Degenerate Representation Functions for $SO(\nu)$, $SU(\nu)$, and $SU(\nu) \otimes SU(\nu)$ and Their Analytical Reductions*

R. DELBOURGO, K. KOLLER, AND RUTH M. WILLIAMS
Physics Department, Imperial College, London, England

(Received 14 June 1968)

We have studied the problem of obtaining the rotation matrix elements $d^N(\theta) = \langle N| e^{-i\theta J_2} |N\rangle$, where $|N\rangle$ refer to a particular degenerate class of basis vectors of a symmetry group G which embraces the rotations $SU(2)_J$ as a subgroup. For $G = SO(\nu)$, $SU(\nu)$, and $SU(\nu) \otimes SU(\nu)$, we prove that these particular representation functions are proportional to the Gegenbauer polynomials $C_N^{\frac{1}{2}(\nu-2)}(\cos\theta)$, $C_N^{\frac{1}{2}(\nu-1)}(\cos\theta)$, and $C_N^{\frac{1}{2}\nu}(\cos\theta)$, respectively. The reduction of such functions into one another according to the formula $C_{N'}^{\lambda'} = \sum_N a_N C_N^{\lambda}$ has been solved in generality for complex values of N and corresponds to the reduction of Regge poles of G into Regge poles of one of its subgroups. The reduction formula for functions of the second type $E_{N'}^{\lambda'} = \sum_N b_N E_N^{\lambda}$ has also been derived; here one simply meets an infinite series.

1. INTRODUCTION

There are three main lessons to be drawn from the nonrelativistic Coulomb problem: (i) the special nature of the dynamics provides the higher symmetry group $SU(2) \otimes SU(2) \approx SO(4)$; (ii) the bound states fall into simple classes of representations of the group, namely of the type (N, N); (iii) the Reggeization procedure for the bound and scattering states is more easily discussed by working directly in the plane of the principal quantum number rather than returning to the conventional angular-momentum plane. In line with these three ideas, a scheme[1] has been recently developed for hadronic matter having the following ingredients: (i) Strong interaction dynamics is evidenced by the existence of an approximate supermultiplet group G which contains the rotations as an $SU(2)_J$ subgroup; (ii) the hadronic states belong to simple, degenerate classes of G representations; and (iii) Reggeization is carried out in the plane of the Casimir operator which labels the physical sequence of G multiplets. Reduction to the angular-momentum plane is necessary only for introducing simple symmetry-breaking corrections to G and for comparison with the results of standard Regge theory.

One of the crucial mathematical problems which is encountered in these Reggeized supermultiplet schemes is, for the degenerate series $|N\rangle$ of G-representations in question, the determination of the rotation matrix element

$$d^N(\theta) = \langle N| e^{-i\theta J_2} |N\rangle,$$

the analog of $d_{00}^J(\theta) = P_J(\cos\theta)$ for $SU(2)_J$. We prove in Sec. 2, for particular most-degenerate sequences of basis vectors (fully specified there), that, when G is $SO(\nu)$, $SU(\nu)$, and $SU(\nu) \otimes SU(\nu)$, the generalized rotation functions are the Gegenbauer polynomials

$$C_N^{\frac{1}{2}(\nu-2)}(\cos\theta), \quad C_N^{\frac{1}{2}(\nu-1)}(\cos\theta),$$

and

$$C_N^{\frac{1}{2}\nu}(\cos\theta),$$

respectively, appropriately normalized. A second important aspect of the Reggeization scheme concerns the reduction of the (above) series of functions pertaining to G into the functions belonging to various subgroups—for instance the reduction of C^1 into $C^{\frac{1}{2}} = P$ for the H-atom case—and this is fully treated in Sec. 3. Formula (29) summarizes the work and expresses the fact that (stated in terms of representations of the first kind) a Regge pole of a higher symmetry group decomposes into a series of Regge poles of a subgroup *plus* a specific background integral. An alternative decomposition (stated in terms of representations of the second kind) is given by formula (32).

2. DEGENERATE REPRESENTATION FUNCTIONS

Let us recall how one proves that the rotation functions $d_{00}^J(\theta) = \langle J0| e^{-i\theta J_2} |J0\rangle$ of $SU(2)$ are the Legendre polynomials. This will serve as a good introduction and guide to the subsequent determination of the rotation functions for the higher groups. There are at least two simple methods for finding the $d^J(\theta)$.

(a) The first utilizes the $SU(2)$ Lie algebra by expressing the commutators $[e^{-i\theta J_2}, \mathbf{J}]$ as commutators $[e^{-i\theta J_2}, J_3]$ and differentials of $e^{-i\theta J_2}$. Sandwiching $[e^{-i\theta J_2}, \mathbf{J}^2]$ between the states in question provides a differential equation whose regular solution is known

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research OAR through the European Office of Aerospace Research, United States Air Force.

[1] A. Salam and J. Strathdee, Phys. Rev. Letters **19**, 339 (1967); R. Delbourgo, A. Salam, and J. Strathdee, Phys. Rev. **172**, 1727 (1968); R. Delbourgo, J. Math. Phys. **9**, 1936 (1968); R. Delbourgo and M. A. Rashid, Phys. Rev. **176**, 2074 (1968).

to be the Legendre function. In fact this technique is no more difficult to apply for finding the full set of $d^J_{MM'}(\theta)$ functions in closed form.

(b) The second makes use of the basic tensor representation q_μ ($\mathbf{q}$ being a unit vector) and the fact that general $|j\rangle$ vectors correspond to forming symmetrized traceless tensor products

$$Q_{\mu_1\cdots\mu_J} = q_{\mu_1}\cdots q_{\mu_J} - \text{trace terms}.$$

The rotation functions come as scalar products

$$Q'_{\mu_1\cdots\mu_J}Q_{\mu_1\cdots\mu_J}, \quad \text{where} \quad q_\mu q'_\mu = \cos\theta.$$

That they are the Legendre polynomials may be established by finding the recurrence relation between three contiguous such functions and assuming regularity at $\theta = 0$. The general $d^J_{MM'}(\theta)$ are then found from the P_J by releasing indices[2] from the scalar product QQ'. They are not immediately writable in closed form.

Let us now turn to a larger group G which contains the rotations as a subgroup and see how to determine the relevant rotation functions. One must, for a start, specify the class $|NW\rangle$ of G-basis vectors which characterize the degenerate series in question; here N refers to Casimir operator labels of G and W to the further subgroup labels which completely specify the states. Now, in practice, when G is one of the groups $SO(\nu)$, $SU(\nu)$, and $SU(\nu) \otimes SU(\nu)$, it is easiest to specify the degeneracy character of the basis vectors by stating precisely the nature of Young tableaux that represent the sequence of representations (the relationship between the number of boxes in various rows of the tableau), which corresponds to stating specific relations between the Casimir operator labels N of G. This already indicates that one should apply the tensor representation technique (b) rather than differential method (a). A more important reason for abandoning method (a) is the fact that the Lie algebra may be quite complicated with the additional difficulty that even if one knew the relations between the (quadratic, cubic, $\cdots$) Casimir (C) operators of G and the commutators $[e^{-i\theta J_2}, C]$ the sandwiching of the identities would provide involved second-, third-, $\cdots$, order equations[3] for

$$d^N_{WW'}(\theta) = \langle NW|\, e^{-i\theta J_2}\, |NW'\rangle \tag{1}$$

which, on the face of it, appear difficult to disentangle and solve. We shall therefore adopt method (b) below to compute the rudimentary functions $d^N_{11}(\theta)$ in which the W-subgroup is trivially represented[4] as a singlet.

A. $G = SO(\nu)$

We assume $\nu \geq 3$ in order that $G \supset SO(3)_J$. The most degenerate class of base vectors $|N\rangle$ is characterized by the Young tableau consisting of N boxes in the first row and no others. It is constructed from the basic unit vector[5] q_μ ($\mu = 1, \cdots, \nu$), $q_\mu q_\mu = 1$, by forming the symmetrized scalar product

$$Q_{\mu_1\cdots\mu_N} = q_{\mu_1}\cdots q_{\mu_N} + a_N \sum_{\mu\ \text{perms}} \delta_{\mu_1\mu_2} q_{\mu_3}\cdots q_{\mu_N} + b_N \sum_{\mu\ \text{perms}} \delta_{\mu_1\mu_2}\delta_{\mu_3\mu_4} q_{\mu_5}\cdots q_{\mu_N} + \cdots$$

and imposing tracelessness: $Q_{\mu\mu\mu_3\cdots\mu_N} = 0$. This provides the coefficients $a_N, b_N, \cdots$ and it is easy to find

$$a_N = -(\nu + 2N - 4)^{-1}.$$

The basic recurrences follow by noticing that

$$q_{\mu_1}Q_{\mu_2\cdots\mu_{N+1}} + q_{\mu_2}Q_{\mu_1\mu_3\cdots\mu_{N+1}} + \cdots + q_{\mu_{N+1}}Q_{\mu_1\cdots\mu_N} \equiv \sum_j^{N+1} q_{\mu_j} Q_{\mu_1\cdots(j)\cdots\mu_{N+1}}$$

is a symmetric polynomial of degree q^{N+1} and vanishes under a double tracing operation. Accordingly it can be expressed as the linear combination

$$\alpha Q_{\mu_1\cdots\mu_{N+1}} + \beta \sum_\mu \delta_{\mu_i\mu_j} Q_{\mu_1\cdots(ij)\cdots\mu_{N+1}},$$

where the coefficients α and β can be found by comparing terms of order q^{N+1} and q^{N-1}. One finds

$$\sum_{j=1}^{N+1} q_{\mu_j} Q_{\mu_1\cdots(j)\cdots\mu_{N+1}} = (N+1)Q_{\mu_1\cdots\mu_{N+1}} + \frac{2(\nu+N-3)}{(\nu+2N-2)(\nu+2N-4)} \times \sum_{i,j}^{N+1} \delta_{\mu_i\mu_j} Q_{\mu_1\cdots(ij)\cdots\mu_{N+1}}. \tag{2}$$

The representation functions[6] (corresponding to W-singlets) are proportional to

$$F_N(\cos\theta) = q'_{\mu_1}\cdots q'_{\mu_N} Q_{\mu_1\cdots\mu_N}, \quad q\cdot q' = \cos\theta. \tag{3}$$

If we contract identity (2) over $q'_{\mu_1}\cdots q'_{\mu_{N+1}}$, we obtain

---

[2] For example, $d^j_{\pm 1 0}(\theta)$ is obtained from $d^j_{00}(\theta)$ by evaluating $\epsilon^{\pm}_\mu(q)[\partial(Q\cdot Q')/\partial q_\mu]$ where ϵ_μ is a polarization vector orthogonal to q.

[3] The representation functions for $SO(3,1)$ satisfy a pair of coupled second-order differential equations which, in general, lead to a quartic one after decoupling.

[4] These are the analogs of $d^j_{00}(\theta)$ wherein the $U(1)$ subgroup is trivially represented.

[5] If we designate the generators by $J_{\mu\nu}$ which obey the standard Lie algebra of the orthogonal group,

$$[J_{\kappa\lambda}, J_{\mu\nu}] = i(\delta_{\kappa\nu}J_{\lambda\mu} + \delta_{\lambda\mu}J_{\kappa\nu} - \delta_{\kappa\mu}J_{\lambda\nu} - \delta_{\lambda\nu}J_{\kappa\mu}),$$

then one may identify the rotations as the subgroup consisting of J_{23}, J_{31}, and J_{12}, and to determine $\langle e^{-i\theta J_2}\rangle$ below one fixes $q_\mu = (0,0,1,0,0,\cdots)$ and $q'_\mu = (\sin\theta, 0, \cos\theta, 0, 0, \cdots)$.

[6] Strictly, $F_N = Q'_{\mu_1\ldots\mu_N}Q_{\mu_1\ldots\mu_N}$ but the tracelessness and symmetry conditions allow it to be rewritten as (3).

the recurrence relation

$$F_{N+1} - \cos\theta F_N + \frac{N(\nu+N-3)}{(\nu+2N-2)(\nu+2N-4)} F_{N-1} = 0,$$

letting

$$F_N = N! \frac{C_N}{(\nu+2N-4)(\nu+2N-2)} \cdots \nu(\nu-2), \tag{4}$$

$$(N+1)C_{N+1} - (\nu+2N-2)\cos\theta\, C_N + (\nu+N-3)C_{N-1} = 0, \tag{5}$$

whence we at once recognize the recurrence relation for the Gegenbauer polynomial $C_N^{\frac{1}{2}(\nu-2)}(\cos\theta)$. Thus,

$$d_{11}^N(\theta) = \frac{\Gamma(N+1)\Gamma(\nu-2)}{\Gamma(\nu+N-2)} C_N^{\frac{1}{2}(\nu-2)}(\cos\theta), \tag{6}$$

which is a well-known result[7]; however, the next two cases are new.[8]

B. $G = SU(\nu)$

Here we take $\nu \geq 2$ in order that $G \supset SU(2)_J$. We shall consider the class of basis vectors consisting of $2N$ boxes in the first row, N boxes in the second row, $\cdots$, and N boxes in the $(\nu-1)$th row, the degenerate Feynman series as it is sometimes termed. This sequence is generated[9] from the fundamental tensor q_A^B $(A, B = 1, \cdots, \nu)$, $q_A^B q_B^A = \nu$, by forming the symmetrized product

$$Q_{A_1\cdots A_N}^{B_1\cdots B_N} = \sum_{A \text{ perms}} q_{A_1}^{B_1} q_{A_2}^{B_2} \cdots q_{A_N}^{B_N} + a_N \sum_{\substack{A,B \\ \text{perms}}} \delta_{A_1}^{B_1}\delta_{A_2}^{B_2} q_{A_3}^{B_3} \cdots q_{A_N}^{B_N} + b_N \sum \delta_{A_1}^{B_1}\delta_{A_2}^{B_2}\delta_{A_3}^{B_3}\delta_{A_4}^{B_4} q_{A_5}^{B_5} \cdots q_{A_N}^{B_N} + \cdots$$

and imposing tracelessness $Q_{CA_2\cdots A_N}^{CB_2\cdots B_N} = 0$. This determines the coefficients

$$a_N = -[2(N-2)!(\nu+2N-3)]^{-1}$$

and so on. Basic recurrences are established by observing that

$$\sum_{i,j}^{N+1} q_{A_i}^{B_j} Q_{A_1\cdots(i)\cdots A_{N+1}}^{B_1\cdots(j)\cdots B_{N+1}}$$

is quadruply traceless and hence can be written in the form

$$\alpha Q_{A_1\cdots A_{N+1}}^{B_1\cdots B_{N+1}} + \beta \sum_{i,j,k,l}^{N+1} \delta_{A_i}^{B_j}\delta_{A_k}^{B_l} Q_{A_1\cdots(jl)\cdots(ik)\cdots A_{N+1}}^{B_1\cdots(jl)\cdots B_{N+1}}.$$

α and β are deducible by comparing terms of order q^{N+1} and q^{N-1}, and we obtain the basic identity

$$\sum_{i,j}^{N+1} q_{A_i}^{B_j} Q_{A_1\cdots(i)\cdots A_{N+1}}^{B_1\cdots(j)\cdots B_{N+1}} = Q_{A_1\cdots A_{N+1}}^{B_1\cdots B_{N+1}} + \frac{(\nu+N-2)}{(\nu+2N-1)(\nu+2N-3)} \times \sum_{i,j,k,l}^{N+1} \delta_{A_i}^{B_j}\delta_{A_k}^{B_l} Q_{A_1\cdots(jl)\cdots(ik)\cdots A_{N+1}}^{B_1\cdots(jl)\cdots B_{N+1}}. \tag{7}$$

Upon contracting over an index pair, formula (7) provides the corollaries

$$\sum_k [q_C^{B_k} Q_{A_1\cdots A_N}^{B_1\cdots(k)\cdots B_N C} + q_{A_k}^C Q_{A_1\cdots(k)\cdots A_N C}^{B_1\cdots B_N}] = \frac{2(\nu+N-2)}{\nu+2N-3} \sum_{i,j}^{N} \delta_{A_i}^{B_j} Q_{A_1\cdots(i)\cdots A_N}^{B_1\cdots(j)\cdots B_N} \tag{7'}$$

and

$$q_D^C Q_{A_1\cdots A_{N-1}C}^{B_1\cdots B_{N-1}D} = \frac{(\nu+2N-2)(\nu+N-2)}{(\nu+2N-3)} Q_{A_1\cdots A_{N-1}}^{B_1\cdots B_{N-1}} \tag{7''}$$

which are used below.

