \mathcal{H}_C + \mathcal{H}_I, \quad (2.1)$$

where $\mathcal{H}_B$ and $\mathcal{H}_C$ are the Hamiltonians of the subsystems and $\mathcal{H}_I$ that of their interaction. Under the assumption that the Hamiltonian $\mathcal{H}_B$ has been completely diagonalized, our method is designed to give an "effective Hamiltonian" for the system C with due regard to its interaction with the subsystem B.

Let the Hilbert spaces appropriate to the subsystems B and C be $\mathfrak{H}_B$ and $\mathfrak{H}_C$, respectively. The state vector of the entire system then belongs to $\mathfrak{H}_B \otimes \mathfrak{H}_C$. The diagonalization of the total Hamiltonian is achieved in two steps. First we find ψ_n and Λ_n, the former being a vector in $\mathfrak{H}_B$ and at the same time an operator in $\mathfrak{H}_C$, and the latter an operator in $\mathfrak{H}_C$, such that

$$(\mathcal{H}_B + \mathcal{H}_C + \mathcal{H}_I)\psi_n = \psi_n\Lambda_n, \quad (2.2)$$

and

$$\psi_n^\dagger\psi_n = \mathbf{1}, \quad (2.3)$$

where $\mathbf{1}$ is the identity operator in $\mathfrak{H}_C$. More precisely, ψ_n has the form

$$\psi_n = \sum_j |u_j)\, L_j, \quad (2.4)$$

where $|u_j\rangle \in \mathfrak{H}_B$ and the coefficients L_j of the linear combination belong to the ring of operators in $\mathfrak{H}_C$.

[18] In view of the previous footnote, the expectation values with respect to the excited states are also $O(1)$.

[19] The precise condition on v is given in Sec. V. See (5.5), (5.15), and the lemma preceding (5.19).

[20] The quantity $\tilde{v}(\mathbf{x})$ is the Fourier transform of $v(\mathbf{k})$.

[21] Note that $v(0) > 0$ when the $\mathbf{k} \neq 0$ stability conditions are satisfied.

[22] K. Huang and A. Klein, Ann. Phys. (N.Y.) **30**, 203 (1964).

[23] J. Gavoret and P. Nozières, Ann. Phys. (N.Y.) **28**, 349 (1964).

[24] A detailed comparison between the present method and the original method of Born and Oppenheimer[13] as applied to the problem of the hydrogen molecule will be presented by one of us (H. E.) in a forthcoming publication.

The product in (2.3) means that

$$\psi_n^\dagger \psi_n = \sum_{i,j} (u_i \mid u_j) L_i^\dagger L_j. \tag{2.5}$$

It should be noted here that Λ_n is a Hermitian operator; in fact, by the isometric property (2.3) of ψ_n, we see from (2.2) that

$$\Lambda_n = \psi_n^\dagger \mathcal{H} \psi_n, \tag{2.6}$$

the product being understood in the sense of (2.5). We call Λ_n an effective Hamiltonian for the subsystem C. The second step in the diagonalization of the total Hamiltonian $\mathcal{H}$ is to diagonalize Λ_n:

$$\Lambda_n |\phi_{n\alpha}) = E_{n\alpha} |\phi_{n\alpha}), \tag{2.7}$$

where $|\phi_{n\alpha})$ is a normalized eigenvector in $\mathfrak{H}_C$,

$$(\phi_{n\alpha} \mid \phi_{n\beta}) = \delta_{\alpha\beta}, \tag{2.8}$$

and $E_{n\alpha}$ is a c number eigenvalue. Note carefully that, because Λ_n stands to the right of ψ_n in (2.2),

$$|n, \alpha\rangle = \psi_n |\phi_{n\alpha}) \quad (\in \mathfrak{H}_B \otimes \mathfrak{H}_C) \tag{2.9}$$

is a normalized eigenvector of $\mathcal{H}$ with eigenvalue $E_{n\alpha}$.

The solution to the system of equations, (2.2) and (2.3), can be obtained at least formally in the framework of our perturbation expansion with respect to the interaction $\mathcal{H}_I$. In the zeroth-order approximation in which we ignore $\mathcal{H}_I$ completely, the solution is simply given by

$$\psi_n^{(0)} = |n), \quad \Lambda_n^{(0)} = W_n^{(0)} + \mathcal{H}_C, \tag{2.10}$$

where $W_n^{(0)}$ and $|n) \in \mathfrak{H}_B$ are an eigenvalue and the corresponding eigenvector of $\mathcal{H}_B$:

$$\mathcal{H}_B |n) = W_n^{(0)} |n). \tag{2.11}$$

Note that $\psi_n^{(0)}$ is just a vector in $\mathfrak{H}_B$ and has no operator character. The formulas for the perturbation corrections are obtained in Appendix A on the assumption that the eigenvalue problem (2.11) has been completely solved. It is worth noting that despite its operator character such an eigenvector of (2.2) as $\psi' = |n) + \sum_j |u_j) L_j$ can be normalized without difficulty when the operator part is small as compared with the zeroth-order c number vector $|n)$; the normalization factor (which is to be multiplied from the right) is obtained in the form of a power series in $L_j^{(\dagger)}$.

III. DIAGONALIZATION OF $\mathcal{H}_B$

Consider N identical bosons confined to a cubic box of volume V. Then, the "pair Hamiltonian" is defined by[15,25]

$$\begin{aligned}\mathcal{H}_P = {} & \sum_{\mathbf{k}} (k^2 - \mu) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + (2V)^{-1} v(0) \\ & \times \left(\sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}\right)\left[\left(\sum_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}}\right) - 1\right] \\ & + (2V)^{-1} \sum_{\mathbf{k},\mathbf{p}(\neq \pm\mathbf{k})} v(\mathbf{p} - \mathbf{k}) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \\ & + (2V)^{-1} \sum_{\mathbf{k},\mathbf{p}(\neq \mathbf{k})} v(\mathbf{p} - \mathbf{k}) a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger a_{\mathbf{p}} a_{-\mathbf{p}},\end{aligned} \tag{3.1}$$

where $v(\mathbf{k})$ is the Fourier transform of the interparticle potential, $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ are the usual boson destruction and creation operators, respectively, for plane wave states satisfying periodic boundary conditions with respect to V. The quantity μ is the chemical potential; it is introduced here because we have to use a representation in which the total particle number is not sharp. Our task is to find the ground state of $\mathcal{H}_P$ that satisfies the requirements (1.3)–(1.5).

On the basis of the discussion of Sec. I we replace the operators $a_0^{(\dagger)}$ by $N_0^{\frac{1}{2}} + c^{(\dagger)}$ assuming that the number of $\mathbf{k} = 0$ particles is macroscopic, $N_0 = O(N)$. The density of the condensate N_0/V is denoted by ρ_0. Then we use the method of Sec. II to split $\mathcal{H}_P$ into three parts:

$$\mathcal{H}_P = \mathcal{H}_B(a_{\mathbf{k}(\neq 0)}) + \mathcal{H}_C(c) + \mathcal{H}_I(a_{\mathbf{k}(\neq 0)}, c), \tag{3.2}$$

where the symbols in the brackets indicate the variables included in each part of the Hamiltonian. It is understood that $a_{\mathbf{k}}, c$ mean $a_{\mathbf{k}}^{(\dagger)}, c^{(\dagger)}$, respectively. Specifically,

$$\begin{aligned}\mathcal{H}_I = {} & V^{-1} B_{01} \sum_{\mathbf{k}}{}' [v(0) + v(\mathbf{k})] B_{\mathbf{k}1} \\ & + (2V)^{-1}\left[B_{02}^\dagger \sum_{\mathbf{k}}{}' v(\mathbf{k}) B_{\mathbf{k}2} + \text{H.c.}\right] - \mathcal{H}_C^R,\end{aligned} \tag{3.3}$$

where H.c. stands for Hermitian conjugate, and the renormalization term has the form

$$\mathcal{H}_C^R = \Delta f_0 c^\dagger c + \tfrac{1}{2}\Delta h_0 (cc + c^\dagger c^\dagger). \tag{3.4}$$

Moreover,

$$\begin{aligned} B_{01} &= a_0^\dagger a_0 - N_0 = c^\dagger c + N_0^{\frac{1}{2}}(c + c^\dagger), \\ B_{02} &= a_0 a_0 - N_0 = cc + 2N_0^{\frac{1}{2}} c, \end{aligned} \tag{3.5}$$

$$\begin{aligned} B_{\mathbf{k}1} &= a_{\mathbf{k}}^\dagger a_{\mathbf{k}} - \xi_{\mathbf{k}}, \\ B_{\mathbf{k}2} &= a_{\mathbf{k}} a_{-\mathbf{k}} - \eta_{\mathbf{k}}. \end{aligned} \quad (\mathbf{k} \neq 0), \tag{3.6}$$

The quantities Δf_0, Δh_0 and $\xi_{\mathbf{k}}$, $\eta_{\mathbf{k}}$ are c numbers to be determined later. Further $\mathcal{H}_B$ is given by[26]

$$\mathcal{H}_B = \mathcal{H}_B^0 + \mathcal{H}_B', \tag{3.7}$$

[25] We use a unit system $\hbar = 2m = 1$, where m is mass of the boson.

[26] The single prime on a double summation sign

$$\sum_{\mathbf{k},\mathbf{p}}{}'$$

means that terms for $\mathbf{k}$ and/or $\mathbf{p} = 0$ are to be omitted.

where

$$\mathcal{H}_B^0 = \overline{W}_0 + \sum_{\mathbf{k}}' \{[\chi + f(\mathbf{k})]a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \tfrac{1}{2}h(\mathbf{k})(a_{\mathbf{k}}a_{-\mathbf{k}} + a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger)\} \quad (3.8)$$

and

$$\mathcal{H}_B' = (2V)^{-1}v(0)\sum_{\mathbf{k},\mathbf{p}}' B_{\mathbf{k}1}^\dagger B_{\mathbf{p}1} + (2V)^{-1}\sum_{\mathbf{k},\mathbf{p}(\neq\pm\mathbf{k})}' v(\mathbf{p}-\mathbf{k})B_{\mathbf{k}1}^\dagger B_{\mathbf{p}1} + (2V)^{-1}\sum_{\mathbf{k},\mathbf{p}(\neq\mathbf{k})}' v(\mathbf{p}-\mathbf{k})B_{\mathbf{k}2}^\dagger B_{\mathbf{p}2}, \quad (3.9)$$

the symbols here being defined as

$$\overline{W}_0 = (2V)^{-1}v(0)\sum_{\mathbf{k},\mathbf{p}}' \xi_{\mathbf{k}}\xi_{\mathbf{p}} - (2V)^{-1}\sum_{\mathbf{k},\mathbf{p}(\neq\pm\mathbf{k})}' v(\mathbf{p}-\mathbf{k})\xi_{\mathbf{k}}\xi_{\mathbf{p}} - (2V)^{-1}\sum_{\mathbf{k},\mathbf{p}(\neq\mathbf{k})}' v(\mathbf{p}-\mathbf{k})\eta_{\mathbf{k}}\eta_{\mathbf{p}} + [\tfrac{1}{2}\rho_0 v(0) - \mu]N_0, \quad (3.10)$$

$$f(\mathbf{k}) = k^2 + \rho_0 v(\mathbf{k}) - V^{-1}\sum_{\mathbf{p}}' v(\mathbf{p})\eta_{\mathbf{p}} + V^{-1}\sum_{\mathbf{p}}' [v(\mathbf{p}-\mathbf{k}) - v(\mathbf{p})]\xi_{\mathbf{p}}, \quad (3.11)$$

$$h(\mathbf{k}) = \rho_0 v(\mathbf{k}) + V^{-1}\sum_{\mathbf{p}}' v(\mathbf{p}-\mathbf{k})\eta_{\mathbf{p}}. \quad (3.12)$$

Further, $\mathcal{H}_C$ is given by

$$\mathcal{H}_C = \mathcal{H}_C^0 + \mathcal{H}_C' + \mathcal{H}_C^R, \quad (3.13)$$

where

$$\mathcal{H}_C^0 + \mathcal{H}_C^R = \chi N_0^{\frac{1}{2}}(c + c^\dagger) + (\chi + f_0)c^\dagger c + \tfrac{1}{2}h_0(cc + c^\dagger c^\dagger), \quad (3.14)$$

$$\mathcal{H}_C' = V^{-1}N_0^{\frac{1}{2}}v(0)(c^\dagger cc + c^\dagger c^\dagger c) + (2V)^{-1}v(0)c^\dagger c^\dagger cc, \quad (3.15)$$

and

$$\chi = -\mu + \left(\rho_0 + V^{-1}\sum_{\mathbf{k}}' \xi_{\mathbf{k}}\right)v(0) + V^{-1}\sum_{\mathbf{k}}' v(\mathbf{k})(\xi_{\mathbf{k}} + \eta_{\mathbf{k}}).$$

Finally,

$$f_0 = f(0) + \Delta f_0, \qquad h_0 = h(0) + \Delta h_0, \quad (3.16)$$

where $f(0)$ and $h(0)$ are given by (3.11) and (3.12), respectively. It may be seen in the following that the condition (1.3) requires the chemical potential to be such that

$$\chi = 0. \quad (3.17)$$

At this point we note only that (3.17) removes the term linear in $c^{(\dagger)}$ from $\mathcal{H}_C$. Note that this linear term contains the large factor $N_0^{\frac{1}{2}}$. We take $\chi = 0$ throughout the following calculations.

Now, in order to carry out the diagonalization procedure as presented in the previous sections, we must first solve the eigenvalue problem for $\mathcal{H}_B$ and this forms the subject matter of the remainder of this section. For this purpose we can use a many-body perturbation theory[27] by taking (3.8) and (3.9) as unperturbed Hamiltonian and perturbation, respectively.

Let us first study the unperturbed Hamiltonian $\mathcal{H}_B^0$. It is convenient to use a new set of variables $\{\zeta_{\mathbf{k}\pm}, \pi_{\mathbf{k}\pm}\}$ which is obtained by two successive canonical transformations,[8]

$$q_{\mathbf{k}} = 2^{-\frac{1}{2}}i(a_{\mathbf{k}} - a_{\mathbf{k}}^\dagger), \qquad p_{\mathbf{k}} = 2^{-\frac{1}{2}}(a_{\mathbf{k}} + a_{\mathbf{k}}^\dagger), \quad (3.18)$$

and

$$\zeta_{\mathbf{k}+} = 2^{-\frac{1}{2}}(q_{\mathbf{k}} + q_{-\mathbf{k}}), \quad \pi_{\mathbf{k}+} = 2^{-\frac{1}{2}}(p_{\mathbf{k}} + p_{-\mathbf{k}}),$$
$$\zeta_{\mathbf{k}-} = 2^{-\frac{1}{2}}(p_{\mathbf{k}} - p_{-\mathbf{k}}), \quad \pi_{\mathbf{k}-} = -2^{-\frac{1}{2}}(q_{\mathbf{k}} - q_{-\mathbf{k}}).$$

The new variables satisfy the usual momentum-coordinate commutation relations,

$$[\pi_{\mathbf{k}+}, \zeta_{\mathbf{k}+}] = -i, \quad [\zeta_{\mathbf{k}+}, \zeta_{\mathbf{k}-}] = 0,$$
$$[\pi_{\mathbf{k}+}, \zeta_{\mathbf{k}-}] = 0, \quad \text{etc.}$$

Then $\mathcal{H}_B^0$ takes on the form of the standard harmonic oscillator Hamiltonian,

$$\mathcal{H}_B^0 = -\sum_{\frac{1}{2}[\mathbf{k}]} f(\mathbf{k}) + \tfrac{1}{2}\sum_{\frac{1}{2}[\mathbf{k}]}' [M_{\mathbf{k}}^{-1}(\pi_{\mathbf{k}+}^2 + \pi_{\mathbf{k}-}^2) + \lambda_{\mathbf{k}}(\zeta_{\mathbf{k}+}^2 + \zeta_{\mathbf{k}-}^2)], \quad (3.19)$$

where $\frac{1}{2}[\mathbf{k}]$ under the summation symbol means that the allowed momenta $\mathbf{k}$ should belong to a half-space only, $k_z > 0$, say. We can now see clearly that $\mathcal{H}_B^0$ has a normalizable ground state if and only if

$$M_{\mathbf{k}}^{-1} \equiv f(\mathbf{k}) + h(\mathbf{k}) > 0, \quad (3.20)$$

$$\lambda_{\mathbf{k}} \equiv f(\mathbf{k}) - h(\mathbf{k}) > 0. \quad (3.21)$$

These are what we call the $\mathbf{k} \neq 0$-stability conditions. How these conditions restrict the interaction potential $v(\mathbf{k})$ is studied in Sec. VI. For the moment, let us assume that the conditions are satisfied. Under this assumption, $\mathcal{H}_B^0$ can readily be diagonalized by a Bogoliubov transformation,

$$a_{\mathbf{k}} = \alpha_{\mathbf{k}}b_{\mathbf{k}} + \beta_{\mathbf{k}}b_{-\mathbf{k}}^\dagger, \quad (3.22)$$

where $b_{\mathbf{k}}^{(\dagger)}$ are new boson operators and their commutation relations are secured by requiring the real c number coefficients $\alpha_{\mathbf{k}}$, $\beta_{\mathbf{k}}$ to satisfy

$$\alpha_{\mathbf{k}}^2 - \beta_{\mathbf{k}}^2 = 1. \quad (3.23)$$

[27] J. Goldstone, Proc. Roy. Soc. (London) A239, 267 (1957); C. Bloch, Nucl. Phys. 7, 451 (1958).

If we choose the coefficients such that[28]

$$\alpha_{\mathbf{k}}^2 = \tfrac{1}{2}\{[f(\mathbf{k})/\epsilon(\mathbf{k})] + 1\},$$
$$\alpha_{\mathbf{k}}\beta_{\mathbf{k}} = -\tfrac{1}{2}h(\mathbf{k})/\epsilon(\mathbf{k}), \tag{3.24a}$$

where

$$\epsilon(\mathbf{k}) = [f^2(\mathbf{k}) - h^2(\mathbf{k})]^{\frac{1}{2}}, \tag{3.24b}$$

then $\mathcal{H}_B^0$ is diagonalized in the new representation, that is

$$\mathcal{H}_B^0 = W_0 + \sum_{\mathbf{k}}{}' \epsilon(\mathbf{k}) b_{\mathbf{k}}^\dagger b_{\mathbf{k}}. \tag{3.25}$$

The quantity $\epsilon(\mathbf{k})$ is real as a consequence of the assumed $k \neq 0$ stability conditions (3.20) and (3.21).[29] The ground state is given by the no-particle state $|0\}$ of $b_{\mathbf{k}}$,

$$b_{\mathbf{k}}|0\} = 0 \quad (\text{all } \mathbf{k} \neq 0), \tag{3.26}$$

and its energy is

$$W_0 = \overline{W}_0 + \tfrac{1}{2}\sum_{\mathbf{k}}{}' [\epsilon(\mathbf{k}) - f(\mathbf{k})]. \tag{3.27}$$

The motivation for the separation (3.7) is that if the $\xi_{\mathbf{k}}$ and $\eta_{\mathbf{k}}$ are chosen as

$$\xi_{\mathbf{k}} = \{0|\, a_{\mathbf{k}}^\dagger a_{\mathbf{k}}\, |0\},$$
$$\eta_{\mathbf{k}} = \{0|\, a_{\mathbf{k}} a_{-\mathbf{k}}\, |0\}, \tag{3.28}$$

then we can show that in the volume limit the Hamiltonian $\mathcal{H}_B$ is essentially equivalent[30] to $\mathcal{H}_B^0$ in the following sense: (1) The eigenvalue spectra are identical, and (2) the expectation values of quantities of $O(1)$ with respect to the corresponding eigenstates are equal in the volume limit. The proof of this claim is given in succeeding paragraphs. Using (3.22)–(3.24) the matrix elements in (3.28) are easily found to be

$$\xi_{\mathbf{k}} = \beta_{\mathbf{k}}^2 = \tfrac{1}{2}\{[f(\mathbf{k})/\epsilon(\mathbf{k})] - 1\},$$
$$\eta_{\mathbf{k}} = \alpha_{\mathbf{k}}\beta_{\mathbf{k}} = -\tfrac{1}{2}h(\mathbf{k})/\epsilon(\mathbf{k}). \tag{3.29}$$

Now we substitute these equations into (3.11) and (3.12) to obtain a set of coupled integral equations for $f(\mathbf{k})$ and $h(\mathbf{k})$:

$$(\mathbf{k}) = k^2 + \rho_0 v(\mathbf{k}) + (2V)^{-1}\sum_{\mathbf{p}}{}' [v(\mathbf{p} - \mathbf{k}) - v(\mathbf{p})] \times \{[f(\mathbf{p})/\epsilon(\mathbf{p})] - 1\} + (2V)^{-1}\sum_{\mathbf{p}}{}' v(\mathbf{p})h(\mathbf{p})/\epsilon(\mathbf{p}), \tag{3.30}$$

$$h(\mathbf{k}) = \rho_0 v(\mathbf{k}) - (2V)^{-1}\sum_{\mathbf{p}}{}' v(\mathbf{p} - \mathbf{k})h(\mathbf{p})/\epsilon(\mathbf{p}). \tag{3.31}$$

In these equations we have set $\chi = 0$ in accordance with (3.17).

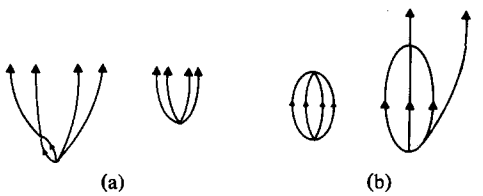
We turn now to the proof of the asymptotic equivalence of $\mathcal{H}_B$ and $\mathcal{H}_B^0$. According to Goldstone[27] the ground state of the total Hamiltonian $\mathcal{H}_B$ can be written as

$$|\Omega_B) = Z^{\frac{1}{2}}\sum_{n=0}^{\infty}[(W_0 - \mathcal{H}_B^0)^{-1}\mathcal{H}_B']_{\text{linked}}^n\, |0\}, \tag{3.32}$$

where $Z^{\frac{1}{2}}$ is a normalization factor and $|0\}$ the unperturbed ground state (3.26). It should be observed here that (3.6) and (3.28) make the operators $B_{\mathbf{k}i}$ the normal products in the b representation as defined by (3.22) and (3.26). Namely, $\{0\,|B_{\mathbf{k}i}|\,0\} = 0$, $(i = 1, 2)$. The vectors $[(W_0 - \mathcal{H}_B^0)^{-1}\mathcal{H}_B']^n\,|0\}$ can be represented in a well-known way by diagrams like those of Figs. 1(a) and 1(b). If every part of a diagram is linked to some external lines, then we call it a linked diagram. If, on the contrary, a diagram has some isolated bubbles, then such a diagram is said to be unlinked. The diagrams in Fig. 1(a) and in Fig. 1(b) are linked and unlinked, respectively. The suffix "linked" in (3.32) means that we should take only those terms which are represented by linked diagrams; this is because all terms represented by unlinked diagrams can be absorbed by the normalization factor $Z^{\frac{1}{2}}$.

The pair Hamiltonian has some special features that greatly assist us in calculating its energy spectrum and various matrix elements with respect to its eigenstates. First of all, the perturbation term $\mathcal{H}_B'$ is made up of the pair operators $B_{\mathbf{k}i}$, so that for any term of the perturbation series the particles in the intermediate states or the final state can be paired off according to their momenta. Now consider a diagram that contributes to (3.32) a term,

$$|u) = \sum_{\mathbf{k}_1,\cdots,\mathbf{k}_s} V^{-\lambda} A(\mathbf{k}_1, \cdots, \mathbf{k}_s) b_{\mathbf{k}_1}^\dagger b_{-\mathbf{k}_1}^\dagger \cdots b_{\mathbf{k}_s}^\dagger b_{-\mathbf{k}_s}^\dagger\, |0\}, \tag{3.33}$$



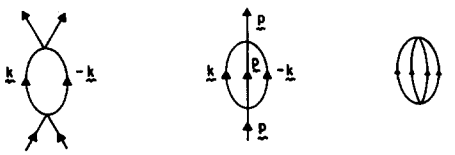
(a) (b)

FIG. 1. (a) Linked diagram. (b) Unlinked diagram.

[28] The stability conditions can be obtained without using the canonical transformation to the standard form of the harmonic oscillator Hamiltonian (3.19). After the Bogoliubov transformation (3.22), $\mathcal{H}_B^0$ becomes a sum of $\epsilon(\mathbf{k})b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + \frac{1}{2}h'(\mathbf{k})(b_{\mathbf{k}}b_{-\mathbf{k}} + b_{\mathbf{k}}^\dagger b_{-\mathbf{k}}^\dagger)$ over $\mathbf{k}$, where $\epsilon(\mathbf{k}) = (\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2)f(\mathbf{k}) + 2\alpha_{\mathbf{k}}\beta_{\mathbf{k}}h(\mathbf{k})$, $h'(\mathbf{k}) = 2\alpha_{\mathbf{k}}\beta_{\mathbf{k}}f(\mathbf{k}) + (\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2)h(\mathbf{k})$. Now, the equation $h'(\mathbf{k}) = 0$ has real solutions for $\alpha_{\mathbf{k}}, \beta_{\mathbf{k}}$ if and only if $f(\mathbf{k}) \pm h(\mathbf{k})$ have the same sign. But, if they were both negative, then $f(\mathbf{k}) < 0$, which implies $\epsilon(\mathbf{k}) < 0$, because by using $h'(\mathbf{k}) = 0$ to eliminate $\alpha_{\mathbf{k}}\beta_{\mathbf{k}}$, we get $\epsilon(\mathbf{k}) = (\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2)[f(\mathbf{k})^2 - h(\mathbf{k})^2]/f(\mathbf{k})$. Thus we reach the stability condition. At the same time we know that the sign of $\alpha_{\mathbf{k}}\beta_{\mathbf{k}}$ must be opposite to the sign of $h(\mathbf{k})/f(\mathbf{k})$.

[29] When $f(\mathbf{k}) > 0$, $\alpha_{\mathbf{k}}\beta_{\mathbf{k}}$ has the sign of $-h(\mathbf{k})$. See footnote 28.

[30] There is no surprise in this equivalence. A beautiful argument has been given for the case of pairwise interacting fermions (BCS model) by R. Haag, Nuovo Cimento **25**, 287 (1962); see also H. Ezawa, J. Math. Phys. **5**, 1078 (1964); H. Umezawa, Y. Takahashi, and S. Kamefuchi, Ann. Phys. (N.Y.) **26**, 336 (1964).



(a) (b) (c)

FIG. 2. Miscellaneous contractions appearing in the expansion (3.32) of the state vector $|\Omega_B)$.

where A is a certain function of the s external momenta *not* involving any volume factor; the volume dependence is factored out as $V^{-\lambda}$ in front. It is not difficult to see that the exponent of the volume factor should satisfy an inequality,

$$\lambda \geq \tfrac{1}{2}s. \tag{3.34}$$

In fact, if the diagram involves no internal lines we have the equality $\lambda = \frac{1}{2}s$, because there is a factor V^{-1} for each product of four b operators, or $B^\dagger_{ki}\, B_{pj}$ in (3.9). Now if in such a diagram we apply a contraction of the form shown in Fig. 2(a), $\Delta S = 2$ pairs of external lines are removed whereas the summation over the internal momentum contributes a factor $V^{\Delta\lambda}$, $\Delta\lambda = 1$. Therefore such a contraction changes λ into $\lambda' = \lambda - 1$ and s into $s' = s - 2$, so that the equality in (3.34) remains valid: $\lambda' = \frac{1}{2}s'$. For the contraction shown in Fig. 2(b) we have $\Delta s = 3$, $\Delta\lambda = 1$ so that $\lambda' = \frac{1}{2}(s' + 1) \geq \frac{1}{2}s'$ in accordance with (3.34). A contraction of the type shown in Fig. 2(c) need not be considered since it is an unlinked diagram. Continuing the process of adding contractions, we can verify (3.34) for all diagrams involved in the perturbation series (3.32). As a consequence of (3.34) we can conclude that the norm $\|u\|$ of the vector (3.33) is of the order one or smaller[31] as $V \to \infty$.

With the aid of (3.34) we can now prove that in the asymptotic sense ($V \to \infty$) the excited states of $\mathcal{H}_B$ are given simply by

$$|\mathbf{p}_1, \cdots, \mathbf{p}_r) = b^\dagger_{\mathbf{p}_1} \cdots b^\dagger_{\mathbf{p}_r} |\Omega_B) \quad (r = 1, 2, \cdots), \tag{3.35}$$

and the corresponding energies are $W_0 + \sum_{i=1}^{r} \epsilon(\mathbf{p}_i)$. Note that this eigenvalue spectrum is the same as that of $\mathcal{H}^0_B$. The proof begins with the observation that the vector

$$|R_1) = [\mathcal{H}'_B, b_{\mathbf{p}}]\, |\Omega_B) \tag{3.36}$$

has a vanishing norm as $V \to \infty$:

$$\|R_1\|^2 = O(V^{-1}). \tag{3.37}$$

[31] This is not the case for the full Hamiltonian because its interaction term

$$\sum_{\mathbf{k},\mathbf{p},\mathbf{q}} v(\mathbf{k}) a^\dagger_{\mathbf{p}+\mathbf{k}} a^\dagger_{\mathbf{q}-\mathbf{k}} a_{\mathbf{p}} a_{\mathbf{q}}$$

involves three independent momenta which are to be summed.

In fact, the commutator applied on a constituent of $|\Omega_B)$,

$$[\mathcal{H}'_B, b^\dagger_{\mathbf{p}}]\, |u) = (2V)^{-1} v(0) \sum_{\mathbf{k}}{}' B^\dagger_{\mathbf{k}1} [B_{\mathbf{p}1}, b^\dagger_{\mathbf{p}}]\, |u) + \cdots$$

increases only one momentum summation at the cost of one volume factor V^{-1}. Hence we get (3.37).[32] By a similar argument we can also prove that

$$\|R_2\|^2 = O(V^{-1}), \quad |R_2) = b_{\mathbf{p}}\, |\Omega_B), \tag{3.38}$$

and therefore that $b^\dagger_{\mathbf{p}}\, |\Omega_B)$ is asymptotically normalized:

$$\|b^\dagger_{\mathbf{p}}\, |\Omega_B)\|^2 = 1 + O(V^{-1}). \tag{3.39}$$

We can now conclude that, as $V \to \infty$ asymptotically, $b^\dagger_{\mathbf{p}}\, |\Omega_B)$ is an eigenstate of $\mathcal{H}_B$, because

$$\begin{aligned} \mathcal{H}_B b^\dagger_{\mathbf{p}}\, |\Omega_B) &= b^\dagger_{\mathbf{p}} \mathcal{H}_B\, |\Omega_B) + [\mathcal{H}^0_B, b^\dagger_{\mathbf{p}}]\, |\Omega_B) + |R_1) \\ &= [W_0 + \epsilon(\mathbf{p})] b^\dagger_{\mathbf{p}}\, |\Omega_B) + |R_1), \end{aligned}$$

and the residual $|R_1)$ has a vanishing norm by (3.37). Repeating a similar argument, we can verify the above statement concerning (3.35). A generalization of (3.38) shows that it is impossible to create a state lying lower than $|\Omega_B)$ by applying the destruction operators $b_{\mathbf{p}}$ on $|\Omega_B)$. From this result we know that the set of vectors of the form (3.35) is asymptotically complete as a basis for the cyclic representation[9] of $b^{(\dagger)}_{\mathbf{k}}$.

Let us turn to the calculation of various matrix elements. We have the formulas

$$(\Omega_B|\, B_{\mathbf{k}i}\, |\Omega_B) = O(V^{-1}), \tag{3.40}$$

$$(\mathbf{p}, -\mathbf{p}|\, B_{\mathbf{k}i}\, |\Omega_B) = \{\mathbf{p}, -\mathbf{p}|\, B_{\mathbf{k}i}\, |0\} + O(V^{-1}), \quad \text{etc.}, \tag{3.41}$$

where

$$|\mathbf{p}, -\mathbf{p}\} = b^\dagger_{\mathbf{p}} b^\dagger_{-\mathbf{p}}\, |0\}. \tag{3.42}$$

To prove (3.41), for example, we have only to observe that the matrix element can involve two kinds of contractions. Namely, the momentum $\mathbf{k}$ of $B_{\mathbf{k}i}$ can be contracted either with the momenta of the virtual pairs contained in the ground state $|\Omega_B)$ or with the momentum $\mathbf{p}$ of the real pair. For the same reason that we obtain (3.38), the former contraction yields only a vanishing contribution of $O(V^{-1})$. The latter leads to $\{\mathbf{p}, -\mathbf{p}|\, B_{\mathbf{k}i}\, |0\}$, which is the main term of (3.41). In general, nonvanishing contributions to any matrix elements $(\psi'|\, B_{\mathbf{p}i} \cdots B_{\mathbf{q}j}\, |\psi)$ come only from contractions of $\mathbf{p}, \cdots, \mathbf{q}$ with the momenta of real pairs involved in the states $|\psi)$ and $|\psi')$. This completes the desired proof that $\mathcal{H}_B$ is asymptotically

[32] This conclusion remains valid if we consider a wave packet state,

$$|f) = \sum_{\mathbf{p}} V^{-\frac{1}{2}} f(\mathbf{p}) b^\dagger_{\mathbf{p}}\, |\Omega_B)$$

for which $\int |f|^2\, d\mathbf{p} < \infty$.

equivalent to $\mathcal{H}_B^0$, and the diagonalization of $\mathcal{H}_B$ is thereby accomplished in the asymptotic sense.

IV. EFFECTIVE CONDENSATE HAMILTONIAN

We are now in the position to construct the effective Hamiltonian Λ_0 for the condensate fluctuation which is associated with the ground state of the particles outside the condensate by using the modified Born–Oppenheimer method of Sec. II. The present calculation makes use of our perturbation method through second order (see Appendix A). The result for Λ_0 actually incorporates the ladder-type corrections which occur in the higher orders of conventional perturbation theory. In the case of the pair Hamiltonian model, which we are going to study in this section, we have good reason to believe that the present result for Λ_0 is actually correct through all orders of our perturbation theory in the volume limit. In fact, the excitation spectrum of the condensate fluctuation as well as the ground-state expectation values of products of the operators $c^{(\dagger)}$ can be calculated in an asymptotically exact fashion by making use of a Green function approach, and all the results therefrom will prove to be in complete agreement with the results from our second-order perturbation calculations. The Green function approach and our method of perturbation are complementary to each other: The former, though the more convenient to handle, rests upon the assumption that the Hamiltonian $\mathcal{H}_P$ has a bona fide ground state in the representation based upon the operator replacement (1.1), while the latter is useful for studying this assumption.

A. Modified Born–Oppenheimer Approach

The total pair-model Hamiltonian $\mathcal{H}_P$ is given by (3.2). For the purpose of applying the perturbation method as developed in Sec. II and Appendix A, it proves convenient to subdivide $\mathcal{H}_I$, the interaction part of $\mathcal{H}_P$, as

$$\mathcal{H}_I = \mathcal{H}_{\text{int}} - \mathcal{H}_C^R, \quad \mathcal{H}_{\text{int}} = \mathcal{H}_I^0 + \mathcal{H}_I', \tag{4.1}$$

where

$$\mathcal{H}_I^0 = N_0^{\frac{1}{2}} V^{-1}(c + c^\dagger) \sum_{\mathbf{k}}{}' [v(0) + v(\mathbf{k})] B_{\mathbf{k}1} + N_0^{\frac{1}{2}} V^{-1}\Big[c \sum_{\mathbf{k}}{}' v(\mathbf{k}) B_{\mathbf{k}2}^\dagger + \text{H.c.}\Big], \tag{4.2}$$

$$\mathcal{H}_I' = V^{-1} c^\dagger c \sum_{\mathbf{k}}{}' [v(0) + v(\mathbf{k})] B_{\mathbf{k}1} + (2V)^{-1}\Big[cc \sum_{\mathbf{k}}{}' v(\mathbf{k}) B_{\mathbf{k}2}^\dagger + \text{H.c.}\Big]. \tag{4.3}$$

It may be seen that $\mathcal{H}_C' + \mathcal{H}_I'$ is negligible as long as $c = o(N^{\frac{1}{2}})$, in particular when the $\mathbf{k} = 0$ stability conditions (1.14) are satisfied.

Let us calculate the effective condensate Hamiltonian Λ_0 associated with the ground state (3.32) of the $\mathbf{k} \neq 0$ particles. We treat $H_1 = \mathcal{H}_{\text{int}}$ and $H_2 = -\mathcal{H}_C^R$ as small perturbations of first and second order, respectively.[33] As is well known, this procedure leads to a low-density expansion[4]; the expansion parameter is $(\rho_0 a^3)^{\frac{1}{2}}$ with a being the zero energy scattering length of the interparticle potential v. According to (A16) the zeroth-order condensate Hamiltonian is

$$\Lambda_0^{(0)} = W_0 + \mathcal{H}_C, \tag{4.4}$$

and the formula (A17) gives the first-order correction,

$$\Lambda_0^{(1)} = (\Omega_B|\, \mathcal{H}_{\text{int}}\, |\Omega_B) = O(N^{-\frac{1}{2}} c, N^{-1} c^2), \tag{4.5}$$

where use has been made of (3.40). It is understood that c in the Landau symbol stands for the operators $c^{(\dagger)}$. In order to calculate the second-order correction by (A18),[34] i.e.,

$$\Lambda_0^{(2)} = \text{H.P.}\,(\Omega_B|\, \mathcal{H}_{\text{int}}\, |\psi_0^{(1)}) - (\Omega_B|\, \mathcal{H}_C^R\, |\Omega_B), \tag{4.6}$$

we have to know the first-order correction $\psi_0^{(1)}$ to the "wavefunction" of the $\mathbf{k} \neq 0$ particles. Inserting a complete set of intermediate states we can write the matrix element in the first term of (4.6) as

$$(\Omega_B|\, \mathcal{H}_{\text{int}}\, |\psi_0^{(1)}) = \sum_n (\Omega_B|\, \mathcal{H}_{\text{int}}\, |n)(n \mid \psi_0^{(1)}). \tag{4.7}$$

Due to the particular features (3.40)–(3.41) of the pair Hamiltonian, the only intermediate states significant in the volume limit are $|\mathbf{k}, -\mathbf{k})$; for notation see (3.35).[35] Using (3.41) we find

$$(\mathbf{k}, -\mathbf{k}|\, \mathcal{H}_I^0\, |\Omega_B) = (2N_0^{\frac{1}{2}}/V)[\gamma_+(\mathbf{k}) c + \gamma_-(\mathbf{k}) c^\dagger] + O(N^{-\frac{3}{2}} c), \tag{4.8}$$

$$(\mathbf{k}, -\mathbf{k}|\, \mathcal{H}_I'\, |\Omega_B) = O(N^{-2} c^2), \tag{4.9}$$

where

$$\gamma_+(\mathbf{k}) = [v(0) + v(\mathbf{k})]\alpha_{\mathbf{k}}\beta_{\mathbf{k}} + v(\mathbf{k})\alpha_{\mathbf{k}}^2, \quad \gamma_-(\mathbf{k}) = [v(0) + v(\mathbf{k})]\alpha_{\mathbf{k}}\beta_{\mathbf{k}} + v(\mathbf{k})\beta_{\mathbf{k}}^2. \tag{4.10}$$

Note that

$$\gamma_+(\mathbf{k}) - \gamma_-(\mathbf{k}) = v(\mathbf{k}). \tag{4.11}$$

The first-order correction $\psi_0^{(1)}$ to the state vector is obtained by solving the operator equation (A19). Let us momentarily ignore the "small" quantities $\mathcal{H}_C'$

[33] One may start with λH_1 and $\lambda^2 H_2$, by employing a parameter λ and putting $\lambda = 1$ after the perturbation calculation. The renormalization counter term is then necessarily $O(\lambda^2)$ if one takes λH_1 for the interaction.

[34] H.P. stands for "Hermitian part of."

[35] The summation over momenta $\mathbf{k}$ of the intermediate states $|\mathbf{k}, -\mathbf{k})$ should be restricted to a half of the momentum space so as to avoid double counting.

and the matrix element (4.9) of $\mathcal{H}_I'$. Then substitute the trial form

$$(\mathbf{k}, -\mathbf{k} \mid \psi_0^{(1)}) = (2N_0^{\frac{1}{2}}/V)(x_\mathbf{k} + y_\mathbf{k} c + z_\mathbf{k} c^\dagger) \quad (4.12)$$

into (A19), where $x_\mathbf{k}$, $y_\mathbf{k}$, and $z_\mathbf{k}$ are c number unknowns. On comparing the coefficients of c, $c^\dagger$ and the c number term on both sides of (A19), we get a system of simultaneous equations for those unknown coefficients, which gives

$$\begin{aligned} x_\mathbf{k} &= 0, \\ y_\mathbf{k} &= [4\epsilon^2(\mathbf{k}) - \omega_0^2]^{-1}\{[-2\epsilon(\mathbf{k}) - f_0]\gamma_+(\mathbf{k}) + h_0\gamma_-(\mathbf{k})\}, \\ z_\mathbf{k} &= [4\epsilon^2(\mathbf{k}) - \omega_0^2]^{-1}\{[-2\epsilon(\mathbf{k}) + f_0]\gamma_-(\mathbf{k}) - h_0\gamma_+(\mathbf{k})\}, \end{aligned} \quad (4.13)$$

where

$$\omega_0^2 = f_0^2 - h_0^2. \quad (4.14)$$

Then, the first term of (4.6) turns out to be

$$\text{H.P.}\,(\Omega_B|\,\mathcal{H}_I\,|\psi_0^{(1)}) = \delta f_0 c^\dagger c + \tfrac{1}{2}\delta h_0(cc + c^\dagger c^\dagger) + O(N^{-\frac{1}{2}}c, N^{-1}c^2, \cdots), \quad (4.15)$$

where

$$\begin{aligned} \delta f_0 &= 2\rho_0 V^{-1} \sum_\mathbf{k}{}' [\gamma_+(\mathbf{k}) y_\mathbf{k} + \gamma_-(\mathbf{k}) z_\mathbf{k}], \\ \delta h_0 &= 2\rho_0 V^{-1} \sum_\mathbf{k}{}' [\gamma_+(\mathbf{k}) z_\mathbf{k} + \gamma_-(\mathbf{k}) y_\mathbf{k}]. \end{aligned} \quad (4.16)$$

If we had included $\mathcal{H}_C'$ and $\mathcal{H}_I'$ in the calculation of $\psi_0^{(1)}$, we would have obtained additional terms in (4.15), some of which are of higher degree in the operators $c^{(\dagger)}$ and are denoted by the dots in the Landau symbol. They are negligible if $c = o(N^{\frac{1}{2}})$. Thus (4.6) becomes

$$\Lambda_0^{(2)} = O(N^{-\frac{1}{2}}c, \cdots) \quad (4.17)$$

if we choose the renormalization coefficients Δf_0 and Δh_0 such that

$$\Delta f_0 = \delta f_0, \quad \Delta h_0 = \delta h_0. \quad (4.18)$$

Combining the results (4.3), (4.4), and (4.17) we get

$$\Lambda_0 = \Lambda_0^0 + \Lambda_0', \quad (4.19)$$

where

$$\Lambda_0^0 = W_0 + f_0 c^\dagger c + \tfrac{1}{2}h_0(cc + c^\dagger c^\dagger), \quad (4.20)$$

$$\Lambda_0' = V^{-1}N_0^{\frac{1}{2}}v(0)(c^\dagger cc + c^\dagger c^\dagger c) + (2V)^{-1}v(0)c^\dagger c^\dagger cc + O(N^{-\frac{1}{2}}c, N^{-1}c^2, \cdots), \quad (4.21)$$

and the coefficients f_0 and h_0 of Λ_0^0, given by (3.16), (4.16), and (4.18), are

$$\begin{aligned} f_0 &= f(0) + 2\rho_0 V^{-1} \sum_\mathbf{k}{}' [\gamma_+(\mathbf{k}) y_\mathbf{k} + \gamma_-(\mathbf{k}) z_\mathbf{k}], \\ h_0 &= h(0) + 2\rho_0 V^{-1} \sum_\mathbf{k}{}' [\gamma_+(\mathbf{k}) z_\mathbf{k} + \gamma_-(\mathbf{k}) y_\mathbf{k}]. \end{aligned} \quad (4.22)$$

But this is a system of coupled linear equations because $y_\mathbf{k}$ and $z_\mathbf{k}$ depend linearly on f_0 and h_0 through (4.13). Decoupling is immediate if we form $f_0 \pm h_0$. Then,

$$f_0 - h_0 = 2D^{-1}\left\{-\frac{1}{V}\sum_\mathbf{k}{}' v(\mathbf{k})\alpha_\mathbf{k}\beta_\mathbf{k} - \frac{\rho_0}{V}\sum_\mathbf{k}{}' \frac{2\epsilon(\mathbf{k})v^2(\mathbf{k})}{4\epsilon^2(\mathbf{k}) - \omega_0^2}\right\}, \quad (4.23)$$

$$f_0 + h_0 = 2D^{-1}\left\{\rho_0 v(0) - \frac{\rho_0}{V}\sum_\mathbf{k}{}' \frac{2\epsilon(\mathbf{k})}{4\epsilon^2(\mathbf{k}) - \omega_0^2} \times [\gamma_+(\mathbf{k}) + \gamma_-(\mathbf{k})]^2\right\}, \quad (4.24)$$

where use has been made of (4.11), and

$$D = 1 + \frac{2\rho_0}{V}\sum_\mathbf{k}{}' v(\mathbf{k}) \frac{\gamma_+(\mathbf{k}) + \gamma_-(\mathbf{k})}{4\epsilon^2(\mathbf{k}) - \omega_0^2}. \quad (4.25)$$

This completes the calculation of the effective condensate Hamiltonian through the second order of our perturbation scheme. The presence of the denominator D in (4.23) and (4.24) shows, however, that some higher-order effects are already taken into account.

The first term Λ_0^0 of (4.19) has exactly the same form as (1.9), whose properties depend on the sign of $f_0 \pm h_0$. When they are both positive,

$$f_0 \pm h_0 > 0, \quad (4.26)$$

so as to satisfy the $\mathbf{k} = 0$ stability condition (1.14), Λ_0^0 is a harmonic oscillator Hamiltonian having a normalizable ground state $|\Omega_C)$. Since the coefficients f_0 and h_0 are $O(1)$, it follows that the matrix elements $(\Omega_C|\,(c^\dagger)^i c^j\,|\Omega_C)$ are all $O(1)$ $(i, j = 0, 1, 2, \cdots)$, verifying the statement that $c = O(1)$, which is stronger than our presupposition $c = o(N^{\frac{1}{2}})$. Further, when (4.26) is satisfied the second term Λ_0' of (4.19), if treated as a perturbation to Λ_0^0, has no effects in the volume limit. In the asymptotic sense, therefore,

$$\Lambda_0 = \Lambda_0^0, \quad \text{(stable case, } V \to \infty). \quad (4.27)$$

The physical meaning of the ω_0 in (4.14) is now clear: it is the frequency of the harmonic oscillator (4.27). Thus in this approximation the excitation energies of the condensate fluctuation are integer multiples of the energy quantum ω_0. The Green function method discussed in the next subsection tells us that this excitation spectrum is asymptotically exact.

In Secs. V and VI we study how the stability conditions (3.20)–(3.21) and (4.26) restrict the form of the potential v and the particle density ρ. We find that there actually exist some potentials that can satisfy all these conditions. In Appendix B we study cases where some of the conditions are violated.

A remark is in order about the denominator $4\epsilon^2(\mathbf{k}) - \omega_0^2$ in (4.23)–(4.25). Clearly we have to require that

$$[2\epsilon(\mathbf{k})]^2 > \omega_0^2, \quad (\text{all } \mathbf{k} \neq 0). \tag{4.28}$$

Physically speaking the role of the inequality (4.28) is to prohibit a condensate excitation from decaying into two $\mathbf{k} \neq 0$-excitations. Here we wish to show that it is a consistent assumption to assume that (4.28) is satisfied if

$$D > \tfrac{1}{2}. \tag{4.29}$$

By consistency we mean that if (4.29) and the $\mathbf{k} \neq 0$ stability conditions, (3.20) and (3.21), are *assumed* satisfied—which in turn requires that (4.28) holds—then in fact (4.28) is satisfied. For this purpose, it suffices to show that $[2\epsilon(0)]^2 > \omega_0^2$ since, according to (3.30), $f(\mathbf{k})$ contains the rapidly increasing term k^2. Now it follows from (4.23) and (4.24) that

$$f_0 - h_0 < -2D^{-1}\frac{1}{V}\sum_{\mathbf{k}}{}' v(\mathbf{k})\alpha_{\mathbf{k}}\beta_{\mathbf{k}} = [f(0) - h(0)]D^{-1}, \tag{4.30}$$

$$f_0 + h_0 < 2D^{-1}\rho_0 v(0) = [f(0) + h(0)]D^{-1}, \tag{4.31}$$

and thus

$$\omega_0^2 < \epsilon(0)^2 D^{-2} < [2\epsilon(0)]^2. \tag{4.32}$$

B. Method of Green Functions

We now proceed to show that our second-order approximation (4.27) gives the complete effective condensate Hamiltonian in the volume limit. The argument is based on the observation that the excitation spectrum of the condensate fluctuation as obtained from (4.27) agrees exactly with the corresponding result from our Green function treatment, which is asymptotically exact as $V \to \infty$. To prove this result within the approach described in Subsection A would be very difficult if not impossible.

In almost all works on low-temperature boson systems the Green function method has been formulated *after* one makes the Bogoliubov replacement $a_0^{(\dagger)} \to N_0^{\frac{1}{2}}$. Recently, Popov and Faddeev[36] proposed that one use the operator replacement (1.1) so as to include the condensate fluctuation within the Green function formalism. All the conventional graph techniques can readily be adapted to the new situation. The only necessary addition to the conventional set of prescriptions is that the chemical potential should be determined so as to eliminate the tadpole diagrams due to c and $c^\dagger$; this condition is equivalent to (1.3).

[36] V. N. Popov and L. D. Faddeev, Zh. Eksperim. i Teor. Fiz. **47**, 1315 (1964) [English transl.: Soviet Phys.—JETP **20**, 890 (1965)].

Now we define one-particle Green functions for the condensate fluctuation by

$$\begin{aligned} g(\omega) &= \mathcal{F}_t \langle\Omega|\, T[c(t), c^\dagger(0)]\, |\Omega\rangle, \\ \hat{g}(\omega) &= \mathcal{F}_t \langle\Omega|\, T[c^\dagger(t), c^\dagger(0)]\, |\Omega\rangle, \end{aligned} \tag{4.33}$$

where $|\Omega\rangle$ is the ground state (assumed to exist) of the pair Hamiltonian $\mathcal{H}_P$ in the representation based upon (1.1), T denotes the Wick chronological operator, $c(t)$ is the Heisenberg operator

$$c(t) = \exp(i\mathcal{H}_P t)c \exp(-i\mathcal{H}_P t),$$

and $\mathcal{F}_t$ denotes the Fourier transform in time

$$\mathcal{F}_t A(t) = (2\pi)^{-1}\int_{-\infty}^{\infty} A(t)\exp(i\omega t)\, dt. \tag{4.34}$$

In order to carry out a perturbation calculation for g and $\hat{g}$ we use[37]

$$H_0 = \sum_{\mathbf{k}}{}' \epsilon(\mathbf{k}) b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + f(0)c^\dagger c \tag{4.35}$$

as the free Hamiltonian to define the interaction picture

$$\tilde{\mathcal{O}}(t) = \exp(iH_0 t)\mathcal{O}\exp(-iH_0 t),$$

where $\mathcal{O}$ is a Schrödinger operator. The rest of the Hamiltonian, $\mathcal{H}_P - H_0$, is regarded as a perturbation. Thus the unperturbed $\mathbf{k} \neq 0$ Green functions are given simply by

$$\begin{aligned} G^{(0)}(\mathbf{k}, k_0) &= \mathcal{F}_t\big(T[\tilde{b}_{\mathbf{k}}(t), \tilde{b}_{\mathbf{k}}^\dagger(0)]\big) = (k_0 - \epsilon(\mathbf{k}) + i\delta)^{-1}, \\ \hat{G}^{(0)}(\mathbf{k}, k_0) &= \mathcal{F}_t\big(T[\tilde{b}_{\mathbf{k}}^\dagger(t), \tilde{b}_{-\mathbf{k}}^\dagger(0)]\big) = 0, \end{aligned} \tag{4.36}$$

and the $\mathbf{k} = 0$ Green functions, which are defined analogously, are given by

$$g^{(0)}(\omega) = (\omega - f(0) + i\delta)^{-1}, \quad \hat{g}^{(0)}(\omega) = 0, \tag{4.37}$$

where the parentheses mean the expectation values with respect to the ground state of (4.35). In this section we assume that $c = o(N^{\frac{1}{2}})$ so that $\mathcal{H}_C'$ in (3.15) may be ignored. The standard equations for the $\mathbf{k} \neq 0$ Green functions[5,6,38] can be immediately modified to apply to the $\mathbf{k} = 0$ Green functions:

$$\begin{aligned} &[(\omega + i\delta) - \sigma_1(\omega)]g(\omega) - \sigma_2(\omega)\hat{g}(\omega) = 1, \\ &-\sigma_2(\omega)g(\omega) + [(-\omega + i\delta) - \sigma_1(-\omega)]\hat{g}(\omega) = 0, \end{aligned} \tag{4.38}$$

where $\sigma_1(\omega)$ and $\sigma_2(\omega)$ are the irreducible self-energy parts of the condensate fluctuation; see Fig. 3. Note in particular that the term $f(0)$ of the condensate

[37] See (3.14) and (3.25).

[38] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, translated by R. A. Silverman (Prentice-Hall, Engelwood Cliffs, New Jersey, 1963).

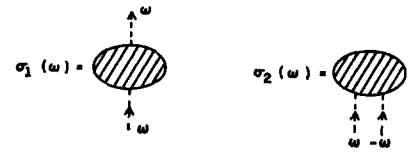


FIG. 3. Irreducible self-energy parts of the condensate fluctuation. To be attached are external lines for excitations of the condensate fluctuation (shown by dotted lines) carrying momentum zero and energy ω.

"kinetic energy" in (4.35) is amalgamated into $\sigma_1(\omega)$. We have written $i\delta$ explicitly in (4.38) since this is needed in the following discussion. The solutions of (4.38) are

$$g(\omega) = [\omega - u(\omega) + s(\omega)]/d(\omega), \qquad \hat{g}(\omega) = -\sigma_2(\omega)/d(\omega), \tag{4.39}$$

where

$$d(\omega) = [\omega - u(\omega)]^2 - \{[s(\omega) - i\delta]^2 - \sigma_2(\omega)^2\}, \tag{4.40}$$

and

$$s(\omega) = \tfrac{1}{2}[\sigma_1(\omega) + \sigma_1(-\omega)], \qquad u(\omega) = \tfrac{1}{2}[\sigma_1(\omega) - \sigma_1(-\omega)]. \tag{4.41}$$

In the case of the pair Hamiltonian (3.1), the irreducible self-energy parts are simply represented by the diagrams in Fig. 4, all further corrections vanishing as $V \to \infty$ if the matrix elements of c are $o(N^{\frac{1}{2}})$, which we assume in this subsection. With the aid of Fig. 4(a) one finds

$$\sigma_1(\omega) = f(0) + \frac{i}{2\pi}\frac{2\rho_0}{V}\sum_{\mathbf{k}}\int dk_0 \times \{G^{(0)}(\mathbf{k}, \tfrac{1}{2}\omega + k_0)G^{(0)}(-\mathbf{k}, \tfrac{1}{2}\omega - k_0)\gamma_+^2(\mathbf{k}) + G^{(0)}(\mathbf{k}, -\tfrac{1}{2}\omega + k_0)G^{(0)}(-\mathbf{k}, -\tfrac{1}{2}\omega - k_0)\gamma_-^2(\mathbf{k})\},$$

where the vertex functions $\gamma_\pm(\mathbf{k})$ are given by (4.10). The method of residues applied to the k_0 integration gives

$$\sigma_1(\omega) = f(0) + \frac{2\rho_0}{V}\sum_{\mathbf{k}}{}' \left[\frac{\gamma_+^2(\mathbf{k})}{\omega - 2\epsilon(\mathbf{k}) + i\delta} - \frac{\gamma_-^2(\mathbf{k})}{\omega + 2\epsilon(\mathbf{k}) - i\delta}\right]. \tag{4.42}$$

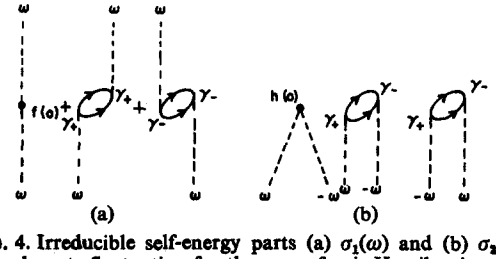


FIG. 4. Irreducible self-energy parts (a) $\sigma_1(\omega)$ and (b) $\sigma_2(\omega)$ of the condensate fluctuation for the case of pair Hamiltonian model. Dotted lines are for the condensate fluctuation (to be attached) and full lines represent excitations of $\mathbf{k} \neq 0$ particles.

In the same way, with the aid of Fig. 4(b), we obtain

$$\sigma_2(\omega) = h(0) - \frac{2\rho_0}{V}\sum_{\mathbf{k}}{}' \frac{2\epsilon(\mathbf{k}) - i\delta}{[2\epsilon(\mathbf{k}) - i\delta]^2 - \omega^2} \times 2\gamma_+(\mathbf{k})\gamma_-(\mathbf{k}). \tag{4.43}$$

Thus according to the definitions (4.41)

$$s(\omega) = f(0) - \frac{2\rho_0}{V}\sum_{\mathbf{k}}{}' \frac{2\epsilon(\mathbf{k})}{[2\epsilon(\mathbf{k}) - i\delta]^2 - \omega^2} \times [\gamma_+^2(\mathbf{k}) + \gamma_-^2(\mathbf{k})], \tag{4.44}$$

$$u(\omega) = \omega\bar{u}(\omega), \tag{4.45}$$

with

$$\bar{u}(\omega) = -\frac{2\rho_0}{V}\sum_{\mathbf{k}}{}' v(\mathbf{k})\frac{[\gamma_+(\mathbf{k}) + \gamma_-(\mathbf{k})]}{[2\epsilon(\mathbf{k}) - i\delta]^2 - \omega^2}. \tag{4.46}$$

In view of (4.33) and (4.39), we know that the one-particle excitation energy of the condensate fluctuation should show up as a pole ω_0 of $g(\omega)$, or, equivalently, of $\hat{g}(\omega)$ in the lower half ω plane:

$$d(\omega_0) = 0, \quad \text{Im}\,\omega_0 < 0. \tag{4.47}$$

Recalling (4.40) and (4.45) we can write ω_0^2 as

$$\omega_0^2 = f_0^2 - h_0^2, \tag{4.48}$$

where

$$f_0 = D^{-1}[s(\omega_0) - i\delta], \qquad h_0 = D^{-1}\sigma_2(\omega_0), \tag{4.49}$$

and

$$D = 1 - \bar{u}(\omega_0). \tag{4.50}$$

The signs of f_0 and h_0 are determined in such a way that f_0 approaches its unperturbed value $f(0)$ as $v(\mathbf{k}) \to 0$. One should notice that the notations in (4.49) are justified because the present results are essentially identical with those obtained from our perturbation method [see (4.23)–(4.25)]. These equations differ only by the imaginary term $i\delta$, which serves to fix the sign of ω_0.

We now locate the solution ω_0 of (4.47) in the complex ω plane under the assumption that $(\rho_0 a^3)^{\frac{1}{2}} \ll 1$. In this case (4.28) is satisfied so that the imaginary parts of s, $\bar{u}$, and σ_2 are $O(\delta)$. Furthermore, because these imaginary parts would all vanish if $v(\mathbf{k}) \to 0$, they are therefore all smaller than the term $i\delta$ explicitly written in (4.49). Therefore,

$$\text{Im}\,(f_0^2 - h_0^2) = -[2s/D^2]_{\delta=0} \cdot \delta. \tag{4.51}$$

In the following, all expressions are understood to be evaluated at $\delta = 0$. Now supposing that

$$f_0^2 - h_0^2 > 0, \tag{4.52}$$

then, according to the criterion Im $\omega_0 < 0$, we know

from (4.51) that

$$\omega_0 = (f_0^2 - h_0^2)^{\frac{1}{2}} \operatorname{sgn} f_0, \tag{4.53}$$

with the square root being taken to be positive. Under the assumption of (4.52) we know that the sign of f_0 is the same as the common sign of $f_0 \pm h_0$, and, therefore, that the result (4.53) is in complete agreement with the perturbation result, namely with the eigenfrequency of the effective Hamiltonian (4.27).

Since the self-energy functions, (4.42) and (4.43), are asymptotically exact as $V \to \infty$, so is the result (4.53) for the one-particle excitation energy. Moreover, we can easily see that apart from the δ function for the over-all energy conservation, the many-particle Green function

$$g(\omega_1, \cdots, \omega_n) = \mathcal{F}_{t_1 \cdots t_n} \langle \Omega | \, T[c(t_1) \cdots c^\dagger(t_n)] \, | \Omega \rangle$$

decomposes into a sum of products of one-particle Green functions in the volume limit. This means that the many-particle excitation energies of the condensate fluctuation are integer multiples of the one-particle excitation energy. Since we know already of the agreement between the values of the one-particle excitation energies as obtained from the Green function method and the modified Born–Oppenheimer method, we can now conclude that our second-order result (4.27) gives the correct energy spectrum in the volume limit. Corresponding results for the matrix elements of $c^{(\dagger)}$ and their products (i.e., agreement of ground state expectation values) can be easily established once one Fourier-transforms the Green functions back into t space. This completes the comparison between the results of the Green function theory and our perturbation theory. From this we conclude that our second-order perturbation approximation (4.27) for the effective condensate Hamiltonian is already exact in the volume limit.

V. IMPLICATIONS OF THE STABILITY CONDITIONS

Having obtained a set of stability conditions for the pair Hamiltonian model, we now determine the restrictions which these conditions place on the particle density ρ and the characteristics of the interparticle potential v. For the reason given at the close of the previous section we restrict our attention to those cases where (4.29) is satisfied[39]:

$$D > \tfrac{1}{2}. \tag{5.1}$$

[39] The discussions in this and in the next section are not affected if we take $D > 0$ in place of (4.29) as long as $[2\epsilon(0)]^2 - \omega_0^2 > 0$.

We further assume that

$$2\epsilon(\mathbf{p}) - \omega_0 > 0 \quad (\text{all } \mathbf{p} \neq 0). \tag{5.2}$$

In order that the $\mathbf{k} = 0$ stability conditions (4.26) be satisfied, we must then require that the numerators of (4.23) and (4.24) are positive. Thus in terms of the self-energy functions,

$$s + \sigma_2 = 2\rho_0 v(0) - \frac{\rho_0}{V} \sum_{\mathbf{p}}{}' \frac{4\epsilon(\mathbf{p})[\gamma_+(\mathbf{p}) + \gamma_-(\mathbf{p})]^2}{4\epsilon(\mathbf{p})^2 - \omega_0^2} > 0 \tag{5.3}$$

and

$$s - \sigma_2 = \frac{1}{V} \sum_{\mathbf{p}}{}' v(\mathbf{p}) \frac{1}{\epsilon(\mathbf{p})} h(\mathbf{p}) - \frac{\rho_0}{V} \sum_{\mathbf{p}}{}' v(\mathbf{p})^2 \frac{4\epsilon(\mathbf{p})}{4\epsilon(\mathbf{p})^2 - \omega_0^2} > 0. \tag{5.4}$$

The $\mathbf{k} \neq 0$ stability conditions are given by (3.20) and (3.21). The functions $f(\mathbf{k})$ and $h(\mathbf{k})$ appearing in these equations are the solutions of a system of coupled integral equations, (3.30) and (3.31). Because $f(\mathbf{k})$ has a rapidly increasing term $\mathbf{k}^2$, the $\mathbf{k} \neq 0$ stability conditions will be satisfied if

$$f(0) + h(0) = 2\rho_0 v(0) > 0, \tag{5.5}$$

$$f(0) - h(0) = V^{-1} \sum_{\mathbf{k}}{}' v(\mathbf{k}) \frac{1}{\epsilon(\mathbf{k})} h(\mathbf{k}) > 0. \tag{5.6}$$

At this juncture we note that the inequalities (5.5) and (5.6) follow from the preceding four. From now on, therefore, we call the set of inequalities (5.1)–(5.4) the stability conditions for the many-boson system. These constitute a set of sufficient conditions for a many-boson system to have a bona fide ground state with a B.E. condensation. Based on the above, the subset (5.1) and (5.2) can be called the $k \neq 0$ stability conditions, as opposed to the old ones (5.5) and (5.6), or (3.20) and (3.21).

It should be kept in mind, in view of the discussions in Appendix B, that if any of $f_0 \pm h_0 > 0$ and $f(0) \pm h(0) > 0$ are violated, either there is no B.E. condensation in the system or the operator replacement (1.1) loses its usefulness.

In the following we show that one of the stability conditions, (5.4), can be rewritten in a simple and suggestive way. As for the other ones, in particular (5.3), we have to be content with an analysis of a weak coupling model, this analysis being the subject of the next section.

In order to analyze (5.4), we rewrite its first term

with the aid of the integral equation (3.31) as

$$\frac{1}{\rho_0 V}\sum_{\mathbf{p}}\frac{v(\mathbf{p})h(\mathbf{p})}{\epsilon(\mathbf{p})} = \sum_{\frac{1}{2}[\mathbf{p}]} 2v(\mathbf{p})\frac{1}{2V\epsilon(\mathbf{p})}2v(\mathbf{p}) - \sum_{\frac{1}{2}[\mathbf{p}],\frac{1}{2}[\mathbf{g}]} 2v(\mathbf{p})\frac{1}{2V\epsilon(\mathbf{p})} \times \big(v(\mathbf{p}-\mathbf{g})+v(\mathbf{p}+\mathbf{g})\big)\frac{1}{2V\epsilon(\mathbf{g})}2v(\mathbf{g}) + \cdots, \tag{5.7}$$

where, as in Sec. III, $\frac{1}{2}[\mathbf{p}]$ under the summation sign means that the sum includes only those vectors $\mathbf{p}$ which belong to a half-space. This expression can be written in a compact way by considering a fictitious problem of a pair of b excitations (see Sec. III) interacting via the two-body potential $\tilde{v}$. A state vector $|\mathbf{p}\rangle$ describing a pair of b excitations with momenta $\mathbf{p}$, $-\mathbf{p}$ are normalized so that

$$\langle \mathbf{p} \,|\, \mathbf{g}\rangle = V\delta_{\mathbf{p},\mathbf{g}}. \tag{5.8}$$

These state vectors can thus be written as

$$\begin{aligned} |\mathbf{p}\rangle &= V^{\frac{1}{2}} b_{\mathbf{p}}^{\dagger} b_{-\mathbf{p}}^{\dagger}|0\}, \quad (\mathbf{p}\neq 0), \\ |0\rangle &= V^{\frac{1}{2}} b_0^{\dagger} b_0^{\dagger}|0\}, \end{aligned} \tag{5.9}$$

with the understanding that $\mathbf{p}\,(\neq 0)$ belongs to the half-space indicated by the summation signs in (5.7).[40] The Hamiltonian describing a single pair is $\mathcal{C}+\mathcal{U}$, where the matrix elements of the operators $\mathcal{C}$ and $\mathcal{U}$ in the momentum representation are

$$\langle \mathbf{p}|\,\mathcal{C}\,|\mathbf{g}\rangle = 2V\epsilon(\mathbf{p})\delta_{\mathbf{p},\mathbf{g}}, \tag{5.10}$$

$$\langle \mathbf{p}|\,\mathcal{U}\,|\mathbf{g}\rangle = v(\mathbf{p}-\mathbf{g}) + v(\mathbf{p}+\mathbf{g}). \tag{5.11}$$

In terms of $\mathcal{C}$ and $\mathcal{U}$ the left-hand side of (5.7) is

$$\frac{1}{\rho_0 V}\sum_{\mathbf{p}} v(\mathbf{p})\frac{h(\mathbf{p})}{\epsilon(\mathbf{p})} = \langle 0|\,\mathcal{U}(\mathcal{C}+\mathcal{U})^{-1}\mathcal{U}\,|0\rangle, \tag{5.12}$$

as is easily verified using the expansion

$$(\mathcal{C}+\mathcal{U})^{-1} = \mathcal{C}^{-1} - \mathcal{C}^{-1}\mathcal{U}\mathcal{C}^{-1} + \cdots, \tag{5.13}$$

inserting as intermediate states the complete set of pair states,[41] and by comparing the result with the right side of (5.7). Likewise, the second term of (5.4) can be written as

$$-\frac{1}{V}\sum_{\mathbf{p}}\frac{4\epsilon(\mathbf{p})v(\mathbf{p})^2}{4\epsilon(\mathbf{p})^2-\omega_0^2} = -\langle 0|\,\mathcal{U}\,\frac{1}{2}\left(\frac{1}{\mathcal{C}-\omega_0}+\frac{1}{\mathcal{C}+\omega_0}\right)\mathcal{U}\,|0\rangle. \tag{5.14}$$

[40] The state vector $|0\}$ is the no-particle state of the $b_{\mathbf{k}}$ operators, here defined for all $\mathbf{k}$, including $\mathbf{k}=0$. Compare with (3.26).

[41] Note that $\mathcal{U}$ connects the state $|0\rangle$ to pair states $|\mathbf{p}\rangle$ only.

Hence the inequality (5.4) can be put in the simple form

$$(1/\rho_0)(s-\sigma_2) = \langle 0|\,\mathcal{U}(\mathcal{C}+\mathcal{U})^{-1}\mathcal{U}\,|0\rangle - \tfrac{1}{2}\langle 0|\,\mathcal{U}[(\mathcal{C}-\omega_0)^{-1}+(\mathcal{C}+\omega_0)^{-1}]\mathcal{U}\,|0\rangle > 0. \tag{5.15}$$

Now suppose that the eigenvalue problem

$$(\mathcal{C}+\mathcal{U})\,|\alpha\rangle = E_\alpha\,|\alpha\rangle \tag{5.16}$$

for a pair of b excitations with zero center-of-mass momentum has been solved. Utilizing the complete set of pair states $|\mathbf{p}\rangle$, the second term in (5.15) itself satisfies the inequality

$$\langle 0|\,\mathcal{U}[(\mathcal{C}-\omega_0)^{-1}+(\mathcal{C}+\omega_0)^{-1}]\mathcal{U}\,|0\rangle > \sum_{(\mathbf{p})}\frac{|\langle \mathbf{p}|\,\mathcal{U}\,|0\rangle|^2}{\epsilon(\mathbf{p})}.$$

Thus (5.15) can be replaced by the following:

$$\sum_{\frac{1}{2}[\alpha]}|\langle\alpha|\,\mathcal{U}\,|0\rangle|^2/E_\alpha > \sum_{\frac{1}{2}[\mathbf{p}]}|\langle\mathbf{p}|\,\mathcal{U}\,|0\rangle|^2/\big(2\epsilon(\mathbf{p})\big), \tag{5.17}$$

where the symbol $\frac{1}{2}[\alpha]$ on the summation sign warns against double counting the eigenstates of (5.16). Because of the identity

$$\langle 0|\,\mathcal{U}^2\,|0\rangle = \sum_{\frac{1}{2}[\alpha]}|\langle\alpha|\,\mathcal{U}\,|0\rangle|^2 = \sum_{\frac{1}{2}[\mathbf{p}]}|\langle\mathbf{p}|\,\mathcal{U}\,|0\rangle|^2, \tag{5.18}$$

the two sums in (5.17) can be interpreted in terms of averages with weight factors $|\langle\alpha|\,\mathcal{U}\,|0\rangle|^2$ and $|\langle\mathbf{p}|\,\mathcal{U}\,|0\rangle|^2$, respectively. The inequality (5.17) suggests that the interparticle potential $\tilde{v}$ must be such that it lowers the energy levels of the fictitious two-particle system (5.16) from their values when $\tilde{v}=0$ and/or shifts the dominant weight factors to the low-energy side. Crudely speaking, this suggests that $\tilde{v}$ must be partially attractive. On the other hand, we recall that (5.5) requires that $v(0)=\int\tilde{v}(\mathbf{x})\,d\mathbf{x}>0$. The following lemma allows a more precise statement to be made.

Lemma: For a low-density system of particles interacting via a potential that is sufficiently weak and everywhere positive in coordinate space, the quantity $s-\sigma_2$ is negative, in violation of the stability condition (5.4).

Keeping only the first two terms of the expansion (5.13) we obtain

$$s-\sigma_2 < -\rho_0\int|\langle\mathbf{p}=0|\,\mathcal{U}\mathcal{C}^{-1}\,|\mathbf{x}\rangle|^2\,\tilde{v}(\mathbf{x})\,d\mathbf{x} < 0, \tag{5.19}$$

which was to be proved. Here $|\mathbf{x}\rangle$ is an eigenvector of the relative position operator $\mathbf{x}$; for the sake of clarity the momentum eigenstate is written as

$|\mathbf{p} = 0\rangle$. Note that when $\epsilon(\mathbf{p})$ is chosen to have the well-known form, $\epsilon(\mathbf{p}) = p(p^2 + 16\pi a\rho_0)^{\frac{1}{2}}$, the main contributions to the sums in (5.4) come from momenta $\lesssim (a\rho_0)^{\frac{1}{2}}$.[4] Then the above expansion of the propagator $(\mathcal{G} + \mathcal{U})^{-1}$ is allowed, and hence the lemma is true when the density parameter is sufficiently small, $(\rho a^3)^{\frac{1}{2}} \ll 1$.

If the above lemma suggests that the pair-Hamiltonian model with a purely positive potential may not have a B.E. condensation, we have to recall that for such a system the assumption of *no* condensation is untenable.[15,42] For the proof, we multiply the integral equation (3.31) for $h(\mathbf{k})$ by $V^{-1}h(\mathbf{k})/\epsilon(\mathbf{k})$ and sum over momenta to obtain

$$V^{-1}\sum_{\mathbf{k}}{}' \epsilon(\mathbf{k})^{-1}h(\mathbf{k})^2 + \tfrac{1}{2}\sum_{\mathbf{k},\mathbf{p}}{}' g(\mathbf{k})v(\mathbf{k} - \mathbf{p})g(\mathbf{p}) = 0, \tag{5.20}$$

where $g(\mathbf{k}) = h(\mathbf{k})/\epsilon(\mathbf{k})$, and we have put $\rho_0 = 0$, i.e., we assume that the system has no macroscopic occupation of the $\mathbf{k} = 0$ state. The first term is nonnegative. By the convolution theorem for Fourier transforms and the assumption that $\tilde{v}(\mathbf{x}) > 0$, the second term is also nonnegative:

$$\tfrac{1}{2}\int_V |\tilde{g}(\mathbf{x})|^2\, \tilde{v}(\mathbf{x})\, d\mathbf{x} \geq 0, \tag{5.21}$$

where $\tilde{g}(\mathbf{x})$ is the Fourier transform of $g(\mathbf{k})$. Thus the only solution to the integral equation (3.31) is $h(\mathbf{k}) = 0$ (all $\mathbf{k}$), which means $\beta_{\mathbf{k}}^2 = 0$ according to (3.24), and therefore

$$\langle\Omega|\sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}\,|\Omega\rangle = 0, \tag{5.22}$$

in contradiction to the fundamental requirement (1.4). Thus the assumption of no condensation is untenable. These results appear to be contradictory. Further, it is ill-advised to try to avoid this contradiction by supposing that $D < 0$, for, according to (4.25), in the low density limit $D \approx 1$. Thus, in conclusion, it appears that the operator replacement (1.1) loses its usefulness for a purely repulsive potential, for example, of the Yukawa type.

VI. A WEAK COUPLING MODEL

In this section we express the stability conditions (5.1)–(5.4) in terms of the parameters of a simple potential of the form

$$v(\mathbf{k}) = A(k^2 + \kappa^2)^{-1} + B(k^2 + \lambda^2)^{-1}, \tag{6.1}$$

corresponding to the configuration space potential $\tilde{v}(\mathbf{r}) = (4\pi r)^{-1}\cdot(Ae^{-\kappa r} + Be^{-\lambda r})$. We find that there exists a range of parameters for which the stability conditions can in fact be satisfied. In accordance with the lemma of the previous section, that a purely repulsive potential cannot satisfy (5.4) and (5.5) simultaneously, we can expect that one of the constants A and B must be negative. Although the precise requirements on A and B are determined later in this section, we can at present require that

$$x + y > 0, \tag{6.2a}$$

$$|x| \approx |y| \lesssim 1, \tag{6.2b}$$

where[43]

$$x = \lambda mA/(4\pi\kappa^2), \quad y = mB/(4\pi\lambda). \tag{6.3}$$

Equation (6.2a) ensures that $v(0) > 0$; $|x| \approx |y|$ is suggested by the above lemma; and the requirement $|x|, |y| \lesssim 1$ ensures that the attractive component of $\tilde{v}(\mathbf{r})$ is not sufficiently strong to allow any two-body bound states. Further, to simplify the later analysis we assume that

$$\lambda \gg \kappa. \tag{6.4}$$

Finally the assumption of "weak coupling" is characterized by the requirement

$$z = m\rho_0 B/\kappa^2\lambda^2 \ll 1. \tag{6.5}$$

We begin the analysis by obtaining approximate solutions of the integral equations (3.30) and (3.31) for $f(\mathbf{k})$ and $h(\mathbf{k})$, respectively. The former equation may be written as

$$f(\mathbf{k}) = k^2/(2m) + \rho_0[v(\mathbf{k}) - v(0)] + f(0) + \delta f(\mathbf{k}), \tag{6.6}$$

where

$$f(0) = \rho_0 v(0) + \tfrac{1}{2}(2\pi)^{-3}\int d\mathbf{p}\, v(\mathbf{p})h(\mathbf{p})/\epsilon(\mathbf{p}), \tag{6.7}$$

$$\delta f(\mathbf{k}) = \tfrac{1}{2}(2\pi)^{-3}\int d\mathbf{p}[v(\mathbf{p} - \mathbf{k}) - v(\mathbf{p})]\left[\frac{f(\mathbf{p})}{\epsilon(\mathbf{p})} - 1\right]. \tag{6.8}$$

[42] C. J. Pethick and D. ter Haar, University of Oxford, Clarendon Laboratory Report, (1965); Phys. Letters **19**, 20 (1965). These authors conclude that $\tilde{v}(\mathbf{x}) > 0$ is a sufficient condition for the pair-Hamiltonian model to undergo a B.E. condensation at sufficiently low temperatures. This conclusion was obtained with reference to a portion H_T of $\mathcal{H}_p$, the former being quadratic in $a_{\mathbf{k}}^{(\dagger)}$ operators. Pethick and ter Haar then presented an argument based on a type of thermodynamic perturbation theory that H_T and $\mathcal{H}_p$ are thermodynamically equivalent in the volume limit. Their derivation closely parallels but corrects an error of an attempted proof of thermodynamic equivalence given earlier by one of us (see Appendix A of Ref. 15). Independently we sketched essentially the same formal proof as that of Pethick and ter Haar. We do not, however, believe the proof to be very convincing for the following reason. Employing the same thermodynamic perturbation theory as we use in the "proof" of thermodynamic equivalence, we find the formal result $\langle\theta(c, c^\dagger)\rangle_p = O(1)$, where $\langle\cdots\rangle_p$ is the ensemble average taken with respect to $\mathcal{H}_p$ and θ is any polynomial in $c^{(\dagger)}$. This result, however, cannot in general be true, for we have shown in Secs. IV and V that if certain stability conditions are not satisfied, $c^{(\dagger)} = O(1)$ is an untenable assumption. We believe that these various formal "proofs" break down because the perturbation series are divergent.

[43] Although throughout this paper we have chosen the particle mass as $m = \frac{1}{2}$, in this section we write the mass as m.

We show very shortly that, as long as (6.2)–(6.5) are satisfied, it is a good approximation to write

$$f(\mathbf{k}) = k^2/(2m) + f(0) + \rho_0[v(\mathbf{k}) - v(0)], \quad (6.9)$$

$$\epsilon(\mathbf{k}) = \epsilon(0) + k^2/(2m). \quad (6.10)$$

Specifically, our procedure in the following is to utilize (6.10) in order to solve (3.31) for $h(\mathbf{k})$, and then using this result, we show that (6.9) and (6.10) are valid.

In Appendix C we show, using (6.10), that when (6.2)–(6.5) are satisfied, the following is a good approximate solution of (3.31):

$$h(\mathbf{k}) = h^{(0)}(\mathbf{k}) + h^{(1)}(\mathbf{k}), \quad (6.11)$$

where

$$h^{(0)}(\mathbf{k}) = \rho_0 v(\mathbf{k}), \quad h^{(1)}(\mathbf{k}) = \frac{-y}{1 + (y/\tau)\, k^2 + (\tau\lambda)^2}\,\frac{\rho_0 B}{}, \quad (6.12)$$

and $\tau = 1.62$. One can then easily verify that

$$\tfrac{1}{2}(2\pi)^{-3}\int d\mathbf{p} v(\mathbf{p}) h^{(0)}(\mathbf{p})/\epsilon(\mathbf{p}) = (2m)^{-1}\kappa^2 zy[1 + O(\kappa/\lambda)], \quad (6.13)$$

$$\tfrac{1}{2}(2\pi)^{-3}\int d\mathbf{p} v(\mathbf{p}) h^{(1)}(\mathbf{p})/\epsilon(\mathbf{p}) = -2(2m)^{-1}\kappa^2 zy^2(\tau + 1)^{-1}(\tau + y)^{-1}[1 + O(\kappa/\lambda)]. \quad (6.14)$$

In the following we avoid writing the error terms since $\lambda \gg \kappa$. Combining (6.7) and (6.11)–(6.14) we have

$$f(0) + h(0) = 2\rho_0 v(0) + \left(\frac{\kappa^2}{2m}\right) zy\,\frac{0.38 + 0.24y}{1.62 + y}, \quad (6.15)$$

$$f(0) - h(0) = \left(\frac{\kappa^2}{2m}\right) zy\,\frac{2.85 + 0.24y}{1.62 + y}. \quad (6.16)$$

The second term on the right-hand side of (6.15) is the error due to the approximations (6.9) and (6.10) [see (5.5)]. Thus to ensure the validity of the present method we require that

$$\frac{B}{\lambda^2}\, y\left(\frac{0.19 + 0.12y}{1.62 + y}\right) \ll 2v(0). \quad (6.17)$$

Finally, according to (3.25), $\epsilon(0)$ is given by $[f(0)^2 - h(0)^2]^{\frac{1}{2}}$.

The validity of (6.10) is now easily established. Using (6.9) and (6.11), for values of $k \ll \kappa$ one finds

$$f(\mathbf{k}) \pm h(\mathbf{k}) = f(0) \pm h(0) + (k^2/2m)[1 + O(z)]$$

so that in this range (6.10) is a good approximation. Further, for values of $k \gtrsim \kappa$, the term $k^2/2m$ in (6.9) dominates all other terms, and thus (6.10) is valid in this range as well.

Concerning the validity of (6.9) as an approximation to the exact expression (6.6), we find that the term $\delta f(\mathbf{k})$ is ignorable as compared to $f(0)$ and $k^2/(2m)$ when $k \lesssim \kappa/5$ and $k \gtrsim \kappa$, respectively. In particular, for $k \lesssim \kappa/5$, direct evaluation of the integral in (6.8) using (6.9) and (6.10) shows that

$$|\delta f(\mathbf{k})| \leq f(0)(k^2/\kappa^2)O(\kappa/\lambda).$$

We now proceed to study the stability conditions (5.1)–(5.4). In the following it suffices to approximate $\gamma_+(k) \pm \gamma_-(k)$ [see (4.10) and (4.11)] by

$$\gamma_+(\mathbf{k}) - \gamma_-(\mathbf{k}) = v(\mathbf{k}) \approx v(0), \quad (6.18)$$

$$\gamma_+(\mathbf{k}) + \gamma_-(\mathbf{k}) \approx v(0)[1 + \Delta/\epsilon(\mathbf{k})], \quad (6.19)$$

where

$$\Delta = f(0) - \epsilon(0) - 2h(0). \quad (6.20)$$

These approximations are valid because of the presence of the factor $[4\epsilon(k)^2 - \omega_0^2]^{-1}$ in (4.25), (5.3), and (5.4) which is a very rapidly decreasing function of k as compared to $v(\mathbf{k})$ or $h(\mathbf{k})$. Thus (4.25) may be written as

$$D = 1 + \frac{m\rho_0 v(0)^2}{2\pi^2}\int_0^\infty dk\left(\frac{1}{2\epsilon(\mathbf{k}) - \omega_0} + \frac{1}{2\epsilon(\mathbf{k}) + \omega_0}\right) \times \left(1 + \frac{\Delta}{\epsilon(\mathbf{k})}\right). \quad (6.21)$$

The integral is readily evaluated using (6.10) [assume for the moment that (5.2) is satisfied so that the integrand is well behaved over the entire range of integration], and the result is

$$D = 1 + O\left[\frac{\kappa}{\lambda} z^{\frac{1}{2}} y^{\frac{3}{4}}\left(\frac{v(0)}{B\lambda^{-2}}\right)^{\frac{1}{4}}\right]. \quad (6.22)$$

Recalling (6.2), (6.4), (6.5), and the fact that $v(0) < |B|\,\lambda^{-2}$, it suffices to take $D = 1$, and thus the stability condition (5.1) is satisfied.

We now turn to the stability condition (5.3). An analysis similar to the above shows that the leading contribution to the integral in (5.3) can be obtained by substituting $\gamma_+(\mathbf{k}) + \gamma_-(\mathbf{k}) = v(\mathbf{k})$ and $\omega_0 = 0$. Combining (4.24), (5.3), and our previous result (6.22) we have

$$f_0 + h_0 \approx 2\rho_0 v(0) - \frac{\rho_0}{V}\sum_{\mathbf{k}}{}' \frac{v(\mathbf{k})^2}{\epsilon(\mathbf{k})}. \quad (6.23)$$

Recalling (6.13) we obtain

$$f_0 + h_0 = 2\rho_0 v(0) - (\kappa^2/m)zy. \quad (6.24)$$

At the same time we note that with the aid of (6.11) and (6.12) we can write (5.4) as

$$f_0 - h_0 \approx \frac{1}{V}\sum_{\mathbf{k}}{}' \frac{v(\mathbf{k}) h^{(1)}(\mathbf{k})}{\epsilon(\mathbf{k})} - \tfrac{1}{4}\omega_0^2\rho_0\,\frac{1}{V}\sum_{\mathbf{k}}{}' \frac{v(\mathbf{k})^2}{\epsilon(\mathbf{k})^3}.$$

Recalling the identity $\omega_0^2 = (f_0 - h_0)(f_0 + h_0)$ and (6.14), it follows that

$$(f_0 - h_0)\left\{1 + O\left[\left(\frac{v(0)}{B\lambda^{-2}}\right)^2 \frac{\sigma}{\lambda}\right]\right\} = -2\frac{\kappa^2}{m}\frac{zy^2}{(\tau + 1)(\tau + y)},$$

where $\sigma = [2m\epsilon(0)]^{\frac{1}{2}}$. Now using (6.15)–(6.17) one easily finds that $\sigma/\lambda \ll \sigma/\kappa \ll 1$. This result, in conjunction with the inequality $v(0) < |B|\,\lambda^{-2}$, leads to

$$f_0 - h_0 = -2(\kappa^2/m)[zy^2/(\tau + 1)(\tau + y)]. \quad (6.25)$$

The $\mathbf{k} = 0$ stability conditions require that both (6.24) and (6.25) are positive. Since the parameters y and z have the same sign, it follows from (6.25) that

$$-\tau < y < 0. \quad (6.26)$$

Equation (6.24) rewritten as

$$f_0 + h_0 = 2(\kappa^2/m)zy^{-1}[x + y(1 - \tfrac{1}{2}y)]$$

provides the added restrictions

$$x > 0, \quad (6.27)$$

$$1 - (1 + 2x)^{\frac{1}{2}} < y < 0. \quad (6.28)$$

The requirements $x > 0$, $y < 0$ imply that the short-range component of the potential $\tilde{v}(\mathbf{x})$ must be attractive and the long-range part repulsive. Note further that because of (6.2b) the lower bound on y is actually provided by (6.28) rather than by (6.26).

Finally, we turn to the last of the stability conditions (5.2). This condition, however, is automatically satisfied when (6.28) is satisfied, as is readily seen using (6.5), (6.16), (6.24), and (6.25).

Summarizing, in the case of the pair Hamiltonian model when the potential is of the form (6.1) and the restrictions (6.2)–(6.5) and (6.27)–(6.28) are satisfied, then the stability conditions (5.1)–(5.4) are satisfied. We have thus explicitly demonstrated that the stability conditions can in certain cases be satisfied.

VII. THE FULL HAMILTONIAN

We now turn to the question of the existence of a B.E. condensation in a system of N pairwise interacting bosons described by the complete Hamiltonian

$$\mathcal{H} = \sum_{\mathbf{p}} (p^2 - \mu)a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + (2V)^{-1}\sum_{\mathbf{k},\mathbf{p},\mathbf{q}} v(\mathbf{k})a_{\mathbf{p}+\mathbf{k}}^\dagger a_{\mathbf{q}-\mathbf{k}}^\dagger a_{\mathbf{q}} a_{\mathbf{p}}. \quad (7.1)$$

Only a small subclass of the interaction terms of (7.1) are included in the pair Hamiltonian (3.1) which we have been studying in the previous sections. Again we ask: Does the system described by (7.1) undergo a B.E. condensation? In addition, what is the order of magnitude of the operators $c^{(\dagger)}$ defined by (1.1)? By tentatively assuming the existence of a condensation, we use the operator replacement (1.1) to follow the program outlined in Sec. II. First, the Hamiltonian $\mathcal{H}_B$ of the particles outside the condensate is diagonalized by truncating it in the manner introduced by Bogoliubov.[3] The effective condensate Hamiltonian Λ_0 is constructed using the perturbation method of Appendix A. Then we use an invariance argument as supplemented by a result from Green function theory to indicate that this approximate result already has the characteristic features one would expect of the exact condensate Hamiltonian. We then find that in the low-density limit the Hamiltonian Λ_0 is diagonal in the occupation number representation of $a_0^\dagger a_0$ but not of $c^\dagger c$. Some discussion is also given regarding systems of moderate density.

A. Perturbation Calculation of Λ_0

Following the program of Sec. II, we split the Hamiltonian (7.1) into three parts $\mathcal{H}_B$, $\mathcal{H}_C$, and $\mathcal{H}_I$ which describe the system of particles outside the condensate, the condensate, and their mutual interaction, respectively. The Hamiltonian $\mathcal{H}_B$ is written as

$$\mathcal{H}_B = \mathcal{H}_B^0 + \mathcal{H}_B', \quad (7.2)$$

where the first term is identical with (3.8), i.e.,

$$\mathcal{H}_B^0 = \overline{W} + \sum_{\mathbf{k}}{}' \{[\chi + f(\mathbf{k})]a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \tfrac{1}{2}h(\mathbf{k})(a_{\mathbf{k}}a_{-\mathbf{k}} + a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger)\}, \quad (7.3)$$

and the second term is

$$\mathcal{H}_B' = N_0^{\frac{1}{2}}V^{-1}\sum_{\substack{\mathbf{p},\mathbf{k}\\(\mathbf{p}+\mathbf{k}\neq 0)}}{}' v(\mathbf{k})(a_{\mathbf{p}+\mathbf{k}}^\dagger a_{\mathbf{p}}a_{\mathbf{k}} + \text{H.c.}) + (2V)^{-1}\sum_{\mathbf{k}}\sum_{\substack{\mathbf{p},\mathbf{q}\\(\mathbf{q}-\mathbf{k},\,\mathbf{p}+\mathbf{k}\neq 0)}}{}' v(\mathbf{k}) :a_{\mathbf{p}+\mathbf{k}}^\dagger a_{\mathbf{q}-\mathbf{k}}^\dagger a_{\mathbf{p}}a_{\mathbf{q}}: . \quad (7.4)$$

Here the symbol $:\cdots:$ means the normal product in the b representation as defined by a Bogoliubov transformation (3.22), whose coefficients are determined below. For instance,

$$:a_{\mathbf{p}}a_{\mathbf{q}}: = \alpha_{\mathbf{p}}\alpha_{\mathbf{q}}b_{\mathbf{p}}b_{\mathbf{q}} + \alpha_{\mathbf{p}}\beta_{\mathbf{q}}b_{-\mathbf{q}}^\dagger b_{\mathbf{p}} + \beta_{\mathbf{p}}\alpha_{\mathbf{q}}b_{-\mathbf{p}}^\dagger b_{\mathbf{q}} + \beta_{\mathbf{p}}\beta_{\mathbf{q}}b_{-\mathbf{p}}^\dagger b_{-\mathbf{q}}^\dagger = a_{\mathbf{p}}a_{\mathbf{q}} - \alpha_{\mathbf{p}}\beta_{\mathbf{p}}\delta_{\mathbf{p},-\mathbf{q}}. \quad (7.5)$$

Further, in Eqs. (3.11) and (3.12), which define the c number coefficients in (7.3), we have to put

$$\xi_{\mathbf{k}} = \beta_{\mathbf{k}}^2, \quad \eta_{\mathbf{k}} = \alpha_{\mathbf{k}}\beta_{\mathbf{k}}. \quad (7.6)$$

The condensate part $\mathcal{H}_C$ of the Hamiltonian (7.1) need not be written down here because it has exactly the same form as its pair Hamiltonian counterpart

(3.13). Finally, the interaction part of $\mathcal{H}$ is given by

$$\mathcal{H}_I = [(J_1 + J_2 + J_3)N_0^{\frac{1}{2}}c + \text{H.c.}] + (J_1 - \Delta f_0)c^\dagger c + [\tfrac{1}{2}(J_2 - \Delta h_0)cc + \text{H.c.}], \quad (7.7)$$

where

$$J_1 = V^{-1}\sum_{\mathbf{k}}{}' [v(0) + v(\mathbf{k})] :a_{\mathbf{k}}^\dagger a_{\mathbf{k}}:, \quad (7.8)$$

$$J_2 = V^{-1}\sum_{\mathbf{k}}{}' v(\mathbf{k}) :a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger:, \quad (7.9)$$

$$J_3 = N_0^{-\frac{1}{2}}V^{-1}\sum_{\substack{\mathbf{k},\mathbf{p}\\(\mathbf{p}+\mathbf{k}\neq 0)}}{}' v(\mathbf{k})a_{\mathbf{p}+\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger a_{\mathbf{p}}. \quad (7.10)$$

Comparison with the pair interaction (3.3) shows that the term carrying J_3 is the new addition in $\mathcal{H}_I$.