The rotation function is proportional to[10]

$$F_N(\cos\theta) = q'^{A_1}_{B_1} \cdots q'^{A_N}_{B_N} Q_{A_1\cdots A_N}^{B_1\cdots B_N}; \quad q_A^B q'^A_B = \nu\cos\theta. \tag{8}$$

Upon contracting (7) over $q'^{A_1}_{B_1} \cdots q'^{A_{N+1}}_{B_{N+1}}$ and making free use of corollaries (7) and (7″), a little work shows that the F_N satisfy the recurrence formula

$$F_{N+1} - (\nu+2N)\cos\theta F_N + \frac{N(\nu+2N)(\nu+2N-2)(\nu+N-1)}{(\nu+2N-1)(\nu+2N-3)} F_{N-1} = 0,$$

whereupon, defining

$$F_N = N! \frac{\nu(\nu+2)\cdots(\nu+2N-2)}{(\nu-1)(\nu+1)\cdots(\nu+2N-3)} C_N, \tag{9}$$

we get

$$(N+1)C_{N+1} - (\nu+2N-1)\cos\theta C_N + (\nu+N-2)C_{N-1} = 0, \tag{10}$$

which allows us to identify $C_N^{\frac{1}{2}(\nu-1)}(\cos\theta)$. Thus for

[7] A. Erdélyi *et al.*, Ed., *Higher Transcendental Functions, Vol. 2* (McGraw-Hill Book Co., Inc., New York, 1954), Eq. 11.2(8).

[8] Actually the formulas for the degenerate series of $SU(\nu) \otimes SU(\nu)$ were given in Ref. 1. However, they were obtained by a much more cumbersome method than the one we have presented here.

[9] If one designates the generators of $SU(\nu)$ by J_A^B, $(J_A^A = 0)$, the Lie algebra is given by the commutation rules $[J_A^B, J_C^D] = \delta_C^B J_A^D - \delta_A^D J_C^B$ and the rotation subgroup $\mathbf{J}$ is given by the generators $J_1^2 + J_3^4 + \cdots$, $J_2^1 + J_4^3 + \cdots$, and $\frac{1}{2}(J_1^1 - J_2^2 - J_3^3 - J_4^4 + \cdots)$, whereupon one identifies the tensors below:

$$q_A^B = \mathbf{q}\cdot(\mathbf{J})_A^B, \quad q'^B_A = \mathbf{q}'\cdot(\mathbf{J})_A^B,$$
$$\mathbf{q} = (0, 0, 1), \quad \mathbf{q}' = (\sin\theta, 0, \cos\theta).$$

[10] Again it is tracelessness and symmetry which allow $F_N = Q'^{A_1\cdots A_N}_{B_1\cdots B_N} Q_{A_1\cdots A_N}^{B_1\cdots B_N}$ to collapse into (8).

the degenerate $SU(\nu)$ series considered,

$$d^N_{11}(\theta) = \frac{\Gamma(N+1)\Gamma(\nu-1)}{\Gamma(N+\nu-1)} C_N^{\frac{1}{2}(\nu-1)}(\cos\theta). \quad (11)$$

Note that when $\nu = 4$ we meet $C_N^{\frac{3}{2}}$ polynomials. In spite of the isomorphism $SU(4) \approx SO(6)$ these differ from the orthogonal functions C^2 for $SO(6)$ because the series of representations are quite different; in one case the space coordinates are assigned to the [6] representation, in the other case to the [15].

C. $G = SU(\nu) \otimes SU(\nu)$

Again we take $\nu \geq 2$ and shall assign the rotations[11] to the diagonal $SU(2)$ subgroup. We consider the degenerate class of vectors described by tableaux $(N, 0, \cdots, 0)$ and $(N, N, \cdots, N)$ for each, respectively, of the $SU(\nu)$, as well as the adjoint series. These are constructed from the basic tensor $(N = 1)$: $q_A^{\hat{B}}$ $[A = 1, \cdots, \nu;\ \hat{B} = 1, \cdots, \nu$; hatted and unhatted indices referring to each of the $SU(\nu)$ groups], $q_A^{\hat{B}} q_{\hat{B}}^C = \nu\delta_A^C$, by forming symmetrical products of these

$$Q_{A_1\cdots A_N}^{\hat{B}_1\cdots\hat{B}_N} = \sum_{B \text{ perms}} q_{A_1}^{\hat{B}_1} \cdots q_{A_N}^{\hat{B}_N}.$$

They are already irreducible representations of $SU(\nu) \otimes SU(\nu)$, though not of the diagonal $SU(\nu)$ (since they are not traceless under AB contractions). Clearly,

$$\sum_{i,j}^{N+1} q_{A_i}^{\hat{B}_j} Q_{A_1\cdots(i)\cdots A_{N+1}}^{\hat{B}_1\cdots(j)\cdots\hat{B}_{N+1}} = (N+1) Q_{A_1\cdots A_{N+1}}^{\hat{B}_1\cdots\hat{B}_{N+1}} \quad (12)$$

and it is also easy to establish directly from the basic definition that

$$q_{\hat{D}}^C Q_{A_1\cdots A_N C}^{\hat{B}_1\cdots\hat{B}_N\hat{D}} = (\nu+N) Q_{A_1\cdots A_N}^{\hat{B}_1\cdots\hat{B}_N}. \quad (12')$$

The fundamental representation function is proportional to

$$F_N(\cos\theta) = q_{\hat{B}_1}^{\prime A_1} \cdots q_{\hat{B}_N}^{\prime A_N} Q_{A_1\cdots A_N}^{\hat{B}_1\cdots\hat{B}_N}; \quad q_A^{\hat{B}} q_{\hat{B}}^{\prime A} = \nu\cos\theta. \quad (13)$$

Consequently, contraction of (12) over $q_{\hat{B}_1}^{\prime A_1} \cdots q_{\hat{B}_{N+1}}^{\prime A_{N+1}}$, making use of (12′), yields the recurrence

$$F_{N+1} - (\nu+2N)\cos\theta F_N + N(N+\nu-1)F_{N-1} = 0$$

or, setting

$$F_N = N!\, C_N, \quad (14)$$

$$(N+1)C_{N+1} - (\nu+2N)C_N + (N+\nu-1)C_N = 0. \quad (15)$$

We recognize here the Gegenbauer function

$$C_N^{\frac{1}{2}\nu}(\cos\theta).$$

Thus for the degenerate $SU(\nu) \otimes SU(\nu)$ series in question,

$$d^N_{11}(\theta) = \frac{\Gamma(N+1)\Gamma(\nu)}{\Gamma(N+\nu)} C_N^{\frac{1}{2}\nu}(\cos\theta). \quad (16)$$

Notice that when $\nu = 2$ we have the $O(4)$ functions C^1 because the series of representations for both groups coincide.

We note a few of the properties of these functions and their relationship with the group theory before we treat the reduction problem in the next section. From the definition

$$C_\alpha^\lambda(z) = \frac{\Gamma(\alpha+2\lambda)}{\Gamma(\alpha+1)\Gamma(2\lambda)} F(\alpha+2\lambda, -\alpha; \lambda+\tfrac{1}{2}; \tfrac{1}{2}-\tfrac{1}{2}z), \quad (17)$$

it is well known that the Gegenbauer function satisfies the symmetry property,

$$C_\alpha^\lambda = -\sin\alpha\pi[C_{-\alpha-2\lambda}/\sin\pi(\alpha+2\lambda)]. \quad (18)$$

Translated in terms of the normalized d functions (6), (11), and (16), we have the "weak equivalence" relations

$$d^N(\theta) = d^{-N-\nu+2}(\theta), \quad \text{for } SO(\nu),$$
$$d^N(\theta) = d^{-N-\nu+1}(\theta), \quad \text{for } SU(\nu),$$
$$d^N(\theta) = d^{-N-\nu}(\theta), \quad \text{for } SU(\nu) \otimes SU(\nu),$$

which simply reflect the fact that the three quadratic[12] Casimir operators $N(N+\nu-2)$, $N(N+\nu-1)$, and $N(N+\nu)$ that label the three sets of degenerate series are invariant under the substitutions $N \to -N-\nu+2$, $-N-\nu+1$, and $-N-\nu$, respectively. Indeed, the positions of the symmetry axes at $N = 1-\frac{1}{2}\nu$, $\frac{1}{2}-\frac{1}{2}\nu$, and $-\frac{1}{2}\nu$, respectively, are associated with the principal unitary infinite-dimensional representations of the noncompact extensions $SO(\nu-1,1)$, $SU(\nu-1,1)$, and $SL(\nu, C)$; and from the symmetry point to $N = 0$ (the first, trivial, finite-dimensional representation) stretch the supplementary series of

[11] The Lie algebra here is provided by

$$[J_A^B, J_C^D] = \delta_C^B J_A^D - \delta_A^D J_C^B,$$
$$[J_A^B, J_{\hat{C}}^{\hat{D}}] = 0,$$
$$[J_{\hat{A}}^{\hat{B}}, J_{\hat{C}}^{\hat{D}}] = \delta_C^B J_{\hat{A}}^{\hat{D}} - \delta_A^D J_{\hat{C}}^{\hat{B}},$$

with $J_A^A = J_{\hat{A}}^{\hat{A}} = 0$. We take the diagonal subalgebra $(\mathbf{J})_A^B + (\mathbf{J})_{\hat{A}}^{\hat{B}}$ with $\mathbf{J}$ defined in Footnote 9 for our rotation subgroup and identify our basic tensors $q_A^{\hat{B}} = \mathbf{q} \cdot (\mathbf{J})_A^{\hat{B}}$ and $q'^{\hat{B}}_A = \mathbf{q}' \cdot (\mathbf{J})_A^{\hat{B}}$.

[12] Stated in terms of the generators given in Footnotes 5, 9, and 11, the Casimir operators are $\frac{1}{2}J_{\mu\nu}J_{\mu\nu}$, $\frac{1}{2}J_A^B J_B^A$, and $\frac{1}{4}(J_A^B J_B^A + J_{\hat{A}}^{\hat{B}} J_{\hat{B}}^{\hat{A}})$, respectively.

representations.[13] Moreover, as the asymptotic behavior of $C_\alpha^\lambda(z)$ is $(z^\alpha + z^{-\alpha-2\lambda})$, the unitary representations $C_{-\lambda}^\lambda(z)$ of the noncompact group extensions have the lowest possible asymptotic behavior and are square-integrable over the corresponding group manifold.

3. ANALYTICAL REDUCTIONS

According to the generalized Regge scheme,[1] a pole belonging to a higher symmetry group G provides an amplitude of the type $\beta(t)\, d^{\alpha(t)}(\theta)/\sin \pi\alpha(t)$, where $d^{\alpha(t)}$ is the generalized rotation function continued along the trajectory to complex $N = \alpha(t)$. One may naturally ask what sequences of ordinary $SU(2)_J$ Regge poles (and trajectories) are implied from the higher group. At integer N values this question is straightforwardly answered since it amounts to a reduction of the higher multiplet (labeled by N) into sets of J multiplets; analytically this corresponds to decomposing the $C_N(\cos\theta)$ polynomials into series of Legendre polynomials $P_J(\cos\theta)$. More generally, one is interested in the reduction problem for complex $N = \alpha$ (i.e., along the whole trajectory) from G to one of several possible subgroups—stated analytically, how $C_\alpha^{\lambda'}$ decomposes as a series of C_α^λ. The solution to this problem is embodied in formula (29) and is valid for arbitrary complex λ and λ', although physically one is concerned only with integer or half-integer values of λ and λ'. We have obtained it in three stages: in the first almost trivial step the decomposition for integer N has been obtained to serve as a boundary check on the succeeding formulas; the second step makes use of the basic recurrence relation between C^λ and $C^{\lambda+1}$ to write $C_\alpha^{\lambda'}$ as a finite series $\sum_k C_{\alpha-2k}^\lambda$ *plus* a finite number of "background terms," valid when λ and λ' differ by an integer; the last stage consists in rewriting the finite remainder as a difference of two infinite series which permit analytic continuation in λ and λ' to values which differ by half integers as well; these "background series" are expressible as integrals over the principal unitary representations, the analog of the background integral at $j = -\frac{1}{2}$ which appears in ordinary Regge theory.

The case of integral N is simply resolved. The polynomial character of $C_N^{\lambda'}(z)$ and its oddness or evenness, according to whether N is odd or even, means that a decomposition of the form

$$C_N^{\lambda'}(z) = \sum_{k=0}^{[N/2]} a_k C_{N-2k}^\lambda(z) \tag{19}$$

is possible for arbitrary λ and λ' with $[N/2]$ denoting the largest integer less than or equal to $N/2$. The coefficients a_k may be found from the usual orthogonality properties of the Gegenbauer functions

$$a_k = \frac{(N-2k+\lambda)\Gamma(N-2k+1)[\Gamma(\lambda)]^2}{2^{1-2\lambda}\pi\Gamma(N-2k+\lambda)} \times \int_{-1}^{1} C_N^{\lambda'}(z) C_{N-2k}^\lambda(z)(1-z^2)^{\lambda-\frac{1}{2}}\, dz \tag{20}$$

for, upon using Rodrigues's formulas,[14]

$$C_N^\lambda(z) = \frac{(-2)^N\Gamma(N+\lambda)\Gamma(N+2\lambda)}{\Gamma(\lambda)\Gamma(N+1)\Gamma(2N+2\lambda)}(1-z^2)^{\frac{1}{2}-\lambda} \times \left(\frac{d}{dz}\right)^N [(1-z^2)^{N+\lambda-\frac{1}{2}}] \tag{21}$$

and

$$\left(\frac{d}{dz}\right)^r C_N^\lambda = \frac{2^r\Gamma(r+\lambda)}{\Gamma(\lambda)} C_{N-r}^{r+\lambda}, \tag{22}$$

we get[15]

$$\begin{aligned} a_k &= 2^{2N-4k+2\lambda-1} \\ &\quad \times \frac{\Gamma(N-2k+\lambda')\Gamma(N-2k+\lambda+1)\Gamma(\lambda)}{\pi\Gamma(2N-4k+2\lambda)\Gamma(\lambda')} \\ &\quad \times \int_{-1}^{1} C_{2k}^{N-2k+\lambda'}(z)(1-z^2)^{N-2k+\lambda-\frac{1}{2}}\, dz \\ &= \frac{(N-2k+\lambda)\Gamma(k+\lambda'-\lambda)\Gamma(N-k+\lambda')\Gamma(\lambda)}{\Gamma(k+1)\Gamma(N-k+\lambda+1)\Gamma(\lambda')\Gamma(\lambda'-\lambda)}. \end{aligned} \tag{23}$$

We may actually continue the series (19) up to $k = \frac{1}{2}(N+n)$ for positive integer $n < 2\lambda$ since the a_k are finite and the C_{-n}^ν vanish identically.[16] But we cannot extend the k summation up to infinity as the Γ functions occurring in a_k do not vanish to permit the series to terminate naturally. Consequently the analytic continuation to complex $N = \alpha$ *cannot* be trivially deduced from formula (19).

We therefore follow a different approach which hinges upon the contiguity relation[17]

$$C_\alpha^{\lambda+1} - C_{\alpha-2}^{\lambda+1} = (\alpha+\lambda)C_\alpha^\lambda/\lambda. \tag{24}$$

Writing (24) for values of α decreasing by units of 2 and summing, we get

$$C_\alpha^{\lambda+1} = \sum_{k=0}^{l} \lambda^{-1}(\alpha+\lambda-2k)C_{\alpha-2k}^\lambda + C_{\alpha-2l-2}^{\lambda+1}. \tag{25}$$

Let us fix l on the right in such a way that $\alpha - 2l$ is as

[13] These statements can be verified for the well-known representations of $SO(2,1) \approx SU(1,1)$ and $SO(3,1) \approx SL(2,C)$.

[14] A. Erdélyi, Ed., *Higher Transcendental Functions, Vol. 1* (McGraw-Hill Book Co., Inc., New York, 1953), Eqs. 3.15 (10), (30).

[15] *Tables of Integral Transforms, Vol. 2*, A. Erdélyi, Ed. (McGraw-Hill Book Co., Inc., New York, 1954), Eqs. 16.3(4). The hypergeometric function ${}_3F_2$ of unit argument is evaluated by Eq. 4.4(6) to give formula (23).

[16] Through the Γ functions appearing in definition (17).

[17] Reference 14, Eqs. 3.15.2 (27), (28).

close as possible to the right of the symmetry point $-\lambda$, i.e., let $0 \leq \alpha - 2l + \lambda < 2$; by this device the remaining function $C^{\lambda+1}_{\alpha-2l-2}$ is of "background size" relative to the C^{λ}, in so far as it is more convergent in its asymptotic $|z|$ behavior; for integer α the background term vanishes automatically and the decomposition (25) reduces to (19). The same trick can be applied successfully to obtain decompositions for λ and λ' values differing by integers (see the Appendix for details). One arrives at

$$C^{\lambda'}_{\alpha} = \sum_{k=0}^{l} a_k C^{\lambda}_{\alpha-2k} + \sum_{r=1}^{\lambda'-\lambda} B_r C^{\lambda+r}_{\alpha-2l-2}, \tag{26}$$

where $\lambda' - \lambda$ is a positive integer, $0 \leq \alpha - 2l + \lambda < 2$, and

$$a_k = \frac{(\alpha - 2k + \lambda)\Gamma(k + \lambda' - \lambda)\Gamma(\alpha - k + \lambda')\Gamma(\lambda)}{\Gamma(k+1)\Gamma(\alpha - k + \lambda + 1)\Gamma(\lambda')\Gamma(\lambda' - \lambda)}, \tag{27}$$

$$B_r = \frac{\Gamma(l + \lambda' - \lambda - r + 1)\Gamma(\alpha + \lambda' - l)\Gamma(\lambda + r)}{\Gamma(l+1)\Gamma(\lambda' - \lambda - r + 1)\Gamma(\alpha + \lambda - l + r)\Gamma(\lambda')}. \tag{28}$$

Formula (26) tells us that a Regge pole of a higher group decomposes as a sum of Regge poles of a lower group plus a number of "background terms"; and the square-integrability[18] of these background terms allows them to be expressed as a *background integral*

$$\int_{\rho=-\infty}^{\infty} B(\rho) C^{\lambda}_{-\lambda+i\rho}(\lambda^2 + \rho^2)\, d\rho$$

over the principal unitary series of the noncompact extension of the lower group. The distribution $B(\rho)$ is perfectly definite, being given[19] by (26), but we shall not belabor the issue by determining it here.

All the above is not quite general enough as we often require decompositions for λ and λ' differing by half integers, the Toller $\rightarrow$ Regge pole decomposition, $C^1 \rightarrow C^{\frac{1}{2}}$, being a case in point. As a matter of fact, we can cover this situation with very little more work. We formally rewrite the finite background sum as the difference of two infinite series

$$\sum_{r=1}^{\lambda'-\lambda} B_r C^{\lambda+r}_{\beta} = \sum_{k=0}^{\infty} b_k C^{\lambda+k+1}_{\beta} + \sum_{k=0}^{\infty} c_k C^{\lambda'+k+1}_{\beta}, \tag{29}$$

with $B_{\lambda'-\lambda+k+1} = b_{\lambda'-\lambda+k} = -c_k$, and analytically continue in $\lambda' - \lambda$. The correctness of this step can be checked by writing each Gegenbauer function as a sum of two hypergeometric functions of argument z^{-2} [formula (31) below] and comparing[20] all powers of z. Our final result therefore is summarized by

$$C^{\lambda'}_{\alpha} = \sum_{k=0}^{l} a_k C^{\lambda}_{\alpha-2k} + \sum_{k=0}^{\infty} b_k C^{\lambda+k+1}_{\alpha-2l-2} + \sum_{k=0}^{\infty} c_k C^{\lambda'+k+1}_{\alpha-2l-2}, \tag{30}$$

where $0 \leq \alpha - 2l + \lambda < 2$, a_k is given by Eq. (27), $b_k = -C_{k+\lambda-\lambda'} = B_{k+1}$, and B_k is given by Eq. (28). Hence, for general λ and λ' we encounter an infinite series of background terms; these cancel to a finite series when $\lambda' - \lambda$ is integral and disappear altogether as they must when α is integer. The reduction formula for the actual representation functions $d^N(\theta)$ is obtained from (30) by renormalization with the factors occurring in formulas (6), (11), and (16).

For the sake of completeness we provide below the reduction formula for representations of the second kind, which appear in the Mandelstam form of the Regge pole amplitude ($\beta Q_{-\alpha-1}/\sin \pi\alpha$ in ordinary Regge theory). These functions of the second kind, $E^{\lambda}_{\alpha}(z)$, in contrast to functions of the first kind, $C^{\lambda}_{\alpha}(z)$, are singular near $z = 1$ and possess no symmetry under the substitution $\alpha \rightarrow -\alpha - 2\lambda$. [For $\lambda = \frac{1}{2}$, $\cos \alpha\pi E_{\alpha}(z)$ coincides with the Legendre functions of the second kind $Q_{\alpha}(z)$.] On the other hand, they have simple asymptotic characteristics. They arise in the break up

$$C^{\lambda}_{\alpha}(z) = \frac{\sin \alpha\pi}{\pi} [E^{\lambda}_{\alpha} + E^{\lambda}_{-\alpha-2\lambda}], \tag{31}$$

$$E^{\lambda}_{\alpha} \equiv \frac{\Gamma(\alpha + 2\lambda)}{\Gamma(\lambda)\Gamma(\alpha + \lambda + 1)} \frac{\pi}{\sin \pi(\alpha + \lambda)} (2z)^{-\alpha-2\lambda} \times F\left(\frac{\alpha + 2\lambda + 1}{2}, \frac{\alpha + 2\lambda}{2}; \alpha + \lambda + 1; \frac{1}{z^2}\right). \tag{32}$$

For $\mathrm{Re}\, \alpha > -\lambda$, the asymptotic behavior as $|z| \rightarrow \infty$ is governed by $E^{\lambda}_{-\alpha-2\lambda}$, which is a series of decreasing powers in z^2 starting from z^{α}. Accordingly, we are interested in a decomposition of the type

$$E^{\lambda'}_{-\alpha-2\lambda'} = \sum_{k=0}^{\infty} a_k E^{\lambda}_{-\alpha-2\lambda+2k},$$

[18] The asymptotic behavior $\sim|z|^{\alpha-2l-2}$ of these background terms, stronger than any single function $C^{\lambda}_{\alpha}(z)$, means that they can only be expressed as a distribution over the most convergent of these, viz., $C_{-\lambda+i\rho}(z)$; these functions pertain to the principal unitary representations of the noncompact group extension and in fact there are standard theorems that any square-integrable function $f(z)$ defined over the group can be decomposed in terms of them. See Footnote 19.

[19] If $B(z)$ satisfies $\int_1^{\infty} |B(z)|^2 (z^2 - 1)^{\lambda-\frac{1}{2}}\, dz < \infty$, then one may write $B(z) = \int C^{\lambda}_{-\lambda+i\rho}(z) B(\rho)(\lambda^2 + \rho^2)\, d\rho$, with the inverse transform $B(\rho) = \int_1^{\infty} B(z) C^{\lambda}_{-\lambda+i\rho}(z)(z^2 - 1)^{\lambda-\frac{1}{2}}\, dz$ according to the standard references [see, for instance, I. M. Gel'fand and M. A. Naimark *Unitäre Darstellungen der klassischen Gruppen* (Berlin, 1957)].

[20] In fact, this was the method by which we originally derived (29) before realizing it could be recovered from (26) by the method described in the text.

which is easily discovered, once it is recognized that the E_α satisfy the same contiguity relations as the C_α; for we can follow the same steps which led to formula (26) to obtain

$$E^{\lambda'}_{-\alpha-2\lambda'} = \sum_{k=0}^{l} a_k E^{\lambda}_{-\alpha-2\lambda+2k} + \sum_{r=1}^{\lambda'-\lambda} B_r E^{\lambda+r}_{-\alpha-2\lambda+2l+2}, \quad (33)$$

except that we no longer have any compulsion to prevent l being as large as we like—there is no longer any well-defined "background line" which governs the asymptotic behavior of representations of the second kind. In fact, since

$$E^{\lambda+r}_{-\alpha-2\lambda+2l+2} = O(|z|^{\alpha-2l-2r-2}),$$

it is clear that all remainder terms on the right (33) can be made as small as we please for $|z| > 1$ by making l sufficiently large. Consequently we can extend the summation over k to infinity for $|z| > 1$ and drop all remainder terms. Thus, for $\operatorname{Re}\alpha > -\lambda$ and $|z| > 1$,

$$E^{\lambda'}_{-\alpha-2\lambda'}(z) = \sum_{k=0}^{\infty} a_k E^{\lambda}_{-\alpha-2\lambda+2k}(z), \quad (34)$$

where the coefficients a_k coincide with the ones obtained previously for representations of the first kind and are given by formula (27).[21]

APPENDIX

This contains the details of derivation of formula (26) from (25). Repeated application of (25), when $\lambda' - \lambda$ is a positive integer, gives

$$C^{\lambda'}_{\alpha} = \sum_{k_{\lambda'-\lambda}=0}^{l} \cdots \sum_{k_1=0}^{l-(k_2+\cdots+k_{\lambda'-\lambda})} C^{\lambda}_{\alpha-2(k_1+\cdots k_{\lambda'-\lambda})} \times \frac{(\alpha+\lambda'-1-2k_{\lambda'-\lambda})}{(\lambda'-1)} \cdots \frac{(\alpha+\lambda-2k_1-\cdots-2k_{\lambda'-\lambda})}{\lambda} + \sum_{r=1}^{\lambda'-\lambda} C^{\lambda+r}_{\alpha-2l-2} \Bigg\{ \sum_{k_{\lambda'-\lambda}=0}^{l} \cdots \sum_{k_{r+1}=0}^{l-2(k_2+\cdots+k_{\lambda-\lambda'})} \times \frac{(\alpha+\lambda'-1-2k_{\lambda'-\lambda})}{(\lambda'-1)} \cdots \frac{(\alpha+\lambda+r-2k_{r+1}-\cdots-2k_{\lambda'-\lambda})}{(\lambda+r)} \Bigg\}.$$

Rearranging the summations,

$$C^{\lambda'}_{\alpha} = \sum_{k=0}^{l} a_{\lambda',k,\lambda} C^{\lambda}_{\alpha-2k} + \sum_{r=1}^{\lambda'-\lambda} B_{\lambda',r,\lambda} C^{\lambda+r}_{\alpha-2l-2}, \quad (A1)$$

where $k = k_1 + k_2 + \cdots + k_{\lambda'-\lambda}$ and

$$a_{\lambda',k,\lambda} = \frac{(\alpha+\lambda-2k)\Gamma(\lambda)}{\Gamma(\lambda')} \tilde{a}_{\lambda',k,\lambda}, \quad (A2)$$

$$B_{\lambda',r,\lambda} = \frac{\Gamma(\lambda+r)}{\Gamma(\lambda')} \sum_{k=0}^{l} (\alpha+\lambda+r-2k) \tilde{a}_{\lambda',k,\lambda+r}, \quad (A3)$$

with

$$\tilde{a}_{\lambda',k,\lambda} = \sum_{k_1=0}^{k} \sum_{k_2=0}^{k-k_1} \cdots \sum_{k_{\lambda'-\lambda-1}=0}^{k-k_1-\cdots-k_{\lambda'-\lambda-2}} \times \{\alpha+\lambda+1-2(k-k_1)\} \cdots \{\alpha+\lambda'-1-2(k-\cdots-k_{\lambda'-\lambda-1})\}. \quad (A4)$$

Careful inspection of (A4) reveals the recurrence relation

$$\tilde{a}_{\lambda',k,\lambda} = (\alpha+\lambda+1-2k)\tilde{a}_{\lambda',k,\lambda+1} + \tilde{a}_{\lambda',k-1,\lambda}$$

which solves as

$$\tilde{a}_{\lambda',k,\lambda} = \frac{\Gamma(\alpha+\lambda'-k)\Gamma(\lambda'-\lambda+k)}{\Gamma(\alpha+\lambda+1-k)\Gamma(\lambda'-\lambda)\Gamma(k+1)}. \quad (A5)$$

The summation in (A3) is easily performed if the recurrences among the $\tilde{a}$ are used and one gets

$$B_{\lambda',r,\lambda} = \frac{\Gamma(\lambda+r)}{\Gamma(\lambda')} \tilde{a}_{\lambda',l,r+\lambda-1} = \frac{\Gamma(\lambda+r)\Gamma(l+\lambda'-\lambda-r+1)\Gamma(\alpha+\lambda'-l)}{\Gamma(\lambda')\Gamma(l+1)\Gamma(\lambda'-\lambda-r+1)\Gamma(\alpha+\lambda-l+r)}. \quad (A6)$$

Hence, expression (26) follows.

[21] In the special case $\lambda' = 1$, $\lambda = \frac{1}{2}$, this agrees with the "Lorentz pole" decomposition into Regge poles; see A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Comments on Kinetic Equation for Autocorrelation Functions

P. RÉSIBOIS, J. BROCAS,* AND G. DECAN
Free University of Brussels

(Received 5 January 1968)

The non-Markoffian kinetic equation for the one-particle momentum autocorrelation function, derived by Zwanzig and studied in great detail recently by Berne, Boon, and Rice, is analyzed in the weak coupling limit. It is shown that, in this limit, this kinetic equation remains non-Markoffian because the kernel which determines the memory effects only decays very slowly. More precisely, it tends to zero over times of the order of the relaxation time itself and *not*, as could be expected, over the much shorter collision time. The comparison with the more traditional approach, based on the solution of a transport equation, is also discussed.