The first step of our program is to diagonalize $\mathcal{H}_B$. Restricting our attention to the low-density limit,

$$\kappa = [\rho_0 v(0)^3]^{\frac{1}{2}} \ll 1, \quad (7.11)$$

and assuming that the potential $v(\mathbf{k})$ is slowly varying for $k \lesssim [\rho_0 v(0)]^{\frac{1}{2}}$, we follow Bogoliubov's procedure[3] of taking as a crude substitute for $\mathcal{H}_B$:

$$H_B = \sum_{\mathbf{k}}{}' [f(\mathbf{k})a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \tfrac{1}{2}h(\mathbf{k})(a_{\mathbf{k}}a_{-\mathbf{k}} + a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger)], \quad (7.12)$$

where

$$f(\mathbf{k}) = k^2 + \rho_0 v(\mathbf{k}), \quad h(\mathbf{k}) = \rho_0 v(\mathbf{k}). \quad (7.13)$$

The chemical potential has been chosen so as to eliminate the terms linear in $c^{(\dagger)}$ in $\mathcal{H}_C$ [see (3.14)]. That is, we choose $\chi = 0$, or

$$\mu = (\rho_0 + V^{-1}\sum_{\mathbf{k}}{}' \xi_{\mathbf{k}})v(0) + V^{-1}\sum_{\mathbf{k}}{}' (\xi_{\mathbf{k}} + \eta_{\mathbf{k}})v(\mathbf{k}). \quad (7.14)$$

Note that two approximations have been made to obtain (7.12): First, we have neglected all terms in $f(\mathbf{k})$ and $h(\mathbf{k})$ of (3.11) and (3.12), respectively, that involve momentum summations; Second, we have ignored $\mathcal{H}'_B$. All the effects of these terms are known to be vanishingly small in the low-density limit.[4] Assuming that the $\mathbf{k} \neq 0$ stability condition,

$$v(\mathbf{k}) > 0 \quad (\text{all } \mathbf{k}), \quad (7.15)$$

is satisfied, the diagonalization of (7.12) is achieved by the Bogoliubov transformation (3.22)–(3.25) in conjunction with (7.13). The ground state $|\Omega_B\rangle$ of H_B is the no-particle state of $b_{\mathbf{k}}$ [see (3.26)], and the excitation energy is given by

$$\epsilon(\mathbf{k}) = k[k^2 + 2\rho_0 v(\mathbf{k})]^{\frac{1}{2}}. \quad (7.16)$$

We now proceed to apply the perturbation theory of Appendix A to construct the effective condensate Hamiltonian Λ_0 associated with the ground state $|\Omega_B\rangle$ for the particles outside the condensate. The calculation is identical to that of Sec. IV for the pair-Hamiltonian model, since in the low-density limit (7.11) we can ignore the term in $\mathcal{H}_I$ involving J_3. This follows because the inclusion of J_3 in the perturbation calculation leads to three-particle intermediate states $|\mathbf{p} + \mathbf{k}, -\mathbf{k}, -\mathbf{p}\rangle$ and thus to twofold momentum summations whose values are small, since when (7.11) applies, most of the particles have low momentum. We can therefore borrow the results of Sec. IV and in particular, the effective condensate Hamiltonian Λ_0 is given by (4.19)–(4.25). When we use (7.13) for $f(\mathbf{k})$ and $h(\mathbf{k})$, Eqs. (4.23) and (4.25) become

$$f_0 - h_0 = -\omega_0^2 \frac{\rho_0}{D}\frac{1}{(2\pi)^3}\int d\mathbf{k}\, \frac{v(\mathbf{k})^2}{\epsilon(\mathbf{k})[4\epsilon(\mathbf{k})^2 - \omega_0^2]}, \quad (7.17)$$

$$D = 1 + 2\rho_0 \frac{1}{(2\pi)^3}\int d\mathbf{k}\, \frac{v(\mathbf{k})^2}{\epsilon(\mathbf{k})}\frac{(k^2 - \rho_0 v(0))}{[4\epsilon(\mathbf{k})^2 - \omega_0^2]}. \quad (7.18)$$

Notice that if $\omega_0 = 0$, the integrals in (7.17) and (7.18) have an infrared divergence, because according to (7.16)

$$\epsilon(\mathbf{k}) \xrightarrow[k\to 0]{} ck,$$

where

$$c = [2\rho_0 v(0)]^{\frac{1}{2}} \quad (7.19)$$

is the sound velocity in the present approximation. Nevertheless, the product of D^{-1} and the integral in (7.17) is finite. Furthermore, according to (4.14), $\omega_0^2 = (f_0 - h_0)(f_0 + h_0)$ so that a consistent solution of (7.17) is

$$f_0 - h_0 = 0, \quad (7.20)$$

and thus

$$\omega_0 = 0. \quad (7.21)$$

A similar analysis of $f_0 + h_0$ [see (4.24)] using (7.21) shows that

$$f_0 + h_0 = c^2 = 2\rho_0 v(0). \quad (7.22)$$

This calculation shows the importance of the denominator D, a quantity which has emerged from our perturbation scheme in a very natural manner. A corresponding treatment of the infrared divergence was given by Gavoret and Nozières.[23] An interesting feature here is that despite the special treatment required for the $\mathbf{k} = 0$ mode, $f_0 + h_0$ is smoothly connected[44] to $f(\mathbf{k}) + h(\mathbf{k}) = k^2 + 2\rho_0 v(0)$, [see (7.12)], for $k \to 0$.

Substitution of the coefficients (7.20) and (7.22) into (4.20) yields as the effective condensate Hamiltonian

$$\Lambda_0 = \Lambda_0^0 + \Lambda_0', \quad (7.23)$$

[44] If we use a better approximation for $\epsilon(\mathbf{k})$ such as $\epsilon(\mathbf{k}) = k(k^2 + c^2)^{\frac{1}{2}}$ with the exact value of the sound velocity c, then the formula (7.21) gives $f_0 + h_0 = c^2$ accordingly. This fact was also observed by Gavoret and Nozières. See Ref. 23.

where

$$\Lambda_0^0 = (4M)^{-1}(c + c^\dagger)^2 + W_0 - (4M)^{-1}, \quad (7.24)$$

$$M^{-1} = c^2 = 2\rho_0 v(0), \quad (7.25)$$

and the lowest-order approximation for Λ_0' is given by (4.21).

A remark is in order about the consistency of the above treatment. Since we have employed crude approximate expressions for $f(\mathbf{k})$ and $h(\mathbf{k})$ in the diagonalization of $\mathcal{H}_B$, one could fear that (7.23)–(7.25), our second-order result for Λ_0, is marred by errors in the zeroth- and the first-order calculations. However, our perturbation method of constructing Λ_0 ensures that the second-order result is the same as the zeroth-order condensate Hamiltonian $\mathcal{H}_C$ of (3.13). Thus the second-order calculation is meaningful.

B. General Considerations

The effective condensate Hamiltonian Λ_0 which obtains in the low-density limit (7.11) is given by (7.23). Among those steps taken to obtain this result we set $\chi = 0$ so as to eliminate the term $\chi N_0^{\frac{1}{2}}(c + c^\dagger)$ of (3.14). If this additional term were present in Λ_0 the ground state expectation value of $c^{(\dagger)}$ would be $O(N_0^{\frac{1}{2}})$ in violation of the assumption underlying the use of the operator replacement (1.1). In this subsection we show that results similar to the above obtain when we remove the restriction to low-density systems. In particular, we show that Λ_0 is given by

$$\Lambda_0 = \Lambda_0^0 + \Lambda_0', \quad (7.26)$$

where the most general form of Λ_0^0 is

$$\Lambda_0^0 = \tfrac{1}{2}M(\rho_0, \mu)^{-1}p^2 + 2^{\frac{1}{2}}\chi(\rho_0, \mu)N_0^{\frac{1}{2}}p + U(g) + w_0(\rho_0, \mu)N_0, \quad (7.27)$$

and where p and g are the canonical variables defined by (1.10), and Λ_0' is such that the coefficients of its operator part vanish in the volume limit ($N_0, N, V \to \infty$; N_0/V, N/V finite). The quantities M, χ, and w_0 are c number functions of $O(1)$ dependent upon ρ_0 and μ, while $U(g)$ is an arbitrary function of g. The "physical" values of ρ_0 and μ are determined by the requirements (1.4) and

$$\chi(\rho_0, \mu) = 0, \quad (7.28)$$

thereby ensuring that (1.3) is satisfied and that $c^{(\dagger)} = o(N_0^{\frac{1}{2}})$.

In order to prove the above proposition, let the result of the exact calculation for Λ_0^0 obtained by using the replacement (1.1) be

$$\Lambda_0^0 = F(p, g; \rho_0, \mu). \quad (7.29)$$

Now suppose that we use $N_0^{\frac{1}{2}} - \xi$ and $c + \xi$ in place of $N_0^{\frac{1}{2}}$ and c in (1.1), that is,

$$N_0^{\frac{1}{2}} \to N_0^{\frac{1}{2}} - \xi, \quad c \to c + \xi, \quad (7.30)$$

where ξ is a c number of $O(1)$. Construction of the effective condensate Hamiltonian must go through in exactly the same way as previously for $\xi = 0$, so that the result will be

$$(\Lambda_0^0)' = F(p + 2^{\frac{1}{2}}\xi, g, \rho_0 - 2\xi N_0^{\frac{1}{2}}V^{-1}, \mu). \quad (7.31)$$

But once we determine ρ_0 and μ by (1.3) and (1.4), the two results, (7.29) and (7.31), must coincide since the transformation (7.30) does not change the operator c nor the c number $N_0^{\frac{1}{2}}$ with regard to their relation to a_0. That is to say, the condensate Hamiltonian must be invariant under the transformation (7.30):

$$F(p, g; \rho_0, \mu) = F(p + 2^{\frac{1}{2}}\xi, g; \rho_0 - 2\xi N_0^{\frac{1}{2}}V^{-1}, \mu) \quad (7.32)$$

when ρ_0 and μ take on their physical values, and when $N_0 \to \infty$.

It can easily be seen that in the limit $N_0 \to \infty$ the invariance condition (7.31) is satisfied by (7.27) if and only if

$$M(\rho_0, \mu)^{-1} = 2\rho_0 \partial\chi(\rho_0, \mu)/\partial\rho_0 = 2\rho_0 \partial^2[\rho_0 w_0]/\partial\rho_0^2, \quad (7.33)$$

$$\partial[\rho_0 w_0(\rho_0, \mu)]/\partial\rho_0 = 0, \quad (7.34)$$

where the functions are to be evaluated for the values of μ and ρ_0 which satisfy (7.28). The last condition (7.34) simply means that the condensate density ρ_0 should be determined so as to minimize $W_0 = w_0\rho_0 V$, the ground state energy of Λ_0^0. Reference to (3.10), (3.14), and the expression for χ following (3.15) shows that in the low-density limit $w_0 = \frac{1}{2}\rho_0 v(0) - \mu$ and $\chi = -\mu + \rho_0 v(0)$. In this case (7.33) and (7.34) are obviously satisfied.

The above discussion shows that a condensate Hamiltonian of the form (7.27) *can* satisfy the invariance requirement (7.32). The point is that the change in the Hamiltonian induced by the transformation $c \to c + \xi$ is a linear form of c which can be counterbalanced by the changes in χ and $w_0 N_0$ induced by $N_0^{\frac{1}{2}} \to N_0^{\frac{1}{2}} - \xi$. In general, a Hamiltonian $\Lambda_0^0 = F(p, g; \rho_0, \mu)$ can satisfy (7.32) only when it is invariant modulo a linear form L of p under the transformation $p \to p + 2^{\frac{1}{2}}\xi$. The coefficients of the linear form L must be independent of q. Hence $\partial^2 F/\partial p^2$ must be invariant under $p \to p + 2^{\frac{1}{2}}\xi$. Thus F must be of the form

$$F = a(g; \rho_0, \mu)p^2 + b(g; \rho_0, \mu)p + c(g; \rho_0, \mu).$$

Finally, the linear form L is independent of g and so the same must be true of the coefficients a and b. This completes the proof that (7.27) is the most general form for the condensate Hamiltonian in the volume limit. Incidentally, (7.33) shows that $a(\rho_0, \mu) = \frac{1}{2}M^{-1}$ is to be identified with the self-energy part $\Sigma_2(\mathbf{k}, \omega)$ at $\mathbf{k} = \omega = 0$; cf. Eqs. (107) and (131) of Ref. 22.

Concerning the potential term $U(g)$ of (7.27), some information is provided by Green function theory, to the effect that in a many-boson system described by the full Hamiltonian, the one- and two-particle excitation energies vanish in the low-momentum limit.[5,23] If the interparticle potential $v(\mathbf{k})$ is not singular at $\mathbf{k} = 0$, it can be shown that the excitation spectrum of the full Hamiltonian is continuous down to zero momentum[45] so that the one- and two-particle condensate excitation energy should be zero. In other words, no energy gap separates the ground state from excited states. On this basis, one could expect that the Hamiltonian Λ_0^0 has no normalizable ground state. The result (7.24) obtained with the aid of our perturbation method is in accordance with this expectation.

It may happen, however, that the "infinitesimal" part Λ_0' of the condensate Hamiltonian supplies some terms which are in the nature of a restoring-force potential and which will make the ground state normalizable and the excitation spectrum discrete. In view of the volume dependence of Λ_0' we can expect that the energy gap between the ground state and the states of one- or two-particle excitations will vanish asymptotically as $V \to \infty$. In the next subsection we illustrate this possibility by employing the lowest-order approximation for Λ_0'.

C. Ground State of the Condensate

In order to obtain some idea about the nature of the ground state of the condensate we restrict our attention to the low-density limit so that Λ_0^0 and Λ_0' are given by (7.24) and (7.25), and (3.15), respectively. Thus we study

$$\Lambda_0 = (4M)^{-1}(c + c^\dagger)^2 + \rho_0 v(0)[N_0^{-\frac{1}{2}}(c^\dagger cc + c^\dagger c^\dagger c) + \tfrac{1}{2}N_0^{-1}c^\dagger c^\dagger cc], \quad (7.35)$$

where we have ignored the c number term $W_0 - (4M)^{-1}$ of (7.24). Using (7.25) we can rewrite (7.35) in terms of the original operators $a_0^{(\dagger)}$ for the $\mathbf{k} = 0$ particles as

$$\Lambda_0 = \tfrac{1}{2}N_0^{-1}\rho_0 v(0)[(a_0^\dagger a_0 - N_0 - \tfrac{1}{2})^2 - \tfrac{1}{4}]. \quad (7.36)$$

The form of (7.36) shows (i) that the stability condition for the condensate is $v(0)$, just as for states $\mathbf{k} \neq 0$, and (ii) that when this condition is satisfied the ground state of Λ_0 is given by the eigenstate of $a_0^\dagger a_0$ with the eigenvalue N_0 or $N_0 + 1$. In general we can write

$$|\Omega_c) = \alpha_1 |N_0 + 1) + \alpha_0 |N_0); \quad (7.37)$$

the notation here is self-explanatory.

With this form for $|\Omega_c\rangle$ it becomes impossible to satisfy (1.3) since

$$(\Omega_c| a_0 - N_0^{\frac{1}{2}} |\Omega_c) = \alpha_0^*\alpha_1(N_0 + 1)^{\frac{1}{2}} - (|\alpha_0|^2 + |\alpha_1|^2)N_0^{\frac{1}{2}}, \quad (7.38)$$

and the right side cannot be made to vanish. In general the role of the requirement (1.3) is to determine the value of the chemical potential μ. However, in the present case μ and ρ_0 have already been chosen in accordance with (7.28) so as to remove the term $N_0^{\frac{1}{2}}\chi(c + c^\dagger)$ linear in $c^{(\dagger)}$ which would otherwise appear in (7.35). Note also that the term in question can be written as $N_0^{\frac{1}{2}}\chi(a_0 + a_0^\dagger - 2N_0^{\frac{1}{2}})$ and thus in the present case setting $\chi = 0$ eliminates the term in Λ_0 linear in $a_0^{(\dagger)}$. Further discussion on this point is given in Sec. VIII in connection with problems of the representation of field operators.

It is important to note that the ground state (7.37) is infinitely degenerate. A normalization condition such as $\langle\Omega_c | \Omega_c\rangle = 1$ merely imposes the restriction $\alpha_1 = \cos\varphi$, $\alpha_0 = \sin\varphi$, but where φ is an arbitrary angle. Further, a condition such as $\langle\Omega_c| a_0^\dagger a_0 |\Omega_c\rangle = N_0$ is identically satisfied and does not provide any restriction on φ. This degeneracy is related to the convergence difficulty discussed in Sec. I.

If Λ_0 were calculated beyond the second-order of our perturbation method the result (7.25) may be expected to change. Also the form (7.35) of the condensate Hamiltonian will very likely change. However, in view of a lack of further information, it might be worthwhile just to examine what will happen if we assume (7.35) but not (7.25). For this purpose it is more convenient to use the variables p and g defined by (1.10). Then (7.35) becomes

$$\Lambda_0 = (2M)^{-1}p^2 + \rho_0 v(0)[2^{-\frac{3}{2}}N_0^{-\frac{1}{2}}(pg^2 + g^2 p + 2p^3 - 4p) + \tfrac{1}{8}N_0^{-1}(p^2 + g^2 - 1)(p^2 + g^2 - 3)]. \quad (7.39)$$

The volume dependence of the terms in (7.39) suggests that a scale transformation is needed to obtain canonical variables of $O(1)$ for the ground state. Let

$$p = N_0^{-\nu}\eta, \quad g = N_0^{\nu}\zeta, \quad (7.40)$$

where

$$[\eta, \zeta] = -i, \quad (7.41)$$

ν is a real number, and it is assumed that η and ζ are

[45] See the remark following (7.22).

$O(1)$. Substitution of (7.40) into (7.39) shows that the only useful choice for ν is $\nu = \frac{1}{6}$, in which case (7.39) becomes

$$\Lambda_0 = N_0^{-\frac{1}{3}}\tilde{\Lambda}_0 + R, \tag{7.42}$$

where

$$\tilde{\Lambda}_0 = (2M)^{-1}\eta^2 + \rho_0 v(0) \times [2^{-\frac{3}{2}}(\eta\zeta^2 + \zeta^2\eta) + \tfrac{1}{8}\zeta^4], \tag{7.43}$$

and the residual term R is negligible as compared with $N_0^{-\frac{1}{3}}\tilde{\Lambda}_0$. Note that the scale transformation with $\nu = \frac{1}{6}$ could not have been used if there had been a potential term $U(g)$ of the form $U(g) \sim g^\beta$, $\beta > -2$.

The form of (7.43) can be simplified by introducing a new momentum variable,

$$\pi = \eta + 2^{-\frac{1}{2}}\rho_0 v(0) M\zeta^2, \tag{7.44}$$

which, because of (7.41), is conjugate to ζ

$$[\pi, \zeta] = -i. \tag{7.45}$$

We then obtain

$$\tilde{\Lambda}_0 = (2M)^{-1}\pi^2 + \lambda\zeta^4, \tag{7.46}$$

where

$$\lambda = \tfrac{1}{8}\rho_0 v(0)[1 - 2\rho_0 v(0)M]. \tag{7.47}$$

It is now clear that the Hamiltonian $\tilde{\Lambda}_0$ has a normalizable ground state if

$$M > 0, \quad \lambda > 0. \tag{7.48}$$

Note that if we use the second-order result (7.25) for M, then we obtain

$$\lambda = 0, \tag{7.49}$$

so that $\tilde{\Lambda}_0$, which has been obtained from (7.35) using (7.40), fails to have a normalizable ground state. Note further that according to (7.33), $M > 0$ means $\partial^2 W_0/\partial\rho_0^2 > 0$.

Now, let us assume the stability conditions (7.48) and check the requirement (1.3). Since the Hamiltonian (7.46) is a symmetric function of ζ, its ground state is such that

$$(\Omega_c|\ \zeta\ |\Omega_c) = 0. \tag{7.50}$$

Hence the ground state expectation values of the operators c and $c^\dagger$ are equal. On the other hand, we know from (7.44) that

$$(\Omega_c|\ \eta\ |\Omega_c) = -2^{-\frac{1}{2}}\rho_0 v(0) M\,(\Omega_c|\ \zeta^2\ |\Omega_c). \tag{7.51}$$

Recalling (7.40) and the definition (1.10) of p and q, we obtain

$$(\Omega_c|\ c\ |\Omega_c) = (\Omega_c|\ c^\dagger\ |\Omega_c) = O(N^{-\frac{1}{6}}), \tag{7.52}$$

which vanishes in the volume limit. In view of the steep potential well ζ^4 in the Hamiltonian (7.46), the ground-state expectation values of any products of ζ and π are finite. Recalling the scale transformation (7.40) we see, therefore, that the operators $c^{(\dagger)}$ behave in general like quantities of $O(N^{\frac{1}{6}})$, and thus are much smaller than $N_0^{\frac{1}{2}}$. Hence, it can be concluded that the Bogoliubov replacement $a_0^{(\dagger)} \to N_0^{\frac{1}{2}}$ is justified if (7.39) is a legitimate version of the actual effective condensate Hamiltonian and if the condition (7.48) is satisfied.

We now briefly discuss the case where the stability conditions (7.48) are violated in such a way that $M\lambda < 0$. Applying the WKB method to (7.46) the eigenfunctions $\varphi(\zeta)$ of Λ_0 are found to behave like

$$\varphi(\zeta) \sim \sum_{\pm} A_\pm \zeta^{-1} \exp\,[\pm i\tfrac{1}{3}(2/M\lambda)^{\frac{1}{2}}\zeta^3], \quad (\zeta \to \infty), \tag{7.53}$$

so that this function is normalizable. Unfortunately, if one uses (7.53) the expectation value (1.3) involves divergent integrals, thereby making the qualification of (7.56) as a ground state wavefunction dubious. Further, if the stability conditions (7.48) are violated in such a way that $M < 0$, $\lambda < 0$ then the energy levels of (7.46) have no lower bound.

VIII. REPRESENTATION OF COMMUTATION RELATIONS

In this section we discuss the relevance of our results to the problem of the canonical commutation relations for dynamical systems having infinitely many degrees of freedom. One purpose in studying nonrelativistic many-body problems in the volume limit $N, V \to \infty$ ($N/V = \rho$:fixed) is that we may get some insight into the possible mathematical structure of relativistic quantum field theories[46] where the infinity of degrees of freedom is one of the essential features apparently demanded by experimental facts (e.g., multiple production of particles at extremely high energy).

We start with a remark concerning the Bogoliubov replacement $a_0 \to N_0^{\frac{1}{2}}$ which is often used on the presupposition that a Bose system undergoes a Bose–Einstein (B.E.) condensation at sufficiently low temperatures. This remark serves at the same time to fix our notation. One might feel uneasy about this replacement because it leads to a violation of the commutation relation $[a_0, a_0^\dagger] = 1$. But, as we now show, this is in general not a problem when one is interested in the volume limit.

[46] An operator replacement essentially the same as our (1.1) was once used in quantum field theory by S. Kamefuchi and H. Umezawa, Nuovo Cimento **31**, 429 (1964). There are numerous field theory papers devoted to the discussion of inequivalent representations.

The commutation relations of a Bose field operator $\phi(\mathbf{x})$, e.g., $[\phi(\mathbf{x}), \phi^\dagger(\mathbf{y})] = \delta(\mathbf{x}-\mathbf{y})$, have to be interpreted in the sense of distributions. Thus we take a set of test functions $\tilde{f}(\mathbf{x})$ to define smeared-out operators,

$$\phi(f) = \int \tilde{f}(\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x}, \tag{8.1}$$

for which the commutation relations are

$$[\phi(f), \phi^\dagger(g)] = \int \tilde{f}(\mathbf{x})\tilde{g}^*(\mathbf{x})\, d\mathbf{x}, \tag{8.2}$$

$$[\phi(f), \phi(g)] = 0. \tag{8.3}$$

For particles confined to a box of volume V we require that $\tilde{f}_V(\mathbf{x})$ vanish outside V. Further, these test functions must be square integrable, so that (8.2) will be sensible, and also integrable,

$$\int \tilde{f}_V(\mathbf{x})\, d\mathbf{x} \leq M(\leq \infty), \quad \text{all } V, \tag{8.4}$$

in order that they be of use in the case of B.E. condensation. If in the volume V one were to take a set of normalized plane waves satisfying periodic boundary conditions, then the smeared-out operators (8.1) are the usual destruction operators $a_\mathbf{k}$.[47] In general we may write

$$\phi(f_V) = V^{-\frac{1}{2}} \sum_\mathbf{k} f_V(\mathbf{k}) a_\mathbf{k}, \tag{8.5}$$

where $f_V(\mathbf{k})$ is the Fourier transform of $\tilde{f}_V(\mathbf{x})$. Now if one takes $a_0^{(\dagger)} = N_0^{\frac{1}{2}}$, the operator (8.5) is replaced by

$$\phi_B(f_V) = \rho_0^{\frac{1}{2}} f_V(0) + V^{-\frac{1}{2}} \sum_\mathbf{k}{}' f_V(\mathbf{k}) a_\mathbf{k}, \tag{8.6}$$

where as usual the prime on the summation symbol indicates that the term $\mathbf{k} = 0$ should be excluded from the sum. Note that the assumption of (8.4) is needed to make the first term of (8.6) well defined. With the aid of (8.4) we immediately see that, when $V \to \infty$, $\phi_B(f_v)$ satisfies the same commutation relations satisfied by $\phi(f_V)$. Thus the Bogoliubov replacement does not violate the commutation relations in the volume limit.

One of the characteristic features of an infinite system is that there are a variety of inequivalent representations of the commutation relations (8.2), (8.3), and not all of them are appropriate for the quantum mechanical description of a dynamical system with a given Hamiltonian.[12] The Hamiltonian chooses a representation, so to speak. It can be proved, for instance, that operators $\phi(\mathbf{x})$ and $\rho_0^{\frac{1}{2}} + \phi(\mathbf{x})$ cannot be related by any unitary transformation if $\rho_0 \neq 0$, $(V \to \infty)$. The Bogoliubov replacement implies the use of a field operator of the form,

$$\phi(\mathbf{x}) = \rho_0^{\frac{1}{2}} + \phi'(\mathbf{x}), \quad \langle\Omega|\, \phi'(\mathbf{x})\, |\Omega\rangle = 0, \tag{8.7}$$

whatever the representation for $\phi'(\mathbf{x})$ may be. Here $|\Omega\rangle$ is the ground state of the dynamical system in question. Such a field operator may be said to be of the Bogoliubov type. But does the existence of a B.E. condensation always demand a field operator of this form? If, furthermore, we ask whether a given Hamiltonian of a many-boson system requires a representation of the form (8.7), there are actually two questions involved. The first is the one asked just above. The second question is whether or not the system described by the given Hamiltonian undergoes a B.E. condensation.

A method for finding a suitable representation has been suggested by Araki and Woods.[7] They proposed that one first calculate a system of Wightman functionals

$$W_V(f_V^1, \cdots, f_V^r, \cdots, g_V^s) = \langle\Omega(V)|\, \phi(f_V^1)\cdots\phi(f_V^r)\phi^\dagger(g_V^1)\cdots\phi^\dagger(g_V^s)\, |\Omega(V)\rangle \quad (r, s = 0, 1, 2, \cdots), \tag{8.8}$$

for a finite dynamical system of volume V by employing its ground state $|\Omega(V)\rangle$, and then analyze their volume limits $(V \to \infty)$ to see what representation is implied. Underlying this procedure is a theorem,[9,48] that once a set of Wightman functionals satisfying certain conditions are given, we can reconstruct a Hilbert space and field operators which reproduce the given set of functionals, and this construction is unique up to unitary transformations (Gel'fand construction). Araki and Woods used this method to find a representation suitable for an infinite *free* Bose gas. They took $|\Omega(V)\rangle$ to be the eigenstate of the number operators $a_\mathbf{k}^\dagger a_\mathbf{k}$, where all the $N = \rho V$ particles are occupying the lowest single-particle states $\mathbf{k} = 0$ (B.E. condensation). Then the limiting representation turns out to be fully reducible: It becomes a direct sum of representation $\vartheta_\alpha (0 \leq \alpha \leq 2\pi)$. The field operator in ϑ_α is of the Bogoliubov type $\phi(\mathbf{x}) = \rho_0^{\frac{1}{2}} e^{i\alpha} + \phi_F(\mathbf{x})$, where $\phi_F(\mathbf{x})$ is a Fock operator.

In the previous sections we have been studying the operator replacement (1.1), that is,

$$a_0 = N_0^{\frac{1}{2}} + c, \tag{8.9}$$

with a particular interest in the order of magnitude

[47] This set of single-particle wavefunctions fails to satisfy (8.4), and thus $\phi_B(f_V)$ does not satisfy (8.2) and (8.3) if the volume limit of the field operator is taken.

[48] A. S. Wightman, Phys. Rev. **101**, 860 (1956); M. A. Naimark, *Normed Rings*, translated by L. F. Baron (P. Nordhoff Ltd., Groningen, The Netherlands, 1959).

of the operator part c as $N_0 \to \infty$. In the case of the pair Hamiltonian model we have found that under certain conditions (stability conditions), restricting the particle density and the interparticle potential, the Hamiltonian has a normalizable ground state $|\Omega(V)\rangle$ such that

$$\langle\Omega(V)|\, c^{(\dagger)}\, |\Omega(V)\rangle = 0 \tag{8.10}$$

and

$$\langle\Omega(V)|\, (c^\dagger)^i (c)^j\, |\Omega(V)\rangle = O(1) \quad (i+j \geq 2). \tag{8.11}$$

This result implies, first of all, that the model system undergoes a B.E. condensation. Because of (8.10) and (8.11), all the terms in the Wightman functionals involving the operators,

$$f_V(0)c/V^{\frac{1}{2}}, \quad g_V^*(0)c^\dagger/V^{\frac{1}{2}}, \tag{8.12}$$

vanish in the volume limit, so that the system (8.8) implies a representation of the field operator of the Bogoliubov type (8.7). Moreover, explicit calculation, in the limit $V \to \infty$, of (8.8) shows that the limiting functionals are identical with what we obtain by taking a linear combination of a Fock operator ϕ_F,

$$\phi(\mathbf{x}) = \rho_0^{\frac{1}{2}} + \int \alpha(\mathbf{x}-\mathbf{y})\phi_F(\mathbf{y})\, d\mathbf{y} + \int \beta(\mathbf{x}-\mathbf{y})\phi_F^\dagger(\mathbf{y})\, d\mathbf{y}, \tag{8.13}$$

for the field operator, and the vacuum state of Fock space for the ground state. The quantities $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are the Fourier transforms of the coefficients of the Bogoliubov transformation (3.22). Therefore, the pair-Hamiltonian model in the volume limit can be described by the representation (8.13) as long as the stability conditions are satisfied. The representation (8.13) is obviously irreducible. It is interesting to recall here that we met with a dilemma in the case where $\tilde{v}(\mathbf{x})$ is weak and everywhere repulsive in coordinate space. In this case, when the operator replacement (8.9) is made, one cannot obtain a normalizable ground state for the condensate, yet we can prove that the assumption of *no* condensation is untenable. Perhaps one has to generalize the concept of B.E. condensation for such cases. The other possibility is that another operator replacement in place of (8.9) will be successful.

In our study of the full Hamiltonian in the low-density limit, we met the interesting situation where the effective condensate Hamiltonian Λ_0 has the form (7.36), that is

$$\Lambda_0 = \tfrac{1}{2}N_0^{-1}\rho_0 v(0)[(a_0^\dagger a_0 - N_0 - \tfrac{1}{2})^2 - \tfrac{1}{4}], \tag{8.14}$$

which is diagonal only in a representation in which $a_0^\dagger a_0$ is itself diagonal. In contrast to the case of the *stable* pair Hamiltonian, it is the occupation-number space of a_0, but not of c, that contains the ground state of Λ_0. The ground state of the present system has a B.E. condensation whenever $v(0) > 0$, yet it fails to admit the Bogoliubov-type representation. In fact (7.38), in conjunction with the normalization condition $\langle\Omega(V)\,|\,\Omega(V)\rangle = 1$, tells us that

$$|\langle\Omega(V)|\, N_0^{-\frac{1}{2}} a_0 - 1\, |\Omega(V)\rangle| \geq \tfrac{1}{2} + 0(\kappa), \tag{8.15}$$

where κ is the density parameter $[\rho_0 v(0)^3]^{\frac{1}{2}}$. This result presents no difficulties as long as $V < \infty$. But, as explained in Sec. I, we meet with a convergence difficulty if we consider the volume limit; the ground state (7.37) keeps rotating in the occupation number space as N_0 increases and there is no convergence. This implies that the occupation number representation is inadequate for the infinite system. Since the ground state of the condensate is given by (7.37) the calculation of the Wightman functional (8.8) proceeds in much the same way as in the case of the free Bose gas.[7] The volume limit of (8.8) using the occupation number representation of $a_0^\dagger a_0$ turns out to be fully reducible. The irreducible representations for the field operator are of the Bogoliubov type $\rho_0^{\frac{1}{2}} e^{i\alpha} + \phi'(\mathbf{x})$ $(0 \leq \alpha < 2\pi)$. But as we have seen in the above, the Hamiltonian cannot be diagonalized in such a representation. In connection with (8.15) we note here that if we wish to satisfy

$$\langle\Omega'|\, a_0 - N_0^{\frac{1}{2}}\, |\Omega'\rangle = 0 \tag{8.16}$$

by

$$|\Omega'\rangle = \sum_n \alpha_n\, |N_0 + n\rangle, \tag{8.17}$$

where $|n\rangle$ is an eigenstate of $a_0^\dagger a_0$ with eigenvalue n, then we must have

$$\cdots = \alpha_{N_0-1} = \alpha_{N_0} = \alpha_{N_0+1} = \cdots, \tag{8.18}$$

equality extending in both right and left directions indefinitely. From this result we see again that the Hamiltonian (8.14) cannot be diagonalized in a Bogoliubov-type representation.

When the density parameter κ is not very small, our discussion in Sec. VIIC concerning the ground state of the full Hamiltonian is based on conjecture. It appears likely that under certain conditions the operator replacement (8.9) gives a normalizable ground state in which

$$\langle\Omega(V)|\, (c^\dagger)^i (c)^j\, |\Omega(V)\rangle = O(N^{\frac{1}{3}(i+j)}) \quad (i+j \geq 2). \tag{8.19}$$

In the volume limit the representation is then of the Bogoliubov type (8.7).

Returning now to the question of the existence of

the B.E. condensation we note here a problem that is left for future investigation. Throughout the above discussions we have taken the point of view that an infinite system should be studied as the limit of a finite system. One might try, however, to consider an infinite system from the outset by taking a field operator of the form of (8.7), whose commutation relations are exactly accounted for by the corresponding relation for $\phi'(\mathbf{x})$, e.g.,

$$[\phi(\mathbf{x}), \phi^\dagger(\mathbf{y})] = [\phi'(\mathbf{x}), \phi'^\dagger(\mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}). \quad (8.20)$$

Let $\mathcal{H}'_B$ denote the Hamiltonian which we obtain by substituting (8.7) into the given Hamiltonian of a system. Then the existence criterion for the B.E. condensation is stated as the existence of a finite condensate density ρ_0 that makes the lowest eigenvalue W'_B of $\mathcal{H}'_B$ extremum[5]:

$$\partial W'_B(\rho_0, \mu)/\partial\rho_0 = 0, \quad (8.21)$$

where μ is a chemical potential. This *ad hoc* requirement added to the eigenvalue problem of $\mathcal{H}'_B$ is motivated by a rather reasonable physical consideration. One would further require, for instance, that the extremum given by (8.21) should actually be a minimum and that the sound velocity be real. It is then interesting to ask how these *ad hoc* requirements compare with the stability conditions of our scheme. We know at least that (8.21) is a necessary consequence of the stability of the condensate. This conclusion was obtained by an invariance argument in Sec. VII. Moreover, the condition as to the reality of the sound velocity is needed also in our scheme because we too have to diagonalize the part $\mathcal{H}_B$ of the Hamiltonian pertaining to particles outside the condensate. But, it is not yet clear whether this is the whole story or not. Judging from the case of the pair Hamiltonian it may well be that the set of stability conditions for the effective condensate Hamiltonian is more stringent than the set of conditions mentioned above. If this is so, then we have to check the existence of the B.E. condensation, case by case, before we can use the Bogoliubov replacement. If, on the contrary, we can show the equivalence of these two sets of conditions it will mean a real justification of the Bogoliubov replacement and furthermore of the method of Hugenholtz–Pines[5] and Beliaev.[6] We hope that some results will be obtained by widening the invariance argument of Sec. VII.

In conclusion, we have obtained certain conditions which when satisfied prove the existence of a Bose–Einstein condensation and justify the Bogoliubov replacement $a_0 = N_0^{\frac{1}{2}}$ for the case of the pair Hamiltonian model. In the case of purely repulsive potentials the operator replacement $a_0 = N_0^{\frac{1}{2}} + c$ does not lead to a stable condensate, yet the assumption of *no* condensation is untenable. This case presents a challenging problem to the representation theory of canonical commutation relations. As for the full Hamiltonian, we have discovered that in the low-density limit it cannot be diagonalized in any irreducible representation of the canonical commutation relations. No clear criterion has been obtained for the existence of a B.E. condensation for the full N-body Hamiltonian, except the one, $v(0) > 0$, in the low-density limit.

ACKNOWLEDGMENTS

We are grateful to the Physics Division of the Aspen Institute for Humanistic Studies which provided us with an opportunity to meet and begin collaboration. Thanks are due to Professor R. Haag for his encouragement at the initial stage of this work. One of the authors (H. E.) started this work while he was staying at the University of Illinois. He wishes to express his gratitude to Professor G. M. Almy for his hospitality at Illinois. He would also like to thank Professor J. Martin of the University of Wisconsin–Milwaukee for his hospitality. His visit to the U.S. was made possible by a Fulbright Travel Grant, which he acknowledges gratefully.

APPENDIX A. PERTURBATION FORMULAS

In this appendix we present a perturbation method to solve the eigenvalue problem formulated in Sec. II:

$$(\mathcal{H}_B + \mathcal{H}_C + \mathcal{H}_I)\psi_n = \psi_n\Lambda_n, \quad (A1)$$

$$\psi_n^\dagger\psi_n = \mathbf{1}, \quad (A2)$$

where $\mathbf{1}$ is the identity operator in $\mathfrak{H}_C$. We refer the reader to that section for the notation to be employed here. In particular the meaning of the product in (A2) is given by (2.5). We assume that the complete set of orthonormal eigenvectors $|n)$ and eigenvalues ϵ_n of $\mathcal{H}_B$ are known. Further, we suppose that $\mathcal{H}_I$ has the form

$$\mathcal{H}_I = \lambda H_1 + \lambda^2 H_2, \quad (A3)$$

where λ is a small real parameter.

Consider an eigenvector $|n)$ of $\mathcal{H}_B$ which we assume to be nondegenerate, $\epsilon_m \neq \epsilon_n$ $(m \neq n)$. We suppose that the perturbed "eigenvector" ψ_n and "eigenvalue" Λ_n can be expanded as

$$\begin{aligned} \psi_n &= (\mathbf{1} + \lambda\mathfrak{U}_n^{(1)} + \lambda^2\mathfrak{U}_n^{(2)} + \cdots)\,|n), \\ \Lambda_n &= \Lambda_n^{(0)} + \lambda\Lambda_n^{(1)} + \lambda^2\Lambda_n^{(2)} + \cdots, \end{aligned} \quad (A4)$$

where the $\mathfrak{U}_n^{(i)}$ are operators in $\mathfrak{H}_B \otimes \mathfrak{H}_C$ and the $\Lambda_n^{(i)}$ are operators in $\mathfrak{H}_C$. In the following we suppress the index n wherever no confusion can arise.

As is the case for ordinary Schrödinger perturbation theory, ψ_n and Λ_n are not uniquely determined by (A1) and (A2). In fact, if ψ_n and Λ_n satisfy these equations, so do ψ_n' and Λ_n' defined by

$$\psi_n' = \psi_n \mathcal{U}, \quad \Lambda_n' = \mathcal{U}^\dagger \Lambda_n \mathcal{U}, \tag{A5}$$

where $\mathcal{U}$ is a unitary transformation in $\mathfrak{H}_C$. Thus, without loss of generality, one can require that

$$(n|\, \mathfrak{U}^{(1)}\, |n) = 0, \tag{A6}$$

and that the operator $(n|\, \mathfrak{U}^{(2)}\, |n)$ defined in $\mathfrak{H}_C$ is Hermitian, i.e.,

$$(n|\, \mathfrak{U}^{(2)}\, |n) = (n|\, \mathfrak{U}^{(2)\dagger}\, |n). \tag{A7}$$

To prove these assertions we begin by noting that the isometry condition (A2) implies that

$$(n|\, \mathfrak{U}^{(1)} + \mathfrak{U}^{(1)\dagger}\, |n) = 0, \tag{A8}$$

$$(n|\, \mathfrak{U}^{(1)\dagger}\mathfrak{U}^{(1)} + \mathfrak{U}^{(2)\dagger} + \mathfrak{U}^{(2)}\, |n) = 0, \quad \text{etc.} \tag{A9}$$

Now if we suppose that $(n|\, \mathfrak{U}^{(1)}\, |n) \neq 0$, then the first-order term of

$$\begin{aligned}\psi_n' &= \psi_n \exp\,[-\lambda\,(n|\, \mathfrak{U}^{(1)}\, |n)] \\ &= [1 + \lambda\mathfrak{U}^{(1)} + \cdots][1 - \lambda\,(n|\, \mathfrak{U}^{(1)}\, |n) + \cdots]\,|n)\end{aligned} \tag{A10}$$

is

$$\mathfrak{U}^{(1)'}\, |n) = [\mathfrak{U}^{(1)} - (n|\, \mathfrak{U}^{(1)}\, |n)]\, |n),$$

so that $(n|\, \mathfrak{U}^{(1)'}\, |n) = 0$ in accord with (A6). Moreover, (A10) is a unitary transformation because

$$\begin{aligned}\{\exp\,[-\lambda\,(n|\, \mathfrak{U}^{(1)}\, |n)]\}^\dagger &= \exp\,[-\lambda\,(n|\, \mathfrak{U}^{(1)\dagger}\, |n)] \\ &= \exp\,[\lambda(n|\, \mathfrak{U}^{(1)}\, |n)],\end{aligned}$$

the second equality following from (A8). Thus we see that it is always possible to fulfill the requirement of (A6).

Turning to (A7), suppose that $(n|\, \mathfrak{U}^{(2)}\, |n)$ is not Hermitian. Then consider another unitary transformation:

$$\psi_n'' = \psi_n \exp\,(-\lambda^2 A), \tag{A11}$$

where

$$A = \tfrac{1}{2}[(n|\, \mathfrak{U}^{(2)}\, |n) - (n|\, \mathfrak{U}^{(2)\dagger}\, |n)]$$

is the anti-Hermitian part of $(n|\, \mathfrak{U}^{(2)}\, |n)$. The second-order term of ψ_n'' is given by

$$\mathfrak{U}^{(2)''}\, |n) = \mathfrak{U}^{(2)}|n) - |n)\, A,$$

so that $(n|\, \mathfrak{U}^{(2)''}\, |n)$ is easily seen to be Hermitian. This completes the proof of (A6) and (A7). Hereafter we assume that (A6) and (A7) are satisfied for the ψ_n in (A4). Finally, note that the isometry condition (A9) now becomes

$$\begin{aligned}(n|\, \mathfrak{U}^{(2)\dagger}\, |n) &= (n|\, \mathfrak{U}^{(2)}\, |n) \\ &= -\tfrac{1}{2}(n|\, \mathfrak{U}^{(1)\dagger}\, \mathfrak{U}^{(1)}\, |n).\end{aligned} \tag{A12}$$

Returning to the eigenvalue problem, (A1) is equivalent to the set of equations

$$(\mathcal{H}_B + \mathcal{H}_C)\, |\psi^{(0)}) = |\psi^{(0)})\, \Lambda^{(0)}, \tag{A13}$$

$$(\mathcal{H}_B + \mathcal{H}_C)\, |\psi^{(1)}) + H_1\, |\psi^{(0)}) = |\psi^{(1)})\, \Lambda^{(0)} + |\psi^{(0)})\, \Lambda^{(1)}, \tag{A14}$$

$$\begin{aligned}&(\mathcal{H}_B + \mathcal{H}_C)\, |\psi^{(2)}) + H_1\, |\psi^{(1)}) + H_2\, |\psi^{(0)}) \\ &\quad = |\psi^{(2)})\, \Lambda^{(0)} + |\psi^{(1)})\, \Lambda^{(1)} + |\psi^{(0)})\, \Lambda^{(2)}, \quad \text{etc.},\end{aligned} \tag{A15}$$

where $|\psi^{(0)}) = |n)$ and $|\psi^{(i)}) = \mathfrak{U}^{(i)}\, |n)$, $i = 1, 2, \cdots$. From (A13) we find

$$\Lambda^{(0)} = \epsilon_n + \mathcal{H}_C. \tag{A16}$$

Furthermore, multiplying (A14) from the left by $(n|$ and using (A6) yields

$$\Lambda^{(1)} = (n|\, H_1\, |n). \tag{A17}$$

The similarity between (A17) and the standard formula for the first-order energy corrections is only formal since, instead of being a c number, $\Lambda^{(1)}$ is an operator in $\mathfrak{H}_C$.

To obtain $\Lambda^{(2)}$ we multiply (A15) from the left by $(n|$ and we find

$$\Lambda^{(2)} = [\mathcal{H}_C, (n\,|\,\psi^{(2)})] + (n|\, H_1\, |\psi^{(1)}) + (n|\, H_2\, |n).$$

But, according to (A7), the operator $(n\,|\,\psi^{(2)})$ is Hermitian in $\mathfrak{H}_C$ so that the commutator term in $\Lambda^{(2)}$ is anti-Hermitian. Since Λ must be Hermitian [see (2.6)] we can conclude that

$$\Lambda^{(2)} = \text{H.P.}\, (n|\, H_1\, |\psi^{(1)}) + (n|\, H_2\, |n), \tag{A18}$$

where H.P. stands for Hermitian part.

In order to find $|\psi^{(1)})$ we multiply (A14) from the left by $(m|$ with the result,

$$\begin{aligned}&(\epsilon_n - \epsilon_m)(m\,|\,\psi^{(1)}) \\ &\quad = (m|\, H_1\, |n) + [\mathcal{H}_C, (m\,|\,\psi^{(1)})], \quad (m \neq n).\end{aligned} \tag{A19}$$

In the case of ordinary perturbation theory, $(m\,|\,\psi^{(1)})$ is a c number so that the commutator term is absent. The method of solving (A19) for $(m\,|\,\psi^{(1)})$ can be found in the text.