1. INTRODUCTION

Time-dependent autocorrelation functions (af) have played an important role in the recent development of nonequilibrium statistical mechanics. In particular, they appear as a major tool in the theory of transport coefficients.[1]

The traditional method for evaluating these quantities is based on the solution of a transport equation of the Boltzmann type for a one-particle distribution function (df), from which the af can be computed (see Ref. 1). However, the possibility exists of deriving a kinetic equation which directly applies to these af. This elegant alternative was first exploited by Zwanzig[2] and recently reconsidered with much detail by Berne, Boon, and Rice.[3]

In particular these authors obtained a kinetic equation for the one-particle momentum af, denoted by $\Psi(t)$ [see Eq. (2.5)], of the following non-Markoffian type:

$$\partial_t \Psi(t) = -\int_0^t dt' K(t-t')\Psi(t'), \qquad (1.1)$$

where the kernel $K(t)$, a function of time only, is defined in terms of the N-particle Hamiltonian with the help of projection operators. The simplicity of Eq. (1.1) is remarkable; in particular, it opens the way to powerful semiphenomenological approximations in cases, like dense fluids, where explicit calculations cannot be performed. This is an important result which cannot be achieved by the transport-equation approach.

Yet the price paid for this apparent simplicity is fairly high; indeed the whole difficulty of the many-body problem is "hidden" in the kernel $K(t)$, which has very complicated analytical properties, even in the simplest cases. It is the aim of the present paper to illustrate this point: we discuss in detail the properties of the kernel K in the limit of weak coupling.

In Sec. 2, we first review briefly previous work leading to Eq. (1.1) and then discuss a very simple approximation leading to a Markoffian kinetic equation, describing an exponential decay for $\Psi(t)$. At first sight, this approximation seems to correspond to the weak coupling limit of Eq. (1.1).

In the third section, we reconsider the same problem from the more traditional point of view: we calculate $\Psi(t)$ from the solution of the Fokker–Planck transport equation which governs the evolution of the one-particle momentum df in a weakly coupled gas. The surprising result is that this calculation runs in conflict with our approximate solution obtained in Sec. 2.

The reason for this discrepancy is analyzed with great detail in Sec. 4. Using the formulation developed by Prigogine and co-workers,[4,5] we show that *even in the weak coupling limit* [$\lambda^2 \to 0$, $t \to \infty$, $(\lambda^2 t)$ finite] *one is not allowed to take the Markoffian approximation to Eq.* (1.1). Indeed, we show that the simplified kernel $K(t)$, which results from this limit, is a distribution which only tends to zero after times of the order of the relaxation time of the system (λ^{-2}); moreover, its explicit form remains rather complicated. When this exact limiting form for $K(t)$ is taken, we recover of course from the kinetic equation (1.1) the result derived before from the transport-equation method. However, it looks as if, for the particular exact model discussed here, the intermediate stage (1.1) is more a step backward than a step forward in the analysis of the behavior of $\Psi(t)$.

This result does not, of course, put any discredit

* Chargé de Recherche au Fonds National Belge de la Recherche Scientifique.

[1] See, for instance, P. Résibois, J. Chem. Phys. **41**, 2979 (1964), and references quoted therein.

[2] R. Zwanzig, Phys. Rev. **124**, 983 (1961).

[3] B. Berne, J. P. Boon, and S. Rice, J. Chem. Phys. **45**, 1086 (1966).

[4] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962).

[5] P. Résibois, in *Many-Particle Physics*, E. Meeron, Ed. (Gordon & Breach Science Publishers, Inc., New York, 1967).

on the above-mentioned semiphenomenological approaches, valid in dense systems, but shows the difficulty of using Eq. (1.1) in exact models.

2. KINETIC EQUATION FOR THE VELOCITY AUTOCORRELATION FUNCTION

We consider a system of N classical particles enclosed in a volume Ω. Its Hamiltonian (mass $m = 1$) is

$$H_N = H_0 + \lambda V = \sum_{i=1}^{N} \tfrac{1}{2}\mathbf{p}_i^2 + \lambda \sum_{i>j=1}^{N} V(|\mathbf{r}_i - \mathbf{r}_j|), \quad (2.1)$$

where $\mathbf{r}_i$ and $\mathbf{p}_i$, respectively, denote the position and the momentum of particle i; $V(|\mathbf{r}_i - \mathbf{r}_j|)$ is the interaction between particles i and j.

The Liouville operator associated with (2.1) is

$$L_N = L_0 + \lambda \delta L \quad (2.2)$$

with

$$L_0 = -i \sum_{j=1}^{N} \mathbf{p}_j \cdot \frac{\partial}{\partial \mathbf{r}_j}, \quad (2.3)$$

$$\lambda \delta L = i\lambda \sum_{i>j=1}^{N} \frac{\partial V}{\partial \mathbf{r}_{ij}} \cdot \left(\frac{\partial}{\partial \mathbf{p}_i} - \frac{\partial}{\partial \mathbf{p}_j} \right). \quad (2.4)$$

The one-particle momentum af is defined by

$$\Psi(t) = \frac{1}{\langle p_1^2 \rangle} \langle \mathbf{p}_1 \exp(-iL_N t)\mathbf{p}_1 \rangle, \quad (2.5)$$

where the bracket denotes the average over all phase space weighted by the canonical distribution

$$\langle \cdots \rangle = \int dr^N \, dp^N \cdots \rho_N^{eq}. \quad (2.6)$$

Here we have used the symbols r^N for $(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$, dr^N for $d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N$, etc.

Moreover, we write for later convenience the canonical equilibrium distribution ρ_N^{eq} as

$$\rho_N^{eq} \equiv P_N^{eq}(r^N) \prod_{i=1}^{N} \varphi_1^{eq}(p_i), \quad (2.7)$$

where $\varphi_1^{eq}(p_i)$ is the normalized one-particle Maxwell–Boltzmann df, while $P_N^{eq}(r^N)$ denotes the configuration distribution

$$P_N^{eq}(r^N) = \frac{\exp(-\beta\lambda V)}{\int dr^N \exp(-\beta\lambda V)}. \quad (2.8)$$

As shown by Zwanzig and by Berne, Boon, and Rice,[6] $\Psi(t)$ obeys the following non-Markoffian kinetic equation:

$$\partial_t \Psi(t) = -\lambda^2 \int_0^t dt' K(t'; \lambda) \Psi(t - t'). \quad (2.9)$$

[6] See, respectively, Refs. 2 and 3.

In this formula, the kernel $K(t; \lambda)$ is defined by

$$K(t; \lambda) = \frac{1}{\langle p_1^2 \rangle} \langle \mathbf{F}_1 \exp[-i(1 - P)L_N t]\mathbf{F}_1 \rangle, \quad (2.10)$$

where $\mathbf{F}_1$ is the total force acting on particle 1, i.e.,

$$\mathbf{F}_1 = -\sum_{j=2}^{N} \frac{\partial V(|\mathbf{r}_1 - \mathbf{r}_j|)}{\partial \mathbf{r}_1}, \quad (2.11)$$

and P is a projection operator defined as follows: for any phase vector $\mathbf{G}(r^N, p^N)$ one has

$$P\mathbf{G}(r^N, p^N) = \frac{\mathbf{p}_1 \rho_N^{eq}}{\langle p_1^2 \rangle} \int dr'^N \, dp'^N \mathbf{p}_1' \cdot \mathbf{G}(r'^N, p'^N). \quad (2.12)$$

The proof of Eq. (2.9) is fairly simple but will not be reproduced here; the reader is referred to the references mentioned above.

We want to consider here a very simple model, where all calculations can in principle be performed exactly; hence we discuss the *weak coupling limit* of Eq. (2.9).

Let us first naively expand the kernel $K(t; \lambda)$ in the coupling parameter λ, so that

$$K(t; \lambda) = K^{(2)}(t) + \sum_{m=3}^{\infty} \lambda^{m-2} K^{(m)}(t). \quad (2.13)$$

Let us assume that this expansion converges for small enough λ and that each term $K^{(n)}(t)$ tends to zero after some time τ_c, independent of λ.

Moreover, we make the hypothesis that the integrals

$$\int_0^\infty dt \, t^m K^{(n)}(t) \quad (m > 0; n \geq 2)$$

converge. We then formally expand (2.9) as

$$\partial_t \Psi(t) = -\lambda^2 \int_0^t dt' \left[K^{(2)}(t') + \sum_{n=1}^{\infty} \lambda^n K^{(n+2)}(t') \right] \times \left[\Psi(t) + \sum_{m=1}^{\infty} \frac{(-t')^m}{m!} \partial_t^m \Psi(t) \right]. \quad (2.14)$$

We also notice that

(1) $\partial_t \Psi(t)$ as well as higher order derivatives are at least of order λ^2;

(2) in the limit $t \gg \tau_c$, we have

$$\int_0^t dt' K^{(n)}(t') t'^m = \int_0^\infty dt' K^{(n)}(t') t'^m, \quad (2.15)$$

a result which is λ-independent.

Thus taking the well-known weak coupling limit $\lambda^2 \to 0$, $t \to \infty$, $(\lambda^2 t)$ finite, we arrive at the following Markoffian form[7]:

$$\partial_t \Psi(t) = -\lambda^2 \zeta^{(2)} \Psi(t); \; \lambda^2 \to 0, \; t \to \infty, \; (\lambda^2 t) \text{ finite}, \quad (2.16)$$

[7] See Ref. 5, as well as H. Terwiel and P. Mazur, Physica **32**, 1813 (1966); P. Mazur and H. Terwiel, Physica **36**, 289 (1967).

where the coefficient $\zeta^{(2)}$ is defined by

$$\zeta^{(2)} = \int_0^\infty dt K^{(2)}(t). \tag{2.17}$$

Deliberately, we have not tried to derive Eq. (2.16) with much rigor because this would require the knowledge of the analytical properties of the kernel $K(t; \lambda)$, which we have not yet investigated. It is, however, interesting to point out here that standard complex variable theory allows to formulate fairly weak sufficient conditions on the Laplace transform of the kernel $K(t; \lambda)$ for Eq. (2.16) to be rigorously valid. This matter is discussed in detail in Appendix A.

Using (2.10), (2.11), and the obvious property

$$PL_0 \cdots = 0, \tag{2.18}$$

we may cast (2.17) in the explicit form

$$\zeta^{(2)} = \frac{1}{\langle p_1^2 \rangle} \int_0^\infty dt \langle \mathbf{F}_1 \exp(-iL_0 t)\mathbf{F}_1 \rangle_0, \tag{2.19}$$

where the subscript 0 means that the average is taken with the unperturbed canonical distribution.

Because $\Psi(0) = 1$, Eq. (2.16) tells us immediately that the decay of the one-particle af is governed by a single exponential

$$\Psi(t) = \exp(-\lambda^2 \zeta^{(2)} t), \quad \lambda^2 \to 0, \quad t \to \infty, \quad (\lambda^2 t) \text{ finite}. \tag{2.20}$$

In the next section, we show that an alternative approach of the weak-coupling behavior of $\Psi(t)$ throws doubt on the validity of this result. As a matter of fact, we see in Sec. 4 that the assumptions made here about the behavior of $K(t; \lambda)$ are incorrect. Yet, at first sight, the analysis given here seems just as satisfactory as what is generally accepted as "correct" in most work on nonequilibrium statistical mechanics.

3. ANALYSIS OF THE af IN TERMS OF THE FOKKER-PLANCK EQUATION

For reasons which will become clear soon, let us formally rewrite the af (2.5) as

$$\Psi(t) = u^{-2} \langle \mathbf{u} \cdot \mathbf{p}_1 \exp(-iL_N t)[1 + (\mathbf{u} \cdot \mathbf{p}_1)/kT] \rangle, \tag{3.1}$$

where $\mathbf{u}$ is an arbitrary parameter with the dimensions of a velocity. The equivalence between (2.5) and (3.1) is readily obtained once the following elementary properties are noticed:

$$\langle \mathbf{p}_1 \rangle = 0, \quad \langle p_1^2 \rangle = 3kT, \tag{3.2}$$

$$\langle p_1^{(\alpha)} \exp(-iL_N t) p_1^{(\beta)} \rangle = \delta_{\alpha\beta}^{Kr} \langle p_1^{(\alpha)} \exp(-iL_N t) p_1^{(\alpha)} \rangle, \tag{3.3}$$

where α and β denote Cartesian components of a vector ($\alpha = x, y, z$; $\beta = x, y, z$).

Using (2.6), Eq. (3.1) can be cast in the following form:

$$\Psi(t) = u^{-2} \int d\mathbf{p}_1 \mathbf{u} \cdot \mathbf{p}_1 \varphi_1(\mathbf{p}_1; t), \tag{3.4}$$

with

$$\varphi_1(\mathbf{p}_1; t) = \int dr^N \, dp^{N-1} \rho_N(r^N, p^N; t) \tag{3.5}$$

and

$$\rho_N(r^N, p^N; t) = \exp[-iL_N t][1 + (\mathbf{u} \cdot \mathbf{p}_1)/kT] \rho_N^{\text{eq}}. \tag{3.6}$$

Clearly, $\rho_N(r^N, p^N; t)$ is a solution of the Liouville equation

$$i\partial_t \rho_N(r^N, p^N; t) = L_N \rho_N(r^N, p^N; t) \tag{3.7}$$

and is normalized to one:

$$\begin{aligned} \int dr^N \, dp^N \rho_N(r^N, p^N; t) &= \int dr^N \, dp^N \rho_N(r^N, p^N; 0) \\ &= \int dr^N \, dp^N [1 + (\mathbf{u} \cdot \mathbf{p}_1)/kT] \rho_N^{\text{eq}} \\ &= 1. \end{aligned} \tag{3.8}$$

It may thus be interpreted as an N-particle df and, from (3.5), $\varphi_1(\mathbf{p}_1; t)$ is a one-particle df.

The motivation for writing $\Psi(t)$ in the form (3.1) is now quite clear: the introduction of the formal parameter $\mathbf{u}$ allows us, with the help of (3.4), to calculate the af from the one-particle df $\varphi_1(\mathbf{p}_1; t)$, which has been extensively studied in the literature (see Refs. 4 and 5, and references quoted therein). In particular, let us consider the weak coupling limit $\lambda^2 \to 0$, $t \to \infty$, $(\lambda^2 t)$ finite. From (3.6) (taken at $t = 0$ and in the limit $\lambda \to 0$), we also deduce the following initial condition:

$$\varphi_1(\mathbf{p}_1; 0) = [1 + (\mathbf{u} \cdot \mathbf{p}_1)/kT] \varphi_1^{\text{eq}}(p_1), \tag{3.9}$$

$$\varphi_1(\mathbf{p}_i; 0) = \varphi_1^{\text{eq}}(p_i), \quad i \neq 1. \tag{3.10}$$

It is then well known that $\varphi_1(\mathbf{p}_1; t)$ obeys the following equation:

$$\partial_t \varphi_1(\mathbf{p}_1; t) = \lambda^2 \Omega^{(2)}(\mathbf{p}_1) \varphi_1(\mathbf{p}_1; t), \tag{3.11}$$

where $\Omega^{(2)}(\mathbf{p}_1)$ denotes the linear Fokker–Planck operator

$$\Omega^{(2)}(\mathbf{p}_1) = \frac{\rho}{8\pi^3} \int d\mathbf{k} \int d\mathbf{p}_2 \, \mathbf{k} V_k \cdot \frac{\partial}{\partial \mathbf{p}_1} \pi \delta[\mathbf{k} \cdot (\mathbf{p}_1 - \mathbf{p}_2)] \times \mathbf{k} V_k \cdot \left(\frac{\partial}{\partial \mathbf{p}_1} - \frac{\partial}{\partial \mathbf{p}_2} \right) \varphi_1^{\text{eq}}(p_2). \tag{3.12}$$

Here, V_k is the Fourier transform of the interaction potential

$$V_k = \int d\mathbf{r} V(r) \exp(-i\mathbf{k}\cdot\mathbf{r}); \tag{3.13}$$

$\delta(x)$ is the Dirac delta function, and ρ is the particle density.

The eigenvalues and eigenfunctions of (3.12) are *not* known; however, it is readily verified that the operator $[\varphi_1^{eq}]^{-1}\Omega^{(2)}(\mathbf{p}_1)$ is self-adjoint and satisfies the conditions which guarantee that the Sturm–Liouville type of problem,

$$[\varphi_1^{eq}]^{-1}\Omega^{(2)}(\mathbf{p}_1)\Phi_n(\mathbf{p}_1) = \Lambda_n[\varphi_1^{eq}]^{-1}\Phi_n(\mathbf{p}_1), \tag{3.14}$$

possesses a complete set of eigenfunctions Φ_n with real eigenvalues Λ_n (Ref. 9).

Moreover, we have the orthonormality property

$$\int d\mathbf{p}_1[\varphi_1^{eq}]^{-1}\Phi_n(\mathbf{p}_1)\Phi_m(\mathbf{p}_1) = \delta_{n,m}^{Kr}. \tag{3.15}$$

It is also known that $\Omega^{(2)}(\mathbf{p}_1)$ is a seminegative definite operator such that (see Refs. 4, 5, and 8)

$$\Lambda_0 = 0, \quad \Phi_0(\mathbf{p}_1) = \varphi_1^{eq}(p_1),$$
$$\Lambda_n < 0 \quad (n > 0). \tag{3.16}$$

If we expand $\varphi_1(\mathbf{p}_1; t)$ in the series of eigenfunctions Φ_n, so that

$$\varphi_1(\mathbf{p}_1; t) = (3/kT)^{\frac{1}{2}} \sum_{\alpha=x,y,z} \sum_{n=0}^{\infty} u^{(\alpha)} C_n^{(\alpha)}(t)\Phi_n(\mathbf{p}_1), \tag{3.17}$$

one obtains easily from (3.11), (3.14)–(3.16), the following convergent formal expansion for the af $\Psi(t)$:

$$\Psi(t) = \sum_{\alpha=x,y,z} \sum_{n>0} |C_n^{(\alpha)}|^2 \exp(-\lambda^2 |\Lambda_n| t). \tag{3.18}$$

The expansion coefficients $C_n^{(\alpha)} \equiv C_n^{(\alpha)}(0)$ are given by

$$C_n^{(\alpha)} = (3/kT)^{\frac{1}{2}} \int d\mathbf{p}_1 p_1^{(\alpha)} \Phi_n(\mathbf{p}_1) \quad (n > 0) \tag{3.19}$$

and obey the sum rule

$$\sum_{\alpha=x,y,z} \sum_{n>0} |C_n^{(\alpha)}|^2 = 1. \tag{3.20}$$

Comparing (3.18) with (2.20), we immediately see that these two expressions are only compatible

(a) if $\zeta^{(2)}$ is one of the eigenvalues, say Λ_{n_1} $(n_1 \neq 0)$ of (3.14);

(b) if simultaneously

$$C_n^{(\alpha)} = \delta_{n,n_1}^{Kr}. \tag{3.21}$$

Let us first show that condition (a) also implies condition (b); then we verify that (3.21) cannot be true.

We write (2.19) explicitly as

$$\begin{aligned}\zeta^{(2)} &= \frac{1}{\langle p_1^2\rangle}\int_0^\infty dt \frac{\rho}{8\pi^3}\int d\mathbf{k}\int d\mathbf{p}_1\int d\mathbf{p}_2 \mathbf{k} V_k \\ &\quad \times \exp[-i\mathbf{k}\cdot(\mathbf{p}_1 - \mathbf{p}_2)t]\mathbf{k}V_k\varphi_1^{eq}(p_1)\varphi_1^{eq}(p_2) \\ &= \frac{\rho}{\langle p_1^2\rangle 8\pi^3}\int d\mathbf{k}\int d\mathbf{p}_1\int d\mathbf{p}_2 \\ &\quad \times \mathbf{k}V_k\pi\delta[\mathbf{k}\cdot(\mathbf{p}_1 - \mathbf{p}_2)]\mathbf{k}V_k\varphi_1^{eq}(p_1)\varphi_1^{eq}(p_2)\end{aligned} \tag{3.22}$$

and, with (3.2) and integration by parts, as

$$\zeta^{(2)} = -\frac{1}{\langle p_1^2\rangle}\int d\mathbf{p}_1[\varphi_1^{eq}]^{-1}[\mathbf{p}_1\varphi_1^{eq}]\Omega^{(2)}(\mathbf{p}_1)[\mathbf{p}_1\varphi_1^{eq}]. \tag{3.23}$$

With the help of (3.15) and (3.19), this also gives

$$\zeta^{(2)} = \sum_{\alpha=x,y,z}\sum_{n>0} |C_n^{(\alpha)}|^2 |\Lambda_n|, \tag{3.24}$$

which immediately shows that condition (a) implies condition (b).

Now, this latter requirement in turn imposes that $\mathbf{p}_1\varphi_1^{eq}(p_1)$ is an eigenfunction of $\Omega^{(2)}(\mathbf{p}_1)$. Direct calculation immediately tells us that this is *not* the case, and we are thus lead to a contradiction.

4. ANALYTICAL BEHAVIOR OF THE KERNEL $K(t, \lambda)$ IN THE WEAK COUPLING LIMIT

The results of the preceding sections lead us to a paradox: we have found a different behavior for $\Psi(t)$, according to the method we have used! If we look back at the various steps of the calculation, we realize that the weak point is in the unchecked assumptions made on the series (2.13), a crucial step for obtaining the simple exponential decay (2.17).

We now show that these assumptions are indeed incorrect: *even in the weak coupling limit, one has to retain an infinite class of terms in the expansion* (2.13).

In order to prove this, it is simpler to consider the Laplace transform $\tilde{K}(z; \lambda)$ of the kernel $K(t; \lambda)$. We write

$$K(t; \lambda) = \frac{-1}{2\pi i}\oint_C dz \exp[-izt]\tilde{K}(z; \lambda), \tag{4.1}$$

where the contour C passes above all singularities of $\tilde{K}(z; \lambda)$. We have from (2.10) and (2.18) that

$$\tilde{K}(z; \lambda) = \frac{1}{\langle p_1^2\rangle}\left\langle \mathbf{F}_1 \frac{1}{L_0 + (1 - P)\lambda\delta L - z}\mathbf{F}_1 \right\rangle. \tag{4.2}$$

[8] See especially, I. Prigogine and R. Balescu, Physica **23**, 555 (1957).

[9] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. I* (Interscience Publishers, Inc., New York, 1953). Note that all equations are written here for a discrete spectrum. Because of the very fast decay of the initial condition (3.9) at infinity, this result appears to be correct in view of recent rigorous analysis [see C. H. Su, J. Math. Phys. **8**, 148 (1967)]. However, our conclusions would be unaffected if the continuous spectrum of $\Omega^{(2)}(\mathbf{p}_1)$ contributed to the evolution of $\varphi_1(\mathbf{p}_1; t)$.

A possible reason for a nontrivial behavior of $\tilde{K}(z;\lambda)$ for small z [and thus of $K(t;\lambda)$ for large t] is apparent in Eq. (4.2); the projection operator P, defined by (2.12), appears in the denominator. This makes it very difficult to prove any statement about the analytical properties of $\tilde{K}(z;\lambda)$; in particular, the theory of Cauchy integrals, which plays such an important role in the theory of Prigogine and coworkers (see Refs. 4 and 5), is not directly applicable to (4.2).

In order to circumvent this difficulty, we establish an equivalent form for $\tilde{K}(z;\lambda)$ which involves in the denominator another projection operator P_0, defined by

$$P_0\mathbf{G}(r^N, p^N) = \Omega^{-N}\int dr'^N \mathbf{G}(r'^N, p^N). \tag{4.3}$$

The following identities can then be shown to hold:

$$\tilde{K}(z;\lambda) = \frac{\tilde{K}^0(z;\lambda)}{1 - z^{-1}\lambda^2\tilde{K}^0(z;\lambda)} \tag{4.4}$$

with

$$\tilde{K}^{(0)}(z;\lambda) = \frac{1}{\langle p_1^2\rangle}\int d\mathbf{p}_1\mathbf{p}_1[1 + z^{-1}\lambda^2\Omega(\mathbf{p}_1, z;\lambda)]^{-1} \times D(\mathbf{p}_1, z;\lambda)\mathbf{p}_1\varphi_1^{eq}(p_1). \tag{4.5}$$

In this last equation, we have introduced the one-particle operators D and Ω, defined respectively by

$$D(\mathbf{p}_1, z;\lambda) = \int dr^N\, dp^{N-1}\delta L \frac{1}{L_0 + \lambda(1 - P_0)\delta L - z} L_N \times P_N(r^N)\prod_{i=2}^{N}\varphi_1^{eq}(p_i) \tag{4.6}$$

and

$$\Omega(\mathbf{p}_1, z;\lambda) = \frac{1}{\Omega^N}\int dr^N\, dp^{N-1}\delta L \frac{1}{L_0 + \lambda(1 - P_0)\delta L - z}\delta L \times \prod_{i=2}^{N}\varphi_1^{eq}(p_i). \tag{4.7}$$

Although this has not been proved rigorously, we assume that $D(\mathbf{p}_1, z;\lambda)$ and $\Omega(\mathbf{p}_1, z;\lambda)$ have a finite value in the thermodynamic limit $N \to \infty$, $\Omega \to \infty$, $N/\Omega = \rho$ finite; this is easily verified to lowest order in λ.

The proof of these identities is straightforward, although fairly long; for these reasons, we have put it in Appendix B.

Equations (4.4) to (4.7) express the kernel $\tilde{K}(z;\lambda)$ in terms of operators which involve only the projector $(1 - P_0)$ in the denominator, but no more $(1 - P)$.

It may be asked, of course, what is the advantage of these new expressions, which look much more complicated than the compact expression (4.2)? The answer is that the analytical properties of operators involving the projector $(1 - P_0)$ are fairly well understood, in the frame of perturbation calculus at least.

Let us consider for instance Eq. (4.7); we have, provisionally assuming convergence,

$$\lambda^2\Omega(\mathbf{p}_1, z;\lambda) = \sum_{n=2}^{\infty}\lambda^n\Omega^{(n)}(\mathbf{p}_1, z) \tag{4.8}$$

with

$$\Omega^{(n)}(\mathbf{p}_1, z) = \frac{(-1)^n}{\Omega^N}\int dr^N\, dp^{N-1}\delta L[(L_0 - z)^{-1}(1 - P_0)\delta L]^{n-1} \times \prod_{i=2}^{N}\varphi_1^{eq}(p_i). \tag{4.9}$$

We can write this quantity in Fourier space with respect to spatial coordinates:

$$\Omega^{(n)}(\mathbf{p}_1, z) = (-1)^n\int dp^{N-1} \times \langle\{0\}|\,\delta L[(L_0 - z)^{-1}(1 - P_0)\delta L]^{n-1}\,|\{0\}\rangle \times \prod_{i=2}^{N}\varphi_1^{eq}(p_i). \tag{4.10}$$

Here we have used the usual Dirac notation, in terms of kets $|\{\mathbf{k}\}\rangle$ and bras $\langle\{\mathbf{k}\}|$, for the Fourier matrix element of an arbitrary function or operator $G(r^N, p^N)$:

$$\langle\{\mathbf{k}\}|\,G\,|\{\mathbf{k}'\}\rangle = \frac{1}{\Omega^N}\int dr^N \exp\left[-i\sum_{j=1}^{N}\mathbf{k}_j\mathbf{r}_j\right] \times G(r^N, p^N)\exp\left[i\sum_{j=1}^{N}\mathbf{k}'_j\mathbf{r}_j\right]. \tag{4.11}$$

It is well known that the operator $(1 - P_0)$ has the following formal representation in Fourier space:

$$(1 - P_0) = \sum_{\{\mathbf{k}\}\neq\{0\}}|\{\mathbf{k}\}\rangle\langle\{\mathbf{k}\}|, \tag{4.12}$$

i.e., it picks up from a complete Fourier expansion the nonvanishing wavenumber components. This property has been studied in detail recently[10] and is also briefly discussed in Appendix C.

For the operator ψ, defined by

$$\psi(z) = \langle\{0\}|\,\delta L[(L_0 - z)^{-1}(1 - P_0)\delta L]^n\,|\{0\}\rangle, \tag{4.13}$$

we may thus write the equivalent form

$$\psi(z) = \langle\{0\}|\,\delta L[(L_0 - z)^{-1}\delta L]^n\,|\{0\}\rangle_{\text{irr}}, \tag{4.14}$$

where the subscript "irr" (irreducible) is taken with the meaning used in the theory of Prigogine and coworkers[4,5]; that is, in the evaluation of (4.14), all

[10] M. Baus, Bull. Acad. Sci. Belg. **53**, 1291 (1967).

intermediate states are to be taken with nonvanishing wavenumbers $\{\mathbf{k}\} \neq \{0\}$.

Operators of the type (4.14) have been extensively studied. We define

$$\psi^+(z) = \psi(z) \quad [z \in S^+],$$

$$\psi^+(z) = [\text{analytical continuation of } \psi^+(z \in S^+)] \quad [z \in S^-]. \tag{4.15}$$

It can then be shown that, in the thermodynamic limit, the only singularities of $\psi^+(z)$ are in the lower half-plane S^-, *at a finite distance from the real axis, independent of the coupling constant* (see also Refs. 4 and 5). This result cannot, however, by any means, be considered as a mathematical theorem, and this without doubt brings a certain weakness in the conclusion to be obtained below. Yet, many model cases can be treated exactly, which confirm the statement made here.[11]

More precisely, it can be safely assumed that the singularities of $\psi^+(z)$ are at a distance from the real axis characterized by

$$\text{Im } z \geq -\tau_c^{-1}, \tag{4.16}$$

where τ_c is some time characterizing the duration of the collision process. We see thus that if $|z| \ll \tau_c^{-1}$, then $\psi^+(z)$ may be expanded in a Taylor series around the origin.

Let us now explicitly assume that the forces are weak ($\lambda \ll 1$) and let us also preclude that we shall later be interested in values of z of the order λ^2 [$(\lambda^2 t)$ finite implies $(\lambda^2 z^{-1})$ finite]. We may then write from (4.8), (4.10), (4.14), and the basic assumption on $\psi^+(z)$, that

$$\begin{aligned}\Omega^+(\mathbf{p}_1, z; \lambda) &= \lambda^2 \Omega^{+(2)}(\mathbf{p}_1; z) + O(\lambda^3); \\ &\qquad \lambda^2 \to 0, \quad (\lambda^2 z^{-1}) \text{ finite} \\ &= \lambda^2 \Omega^{+(2)}(\mathbf{p}_1; 0) + O(\lambda^3); \\ &\qquad \lambda^2 \to 0, \quad (\lambda^2 z^{-1}) \text{ finite},\end{aligned} \tag{4.17}$$

and a similar result holds for $D^+(\mathbf{p}_1, z; \lambda)$. Here, the superscript "+" has the same meaning as in (4.15).