APPENDIX B. VIOLATION OF STABILITY CONDITIONS

We briefly discuss what happens in the pair-Hamiltonian model if the $\mathbf{k} = 0$ stability conditions are violated. The $\mathbf{k} \neq 0$ stability conditions are assumed to be satisfied; in particular we assume $v(0) > 0$. Thus we study

$$\begin{aligned}\Lambda_0 &= f_0 c^\dagger c + \tfrac{1}{2}h_0(cc + c^\dagger c^\dagger) \\ &\quad + \rho_0 v(0)[N_0^{-\frac{1}{2}}(c^\dagger cc + \text{H.c.}) + \tfrac{1}{2}N_0^{-1}c^\dagger c^\dagger cc]\end{aligned} \tag{B1}$$

for several choices of $f_0 \pm h_0$.

Case 1: $f_0 + h_0 > 0$ and $f_0 - h_0 = 0$. In this case the situation is the same as the one we discussed in Sec. VII. If, (a), $f_0 > \rho_0 v(0)$, then the condensate is stable and $c = O(N^{\frac{1}{2}})$. If on the contrary, (b), $f_0 < \rho_0 v(0)$, then the condensate Hamiltonian formally constructed has a normalizable ground state, but the ground-state expectation values of $c^{(\dagger)}$ involve divergent integrals in contradiction to the fundamental requirement (1.3). If, (c), $f_0 = \rho_0 v(0)$, then Λ_0 has the form of (7.39).

Case 2: $f_0 + h_0 < 0$ and $f_0 - h_0 = 0$. In this case the spectrum of the condensate Hamiltonian Λ_0 is not bounded below.

Case 3: $f_0 + h_0 = 0$ and $f_0 - h_0 \neq 0$. The method of scale transformation as applied to the variables p, q of (1.10) leads to

$$\Lambda_0 = N_0^{-\frac{1}{2}}[2f_0\xi^2 + 2^{-\frac{1}{2}}\rho_0 v(0)\eta^3], \tag{B2}$$

where we have put $q = N_0^{-\frac{1}{4}}\xi$ and $p = N_0^{\frac{1}{4}}\eta$. The Hamiltonian (B2) has no normalizable ground state irrespective of the sign of f_0.

Case 4: $f_0 + h_0$ and $f_0 - h_0$ have opposite signs or are both negative. In this case the finite part of the Hamiltonian (B1),

$$\Lambda_0^0 = \tfrac{1}{2}(f_0 + h_0)p^2 + \tfrac{1}{2}(f_0 - h_0)q^2 - f_0, \tag{B3}$$

has no normalizable ground state. The "infinitesimal" part $\Lambda_0 - \Lambda_0^0$ may act as a restoring force, but it can compete with Λ_0^0 only when $p, q = O(N^{\frac{1}{2}})$. Whenever cases (2)–(4) apply, either there is no B.E. condensation in the system or there is a condensation but the operator replacement (1.1) loses its usefulness.

APPENDIX C. THE FUNCTION $h(\mathbf{p})$

Before discussing the solution of the integral equation (3.31) for the case of a superposition of Yukawa potentials [see (6.1)], we briefly describe how to solve that equation when $v(\mathbf{k})$ describes a single Yukawa potential,

$$v(\mathbf{k}) = B/(k^2 + \lambda^2). \tag{C1}$$

In accordance with (6.10) we assume that

$$\epsilon(\mathbf{k}) = \epsilon(0) + k^2/(2m). \tag{C2}$$

Thus the integral equation (3.31) becomes

$$\bar{h}(\mathbf{p}) = v(\mathbf{p}) - \int d\mathbf{k} G(\mathbf{p}, \mathbf{k})\bar{h}(\mathbf{k}), \tag{C3}$$

where $\bar{h} = \rho_0^{-1}h$ and the kernel is given by

$$G(\mathbf{p}, \mathbf{k}) = (2\pi)^{-3}mB[(\mathbf{p} - \mathbf{k})^2 + \lambda^2]^{-1}(k^2 + \sigma^2)^{-1}, \tag{C4}$$

and $\sigma^2 = 2m\epsilon(0)$.

The kernel has the remarkable property that in a good approximation it reproduces functions of Yukawa type; that is,

$$\int G(\mathbf{p}, \mathbf{k})(k^2 + \xi^2)^{-1}\, d\mathbf{k} \simeq \beta(\xi)(p^2 + M^2)^{-1}, \tag{C5}$$

where

$$M^2 = (\lambda + \sigma)(\lambda + \xi), \tag{C6}$$

$$\beta(\xi) = (4\pi)^{-1}(\xi + \sigma)^{-1}mB. \tag{C7}$$

For the proof, we evaluate the integral of (C5) by invoking the convolution theorem for Fourier transforms; the integral is transformed to a sum of two integrals, each containing the product of two Yukawa potentials in coordinate space. We obtain

$$\int d\mathbf{k} G(\mathbf{p}, \mathbf{k})(k^2 + \xi^2)^{-1} = mB[4\pi p(\xi^2 - \sigma^2)]^{-1}\tan^{-1}[(\xi - \sigma)p/(p^2 + M^2)]. \tag{C8}$$

But the inequality $pM \leq \frac{1}{2}(p^2 + M^2)$ shows that the argument of $\tan^{-1}$ never exceeds

$$\tfrac{1}{2}(\xi - \sigma)(\lambda + \xi)(\lambda + \sigma)^{\frac{1}{2}},$$

which, in turn, is less than $\frac{1}{2}$ in the following calcu lations. Thus in the following (C5) is in fact a good approximation to (C8).

In view of this reproducing property of the kernel, we try to obtain the solution of (C3) in the form

$$\bar{h}(\mathbf{p}) = \sum_{n=0}^{\infty} a_n(p^2 + \lambda_n^2)^{-1}, \tag{C9}$$

where $0 < \lambda_0 < \lambda_1 < \cdots$ as well as the a_n are parameters to be determined. In particular, the first term of this sum is taken to be $v(\mathbf{p})$, so that $a_0 = B$, $\lambda_0 = \lambda$. Substituting (C9) into (C3) we obtain the recurrence relations

$$a_{n+1} = -\beta(\lambda_n)a_n, \tag{C10}$$

$$\lambda_{n+1}^2 = (\lambda + \lambda_n)(\lambda + \sigma). \tag{C11}$$

In terms of $\gamma_n = \lambda_n/(\lambda + \sigma)$, the latter relation becomes

$$\gamma_{n+1}^2 = \gamma_n + \gamma_0, \quad \gamma_0 = \lambda/(\lambda + \sigma), \tag{C12}$$

and the the sequence γ_n converges to

$$\gamma_\infty = \tfrac{1}{2}[1 + (1 + 4\gamma_0)^{\frac{1}{2}}]. \tag{C13}$$

Every γ_n lies in the interval $[\gamma_0, \gamma_\infty)$, and the length

of the interval decreases as σ decreases. A good feature of the sequence γ_n is that it converges very rapidly. For instance, when $\sigma = 0$, $\gamma_0 = 1$, the sequence is as follows: 1, $2^{\frac{1}{2}}$, 1.55, 1.60, $\cdots$, and eventually $\gamma_\infty = 1.62$. The rapid convergence of γ_n suggests that it would be a good approximation even if we set $\gamma_1 = \gamma_2 = \cdots = \gamma_\infty$. Then,

$$\bar{h}(\mathbf{p}) = v(\mathbf{p}) - \bar{\beta}B \cdot (p^2 + \lambda_\infty^2)^{-1}, \tag{C14}$$

where

$$\lambda_\infty = (\lambda + \sigma)\gamma_\infty \tag{C15}$$

and

$$\bar{\beta} = \beta(\lambda)/[1 + \beta(\lambda_\infty)]. \tag{C16}$$

Now we turn to obtaining the solution of (3.31) when $v(\mathbf{k})$ has the form (6.1); that is,

$$v(\mathbf{k}) = B/(k^2 + \lambda^2) + A/(k^2 + \kappa^2). \tag{C17}$$

In this case a good approximate solution is

$$\bar{h}(\mathbf{p}) = B/(p^2 + \lambda^2) + A/(p^2 + \kappa^2) - \bar{\beta}B/(p^2 + \lambda_\infty^2). \tag{C18}$$

This result follows from the fact that if (3.31) is solved by iteration with (C18) as the lowest order form for $\bar{h}(\mathbf{p})$, the correction terms are very small in view of (6.2b) and (6.4).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Renormalized Coupling Constants and the Inversion Problem

LUCIANO FONDA AND GIAN CARLO GHIRARDI
Istituto di Fisica Teorica dell'Università, Trieste, Italy

(Received 19 May 1966)

It is shown how, starting from the (experimental) knowledge of scattering phase shift, energies of bound states and renormalized coupling constants, one is able to determine completely the parameters of a quantum field theoretic model previously considered by the authors, the so-called Dyson model. The very same conclusion holds for the case of potential scattering, which is also briefly considered.

1. INTRODUCTION

THE problem of determining the parameters of some field theoretic models from the (experimental) knowledge of the scattering phase shift and the energies of the bound states has been previously considered by the authors.[1] In I it was found that the problem did not have, in general, a unique solution. In this paper we again consider this matter by also assuming the (experimental) knowledge of the renormalized coupling constants.

It is shown in Sec. 2 how this knowledge, together with that of the phase shift over the full energy range and that of the bound state energies, is able to determine completely the parameters of the models.

In Sec. 3 the concept of renormalized coupling constant is introduced in potential scattering[2,3] and it is briefly shown how also in this case the potential is uniquely determined.

2. FIELD THEORETIC MODEL

We start by recalling the main results of I. It has been shown there that if we assign the phase shift $\delta(\omega)$ in such a way that it has one of the behaviors A, B', or B'' shown in Fig. 3 of that paper, and the number n and the energies of the bound states, we can reproduce this situation by means of the Dyson model, whose Hamiltonian is[1]

$$H = H_0 + H_I,$$

$$H_0 = m_N \psi_N^\dagger \psi_N + \sum_{i=1}^{N_0} (m_N + \omega_i^{(0)}) \psi_i^\dagger \psi_i + \int \omega(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k})\, d^3k, \quad (2.1)$$

$$H_I = (2\pi)^{-\frac{3}{2}} \int \frac{f(\omega)}{(2\omega)^{\frac{1}{2}}} \left[\sum_{i=1}^{N_0} g_i \psi_i^\dagger \psi_N a(\mathbf{k}) + \text{h.c.} \right] d^3k,$$

by choosing in a proper way the parameters appearing in it.

The *a priori* unknown quantities of the problem are:

(a) the number N_0 of excited states of the scatterer;

(b) the N_0 energies $\omega_i^{(0)}$ of these excited states;

(c) the N_0 unrenormalized coupling constants g_i;

(d) the form factor $f(\omega)$, which is subject to the condition $f(\mu) = 1$.

The quantities which are known are, the number n and the energies of the bound states of the system and the scattering phase shift $\delta(\omega)$ over the full energy range, i.e., from $\omega = \mu$ to $\omega = +\infty$.

As shown in I, the number N_0 is immediately obtained; it is equal to;

$$N_0 = n + P_0 \quad (2.2a)$$

if the phase shift starts decreasing (cases A and B''), and equal to

$$N_0 = n + P_0 + 1 \quad (2.2b)$$

if the phase shift starts increasing (case B') being P_0 the number of times that $\sin \delta(\omega)$ vanishes in the interval $\mu < \omega < +\infty$, i.e., the number of times $\delta(\omega)$ goes through a multiple of π excluding the points $\omega = \mu$ and $\omega = +\infty$. We outline again here that the phase shift cannot in any case go through the same value multiple of $-\pi$ more than once, otherwise the assigned situation cannot be reproduced by this model. The fact that N_0 is connected to n and P_0 through (2.2a) and (2.2b) is an immediate consequence of the fact that for this model the Levinson theorem holds.

Instead of using the $2N_0$ parameters $\omega_i^{(0)}$ and g_i we can use the $\omega_i^{(0)}$'s, $\sum_{i=1}^{N_0} g_i^2$ and the $N_0 - 1$ CDD poles $z_i^{(0)}$ of the model. The relation between these two sets of parameters is obtained through the identity

$$\left[\sum_{i=1}^{N_0} \frac{g_i^2}{z - \omega_i^{(0)}} \right]^{-1} = \frac{\prod_{i=1}^{N_0} (z - \omega_i^{(0)})}{\sum_{i=1}^{N_0} g_i^2 \prod_{j=1}^{N_0 - 1} (z - z_j^{(0)})}. \quad (2.3)$$

[1] L. Fonda and G. C. Ghirardi, J. Math. Phys. 7, 906 (1966). We refer to this paper as I in what follows.

[2] For the treatment of the inversion problem in potential scattering, see the review article by R. G. Newton, J. Math. Phys. 1, 319 (1960), and V. de Alfaro and T. Regge, *Potential Scattering* (North-Holland Publishing Company, Amsterdam, 1965).

[3] For the definition of the renormalized coupling constants in potential scattering, see: L. Bertocchi, M. McMillan, E. Predazzi, and M. Tonin, Nuovo Cimento 31, 1352 (1964); L. Bertocchi, S. Fubini, and G. Furlan, *ibid.* 32, 745 (1964).

The form factor is then obtained in terms of the free parameters by

$$f^2(\omega) = -\frac{4\pi R(\omega)}{(\omega^2-\mu^2)^{\frac{1}{2}}} \times \exp\left[-\frac{P}{\pi}\int\frac{\delta(y)\,dy}{y-\omega}\right]\sin\delta(\omega), \quad (2.4)$$

where $R(\omega)$ is given in terms of the bound states energies ω_i and the free parameters

$$\sum_{i=1}^{N_0} g_i^2,\ z_j^{(0)}$$

by

$$R(z) = \frac{\left[\prod_{i=1}^{n}(z-\omega_i)\right](z-\mu)^{N_0-n}}{\sum_{i=1}^{N_0} g_i^2 \prod_{j=1}^{N_0-1}(z-z_j^{(0)})}. \quad (2.5)$$

In order to determine the $2N_0$ parameters

$$\omega_i^{(0)}, \sum_{i=1}^{N_0} g_i^2$$

and $z_j^{(0)}$ we have made use in I of the following relations:

(1) The condition $f^2(\mu) = 1$.

(2) The n eigenvalue equations for the bound states of the model.

(3) The equations for the resonances.

As extensively discussed in I, there are only $N_0 - n$ independent relations of this kind if the number of resonances is greater than this. More precisely when the phase goes from $-m\pi$ to $-(m-1)\pi$, also if it crosses the value $-m\pi + \frac{1}{2}\pi$ more than once, we get only one independent equation for these resonances.

(4) As discussed in I, the positions of the P_0 CDD poles falling above n are fixed by the zeroes of $\sin\delta(\omega)$ as it is necessary in order to have a positive $f^2(\omega)$.

Summarizing, we have $N_0 + 1 + P_0$ independent relations among the $2N_0$ free parameters. There follows that, in general, the assignment of the phase shift and of the bound state energies is not sufficient to determine completely the model.

We come now to the discussion of the use of the renormalized coupling constants. Let us consider the Low equation

$$(\phi_{\mathbf{k}_f}, H_I\Psi_{\mathbf{k}_i}^{(+)}) = (\phi_{\mathbf{k}_f}, H_I\phi_{\mathbf{k}_i}) + \int\frac{(\phi_{\mathbf{k}_f}, H_I\Psi_{\mathbf{k}}^{(+)})(\Psi_{\mathbf{k}}^{(+)}, H_I\phi_{\mathbf{k}_i})}{E_i - E + i\epsilon}\,d^3k + \sum_{b=1}^{n}\frac{(\phi_{\mathbf{k}_f}, H_I\Psi_b)(\Psi_b, H_I\phi_{\mathbf{k}_i})}{E_i - E_b} \quad (2.6)$$

$\phi_{\mathbf{k}}$ being the state describing free motion, and $\Psi_{\mathbf{k}}^{(+)}$ the scattering state. They are normalized as follows:

$$(\phi_{\mathbf{k}}, \phi_{\mathbf{k}'}) = (\Psi_{\mathbf{k}}, \Psi_{\mathbf{k}'}) = \delta^3(\mathbf{k}-\mathbf{k}').$$

Ψ_b is the normalized bound state eigenvector.

For our model the first term at the right-hand side of (2.6) vanishes. We recall now the expression for the T matrix

$$(\phi_{\mathbf{k}_f}, H_I\Psi_{\mathbf{k}_i}^{(+)}) = \frac{f(\omega_i)f(\omega_f)}{16\pi^3(\omega_i\omega_f)^{\frac{1}{2}}}[D(\omega_i + i0)]^{-1}, \quad (2.7)$$

where

$$D(z) = \frac{\prod_{i=1}^{N_0}(z-\omega_i^{(0)})}{\sum_{i=1}^{N_0} g_i^2 \prod_{j=1}^{N_0-1}(z-z_j^{(0)})} + \frac{1}{\pi}\int_\mu^\infty\frac{\varphi(y)}{y-z}\,dy$$

and

$$\varphi(\omega) = [(\omega^2-\mu^2)^{\frac{1}{2}}/4\pi]f^2(\omega).$$

Moreover one immediately gets

$$(\phi_{\mathbf{k}_f}, H_I\Psi_b) = \frac{1}{(2\pi)^{\frac{3}{2}}}\frac{f(\omega_f)}{(2\omega_f)^{\frac{1}{2}}}\sum_{i=1}^{N_0} g_i(\Psi_i^{(0)}, \Psi_b), \quad (2.8)$$

where

$$|\Psi_i^{(0)}\rangle = \psi_i^\dagger\,|0\rangle$$

is the ith bare excited state, and $|0\rangle$ the vacuum state.

Substitution of (2.7) and (2.8) into (2.6) yields

$$\frac{1}{D(\omega+i0)} = \frac{1}{4\pi^2}\int_\mu^\infty\frac{k\,d\omega}{\omega_i-\omega+i\epsilon}\cdot\frac{f^2(\omega)}{|D(\omega+i0)|^2} + \sum_{b=1}^{n}\frac{\left|\sum_{i=1}^{N_0} g_i(\Psi_i^{(0)}, \Psi_b)\right|^2}{\omega_i-\omega_b}. \quad (2.9)$$

The residues of the function $[D(z)]^{-1}$ at the position of the bound states are the squares of the renormalized coupling constants g_b^R of the model

$$\left|\sum_{i=1}^{N_0} g_i(\Psi_i^{(0)}, \Psi_b)\right|^2 \equiv \lim_{z\to\omega_b}(z-\omega_b)\frac{1}{D(z)} = (g_b^R)^2. \quad (2.10)$$

We recall that $D(z)$, in our case, is known in terms of the parameters of the model, the bound state energies and the phase shift:

$$D(z) = \frac{\left[\prod_{i=1}^{n}(z-\omega_i)\right](z-\mu)^{N_0-n}}{\sum_{i=1}^{N_0} g_i^2 \prod_{j=1}^{N_0-1}(z-z_j^{(0)})} \times \exp\left[-\frac{1}{\pi}\int\frac{\delta(y)}{y-z}\,dy\right]. \quad (2.11)$$

The quantities g_b^R are measurable quantities since they give the strength of the coupling of a ϑ particle with the bth bound state of the system.

Substitution of (2.11) into (2.10), therefore, yields n new relations for the parameters of the model. These relations are all independent since the Low equation, which seems to connect the g_b^R among themselves, is identically satisfied for any choice of the free parameters. If we consider, together with the previously obtained relations among the parameters of the model, the n so obtained relations, we have a total of $N_0 + 1 + P_0 + n$ relations among the $2N_0$ parameters of the model. Let us compare these two numbers.

In case B' we have from Eq. (2.2b) $N_0 = n + P_0 + 1$. Therefore we have in this case exactly $2N_0$ independent relations among the $2N_0$ parameters of the model.

In cases A and B'', $N_0 = n + P_0$ and we get a total of $2N_0 + 1$ relations among the $2N_0$ parameters. This means that, in general, the knowledge of only $n - 1$ of the renormalized coupling constants allows the complete determination of the model. The value of the last renormalized coupling constant is then fixed automatically by the model, and only if it meets the experimentally observed value, the physical situation can be described in the framework of this model. This is, for example, the situation which arises in the Lee model when there is a normalizable state. As shown in I, the parameters of the Lee model are in any case completely determined by the assignment of the phase shift and of the energy of the eventual bound state. Much the same holds also for the multichannel separable model considered in Sec. 5 of I.

3. POTENTIAL SCATTERING

In potential scattering it is known that,[2] for potentials whose first and second absolute moments are convergent, from the knowledge of the phase shift δ_l over the full energy range and of the energies E_b of the n bound states, one obtains an n-parameter family of potentials all of which lead to the same δ_l and E_b. In particular, this freedom is due to the fact that the normalization constants of the bound states

$$N_b = \varphi_l(-ik_b, r)/\psi_l^{(b)}(r), \quad (k_b > 0) \tag{3.1}$$

are not determined from the knowledge of δ_l and E_b. Here $\psi_l^{(b)}(r)$ is the normalized wavefunction [i.e., $\int_0^\infty dr(\psi_l^{(b)}(r))^2 = 1$] of the considered l-wave bound state, and $\varphi_l(-ik_b, r)$ is the regular solution of the l-wave radial Schrödinger equation at the energy of the bound state, determined by the boundary condition:

$$\lim_{r\to 0} (2l+1)!!\, r^{-l-1}\varphi_l(-ik_b, r) = 1. \tag{3.2}$$

Also in this case, however, the knowledge of the renormalized coupling constants yields uniquely all the normalization constants N_b. We see this for the case in which the Jost function, $F_l(k)$, which is determined uniquely by δ_l and E_b, is analytic in the whole complex plane, but the interval $(i|a|, i\infty)$ and a pole in the origin. This is, for example, the case of superposition of Yukawa potentials with range $2|a|$.

Defining the renormalized coupling constants through the Low equation (2.6), we write

$$Y_{lm}(\hat{\mathbf{k}})g_b^R = -\frac{\mu(2\pi)^{\frac{1}{2}}}{\hbar^2} \lim_{k\to ik_b} (\phi_{\mathbf{k}}, H_I\Psi_b). \tag{3.3}$$

It is straightforward to see that the so-defined g_b^R appears as coefficient of the asymptotic bound state wavefunction:

$$\psi_l^{(b)}(r) \underset{r\to\infty}{\sim} g_b^R e^{-k_b r}. \tag{3.4}$$

On the other hand, the asymptotic form of $\varphi_l(-ik_b, r)$ can immediately be derived from the equation of definition of the Jost function

$$\varphi_l(k, r) = F_l(k)f_l(-k, r) + F_l(-k)f_l(k, r). \tag{3.5}$$

We analytically continue this equation up to the point $k = -ik_b$. Since $F_l(-ik_b) = 0$ and $f_l(-ik_b, r) \underset{r\to\infty}{\sim} i^l e^{-k_b r}$, we get

$$\varphi_l(-ik_b, r) \underset{r\to\infty}{\sim} F_l(ik_b)i^l e^{-k_b r}, \tag{3.6}$$

and therefore,

$$N_b = i^l F_l(ik_b)(g_b^R)^{-1}. \tag{3.7}$$

We note that ik_b may even lie on the interval $(i|a|, i\infty)$ since, for $F_l(-ik_b) = 0$, $F_l(k)$ is not discontinuous on crossing the imaginary axis at the point ik_b.

Equation (3.7) expresses N_b as a function of δ_l, E_b and the renormalized coupling constant g_b^R. There follows that the potential $H_I(r)$ is uniquely determined from the knowledge of phase shift, binding energies, and renormalized coupling constants.

ACKNOWLEDGMENTS

This work was supported in part by the Istituto Nazionale di Fisica Nucleare.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Remarks on Relativistic Statistical Mechanics. I

RÉMI HAKIM
Laboratoire de Physique Théorique et Hautes Energies, Faculté des Sciences, Orsay, France

(Received 22 September 1966)

The modifications introduced by the specific forms of relativistic dynamics of many-particle systems are shown to give rise to a different (with respect to the nonrelativistic case) manner to set the problems involved in a tentative construction of relativistic statistical mechanics. Although the difficult problems of relativistic dynamics are not solved, it is possible to define relativistic generalizations of phase space, distribution functions, Gibbs ensembles, and average values. In particular, phase space is chosen for convenience and is no longer related (as is usually the case) to the "initial data," whose nature is yet unknown. As a consequence, only those observables which depend on the variables characterizing phase space give rise to easily computed average values. However, it is possible to enlarge at will the basic phase space and to define subsequent densities from which average values may be calculated. [Example: The calculation of average values of observables $A(\cdots x_i^\mu, u_i^\mu \cdots)$ needs only densities of the form $\mathcal{N}(\cdots x_i^\mu, u_i^\mu \cdots)$ and observables involving acceleration variables need the enlarging of phase space so as to include the latter. On this enlarged phase space, densities of the form $\mathcal{N}(\cdots x_i^\mu, u_i^\mu, \gamma_i^\mu \cdots)$ may be defined and are used to compute average values, etc.] The notion of equilibrium is discussed and suggestions for reaching the solution of this unsolved problem are made.

1. INTRODUCTION

In the past few years, a large number of papers on relativistic statistical mechanics and connected topics have been published. There are many reasons for such proliferation. One of the principal reasons may be found in the probable appearance (and this in the near future) of very high-temperature plasmas ($T \sim 0.5 \times 10^9$ deg) needed for thermonuclear fusion. Furthermore, relativistic plasmas *do exist* in stars (at least the electron component is relativistic as indicated by synchrotron radiation) and thus a relativistic statistical mechanics is necessary to treat them in a suitable way. However, even a nonrelativistic plasma needs a relativistic treatment when radiation phenomena are considered. Indeed, Schwinger[1,2] has shown that, in problems involving radiation phenomena, the relativistic corrections are much more important than the quantal ones, and this in a large domain of frequencies (from radio wavelength to the far infrared). Another important case where relativistic effects should be taken into account is that of the degenerate electron gas[3] whose Fermi energy is of the order of mc^2. Such a case, encountered in astrophysics when dealing with white dwarfs, occurs even at zero temperature. However, it also requires a quantal treatment, while throughout this paper we limit ourselves to a *classical theory*.

However, the above "practical" reasons should, in general, not be considered too seriously. Indeed, the domain of validity of nonquantal relativistic statistical mechanics is not extremely large: from $\sim 10^9$ deg to 2×10^9 deg (the latter number referring to the apparition of electron pairs which demand a quantum theory). So, we believe that problems of classical plasmas and radiation phenomena could be perfectly treated only with *relativistic corrective* terms taking "large velocities" into account. In this way, it is possible to make use of the Darwin Hamiltonian[4] and this possibility has been exploited by Krizan and Havas.[5]

Fortunately, besides the pragmatical reasons given above, there exist theoretical possibilities which are, in our opinion, sufficient to justify relativistic statistical mechanics. More particularly, if we consider that the special theory of relativity is a part of the laws of Nature, then it appears to be necessary to generalize in a suitable way the totality of Newtonian physics and hence statistical mechanics. However, besides these philosophical demands, it should be remarked that the theory of relativity implies a number of qualitative features which do not exist in Newtonian physics. From these particular features one may expect new phenomena in the statistical domain. For instance, the equivalence between mass and energy or the noninstantaneous character of relativistic actions could perfectly well be at the origin of new phenomena. We see that this is actually the case (see Papers II and III).

The various papers devoted to relativistic statistical

[1] J. Schwinger, Phys. Rev. **75**, 1912 (1949).
[2] J. L. Delcroix, in *La théorie des gaz neutres et ionisés* (Dunod Cie, Paris, 1960).
[3] B. Jancovici, Nuovo Cimento **25**, 428 (1962).
[4] C. G. Darwin, Phil. Mag. **39**, 357 (1920).
[5] J. E. Krizan and P. Havas, Phys. Rev. **128**, 2916 (1962); E. J. Krizan, *ibid.* **140A**, 1155 (1965).

mechanics can be *grosso modo* divided into two classes. In the first class we find the theories which are not manifestly covariant, either because they do not actually satisfy the principle of relativity (and they are generally *approximations*) or because they need a special proof of their effective covariance. In the second class fall the theories whose covariance is obvious at each step of the derivation.

In the former class we can include the classical works by Jüttner[6] dealing with the equilibrium of a relativistic gas (classical, Bose or Fermi gas). In 1939 a first attempt to build relativistic statistical mechanics was made by Van Dantzig.[7] However, this attempt was inconclusive, presumably because at that time relativistic dynamics of many-particle systems was not sufficiently developed. The first results in relativistic kinetic theory[8] are due to Lichnérowicz and Marrot[9] (relativistic Boltzmann equation). However, it is only very recently (and especially because of the developments of plasma physics) that attempts to build relativistic statistical mechanics have been undertaken. These theories start with the relativistic expression of the energy, i.e.,

$$H_i = \{[\mathbf{p}_i - e_i\mathbf{A}(\mathbf{x}_i, t)]^2 + m_i^2\}^{\frac{1}{2}} + V(\mathbf{x}_i, t), \qquad i = 1 \cdots N$$

from which is derived the Hamiltonian

$$H = \sum_i H_i + \frac{1}{8\pi}\int (\mathcal{E}^2 + \mathcal{H}^2)\, d_3x,$$

where the second term on the right-hand side of the preceding equation refers to the field energy. Using this generalized Hamiltonian, it is possible to derive a relativistic Liouville equation for a pseudodensity involving *both* fields and particles. Next, the theory follows the classical developments and in particular diagrammatic methods are used. These theories have been studied extensively by Balescu, de Gottal, Hénin, Mangeney, and Prigogine.[10] However, besides a number of conceptual difficulties (both mathematical and physical) the transformation properties of these theories are not yet completely clear and we no longer consider them in this paper.

We now limit ourselves to the second class, those papers written in an explicit covariant manner. Indeed, only the latter works contain mainly the *substantifique moelle* of the methods which are used throughout this series of papers. In this second class, there are only a few papers devoted to relativistic statistical mechanics.[8] In fact, they rather deal with relativistic kinetic theory[8] or with equilibrium. The methods which lead to such theories have essentially been set up by Bergmann, Chernikov, Ţiţeica, Synge, Tauber, and Weinberg.[11] They have been constructed so as to suit the relativistic hydrodynamics first established by Eisenhart, Synge, Lichnérowicz, Eckart, and Taub.[12] In such a way, various relativistic kinetic equations have been obtained: Vlasov equation,[13] Boltzmann equation,[14] Landau equation,[15] and Fokker–Planck equation.[16] Unfortunately, the greater part of these kinetic equations can hardly be justified[17] on the basis of relativistic statistical mechanics and they are only *ad hoc* semiphenomenological equations.

Let us now briefly glance through the various problems which arise in a relativistic generalization of statistical mechanics. A first problem deals with the *covariance of the theory*. We like to obtain a theory which would be explicitly covariant at each step of the derivation and of the calculations. In our opinion such a procedure does not merely arise from esthetic considerations. Indeed, in so doing we avoid an explicit proof (and even several proofs: one by manipulation) that the theory actually satisfies the

[6] F. Jüttner, Ann. Physik. **34**, 856 (1911a); **35**, 145 (1911b); Z. Physik **47**, 542 (1928).

[7] D. Van Dantzig, Nederl. Akad. Weetnsh. Proc. **42**, 601 (1993b).

[8] By "kinetic theory" we mean a theory based on a one-particle distribution function and a kinetic equation.

[9] A. Lichnérowicz and R. Marrot, Compt. Rend. **210**, 759 (1940).

[10] R. Balescu, in *1964 Cargèse Summer School* (Gordon & Breach, New York, to be published); Ph. de Gottal, thesis and Physica **32**, 126 (1966) (See Ref. p. 72); F. Hénin, Physica **29**, 1233 (1963); A. Mangeney, thesis, Paris, (1964) and Ann. Phys. (Paris) **10**, 191 (1965); I. Prigogine, in *Statistical Mechanics of Equilibrium and Non-Equilibrium*, J. Meixner, Ed. (North-Holland Publishing Company, Amsterdam, 1965).

[11] (a) P. G. Bergmann, Phys. Rev. **84**, 1026 (1951); (b) *Handbuch der Physik*, S. Flügge, Ed. (Springer-Verlag, Berlin, 1960), Vol. IV; (c) N. A. Chernikov, Dokl. Akad. Nauk SSSR **1**, 103 (1956); (d) S. Ţiţeica, Studii Şi Cercetari de Fizică **7**, 7 (1956); (e) J. L. Synge, *The Relativistic Gas* (North-Holland Publishing Company, Amsterdam, 1957); (f) G. E. Tauber and J. W. Weinberg, Phys. Rev. **122**, 1342 (1961).

[12] L. P. Eisenhart, Trans. Am. Math. Soc. **26**, 205 (1924); J. L. Synge, Trans. Roy. Soc. Canada **28**, 127 (1934); Proc. London Math. Soc. **43**, 376 (1937); A. Lichnérowicz, Compt. Rend. **211**, 117 (1940); **219**, 270 (1944); Ann. Sci. Ecole Normale Sup. **58**, 285 (1941); C. Eckart, Phys. Rev. **58**, 919 (1940); A. H. Taub, *ibid.* **74**, 328 (1948).

[13] P. C. Clemmow and A. J. Wilson, Proc. Cambridge Phil. Soc. **53**, 222 (1957); B. Kurşunoğlu, Nucl. Fusion **1**, 213 (1961); Yu. L. Klimontovich, Zh. Eksperim. i Teor. Fiz. **37**, 535 (1959) [English transl.: Soviet Phys.—JETP. **10**, 524 (1960)]; K. Goto, Progr. Theoret. Phys. (Kyoto) **20**, 1 (1958); G. E. Tauber and J. W. Weinberg, Ref. 11(f); R. Hakim, Compt. Rend. **260**, 3861 (1965).

[14] N. A. Chernikov, Soviet Phys. Dokl. **2**, 248 (1957); **5**, 764, 786 (1960); **7**, 397, 414, 428 (1962); Acta Phys. Pol. **23**, 629 (1963); **26**, 1069 (1964); **27**, 465 (1964); Phys. Letters **5**, 115 (1963); W. Israel, J. Math. Phys. **4**, 1163 (1963); R. W. Lindquist, Ann. Phys. **37**, 487 (1966); G. E. Tauber and J. W. Weinberg, Ref. 11(f).

[15] S. T. Beliaev and G. I. Budker, Soviet Phys.—Dokl. **1**, 218 (1956); Yu. L. Klimontovich, Zh. Eksperim. i Teor. Fiz. **38**, 1212, (1960) [English transl.: Soviet Phys.—JETP **11**, 876 (1960)].

[16] Yu. L. Klimontovich, Ref. 15; R. Hakim, Orsay report Th/68 (1964).

[17] This is not entirely true. Indeed, Klimontovich[13,15] justifies his equations with the help of a very interesting method which will be used extensively in Paper II. However, his definitions and methods are neither completely correct nor well established. We shall give the correct proofs of these equations in Paper II.

relativity principle. Furthermore, it is in the essence of the special theory of relativity to work in the frame of Minkowski geometry without any need to introduce objects extraneous to this geometry and to the one of the system under study. For instance, no objects such as spacelike 3-planes, $t = \text{const}$, need be considered. All is contained in the geometries of space–time and of the system. A second problem deals with the *evolution of the system.* In statistical mechanics, time indeed plays a very particular role as, for instance, in problems involving time scales, stochastic processes, perturbations, etc. However, we see that *there is no canonical notion of evolution of a relativistic system of particles* so that we must find a scheme to treat such problems as the ones referred to above. The third kind of questions concerns relativistic dynamics from which arise the main problems. These problems have been discussed at length by Havas,[18] whose paper contains also a quasi-exhaustive list of references on relativistic statistical mechanics and connected topics [R. Hakim, Ann. Inst. H. Poincaré **6**, 225 (1967)]. It is commonly admitted that relativistic statistical mechanics should treat both fields and particles in the same way. However, this point of view neglects other possibilities, such as the action-at-a-distance formalism, and we see that it leads (as expected) to infinities very difficult to drop into a statistical framework. At this point, we want to emphasize strongly that we have not solved (nor tried to solve) the dynamical problems. We have only explicitly assumed a number of plausible hypotheses concerning them. We want to point out that these basic assumptions also exist (although implicitly) in previous works.[10]

This paper is devoted to a discussion of the basis of possible ways out and difficulties of the problems involved. We first start with a discussion of relativistic kinetic theory, of which we give a helpful alternative formalism (Sec. 2). Next, a brief (and necessary) summary of relativistic dynamics is given (Sec. 3). Section 4 deals with the statement of the main statistical notions to be put (in our opinion) at the beginning of relativistic statistical mechanics. In Sec. 5, densities on the phase space previously considered are defined, while Sec. 6 is devoted to average values. Finally, in order that the reader not be mystified, in Sec. 7 we discuss the unsolved but important problem of equilibrium.

Paper II will contain several hierarchies of equations for the densities defined in Sec. 5 of this paper (exactly, for more general densities). In particular, a possible treatment of nonquantal radiation problems is given and applications to simple kinetic equations are included. Paper III will be devoted to more detailed applications and, in particular, to relativistic hydrodynamics of plasmas.

We now list some of the notations and conventions used in this paper. Throughout this paper we take $c = 1$ (c is the speed of light). The Einstein summation convention is used for the Greek indices. The latter run from 0 to 3 while the Latin indices (when corresponding to tensorial indices) run from 1 to 3; otherwise they can number particles (from 1 to N). Capital Latin letters are used in order to number the components of eight-vectors *or* of $8N$ vectors or tensors. The Minkowski space–time is endowed with the metric $+ - - -$. Furthermore, in addition to the usual mathematical symbols $\forall$, $\exists$, $\varnothing$, $\subset$, etc..., we use the following notations:

$\otimes$, tensorial product;
$\oplus$, direct sum;
$\wedge$, exterior product;
$\prec$, order;
$\boldsymbol{\alpha}$, tensor whose components are $\alpha \cdots$;
δ, Dirac measure:
$$\delta = \delta(x_0) \otimes \delta(x_1) \otimes \delta(x_2) \otimes \delta(x_3);$$
R^+, positive real numbers;
δ_Σ, superficial uniform measure $+1$ on the surface Σ;
$\Delta^{\mu\nu}(n_\lambda)$, projector on the 3-space orthogonal to the timelike 4-vector n_λ,
$$\begin{cases} \Delta^{\mu\nu}(n_\lambda) = g^{\mu\nu} - n^\mu n^\nu, \\ n^\lambda n_\lambda = 1; \end{cases}$$
p^μ, designate indifferently $m_0 u^\mu$ or (when existing) the canonically conjugated momentum of x^μ. There should not be ambiguities either because of the context or of further specification;
$\theta(x)$, heaviside step function.

As usual, we denote by boldface letters the spatial components of 4-vectors. There should not be ambiguities with the notation used for tensors *per se*.

2. PRELIMINARY REMARKS

Before considering relativistic statistical mechanics, it seems worthwhile to discuss briefly relativistic kinetic theory.

Outline of Relativistic Kinetic Theory

The relativistic one-particle phase-space[19] (the so-called μ space) is an eight-dimensional space; its

[18] P. Havas, in *Statistical Mechanics of Equilibrium and Non Equilibrium*, J. Meixner, Ed. (North-Holland Publishing Company, Amsterdam, 1965).

[19] We should say "state space" rather than "phase space." This point is explained in the remarks below.

points have four space-time coordinates and four momentum coordinates.[20] The distribution function $\mathcal{N}(x_\mu, p_\mu)$ is defined in such a way that the expression

$$j^\mu(x_\mu) = \int d_4p \mathcal{N}(x_\nu, p_\nu) \cdot u^\mu 2m \cdot \theta(p^0) \cdot \delta(p^\mu p_\mu - m^2) \tag{2.1}$$

is the particle four-current. [In Eq. (2.1) we have $p^\mu = mu^\mu$ and, of course, $2m\theta(p^0) \cdot \delta(p^\mu p_\mu - m^2)\, d_4p = m(d_3p/p^0(\mathbf{p}))$.] From Eq. (2.1) it is easily seen that $\mathcal{N}(x_\nu, p_\nu)$ is normalized by

$$\int_\Sigma j^\mu(x_\nu)\, d\Sigma_{\mu'} = N, \quad \forall\Sigma, \tag{2.2}$$

where Σ is an arbitrary spacelike three-surface and N is the number of particles constituting the gas. [In Eq. (2.2) $d\Sigma_\mu$ is the differential form with vectorial values

$$d\Sigma_\mu = \frac{1}{3!} \epsilon_{\mu\nu\rho\lambda}\, dx^\nu \wedge dx^\rho \wedge dx^\lambda,$$

where

$$\epsilon_{\mu\nu\rho\lambda} = \begin{cases} +1 \text{ when } (\mu\nu\rho\lambda) \text{ is an even permutation of (0123),} \\ -1 \text{ when } (\mu\nu\rho\lambda) \text{ is an odd permutation of (0123),} \\ 0 \text{ otherwise.} \end{cases}$$

We denote by $d\Sigma$: $d\Sigma = n^\mu_\Sigma\, d\Sigma_\mu$, where n^μ_Σ is the normal unit to Σ.] Condition (2.2) iimplies the conservation relation $\partial_\mu j^\mu(x_\nu) = 0$. The expression

$$2m\theta(p^0) \cdot \delta(p^\mu p_\mu - m^2) \cdot \mathcal{N}(x_\nu, p_\nu) u^\mu\, d\Sigma_\mu\, d_4p$$

occurring implicitly in the normalization condition (2.2) can be interpreted as the number of world lines which will cut $d\Sigma$ centered at point x_ν with the energy-momentum four-vector p_ν (modulo d_4p). The reasons for defining $\mathcal{N}(x_\nu, p_\nu)$ through a current have been explained by Bergmann[11] (see also below).

The relativistic one-particle Liouville equation can be obtained as follows. Let us denote by η the eight-vector whose components are $(u^\mu, F^\mu/m)$, where F^μ is the external four-force acting upon the gas. The eight-vector η (whose components are denoted by η^A, $A = 1 \cdots 8$) has a velocity character in μ space; hence an eight-current in μ space can be defined by[21,22]

$$J^A(x_B) = \mathcal{N}(x_B) \cdot \eta^A(x_B), \tag{2.3}$$

which satisfies the conservation relation[23] (see Bergmann, Ref. 11a)

$$\nabla_A J^A(x_B) = \nabla_A\{\mathcal{N}(x_B) \cdot \eta^A(x_B)\} = 0 \tag{2.4}$$

(where ∇_A is the covariant derivative in an arbitary coordinate system in μ-space); when condition

$$\nabla_A \eta^A(x_B) = 0 \tag{2.5}$$

holds, the relativistic one-particle Liouville equation follows and it reads

$$\eta^A(x_B) \cdot \partial_A \mathcal{N}(x_B) \equiv (d/d\tau)\mathcal{N}(x_B) = 0, \tag{2.6}$$

where τ is the proper time. Using now the conventional coordinates (x^μ, p^μ), Eq. (2.6) becomes

$$u^\mu \partial_\mu \mathcal{N}(x_\nu, p_\nu) + F^\mu(x_\nu, u_\nu)(\partial/\partial p^\mu)\mathcal{N}(x_\nu, p_\nu) = 0. \tag{2.7}$$

Condition (2.5) is valid when (for instance) the equations of motion of the particle can be cast into a Hamiltonian form; this is the case when the external force is due to an electromagnetic field or a "mesic" (i.e., scalar) force field. In all that follows we always bear in mind these two important cases.

The most general form of a kinetic equation that $\mathcal{N}(x_\nu, p_\nu)$ should satisfy is obtained by adding an *ad hoc* phenomenological term to the right-hand side of Eq. (2.7)

$$u^\mu \partial_\mu \mathcal{N} + F^\mu \cdot (\partial/\partial p^\mu)\mathcal{N} = C(\mathcal{N}), \tag{2.8}$$

where F^μ is an external four-force. The right-hand side of Eq. (2.8) is a collision term, which is nothing but the variation of $\mathcal{N}$ per unit proper time which is due to collisions. This term must be such that p^μ and m are collisional invariants.[24] The left-hand side of Eq. (2.8) is the variation of $\mathcal{N}$ per unit proper time which is due to the streaming. In order that Eq. (2.8) actually be a kinetic equation, the collision term must verify a number of requirements.[25] In particular, the relativistic form of the Maxwell–Boltzmann equilibrium distribution function, i.e., the Jüttner[6]–Synge distribution

$$\mathcal{N}(x_\nu, p_\nu) = \frac{N\rho(x_\nu)}{4\pi m^2 K_2(m\xi)} \exp\left[-\xi^\mu \cdot p_\mu\right] \tag{2.9}$$

should be a solution of $C(\mathcal{N}) = 0$, when $F^\mu = 0$. [In Eq. (2.9) the notations are those used by Synge[11]; ξ^μ is the reciprocal temperature four-vector, K_2 is a Kelvin function of order 2. Instead of Synge's notation

[20] In fact, the μ space is a seven-dimensional space because of the constraint $p^\mu p_\mu = m^2$ (or any other one). However, it is preferable to work in a flat eight-dimensional space rather than in a curved seven-dimensional one. In the former case, the constraint $p^\mu p_\mu = m^2$ is taken into account by simply adding a δ factor as, e.g., in Eq. (2.1) below.

[21] x_B ($B = 1 \cdots 8$) designates the coordinates of a point in μ space. When the coordinates chosen are (x_ν, p_ν), then $x_B = x_\nu$ for $B = 1 \cdots 4$ and $x_B = p_\nu$ for $B = 5 \cdots 8$.

[22] η^A depends on x_B through its definition.

[23] With $\partial_A \equiv \partial/\partial x^A$

[24] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, Cambridge, England, 1960).

[25] See, e.g., D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory* (McGraw-Hill Book Company, New York, 1964), p. 85.

N_0 (which could be confusing) we have used $N\rho(x_\nu)$ for the invariant world density of the fluid.]

In the same way $C(\mathcal{N})$ must be such that there exists an H theorem. These requirements are satisfied by the various kinetic equations indicated above (except of course the relativistic Vlasov equation). The connection of the covariant notations used in Eq. (2.8) and the three-dimensional ones is straightforward and can be found in the literature.[26]

From the distribution functions obtained, for instance, from a kinetic equation of the form (2.8), one can compute the energy-momentum tensor of the relativistic fluid under study as

$$T^{\mu\nu}(x_\rho) = \int d_4p 2m\theta(p^0) \cdot \delta(p^\mu p_\mu - m^2) m u^\mu u^\nu \mathcal{N}(x_\nu, p_\nu). \tag{2.10}$$

In the absence of an external force field this tensor satisfies the conservation relation[27]

$$\partial_\mu T^{\mu\nu}(x_\rho) = 0, \tag{2.11}$$

which expresses the conservation of the momentum four-vector for the fluid. In the case of an external electromagnetic field $F^{\mu\nu}(x)$, this relation should be replaced by

$$\partial_\mu T^{\mu\nu}(x_\rho) = j_\alpha(x_\rho) F^{\alpha\nu}(x_\rho). \tag{2.12}$$

When the external force field is "mesic," Eqs. (2.11) or (2.12) should be replaced by

$$\partial_\mu T^{\mu\alpha}(x_\rho) = \lambda\rho(x_\rho) \cdot \partial^\alpha \phi(x_\rho), \tag{2.13}$$

where λ is the coupling constant of the scalar field ϕ and $\rho(x_\rho)$ the invariant world density of the fluid (more generally ρ is the source term of the scalar field). The explicit knowledge of $T^{\mu\nu}$ implies the equations of relativistic hydrodynamics. Moreover, the various relativistic fluids can be classified according to the form of $T^{\mu\nu}$.[28] The generalization to the case of an external gravitational field is straightforward and has already been given by Tauber and Weinberg[11f] and by Chernikov.[14]

An Alternative Approach to Relativistic Kinetic Theory

Here we want to give an equivalent treatment of relativistic kinetic theory so as to illustrate some methods used in a forthcoming section. This section is principally intended for pedagogical reasons, and outlines in a simple case what is done in N-particle problems.