Moreover, from (4.6), (4.7), and (2.8), we see that

$$D^{+(2)}(\mathbf{p}_1, z) = \Omega^{+(2)}(\mathbf{p}_1, z) \tag{4.18}$$

and

$$\Omega^{+(2)}(\mathbf{p}_1, z) = \lim_{\epsilon \to 0} \Omega^{-N} \int dr^N \, dp^{N-1} \delta L \frac{1}{L_0 - i\epsilon} \delta L \times \prod_{i=2}^{N} \varphi_1^{eq}(p_i). \tag{4.19}$$

[11] A pathological case is found when one tries to expand $\psi^+(z)$ as a virial series, in power of the density ρ; the well-known divergence in the virial expansion of the transport coefficients is related to a logarithmic singularity of $\psi^+(z)$ at the origin $z = 0$.

With the help of (2.3), (2.4), (3.13), and the well-known formula

$$\lim_{\epsilon \to 0} [i(x - i\epsilon)]^{-1} = \pi\delta(x) - iP(x^{-1}), \tag{4.20}$$

it is easily established from (4.19) that

$$\Omega^{+(2)}(\mathbf{p}_1, 0) = -i\Omega^{(2)}(\mathbf{p}_1), \tag{4.21}$$

where the Fokker–Planck operator $\Omega^{(2)}(\mathbf{p}_1)$ was defined in (3.12).

It is now a matter of some elementary algebra to derive from the formulas established in this section the following expression for the kernel $\tilde{K}(z, \lambda)$, valid in the weak coupling limit $\lambda^2 \to 0$, $z \to 0$, $(\lambda^2 z^{-1})$ finite:

$$\lambda^2 \tilde{K}(z; \lambda) = \left[\frac{1}{\langle p_1^2 \rangle} \int d\mathbf{p}_1 \mathbf{p}_1 \frac{1}{z - i\lambda^2 \Omega^{(2)}(\mathbf{p}_1)} \mathbf{p}_1 \varphi_1^{eq}(p_1)\right]^{-1} - z. \tag{4.22}$$

It is also easy to establish that, when $\lambda^2 \to 0$, all the singularities of $\tilde{K}(z; \lambda)$ other than those described by (4.22) (i.e., for $|z|$ finite) are located at a distance from the real axis given by (4.16).

The difficulty with the assumptions made about expansion (2.13) is now easily understood. Indeed, if we *formally* expand (4.23) in a power series of λ^2, we get

$$\lambda^2 \tilde{K}^+(z; \lambda) = \sum_{n=1}^{\infty} \lambda^{2n} \tilde{K}^{+(2n)}(z); \quad \lambda \to 0, \quad z \to 0, \quad (\lambda^2 z^{-1}) \text{ finite}, \tag{4.23}$$

and the coefficients $\tilde{K}^{+(2n)}(z)$ have the form

$$\tilde{K}^{+(2n)}(z) = (i^n \beta_{2n})/z^{n-1}. \tag{4.24}$$

Here the coefficients β_{2n} are independent of both z and λ, and the factor i^n is introduced for convenience. The explicit form of these β_{2n} is complicated but not needed here, except for β_2 which is immediate to calculate:

$$\beta_2 = -\frac{1}{\langle p_1^2 \rangle} \int d\mathbf{p}_1 \mathbf{p}_1 \Omega^{(2)}(\mathbf{p}_1) \mathbf{p}_1 \varphi_1^{eq}(p_1) = \zeta^{(2)}. \tag{4.25}$$

Inserting (4.23) into (4.1) and performing the trivial residue evaluation at $z = 0$, we obtain the following expression for the time-dependent kernel [see also the argument after (5.1)]:

$$\lambda^2 K(t; \lambda) = 2\lambda^2 \zeta^{(2)} \delta(t) + \lambda^2 K'(t; \lambda) + O[\exp(-t/\tau_c)]; \quad \lambda^2 \to 0, \quad t \to \infty, \quad (\lambda^2 t) \text{ finite}, \tag{4.26}$$

with

$$\lambda^2 K'(t; \lambda) = \sum_{n=2}^{\infty} \lambda^{2n} \beta_{2n} \frac{t^{n-2}}{(n-2)!}; \quad \lambda^2 \to 0, \quad t \to \infty, \quad (\lambda^2 t) \text{ finite}. \tag{4.27}$$

The full complexity of Eq. (4.26) should be appreciated. First, let us notice that the leading term $2\lambda^2\zeta^{(2)}\delta(t)$, when inserted into (2.9), would immediately lead to the simple exponential decay (2.20); moreover, the singular nature of this term is not related to the formal nature of expansion (4.23). As we show in Appendix D, this term remains when the compact form (4.22) is Laplace-inverted: it is related to the fact that the first derivative of $\Psi(t)$ does not vanish at $t = 0$ when the limit $t \to 0$ is taken *after* one goes first to the weak coupling approximation [see (3.18)].

Next, we notice that the remainder $\lambda^2 K'(t; \lambda)$ is of order λ^4 for times of the order λ^{-2} (assuming convergence). Nevertheless, *it does not satisfy the requirement of tending to zero over a time independent of* λ. In other words, its total contribution for times of the order λ^{-2} is

$$\int_0^t dt' K'(t'; \lambda) = \lambda^2 \sum_{n=2}^{\infty} \beta_{2n} \frac{(\lambda^2 t)^{n-1}}{(n-1)!} = O(\lambda^2); \quad \lambda^2 \to 0, \quad t \to \infty, \quad (\lambda^2 t) \text{ finite}, \quad (4.28)$$

a quantity of the same order as the integral of the first term in (4.26). We are thus not allowed to neglect this remainder, and the derivation leading to (2.16) is in error.

We shall not pursue further the analysis of the expansion (4.24), because this would require a detailed calculation of the coefficients β_{2n}. Moreover, in the next section, we derive the correct behavior for the af $\Psi(t)$ by a direct method, independent of any perturbation expansion. Also a detailed model calculation, together with a few general properties of the kernel $K(t; \lambda)$, are presented in Appendix D.

5. CONCLUSION

In the preceding section, we have established the correct form for the kernel $\tilde{K}^+(z; \lambda)$ in the weak coupling limit [see Eq. (4.22)] and we have shown why the naive reasoning which leads to a simple exponential decay for the af $\Psi(t)$ is in error.

We still have to prove that Eq. (4.22) for $\tilde{K}^+(z; \lambda)$ solves the apparent paradox we have found and leads to (3.18) for the solution $\Psi(t)$.

Let us first take the Laplace transform of Eq. (2.9); we obtain

$$\Psi(t) = \frac{1}{2\pi i} \oint_C dz \frac{\exp(-izt)}{z + \lambda^2 \tilde{K}(z; \lambda)}. \quad (5.1)$$

The contour C is chosen as a parallel to the real axis in the upper half-plane, closed by an infinite semicircle in the lower half-plane; because the contribution to (5.1) coming from this semicircle vanishes, we may as well replace $\tilde{K}(z; \lambda)$ by its determination $\tilde{K}^+(z; \lambda)$, defined by a formula analogous to (4.14). Moreover, in the weak coupling limit, we may use expression (4.22) for the kernel $\tilde{K}^+(z; \lambda)$, because in Eq. (5.1) we are only interested in the singularities of the integrand located at a distance of the order λ^2 from the real axis; all other singularities give asymptotically negligible contributions. After some straightforward algebra, we get

$$\Psi(t) = \frac{1}{2\pi i} \oint_C dz \exp(-izt) \left[\frac{1}{\langle p_1^2 \rangle} \int d\mathbf{p}_1 \mathbf{p}_1 \times \frac{1}{z - i\lambda^2 \Omega^{(2)}(\mathbf{p}_1)} \mathbf{p}_1 \varphi_1^{eq}(p_1) \right]; \quad \lambda^2 \to 0, \quad t \to \infty, \quad (\lambda^2 t) \text{ finite}. \quad (5.2)$$

The bracketed term in the integrand can be expressed as a series expansion in the eigenfunctions of $\Omega^{(2)}(\mathbf{p}_1)$; using (3.14) and (3.19), we get immediately

$$\Psi(t) = \frac{1}{2\pi i} \oint_C dz \exp(-izt) \left[\sum_{\alpha=x,y,z} \sum_{n>0} \frac{|C_n^{(\alpha)}|^2}{z + i\lambda^2 |\Lambda_n|} \right].$$

Commuting the sum and the integral, we get then the required result (3.18) by application of residue theorem.

This shows the complete equivalence between the transport-equation method, as developed in Sec. 3, and the approach based on a kinetic equation for the af when the weak coupling limit is correctly taken.

However, we see that for explicit detailed calculations, the apparent simplicity of the kinetic equation (2.9) is paid for by the great complication of the kernel, even in a simple limiting situation [see Eq. (4.22)].

This of course does not prevent these kinetic equations from being extremely useful in the semi-phenomenological treatment of systems which are beyond the power of quantitative treatment; a nice example is furnished by the Berne-Boon and Rice analysis of Rahman's computer experiments.[12] Yet the present derivation shows very clearly the great care that has to be exercised when exact calculations are developed starting from these kinetic equations. Such difficulties are usually avoided in the more traditional treatment based on transport equations; as we exemplified in Sec. 3, the calculations, although less elegant, are then generally straightforward.

ACKNOWLEDGMENTS

We gratefully acknowledge pertinent comments on this paper by Dr. G. Severne and M. Baus.

[12] See Ref. 3 and A. Rahman, Phys. Rev. **136**, A405 (1964).

APPENDIX A: MARKOFFIAN LIMIT OF A NON-MARKOFFIAN KINETIC EQUATION

We formulate[13] here rigorously sufficient conditions on the Laplace transform of the kernel $K(t;\lambda)$ which guarantee that the weak coupling limit of Eq. (2.9) is the Markoffian equation

$$\partial_t \Psi(t) = -\lambda^2\left[\int_0^\infty dt' K^{(2)}(t')\right]\Psi(t). \tag{A1}$$

This analysis is based on the Lagrange theorem,[14] which we recall here without proof: If $\tilde{K}(z;\lambda)$ is holomorphic in a domain including a circle on which

$$\lambda^2 \frac{|\tilde{K}(z;\lambda)|}{|z|} < 1 \tag{A2}$$

and if

$$|\tilde{K}(z;\lambda)| \neq 0 \tag{A3}$$

inside this circle, then the equation

$$z + \lambda^2 \tilde{K}(z;\lambda) = 0 \tag{A4}$$

has a single zero, z_0, inside this circle. Moreover, this zero is given by

$$z_0 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\left[\frac{d^{n-1}\tilde{K}^n(z;\lambda)}{dz^{n-1}}\right]_{z=0} \lambda^{2n}. \tag{A5}$$

From this result we want to prove the following, more general result:

Theorem: Assume that the Laplace transform $\tilde{K}(z;\lambda)$ of the kernel [see (4.1)] exists and satisfies the following requirements:

(1) $\tilde{K}(z;\lambda)$ is holomorphic in z in the semi-infinite plane

$$\text{Im } z > -\tau_c^{-1} \quad (\tau_c > 0, \lambda\text{-independent}) \tag{A6}$$

and holomorphic in λ for λ smaller than some finite value λ_c;

(2) $\tilde{K}(0;\lambda)$ is finite for all $|\lambda| < |\lambda_c|$, so that

$$|\tilde{K}(0;\lambda)| > B > 0 \quad (B \text{ is } \lambda\text{-independent}); \tag{A7}$$

(3) $|\tilde{K}(z;\lambda)|$ tends to zero at least as fast as $|z|^{-1}$ for $z \to \infty$.

Consider the solution of Eq. (2.9), which in general is

$$\Psi(t) = \frac{1}{2\pi i}\oint_C dz \frac{\exp(-izt)}{z + \lambda^2 \tilde{K}(z;\lambda)}, \tag{A8}$$

where the contour C lies above all the singularities of the integrand. In the weak coupling limit $\lambda^2 \to 0$, $t \to \infty$, $(\lambda^2 t)$ finite, we then have the following asymptotic result:

$$\Psi(t) = \exp[i\lambda^2 \tilde{K}(0;0)t]. \tag{A9}$$

First we establish a lemma.

Lemma: Consider a region

$$\text{Im } z \geq -r > -\tau_c^{-1}, \tag{A10}$$

where r is an arbitrary positive number, smaller than τ_c; and consider values of λ^2 such that[15]

$$\lambda^2 < rB/\{A(r)[M(r) + B]\} \tag{A11}$$

where $M(r)$ and $A(r)$, respectively, denote the least upper bound of the modulus of $\tilde{K}(z;\lambda)$ on a circle of radius r centered around the origin and in the whole region (A10) [clearly $M(r) \leq A(r)$].

Then, under the conditions of the theorem, $z + \lambda^2 \tilde{K}(z;\lambda)$ has only one zero in the region (A10) and this zero is given by Lagrange formula (A5).

Proof: The proof runs in three steps:

(a) From (A11), the zeros z_0 of $z + \lambda^2 \tilde{K}(z;\lambda)$ in (A10) are such that

$$|z_0| = |\lambda^2 \tilde{K}(z_0;\lambda)| < \lambda^2 A(r) < rB/[M(r) + B]. \tag{A12}$$

(b) Inside and on the circle of radius R, defined by

$$R = rB/[M(r) + B], \tag{A13}$$

the function $\tilde{K}(z;\lambda)$ does not vanish. Indeed, $\tilde{K}(z;\lambda)$, being holomorphic inside the radius r, has the expansion

$$\tilde{K}(z;\lambda) = \sum_{n=0}^{\infty} Q_n(\lambda) z^n, \quad |z| \leq r. \tag{A14}$$

Then

$$|\tilde{K}(z;\lambda) - \tilde{K}(0;\lambda)| \leq \sum_{n=1}^{\infty} |Q_n(\lambda)|\,|z|^n. \tag{A15}$$

Moreover, by the Cauchy principle, we have

$$|Q_n(\lambda)| \leq [M(r,\lambda)]/r^n < [M(r)]/r^n, \tag{A16}$$

where $M(r,\lambda)$ denotes the maximum modulus of $\tilde{K}(z,\lambda)$ on the circle of radius r. Thus

$$|\tilde{K}(z;\lambda)| \geq |\tilde{K}(0,\lambda)| - |\tilde{K}(z;\lambda) - \tilde{K}(0;\lambda)| \tag{A17}$$

$$> B - M(r)z(r - z)^{-1}, \quad |z| < r. \tag{A18}$$

The second term increases when $|z|$ increases, but clearly

$$|\tilde{K}(z;\lambda)| > 0 \tag{A19}$$

when $|z| \leq R$ [see (A13)].

[13] It is gratefully acknowledged that the matter of this appendix is the result of fruitful discussions with Dr. M. De Leener.

[14] See, for instance, G. Sansone and J. Gerretsen, *Lecture on the Theory of Functions of a Complex Variable* (P. Noordhoff Ltd., Groningen, The Netherlands, 1960).

[15] We assume here $rB/A(r)[M(r) + B] < \lambda_c^2$; strictly speaking, instead of (A11) we have $\lambda^2 < \min[rB/A(r)[M(r) + B], \lambda_c^2]$; the case $rB/A(r)[M(r) + B] \geq \lambda_c^2$ is, however, trivial to treat.