A particle of proper mass m is described by its equations of motion assumed to be of the form[29]

$$\begin{aligned} m(du^\mu/d\tau) &= F^\mu(x_\rho, u_\rho), \\ dx^\mu/d\tau &= u^\mu. \end{aligned} \tag{2.14}$$

Let us now consider an "observer" and assume that he locates the position of the particle in μ space as being (x_0^μ, u_0^μ). If the Cauchy problem corresponding to Eq. (2.14) is well set, then there exists a unique trajectory in μ space

$$x^B(\tau) \equiv \begin{cases} x^\mu(\tau) = x^\mu(\tau; x_0^\mu, u_0^\mu), \\ u^\mu(\tau) = u^\mu(\tau; x_0^\mu, u_0^\mu), \end{cases} \quad \tau \gtrless 0 \quad \text{or} \quad \tau < 0 \tag{2.15}$$

such that $x^\mu(0) = x_0^\mu$ and $u^\mu(0) = u_0^\mu$. In μ space this trajectory is determined by the sequence of points $x^B(\tau)$ (τ varying), and can be represented by the density

$$R(x_\nu, u_\nu; \tau) = \delta[x_\nu - x_\nu(\tau; x_0^\mu, u_0^\mu)] \otimes \delta[u_\nu - u_\nu(\tau; x_0^\mu, u_0^\mu)] \tag{2.16}$$

normalized through the obvious condition

$$\iint_\mu R(x_\nu, u_\nu; \tau)\, d_4x\, d_4u = 1. \tag{2.17}$$

Let us now assume that the measures of the "observer" are not very accurate and hence that the initial data of the particles are random, their repartitions in μ space having the density $D_0(x_0^\mu, u_0^\mu)$ normalized[30] by

$$\int_\mu D_0(x_0^\mu, u_0^\mu)\, d_4x_0\, d_4u_0 = 1. \tag{2.18}$$

[Equivalently, let us consider an *ensemble* (in the Gibbs sense) of similar systems, i.e., with the same equations of motion as Eq. (14), and ensure that the initial measures are distributed according to $D_0(x_0^\mu, u_0^\mu) \cdots$.] It follows that, at a given proper time τ, the density in μ space is no longer $R(x_\nu, u_\nu; \tau)$ (i.e., we have a cloud of points rather than one point) but is rather defined as its average value.

$$R(x_\nu, u_\nu; \tau) \equiv R(x_\nu, u_\nu; x_{\nu 0}, u_{\nu 0}; \tau)$$

is now a random function because of the random character of $(x_{\nu 0}, u_{\nu 0})$. Thus the density in μ space is

[26] See, e.g., P. C. Clemmow and A. J. Wilson, B. Kurşunğolu, Yu. L. Klimontovich,[18] or Y. Aboniy, Cahiers de Phys. **18**, 460 (1964).

[27] In a recent paper, De Gottal and Prigogine [Physica **31**, 677 (1965)] have suggested that Eq. (11) would no longer be valid as a consequence of the occurrence of the interaction between particles and should be replaced by $\partial_\mu T^{\mu\nu} = S^\nu$, where S^ν is the contribution of the field. In fact, this equation seems to be clear as a consequence of the conservation of the energy-momentum tensor of the system particles + fields:

$$\partial_\mu \{T^{\mu\nu}_{\text{part}} + T^{\mu\nu}_{\text{fields}}\} = 0.$$

[28] A. Lichnérowicz, in *Les Théories relativistes de l'électromagnétisme et de la gravitation* (Masson et Cie., Paris, 1955).

[29] See the remarks below. In particular this form excludes radiating particles. We return to this question in Paper II.

[30] Of course, we implicitly assume that the constraint $u_0^\mu u_\mu^0 = 1$ is included in D_0.

defined by

$$D(x_\nu, u_\nu; \tau) = \langle R(x_\nu, u_\nu; \tau)\rangle = \int_\mu R(x_\nu, u_\nu; x_{\nu 0}, u_{\nu 0}; \tau) \times D_0(x_{\nu 0}, u_{\nu 0})\, d_4x_0\, d_4u_0. \quad (2.19)$$

As a consequence, $D(x_\nu, u_\nu; \tau)$ is normalized on the whole μ space and not, as $\mathcal{N}(x_\nu, u_\nu)$, on a submanifold[31]

$$\int_\mu D(x_\nu, u_\nu; \tau)\, d_4x\, d_4u = 1. \quad (2.20)$$

The cloud of points represented by $D(x_\nu, u_\nu; \tau)$ is nothing but the cloud obtained by transforming the initial cloud by the laws of motion. In other words,

$$D(x_\mu, u_\mu; \tau) = T_\tau D(x_\mu, u_\mu; 0), \quad (2.21)$$

where $\{T_\tau\}$ is a representation in the space of densities on μ space of the transformation group (or semigroup if we limit ourselves to $\tau \geq 0$) equivalent to the laws of motion.

Deriving directly $R(x_\nu, u_\nu; \tau)$ and taking the average value of the result obtained, it is easy to find a one-particle relativistic equation satisfied by $D(x_\nu, u_\nu; \tau)$

$$\frac{\partial}{\partial\tau} D(x_\nu, u_\nu; \tau) + u^\mu \partial_\mu D(x_\nu, u_\nu; \tau) + \frac{F^\mu}{m}(x_\nu, u_\nu)\frac{\partial}{\partial u^\mu} D(x_\nu, u_\nu; \tau) \equiv \frac{d}{d\tau} D(x_\nu, u_\nu; \tau) = 0. \quad (2.22)$$

Of course, $R(x_\nu, u_\nu; \tau)$ also satisfies Eq. (2.22), in the derivation of which we have implicitly assumed that[32]

$$(\partial/\partial u^\mu) F^\mu(x_\nu, u_\nu) = 0.$$

This assumption is verified in the case of electromagnetic forces.

From the formal solution of Eq. (2.22), one easily sees that

$$T_\tau = \exp\left\{-\tau\left[u^\mu\partial_\mu + \frac{F^\mu}{m}(x_\nu, u_\nu)\frac{\partial}{\partial u^\mu}\right]\right\}. \quad (2.23)$$

By adding to the right-hand side of Eq. (2.22) an *ad hoc* collision term several kinetic equations (for instance, Boltzmann, Landau, Fokker–Planck equations, etc.) may be obtained.

Connection Between the Two Formalisms

Now, we prove two lemmas which establish this connection.

[31] $\mathcal{N}$ or D could indifferently be normalized to 1 or N.
[32] This is nothing but Eq. (5).

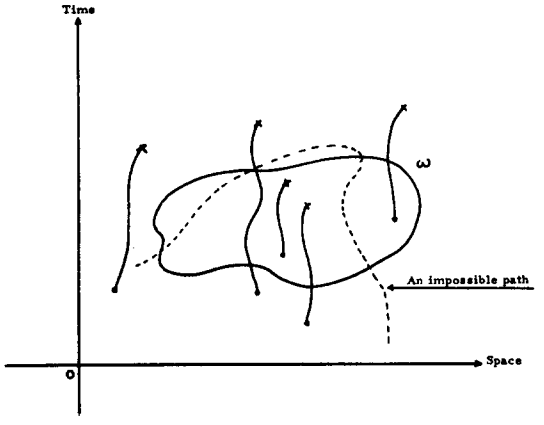


FIG. 1. Visualization of $\lim D(\tau) = 0$.

Lemma 1: The distribution function $D(x_\mu, u_\mu; \tau)$ verifies

$$\lim_{\tau\to\pm\infty} D(x_\mu, u_\mu; \tau) = 0 \quad (2.24)$$

in the sense of Lebesgue measure.

Proof. Let us consider the local "instantaneous" density

$$\rho(x_\nu, \tau) = \int d_4u\, D(x_\nu, u_\nu; \tau).$$

From this density, we can calculate the number of particles[33] within a *finite* Lebesgue-measurable four-volume $\omega \subset \mathcal{M}^4$, and whose proper time is τ:

$$n_\omega(\tau) = \int_\omega d_4x\, \rho(x_\nu, \tau).$$

Then, because of *causality* which implies that, having crossed ω, a particle cannot return within (see Fig. 1), we have

$$\lim_{\tau\to\pm\infty} n_\omega(\tau) = 0, \quad \forall\omega \subset \mathcal{M}^4,$$

from which Eq. (2.4) follows. Q.E.D.

It is of course *not* the value of D which tends towards zero. The convergence of D towards zero is intended to be a convergence in the sense of measures.

Lemma 2: The distribution functions $\mathcal{N}(x_\mu, u_\mu)$ and $D(x_\mu, u_\mu; \tau)$ are connected through the relation

$$\mathcal{N}(x_\mu, u_\mu) = \int_{-\infty}^{+\infty} D(x_\mu, u_\mu; \tau)\, d\tau. \quad (2.25)$$

Proof. Let us first remark that from the random density

$$R(x_\mu, u_\mu; \tau) \equiv R(x_B; \tau)$$

we can define a random current similar to the one used

[33] It is preferable to speak about the "number of particles within ω" rather than about the "percentage of particles within ω."

in relativistic classical electrodynamics (i.e., the Feynman current) by

$$J^A_{\text{stoch}}(x_B) = \int_{\text{along the traj.}} dx^A \cdot \delta[x_B - x_B(\tau, x_{B0})] = \int_{-\infty}^{+\infty} d\tau \delta[x_B - x_B(\tau, x_{B0})]\eta^A(\tau). \quad (2.26)$$

Using now the fact that (a) $\eta^A(\tau)$ depends on τ *only* through the intermediary of $x_B(\tau)$ and (b) δ functions occur in Eq. (2.26), we get

$$J^A_{\text{stoch}}(x_B) = \eta^A \cdot \int_{-\infty}^{+\infty} d\tau \delta[x_B - x_B(\tau, x_{B0})] = \eta^A \cdot \int_{-\infty}^{+\infty} R(x_B, \tau)\, d\tau. \quad (2.27)$$

Taking the average value of both sides of this last equality, we obtain

$$J^A(x_B) = \eta^A \cdot \int_{-\infty}^{+\infty} D(x_B, \tau)\, d\tau \equiv \eta^A \cdot \mathcal{N}(x_B), \quad (2.28)$$

from which Eq. (2.25) follows. Q.E.D.

It is now easy to show that the two given forms of the one-particle relativistic Liouville equations are consistent. Indeed, it is sufficient to integrate Eq. (2.22) over τ and to take Lemma 1 into account. Then we get Eq. (2.7).

Discussion and Remarks 1

Here we discuss the conventional approach to relativistic kinetic theory and indicate its main characteristic features. The latter is found again later when dealing with relativistic statistical mechanics.

(1) $\mathcal{N}(x_\nu, p_\nu)$ is actually not a density of probability. Indeed, one can easily see that the zeroth-order moment (in momentum space) has no direct physical meaning and does not correspond to a normalization of $\mathcal{N}$; it is only through the first moment (i.e., the current) that the one-particle distribution function is normalized. In connection with this point it should be emphasized [contrary to what is asserted by Goto, Ref. 13, Eq. (3-9)] that the expression

$$\int \mathcal{N}(x_\nu, p_\nu)\delta(p^\nu p_\nu - m^2)2m\theta(p^0)\, d_4p \quad (2.29)$$

is never the invariant world density of the fluid under consideration. This can be verified in the case of the distribution function of the simple gas at local equilibrium; we have

$$\int \frac{N\xi\rho(x_\nu)}{4\pi m^2 K_2(m\xi)} 2m\theta(p^0) \exp(-\xi^\mu p_\mu) \cdot \delta(p^\mu p_\mu - m^2)\, d_4p = N\rho(x_\nu)\frac{K_1(m\xi)}{K_2(m\xi)} \neq N\rho(x_\nu). \quad (2.30)$$

Of course, at the Newtonian limit, the right-hand side and the left-hand side of the $\neq$ in Eq. (2.30) are equal since $K_1 \sim K_2$ when $\xi \to \infty$ (which is equivalent to $c \to \infty$: see Synge[11e]). Had we normalized $\mathcal{N}$ to $\rho(x_\nu)$, the current would no longer have had the usual form of

$$j^\mu(x_\nu) = N\rho(x_\nu)(\xi^\mu/\xi) = N\rho(x_\nu)\bar{u}^\mu(x_\nu).$$

The fundamental reason for these features is that relativistic kinetic theory is a statistics of curves rather than a statistics of points. This has already been remarked by Bergmann[11a,b] in his "generalized statistical mechanics." To specify this question more precisely we have to discuss the question of phase space. To this end let us consider the normalization condition satisfied by $\mathcal{N}(x_\nu, p_\nu)$,

$$\int\int_\Sigma 2m\theta(p^0) \cdot \delta(p^\mu p_\mu - m^2) \cdot \mathcal{N}(x_\nu, p_\nu)u^\mu\, d\Sigma_\mu\, d_4p = N. \quad (2.31)$$

Equation (2.31) shows clearly that the distribution function is normalized on a six-dimensional manifold, which in fact is the effective phase space. This manifold has not an invariant meaning because of the arbitrariness of Σ. This effective phase space corresponds to a *possible* set of initial data as in Newtonian physics. In nonrelativistic statistical mechanics phase space is indeed defined as the set of initial data, and it has an invariant meaning with respect to the Galilei group so that the possible trajectories of the system lie therein. From a relativistic point of view the possible trajectories lie in the *state space*, which is just the "minimum" invariant space containing all possible effective phase space. In the sequel this state space is referred to as the phase space. It is now clear that a relativistic ensemble (in the Gibbs sense) is nothing but (a) the set of all possible trajectories of the system (this manifold has actually six dimensions as it should have), and (b) an invariant measure on this set.[34]

To sum up, the distribution function is not a density of probability but it allows one to obtain a density of probability over each possible effective phase space.

Though $\mathcal{N}(x_\nu, p_\nu)$ is not a density, a consistent probability theory can be given in order to develop this statistics of curves. Indeed, consider an open subset A of an arbitrary spacelike three-surface Σ and call C_A the set of trajectories intersecting A. C_A generates a σ-field in the space of trajectories and we can define a measure on this σ-field according to

$$\text{Mes}\,(C_A) = \int_{A \subset \Sigma} j^\mu(x_\nu)\, d\Sigma_\mu,$$

[34] Invariant with respect to Lorentz group.

and this measure is independent of Σ provided the current $j^\mu(x_\nu)$ be conservative. The same argument can also be applied in μ space.

(2) Another remark deals with average values. Following Bergmann[11a,b] the average value of a given function with tensorial values, namely $\psi(x_\nu, p_\nu) \equiv \psi(x_A)$, is defined as

$$\langle\psi(x_A)\rangle = \int_{\Xi^7 \subset \mu} \psi(x_A)\mathcal{N}(x_A)\eta^B(x_A)\, d\Xi_B \quad (2.32)$$

or with the more usual system of coordinates (x_ν, p_ν):

$$\langle\psi(x_\nu, p_\nu)\rangle = \int\!\!\int_\Sigma \psi(x_\nu, p_\nu)u^\mu\mathcal{N}(x_\nu, p_\nu) \times 2m\theta(p^0)\delta(p^\mu p_\mu - m^2)\, d\Sigma_\mu\, d_4p \quad (2.33)$$

and hence $\langle\psi\rangle$ is *biased* by the choice of Σ. Consequently, average values have, in general, not a well-determined variance.[11,35,36]

It should be emphasized that Eq. (2.32) or (2.33) represents the *flux* of the current of "property ψ" through Ξ^7 or Σ, this flux being constant and independent of Ξ^7 or Σ only when this current is conservative, which is a very particular case.

Now if we look at Eq. (2.33) more closely we remark that, although $\langle\psi\rangle$ depends on Σ, the current of property ψ does not. This would suggest that local average of ψ could be defined in an invariant way. Unfortunately, this is not the case and local averages are again biased by Σ. Indeed, according to Synge [Ref. 11, Eq. (241)] they are defined by

$$\overline{\psi}(x_\alpha) = J^\mu(\{\psi\}, x_\alpha) \cdot n^\Sigma_\mu(x_\alpha)/j^\mu(x_\alpha) \cdot n^\Sigma_\mu(x_\alpha), \quad (2.34)$$

where J^μ is the current of property ψ and $n^\mu_\Sigma(x^\alpha)$ is the normal unit to an arbitrary spacelike three-surface Σ. Equation (34) shows that even local averages are biased by Σ. This is of course in contrast with the definition by Goto[13] [Eq. (3-8)], itself due to the incorrect normalization of $\mathcal{N}$.

However, as noted by Synge[11e] a local average could be defined

$$\psi(x_\alpha) = J^\mu(\{\psi\}, x_\alpha) \cdot j_\mu(x_\alpha)/\{j^\mu(x_\alpha)j_\mu(x_\alpha)\}. \quad (2.35)$$

In fact, it could be argued that we do not need a general theory of relativistic average values. Indeed, what are needed are rather the currents of the various properties ψ, e.g., j^μ or $T^{\mu\nu}$. In other words, we mainly need hydrodynamical quantities, but this only shifts the problem. In particular, there remain ambiguities in the definition of the average four-velocity of a relativistic fluid.

It would certainly be interesting to have a relativistic statistical mechanics where densities are actually densities of probability and where mean values have a well-determined variance. Unfortunately this seems to be hardly possible.

(3) In the relativistic framework there is no canonical notion of evolution of a physical system. This is in fact a counterpart of the arbitrariness of the choice of "physical space," i.e., of a spacelike hypersurface. The point of view adopted is a global one in which a *sub specie aeternitatis* description of the system is given. The history of the system is written in Minkowski space–time once and for all, while "physical observations" are related to spacelike three-cuts. As we have already noted, this point of view leads to building a statistics of trajectories (which constitute the fundamental element representing the system) rather than a statistics of points (which would be as arbitrary as the chosen cut of the system). Of course, in order to avoid difficulties occurring because of the ambiguous[37] nature of the relativistic notion of simultaneity, we could choose a well-determined family of spacelike hypersurfaces indexed by a parameter (as, for instance, the family of three-planes $t = \text{const}$). Doing so, a conventional notion of evolution (and hence a conventional form for relativistic statistical mechanics) would be preserved. However, the theory would include an element extraneous to the geometry of space–time and of the system and could no longer be fully covariant.

This absence of a natural and invariant notion of evolution gives rise to conceptual difficulties[38] when one wants to consider relativistic stochastic processes.[39] We show in another paper[40] how it is possible to obtain such a relativistic theory.

This unfortunate circumstance also renders difficult a tentative description of disintegrating gases occurring, for instance, in high-energy astrophysics. A quite different reason may be found in the derivation of kinetic equations. Indeed, the derivation of kinetic equations generally implies a change of scale of

[35] F. Halbwachs, *Théorie relativiste des fluides à spin* (Gauthier-Villars, Paris, 1960).

[36] R. Hakim, J. Math. Phys. 6, 1482 (1965).

[37] The set of all spacelike hypersurfaces of $\mathcal{M}^4$ is not totally ordered (for a "natural" order, of course). The word "ambiguous" could be replaced by "relative."

[38] In particular, the use of the time coordinate as a parameter describing the evolution of the process, seems to be doubtful. Indeed, such a notion would only define a stochastic process in a given system of coordinates, while the definition would no longer be the same in another system. This can be shown easily by the mathematical definition of a stochastic process.

[39] Such a study is necessary if we want to give a probabilistic meaning to the various Fokker–Planck equations studied in the literature. It is also needed if we want to establish a relativistic theory of irreversible processes and in particular if we want to obtain a covariant generalization of Onsager relations.

[40] R. Hakim (to be published). See also Sec. 7 of the report Orsay Th/107 (1965).

time[41] (i.e., long time assumptions) and it would be most surprising that such a change depends on the method used to locate the temporal evolution of the system. Therefore, we need a temporal notion of evolution which would be *intrinsically attached* to the system. We shall come back to these points in the following and in Papers II and III.

Remarks and Discussion 2

(1) In the conventional approach to relativistic kinetic theories, we essentially start with a *congruence of curves* in μ space, whereas in the proper time approach we deal with points. The second approach, which is completely equivalent to the conventional one, is much closer to the classical considerations and allows the application of *standard methods.*

It is clear, however, that the physically interesting results such as the hydrodynamic quantities, will be obtained from $\mathcal{N}(x_A)$ and *not* from $D(x_A, \tau)$. So, the distribution $D(x_A, \tau)$ is nothing but a helpful intermediary in the calculations. Nevertheless, the fact that $D(x_A, \tau)$ is actually a density of probability allows the obtaining of results concerning $\mathcal{N}(x_A)$ which could hardly be found by reasoning directly on this last distribution.

(2) In deriving $D(x_A, \tau)$ from $R(x_A, \tau)$ we have assumed the existence of a distribution of the initial values $D_0(x_0^\nu, u_0^\nu)$. In fact, it is sufficient to assume the existence of an averaging operation $\langle\ \rangle$ such that

$$D(x_A, \tau) = \langle R(x_A, \tau) \rangle, \tag{2.36}$$

which need not be specified further. In the sequel we see that, in the absence of a complete solution of the problems raised by many-particle systems, more general densities can be defined without an explicit knowledge of the nature of the Cauchy data; for instance, it will be sufficient to *assume* the existence of such brackets $\langle\ \rangle$.

(3) Note that the initial data (x_0^μ, p_0^μ) can be considered as distributed either in the whole μ space or only on an arbitrary hypersurface. Mathematically the former distribution would correspond to a *discontinuous cut* of the congruence of trajectories, while the latter would correspond to a continuous cut. Physically both cases are possible because initial measures on the ensemble could perfectly be performed nonsimultaneously.

(4) Let us emphasize that in R or D, (τ, x_μ, u_μ) are independent variables. Indeed, to see this property we should bear in mind the definition of R:

$$R(x_\mu, u_\mu; \tau) = \delta[x_\mu - x_\mu(\tau, x_{\mu 0}, u_{\mu 0})] \otimes \delta[u_\mu - u_\mu(\tau, x_{\mu 0}, u_{\mu 0})].$$

However, in the same way as in conventional statistical mechanics, we could say in a fictitious sense that $u^\mu = dx^\mu/d\tau$ [instead of $dx^\mu(z)/dz = u^\mu(z)$].

(5) Let us note that it is always possible to set

$$D(x_\nu, u_\nu; \tau) = \hat{D}(x_\nu, u_\nu; \tau) \cdot \delta(u^\mu u_\mu - 1)$$

while $\hat{D}$ also verifies the Liouville equation (2.22). This is easily shown by introducing relation (2.36) in Eq. (2.22) and taking into account $u^\mu F_\mu = 0$. This is, of course, due to the fact that the constraint $u^\mu u_\mu = 1$ is also a constant of motion.

(6) In the case where we are only interested in the future[42] behavior of the system (i.e., $\tau > 0$), we have no longer a conservation relation of the form $\partial_\mu j^\mu(x_\nu) = 0$. Let us assume indeed that the particles of the system are for instance created at some points in space–time which are considered as initial data. Then the conservation of the number of particles read

$$\begin{cases} \dfrac{\partial}{\partial\tau} \displaystyle\int D(x_\mu, u_\mu; \tau)\, d_4u + \partial_\mu \int D(x_\mu, u_\mu; \tau) u^\mu\, d_4u = 0, \\ \tau \geq 0, \end{cases} \tag{2.37}$$

which, integrated over τ from zero to infinity, leads to

$$\partial_\mu j^\mu(x_\nu) = \int D_0(x_\nu, u_\nu)\, d_4u, \tag{2.38}$$

where we have taken Lemmas 1 and 2 into account. More generally, this point of view leads to source terms in the various equations considered. Note that it is only in such a case that we actually have a semigroup $\{T_\tau\}$ and *not* a group. This point of view is very helpful when we deal with a disintegrating gas consisting of excited subsystems (e.g., atoms).

Let us now examine the case where the initial data are distributed on a given spacelike hypersurface Σ. We have the "nonconservation" relation

$$\partial_\mu j^\mu(x_\nu) = \rho_0(x_\nu)\delta_\Sigma, \tag{2.39}$$

which at first sight could be most surprising. However, it is easy to see that relation (2.39) does not violate the conservation of the number of particles. Indeed let us consider (see Fig. 2) a spacelike hypersurface S *situated in the future* of Σ and thus such that $S \cap \Sigma = \varnothing$. As a result we have

$$\partial_\mu j^\mu(x_\nu) \big|_{x_\nu \in S \subset \text{future of } \Sigma} = 0 \tag{2.40}$$

since $\delta_\Sigma = 0$ for all points $x \notin \Sigma$. As a consequence the number of particles crossing any set $A \subset S$ is, as

[41] See, e.g., R. Zwanzig, in *1960 Boulder Summer School* (Interscience Publishers, Inc., New York, 1961), Vol. 3.

[42] "Future" with respect to the initial data.

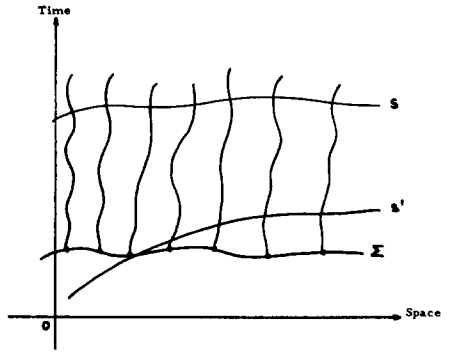


FIG. 2. Apparent lack of conservation of particles.

expected, equal to the number of particles crossing the similar set on Σ. Of course, this is no longer the case when $S \cap \Sigma \neq \varnothing$ (see Fig. 2).

(7) In writing Eq. (2.14) we have neglected the radiation reaction (in the case of a charged particle). If we take this last effect into account, instead of Eq. (2.14) we would write the Abraham–Lorentz–Dirac equations

$$m_0 \frac{du^\mu}{d\tau} = eF^{\mu\nu}_{\text{ext}} u_\nu + \tfrac{2}{3}e^2\left\{\frac{d^2u}{d\tau^2} - \left(\frac{du_\nu}{d\tau}\right)\left(\frac{du^\nu}{d\tau}\right)u^\mu\right\}$$

while the μ space would be a 12-dimensional space: $\mu = (\{x_\rho\}, \{u_\rho\}, \{\gamma_\rho\})$. In this generalized μ space the one-particle random density is of the form

$$R_1(x_\mu, u_\nu, \gamma_\nu; \tau) = \delta[\gamma_\mu - \gamma_\mu(\tau, x_0 u_0)] \otimes R(x_\nu, u_\nu; \tau),$$

where we have not indicated a γ_0^ν dependence for reasons which is given in a precise manner in Paper II. Hence the generalized R_1 is also averaged by means of the *same* $D_0(x_0^\nu, u_0^\nu)$ and yields a generalized one-particle distribution function $D^G(x^\nu, u^\nu, \gamma^\nu; \tau)$ which verifies an equation similar to Eq. (2.22). However, D^G no longer satisfies a Liouville theorem because of the form of the equations of motion. From this D^G we can obtain, after an integration over the proper time, a $\mathcal{N}^G$. The normalization of this $\mathcal{N}^G$ is given in Paper II where this formalism is studied in detail; in particular, it allows (for instance) the construction of an acceleration mechanism for the charged component of cosmic rays by random force fields which takes into account the energy loss due to radiation.

3. SUMMARY OF RELATIVISTIC DYNAMICS

In this section we briefly recall the starting points and the main formulas of dynamics of relativistic particles.

Before performing this program, we should bear in mind that, within the framework of relativistic *nonquantal* physics, *only three kinds of fields of force are actually known.* Only scalar (or pseudoscalar), vector, and tensor fields are known, which can be more or less considered as describing respectively classical nuclear forces, electromagnetic forces, and gravitation. As usual, relativistic forces may be divided into external forces and interaction ones. Among the former, there is no example of an external scalar field of force, though we could perhaps consider some collective effects of nuclear forces in the so-called neutron stars.[43] We cannot consider the gravitational forces as interactions because they are generally very weak (although this assertion be trivially false in the hypothetical case of collapsing stars[44] or even for superdense stars). However, they may have some important collective effects and hence generally act as "external forces." For the sake of simplicity we discard gravitational *collective* forces: they introduce only minor modifications of the given formalism. For instance, ∂_μ has to be replaced by ∇_μ which involves derivatives of the metric tensor $g^{\mu\nu}$ itself depending functionally on the one-particle reduced density (this last statement is equivalent to considering only collective effects of the gravitational field).

One-Particle Dynamics

The equations of motion of a particle imbedded both in an *external* electromagnetic field and an *external* "mesic" field are

$$\frac{d}{d\tau}\{m_0 + \lambda\phi(x_\alpha)\}u^\mu = \lambda\partial^\mu\phi(x_\alpha) + eF^{\mu\nu}(x_\alpha)u_\nu \tag{3.1}$$

or

$$\{m_0 + \lambda\phi(x_\alpha)\}\frac{d}{d\tau}u^\mu = \lambda\Delta^{\mu\nu}(u_\alpha)\partial_\nu\phi(x_\alpha) + eF^{\mu\nu}(x_\alpha)u_\nu \tag{3.2}$$

and they can be cast into a Hamiltonian form (see the review article by Schay[45]).

A *possible* Hamiltonian is[46]

$$H = [(p^\mu - eA^\mu)^2/2(m_0 + \lambda\phi)] - \tfrac{1}{2}\lambda\phi \tag{3.3}$$

and hence the canonical equations read

$$\frac{dx^\mu}{d\tau} = \frac{\partial H}{\partial p_\mu} = \frac{p^\mu - eA^\mu}{m_0 + \lambda\phi},$$
$$\frac{dp^\mu}{d\tau} = -\frac{\partial H}{\partial x_\mu} = +\lambda\partial^\mu\phi + eu^\rho\partial_\rho A^\mu. \tag{3.4}$$

[43] G. Szamosi, in *Varenna Summer School: High Energy Astrophysics, 1965* (to be published). See the reference quoted therein.

[44] K. Thorn, in *Varenna Summer School: High Energy Astrophysics, 1965* (to be published).

[45] G. Schay, Jr., Nuovo Cimento Suppl. **26**, 291 (1962). See also A. Peres and N. Rosen (quoted by G. Schay, Jr.).

[46] This Hamiltonian, which is merely formal and thus has no definite physical sense, is, of course, not unique.

In the above equations (3.1)–(3.4), ϕ is the "mesic" field (coupling constant λ) and $F^{\mu\nu}$ is the electromagnetic field ($F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, coupling constant e). We have recalled these well-known properties only to introduce canonically conjugated variables (x_μ, p_μ). In fact we rather use the variables (x_μ, u_μ). Let us also mention that the expression

$$H = \sum_{i=1}^{i=N} H_i \tag{3.5}$$

constitutes a *formal Hamiltonian* for a system of noninteracting particles [H_i is given by Eq. (3.3), where all symbols are indexed by i].

Many-Particle Dynamics

Dynamics of relativistic interacting particles are exposed in detail in three complementary books by Barut,[47] Rohrlich,[48] and Rzewuski[49] (see also Bergmann[11]).

Essentially two different points of view can equally be adopted. They are the action-at-a-distance and the field point of view (for a comparison and a discussion, see Havas[50]).

In the field point of view, we start with the equations of motion derived for instance from the Hamiltonian (3.5), to which equations for the fields are added; i.e.,

$$\Box A^\mu(x_\rho) = \sum_{i=1}^{i=N} 4\pi e_i \int_{-\infty}^{+\infty} \delta[x_\rho - x_{\rho i}(\tau_i)]\, dx_i^\mu, \tag{3.6}$$

$$\partial_\mu A^\mu(x_\rho) = 0, \tag{3.7}$$

for the electromagnetic field and (for instance)

$$\Box\phi(x_\rho) + M_0^2\phi(x_\rho) = \sum_{i=1}^{i=N} \lambda_i \int_{-\infty}^{+\infty} \delta[x_\rho - x_{\rho i}(\tau_i)]\, d\tau_i \tag{3.8}$$

for the "mesic" field.[51] Now, Eqs. (3.6), (3.7), and (3.8) contain both field and particle variables. However, these equations also involve the self-fields and therefore include divergences which have to be eliminated with the help of the so-called "renormalization of mass." As we see in Paper II, in a statistical treatment it is not easy to separate the effects due to mass renormalization from other terms.[52] It seems also hardly possible to split the fields into self-fields, radiation fields, interaction fields, etc.; we have only one entity, the total field. Note also that in the field point of view the mass m_0 appearing in the equations of motion for the particles is the bare mass.

In the action-at-a-distance formalism there is no field to support interactions and the equations of motion are given *a priori*. For instance, in the electromagnetic case one may start from the Fokker variational principle[53]

$$\delta S = \delta\Big\{ m_0 \sum_{i=1}^{i=N} \int (u_i^\mu u_\mu^i)^{\frac{1}{2}}\, d\tau + \tfrac{1}{2} \sum_{i,j} e^2 \iint u_\mu^i u_j^\mu G(x_i^\mu - x_j^\mu)\, d\tau_i\, d\tau_j \Big\} = 0, \tag{3.9}$$

where $G(x_\rho)$ is a kernel specifying the interaction[54] under consideration. As a result the equations of motion are very complicated integrodifferential ones and hence most difficult to solve. For the sake of comparison with the field point of view we adopt the ideas developed by Bergmann[11] according to which in Eq. (3.9) m_0 is again the bare mass.[55]

Both approaches lead to the Lorentz–Dirac equations and their generalizations to interacting particles:

$$m(du_i^\mu/d\tau_i) = \Gamma_i^\mu + e\overset{(i)}{F}{}^{\mu\nu}_{\text{int}} u_{\nu i}, \quad i = 1 \cdots N, \tag{3.10}$$

where Γ_i^μ is the radiation reaction term

$$\Gamma_i^\mu = \left\{ m\tau_0 \frac{d^2 u_i^\mu}{d\tau_i^2} + \frac{du_i^\rho}{d\tau_i} \cdot \frac{du_\rho^i}{d\tau_i}\, u_i^\mu \right\}, \tag{3.11}$$

m being the *finite renormalized mass* and τ_0 the so-called "noncausality" time[48]

$$\tau_0 = \tfrac{2}{3}e^2/m. \tag{3.12}$$

Note that Eq. (3.11) can also be rewritten as

$$\Gamma_i^\mu = m\tau_0 \Delta^{\mu\nu}(u_{\rho i})\, du_\nu^i/d\tau_i. \tag{3.13}$$

In Eq. (3.10) $\overset{(i)}{F}{}^{\mu\nu}_{\text{int}}$ is the field due to all other particles but the ith; it satisfies[56]:

$$\begin{cases} \partial_\nu \overset{(i)}{F}{}^{\mu\nu *}_{\text{int}} = 0, \\ \partial_\mu \overset{(i)}{F}{}^{\mu\nu}_{\text{int}} = \sum_{j \neq i} e \int_{-\infty}^{+\infty} \delta[x_i^\mu - x_j^\mu(\tau_j)]\, dx_j^\nu. \end{cases} \tag{3.14}$$

[47] A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (The Macmillan Company, New York, 1964).

[48] F. Rohrlich, *Classical Charged Particles* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1965).

[49] J. Rzewuski, *Field Theory* (P. W. N. Publ., Warsaw, 1964), Part I.

[50] P. Havas, Phys. Rev. **74**, 939 (1948).

[51] The equations of motion for fields and particles can also be obtained from a variational principle.

[52] E.g., A. Mangeney (Ref. 10) found a supplementary term in the study of bremsstrahlung. However, it has been shown that this term is not observable because it corresponds, in fact, to a finite term which should be included in the experimental mass. [A. Mangeney, (private communication).]

[53] A. D. Fokker, Z. Physik **58**, 386 (1929).

[54] In all the following and in both points of view we consider only retarded interactions. The questions of symmetry past–future, which are discussed in classical papers (see Ref. 50 and quoted papers therein) is irrelevant from statistical considerations. Furthermore, they could perfectly be included in this paper. Therefore they are dropped only for the sake of simplicity. Consequently we have $G(x_\mu) = \theta(x_0)\delta(x^\mu x_\mu)$ or $G(x_\mu) = \theta(x_0)F(x^\mu x_\mu)$ if we consider "extended" particles.

[55] We could avoid the procedure of renormalization.

[56] In the action-at-a-distance point of view $\overset{(i)}{F}{}^{\mu\nu}_{\text{int}}$ is such that the arbitrary homogeneous solution of Eq. (14) is chosen as being zero.

Equation (3.10) is obtained by splitting the self-field due to the ith particle into

$$(F^{\mu\nu}_{\text{self}})^{\text{ret}} = \tfrac{1}{2}\{(F^{\mu\nu}_{\text{self}})^{\text{ret}} + (F^{\mu\nu}_{\text{self}})^{\text{adv}}\} + \tfrac{1}{2}\{(F^{\mu\nu}_{\text{self}})^{\text{ret}} - (F^{\mu\nu}_{\text{self}})^{\text{adv}}\}, \quad (3.15)$$

the first term of the right-hand side of Eq. (3.15) leading to the renormalization of mass while the second term yields *rigorously* Γ^{μ}_{i}.

Had we used symmetrical actions, then Eq. (3.10) would have appeared as the result of the action of an "*absorber*" [57] (complete or incomplete[48]).

At this stage we want to emphasize that the Lorentz–Dirac equations are *not* approximate equations, valid to a given order in e^2, as is sometimes stated. We must also note that Eq. (3.10) are not the correct equations of motion since they lead to the well-known nonphysical "runaway solutions." To eliminate the latter we can impose an asymptotic condition (or several ones):

$$\lim_{\tau_i \to \pm\infty} \left(\frac{du^{\mu}_{i}}{d\tau_i}\right) = 0, \quad i = 1 \cdots N. \quad (3.16)$$

In Paper II we exploit Eq. (3.10), to which conditions (3.16) will be added.

At this step it is absolutely necessary to note that very important problems in either approach are not solved.[18] For instance, the existence and uniqueness of solutions of the equations of motion are not known, the exact nature of Cauchy data is not known, etc.

4. STATEMENT OF THE MAIN STATISTICAL PROBLEMS

Basic problems in relativistic statistical mechanics are of several sorts. Firstly, they are of dynamical order: What are initial data (and hence phase space)? To what extent does the non-Hamiltonian character of relativistic dynamics permit the transposition of Newtonian statistical mechanics? Secondly, they are of statistical order: What may be called a "relativistic Gibbs ensemble"? How does one treat in a covariant way the random character of the possibly existing fields? Thirdly they are also relevant to measure theory (in the so-called "operational" sense): given initial data, how is it possible to obtain them experimentally (i.e., with a *gedanken* experiment)? In this section we set these problems and try to give some insight into their possible (and future?) solution.

Initial Data—Observations and Measures

In Newtonian statistical mechanics probabilities are introduced with initial data of the subjacent dynamical problem. These probabilities are introduced so as to take into account the "nonaccuracy of the measurements" of an observer at an instant considered as being the origin of times. In the preceding sentence are two key expressions which we should investigate in order to see whether they could also be used with the same meaning in the relativistic framework: "measurement at a given time" and "observer." In Newtonian physics an "observer" is, at least in principle, able to perform instantaneous measurements and hence to have a global knowledge of the whole system at a given time. Is this situation prevailing in special relativity? It is one of the aims of this section to answer this question.

Let us now consider an observer[58] (i.e., a "punctual physicist" moving along a *timelike curve* in space–time and hence using "punctual physical apparatus"!) and a many-particle system on which this observer wants to undertake miscellaneous experiments. Naturally this observer knows the nature of the particles of the system, the laws of motion of these particles and possibly of the fields. Before answering the question of knowing how he proceeds, let us first remark that the observer is *not necessarily* a Galilean observer (i.e., his trajectory in Minkowski space–time is not necessarily a timelike straight line). Moreover, let us assume, as is conventionally done, that this observer receives information only through electromagnetic signals (or more generally through signals propagating in space–time with the velocity of light). Of course, this is not an essential hypothesis and we see below what kind of modifications would be implied by signals with a velocity less than that of light.

At a given instant of his clock taken as the origin of times, this observer receives the signal emitted—or reflected—(see Fig. 3) by the different particles of the system and locates them in space–time. It is indeed possible to locate space–time events only with time measurements and using signals traveling with the speed of light.[59,60] The observer may receive information on the state of the fields[61] on the backward null cone $\Gamma^{-}(0)$. Such a process is referred to as an "observation." Although it seems hardly possible to imagine a thought experiment which gives an observation of the fields, we assume, however, that this is

[57] J. W. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945); **21**, 425 (1949).

[58] The notion of "observer" is not so clear as one generally believes. For some people an observer is simply a Lorentzian system of coordinates. For others, it reduces to a family of spacelike 3-planes, etc.

[59] J. L. Synge, *Relativity: The Special Theory* (North-Holland Publishing Company, Amsterdam, 1958), Chap. 1.

[60] Figure 5 shows that the observation of the event M has a finite duration: in 0′ the observer emits signals which are reflected by M towards 0. See Ref. 59 for a detailed analysis.

[61] See the remark below.

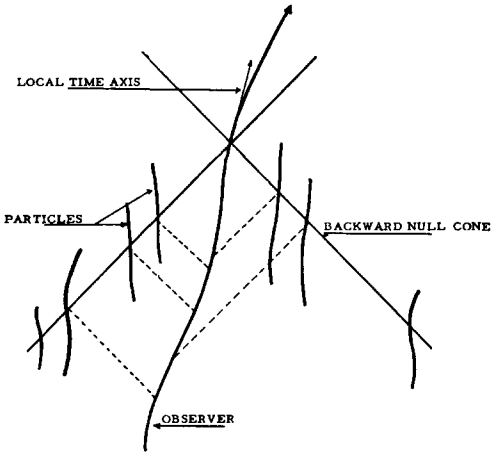


FIG. 3. Location of particles in space–time.

actually possible. If we did not make such an assumption, it would be impossible to describe completely relativistic statistical systems. The process of observation is by essence nonlocal. On the contrary, the measures affected by the observer are local processes relative to the event of space–time where the observer experiments. For instance, the field at point O (see Fig. 5) can be *measured* (e.g., with a test particle, whose interactions with the field are sufficiently weak in order that they have practically no influence on the system under consideration).

At this point, we may note that neither observations nor measures can be performed instantaneously.[60,62] In fact, the processes of observations and measures start slightly before the event O taken as the origin, say O' (see Fig. 5). The point O' can even be rejected towards the far past when the system is infinite. At any rate, the final results of observations or measures give initial data lying on a null cone. This shows that, given an observer, there exists a canonical slicing of space–time by a family of backward cones whose origins last on the trajectory of this observer (see Fig. 4). Therefore, the physical space at a given instant (proper time) of a definite observer is *never* a spacelike three-plane or even a spacelike three-surface but rather a backward null cone. As a consequence, an event M may be located by (see Fig. 5):

(a) the direction of the 4-vector OM;

(b) the proper-time interval between O and O'.

Unfortunately, these intrinsic notions (intrinsic modulo the observer who constitutes the only data extraneous to the geometry of the system *and* of Minkowski space–time—however, in an *operational* point of view, it is necessary to introduce an observer) are of poor

[62] E.g., a measure of the field intensity involves a test particle and the observation of its motion during a finite (although possibly small) interval of time.

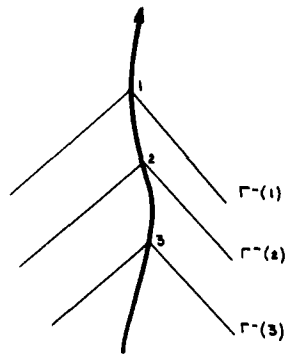


FIG. 4. A canonical "slicing" of space–time.

interest in the important case of electromagnetic interactions. Indeed, in that case the null cone is a *characteristic surface* of the Maxwell equations. Consequently, the data of the electromagnetic field on this surface are not sufficient to determine the future field completely (and uniquely). This is a well-known mathematical property: the Cauchy problem is not *well set* when the initial data lie on a characteristic surface. Physically, this could mean that such observations are incomplete by nature. We are thus led to give up such an intrinsic slicing of space–time and should use an arbitrary hypersurface to locate the initial data of the statistical system under consideration.

Let us also note that in the preceding discussion we have implicitly assumed that the observer and the physical system are not imbedded in a refractive medium. If it were not the case, the situation would be worse. Indeed, instead of obtaining initial data on a null cone the observer would only get initial data on a timelike conoid since light would propagate with a lesser velocity (see Fig. 6). As a consequence, the observer *never* has a complete knowledge of the system (see Fig. 6). Furthermore, it is not sure at all that the Cauchy problem on a timelike hypersurface for the fields has a solution or even a meaning.

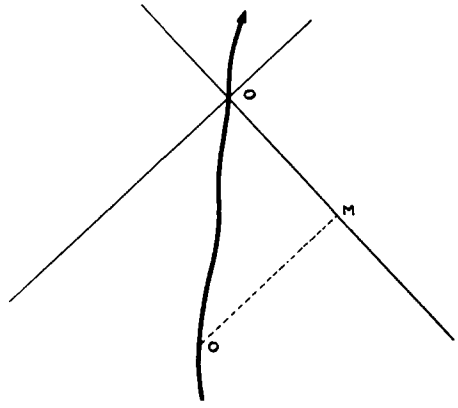


FIG. 5. Location of an event.

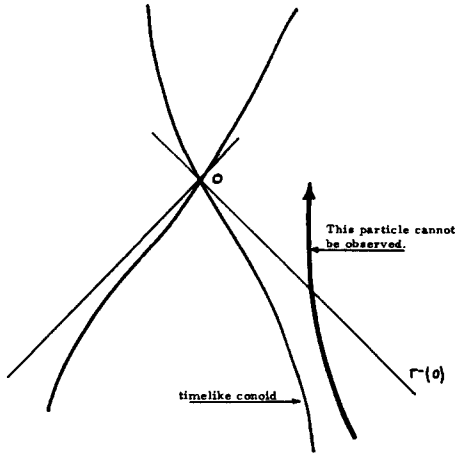


FIG. 6. Example of a nonobservable particle when using electromagnetic signals (in a dispersive medium).

At the beginning of this section we limited ourselves to the use of signals traveling with the speed of light. Had we used signals with a lesser velocity, the conventional notion of a physical three-space (see Fig. 7) would no longer have been preserved: initial data would have been found *in* (and also *on*) the null cone $\Gamma^-(0)$ and hence would last in a four-dimensional space.

Among the consequences of the preceding (brief) discussion, we see that in the most important case of electromagnetic interactions, it is necessary to introduce an arbitrary initial spacelike three-surface on which initial data are given. However, this scheme leads to the notion introduced by Cattaneo,[63] of an extended observer, i.e., an infinite family of local observers (as the one previously considered) whose trajectories are orthogonal[64] to the initial spacelike hypersurface. In other words, we could use a fluid of

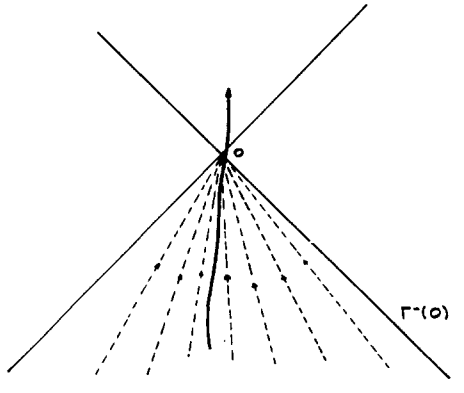


FIG. 7. Use of signals propagating with velocity less than that of light.

[63] C. Cattaneo, Nuovo Cimento **10**, 318 (1958).