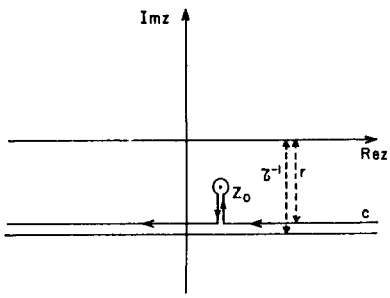


FIG. 1. The contour of integration in Eq. (A8).

(c) On the circle of radius R, one has [see (A11) and (A13)]

$$\lambda^2 \frac{|\tilde{K}(z;\lambda)|}{|z|} < \frac{rB}{[M(r)+B]A(r)} \frac{M(r)}{rB[M(r)+B]^{-1}} < 1. \tag{A20}$$

By (A19) and (A20), we see that the circle R satisfies the conditions of Lagrange theorem; we have thus one zero inside this circle, which is given by (A5). From (A12), there can be no other zero inside the region (A10). This completes the proof of the lemma.

With this lemma the proof of the main theorem is immediate. Indeed we can choose the contour of integration in (A8) as is indicated in the Fig. 1, because the integrand of (A8) has $z_0(\lambda)$ as only singularity in the region (A10). As the two segments parallel to the imaginary axis do not contribute, we have

$$\Psi(t) = \exp[-iz_0(\lambda)t] + \frac{1}{2\pi i}\exp(-tr) \times \int_{-\infty-ir}^{\infty-ir} d\omega \frac{\exp[-i\omega t]}{(\omega - ir) + \lambda^2\tilde{K}(\omega - ir;\lambda)}. \tag{A21}$$

The integral in the second member exists because the integrand is regular and the condition imposed on $|\tilde{K}(z;\lambda)|$ at infinity insures that the limit $\omega \to \infty$ poses no difficulty (in the Cauchy principal value sense).

Thus

$$\int_{-\infty-ir}^{\infty-ir} d\omega \frac{\exp[-i\omega t]}{(\omega - ir) + \lambda^2\tilde{K}(\omega - ir;\lambda)} < C.$$

As $|z_0| < R < r$, we see that in the asymptotic limit $t \to \infty$, one has

$$\Psi(t) = \exp -iz_0(\lambda)t. \tag{A22}$$

Note that this result is valid for any λ satisfying (A11).[16] As the right side of this latter equation is independent of λ, we see that this condition is automatically fulfilled when $\lambda^2 \to 0$. Moreover, the regularity of $\tilde{K}(z;\lambda)$ with respect to λ also insures that

$$\lim_{\lambda\to 0} z_0(\lambda) = -\lambda^2\tilde{K}(0;0), \tag{A23}$$

and we thus get (A9) in the weak coupling limit.

Note that *neither* (A6) *nor the asymptotic behavior of* $|\tilde{K}(z;\lambda)|$ for $|z| \to 0$ *are satisfied by the kernel* (4.22).

Let us finally point out that the proof was possible because $K(t;\lambda)$ is an ordinary function. In transport theory, equations similar to (A5) are often used (see Refs. 4 and 5), but no comparable rigor can be achieved in their derivation, because in this latter case the kernel is an operator.

APPENDIX B: PROOF OF EQUATIONS (4.4) TO (4.7)

With the help of (2.4), (2.6), and (4.2), we can write

$$\lambda^2\tilde{K}(z;\lambda) = \frac{1}{\langle p_1^2\rangle}\int dr^N\, dp^N \mathbf{p}_1\lambda\delta L \times \frac{1}{L_0 + \lambda(1-P)\delta L - z} L_N\mathbf{p}_1\rho_N^{eq}. \tag{B1}$$

We use the identity

$$[L_0 + \lambda(1-P)\delta L - z]^{-1} = [L_0 + \lambda\delta L - z]^{-1} + [L_0 + \lambda\delta L - z]^{-1}P\lambda\delta L[L_0 + \lambda(1-P)\delta L - z]^{-1} \tag{B2}$$

and insert it in (B1). Using (2.12), we obtain

$$\tilde{K}(z;\lambda) = \tilde{K}^0(z;\lambda) + (\lambda^2/z)\tilde{K}^0(z;\lambda)\tilde{K}(z;\lambda), \tag{B3}$$

where

$$\tilde{K}^{(0)}(z;\lambda) = \frac{1}{\langle p_1^2\rangle}\int dr^N\, dp^N \mathbf{p}_1\delta L \times \frac{1}{L_0 + \lambda\delta L - z} L_N\mathbf{p}_1\rho_N^{eq}. \tag{B4}$$

The identity (B3) was in fact already proved in Ref. 3.

In a second step, we transform $\tilde{K}^0(z;\lambda)$ by using the projection operator P_0 defined by (4.3). Let us first introduce the auxiliary operator $A(r^N, p^N, z;\lambda)$:

$$A(r^N, p^N, z;\lambda) = \delta L[L_0 + \lambda\delta L - z]^{-1}L_N = \delta L[L_0 + \lambda P_0\delta L + \lambda(1-P_0)\delta L - z]^{-1}L_N. \tag{B5}$$

An identity similar to (B2), but involving now P_0, leads us to the following result:

$$A(r^N, p^N, z;\lambda)\mathbf{p}_1\rho_N^{eq} = \psi(r^N, p^N, z;\lambda)\mathbf{p}_1\rho_N^{eq} - \frac{1}{z}\psi(r^N, p^N, z;\lambda)\frac{1}{\Omega^N} \times \int dr'^N A(r'^N, p^N, z;\lambda)\mathbf{p}_1\rho_N^{eq}, \tag{B6}$$

[16] We see here that the asymptotic result (A22) is valid for λ^2 *finite*, satisfying (A11). We do not know, however, whether this condition (A11) is very restrictive or not.

where we have introduced the operator

$$\psi(r^N, p^N, z; \lambda) = \delta L[L_0 + \lambda(1 - P_0)\delta L - z]^{-1} L_N. \tag{B7}$$

We decompose the equilibrium distribution ρ_N^{eq} according to (2.7) and integrate both members of Eq. (B6) over all positions and all momenta except $\mathbf{p}_1$. In doing this, we take into account that, in the product term on the right side of (B6), only particle 1 appears *explicitly* in both $\psi(r^N, p^N, z; \lambda)$ and $A(r^N, p^N, z; \lambda)$. If any other particle were common to these two operators, the corresponding contribution would vanish *in the thermodynamic limit*. More precisely, it is immediately verified that, after integrating (B6) over r^N and p^{N-1}, this product term vanishes except when $\mathbf{p}_1$ appears in the two factors of the product; imposing one more particle to be common to these two factors reduces the contribution by a factor $1/N$, a negligible contribution when

$$N \to \infty, \quad \Omega \to \infty, \quad N/\Omega = \rho \text{ finite}.$$

Using then a familiar trick,[17] we may then formally write

$$\begin{aligned}
&\int dr^N\, dp^{N-1} \psi(r^N, p^N, z; \lambda) \frac{1}{\Omega^N} \\
&\quad \times \int dr'^N A(r'^N, p^N, z; \lambda) \mathbf{p}_1 \prod_{i=1}^{N} \varphi_1^{eq}(p_i) P_N^{eq}(r'^N) \\
&= \int dr^N\, dp^{N-1} \psi(r^N, p^N, z; \lambda) \prod_{i=2}^{N} \varphi_1^{eq}(p_i) \frac{1}{\Omega^N} \\
&\quad \times \int dr'^N\, dp'^{N-1} A(r'^N, p_1 p'^{N-1}, z; \lambda) \mathbf{p}_1 \varphi_1^{eq}(p_1) \\
&\quad \times \prod_{i=2}^{N} \varphi_1^{eq}(p_i') P_N^{eq}(r'^N) + O\left(\frac{1}{N}\right).
\end{aligned} \tag{B8}$$

We get then from (B6)–(B8) the required Eqs. (4.5)–(4.7).

APPENDIX C: PROOF OF EQUATION (4.12)

Take an arbitrary function $G(r^N, p^N)$; its Fourier expansion is

$$G(r^N, p^N) = \sum_{\{\mathbf{k}\}} G_{\{\mathbf{k}\}}(p^N) \exp\left[i \sum_{j=1}^{N} \mathbf{k}_j \mathbf{r}_j\right]. \tag{C1}$$

From (4.3), we immediately get

$$\begin{aligned}
(1 - P_0) G(r^N, p^N) &= \sum_{\{\mathbf{k}\}} G_{\{\mathbf{k}\}}(p^N) \exp\left[i \sum_{j=1}^{N} \mathbf{k}_j \mathbf{r}_j\right] - G_{\{0\}} \\
&= \sum_{\{\mathbf{k}\} \neq \{0\}} G_{\{\mathbf{k}\}}(p^N) \exp\left[i \sum_{j=1}^{N} \mathbf{k}_j \mathbf{r}_j\right].
\end{aligned} \tag{C2}$$

Following Dirac,[18] we may consider $G(r^N, p^N)$ as the configurational representation of an abstract vector $|G(p^N)\rangle$; we write then (C2) in abstract form as

$$\langle r^N | (1 - P_0) | G(p^N) \rangle = \sum_{\{\mathbf{k}\} \neq \{0\}} \langle r^N | \{\mathbf{k}\} \rangle \langle \{\mathbf{k}\} | G(p^N) \rangle,$$

which also implies

$$(1 - P_0) | G(p^N) \rangle = \sum_{\{\mathbf{k}\} \neq \{0\}} | \{\mathbf{k}\} \rangle \langle \{\mathbf{k}\} | G(p^N) \rangle.$$

This proves Eq. (4.12).

APPENDIX D: FURTHER CONSIDERATIONS ON THE WEAK COUPLING LIMIT OF KERNEL $K(t; \lambda)$

Using Eqs. (4.22), (3.14), and (3.19), one obtains easily the following form for the weak coupling limit of $\tilde{K}^+(z; \lambda)$:

$$\lambda^2 \tilde{K}^+(z; \lambda) = \frac{i \sum_{\alpha; n>0} \lambda^2 |C_n^{(\alpha)}|^2 |\Lambda_n| / z + i\lambda^2 |\Lambda_n|}{\sum_{\alpha; n>0} |C_n^{(\alpha)}|^2 / z + i\lambda^2 |\Lambda_n|}. \tag{D1}$$

Unfortunately, this formula does not allow us to calculate easily the time-dependent kernel $K(t; \lambda)$ in closed form, although, as seen below, some general properties of $K(t; \lambda)$ can be deduced.

In order to have a better understanding of Eq. (D1), let us first consider a model calculation, in which only two coefficients $C_n^{(\alpha)}$ do not vanish. Thus we write

$$\begin{aligned}
&\lambda^2 \tilde{K}^+_{\text{model}}(z; \lambda) \\
&= i\lambda^2 \frac{|\Lambda_1|\, |C_1|^2 / (z + i\lambda^2 |\Lambda_1|) + |\Lambda_2|\, |C_2|^2 / (z + i\lambda^2 |\Lambda_2|)}{|C_1|^2 / (z + i\lambda^2 |\Lambda_1|) + |C_2|^2 / (z + i\lambda^2 |\Lambda_2|)}
\end{aligned} \tag{D2}$$

with [see (3.20)]

$$|C_1^2| + |C_2^2| = 1. \tag{D3}$$

Thus we have

$$\lambda^2 \tilde{K}^+_{\text{model}}(z; \lambda) = i\lambda^2 \frac{[\beta z + i\lambda^2 |\Lambda_1|\, |\Lambda_2|]}{z + i\lambda^2 \alpha} \tag{D4}$$

with

$$\alpha = |C_1^2|\, |\Lambda_2| + |C_2^2|\, |\Lambda_1|,$$

$$\beta = |C_1^2|\, |\Lambda_1| + |C_2^2|\, |\Lambda_2|.$$

Some care has to be exercised in calculating the inverse kernel by (4.1). Indeed $\tilde{K}^+_{\text{model}}(z; \lambda)$ tends to a constant when $|z| \to \infty$. We may, however, write it as

$$\lambda^2 \tilde{K}^+_{\text{model}}(z; \lambda) = i\lambda^2 \left[\beta + i\lambda^2 \frac{|\Lambda_1|\, |\Lambda_2| - \alpha\beta}{z + i\lambda^2 \alpha}\right]. \tag{D5}$$

The first term gives a delta function by Laplace inversion, while the second may be calculated by

[17] P. Résibois and H. T. Davis, Physica **30**, 1077 (1964).

[18] P. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, London, 1947), 3rd ed.

residue theorem. We obtain[19]

$$\lambda^2 K_{\text{model}}(t;\lambda) = 2\lambda^2\beta\delta(t) + \lambda^2 K'_{\text{model}}(t;\lambda) \quad \text{(D6)}$$

with

$$\lambda^2 K'_{\text{model}}(t;\lambda) = \lambda^4(|\Lambda_1|\,|\Lambda_2| - \alpha\beta)\exp -\lambda^2\alpha t. \quad \text{(D7)}$$

These results, in agreement with the discussion of Sec. 4, show that:

(1) The singular term in the kernel (D6) is independent of any λ expansion; it just expresses the fact that

$$\partial_t \psi(t)|_{t=0} \neq 0 \quad \text{(D8)}$$

when one *first* takes the weak coupling limit and *then* lets $t \to 0$.

(2) The remainder $\tilde{K}'_{\text{model}}(t;\lambda)$ is a well-defined expression of order λ^4. However, it only tends to zero for times of the order λ^{-2}. In particular, the integral

$$\lambda^2\int_0^\infty dt' K'_{\text{model}}(t';\lambda) = \lambda^2\frac{(|\Lambda_1|\,|\Lambda_2| - \alpha\beta)}{\alpha} \quad \text{(D9)}$$

is of the same order as the corresponding integral of the singular part.

Let us now come back to the complete expression (D1). In analogy with (D5) and using (3.24), we write it as

$$\lambda^2\tilde{K}^+(z;\lambda) = i\lambda^2\zeta^{(2)} + \lambda^2\tilde{K}'(z;\lambda) \quad \text{(D10)}$$

with

$$\lambda^2\tilde{K}'(z;\lambda) = i\lambda^2\left[\sum_{\alpha,n}\frac{|C_n^{(\alpha)}|^2(|\Lambda_n| - \zeta^{(2)})}{z + i\lambda^2|\Lambda_n|}\right] \times \left[\sum_{\alpha,n}\frac{|C_n^{(\alpha)}|^2}{z + i\lambda^2|\Lambda_n|}\right]^{-1}. \quad \text{(D11)}$$

By Laplace inversion, the first term gives the distribution $2\lambda^2\zeta^{(2)}\delta(t)$, but the remainder

$$\lambda^2 K'(t;\lambda) = -\frac{1}{2\pi i}\oint_C dz \exp(-izt)\lambda^2\tilde{K}'(z;\lambda) \quad \text{(D12)}$$

cannot be evaluated easily in closed form. Yet it is easy to verify that $\tilde{K}'(z;\lambda)$ is a meromorphic function, with poles in the lower half-plane.

Then one shows readily that

$$\lambda^2\int_0^\infty dt K'(t;\lambda) = -i\lambda^2\tilde{K}'(0;\lambda) = \lambda^2\left[\sum_{\alpha,n}\frac{|C_n^{(\alpha)}|^2(|\Lambda_n| - \zeta^{(2)})}{|\Lambda_n|}\right] \times \left[\sum_{n,\alpha}\frac{|C_n^{(\alpha)}|^2}{|\Lambda_n|}\right]^{-1}, \quad \text{(D13)}$$

which again is of order λ^2.