[64] This condition is not absolutely necessary. It implies the regularity of the initial hypersurface Σ_0, i.e., $\partial\mu\Sigma_0 \neq 0$.

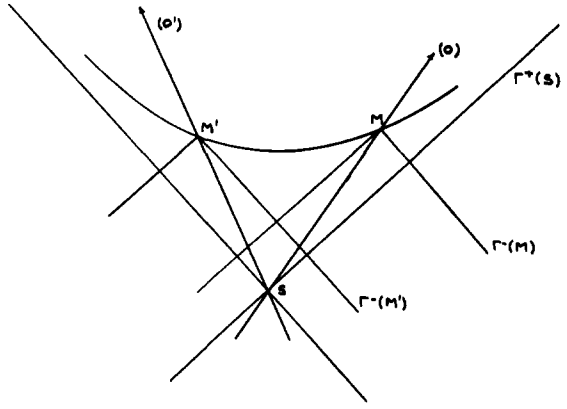


FIG. 8. Two Lorentzian observers.

reference.[65] Let us note that this notion of an extended observer renders irrelevant the distinction between observation and measure. Furthermore, from an operational point of view, it seems difficult to relate measures or observations of an extended observer to usual notions using clocks, rods, and signals.

In the cases where the null cones are not characteristic surfaces of the field equations, let us try to relate the observations or measures of two Galilean observers. We assume, without loss of generality, that they pass through the same event S (see Fig. 8) where they have synchronized their clocks. Let us assume now that M and M' are the origins of the proper times for the two observers O and O', respectively. First, we remark that M and M' are related by a proper orthochronous Lorentz transformation. Therefore the measure of a local quantity $\mathbf{F}\cdots(M)$ by observer O is related to the measure of the same local quantity $\mathbf{F}\cdots(M')$ by a Lorentz transformation.[66] On the contrary, the results of observations are global quantities which cannot easily be interrelated since they refer to different cuts of the subjacent physical system. Let us note that when Lorentz transformations are considered, it is intended that the point S is *the* fixed point.

Let us now draw some conclusions from this brief analysis.

(a) At any rate the Cauchy data *measured* by an observer are of the same nature as in Newtonian physics (i.e., lying on a spacelike 3-plane $t = \text{const}$ or even on a spacelike 3-surface; in a following paragraph we see that the mathematical Cauchy data could even be much more complicated).

(b) The initial data measured by an observer are not always sufficient to fully determine the ulterior behavior of the physical system.

[65] With some conditions on the congruence of "observers."

[66] Only when $\mathbf{F}\cdots(SM)$ is invariant in the strict sense, i.e., $\mathbf{F}\cdots = \mathbf{F}\cdots(SM)^2$.

(c) It is not at all clear whether the mathematical Cauchy data[67] can actually be measured.

(d) Points (a), (b), and (c) above suggest a choice, as fundamental random data to be introduced in relativistic statistical mechanics, of the trajectories of the particles and possibly the fields rather than "initial data." Indeed, a fully invariant viewpoint should be independent of "observers," systems of coordinates, etc. As a consequence, only the complete trajectories of the particles of the system are at our disposal to be chosen at random. Furthermore, such a point of view is in agreement with the one considered in the relativistic kinetic theory. However, one runs into the trouble that it cannot be *directly* related to conventional ideas according to which probabilities have to be introduced *ab initio*.[68] We come back to these questions in the following.

Statement of the Fundamental Statistical Problem

The preceding discussion has suggested that from an operational viewpoint "initial data" of the system are not always complete. If the basic statistical problem is to be well set, then what should be "randomized" is the "mathematical[69] Cauchy data." Henceforth we limit ourselves to these "mathematical Cauchy data," always bearing in mind that the problem of knowing *how* they could actually be measured remains open. In this section we mainly deal with the field point of view although we do not completely forget action-at-a-distance.

For the sake of discussion we consider only the example of a system of identical particles interacting through a scalar potential ϕ satisfying Eq. (3-8). Naturally the main results of the discussion remain valid for electromagnetic interactions. Setting now (see Sec. 5)

$$R_1(x_\mu, u_\mu; \tau) = \sum_{i=1}^{i=N} \delta[x_\mu - x_{\mu i}(\tau)] \otimes \delta[u_\mu - u_{\mu i}(\tau)], \tag{4.1}$$

Equation (3.8) can be rewritten as

$$\Box\phi + M_0^2\phi = \lambda \int_{-\infty}^{+\infty} d\tau \int d_4u R_1(x_\mu, u_\mu; \tau), \tag{4.2}$$

while the equations for the particles are again

$$m(du_i^\mu/d\tau_i) = \lambda\Delta^{\mu\nu}(u_{\alpha i})\partial_\nu\phi(x_{\alpha i}), \quad i = 1 \cdots N. \tag{4.3}$$

Now let us come back to the initial-value problem and try to set it in a form similar to the Newtonian one. To this end we *assume* that initial data are given *on* a spacelike three-surface.[70] If ϕ were an external field then the solutions of Eq. (4.3) would depend on $6N$ initial data (x_{i0}^μ, u_{i0}^μ). In the same way, if the motion of the particles were completely known [i.e., if the right-hand side of Eq. (4.2) were known], it seems *a priori* that we should have to solve a relatively simple Cauchy problem for Eq. (2). In fact it is not so. Indeed the source term of Eq. (4.2) does not vanish in the past of Σ_0 and therefore the Cauchy problem has no meaning![71] A way out of this difficulty is the following. Since the physicist is interested only in the future of Σ_0, one may assume that interactions are switched on on Σ_0. Therefore the source term actually vanishes in the past of Σ_0 and hence the Cauchy problem has a meaning. Such an assumption is therefore *necessary* if we want to preserve a form similar to the Newtonian one for the initial-value problem. Such a drastic assumption is implicitly contained in the work quoted in Ref. 10, where electromagnetic interactions are dealt with.

For the sake of a further discussion we assume the validity of the two following hypotheses which are themselves studied later.

Assumption 1: Interactions are switched on on Σ_0, the source of interactions vanishing "before" Σ_0.

Assumption 2: The initial data of the system consists (on Σ_0) of the $6N$ particular data (x_{i0}^μ, u_{i0}^μ) and of the usual field data $(\phi_0, \partial\phi_0)$ (where $\partial\phi_0 \equiv n_{\Sigma_0}^\mu \partial_\mu\phi_0$ is the normal derivative of the field on ϕ_0). Note that this last hypothesis is only a plausible one.

Under these assumptions the Cauchy problem for Eq. (4.2) is easily solved and we get

$$\begin{aligned}\Phi(x_\rho) = &\int_{\Sigma_0} \{\Delta(x_\rho - x'_\rho)\partial\Phi_0(x'_\rho) \\ &- \partial\Delta(x_\rho - x'_\rho)\Phi_0(x'_\rho)\}\, d\Sigma_0 \\ &+ \lambda \int_0^{+\infty}\!\!\int\!\!\int \Delta(x_\rho - x'_\rho) R_1(x'_\nu, u'_\nu; \tau')\, d_4x'\, d_4u'\, d\tau',\end{aligned} \tag{4.4}$$

[67] In this paragraph we have made some hypotheses on the nature of Cauchy data (i.e., data of the fields and of the particle variables on a spacelike 3-surface, for instance). They are discussed below.

[68] We must note that in Newtonian physics trajectories are also random although their statistical character is related to the random initial data.

[69] So far, nobody has been able to specify the nature of these data so that what follows consists of merely plausible assumptions. However, we see in the following that for the sake of statistics, the specific form of the Cauchy data is not absolutely necessary. What is needed is an averaging operation or more precisely the assumption of its existence. This assumption is in fact a very weak one, existing more or less implicitly in all statistical mechanics, and is independent of the precise nature of initial data. Of course, in Newtonian statistical mechanics, the explicit knowledge of the form of initial data renders explicit the averaging operation.

[70] For a sufficiently regular spacelike hypersurface, there always exists a system of coordinates such that its equation is $x^0 = 0$. Because of the fact that physics is independent of the chosen coordinate system (provided it preserves the normal hyperbolic character of the metric of space–time), the most general "initial physical space" is an arbitrary spacelike hypersurface and thus there is no reason why we should limit ourselves to spacelike 3-planes.

[71] This is a mere mathematical property. Mrs. Y. Bruhat-Choquet (private communication).

where $\Delta(x_\rho)$ is an appropriate Green function of Eq. (4.2).

At this stage we can adopt several viewpoints and they are, of course, somewhat related.

(a) *The initial field on Σ_0 and its normal derivative are given and not at random.* In such a case we have only to deal with the particle aspect of relativistic statistical mechanics since the field variables can be completely eliminated. This point of view leads to a relativistic Klimontovich hierarchy[15] (for the scalar interaction, of course) which is studied in Paper II. Equation (4.4) shows that the field $\phi(x_\rho)$ is a random field because it functionally depends on the random[72] density R_1. This field, although initially not correlated with the initial particular data, does not remain uncorrelated with the particle variables. It can also be remarked that the first term of the right-hand of Eq. (4.4) (which involves the dependence on ϕ_0 and $\partial\phi_0$) is a solution of the *homogeneous* Klein–Gordon Eq. (4.2) and hence will play the role of an *external force field* in the equations of motion of the particles.

(b) *The initial field on Σ_0 and its normal derivative are chosen at random.* In such a case the field ϕ is random, both due to the random distribution R_1 and to the initial field data. This point of view has been adopted by a large number of authors. It can be summarized by saying that it deals with a random Cauchy problem. Here also, there are several possibilities: either the initial data of the field and of the particles are correlated or they are uncorrelated. In the latter case everything occurs as if we were considering the preceding point of view except that *the system is embedded in a random external force field.* More precisely, let us assume that our ultimate aim is to obtain a kinetic equation; then we first derive (with the help of various approximations and hypotheses) a kinetic equation in a given external force field and next remember that this force field is random. Thus we have *random kinetic equations* which can be solved (at least formally) with different techniques.[73]

If, in order to be more "realistic," we want to get rid of assumption 1 and to know the behavior (statistical or not) of the system in the future of Σ_0, we must give up the conventional form (i.e., similar to the Newtonian one) of the initial-value problem. We may, as above, assume that the initial data of the particles are distributed on Σ_0 while the field (and its normal derivative) can be written as

$$\phi(x_\rho) = \phi^+(x_\rho) + \phi^-(x_\rho) + \phi_{\text{in}}(x_\rho),$$

where ϕ^+ is the contribution to ϕ of R_1 for $\tau \geq 0$, ϕ^- is the contribution due to the past ($\tau < 0$), and where ϕ_{in} is an arbitrary solution of the homogeneous equation (4.2) and represents more or less an initial radiation field.[74] More precisely,

$$\phi^+(x_\rho) = \Delta * \int_0^{+\infty} \int d\tau\, d_4u R_1(x_\rho, u_\rho; \tau),$$

$$\phi^-(x_\rho) = \Delta * \int_{-\infty}^{0} \int d\tau\, d_4u R_1(x_\rho, u_\rho; \tau).$$

In this scheme what is random? First of all the initial data of the particles are again chosen at random on Σ_0. Secondly, the past of the system can be chosen at random. For instance, if we assume that there is no initial field, then this "random past" is entirely included in the random function $R_1(x_\rho, u_\rho; \tau)$ (with $\tau < 0$). In order to be compatible with the initial particular data, this random function should satisfy some consistency conditions; e.g., R_1 should be such that

$$\lim_{\substack{\tau \to 0 \\ \tau < 0}} \langle R_1(x_\rho, u_\rho; \tau) \rangle = D_1(x_\rho, u_\rho)\big|_{x_\rho \in \Sigma_0}$$

be verified.[75] Note that this function is random not only because of the random character of the initial particular data but *also* because of the random character of the past of the system. ϕ_{in} may also be chosen at random with or without correlations with the Cauchy data of the particles. Whether this kind of initial-value problem is consistent with the one developed above or not is questionable, as remarked by Havas and Rohrlich. In our opinion, this last point of view corresponds more deeply to the principal features of relativistic interactions, and in particular to nonlocality. Furthermore, it appears to be more physical.

Finally the preceding discussion *suggests* the abandoning of the idea to set the initial-value problem in a form similar to the Newtonian one. For instance, if we want to remove assumption 1, then the Cauchy problem for the field can be set only at infinity.[47,48] In that case, we get

$$\phi(x_\rho) = \lambda \int_{-\infty}^{+\infty} \iiint \Delta(x_\rho - x'_\rho) R_1(x'_\rho, u'_\rho; \tau') \times d\tau'\, d_4x'\, d_4u' + \phi_{\text{in}}(x_\rho) \quad (4.5)$$

[72] "Random" because of the random character of the "initial data" of the particles.

[73] R. Kubo, J. Math. Phys. **4**, 174 (1963); G. Adomian, Rev. Mod. Phys. **35**, 185 (1963); See also the simple example treated in Sec. 7 of the preliminary report, Orsay Th/107.

[74] It would be a radiation field only when dealing with electromagnetic phenomena.

[75] For the moment there is no need to specify more precisely D_1, $\langle\,\rangle$ etc., (see Secs. 5 and 2). Furthermore the meaning of these quantities is intuitively clear.

which clearly shows that only two elements are at our disposal to be chosen at random: (a) the incident field ϕ_{in} and (b) R_1, or more precisely the entire trajectories of the particles of the system. Of course, these elements can be correlated or not. It seems, however, that it is more "physical" to consider ϕ_{in} and R_1 as being uncorrelated.[76] It may be remarked that this last point of view also includes action-at-a-distance (with $\phi_{\text{in}} = 0$). In our opinion only this last point of view has a fully invariant meaning and therefore we adopt it henceforth.[77] It has also the great advantage to avoid the explicit knowledge of the solution of the initial-value problem.

Remarks on the Statistical Treatment of Fields

The preceding paragraph indicates that in general, random fields are dealt with (or random variables in functional spaces when dealing with R_1). For instance, ϕ_0 and $\partial\phi_0$ may be considered as random variables in the Hilbert space $\mathfrak{L}^2(\mathcal{M}^4)$, e.g., when the field energy-momentum is finite. [Indeed, in that case we have

$$P^\mu(\Sigma_0) = \int_{\Sigma_0} T^{\mu\nu}(\phi_0, \partial\phi_0)\, d\Sigma_\nu < \infty,$$

where $T^{\mu\nu}$, the momentum-energy tensor, is a positive definite quadratic form in ϕ_0 and $\partial\phi_0$, which property implies that $(\phi_0, \partial\phi_0) \in \{\mathfrak{L}^2(\mathcal{M}^4)\}^{\times 2}$.] In the same way $R_1(x_\mu, u_\mu; \tau)$ is a random process[78] in the latticed Banach space of positive measures[79] in $\mathcal{M}^4$.

Unfortunately only very little is known on measures in functional spaces. For instance nobody has yet been able to obtain a nontrivial[80] Gaussian measure in Hilbert space.[81] A usual treatment consists in doing a statistics of field oscillators, at least in the electromagnetic case.[82] However, this procedure is mathematically not well defined. [For instance, this point of view leads naturally to a Gaussian measure in Hilbert space, which measure is known to be such that $\mu(\mathfrak{L}^2) = 0$.[81] However, it would be most interesting to find the reasons why such a point of view (which is obviously incorrect in the case of a Gaussian measure) leads to physical results. The problem is open. An element of answer would perhaps be the following: practically "reduced measures" involving one or two or a few oscillators variables are actually used to obtain physical results and never the complete "measure" (which has no mathematical meaning).] Another treatment, which is used in what follows, is inspired from turbulence theory. It consists in assuming that the complete statistical properties of a random variable in a functional space $\mathcal{K}$ are known when all the "moments" are given. In other words, the statistical properties of the random variable $f \in \mathcal{K}$ are *assumed* to be known when the infinite sequence

$$\langle f\rangle \in \mathcal{K}, \quad \langle f \otimes f\rangle \in \mathcal{K}^{\otimes 2}, \cdots, \langle f \otimes \cdots \otimes f\rangle \in \mathcal{K}^{\otimes k}, \cdots$$

is given. This treatment rests on the *analogy* with random variables[83] in R^N. This method has been used to a large extent in turbulence theory[84] and by Klimontovich in dealing with electromagnetic fields.[85] However, it is not sure at all that the above sequence determines a random element in $\mathcal{K}$. This is also an open problem.

So, we *assume* that the knowledge of the sequence

$$\langle R_1^{\otimes p} \otimes \phi_{\text{in}}^{\otimes q}\rangle; \quad p, q = 1, 2 \cdots$$

in the field point of view (or of the sequence

$$\langle R_1^{\otimes q}\rangle; \quad q = 1, 2 \cdots$$

when dealing with action-at-a-distance) is *sufficient* for the complete characterization of the statistics occurring in the problem.

Phase Space—Gibbs Ensemble

In this paragraph we mainly discuss the action-at-a-distance point of view. As a consequence, the results obtained remain valid in the field case but *only for each realization of the incident field.*

(1) Let us first examine the question of phase space. The same remarks as the one effected in Sec. 2 on μ space are valid *mutatis mutandis*. In particular, phase space will be a state space rather than the space of initial data as usual. It is a space which renders easy the description of the system; it is suggested by (a) the μ space used in relativistic kinetic theory and (b) the form of the equations of motion. Denoting by Γ

[76] The discussion on this question given above in points (a) and (b) can equally be repeated.

[77] Unless there is an explicit statement of the contrary.

[78] Note that it is not R_1 which is physically interesting but rather $\int R_1\, d\tau$ and that this last quantity does not constitute a random process in the above Banach space but a random variable.

[79] See, e.g., P. Courrege, *Théorie de la mesure* (Centre de Documentation Universitaire, Paris, 1964).

[80] I.e., such that $\mu(\mathfrak{L}^2) \neq 0$.

[81] See, e.g., M. Zerner, in *1965 Cargèse Summer School* (W. A. Benjamin, New York, to be published). (See also the references quoted there.)

[82] As, e.g., in the first work on these questions: W. E. Brittin, Phys. Rev. **106**, 843 (1957). See also the papers quoted in Ref. 10.

[83] It is well known that the moments of a random variable in R^N determine (modulo weak conditions almost always verified in practice) its density of probability (with respect to a given measure). However, a random variable in R^N does not necessarily possess moments of all order. Consequently if we limit ourselves to random variables possessing moments of all order, we would eliminate an important class of random variables. Hence, it is probable that the "moment method" in the case of random elements in functional spaces, is not completely general.

[84] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge University Press, Cambridge, England, 1953).

[85] Yu. L. Klimontovich, Zh. Eksperim. i Teor. Fiz. **34**, 173 (1958) [English transl.: Soviet Phys.—JETP **7**, 119 (1958)].

this phase space, we thus have

$$\Gamma = \mu^N, \tag{4.6}$$

where

$$\mu = \mathcal{M}^4 \times U^4 \tag{4.7}$$

or

$$= \mathcal{M}^4 \times V^+, \tag{4.8}$$

with

$$V^+ = \left\{ u_\mu : \begin{matrix} u^\mu u_\mu = +1 \\ u^0 > 0 \end{matrix} \right\} \tag{4.9}$$

according to whether the constraints (4.9) are included in the densities or not. Definition (7) is preferable[86] because it implies a flat μ space and hence a flat Γ space. [In the case where gravitation is also taken into account, it has been recognized by Chernikov that the general relativistic μ space is a fibre bundle

$$\mu = \bigcup_{x \in \gamma^4} P(x),$$

where γ^4 is the Riemannian space–time manifold and $P(x)$ is the fibre above x, i.e., the particle momentum space. Note also that the structure group of the $P(x)$'s is nothing but the orthochronous proper Lorentz group. Even in the case of special relativity, μ space (and also Γ space) may be considered as a fibre bundle. However, this structure plays no important role. In Newtonian physics, phase space is also a fibre bundle: the tangent bundle to the manifold "configuration space."]

(2) As for the *Gibbs ensembles*, they are defined exactly as in Sec. 3. They consist of (a) the manifold of solutions of the equations of motion and (b) a positive measure of total mass one over this manifold (or more precisely over a σ-field defined on this manifold).

At first sight this definition of a Gibbs ensemble seems to imply an *a priori* knowledge of the solutions of the equations of motion. Actually, it is not so. For instance, if we deal with the random density[87] $R_1(x_\mu, u_\mu; \tau; \omega)$, we should bear in mind that the physically interesting quantity is (see Sec. 5):

$$\langle R_1(x_\mu, u_\mu; \tau; \omega) \rangle \equiv \int R_1(x_\mu, u_\mu; \tau; \omega)\, d\mu(\omega), \tag{4.10}$$

where $\mu(\omega)$ is the measure over the manifold of solutions. Consequently we try to obtain equations verified for each realization of the process R_1 [or equivalently $x_i^\mu(\tau_i, \omega)$, $i = 1 \cdots N$] and next take the average value so as to get equations satisfied by the moments of R_1. It is therefore not necessary to specify precisely the averaging operation $\langle\ \rangle$ or, equivalently, the measure $\mu(\omega)$. It is sufficient to assert their existence. This is the very essence of the Klimontovich method,[88] which we use in Paper II.

Here again the comparison with Newtonian statistical mechanics is similar to that performed in the discussion of Sec. 2. The point of view developed here is not so far from the Newtonian one as one could believe. In both cases a Gibbs ensemble is defined as above. However, in the classical case $\mu(\omega)$ (and ω) can be explicitly given.

(3) In the Γ space of the system the entire history of the collection of the N-interacting particles is constituted by a N-dimensional manifold. This manifold consists of the intersection of the N cylinders generated by the N trajectories of the particles. Let M_N be this manifold. It can be represented in a parametric way with the help of N parameters, namely: $s_1 \cdots s_N$. This situation is by no means similar to the classical case where the system under consideration is represented in the nonrelativistic Γ space by a one-dimensional manifold. In the relativistic framework this circumstance is due to the absence of a universal time. Let us specify this point more precisely. The N trajectories of the particles are described, as we have already noted, by N parameters. Therefore we can write

$$x_1^0 = t_1, \cdots, x_N^0 = t_N,$$
$$\mathbf{x}_1 = \mathbf{x}_1(t_1), \cdots, \mathbf{x}_N = \mathbf{x}_N(t_N),$$

where we have chosen the N times of the N particles as parameters. However, within the framework of Newtonian physics, there exists a universal canonical parameter, which is the same for all the particles. This parameter is the usual time. This is equivalent to imposing the $(N - 1)$ relations:

$$t_1 = t_2 = \cdots = t_N = t \tag{11}$$

and hence the representative manifold is simply a curve.

As a consequence, this shows clearly that within the relativistic framework there is no canonical notion of the evolution of a system. [Imposing relations (4.11) as is done in Ref. 10 or any other ones amounts to choosing a curve on the manifold M_N describing the system. However, this choice is as arbitrary as the curve chosen.] Note also that relativistic statistical mechanics may be considered as a statistics of manifolds exactly as relativistic kinetic theory is a statistics of curves.

[86] We could use the 4-momentum $p^\mu = mu^\mu$ instead of the 4-velocity.

[87] Let $x_i^\mu(\tau_i, \omega)$ be a solution of the equation of motion of the ith particle. ω denotes an element of the so-called "sample space" of the theory of probabilities. For instance, ω may represent the Cauchy data. There is no need to specify ω further. In R_1 we have made the ω dependence explicit.

[88] Yu. L. Klimontovich, Zh. Eksperim. i Teor. Fiz. 33, 982 (1957) [English transl.: Soviet Phys.—JETP 6, 753 (1958)].

(4) Let us now briefly discuss the field point of view. For each realization of ϕ_{in} the preceding discussion remains valid. Here again a Gibbs ensemble will be (a) the manifold of solutions of equations of motion (including those for the fields) and (b) a positive measure of total mass unity over this manifold. As to phase space, since (in our approach) it is chosen for reasons of practical order, we may either adopt again definition (6) or enlarge it so as to include the fields:

$$\text{phase space} = \Gamma \times \{\text{space of the fields}\}. \quad (4.12)$$

This is just a matter of convenience. Now ω appearing in Eq. (4.10) also involves the random character of the incident field.

(5) Finally we must point out that in either approach our phase space has been chosen for reasons of convenience which are discussed in the following sections. It is not imposed by initial data. In particular, phase space could perfectly be enlarged[89] and this possibility is exploited in Paper II.

On the Non-Hamiltonian Character of Relativistic Dynamics

In our opinion,[90] it is absolutely unnecessary to have a Hamiltonian formalism in order to build statistical mechanics (Newtonian or relativistic). A Hamiltonian formalism only introduces *simplifications* and a Liouville theorem from which a Liouville equation is derived. To establish statistical mechanics, it is *sufficient* to start with (a) the equations of motion, (b) the conservation of the number of particles in an *ad hoc* phase space. These ideas are illustrated in Paper II. It seems that the main problem in the absence of a Hamiltonian formalism[91] concerns the definition of statistical equilibrium. However, it seems to us that this problem is not typically a relativistic one and also exists in classical statistical mechanics. We return to the definition of statistical equilibrium in a later paper.

5. DENSITIES AND RELATED QUESTIONS

In this section we define densities on the Γ space considered above and examine some properties (normalization, equations, etc.) which they verify. In Sec. 2 we have given two formalisms, first a formalism of geometrical character and next a proper time-dependent one. Here we rather consider the inverse order. Both formalisms are mutually illuminating.

Proper Time-Dependent Densities

The k-particle microscopic *random density* is defined as

$$R_k(x_1^\mu, u_1^\mu; \cdots; x_k^\mu u_k^\mu; \tau_1 \cdots \tau_k) = \sum_{i_1, \cdots, i_k} \prod_{j=1}^{j=k} \delta[x_j^\mu - x_{i_j}^\mu(\tau_j)] \otimes \delta[u_j^\mu - u_{i_j}^\mu(\tau_j)] \quad (5.1)$$

with

$$i_\alpha \neq i_\beta, \quad \forall \alpha \neq \beta \leq k, \quad i_\alpha = 1 \cdots N.$$

Obviously, R_k depends on the "initial data" ω through the factors $x_i^\mu(\tau)$ and $u_i^\mu(\tau)$ (which represent the trajectory of the ith particle in its own μ_i space). The "initial data" being chosen at random, R_k is a random function. R_k represents the random joint density of particles reaching $(x_i^\mu u_i^\mu)$ at "times" τ_i (with $i \leq k$).

Let us remark that definition (5.1) is valid in the relativistic case as well as in the classical one. In the latter, we obtain the multi-temporal distributions already considered[92] by Klimontovich (Greek indices then go from 1 to 3 while the τ_j's are to be considered as usual times). The density R_k is normalized through

$$\int_{\mu_1 \times \cdots \times \mu_k} R_k(x_1^\mu, u_1^\mu; \cdots; x_k^\mu, u_k^\mu; \tau_1 \cdots \tau_k) \times \prod_{i=1}^{i=k} d\mu_i = k!\, C_N^k. \quad (5.2)$$

Reduced distribution functions of order k are now easily defined as "average values":

$$k!\, C_N^k D_k(x_1^\mu; u_1^\mu; \cdots; x_k^\mu, u_k^\mu; \tau_1, \cdots, \tau_k) = \langle R_k(x_1^\mu, u_1^\mu; \cdots; x_k^\mu, u_k^\mu; \tau_1, \cdots, \tau_k)\rangle, \quad k \leq N, \quad (5.3)$$

where the brackets $\langle\ \rangle$ have been discussed in Sec. 4. The factor $k!\, C_N^k$ has been chosen so that the D_k's are normalized to unity. For instance, one has[93]

$$N!\, D_N(x^A; \tau_1 \cdots \tau_N) = \langle R_N(x^A; \tau_1 \cdots \tau_N)\rangle \quad (5.4)$$

[89] In nonrelativistic mechanics, phase space is generally considered as being the space of all possible initial data. Fortunately it is *also* the most convenient one to *describe* the system. However, we could enlarge it so as to include acceleration (or other) variables. As a consequence, densities including acceleration variables would be needed, but by virtue of the *form* of the equations of motion, this dependence would be trivial: it would be only through δ factors such as $\delta(\mathbf{F} - m\boldsymbol{\gamma})$. It is not to such a trivial possibility that we refer (see Sec. 5 and Paper II).

[90] This opinion is not universally accepted. For instance, P. Havas and R. Balescu (private discussions) do not completely agree with this point of view. Conversely, R. Kurth, *Axiomatics of Classical Statistical Mechanics* (Pergamon Press, Oxford, 1960), agrees with his opinion.

[91] By "Hamiltonian formalism" we mean that (a) the equations of motion have a Hamiltonian form and (b) there exist canonical variables such that the Hamiltonian has actually the meaning of the energy of the system. Indeed P. Havas [Nuovo Cimento Suppl. 5, 363 (1957)] has shown that the equations of motion of non-Hamiltonian systems (such as the one consisting of a particle submitted to a friction force $-\beta v$) may sometimes be cast into a Hamiltonian form. However, the Hamiltonian obtained has in general no physical meaning.

[92] Although not so explicitly.

[93] x^A stands for $(x_1^\mu, u^\mu, \cdots, x_N^\mu, u_N^\mu)$ or any other system of coordinates in Γ space. The slight ambiguity of notations (the same as in Sec. 2) is not confusing.

from which one may verify[94] that

$$D_k(x_1^\mu, u_1^\mu; \cdots; x_k^\mu, u_k^\mu; \tau_1 \cdots \tau_k) = \int_{8(N-k)\text{ times}} \cdots \int D_N(x^A; \tau_1 \cdots \tau_N)\, d\mu_{k+1} \cdots d\mu_N. \tag{5.5}$$

We also have

$$ND_1(x^\mu, u^\mu; \tau) = \langle R_1(x^\mu, u^\mu; \tau)\rangle = \left\langle \sum_{i=1}^{i=N} \delta[x^\mu - x_i^\mu(\tau)] \otimes \delta[u^\mu - u_i^\mu(\tau)] \right\rangle \tag{5.6}$$

as expected. In the same way one can obtain another important density

$$N(N-1)D_2(x_1^\mu, u_1^\mu; x_2^\mu, u_2^\mu; \tau_1, \tau_2) = \langle R_2(x_1^\mu, u_1^\mu; x_2^\mu, u_2^\mu; \tau_1, \tau_2)\rangle = \left\langle \sum_{i \neq j} \delta[x_1^\mu - x_i^\mu(\tau_1)] \otimes \delta[u_1^\mu - u_i^\mu(\tau_1)] \otimes \delta[x_2^\mu - x_j^\mu(\tau_2)] \otimes \delta[u_2^\mu - u_j^\mu(\tau_2)] \right\rangle. \tag{5.7}$$

More generally, D_k may be calculated either from R_k or from D_{k+p} $(p > 0,\ k + p \leq N)$.

Remarks

It is possible to define other kinds of distributions by considering mean values of random products of the form

$$R_l(x_1^\mu u_1^\mu \cdots x_l^\mu u_l^\mu; \tau_1 \cdots \tau_l) \otimes R_q(x_1^\mu u_1^\mu \cdots x_q^\mu u_q^\mu; \tau_1 \cdots \tau_q).$$

In this way we could generate the reduced distributions D_{q+l} *plus* "mixed" distributions such as the density of probability that we have a particle in the state $(x_1^\mu u_1^\mu)$ at "time" $\tau_1, \cdots$, while particle k is in the state $(x_{k_0}^\mu u_{k_1}^\mu)$ at "time" τ_{k_0} and undergoes transitions through states $(x_{k_1}^\mu u_{k_1}^\mu)$ at "time" $\tau_{k_1}, \cdots$, $(x_{k_j}^\mu u_{k_j}^\mu)$ at "time" τ_{k_j}, etc. The latter distributions are used in Paper II. Let us now give a simple and important example. To this end let us consider the average value $\langle R_1 \otimes R_1\rangle$. We have[95]

$$\langle R_1 \otimes R_1\rangle = \left\langle \sum_{i,j} \right\rangle = \left\langle \sum_{i \neq j} \right\rangle + \left\langle \sum_i \right\rangle. \tag{5.8}$$

The first term of the right-hand side of Eq. (5.8) is nothing but $N(N-1)$ times D_2 while the second term is

$$NP_{2_{\text{def}}} \equiv \left\langle \sum_i \right\rangle \equiv \left\langle \sum_i \delta[x_1^\mu - x_i^\mu(\tau_1)] \otimes \delta[u_1^\mu - u_i^\mu(\tau_1)] \otimes \delta[x_2^\mu - x_i^\mu(\tau_2)] \otimes \delta[u_2^\mu - u_i^\mu(\tau_2)] \right\rangle \tag{5.9}$$

and represents N times the probability that a given particle be in the state 1 and next the *same* particle be in the state 2.

(2) It must be pointed out that the various properties occurring in this paragraph are merely formal since the Dirac measure is not absolutely continuous with respect to Lebesgue measure and hence has no density. However, these formal definitions can easily be justified on the basis of correct mathematics.

(3) The "physical meaning" of these densities could be specified as follows. Let Δ be an elementary[96] Lebesgue–measurable subset of Γ. Then,

$$n_\Delta(\tau_1 \cdots \tau_N) = \int_\Delta D_N(\tau_1 \cdots \tau_N)\, d\Gamma$$

[where we have used the shorthand

$$D_N(x^A; \tau_1 \cdots \tau_N) = D_N(\tau_1 \cdots \tau_N)]$$

represents the probability that particle 1 be in $\delta_1 \subset \mu_1$, particle 2 be in $\delta_2 \subset \mu_2, \cdots$, particle N be in $\delta_N \subset \mu_N$, while their proper times are respectively $\tau_1, \cdots, \tau_N$.

Let us emphasize, however, that, as in relativistic kinetic theory, the proper time-dependent densities are not "physical." The "physical" densities are rather the proper time-independent ones studied in the following.

(4) Let us assume for a moment that the 6 N usual initial data $(x_{i0}^\mu, u_{i0}^\mu;\ i = 1 \cdots N;\ x_i^\mu \in \Sigma_0)$ are actually sufficient to determine (in an action-at-a-distance formalism) the complete behavior of the system.[97] In other words, by a given point x^A in Γ space passes only one manifold $M_N(x^A)$. Consequently, there exists a one-to-one correspondence between couples of points belonging to the same manifold M_N, when using the proper time parametrization, i.e.,

$$x_1^A(\tau_1 \cdots \tau_N) = T_{\tau_1 \cdots \tau_N} x_0^A;\ x_0^A, x_1^A \in M_N,$$

where $\{T_{\tau_1 \cdots \tau_N}\}$ is a group (or semigroup) equivalent to the laws of motion. This is completely similar to the one-particle case (see Sec. 2). Now, if we denote by

[94] To show this property, it is sufficient to pass through the intermediary of the random densities and to take the average value of the final result. Of course, we use the mathematical property

$$\langle \int \rangle = \int \langle\ \rangle$$

which is assumed to be valid.

[95] Let us indicate that the similar formula given by Klimontovich [Ref. 15, just after Eq. (5.6)] is incorrect. For comparison it is necessary to integrate Eq. (5.9) over τ_1 and τ_2.

[96] By "elementary," we mean: $\Delta = \delta_1 \times \cdots \times \delta_N$, where δ_i is a Lebesgue-measurable set in μ_i space.

[97] This seems to have been shown by J. Rzewuski *et al.* (see Ref. 49) in a nonphysical case, the case which consists in considering finite world lines. See also the remark by P. Havas (Ref. 18).

$D_N(x^A; 0 \cdots 0)$ the distribution of initial data, then (as usual) $D_N(x^A; \tau_1 \cdots \tau_N)$ can be defined by the following conditions:

(a) $D_N(x^A, \tau_1 \cdots \tau_N) = T_{\tau_1 \cdots \tau_N} D_N(x_0^A, 0 \cdots 0)$,

(b) $D_N(x^A, \tau_1 \cdots \tau_N) \geq 0, \quad \forall (\tau_1 \cdots \tau_N) \in R^{+N}$,

(c) $$\int_\Gamma D_N(x^A, \tau_1 \cdots \tau_N)\, d\Gamma = 1, \qquad \forall (\tau_1 \cdots \tau_N) \in R^{+N}.$$

One can show that, under the validity of the assumption effected in this paragraph, this definition is consistent with the one given above.

(5) Let us now consider whether there exists a relativistic Liouville equation for a system of N particles. To this end let us consider the case of a system of independent particles. In the same way as in Sec. 2, we obtain by derivation of R_k and after taking the average value,

$$\frac{\partial}{\partial \tau_i} D_k + u_i^\mu \partial_{\mu_i} D_k + \frac{F^\mu}{m_i}(x_i^\nu, u_i^\nu) \frac{\partial}{\partial u_i^\mu} D_k = 0, \qquad i = 1 \cdots k, \quad k = 1 \cdots N, \tag{5.10}$$

where F^μ is the external force field. In particular, this shows that D_N does not verify a Liouville equation but rather N one-particle Liouville equations. We find this property again by using the more sophisticated geometrical techniques given in the next paragraphs. The reason why this property occurs is the absence of a universal time. Indeed, in nonrelativistic physics, instead of $R_N(x_A, t_1 \cdots t_N)$ we have $R_N(x_A, t \cdots t)$, and thus the derivation with respect to now only one t connects all the one-particle Liouville equations between them.[98] Had we used the covariant notion of evolution defined by

$$\tau_1 = \tau_2 = \cdots = \tau_N = \tau,$$

then we should have found a nontrivial Liouville equation. More precisely we should have found an equation of the form

$$\frac{\partial}{\partial \tau} D_N(x_A; \tau \cdots \tau) + \sum_i \left\{ u_i^\mu \partial_{\mu i} + \frac{F_i^\mu}{m} \frac{\partial}{\partial u_i} \right\} \times D_N(x_A; \tau \cdots \tau) = 0.$$

However, in doing so we lose information on the system and the knowledge of the new D_N is not completely equivalent to the knowledge of the details of the motion of the individual particles.

[98] In classical physics the same situation occurs. Only densities involving one-time are generally used. This is due (in part) to the instantaneous character of Newtonian forces. However, in problems of fluctuations, many-time distributions may be used [see, e.g., N. Rostoker, Nucl. Fusion 1, 101 (1960)].

Heuristic Considerations

Let us assume provisionally that the system under study is constituted by a collection of N noninteracting particles. In that case we have

$$D_N(x^A; \tau_1 \cdots \tau_N) = D_1(x_1^\mu, u_1^\mu; \tau_1) \otimes \cdots \otimes D_1(x_N^\mu, u_N^\mu; \tau_N).$$

Integrating now *each* D_1 over its own proper time, we obtain

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\tau_1 \cdots d\tau_N D_N(x^A; \tau_1 \cdots \tau_N) = \bigotimes_{i=1}^{i=N} \mathcal{N}_1(x_i^\mu, u_i^\mu) \underset{\text{def}}{\equiv} \mathcal{N}_N(x^A). \tag{5.11}$$

It follows that the normalization of $\mathcal{N}_N(x_A)$

$$\int_{\Sigma_1 \times V_1^+} \cdots \int_{\Sigma_N \times V_N^+} \mathcal{N}_N(x^A) u_1^{\mu_1} \cdots u_N^{\mu_N} \times d\Sigma_{\mu_1} \cdots d\Sigma_{\mu_N}\, d_4u_1 \cdots d_4u_N = 1, \tag{5.12}$$

where Σ_i $(i = 1, \cdots, N)$ is an arbitrary spacelike three-surface imbedded in Minkowski space-time $\mathcal{M}_i^4$.

From this example we could define the "time"-independent densities as

$$\mathcal{N}_k(x_1^\mu, u_1^\mu \cdots x_k^\mu, u_k^\mu) = \int_{k \text{ times}} \cdots \int D_k(x_1^\mu u_1^\mu \cdots x_k^\mu, u_k^\mu, \tau_1 \cdots \tau_k) \times d\tau_1 \cdots d\tau_k, \tag{5.13}$$

where the $\mathcal{N}_k(x_1^\mu u_1^\mu \cdots x_k^\mu u_k^\mu)$ are normalized by

$$\int_{\Sigma_1 \times V_1^+} \cdots \int_{\Sigma_k \times V_k^+} \mathcal{N}_k(x_1^\mu u_1^\mu \cdots x_k^\mu u_k^\mu) \times u_1^{\mu_1} \cdots u_k^{\mu_k}\, d\Sigma_{\mu_1} \cdots d\Sigma_{\mu_k}\, d_4u_1 \cdots d_4u_k = 1. \tag{5.14}$$

Note that the reduced "time"-independent densities could as well have been defined through the relation

$$\mathcal{N}_k(x_1^\mu u_1^\mu \cdots x_k^\mu u_k^\mu) = \int_{\Sigma_{k+1} \times V_{k+1}^+} \cdots \int_{\Sigma_N \times V_N^+} \mathcal{N}_N(x^A) u_{k+1}^{\mu_{k+1}} \cdots u_N^{\mu_N} \times d\Sigma_{\mu_{k+1}} \cdots d\Sigma_{\mu_N}\, d_4u_{k+1} \cdots d_4u_N. \tag{5.15}$$

The consistency of definitions (5.13) and (5.15) is easily verified by passing through the intermediate step of the random densities and taking into account the properties of the Dirac distribution.

The normalization condition (5.14) considered for $k = N$, shows that, as expected, the actual phase space of the system is a $6N$-dimensional manifold

which involves N arbitrary spacelike three-surfaces. Moreover, this shows that the $\mathcal{N}_k$'s may be considered as the distribution functions which generalize to many particles, the ones ordinarily used in relativistic kinetic theory.

In order that the reduced "time"-independent densities should not actually depend on the arbitrary Σ_i's, it is necessary (and also *locally* sufficient) that the tensors

$$J^{\mu_{k+1}\cdots\mu_r}(x_1^\mu u_1^\mu \cdots x_k^\mu u_k^\mu; x_{k+1}^\mu \cdots x_r^\mu) = \int_{V_{k+1}^+} \cdots \int_{V_r^+} \mathcal{N}_r(x_1^\mu u_1^\mu \cdots x_r^\mu u_r^\mu) \times u^{\mu_{k+1}} \cdots u^{\mu_r}\, d_4u_1 \cdots d_4u_r \quad (5.16)$$

should verify a number of conservation relations. These tensors can be called the generalized currents. The relations satisfied by these tensors are obviously

$$\partial_{\mu_i} J^{\mu_{k+1}\cdots\mu_i\cdots\mu_r}(x_1^\mu u_1^\mu \cdots x_k^\mu u_k^\mu; x_{k+1}^\mu \cdots x_r^\mu) = 0 \quad (5.17)$$

with

$$k < i \leq r \leq N.$$

Relations (5.17) are in fact integrability conditions of the differential forms

$$J^{\mu_{k+1}\cdots\mu_i\cdots\mu_r}\, d\Sigma_{\mu_i}$$

(without summation on the index i).

These conditions are equivalent to

$$(\delta/\delta\Sigma_k)\mathcal{N}_i = 0, \quad \text{with} \quad i < k \leq N.$$

Furthermore, they express the fact that the $\mathcal{N}_i$'s are independent of the chosen Σ_k's and of the way from which they have been calculated ($k > i$).

Geometrical Definitions[99]

In order to deal with a more general point of view, let us start directly with the relativistic ensemble of manifolds M_N previously considered (at the end of Sec. 4). As in the relativistic kinetic theory, it is possible to give an invariant definition of $\mathcal{N}_N(x^A)$ only through the intermediate step of a generalized current in Γ space, namely

$$J^{A_1\cdots A_N}(x^A) = \mathcal{N}_N(x^A) \cdot \xi^{A_1\cdots A_N} \quad (5.18)$$

with

$$\xi^{A_1\cdots A_N}(x^A) = \xi_1^{A_1} \otimes \cdots \otimes \xi_N^{A_N}$$

and where ξ_i^A is the Ath component of the $8N$-vector

$$\boldsymbol{\xi}_i = \mathbf{0} \oplus \mathbf{0} \cdots \oplus \boldsymbol{\eta}_i \oplus \cdots \oplus \mathbf{0} \quad (N \text{ factors}),$$

where $\boldsymbol{\eta}$ has already been defined in Sec. 2 while the index i of $\boldsymbol{\eta}_i$ refers to the ith particle.

The tensor $\xi^{A_1\cdots A_N}$ is, in a sense, a generalized velocity in Γ space since it is "tangent" to the manifold M_N passing through the point $x^A \in \Gamma$.

Let now S^{7N} be an *arbitrary* $7N$-dimensional manifold cutting all the manifolds M_N of the ensemble. Moreover S^{7N} must be such that it cuts each M_N in only one point. Hence $\mathcal{N}_N(x^A)$ is normalized to 1 through

$$\int_{S^{7N}} \mathcal{N}_N(x^A) \cdot \xi^{A_1\cdots A_N}\, dS_{A_1\cdots A_N} = 1, \quad (5.19)$$

where $dS_{A_1\cdots A_N}$ is the element of a $7N$-dimensional surface imbedded in an $8N$-dimensional space (i.e., in Γ space).

At first sight it seems that condition (5.19) would imply a $7N$-dimensional actual phase space. However, we must bear in mind that, in general, we have also N relations between the components of the momenta. In such cases, the actual phase space is, as expected, a $6N$-dimensional manifold. Equation (5.19) is written in a fully covariant form (i.e., covariant with respect to arbitrary changes of coordinated in Γ space taking the $(+\ -\ -\ -)$ character of the metric of $\mathcal{M}^4$ into account. However, it reduces to Eq. (5.12) with the coordinates $[\cdots x_i^\mu, u_i^\mu \cdots)]$.

In order to express the conservation of particles in Γ space let us now state and prove Proposition 1.

Proposition 1: The differential form "numerical flux of particles" $J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N}$ is a closed form:

$$d\{J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N}\} = 0.$$

Proof: Let us consider the tube $\mathcal{C}$ generated by an open connected set[100] Δ_1 of an arbitrary surface S_1^{7N} satisfying the above conditions, and the manifolds M_N cutting Δ_1. This tube comprises an $8N$-volume. The frontier of this tube is a $(8N-1)$ manifold. The quantity

$$\int_{\Delta_1} J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N},$$

which is the *flux* through Δ_1 of the generalized current, represents the probability that the system be in a state characterized by a point in Γ space belonging to Δ_1. Let us now cut $\mathcal{C}$ by another $7N$-manifold S_2^{7N} and let Δ_2 be

$$\Delta_2 = \mathcal{C} \cap S_2^{7N}.$$

The physical condition that Δ_2 and Δ_1 contain the

[99] In this section we adopt the field point of view but we reason once a realization of the field is given. See the remarks after this paragraph. Note that the results obtained are also valid in the action-at-a-distance formalism.

[100] For the topology induced on S^{7N} by that of Γ.

same number of systems reads[101]

$$\int_{\Delta_1} J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N} = \int_{\Delta_2} J^{A_1\cdots A_N}(x_A)\, dS_{A_1\cdots A_N}. \quad (5.20)$$

Note that the integral occurring in the right-hand side of Eq. (5.20) has a meaning since Δ_2 is open (for the topology induced by that of Γ on S_2^{7N}) and hence is measurable (for the measure induced on S_2^{7N} by Lebesgue measure in Γ space).

Let now Σ^{7N} be an arbitrary measurable manifold imbedded in the frontier of $\mathcal{C}$ [this is of course possible since the frontier of $\mathcal{C}$ is a $(8N-1)$ manifold] in such a way that the $7N$-dimensional manifold

$$\Delta_1 \cup \Sigma^{7N} \cup \Delta_2 \underset{\text{def}}{\equiv} V^{7N}$$

encloses a $(7N+1)$ volume. Let us assume for a moment that such Σ^{7N}'s can be found (this is in general not true). Then condition (5.20) can be written as

$$\int_{V^{7N}} J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N} = 0 \quad (5.21)$$

since

$$\int_{\Sigma^{7N}} J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N} = 0.$$

Indeed, the differential form induced by $J^{A_1\cdots A_N} \times dS_{A_1\cdots A_N}$ on Σ^{7N} vanishes identically since $J^{A_1\cdots A_N}$ is "tangent" to Σ^{7N} while $dS_{A_1\cdots A_N}$ is "orthogonal" to it. Therefore condition (5.21) implies that the differential form $J^{A_1\cdots A_N}\, dS_{A_1\cdots A_N}$ be a closed form:

$$d\{J^{A_1\cdots A_N}(x^A)\, dS_{A_1\cdots A_N}\} = 0. \quad (5.22)$$

This latter condition is merely a local condition so that we can get rid of the assumption of the existence of Σ^{7N} in the following way. To this end it is sufficient to take for Δ_1 a small neighborhood of a point x^A in Γ space, and a small Δ_2 near Δ_1. Then, it is easy to show that there exists an *infinitesimal* Σ^{7N} with the required properties.[102] Q.E.D.

Equation (5.22) is equivalent to the conditions

$$\nabla_{A_i} J^{A_1\cdots A_i\cdots A_N}(x^A) = 0, \quad i = 1\cdots N, \quad (5.23)$$

where the caret denotes the antisymmetrical part of $J^{A_1\cdots A_N}$. Equation (5.23) can equivalently be written as

$$\nabla_{A_i}\{\mathcal{N}_N(x_A)\xi_1^{A_1}(x_A) \wedge \cdots \wedge \xi_N^{A_N}(x_A)\} = 0, \quad i = 1\cdots N \quad (5.24)$$

or

$$\sum_{j=1}^{j=N} (-)^{j+1}\mathcal{N}_N(x_A)\nabla_{A_i}\xi_j^{A_j}(x_A) \wedge \xi_1^{A_1}(x_A) \wedge \cdots \wedge \xi_N^{A_N}(x_A) + \partial_{A_i}\mathcal{N}_N(x_A)\xi_1^{A_1}(x_A) \wedge \cdots \wedge \xi_i^{A_i}(x_A) \wedge \cdots \wedge \xi_N^{A_N}(x_A) = 0. \quad (5.25)$$

Remarks and Discussion

(1) The requirement of condition (5.22) or of the equivalent Eqs. (5.23), (5.24), and (5.25) is nothing but the relativistic form of the conservation of the number of particles of the system [i.e., Eqs. (5.24) are nothing but continuity equations in Γ space]. Therefore we see a first important difference with the nonrelativistic case.

(2) Instead of conditions (5.23) we could have imposed the simpler relations

$$\partial_{A_i} J^{A_1\cdots A_i\cdots A_N}(x_A) = 0, \quad i = 1\cdots N. \quad (5.26)$$

However, it would be stronger than Eq. (5.23) since it would also imply the vanishing of the divergence of the symmetrical part of $J^{A_1\cdots A_N}$. In fact, only the antisymmetrical part of $J^{A_1\cdots A_N}$ is physically relevant[103] since the differential form $dS_{A_1\cdots A_N}$ is (by construction) completely antisymmetrical. We might perfectly have defined the generalized current as being

$$J^{A_1\cdots A_N} = \mathcal{N}_N \xi_1^{A_1} \wedge \cdots \wedge \xi_N^{A_N}$$

instead of Eq. (5.18). [Note also that despite the symmetrical definition of $J^{A_1\cdots A_N}$, this quantity is actually not symmetrical. The only symmetry property of $J^{A_1\cdots A_N}(x_A)$ is the following[104]:

$$J^{A_1\cdots A_i\cdots A_j\cdots A_N}(x_1^\mu u_1^\mu \cdots x_i^\mu u_i^\mu \cdots x_j^\mu u_j^\mu \cdots x_N^\mu u_N^\mu) = J^{A_1\cdots A_j\cdots A_i\cdots A_N}(x_1^\mu u_1^\mu \cdots x_j^\mu u_j^\mu \cdots x_i^\mu u_i^\mu \cdots x_N^\mu u_N^\mu) \quad \forall(i,j) \leq N,$$

which is *not*

$$J^{A_1\cdots A_i\cdots A_j\cdots A_N}(x_A) = J^{A_1\cdots A_j\cdots A_i\cdots A_N}(x_A).]$$

Relations (5.23) or (5.26) may be simplified further and yield the N relations

$$\nabla_{A_i}\{\mathcal{N}_N \xi_i^{A_i}\} = 0; \quad A_i = 1\cdots 8N, \quad i = 1\cdots N, \quad (5.27)$$

[101] Because of the fact that Δ_1 and Δ_2 are cuts of $\mathcal{C}$, which is invariant under $\{T_{\tau_1\cdots\tau_N}\}$. Note that in the field point of view $\{T_{\tau_1\cdots\tau_N}\}$ actually exists but for each realization of the fields. In other words $\{T_{\tau_1\cdots\tau_N}\}$ is, in the field point of view, a random group (or semigroup).

[102] In order to show this possibility, it is sufficient to give sets of R^{8N} having the same properties as Δ_1, Δ_2, s_1^{7N}, S_2^{7N}, $\mathcal{C}$ etc., and next to take a homeomorphism and map these sets into Δ_1, Δ_2, S^{7N}, $\mathcal{C}$ etc. Fortunately, in R^{8N} we can easily find such sets by taking "cubes" and sections by planes.

[103] See Sec. 6.

[104] Of course, this symmetry property holds only when we deal with identical particles.

which [in the system of coordinates $(\cdots x_i^\mu u_i^\mu \cdots)$] reduces to

$$\partial_{A_i}\{\mathcal{N}_N \eta_i^{A_i}\} = 0; \quad A_i = 1 \cdots 8, \quad i = 1 \cdots N. \tag{5.28}$$

When relations

$$\partial_{A_i}\eta_i^{A_i} = 0; \quad A_i = 1 \cdots 8, \quad i = 1 \cdots N \tag{5.29}$$

holds, then the N equations (5.28) reduce to N one-particle Liouville equations. Therefore, we can never have a Liouville equation, though a relativistic Liouville theorem be true. Indeed, we have

$$d\mathcal{N}_N = \sum_{i=1}^{i=N} \eta_i^{A_i} \cdot \partial_{A_i}\mathcal{N}_N \, d\tau_i = 0 \tag{5.30}$$

since each term of the preceding sum vanishes identically. Therefore we find again results obtained in a much simpler manner with the proper time-dependent formalism.

(3) In the preceding paragraph, we used the most general $7N$ manifold cutting the various M_N of the ensemble. As a consequence we obtained N equations of continuity in Γ space. Nevertheless, we might as well have chosen as S^{7N} a manifold of the form

$$S^{7N} = \Xi^{3N} \times U_1^4 \times \cdots \times U_N^4$$

with $\Xi^{3N} \subset \mathcal{M}^{4N}$. In such a case, continuity equations read

$$\partial_{\mu_i} \hat{J}^{\mu_1 \cdots \mu_N}(x_1^\mu, \cdots, x_N^\mu) = 0, \quad i = 1 \cdots N \tag{5.31}$$

or

$$\partial_{\mu_i} J^{\mu_1 \cdots \mu_N}(x_1^\mu, \cdots, x_N^\mu) = 0, \quad i = 1 \cdots N \tag{5.32}$$

according to the choice of Eq. (5.23) or (5.26), respectively.[105]

On a surface Ξ^{3N} of the type

$$\Xi^{3N} = \Sigma_1 \times \cdots \times \Sigma_N$$

$\mathcal{N}_N(x_A)$ is, in a sense, the relativistic analog of a classical many-time distribution. In the case where

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_N = \Sigma,$$

then, on $\Xi^{3N} = (\Sigma)^N$, $\mathcal{N}_N$ is similar to a single time distribution. These similarities are much more convincing when the surfaces Σ_i are chosen to be spacelike 3-planes.

(4) All the preceding results are valid in the field formalism when we are given a realization of the incident field. However, they are not completely general and we see in Paper II how to modify them.

[105] In Eqs. (5.31) and (5.32) we have used for the currents in $\mathcal{M}^{4N}$ the same symbol J already employed for the current in Γ space. However, the argument of these tensors is sufficient to avoid a possible confusion.

Reduced Densities

Exactly as in the first paragraph of this section, we can obtain various reduced densities. For instance we have

$$\mathcal{N}_{N-1} = \int_{\Sigma^{(7)} \subset \mu^{(8)}(N)} \mathcal{N}_N(x_A)\eta^A \, d\Sigma_A(N) \tag{5.33}$$

(where the index A goes from 1 to 8 and where η^A has been defined in Sec. 2). Note that in Eq. (5.33) $d\Sigma_A(N)$ is the differential form "element of surface" corresponding to a surface imbedded in $\mu^{(8)}(N)$, itself corresponding to the variables referring to the Nth particle.

In order that $\mathcal{N}_{N-1}$ should not depend on $\Sigma^{(7)}$, the following conservation relations must hold:

$$\nabla_{A_i}\{\mathcal{N}_N(x_A)\eta^{A_i}(x_i^\mu u_i^\mu)\} = 0 \tag{5.34}$$

$(A_i = 1 \cdots 8; \; i = 1 \cdots N)$, and they are nothing but Eq. (5.28). Of course, other reduced densities can be obtained in the same way. More generally, one has

$$\mathcal{N}_k(x_1^\mu, u_1^\mu; \cdots; x_k^\mu, u_k^\mu) = \int \mathcal{N}_l(x_1^\mu, u_1^\mu; \cdots; x_l^\mu, u_l^\mu) \bigwedge_{i=1}^{i=l-k} \phi_i^{A_i} \, d\Sigma_{A_1 \cdots A_{l-k}} \times \Sigma^{7(l-k)} \subset \mu^{8(l-k)}(l, \cdots, l-k),$$

with

$$A_i = 1, \cdots, 8(-k); \quad i = 1, \cdots, (l-k)$$

and where the $\phi_i^{A_i}$'s are defined in a way similar to that of $\xi_i^{A_i}$, the difference being only in the number of components.

The independence of the $\mathcal{N}_k$'s (a) from the $\mathcal{N}_l$'s through which they have been obtained, (b) of the various surfaces $\Sigma^{7(l-k)}$, implies a number of conservation relations; these relations are, however, automatically verified when Eq. (5.28) holds. In the same way, one may find the form of integrability conditions in $\mathcal{M}^{4N}$: they are Eq. (5.17).

Currents in Γ Space Connection with the "Time"-Dependent Densities

In the relativistic phase space we can define the *generalized random currents* as being

$$\begin{aligned} J^{A_1 \cdots A_N}_{\text{Random}}(x_A) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\tau_1 \cdots d\tau_N (N!)^{-1} \\ &\quad \times \xi_1^{A_1}(\tau_1) \otimes \cdots \otimes \xi_N^{A_N}(\tau_N) \\ &\quad \times \sum_{i=1}^{i=N} \prod_{i_j=1}^{i_j=N} \delta[x_i^\mu - x_{i_j}^\mu(\tau_{i_j})] \\ &\quad \otimes \delta[u_i^\mu - u_{i_j}^\mu(\tau_{i_j})] \qquad (5.35) \\ &\equiv \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\tau_1 \cdots d\tau_N R_N(x_A; \tau_1 \cdots \\ &\quad \cdots \tau_N) \xi_1^{A_1}(\tau_1) \otimes \cdots \otimes \xi_N^{A_N}(\tau_N). \end{aligned}$$

Then, using the properties of the δ's occurring in Eq. (5.35), we get

$$J^{A_1\cdots A_N}_{\text{Random}}(x_A) = \xi_1^{A_1} \otimes \cdots \otimes \xi_N^{A_N} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\tau_1 \cdots d\tau_N R_N(x_A, \tau_1 \cdots \tau_N)$$

which proves, after taking the average value of both sides of this last equality and comparing the result with the previous definition of $J^{A_1\cdots A_N}$, that

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} D_N(x_A, \tau_1 \cdots \tau_N)\, d\tau_1 \cdots d\tau_N = \mathcal{N}_N(x_A)$$

as is expected. Of course, this is the same argument which has been applied to distributions D_1 and $\mathcal{N}_1$ at the end of Sec. 2.

6. AVERAGE VALUES

In this section we generalize the average values discussed at the end of Sec. 2. The same troubles as those already indicated also occur (*mutatis mutandis*). Once average values are defined, entropies for the reduced densities are obtained in a straightforward way. This allows us to derive the canonical distribution of a relativistic simple gas as an *illustration* of the whole formalism.

Assume that the system possesses a property represented by a certain tensor $B^{\mu_1\cdots\mu_\alpha}_{\nu_1\cdots\nu_\beta}(x^A)$. Then we can define the *local flux* $\mathfrak{B}^{\mu_1\cdots\mu_\alpha\rho_1\cdots\rho_N}_{\nu_1\cdots\nu_\beta}(x_1^\mu \cdots x_N^\mu)$ as

$$\mathfrak{B}^{\mu_1\cdots\mu_\alpha\rho\cdots\rho_N}_{\nu_1\cdots\nu_\beta}(x_1^\mu \cdots x_N^\mu) = \int_{V^{+N}} B^{\mu_1\cdots\mu_\alpha}_{\nu_1\cdots\nu_\beta}(x_A)\mathcal{N}_N(x_A) \times u_1^{\rho_1}\cdots u_N^{\rho_N}\, d_4u_1 \cdots d_4u_N. \quad (6.1)$$

In fact, we should skew-symmetrize the indices $\rho_1 \cdots \rho_N$ of the tensor $\mathfrak{B}^{\mu_1\cdots\nu_\alpha\rho_1\cdots\rho_N}_{\nu_1\cdots\nu_\beta}(x_1^\mu \cdots x_N^\mu)$. However, as we have already indicated in Sec. 5, the symmetrical part of $\mathfrak{B}$ (in the indices $\rho_1 \cdots \rho_N$) *plays no role* in the following integral (6.2). The total amount of the quantity $B^{\mu_1\cdots\mu_\alpha}_{\nu_1\cdots\nu_\beta}(x_A)$ through the $3N$-dimensional surface Ξ^{3N}, will be the total flux[106] through this arbitrary surface[107] of the tensor $\mathfrak{B}$:

$$\langle B^{\mu_1\cdots\mu_\alpha}_{\nu_1\cdots\nu_\beta}(x_A)\rangle_{\Xi^{3N}} = \int_{\Xi^{3N}} \mathfrak{B}^{\mu_1\cdots\mu_\alpha\rho_1\cdots\rho_N}_{\nu_1\cdots\nu_\beta}\, d\Xi_{\rho_1\cdots\rho_N} \quad (6.2)$$

and, in general, will depend on Ξ^{3N}. When the property $B^{\mu_1\cdots\mu_\alpha}_{\nu_1\cdots\nu_\beta}(x_A)$ is a permanent property of the system, then the local flux $\mathfrak{B}$ satisfies the integrability conditions:

$$\partial_{\rho_i}\mathfrak{B}^{\mu_1\cdots\mu_\alpha\rho_1\cdots\rho_i\cdots\rho_N}_{\nu_1\cdots\nu_\beta}(x_1^\mu \cdots x_N^\mu) = 0 \quad (i = 1\cdots N). \quad (6.3)$$

As an example, let us consider the local flux of the energy-momentum four-vector of a fluid consisting of identical particles. We have

$$B^\mu(x_A) = \sum_{i=1}^{i=N} mu_i^\mu$$

$$\mathfrak{B}^{\mu\rho_1\cdots\rho_N}(x_1^\mu \cdots x_N^\mu) = \sum_{i=1}^{i=N} \int mu_i^\mu u_1^{\rho_1}\cdots u_N^{\rho_N}\mathcal{N}_N(x_A)\, d_4u_1 \cdots d_4u_N = N\int mu_1^\mu u_1^{\rho_1}u_2^{\rho_2}\cdots u_N^{\rho_N}\mathcal{N}_N(x_A)\, d_4u_1\cdots d_4u_N$$

and the total energy-momentum four-vector of the fluid considered on a $3N$-dimensional surface Ξ^{3N} is

$$\langle p^\mu\rangle_{\Xi^{3N}} = N\iint_{\Xi^{3N}} mu_1^\mu u_1^{\rho_1}\cdots u_N^{\rho_N}\mathcal{N}_N(x_A)d\Xi_{\rho_1\cdots\rho_N}d_4u_1\cdots d_4u_N = N\iint_{\Sigma_1} mu_1^\mu u_1^{\rho_1}\mathcal{N}_1(x_1^\mu p_1^\mu)\, d\Sigma_{\rho_1}\, d_4u_1 = \int_{\Sigma_1} T^{\mu\rho_1}(x_1^\mu)\, d\Sigma_{\rho_1}. \quad (6.4)$$

[In deriving Eq. (6.4) we have used the properties of "reduction" of $\mathcal{N}_N$ to $\mathcal{N}_1$: Σ_1 is a three-dimensional spacelike surface imbedded in Ξ^{3N} and depending on the variable x_1^μ only.]

Entropies

Given a reduced distribution function at fixed t and of order k, its entropy at time t is generally defined as being the average value of its logarithm.[108]

Consequently, we take as the entropy of the distribution $\mathcal{N}_k$, the average value of $\log \mathcal{N}_k$ on a surface Ξ^{3k}:

$$S_k(\Xi^{3k}) = -\lambda\int_{\Xi^{3k}} \sigma^{\rho_1\cdots\rho_k}\, d\Xi_{\rho_1\cdots\rho_k}, \quad \lambda > 0, \quad (6.5)$$

[106] The conventional definition of average values is, of course, found anew when we take as Ξ^{3N} the $3N$ plane. $t_1 = \text{const}; \cdots; t_N = \text{const}$.

[107] Arbitrary, but "spacelike."

[108] We say "generally" because the entropy (in the sense of information theory) is not defined unambiguously [either for a continuous random variable or even for a discrete one which depends on an external parameter: see, e.g., B. Mandelbrojt, IBM research note NC-107 (1962)]. Furthermore the entropy of a measure is always defined with respect to a subjacent measure. In statistical mechanics this subjacent measure is almost always the Lebesgue measure. Here we are concerned with this last point of view, a possible generalization being straightforward.

where λ is an arbitrary positive constant[109] and where $\sigma^{\rho_1 \cdots \rho_k}$ is the *k-entropy flux tensor* given by

$$\sigma^{\rho_1 \cdots \rho_k}(x_1^\mu \cdots x_k^\mu) = \int \log \mathcal{N}_k \times \mathcal{N}_k u_1^{\rho_1} \cdots u_k^{\rho_k} \, d_4u_1 \cdots d_4u_k \quad (6.6)$$

as it should be. Hence, the entropy of $\mathcal{N}_N$ is

$$S_N(\Xi^{7N}) = -\lambda \int_{\Xi^{7N}} \log \mathcal{N}_N(x_A) \times \mathcal{N}_N(x_A) \xi^{A_1 \cdots A_N} \, d\Sigma_{A_1 \cdots A_N}. \quad (6.7)$$

When the k entropy of a physical system is constant, then the associated entropy flux tensor satisfies conservation relations:

$$\partial_{\rho_i} \sigma^{\rho_1 \cdots \rho_k}(x_1^\mu \cdots x_k^\mu) = 0, \quad i \leq k. \quad (6.8)$$

Equation (6.5) and (6.6) are the generalization of the entropy and entropy current given by Tauber and Weinberg, by Chernikov, and by Israel:

$$S_1(\Sigma) = -\lambda \int_\Sigma \sigma^\mu \, d\Sigma_\mu, \quad (6.9)$$

$$\sigma^\mu(x_\nu) = \int \log \mathcal{N}_1 \cdot \mathcal{N}_1 u^\mu \, d_4u. \quad (6.10)$$

So an H theorem is expressed by stating

$$\partial_\mu \sigma^\mu < 0. \quad (6.11)$$

Remarks and Discussion

(1) Local average values may be defined as being

$$\langle B^{\mu_1 \cdots \mu_\alpha}_{\nu_1 \cdots \nu_\beta} \rangle_{x_\nu} = \frac{\mathcal{B}^{\mu_1 \cdots \mu_\alpha \rho_1 \cdots \rho_N} J_{\rho_1 \cdots \rho_N}}{\{J^{\rho_1 \cdots \rho_N} J_{\rho_1 \cdots \rho_N}\}} \Bigg|_{x_{\nu_i} = x_{\nu_i}, \forall i \leq N}. \quad (6.12)$$

This definition is the natural generalization of Eq. (2.35) and it suffers the same troubles. In particular, it seems to be hardly possible to give them a form similar to the nonrelativistic one.

(2) In Sec. 4 we mentioned that our phase space might be insufficient and thus should be enlarged.[110] Let us now explain this point. It is clear that our preceding statistical notions (Γ space, densities, average values) only permit the calculation of average values of those quantities which depend on the variables ($\cdots x_i^\mu$, $u_i^\mu \cdots$). If we want to calculate the average value of a quantity depending, for instance, on acceleration variables, then phase space would have to be enlarged so as to include them. It would be a $12N$-dimensional space. Of course, subsequent densities on it would have to be defined: they are straightforward generalizations of the definitions given in Sec. 5. In practice we know only quantities depending on ($\cdots x_i^\mu$, $u_i^\mu \cdots$). However, when dealing with radiation phenomena, acceleration variables should be taken into account.[111] These questions are studied in a detailed manner in Paper II.

(3) Because of the fact that $D_N(x_A; \tau_1 \cdots \tau_N)$ is a true probability distribution, one might think that average values could be defined as usual: i.e., as

$$\langle B^{\mu_1 \cdots \mu_\alpha}_{\nu_1 \cdots \nu_\beta}(x_1^\mu, u_1^\mu \cdots x_N^\mu, u_N^\mu) \rangle_{(\tau_1 \cdots \tau_N)} = \int_\Gamma B^{\mu_1 \cdots \mu_\alpha}_{\nu_1 \cdots \nu_\beta}(x_A) D_N(x_A; \tau_1 \cdots \tau_N) \, d\Gamma.$$

However, these mean values depend on the adopted parametrization. Therefore they must be considered as unphysical. However, they have some interesting consequences. For instance we can compute the mean value of x^A ($A = 1 \cdots 8N$); it is a function $\langle x^A \rangle_{(\tau_1 \cdots \tau_N)}$ of the N proper times, which determines a mean manifold in Γ space. In the same way we can define the "center of mass" [112] of a relativistic fluid when we know its "instantaneous" distribution $D_1(x_\mu p_\mu; \tau)$, by

$$\langle x^\mu \rangle_\tau = \int_\mu x^\mu D_1(x^\mu, p^\mu; \tau) \, d\mu.$$

In $\mathcal{M}^4$ this equation defines a *mean world line* whose points are the centers of mass of the fluid at different "times." Of course, this mean world line is not unique; it depends on the initial data and on the parametrization. Let us remark that this lack of uniqueness of the center of mass, due to the arbitrariness of the parametrization, is a kind of "temporal" counterpart of the arbitrariness of the "physical space," i.e., spacelike 3-manifolds. Indeed it is well known[113] that, in general, the center of mass of a relativistic fluid depends on the three-surface where it is computed and hence is largely arbitrary.

In this connection it is interesting to note that we have a kind of "Ehrenfest theorem" which reads

$$(d/d\tau)\langle x^\mu \rangle_\tau = \langle u^\mu \rangle_\tau,$$
$$(d/d\tau)\langle u^\mu \rangle_\tau = \left\langle \frac{F^\mu}{m} \right\rangle_\tau,$$

and which may easily be proved in the one-particle case.[114]

---

[109] λ depends on the system of units chosen. It is in general equal to the Boltzmann constant k. It also depends on the basis used for the log.

[110] In such a case, densities are also insufficient and should be generalized so as to be densities on the enlarged phase space.

[111] For instance, the radiation field is proportional to $F^{\mu\nu}_{\text{rad}} \alpha \{u^\mu \dot{\gamma}^\nu - u^\nu \dot{\gamma}^\mu\}$, where $\dot{\gamma}^\mu$ can be expressed in terms of $(x_\nu, u_\nu, \gamma_\nu)$ (see Paper II).

[112] Neither the expression "center of mass" nor "center of gravity" nor "barycenter" is appropriate. See, e.g., J. L. Synge, Ref. 59.

[113] See, e.g., C. Møller, *The Theory of Relativity* (Clarendon Press, Oxford, 1952), or J. L. Synge, Ref. 59.

[114] In the case of interacting particles, its proof needs some weak conditions on the nature of interaction.

An Illustration of the Formalism

As a simple application of the above formalism let us derive the canonical distribution of a relativistic gas of noninteracting particles, at local equilibrium.[115] To this end, we assume[116] that such a distribution can be obtained (a) from the maximization of the entropy while (b) the average total momentum-energy and (c) the number of particles within the gas, are given. Since we are dealing with local equilibrium conditions, (a), (b), and (c) have to be replaced by their local equivalents. They are

$$\begin{aligned} &\text{(a)} \quad \delta\langle \log \mathcal{N}_N \rangle_{x_\nu} = \delta\{\sigma^{\mu_1 \cdots \mu_N} U_{\mu_1 \cdots \mu_N}\} = 0, \\ &\text{(b)} \quad \langle P^\mu \rangle_{x_\nu} = \mathcal{C}^{\mu\mu_1 \cdots \mu_N} U_{\mu_1 \cdots \mu_N}, \\ &\text{(c)} \quad \langle n \rangle_{x_\nu} = J^{\mu_1 \cdots \mu_N} U_{\mu_1 \cdots \mu_N}, \end{aligned} \tag{6.13}$$

where

$$U^{\mu_1 \cdots \mu_N} \equiv J^{\mu_1 \cdots \mu_N} \{J^{\mu_1 \cdots \mu_N} J_{\mu_1 \cdots \mu_N}\}^{-\frac{1}{2}} \Big|_{x_{\nu_i} = x_\nu}, \qquad \forall i \leq N,$$

$$\mathcal{C}^{\mu\nu_1 \cdots \nu_N} = \sum_{i=1}^{i=N} m \int d_4 u_1 \cdots d_4 u_N u_i^\mu u_1^{\nu_1} \cdots u_N^{\nu_N} \mathcal{N}_N.$$

Introducing now five x_ν-dependent Lagrange multipliers $\log A$, ξ_μ, conditions (a), (b), and (c) lead to

$$\delta\left\{U_{\mu_1 \cdots \mu_N} \int d_4 u_1 \cdots d_4 u_N u_1^{\mu_1} \cdots u_N^{\mu_N} \mathcal{N}_N \times \left[\log \mathcal{N}_N - \log A + m \sum_{i=1}^{i=N} u_i^\mu \xi_\mu\right]\right\} = 0 \tag{6.14}$$

and finally we obtain

$$\mathcal{N}_N(x_A) = A \exp\left\{-\xi_\mu \sum_{i=1}^{i=N} m u_i^\mu\right\}. \tag{6.15}$$

This result is, as expected, consistent with the Jüttner–Synge distribution.

Let us remark that contrary to the nonrelativistic case we can define another kind of canonical distribution: instead of the conservation of the four-momentum of the gas we can impose the conservation of the total mass of the gas. Then instead of relation (6.15), we should find

$$\mathcal{N}_N(x_A) = A \exp\left[-\xi\left\{\sum_{i=1}^{i=N} m_0 u_1^\mu\right\}^2\right]. \tag{6.16}$$

We discuss the case of interacting particles in Sec. 7.

[115] Since the notion of a box is not a covariant concept, the uniform density in configuration space is not normalizable. Consequently we limit ourselves (and this is not an essential limitation) to the case of local equilibrium, which case allows the normalization of the distribution function.

[116] In general, the canonical distribution may be obtained from the microcanonical one. However, on the basis of information theory arguments [see, e.g., E. T. Jaynes, Phys. Rev. **106**, 620 (1957)] the canonical distribution may also be derived.

7. AN UNSOLVED PROBLEM: EQUILIBRIUM

In this section we discuss the possibility of defining relativistic equilibrium states. We also suggest elements for a possible solution of this unsolved problem.

A Preliminary Unsolved Problem: What is Equilibrium?

(1) Let us consider a nonrelativistic system of particles whose equations of motion are given

$$m_i(d\mathbf{v}_i/dt) = \mathbf{F}(\mathbf{x}_i, \mathbf{v}_i; t), \quad i = 1 \cdots N. \tag{7.1}$$

Assume now that the initial data $(\cdots \mathbf{x}_{i0}, \mathbf{v}_{i0}, \cdots)$ are chosen at random so that we actually deal with statistical mechanics. When may such a classical system be considered as being in an "equilibrium state"? If the equation of motion (7.1) can be cast into a Hamiltonian form,[91] then the usual definitions of equilibrium apply[117]: We first define a *microcanonical* distribution with the help of the physically interesting constant of motion, next the *canonical* distribution is derived for a subsystem.[118] At this point it should be emphasized that the equilibrium distributions obtained are both integrals of motion and solutions of the continuity equation in phase space (by virtue of the Liouville theorem). It should also be pointed out that the notion of temperature is closely related to the choice of the energy as constant of motion to be introduced in the microcanonical distribution.

(2) When Eq. (7.1) cannot be set into a Hamiltonian form,[91] then the situation is by no means so "simple." It seems indeed hardly possible to define equilibrium states as is shown from the following. [Let us consider the trivial example[119] of *one* particle acted on by a friction force and performing (for simplicity) a one-dimensional motion:

$m(dv/dt) + \beta v = 0$; $dx/dt = v$ (β is the friction coefficient).

[117] D. Massignon, *Mécanique statistique des fluides* (Dunod Cie, Paris, 1957).

[118] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover Publication, Inc., New York, 1949).

[119] This example *only* illustrates this point. One might argue that nonconservative forces do not exist in nature. However, this is not completely correct (cf. the Lorentz equation taking radiation reaction into account). One might also say that nonconservative forces are nonconservative only at a macroscopic level and that there will always exist either a conservative model at a lower level or possibly a larger system responsible for the energy nonconservation. In our opinion, these arguments are irrelevant when dealing with a theoretical problem: Newton's second law allows any choice of velocity dependent forces and thus we have to face (from a theoretical point of view) such a general case. Furthermore such systems *do exist* in nonquantal relativity and hence the problem should be solved.

The main properties of this system are

(a) solutions:

$$\begin{cases} v(t) = v_0 \exp\{-\beta t\}, \\ x(t) = x_0 - \beta^{-1} v_0 [\exp\{-\beta t\}, -1]. \end{cases}$$

(b) Phase space: $\{x_0\} \times \{v_0\}$ (two dimensions).

(c) Constant of motion: $mv + \beta x = \text{const}$.

(d) Ergodism:

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T [v(t)]^q \, dt = 0, \quad q > 0.$$

Microcanonical distribution cannot be defined because the constant of motion (c) has no direct physical meaning. The explicit form (a) of the solution of the equations of motion indicates that, anyway, the velocity tends to zero at infinity in times; consequently all possible densities at $t = +\infty$ will have the form

(e) $\rho(x, v)$

$$= \{\text{arbitrary positive function of } x\} \times \delta(v),$$

which agrees with property (d). However, although $\rho(x, v)$ satisfies the continuity equation

$$\text{(f)} \qquad \frac{\partial}{\partial t} \rho + v \frac{\partial}{\partial x} \rho + \frac{\partial}{\partial v} \{-\beta v \rho\} = 0$$

in phase space, it is no longer a constant of motion, i.e., $(d/dt)\rho \neq 0$. Note also that energy is not well defined.]

Furthermore energy is not always defined and so the notion of temperature fails. The above example shows, however, that statistical equilibria (if not thermodynamical)[120] might be defined.

Finally we stress that the question of equilibrium is not a specifically relativistic one and should be solved first in a classical framework.[121]

(3) What might be expected from an equilibrium state? First an equilibrium distribution should satisfy the continuity equation whether there exists a Liouville theorem or not. Next it must be stationary in time and perhaps also invariant under space translations. [Let us come back to the above example. These conditions only imply that ρ should verify

$$(\partial/\partial v)\{\beta v \rho\} = 0$$

or

$$\rho v = \text{const},$$

which yields

$$\rho = \text{const}/v + (\text{another const}) \times \delta(v)$$

and finally since ρ is to be positive and normalized to unity, it follows that $\rho = \delta(v)$.] However, these conditions are not sufficient by themselves in spite of the chosen example, which is much too simple. Unfortunately we do not know what conditions should be added.[122] Anyway an equilibrium distribution is in general *not* a first integral of the motion.[123]

Relativistic Equilibrium in the Field Point of View[124]

The notion of equilibrium used in the field viewpoint by a number of authors[10,125] is far from being completely clear. Indeed, apart from usual infinities which occur in dealing with electromagnetic phenomena,[126] there always remain the divergences due to the self-fields. Furthermore, infinite bare masses are involved in the equilibrium pseudodistribution obtained and therefore its physical meaning seems to be troublesome.[127]

However, the use of the so-called field oscillators may be considered as a heuristic tool. In particular, it is interesting to note that the canonical distribution for a blackbody indicates that the electromagnetic field is a Gaussian random process with zero average value and a spectrum given by Planck's law.

This suggests a definition of the equilibrium for the radiation field[128] by assuming:

(a) $F^{\mu\nu}_{\text{rad}}$ is a Gaussian random process invariant under space–time translations.

(b) with: $\langle F^{\mu\nu}_{\text{rad}} \rangle = 0$, $\partial\mu \langle F^{\mu\nu}_{\text{rad}} \rangle = 0$, $\partial\mu \langle F^{\mu\nu *}_{\text{rad}} \rangle = 0$.

(c) with: Fourier transform of $\langle F^{\mu\nu}_{\text{rad}} \otimes F^{\mu\nu}_{\text{rad}} \rangle \sim$ {blackbody spectrum}.

In actuality, the notion of thermal equilibrium seems to have only a weak meaning when radiation is considered especially due to the absence of classical photons. It is, however, usual to speak about "the equilibrium of the field," etc., notions which might

[120] We distinguish between "statistical equilibrium" and "thermodynamical equilibrium," so the microcanonical equilibrium is a statistical equilibrium (arising only from mechanical considerations) while the canonical one arises also from considerations extraneous to mechanics (occurrence of macroscopic parameters).

[121] The situation can be even worse if we consider the possible existence (at least theoretical) of the so-called "hereditary systems" whose study was first initiated by Volterra [V. Volterra, *Sur les fonctions de ligne* (Gauthier-Villars, Paris, 1928), and J. Math. Pures Appl. 7, 249 (1928)]. The action-at-a-distance formalism furnishes an excellent example of such a system.

[122] Perhaps a condition of maximization of entropy would be sufficient to select the "equilibrium distribution" among the solutions of $(\partial/\partial v)\{\mathbf{F}f\} = 0$?

[123] We return to the discussion of equilibrium in a future paper.

[124] In all that follows we consider only the case of electromagnetic interactions.

[125] Ph. de Gottal, Physica. 32, 548 (1966).

[126] They are irrelevant here since they also occur in the classical case.

[127] Note that the microcanonical ensemble for the system (particles + field) cannot be defined, both because of mathematical impossibilities and because of the invariance requirements.

[128] And not for the total field. In fact, it is difficult to separate all different kinds of fields. We show, however, in Paper II that this is actually possible for the radiation field.

have a sense only within the framework of quantum electrodynamics.

Equilibrium in the Action-at-a-Distance Formalism[124]

(1) In Paper II, we give several hierarchies either verified by the distribution functions $\mathcal{N}_1$, $\mathcal{N}_2$, etc., or by generalized similar densities; i.e., relations of the form

$$A\mathcal{N}_1 = B\mathcal{N}_2 + EP_2,$$
$$C\mathcal{N}_2 = D\mathcal{N}_3 + \cdots,$$
$$\cdots,$$

relating these densities ($A, \cdots, E$ are operators). These hierarchies are generalizations of the well-known BBGKY chain. If we impose now that these densities should be invariant under space–time translations, if we impose that $\mathcal{N}_1$ should be the Jüttner–Synge distribution, then the hierarchy obtained *might* be considered as an equilibrium hierarchy similar to the classical one. In so doing, we have computed the relativistic correlation function[129] at order 1 in e^2 of a gas at equilibrium. However, it is not clear whether or not higher orders can also be calculated (see Paper II for a discussion and the details).

(2) A possible way out of these difficulties consists in assuming that canonical equilibrium[116] is obtained as in Sec. 6 from a maximization of entropy (subject to the constraints arising from the knowledge of the number of particles of the system and its *total* momentum-energy). However, in so doing, we obtain a very complicated functional expression from which we have not yet been able to derive the equilibrium densities.

In conclusion, (a) the notion of equilibrium gives rise to unsolved problems even in a classical theory, (b) equilibrium densities should verify one of the hierarchies given in Paper II and which are, in a sense, something similar to the classical continuity equation, (c) equilibrium densities have to be invariant under space-time translations,[130] (d) it is perhaps possible to start directly with canonical equilibrium. In such case, temperature[131] is defined as in Jüttner–Synge distribution.

[129] This calculation is needed in order to obtain a relativistic generalization of Guernsey kinetic equation: R. L. Guernsey, Phys. Fluids 7, 792, 1600 (1964).

[130] At local equilibrium we have to impose a weaker condition: their invariance under a one-parameter transformation group whose orbits are timelike.

[131] It is not at all sure that the relativistic notion of temperature has a sense. In particular, its definition through Jüttner–Synge distribution may be questioned since this density has never (and for good reasons!) been *proved on the basis of microscopical* considerations as it is the case for the usual Maxwell–Boltzmann distribution.

ACKNOWLEDGMENTS

The author is indebted to Professor B. Jancovici, Professor F. Lurçat, Professor P. Mazur, Dr. G. Flamand, Dr. J. Ginibre, and Dr. J. M. Lévy-Leblond for stimulating criticisms and discussions.

APPENDIX. MICROCANONICAL ENSEMBLE (CASE OF THE PERFECT GAS)

Let us consider a perfect gas in the absence of external forces. Its random distribution function is[132]

$$\mathcal{R}_N(x_{\mu 1}, p_{\mu 1}; \cdots; x_{\mu N}, p_{\mu N}) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\tau_1 \cdots d\tau_N \times \sum_{i=N}^{i=1} \prod_{j=1}^{j=N} \delta\{x_{\mu i} - x_{\mu 0 i_j} - m^{-1} p_{\mu i} \tau_j\} \otimes \delta\{p_{\mu 0 i_j} - p_{\mu i}\}. \tag{A1}$$

Such a system is conservative so that the total momentum-energy 4-vector of the gas does not depend on the manner through which it is calculated. Therefore we have

$$P^\mu = \sum_{i=1}^{i=N} p^\mu_{0i} = \text{const.} \tag{A2}$$

It follows that the microcanonical distribution is obtained by averaging R_N over the initial data $(\cdots x^\mu_0, p^\mu_0 \cdots)$. The average operation should be uniform in configuration space $\mathcal{M}^{4N}$ and in the accessible momentum space. We then get

$$\mathcal{N}_N(x^\mu_1, p^\mu_1; \cdots x^\mu_N, p^\mu_N) = \text{const}\, \delta\left\{P^\mu - \sum_{i=1}^{i=N} p^\mu_i\right\} \tag{A3}$$

or, taking into account the N constraints $p^\mu_i p^i_\mu = m^2$, we have

$$\mathcal{N}_N(x^\mu_1, p^\mu_1; \cdots; x^\mu_N, p^\mu_N) = \text{const}\, \delta\left\{P^\mu - \sum_{i=1}^{i=N} p^\mu_i\right\} \otimes \prod_{i=1}^{i=N} \delta\{p^\mu_i p^i_\mu - m^2\} 2\theta(p^0_i), \tag{A4}$$

where the constant is to be fixed by the normalization condition

$$J^{\mu_1 \cdots \mu_N} = \int \cdots \int \mathcal{N}_N \frac{p^{\mu_1}_1}{m} \cdots \frac{p^{\mu_N}_N}{m} d_4 p_1 \cdots d_4 p_N, \tag{A5}$$

$$\{J^{\mu_1 \cdots \mu_N} \cdot J_{\mu_1 \cdots \mu_N}\}^{1/2N} = \rho_0 = \text{constant density.}$$

[Since $J^{\mu_1 \cdots \mu_N}$ is a constant tensor, the usual normalization

$$\int J^{\mu_1 \cdots \mu_N} d\Sigma_{\mu_1 \cdots \mu_N} = 1$$

is no longer valid. We might, however, define a

[132] In the following we use the variables $(\cdots x^\mu_i, p^\mu_i = m u^\mu_i \cdots)$ instead of $(\cdots x^\mu_i, u^\mu_i \cdots)$.

"local microcanonical ensemble" which would be normalizable; to this end it is sufficient to make the five constants which are at our disposal (i.e., ρ_0, P^μ), x_ν-dependent.]

Equation (A3) essentially agrees with the one already given by Lurçat and Mazur[133] in another context. However, their distribution (which is in fact a phase space factor, useful in the statistical model of pions production) is not correctly normalized (i.e., through a current). This is the reason why their final result (the canonical distribution) is not correct, if considered from the point of view of relativistic statistical mechanics.[134]

Let us check that $\mathcal{N}_N$ actually leads to P^μ as the total energy momentum 4-vector for the gas. As in Sec. 6, we calculate the generalized momentum-energy tensor:

$$\mathcal{T}^{\mu_1 \cdots \mu_N \nu} = \int \cdots \int \mathcal{N}_N \cdot \left\{ \sum_{i=1}^{i=N} p_i^\nu \right\} \cdot \prod_{i=1}^{i=N} \left(\frac{P_i^{\mu_i}}{m} \right) d_4 p_i$$
$$= P^\nu \cdot J^{\mu_1 \cdots \mu_N} \qquad \text{(A6)}$$

from which follows the total energy-momentum[135]

$$P^\mu_{\text{tot}} = P^\mu \int J^{\mu_1 \cdots \mu_N} d\Sigma_{\mu_1 \cdots \mu_N}$$
$$= P^\mu \times (\sim 1) \qquad \text{(A7)}$$

as expected.

Note that the averaging operation $\langle\ \rangle$ is simply obtained from $\mathcal{N}_N$. Had we considered the case of interacting particles, the situation would have been inextricable and another difficulty would arise. Indeed, instead of Eq. (A2), we should have a much more complicated expression [see Ref. 49, Eq. (7-96) and following] and furthermore the averaging operation would not be simply obtained from $\mathcal{N}_N$ because of the richer content of $\langle\ \rangle$. Indeed it seems that, at any rate, Cauchy data are not merely $(\cdots x_{i0}, u_{i0} \cdots)$ for interacting particles.

[133] F. Lurçat and P. Mazur, Nuovo Cimento 31, 140 (1964).

[134] In particular their normalization leads to an incorrect factor K_1 in the Jüttner–Synge distribution (instead of K_2).

[135] Note that this average value coincides with the local one. The difficulty of the nonnormalizable character of $\mathcal{N}_N$ is only apparent and is due to the inadequacy of considering a uniform measure over R^N as possessing a density with respect to Lebesgue measure. Indeed, a uniform probability measure is such that: $\mu(R^N) = 1$, $\mu(A) = 0, \forall A \subset R^N$ (with A: μ measurable). Anyway, we may always assume a x_ν dependence in ρ_0.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 6 JUNE 1967

Basic Algebra of Antilinear Operators and Some Applications. I

FEDOR HERBUT AND MILAN VUJIČIĆ*
Institute of Nuclear Sciences "Boris Kidrich," Belgrade, Yugoslavia

(Received 13 October 1966)

The primary aim of this work is to find the canonical (i.e., simplest) form of antilinear operators which is the analog of the diagonal form of linear ones, as well as to obtain that class of antilinear operators which corresponds to the class of normal, i.e., diagonalizable, linear ones. To achieve this aim two basic tools are used: polar factorization of an arbitrary antilinear operator into a linear, Hermitian, positive, semidefinite, and anti-unitary operator $\hat{A}_a = \hat{H}_1\hat{U}_a = \hat{U}_a\hat{H}_2$, and representation of antilinear operators by antilinear matrices, which are products of a matrix factor transforming by unitary congruence transformations and the operation of conjugation which is the same for all antilinear operators and all bases. The canonical form is defined as the simplest form of the matrix factor. The criteria of simplicity are: a quasi-diagonal form with smallest possible submatrices, a maximal number of zeros in them, and as many positive numbers as possible among the nonzero elements. It is found that the analog of the diagonal form of linear operators is the second-order canonical form consisting of a diagonal part with nonnegative elements, which is as large as possible, and of two-by-two submatrices with zeros on the diagonal and with at least one positive element. The operators having this form are those whose polar factors can be simultaneously canonical, taking for the anti-unitary factor, essentially the Wigner canonical form. These antilinear operators are called normal ones, and they can also be defined by the following relations between the polar factors: $[\hat{H}_1, \hat{H}_2]_- = 0$ and $[\hat{H}_1, \hat{U}_a^2]_- = 0$ or by the single commutator, $[\hat{A}_a, (\hat{A}_a^\dagger)^2]_- = 0$. A simple procedure to obtain the canonical form of a given normal antilinear operator is developed. A few applications of the results obtained are outlined. They belong to different fields such as electric network theory, quantum mechanics, and self-consistent Hartree–Bogoliubov theory.

INTRODUCTION

THE only well-known example of an antilinear operator (hereafter AO) used in quantum physics is that of time reversal. Therefore, AO's are not a standard tool of theoretical physicists, and it seems desirable to begin with a short summary of basic definitions.

AO's are those operators in a complex vector space which anticommute with pure imaginary constants, commute with real constants, and preserve summation, i.e.,

$$\hat{A}_a(\alpha\,|a\rangle + \beta\,|b\rangle) = \alpha^*\hat{A}_a\,|a\rangle + \beta^*\hat{A}_a\,|b\rangle,$$

where α, β are complex numbers, $|a\rangle$, $|b\rangle$ are vectors, and the index a on $\hat{A}_a$ denotes antilinearity.

In a unitary space, the adjoint operator of $\hat{A}_a$, $\hat{A}_a^\dagger$, which is also antilinear, is defined by[1]

$$\langle a|\,(\hat{A}_a^\dagger\,|b\rangle) = \langle b|\,(\hat{A}_a\,|a\rangle) = [(\langle b|\,\hat{A}_a)\,|a\rangle]^*.$$

As in the case of linear operators, Hermitian and skew-Hermitian, AO's satisfy $\hat{A}_a^\dagger = \hat{A}_a$ and $\hat{A}_a^\dagger = -\hat{A}_a$, respectively. Unitary AO's (more often called anti-unitary operators) are defined by

$$(\langle a|\,\hat{U}_a^\dagger)(\hat{U}_a\,|b\rangle) = [\langle a|(\hat{U}_a^\dagger\hat{U}_a\,|b\rangle)]^* = [\langle a\,|\,b\rangle]^*,$$

which leads to $\hat{U}_a^\dagger = \hat{U}_a^{-1}$.

The theory of anti-unitary operators has been given an equally firm ground as that of unitary linear operators in a paper by Wigner.[2]

Studying the variational objects of the Hartree–Bogoliubov self-consistent theory in nuclear physics, we found it necessary to deal with AO's which are not unitary, but skew-Hermitian.[3] Hence, we have investigated a few basic problems in the algebra[4] of AO's, inspired by Wigner's paper[2] and making use of the analogy with linear algebra.[5]

The treatment of this paper has been restricted to finite dimensional complex unitary vector spaces V_n for the sake of simplicity.

I. POLAR FACTORIZATION OF ANTILINEAR OPERATORS

Generalizing the polar factorization of a complex number into a nonnegative number and a phase factor, *one may write every AO in polar form, i.e., as the product of a linear Hermitian, positive, semidefinite, and anti-unitary operator:*

$$\hat{A}_a = \hat{H}_1\hat{U}_a. \tag{1}$$

Polar factorization in the reverse order is also possible

* Present address: Institut de Physique Nucléaire, Faculté des Sciences, 91 Orsay, France.

[1] For notation, see A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1962), Vol. II.

[2] E. P. Wigner, J. Math. Phys. 1, 409 (1960).

[3] F. Herbut and M. Vujičić, Antilinear Operators in Hartree–Bogoliubov Theory (1967) (to be published.).

[4] We use the word algebra here in the same sense as it is used in the term linear algebra, and not as a closed algebraic structure.

[5] A. I. Mal'cev, *Basic Linear Algebra* (State Technical Press, Moscow, 1956).

with the same anti-unitary factor:

$$\hat{A}_a = \hat{U}_a \hat{H}_2. \tag{1'}$$

First, we give a proof for (1'). We observe that $\hat{H}_2^2$ is uniquely determined as $\hat{A}_a^\dagger \hat{A}_a$, which is a linear Hermitian, positive, semidefinite operator, i.e., one with nonnegative eigenvalues. It is known that there exists a unique operator of the same kind which is its square root,

$$\hat{H}_2 = (\hat{A}_a^\dagger \hat{A}_a)^{\frac{1}{2}}. \tag{2}$$

It is easy to see that $\hat{H}_2$ and $\hat{A}_a$ are isometric, i.e.,

$$\|\hat{H}_2 |x\rangle\| = \|\hat{A}_a |x\rangle\|, \quad |x\rangle \in V_n,$$

where $\| |x\rangle\| = (\langle x | x\rangle)^{\frac{1}{2}}$ is the norm of a vector. Isometry of a linear and an antilinear operator always implies the existence of an anti-unitary operator which connects them in the sense of Eq. (1'). To see this, one has to notice that isometry requires coincidence of the zero-eigensubspaces (null spaces) of $\hat{A}_a$ and $\hat{H}_2$, i.e., $V(\hat{A}_a = 0) = V(\hat{H}_2 = 0)$, and, consequently, $\hat{A}_a$ and $\hat{H}_2$ define the same inverse classes (we call an inverse class the set of all those elements which a linear or antilinear operator maps into the same image). The set of inverse classes is a linear space isomorphic to the subspace of images (range) of $\hat{H}_2$, $R(\hat{H}_2)$, and anti-isomorphic to the range of $\hat{A}_a$, $R(\hat{A}_a)$. As a consequence, we have a unique anti-isomorphism, $\hat{U}_a^{(1)}$, mapping $R(\hat{H}_2)$ onto $R(\hat{A}_a)$: if $|a\rangle \in R(\hat{H}_2)$ then there is a vector $|b\rangle$ such that $|a\rangle = \hat{H}_2 |b\rangle$, and by definition $\hat{U}_a^{(1)} |a\rangle = \hat{A}_a |b\rangle$. The isometry between $\hat{A}_a$ and $\hat{H}_2$ has the further consequence that $\hat{U}_a^{(1)}$ preserves the norm, i.e., it is a unitary anti-isomorphism.

If $\hat{A}_a$ is nonsingular, then $R(\hat{A}_a) = R(\hat{H}_2) = V_n$, and $\hat{U}_a$ of Eq. (1') is unique and equal to $\hat{U}_a^{(1)}$. If $\hat{A}_a$ is singular, then we take an arbitrary unitary anti-isomorphism, $\hat{U}_a^{(2)}$, mapping the orthogonal complement of $R(\hat{H}_2)$ [i.e., $V(\hat{H}_2 = 0)$] onto that of $R(\hat{A}_a)$, and define $\hat{U}_a$ so that $\hat{U}_a |x\rangle = \hat{U}_a^{(1)} |x^{(1)}\rangle + \hat{U}_a^{(2)} |x^{(2)}\rangle$, where $|x\rangle \in V_n$, and $|x^{(1)}\rangle$ and $|x^{(2)}\rangle$ are its components in $R(\hat{H}_2)$ and in $V(\hat{H}_2 = 0)$, respectively. Actually, in Eq. (1') $\hat{U}_a^{(2)}$ does not act at all, and one could write Eq. (1') as $\hat{A}_a = \hat{U}_a^{(1)} \hat{H}_2$. Though $\hat{U}_a^{(2)}$ is not determined by $\hat{A}_a$, it is introduced in order to obtain $\hat{U}_a$ defined in the whole space.

Having proved (1'), we now turn to Eq. (1) by writing (1') in the form $\hat{A}_a = (\hat{U}_a \hat{H}_2 \hat{U}_a^\dagger)\hat{U}_a$ and observing that $\hat{U}_a \hat{H}_2 \hat{U}_a^\dagger$ is a linear Hermitian, positive, semidefinite operator which we denote by $\hat{H}_1$, i.e.,

$$\hat{H}_1 = \hat{U}_a \hat{H}_2 \hat{U}_a^\dagger. \tag{3}$$

Though, in general, $\hat{U}_a$ is not unique, $\hat{H}_1$ always is, because a more detailed analysis could show that $\hat{U}_a^{(2)}$ actually gives no contribution in Eq. (3). Anyway, the uniqueness of $\hat{H}_1$ is obvious from the fact that

$$\hat{H}_1^2 = \hat{A}_a \hat{A}_a^\dagger, \tag{4}$$

which follows from Eq. (1).

In order to clarify the relation of $\hat{U}_a^{(1)}$ and $\hat{U}_a^{(2)}$ to $\hat{H}_1$, we notice that $R(\hat{A}_a)$ and $R(\hat{H}_1)$ coincide, which means that $\hat{U}_a^{(1)}$ maps $R(\hat{H}_2)$ onto $R(\hat{H}_1)$, and $\hat{U}_a^{(2)}$ maps $V(\hat{H}_2 = 0)$ onto $V(\hat{H}_1 = 0)$.

We have discussed the factorizations (1) and (1') in the whole space. For application below, let us point out that if a subspace and its orthogonal complement are invariant for $\hat{A}_a$, then they are also invariant for $\hat{A}_a^\dagger$. This follows from the fact that if a subspace is invariant for a linear or antilinear operator, then its orthogonal complement is invariant for the adjoint operator.

We have an immediate corollary when $\hat{A}_a$ commutes with a linear Hermitian operator $\hat{H}$ and thus reduces in its eigensubspaces, i.e., when each eigensubspace of $\hat{H}$ is invariant for $\hat{A}_a$. In this case the polar factorizations (1) and (1') can be performed separately in each eigensubspace, and then also the polar factors commute with $\hat{H}$.

At last, we want to point out that having reduced a complex object like an AO to the product of two simple and well-understood operators, one is able to analyze the complexity of the AO through the relation that exists between the polar factors. Another tool which we make use of in our analysis below is the representation of AO's by antilinear matrices.

II. REPRESENTATION OF ANTILINEAR OPERATORS BY ANTILINEAR MATRICES

Now we are going to use the standard factorization[1] of an AO, $\hat{A}_a$, into a linear and a conjugation operator to obtain a representation of $\hat{A}_a$ in the space of column vectors. For that purpose we write

$$\hat{A}_a = (\hat{A}_a \hat{K}_a^{(Q)}) \hat{K}_a^{(Q)}, \tag{5}$$

where $\hat{K}_a^{(Q)}$ is defined as that AO for which all the vectors of a given basis (Q) are invariant. $\hat{K}_a^{(Q)}$ is obviously an involution, i.e., $(\hat{K}_a^{(Q)})^2 = 1$. If (Q) is an orthonormal basis, $\hat{K}_a^{(Q)}$ is anti-unitary.

The reverse factorization is also possible:

$$\hat{A}_a = \hat{K}_a^{(Q)} (\hat{K}_a^{(Q)} \hat{A}_a). \tag{5'}$$

The factorization in Eq. (5) is basis-dependent in the sense that the conjugation operator $\hat{K}_a^{(Q)}$ is defined by the basis (Q). In this basis the linear factor $\hat{A}_a \hat{K}_a^{(Q)}$ is represented by the matrix $(\hat{A}_a \hat{K}_a^{(Q)})_Q$. The antilinear factor $\hat{K}_a^{(Q)}$ is represented by an operation K which consists in complex conjugating all matrices to the right in matrix multiplications.

$\hat{A}_a$ itself is represented by $(\hat{A}_a\hat{K}_a^{(Q)})_Q K$, which is more complicated than an ordinary matrix, because K, though a very common operator in the space of column vectors, is not expressible as a matrix. We refer to such products of a matrix factor and conjugation as *antilinear matrices*. Since K is one and the same for all AO's, every AO is essentially represented by its matrix factor. The latter can be put in polar form,[5] so that its Hermitian and its unitary factor (the latter together with K) represent the corresponding polar factors of $\hat{A}_a$.

The set of all linear and antilinear operators in the abstract space is represented by the set of all ordinary and antilinear matrices in the space of column vectors, when a basis (Q) is specified. One should notice the following peculiarities when dealing with such an extended set of operators: Multiplication is always defined and the product belongs to the set; the sum, though always defined, belongs to the set if and only if both terms are of the same kind, i.e., linear or antilinear. The direct or Kronecker product, however, is not uniquely defined unless all factors are of the same kind. Namely, the direct product of an AO, $\hat{A}_a$, and of a linear operator, $\hat{B}$, $\hat{A}_a \otimes \hat{B}$, acting on a vector $|x\rangle \otimes |y\rangle = 1/\rho\,|x\rangle \otimes \rho\,|y\rangle$, where $\rho = |\rho|\, e^{i\varphi/2}$ is any complex number other than zero, would give $\hat{A}_a\,|x\rangle \otimes \hat{B}|y\rangle$, and also $e^{i\varphi}(\hat{A}_a\,|x\rangle \otimes \hat{B}|y\rangle)$, which shows that the image has a completely arbitrary phase.

In the rest of this paper, we consider representations of AO's only in orthonormal bases; therefore, we now discuss a few specific properties of antilinear matrices in connection with such bases. Hereafter, every basis, unless otherwise stated, is assumed to be orthonormal.

The elements of the matrix $(\hat{A}_a\hat{K}_a^{(Q)})_Q$ are obtained by the following formula:

$$\langle m|\,(\hat{A}_a\hat{K}_a^{(Q)})\,|n\rangle = \langle m|\,(\hat{A}_a\,|n\rangle), \tag{6}$$

where $|m\rangle$, $|n\rangle \in (Q)$. This is due to $\hat{K}_a^{(Q)}\,|n\rangle = |n\rangle$. These matrix elements are obtained from $\hat{A}_a$ in the same way as those of a matrix representing a linear operator.

In transition from one basis to another, the whole antilinear matrix transforms by a unitary similarity transformation:

$$\begin{aligned} S(\hat{A}_a\hat{K}_a^{(Q)})_Q KS^{-1} &= S(\hat{A}_a\hat{K}_a^{(Q)})_Q KS^{\dagger} \\ &= S(\hat{A}_a\hat{K}_a^{(Q)})_Q \tilde{S}K, \end{aligned}$$

where S is the matrix of the transition operator from the new basis to the basis (Q), and $\tilde{S}$ is its transpose. The matrix factor itself transforms by a unitary congruence transformation

$$S(\hat{A}_a\hat{K}_a^{(Q)})_Q\tilde{S}. \tag{7}$$

The matrix representing a linear operator transforms as a mixed second-rank tensor in arbitrary bases, whereas the matrix factor of an AO transforms as a twice contravariant tensor only in orthonormal bases.

There is no need to investigate the transformation properties of K because it is the same operation in the space of column vectors for any choice of a basis in the abstract space.

AO's can also be represented by antilinear matrices in the space of bras. For this purpose Eq. (5′) is more convenient.

We mention a few most important examples of antilinear matrices. If two AO's are mutually adjoint, their matrix factors are mutually transposed. If an AO is nonsingular, then the matrix factor of its inverse operator is the conjugate inverse of the matrix factor of the AO. The matrix factors of Hermitian and skew-Hermitian AO's are symmetric and skew-symmetric, respectively, unlike the representing matrices of the linear operators of the same kind. It should be noted that Hermitian and skew-Hermitian AO's give a geometrical, i.e., basis-independent meaning to symmetric and skew-symmetric matrices. The matrix factor of an anti-unitary operator is unitary, like the matrix of a linear unitary operator.

Dealing with antilinear matrices is almost as simple as with ordinary matrices, because all the difference is due to the unique operation K, which must be taken into account. In the following, we find representation of AO's by antilinear matrices as useful as matrix representation of linear operators usually is.

III. CANONICAL FORM OF ANTILINEAR OPERATORS

One of the most important problems in linear algebra is to find the spectral form of linear operators. Essentially it means replacing the operator by the direct sum of constants. This can be done for normal operators, i.e., for those which commute with their adjoints, and only for them. An alternative and more convenient approach to this problem is to diagonalize the representing matrix by a suitable unitary similarity transformation. Most important applications of linear algebra in physics are based on this reduction to the simplest, i.e., canonical, form.

It is very likely that the canonical form of AO's will also play an important role in applications. A few examples to support this are given in Sec. V and in subsequent papers.

An AO has *the canonical form*, i.e., acts in the simplest way on a basis, if and only if the representing antilinear matrix in this basis has the simplest form. We call such a basis *canonical*.

Since the conjugation K is always the same, the simplicity of an antilinear matrix depends only on its matrix factor. Therefore, the problem amounts to finding a suitable unitary congruence transformation which will make the matrix factor canonical.

We want to stress that we are looking for the simplest form of the matrix factor with respect to a limited class of bases, namely, the orthonormal ones. Taking into account the larger class of all bases, i.e., the transformations $S(\hat{A}_a\hat{K}_a^{(Q)})_Q(S^{-1})^*$ (S arbitrary, nonsingular), one would obtain for some AO's a simpler form of the matrix factor, but that is beyond the scope of this paper.

It seems natural to take the following *three criteria for the canonical form of the matrix factor*: firstly, it should be quasi-diagonal with as small submatrices as possible (the submatrices on the diagonal correspond to invariant subspaces for the AO); secondly, within these nondiagonal submatrices, the largest possible number of zeros should appear; thirdly, among the nonzero elements, one should have as many positive numbers as possible.

We are going to approach the problem of finding the canonical form of an AO by using the canonical forms of its polar factors. In doing this, two steps have to be distinguished: first, finding the simplest form of the factors, and secondly, the analysis of their mutual relationship, which decisively affects the canonical form of the AO itself.

The simplest form of the linear Hermitian polar factor is, of course, the diagonal form. As to the anti-unitary polar factor, its normal form has been given by Wigner.[2]

To express Wigner's result by an antilinear matrix, we use Eq. (6) and obtain that the matrix factor in the normal form is quasi-diagonal having on its diagonal a unit submatrix (corresponding to the subspace where the anti-unitary operator $\hat{U}_a$ is an involution) and two-by-two submatrices of the form

$$\begin{pmatrix} 0 & (u^*)^{\frac{1}{2}} \\ u^{\frac{1}{2}} & 0 \end{pmatrix}, \tag{8}$$

corresponding to the subbases $|u^*\rangle$, $|u\rangle$, which consist of eigenvectors of $\hat{U}_a^2$ with the eigenvalues u^*, u, always appearing in pairs. When $u = -1$, then, by convention, $(-1^*)^{\frac{1}{2}} = -i$.

Merely changing the phases of some vectors of the basis giving the above matrix factor, one can achieve the canonical form of $\hat{U}_a$, which instead of (8) has

$$\begin{pmatrix} 0 & 1 \\ u & 0 \end{pmatrix}, \tag{9}$$

the unit submatrix being unchanged.[6] This form we call *the Wigner canonical form of anti-unitary operators.*

Wigner's method is based on the diagonalization of $\hat{U}_a^2$ prior to the selection of the canonical basis for $\hat{U}_a$. Among the eigenbases of $\hat{U}_a^2$ there is none which could give a simpler form than Wigner's canonical one. Namely, in any eigenbasis of $\hat{U}_a^2$ only those vectors which correspond to the eigenvalue 1 may give nonzero diagonal elements in the matrix factor of $\hat{U}_a$, as follows from

$$\langle u|\,(\hat{U}_a\,|u\rangle) = u^*\,\langle u|\,(\hat{U}_a\,|u\rangle).$$

Therefore, wherever $u \neq 1$, one must have nothing but zeros on the diagonal.

It remains to be shown that only eigenbases of $\hat{U}_a^2$ can be canonical for $\hat{U}_a$. To that purpose we derive the Wigner canonical form using the above general criteria.

By the first criterion, the canonical form must be diagonal wherever possible. Changing the phases of the basis vectors and using antilinearity, the diagonal elements can be made positive, and because of unitarity they all have to be 1. Thus, the corresponding basis elements are eigenvectors of $\hat{U}_a^2$ with the eigenvalue 1. So, two-by-two nondiagonal submatrices are allowed only in the orthogonal complement of the subspace where $\hat{U}_a$ is an involution. Since these submatrices are unitary, they can have at most two zeros and both of them must be on the diagonal. If we required both nonzero elements to be positive, they would both be 1, which would imply that the corresponding vectors are still in the subspace where $\hat{U}_a^2 = 1$. Therefore, the two-by-two submatrices in the canonical form cannot be simpler than (9). Squaring the antilinear matrix so obtained, one concludes that $\hat{U}_a^2$ is diagonal, i.e., the canonical basis is necessarily an eigenbasis of $\hat{U}_a^2$.

To see how one should select a canonical basis for $\hat{U}_a$ out of the eigenbases of $\hat{U}_a^2$, we briefly describe, following Wigner,[2] a procedure consisting of three parts.

(1) In the subspace where $\hat{U}_a$ is an involution, one takes an arbitrary normalized vector, and adds to it its image by $\hat{U}_a$ with subsequent normalization, unless the sum is zero, when one takes the vector itself multiplied by i. If the subspace is more than one dimensional in the orthogonal complement of the

[6] One can also achieve the form having

$$\begin{pmatrix} 0 & u^* \\ 1 & 0 \end{pmatrix}$$

instead of (8), which is as simple as (9). We choose (9) because for $u = -1$, it is the known canonical form of skew-symmetric unitary matrices under unitary congruence transformations (cf. Sec. VA).

first-basis vector so obtained, the procedure is repeated, etc.

(2) In the subspace where $\hat{U}_a^2 = -1$, i.e., where $\hat{U}_a$ is a skew involution, an arbitrary normalized vector is taken and it is paired with its image by $\hat{U}_a$ in reversed order. If the subspace is more than two dimensional, another normalized vector orthogonal to both previous ones is taken, and the above procedure is repeated, etc. This subspace is necessarily even dimensional.

(3) When $u \neq 1, -1$, a basis is chosen in the eigensubspace of $\hat{U}_a^2$ with the eigenvalue u, and each element of it is paired with its image by $\hat{U}_a$, which necessarily belongs to the eigensubspace with the eigenvalue u^*. These pairs in the order $|u^*\rangle$, $|u\rangle$ give the submatrices (9). Though u and u^* play essentially symmetrical roles, we have only one of them in the submatrices, depending in whose eigensubspace we make the arbitrary choice of an orthonormal basis.

It should be noted that the choice of a canonical basis for $\hat{U}_a$ is not unique. The group of transformations connecting one canonical basis with all the others has been given by Wigner.[2] Every canonical basis gives one- and two-dimensional invariant subspaces for $\hat{U}_a$, which are also nonunique. The only unique invariant subspaces for $\hat{U}_a$, made use of in the above procedure, are the eigensubspaces of $\hat{U}_a^2$ with the eigenvalues 1 and -1, and the direct sums of the eigensubspaces with the eigenvalues u and u^*, when $u \neq 1, -1$.

The Wigner canonical form itself is unique except for the order of the submatrices on the diagonal and conjugation within each submatrix.

Having discussed the Wigner canonical form of anti-unitary operators in sufficient detail, we may turn now to the problem of finding the canonical form of more general AO's. Using *the method of simultaneously canonical polar factors*, we show that this is possible for AO's whose polar factors are simply related to each other.

The simplest mutual relation of the polar factors is their commutation. It is obviously equivalent to $\hat{H}_1 = \hat{H}_2$, as well as to $[\hat{A}_a, \hat{A}_a^\dagger]_- = 0$ [see Eqs. (1), (1′), (2), and (4)]. The class of AO's satisfying this relation we call *proper normal AO's.*

From the commutation of the polar factors of a proper normal AO it follows that every eigensubspace of the Hermitian factor is invariant for the anti-unitary one. In each eigensubspace, one may find a Wigner canonical basis independently. All these together form a canonical basis in the whole space. In this way, *for a given proper normal AO, a canonical basis is that one which is simultaneously an eigenbasis for the Hermitian factor and a Wigner canonical basis for the anti-unitary factor.*

The canonical matrix factor of a proper normal AO is in general the direct sum of a diagonal matrix and of several two-by-two matrices. The diagonal submatrix corresponds to the subspace, where $\hat{U}_a$ is an involution and has only nonnegative diagonal elements, which are eigenvalues of the Hermitian factor. The two-by-two submatrices are of the form

$$\begin{pmatrix} 0 & h \\ hu & 0 \end{pmatrix}, \tag{10}$$

where h is an eigenvalue of the Hermitian factor, and $u \neq 1$ is an eigenvalue of $\hat{U}_a^2$.

So far we have not yet proved that this form is indeed canonical for proper normal AO's. This proof is given at the end of Sec. IV for a class of AO's which contains the proper normal ones.

To the proper normal AO's ($\hat{H}_1 = \hat{H}_2$) belong as special cases, the anti-unitary operators (for which $\hat{H}_1$ is the identity operator), the Hermitian AO's (whose $\hat{U}_a^{(1)}$ is an involution) and the skew-Hermitian AO's ($\hat{U}_a^{(1)}$ skew-involution).

For all proper normal AO's, $\hat{U}_a$ is the direct sum of $\hat{U}_a^{(1)}$, which acts in $R(\hat{H}_1)$, and $\hat{U}_a^{(2)}$ acting in $V(\hat{H}_1 = 0)$. It is convenient to choose $\hat{U}_a^{(2)}$ always to be an involution (see Sec. IV).

We have defined the proper normal AO's as those which commute with their adjoints. In this respect they are the analogs of the linear normal operators. But the proper normal AO's can also be defined by their canonical form, and regarding this form they do not correspond to the diagonalizable linear operators. Because we consider the nature of the canonical form as the basic property of a class of operators, we now search for the analog of the linear normal ones.

In linear algebra the diagonal canonical form defines the normal linear operators. Since their diagonal elements are complex in general, it is possible to obtain the most important subclasses of linear normal operators (Hermitian, skew-Hermitian and unitary ones) by restricting the diagonal elements to be real, pure imaginary, or of unit modulus, respectively. In contrast to this, in the case of AO's, only Hermitian ones can be diagonalized (as follows from $\hat{U}_a^2 = 1$). Furthermore, their canonical form consists of nonnegative elements exclusively, and it cannot be further restricted. This form we call *the first-order canonical* one. Since the corresponding class of AO's is too narrow, one should not consider the Hermitian AO's as the analogs of the normal linear operators, nor the first-order canonical form as the analog of the diagonal one for linear operators.

As a starting point in finding this analog, one should take the Wigner canonical form of anti-unitary operators, because the polar factorization reveals their fundamental role. With respect to the three criteria for the canonical form, the Wigner one exhibits the following properties.

It is first-order canonical wherever possible. The nondiagonal part consists merely of two-by-two submatrices. Each of these has both diagonal elements equal to zero, and at least one off-diagonal element is positive [cf. submatrix (9)].

Beside these, the Wigner canonical form has two more properties, which are further restrictions due to the specific nature of anti-unitary operators and are not imposed by the three criteria.

All the diagonal elements in the first-order canonical part are equal to 1, and so are the off-diagonal elements in the first rows of the two-by-two submatrices; the off-diagonal elements in the second rows u are of unit modulus, which cannot take the value 1.

Abolishing only the latter two specific restrictions, i.e., allowing any nonnegative numbers instead of the 1's and any complex numbers instead of the u's, we define *the second-order canonical form* as the following quasi-diagonal matrix:

$$\begin{pmatrix} h_1 & & & & & & \\ & \ddots & & & & & \\ & & h_j & & & & \\ & & & 0 & h_1' & & \\ & & & h_1''u_1 & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \quad h_k' \\ & & & & & & h_k''u_k \quad 0 \end{pmatrix}, \tag{11}$$

where $j + 2k = n$, all $|u| = 1$, and all h, h', $h'' \geq 0$. Notice that if any of the h' equals zero, then the corresponding $h''u > 0$. This is required by criterion 3 and can always be achieved. The matrix elements are written in polar form to anticipate their origin as the products of eigenvalues of $\hat{H}_1$ or $\hat{H}_2$ and $\hat{U}_a^2$.

Since (11) is not more complicated, according to the criteria, than the Wigner canonical form, and, on the other hand, contains the canonical forms of all proper normal AO's as special cases, we consider it as the natural analog of the diagonal form of linear normal operators.[7]

We call *normal* those AO's which have the second-order canonical form as their simplest form. We study their properties in the next section.

IV. NORMAL ANTILINEAR OPERATORS

Having defined the normal AO's by their canonical form, we now search for an alternative definition in terms of the polar factors.

The Hermitian polar factor $\hat{H}_1$ of (11) is uniquely determined and turns out to be diagonal. If (11) is singular, the unitary polar factor is not unique. We use this nonuniqueness in order to make it as simple as possible. To that purpose, we take every h, h', or h'' which is zero with a phase factor 1. In this way one immediately obtains the unitary factor in a form which may differ from the Wigner canonical one only by having submatrices (9) with $u = 1$ as well. This generalization of the Wigner canonical form, which we call the second-order canonical form of an anti-unitary involution, is a consequence of our definition of the second-order canonical form in general, in which we have given up all accidental restrictions not contained in the three criteria. This is justified below, where we show that the normal AO's can be defined in a very simple way.

Referring to the above slight generalization of the Wigner form as to the canonical form of $\hat{U}_a$, we say that *the class of normal AO's is the widest class of AO's whose polar factors can be put simultaneously in canonical form.*

We now derive the necessary conditions for the polar factors of a normal AO.

Since

$$\begin{pmatrix} 0 & h' \\ h''u & 0 \end{pmatrix} = \begin{pmatrix} h' & 0 \\ 0 & h'' \end{pmatrix}\begin{pmatrix} 0 & 1 \\ u & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ u & 0 \end{pmatrix}\begin{pmatrix} h'' & 0 \\ 0 & h' \end{pmatrix},$$

H_2 is diagonal in the same basis, and so

$$[\hat{H}_1, \hat{H}_2]_- = 0. \tag{12}$$

Similarly,

$$[\hat{H}_1, \hat{U}_a^2]_- = 0, \tag{13}$$

$$[\hat{H}_2, \hat{U}_a^2]_- = 0. \tag{14}$$

As a consequence, besides Eq. (3) we also have

$$\hat{H}_2 = \hat{U}_a \hat{H}_1 \hat{U}_a^\dagger. \tag{15}$$

The matrix representing $\hat{A}_a^2$ is diagonal and to the

[7] We went beyond the canonical form of the proper normal AO's because, due to the commutation of the polar factors, it also has two specific restrictions with respect to the second-order canonical form. Namely, in the submatrices (10) the two eigenvalues of the Hermitian polar factor coincide, and $u \neq 1$.

two-by-two submatrices of (11) corresponds

$$\begin{pmatrix} h'h''u^* & 0 \\ 0 & h'h''u \end{pmatrix}.$$

Obviously, it follows that the polar factors[5] $\hat{H}$ and $\hat{U}$ of $\hat{A}_a^2$ satisfy $\hat{H} = \hat{H}_1\hat{H}_2$, $\hat{U} = \hat{U}_a^2$, and

$$[\hat{H}, \hat{U}]_- = 0, \tag{16}$$

i.e., $\hat{A}_a^2$ is a linear normal operator.

At last, we have

$$[\hat{H}_1\hat{H}_2, \hat{U}_a]_- = 0, \tag{17}$$

which follows from Eqs. (3), (15), and (12).

Now we show that the conditions (12)–(15) are also sufficient for an AO to be normal.

In order to obtain the Wigner canonical form of an anti-unitary operator $\hat{U}_a$, one first diagonalizes $\hat{U}_a^2$. In case of a proper normal AO, $\hat{A}_a = \hat{H}_1\hat{U}_a$, $\hat{H}_1 = \hat{H}_2$, one diagonalizes $\hat{H}_1$ and $\hat{U}_a^2$ as a first step. Now we begin by diagonalizing $\hat{H}_1$, $\hat{H}_2$, and $\hat{U}_a^2$, simultaneously, which is possible because they all commute [see Eqs. (12), (13), and (14)]. Actually, we need only the common eigensubspaces of these operators, in which we proceed to find the vectors of a basis canonical for $\hat{U}_a$ as well.

To that purpose we break up the whole space into two mutually orthogonal subspaces. The first is the direct sum of all those common eigensubspaces where the eigenvalues of $\hat{H}_1$ and $\hat{H}_2$ coincide, i.e., where $\hat{A}_a$ is proper normal. The second is its orthogonal complement, and we say that there $\hat{A}_a$ is *improper normal.* In the first subspace, the procedure is the same as described in the previous section. To obtain a procedure in the second subspace, it is important to notice that, besides the fact that $\hat{U}_a$ of any AO maps an eigenvector of $\hat{H}_2$ into one of $\hat{H}_1$ with the same eigenvalue [which follows from Eq. (3)] for those AO's which satisfy Eq. (15), it is also true the other way round, i.e., $\hat{H}_1|x\rangle = h'|x\rangle$ implies $\hat{H}_2(\hat{U}_a|x\rangle) = h'(\hat{U}_a|x\rangle)$. Thus, a common eigenvector of $\hat{H}_1$ and $\hat{H}_2$ with the eigenvalues h' and h'', respectively, is taken by $\hat{U}_a$ into another common eigenvector of the same operators, now corresponding to the eigenvalues h'' and h', respectively. This has the consequence that to each common eigensubspace corresponds another with exchanged eigenvalues of $\hat{H}_1$ and $\hat{H}_2$ and with the conjugate eigenvalue of $\hat{U}_a^2$. Now we choose an arbitrary basis in one of them, e.g. $|h'h''u\rangle_i$, $i = 1, 2, \ldots, m$, and in the other we take that basis which is the image by $\hat{U}_a$ of the first one, i.e., $|h''h'u^*\rangle_i$, $i = 1, 2, \ldots, m$.

It is easy to see that in the basis obtained in this way, when its elements are arranged so that each of them is next to its image by $\hat{U}_a$, e.g. $|h''h'u^*\rangle_i$, $h'h''u\rangle_i$, the matrix factor of $\hat{A}_a$ has the desired canonical form.

It should be observed that our procedure for the improper normal part of $\hat{A}_a$ is fully analogous to Wigner's procedure[2] for $u \neq 1, -1$, though here both $u = 1$ and $u = -1$ may occur. Besides, in this part of the space the diagonal form of $\hat{A}_a$ is never achieved, not even where $u = 1$.

It is of interest to find among the above four necessary and sufficient conditions [Eqs. (12)–(15)] the smallest number of independent ones.

One immediately obtains (14) and (15) from (12) and (13). To show the mutual independence of the latter two, one may take two simple examples of antilinear matrices, e.g.,

$$A_a = \begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & h' \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}K,$$

$$A_a' = 2^{-\frac{1}{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} h & 0 \\ 0 & h' \end{pmatrix}K,$$

where $h \neq h'$ in both cases, A_a satisfies (12) and not (13), and A_a' vice versa.

Therefore, one can say that *the polar factors of an AO can be brought simultaneously into the canonical forms if and only if they satisfy* (12) *and* (13). We have thus obtained a second definition of normal AO's.

It is important to replace the above two relations by conditions on $\hat{A}_a$ itself, because one would like to recognize whether a given AO is normal or not without having first to calculate its polar factors.

We are now going to show that normal AO's can be defined as those which satisfy only one relation,

$$[\hat{A}_a, (\hat{A}_a^\dagger)^2]_- = 0. \tag{18}$$

This condition can be easily obtained from Eqs. (12) and (13) if we write $\hat{A}_a$ in (18) in polar form. Actually, we use Eqs. (12) and the adjoint of (13), together with (17).

In order to show that (12) follows from (18), we take the latter and its adjoint to obtain an equation which is equivalent to the former,

$$[\hat{A}_a\hat{A}_a^\dagger, \hat{A}_a^\dagger\hat{A}_a]_- = [\hat{H}_1^2, \hat{H}_2^2]_- = 0.$$

To derive Eq. (13) from Eq. (18) we use the following argument. We first show that

$$[\hat{H}_1, \hat{A}_a^2]_- = 0, \tag{19}$$

and that $\hat{A}_a^2$ is a linear normal operator. Therefore, (13) can be derived separately in the range of $\hat{A}_a^2$ and separately in its zero manifold. In the former $\hat{U}_a^2$ is the

unitary polar factor of $\hat{A}_a^2$ and, as a consequence of Eq. (19), it commutes with $\hat{H}_1$. In the zero manifold of $\hat{A}_a^2$, we are able to achieve $\hat{U}_a^2 = 1$ which will obviously commute with $\hat{H}_1$.

To prove Eq. (19), we observe that its equivalent $[\hat{H}_1^2, \hat{A}_a^2]_- = 0$ follows immediately from Eq. (4) and the adjoint of Eq. (18). The linear operator $\hat{A}_a^2$ is normal if $[\hat{A}_a^2, (\hat{A}_a^\dagger)^2]_- = 0$, and this is a corollary of Eq. (18). Consequently, the range and the zero manifold of $\hat{A}_a^2$ coincide with those of its Hermitian polar factor $\hat{H}$, i.e., with $R(\hat{H})$ and $V(\hat{H} = 0)$, respectively. Using Eqs. (18) and (12) we derive $\hat{H}^2 = \hat{A}_a^2(\hat{A}_a^2)^\dagger = \hat{A}_a\hat{A}_a^\dagger\hat{A}_a^\dagger\hat{A}_a = \hat{H}_1^2\hat{H}_2^2 = (\hat{H}_1\hat{H}_2)^2$, i.e.,

$$\hat{H} = \hat{H}_1\hat{H}_2 = \hat{H}_2\hat{H}_1. \tag{20}$$

Furthermore, Eq. (18) implies that $\hat{A}_a$ commutes with $\hat{H}^2$, thus

$$[\hat{A}_a, \hat{H}]_- = 0. \tag{21}$$

From the argument near the end of Sec. I, we also have

$$[\hat{U}_a, \hat{H}]_- = 0. \tag{22}$$

From Eqs. (19) and (22) we conclude that $R(\hat{H})$ and $V(\hat{H} = 0)$ are invariant for both $\hat{H}_1$ and $\hat{U}_a^2$.

In $R(\hat{H})$ the unitary polar factor of $\hat{A}_a^2$ is unique and can be obtained with the help of Eqs. (20) and (22), $\hat{A}_a^2 = (\hat{U}_a\hat{H}_2)(\hat{H}_1\hat{U}_a) = \hat{H}\hat{U}_a^2$, i.e.,

$$\hat{U}_a^2 = \hat{H}^{-1}\hat{A}_a^2. \tag{23}$$

Since $\hat{H}_1$ commutes with both factors in Eq. (23), Eq. (13) is valid in $R(\hat{H})$.

Turning now to $V(\hat{H} = 0)$, we notice that the arbitrary $\hat{U}_a^{(2)}$ has both its domain $V(\hat{H}_2 = 0)$ and its range $V(\hat{H}_1 = 0)$ inside this subspace, because of $\hat{H} = \hat{H}_1\hat{H}_2$. We now restrict the arbitrariness of $\hat{U}_a^{(2)}$ so that $\hat{U}_a^2$ in $V(\hat{H} = 0)$ will become the identity operator. To this purpose we break up $V(\hat{H} = 0)$ into three mutually orthogonal subspaces,

$$V(\hat{H} = 0) = V(\hat{H}_1 = 0 \neq \hat{H}_2) + V(\hat{H}_1 \neq 0 = \hat{H}_2) + V(\hat{H}_1 = \hat{H}_2 = 0).$$

From Eq. (22) one may conclude that $\hat{U}_a^{(1)}$ also reduces in $V(\hat{H} = 0)$. Since it always maps $R(\hat{H}_2)$ onto $R(\hat{H}_1)$, here it maps $V(\hat{H}_1 = 0 \neq \hat{H}_2)$ onto $V(\hat{H} \neq 0 = \hat{H}_2)$. All we have to do now is to choose $\hat{U}_a^{(2)}$ to act as $(\hat{U}_a^{(1)})^{-1}$ in mapping $V(\hat{H}_1 \neq 0 = \hat{H}_2)$ onto

$$V(\hat{H}_1 = 0 \neq \hat{H}_2),$$

and separately in $V(\hat{H}_1 = \hat{H}_2 = 0)$ as an arbitrary involution. Thus, Eq. (13) is obtained in the whole space.

We want to point out that the above choice of $\hat{U}_a^{(2)}$ amounts to taking the undetermined phase factors equal to 1 when any one of the h, h', h'' is zero, because $V(\hat{H} = 0)$ corresponds exactly to the totality of these submatrices.

It is now established that *commutation of an AO with the square of its adjoint*, i.e., Eq. (18), *is a necessary and sufficient condition for the AO to be normal.*

Obtaining the second-order canonical form implies finding one- and two-dimensional invariant subspaces for a normal $\hat{A}_a$, and essentially $\hat{A}_a$ is replaced by the direct sum of its components in these subspaces. However, these subspaces are not unique. The unique ones are firstly, the common eigensubspaces of $\hat{H}_1$, $\hat{H}_2$ and $\hat{U}_a^2$, $V(h, h', u)$, with $h = h'$ and u real, and secondly, the direct sums $V(h, h', u) + V(h', h, u^*)$ in all other cases. We have based our procedure for selecting a canonical basis for $\hat{A}_a$ on breaking up the space into these unique, invariant, and mutually orthogonal subspaces, in analogy with Wigner's procedure[2] for anti-unitary operators.

At last, it remains only to be shown that the second-order canonical form of a normal $\hat{A}_a$, achieved through the canonical forms of its polar factors, is indeed its simplest possible form according to the general criteria stated in the previous section.

The diagonal submatrix of the second-order canonical form with nonnegative numbers corresponds to that invariant subspace where $\hat{A}_a$ is Hermitian, i.e., to

$$V_1 = \sum_{h \geq 0} V(h = h', u = 1).$$

This subspace is uniquely associated with $\hat{A}_a$ and any diagonal submatrix of $\hat{A}_a$ achieved by whatever method necessarily corresponds to a part of V_1.

Further, we have a number of nondiagonal two-by-two submatrices with three zeros in them. Their totality corresponds to

$$V_2 = \sum_{h > 0} [V(h, 0, u = 1) + V(0, h, u = 1)].$$

The nonzero elements are positive, and such matrices cannot be made simpler. One cannot achieve a larger number of such submatrices by some other method, because any of them has to correspond to a part of V_2, which is immediately seen when these submatrices are written in polar form.

The totality of the remaining nondiagonal two-by-two submatrices corresponds to the orthogonal complement of $V_1 + V_2$, where $\hat{A}_a$ is necessarily nonsingular, as it is clear from the fact that all determinants of these two-by-two submatrices are nonzero. Namely, although the determinant of a matrix is not invariant under the unitary congruence transformations, its being zero or not is an invariant

property, and it can be taken as a necessary and sufficient criterion for singularity or nonsingularity of $\hat{A}_a$ as it is the case for linear operators. Therefore, in this part of the space, two-by-two submatrices obtained by any method cannot have more than two zeros and they must be both on the diagonal. The polar factors of such matrices, which are immediately obtained by writing the matrix elements in polar form, are seen to be canonical, and so the nonzero elements cannot be made simpler.

This analysis shows that polar factorization, leading to the diagonalization of $\hat{H}_1$, $\hat{H}_2$, and $\hat{U}_a^2$, is the natural way to obtain the canonical form of a normal AO.

V. APPLICATIONS

We outline a few applications of AO's in different branches of physics, mathematics, and electric network theory. In most of them, polar factorization and the canonical form of normal AO's prove very useful tools.

A. Electric Network Application

According to Youla,[8] "The problem of finding a canonic form of an arbitrary matrix under the group of unitary congruence transformations is not only of mathematical interest but of the utmost importance for applied electrical engineering network theory . . . ."

We have already seen that one can associate a basis independent object with any matrix transforming under the unitary congruence transformations. This object is an AO in the abstract space V_n (see Sec. II). Therefore, AO's are the natural geometrical interpretation of such matrices. Since we have defined the canonical form of an AO as the canonical form of its matrix factor, all our results are directly valid for the matrices under unitary congruence transformations.

We have mentioned that to Hermitian and skew-Hermitian AO's correspond symmetric and skew-symmetric matrices, respectively. Our results coincide with the classical ones[9,10] on the canonical forms of these matrices. The results of this paper on the second-order canonical form give the solutions of the above-mentioned network theory problem for considerably larger classes of matrices.

The Wigner canonical form should be interpreted as the canonical form of a unitary matrix under unitary congruence transformations.

To proper normal AO's, correspond matrices satisfying the condition

$$AA^\dagger = (A^\dagger A)^*, \tag{24}$$

which defines the class of *conjugate-normal* matrices, as we may call them. It should be observed that this property is unitary congruence invariant in the sense that it is valid for any unitary congruence transform of A, in contrast to the property $AA^\dagger = A^\dagger A$, which is unitary similarity invariant. Unitary, symmetric and skew-symmetric matrices belong to the class of conjugate normal ones.

To normal AO's, correspond matrices for which the following relation is valid:

$$AA^\dagger\tilde{A} = \tilde{A}A^\dagger A, \tag{25}$$

and we may call them *congruence-normal matrices.* Conjugate-normal matrices belong to them as a special case. The class of congruence-normal matrices is the largest class of matrices for which we have found the canonical form, and they may be considered as the analogs of the class of normal matrices under unitary similarity transformations, because *relation* (25) *is a necessary and sufficient condition for a matrix to have the second-order canonical form under the unitary congruence transformations.* [See the end of Sec. III and Eq. (18).]

We intend to give a more detailed discussion of our results on the canonical form of congruence-normal matrices in matrix language, and applications to electric network theory in a separate paper later.

B. Connection of Antilinear Operators with Two-Particle Wave Vectors

A column vector representing a two-particle wave vector $|\psi^{(1,2)}\rangle$ in a basis which is obtained as a direct product of a one-particle basis with itself, can be written as a matrix which transforms as a twice contravariant tensor, i.e., under unitary congruence transformations. This enables us to establish a one-to-one correspondence between the unitary space of all AO's and the unitary space of all two-particle wave vectors. This correspondence turns out to be, what we call, an *extended isomorphism,* because it connects all the known operations in these two sets, including even those for which they are not closed. This extended isomorphism permits transfer of any problem from one set to the other, and consequently, its solution where it is more convenient. The canonical forms of two-fermion and two-boson wave vectors are thus readily obtained from the canonical forms of skew-Hermitian and Hermitian AO's, respectively.

This extended isomorphism, and the use of AO's in defining orthogonal and symplectic subgroups of the unitary group, as well as application of AO's to the problem of the canonical form of a two-particle interaction are described in a separate publication.[11]

[8] D. C. Youla, Can. J. Math. **13**, 694 (1961).
[9] I. Schur, Am. J. Math. **67**, 472 (1942).
[10] L. K. Hua, Am. J. Math. **66**, 470 (1944).
[11] F. Herbut and M. Vujičić (to be published).

C. Transposition and Conjugation of Linear Operators

We have analyzed two factorizations of AO's, the polar and the standard one (see Secs. I and II, respectively). In general, the standard factorization cannot be of any help in finding a canonical basis for a normal AO, because its antilinear factor is defined as an arbitrary involution, with no regard to the nature of the AO. On the contrary, in the polar factorization, both factors depend crucially on the AO. For Hermitian AO's and only for them, the polar factorization is a special case of the standard one. Now we discuss a widely used example of these operators: the adjoining of linear operators, in order to illustrate the relation between the two factorizations.

Let V_{n^2} be the space of all linear operators which act in V_n. The operation of adjoining, $\hat{O}_a$, in V_{n^2}, is obviously an antilinear operator and furthermore an involution. Unitary metrics in V_{n^2} can be defined by introducing the following scalar product[12]:

$$(\hat{A}, \hat{B}) = \operatorname{tr} \hat{A}^\dagger \hat{B}, \quad \text{for all} \quad \hat{A}, \hat{B} \in V_{n^2}. \tag{26}$$

Having introduced unitary metrics in V_{n^2}, one can easily see that $\hat{O}_a$ is anti-unitary. Being also an involution, it is a Hermitian AO. Its canonical basis is any set of n^2 orthonormal [in the sense of Eq. (26)] Hermitian linear operators acting in V_n.

It is interesting to note that the same metrics in V_{n^2} can be obtained by the requirement that the basis in V_{n^2}

$$|e_i\rangle\langle e_j|, \quad i,j = 1, \cdots, n, \tag{27}$$

associated with an orthonormal basis $|e_1\rangle, \cdots, |e_n\rangle$ in V_n should also be orthonormal. None of the associated bases can be canonical for $\hat{O}_a$, since $(|e_i\rangle\langle e_j|)^\dagger = |e_j\rangle\langle e_i|$ for $i \neq j$ are not self-adjoint.

However, the associated bases (27) (and no others) can be used to generalize the matrix concepts of transposition and conjugation to the corresponding operations for abstract linear operators. Namely, if we choose a basis in V_n, e.g., (q), and form its associated basis (Q) in V_{n^2}, then the standard factors [Eq. (5)] of $\hat{O}_a$ for the basis (Q), when represented in (Q), are the transposition and the conjugation of matrices which represent linear operators from V_{n^2} in the basis (q). Therefore, it is natural to interpret the standard factors of $\hat{O}_a$ themselves as transposition and conjugation of linear operators. Obviously, these concepts are basis-dependent, as it is always the case with the standard factorization.

D. A One-Particle Operator Approach to Hartree–Bogoliubov Theory

The results of this work have another interesting application in the variational theory of Hartree–Bogoliubov.[3] Namely, because of its transformation properties, the pairing tensor in this self-consistent theory can be interpreted as the matrix factor of a skew-Hermitian antilinear operator $\hat{t}_a$. We call the anti-unitary polar factor of $\hat{t}_a$ the *pairing operator*, and it is a generalization of the time-reversal operator as far as its role in this theory is concerned, because in the special case of the BCS theory, the anti-unitary polar factor of $\hat{t}_a$ is just the time-reversal operator. The Hermitian polar factor of $\hat{t}_a$ is a simple function of the one-particle density operator. So, the polar factors of $\hat{t}_a$ become the main variational objects of Hartree–Bogoliubov theory. Since the pairing operator is a skew-involution, the elements of its canonical basis display the well-known pairing property, which is inherent in every Bogoliubov–Valatin transformation.[13]

Our operator treatment of the kinematical and the dynamical quantities (self-consistent fields) should offer new possibilities for finding workable approximations within the theory of general linear canonical transformations. These problems are discussed in a separate paper.[3]

ACKNOWLEDGMENTS

The authors are very grateful to Professor E. P. Wigner for having read a preliminary version of this work, as well as for his stimulating comments and encouragement. Thanks are also due to Dr. Dj. Živanović for several discussions.

[12] J. von Neumann, Ann. Math. **41**, 94 (1940).

[13] J. G. Valatin, Phys. Rev. **122**, 1012 (1961).