Finally, let us point out that we have proved that the poles γ_i of $K'(z;\lambda)$ are purely imaginary, simple, and ordered in such a way that

$$|\Lambda_i| < |\gamma_i| < |\Lambda_{i+1}|, \quad i = 1, 2, \cdots, n, \cdots. \quad \text{(D14)}$$

We shall not, however, reproduce this calculation in detail here.

[19] The factor 2 in front of the delta function arises because we adopt the convention $\int_0^\infty 2\delta(t)f(t)\,dt = f(0_+)$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Gel'fand–Kirillov Conjecture on the Lie Field of an Algebraic Lie Algebra*†

YUTZE CHOW
Department of Physics, University of Wisconsin, Milwaukee, Wisconsin

(Received 11 March 1968)

This is an article written to review with sufficient detail the so-called Gel'fand–Kirillov conjecture concerning the isomorphisms between the quotient fields of the algebras generated by canonical variables $[p_i, q_j] = \delta_{ij}1$ and the quotient fields of the universal enveloping algebras of algebraic Lie algebras. This conjecture sheds new light on the relation between the universal enveloping algebra of an algebraic Lie algebra and the Lie algebras of dynamic groups in quantum mechanics.

I. INTRODUCTION

The Gel'fand–Kirillov conjecture concerns the quotient field associated with the universal enveloping algebra of an algebraic Lie algebra. This conjecture is of some interest to physicists due mainly to the fact that it relates, through isomorphism, the field of quotients constructed out of the universal enveloping algebra to the field of quotients constructed out of some associative algebra generated by the $2n$ generators

$$p_1, \cdots, p_n, \quad q_1, \cdots, q_n$$

over the ground field $K[x]$ which is just the polynomial ring on a set of indeterminates

$$x_1, \cdots, x_k$$

over the field (i.e., a commutative division ring) K. We note K is also the ground field of the Lie algebra of concern. The $p_1, \cdots, p_n, q_1, \cdots, q_n$ mentioned here are required to satisfy

$$p_i \cdot q_j - q_j \cdot p_i = \delta_{ij} 1_{\mathrm{UL}} \tag{1}$$

where 1_{UL} denotes the unit element of the universal enveloping algebra UL of an algebraic Lie algebra L.

We see that (1) is nothing but the quantum-mechanical commutation relations of canonically conjugate operators. In order to state the conjecture we first define "algebraic" Lie algebras. To make the discussion more or less self-contained, we mention briefly some of the basic concepts and definitions of ring theory that are relevant to the conjecture. Then we shall discuss some properties of the universal enveloping algebra before we go into a full exposition of the conjecture. It is rather unfortunate that the conjecture cannot be fully understood without the mathematical machinery that may seem to be heavy to many physicists.

* Four talks given at the Department of Physics, Syracuse University, in June and December 1967.

† Work supported in part by the Graduate School of the University of Wisconsin, Milwaukee, Wisconsin.

II. NOTATION

VS	Vector space
AA	Associative algebra; and when it is used as a subscript we simply mean to concentrate on the natural AA structure of the object
LA	Lie algebra; and if it is used as a subscript like $\mathcal{A}_{\mathrm{LA}}$ (where $\mathcal{A}$ is an AA) we mean that the AA $\mathcal{A}$ is now made into an LA by means of $[a, a'] \equiv a \cdot a' - a' \cdot a$, where $a \cdot a'$ denotes the AA multiplication
$\mathrm{End}_K V$	The set of all K-endomorphisms on a vector space V over a ground field K ($\Phi \equiv \mathrm{End}_K V$ is also used for short)
Aut V	The set of all automorphisms on V
$\mathrm{Hom}_K(\Phi, K)$	The set of all K-homomorphisms from Φ to K when both are considered as AA's over K
$K[\Phi]$	The set of all polynomial functions on Φ (i.e., on $\mathrm{End}_K V$) with coefficients in K
$A \Subset B$	A is an ideal of B
$\Omega_\square$	A composition; e.g., it sends (a, b) into $a \square b$

III. CONCEPT OF ALGEBRAIC LIE ALGEBRAS

To define an algebraic Lie algebra we define first the concept of an algebraic group.

Def: Algebraic group (i.e., "algebraic linear group" sometimes): An algebraic group G is defined as a subgroup of Aut V (where V is a vector space over a field K) satisfying

$$\exists S \underset{\text{set}}{\subset} K[\Phi]: G = \{\eta \mid \eta \in \mathrm{Aut}\ V, \pi(\eta) = 0, \forall \pi \in S\}. \tag{2}$$

Remarks:

(i) S is called the *defining set* of G. S is obviously an ideal (w.r.t. associative multiplication!) of

$K[\Phi]$; therefore, S is also called the "ideal of polynomial functions associated with G."

(ii) It is obvious that an algebraic group is a Lie group.

(iii) $K[\Phi]$ has the structure of an associative algebra; $K[\Phi]$ is the associative algebra generated by $\mathrm{Hom}_K(\Phi, K)$ and the constant functions.

Def: $K[\Phi]$ *as a* Φ*-module (bimodule):*

(i) $K[\Phi]$ can be defined as a *left* Φ-module

$$\Omega_{\blacksquare}: \Phi \times K[\Phi] \to K[\Phi],$$

i.e.,

$$\Omega_{\blacksquare}: (\varphi, \pi) \mapsto \varphi \blacksquare \pi$$

defined by

$$(\varphi \blacksquare \pi)\phi = \pi(\phi \cdot \varphi), \quad \forall \phi \in \Phi, \tag{3}$$

where $\mapsto$ denotes element-wise mapping.

(ii) $K[\Phi]$ can be defined as a *right* Φ-module

$$\Omega_{\square}: P \times \Phi \to P,$$

i.e.,

$$\Omega_{\square}: (\pi, \varphi) \mapsto \pi \square \varphi$$

defined by

$$(\pi \square \varphi)\phi = \pi(\varphi \cdot \phi), \quad \phi \in \Phi. \tag{4}$$

Def: The derivation d *on* $K[\Phi]$. $d_\varphi, \forall \varphi \in \mathrm{End}_K V$, is defined as the *derivation* on $K[\Phi]$ satisfying

$$d_\varphi K = 0 \tag{5}$$

(K here means "constant functions") and

$$d_\varphi f = \varphi \blacksquare f, \quad \forall f \in \mathrm{Hom}_K(\Phi, K). \tag{6}$$

Proposition: Let

$$L_G \equiv \{\varphi \mid \varphi \in \mathrm{End}_K V, d_\varphi S \subset S\}; \tag{7}$$

then

$$[L_G]_{\mathrm{LA}} \underset{\mathrm{LA}}{\subset} [\mathrm{End}_K V]_{\mathrm{LA}},$$

where Lie algebra composition is the true bracket (i.e., $[\varphi, \varphi'] \equiv \varphi \cdot \varphi' - \varphi' \cdot \varphi$) and S is the defining set of some algebraic group G.

Def: L_G is called the "Lie algebra of the algebraic group G."

Def: a sub-Lie algebra, L, of $[\mathrm{End}_K V]_{\mathrm{LA}}$ is an algebraic Lie algebra if it is the Lie algebra of an algebraic group.

Remark: There are a number of equivalent ways of defining algebraic Lie algebra; for instance, one could define it in terms of the so-called "replica." [1–3]

[1] C. Chevalley, Am. J. Math. **65**, 521 (1943).
[2] C. Chevalley, Ann. Math. **48**, 91 (1947).
[3] Seminaires S. Lie, Ecole Normal Superior, 1954–55.

The reason for such an alternative, but equivalent, definition is due to the following "necessary-and-sufficient" type of theorem: Let

$$L \underset{\mathrm{LA}}{\subset} [\mathrm{End}_K V]_{\mathrm{LA}},$$

then L is *algebraic* $\Leftrightarrow$ every replica of any element of L belongs to L.

IV. SOME IMPORTANT THEOREMS ON ALGEBRAIC LIE ALGEBRA AND ALGEBRAIC GROUPS

Proposition 1: Let

$$X, X' \underset{\mathrm{VS}}{\subset} \mathrm{End}_K V$$

and

$$G = \{g \mid g \in \mathrm{End}_K V, (g^{-1}Xg - X) \in X' \text{ for any } x \in X\}, \tag{8}$$

then G is an algebraic subgroup of the Lie group $GL(n, K)$.

Remark: $[\mathrm{End}_K V]_{\mathrm{LA}} = gl(n, K)$ if $\dim V = n$ and if we use a basis for V. Here we write $gl(n, K)$ to denote the Lie algebra of $GL(n, K)$.

Proposition 2:

$$L_G = \{y \mid y \in \mathrm{End}_K V, [x, y] \in X', \text{any } x \in X\}, \tag{9}$$

where G means the Lie algebra of an algebraic group G.

Proposition 3: (char $K = 0$): Let

$$L \underset{\mathrm{LA}}{\subset} [\mathrm{End}_K V]_{\mathrm{LA}},$$

then

$$[G_a]_{\mathrm{LA}} \underset{\mathrm{LA}}{\subset} L, \quad \text{any } a \in L \Rightarrow L \text{ is algebraic}, \tag{10}$$

where G_a is the intersection of all algebraic subgroups of $GL(n, K)$ whose Lie algebras contain a. [G_a is the unique smallest algebraic subgroup of $GL(n, K)$ for char $K = 0$ (char $\equiv$ characteristic).]

Proposition 4: (char $K = 0$):

$$L \underset{\mathrm{LA}}{\subset} [\mathrm{End}_K V]_{\mathrm{LA}} \Rightarrow L^{(1)} \text{ is algebraic}, \tag{11}$$

where $L^{(1)} \equiv [L, L]$.

Proposition 5: (char $K = 0$):

$$\text{semisimple } L \underset{\mathrm{LA}}{\subset} [\mathrm{End}_K V]_{\mathrm{LA}} \Rightarrow L \text{ is algebraic}. \tag{12}$$

Proposition 6: Let $L_{i\in I}$ be a collection of subalgebraic Lie algebras of $[\mathrm{End}_K V]_{\mathrm{LA}}$, then

$$L \equiv \bigcap_{i\in I} L_i$$

is also a subalgebraic Lie algebra.

If G_i is an algebraic group of automorphisms of V such that L_i is the Lie algebra of G_i, $\forall i \in I$, then L is the Lie algebra of the group

$$G \equiv \bigcap_{i\in I} G_i.$$

Proposition 7: Let $\{L_i\}_{i\in I}$ be defined as above, then the Lie algebra *generated* by all the L_i, $i \in I$, is also algebraic.

Proposition 8: If L^K is a sub-Lie algebra of $\mathrm{End}_K V$, and if K' is a subfield of K, then

$$L^K \text{ is algebraic} \Rightarrow L^{K'} \text{ is algebraic.} \qquad (13)$$

Proposition 9: Let V be any nonassociative algebra, then the set Der V, of all the derivations on V, is a subalgebraic Lie algebra of $[\mathrm{End}_K V]_{\mathrm{LA}}$; i.e., Der V is the Lie algebra of the group Aut V.

Proposition 10: (char $K = 0$): Let $G_1, \cdots, G_m$ be closed connected subgroups of $GL(n, K)$ and let G be the smallest closed subgroup of $GL(n, K)$ such that

$$G \supset G_i, \quad i = 1, \cdots, m,$$

then G is the sub-Lie algebra of $[\mathrm{End}_K V]_{\mathrm{LA}}$ generated by G_i.

Remark: In spite of the formal definition, it can be shown that the Lie algebra of an algebraic group *coincides* with the ordinary definition of Lie algebra when the algebraic group is treated as an ordinary Lie group.[3]

V. SOME EXAMPLES ON ALGEBRAIC LINEAR GROUPS

(A) For any finite-dimensional vector space V, Aut V is an algebraic linear group.

Observations: This is obvious; we take the polynomial as the product of all characteristic polynomials for Aut V.

(B) For any fixed $h \in \Phi$, the set

$$G = \{g \mid g \in \mathrm{Aut}\ V, h \cdot g - g \cdot h = 0\} \qquad (14)$$

is a subgroup of Aut V, and is an algebraic linear group (remembering that we denote $\Phi \equiv \mathrm{End}_K V$).

Observations:

(i) That G is a group is obvious, since

$$h \cdot g_1 - g_1 \cdot h = 0, \quad \text{any } g_1 \in G,$$
$$h \cdot g_2 - g_2 \cdot h = 0, \quad \text{any } g_2 \in G;$$

thus,

$$\begin{aligned} h \cdot (g_1 \cdot g_2) &= (h \cdot g_1) \cdot g_2 \\ &= (g_1 \cdot h) \cdot g_2 \\ &= g_1 \cdot (h \cdot g_2) \\ &= g_1 \cdot (g_2 \cdot h) \\ &= (g_1 \cdot g_2) \cdot h, \end{aligned}$$

i.e., $h \cdot (g_1 \cdot g_2) - (g_1 \cdot g_2) \cdot h = 0$, thus $g_1 \cdot g_2 \in G$.

(ii) We now ask what is the defining set of G. To answer this we shall digress to discuss the following definitions:

Digression: Let dim $V = n$ and $\{v\} \equiv \{v_1, \cdots, v_n\}$ be a basis of V. Now we introduce a set of n^2 elements $g_{ij} \in \Phi$ such that $g_{ij}: v_k \mapsto v_i \delta_{jk}$, $i, j, k = 1, \cdots, n$. Then $h \in \Phi$ can be written as

$$h \equiv \sum_{i,j=1}^{n} a_{ij}(h) g_{ij}, \quad a_{ij}(h) \in K,$$

where $a_{ij}(h)$ is clearly a K-linear mapping, with

$$a_{ij}: \Phi \to K$$

defined by

$$a_{ij}: h \mapsto a_{ij}(h).$$

Def: The $a_{ij}(h)$, $i, j = 1, \cdots, n$, are called the *coordinates* of h (w.r.t. the basis $\{v\}$).

Def: The set of n^2 K-linear mappings, $\{a_{ij}\}$, $i, j = 1, \cdots, n$, is said to form *a system of coordinate functions.*

We now return to the question of the *defining set* of G. Let $\{a_{ij}\}$ be a system of coordinate functions on Φ. We know the mappings

$$g \mapsto a_{ij}(h \cdot g),$$
$$g \mapsto a_{ij}(g \cdot h), \quad \text{any } g \in \mathrm{Aut}\ V,$$

are obviously K-linear. Obviously, then, G is the set of all automorphisms given by all $g \in \mathrm{Aut}\ V$ satisfying:

$$a_{ij}(g \cdot h) - a_{ij}(h \cdot g) = 0.$$

Therefore, the set of all mappings

$$h \mapsto a_{ij}(g \cdot h) - a_{ij}(h \cdot g) = 0$$

is the "defining set" of G.

(C) For any given (fixed) $\pi \in K[\Phi]$, the set

$$G \equiv \{g \mid g \in \text{Aut } V, \pi(g \cdot h) = \pi(h), \text{ any } h \in \Phi\} \quad (15)$$

is an algebraic linear group.

Justification:

(i) For any $g_1, g_2 \in G$, any $h \in \Phi$,

$$\begin{aligned}\pi((g_1 \cdot g_2) \cdot h) &= \pi(g_1 \cdot (g_2 \cdot h)) \\ &= \pi(g_2 \cdot h) \\ &= \pi(h).\end{aligned}$$

(ii) We simply consider the mapping

$$\text{End}_K V \to K,$$

defined by

$$g \mapsto \pi(g \cdot h) - \pi(h) = 0,$$

which furnishes the defining set of G. (We use Φ and $\text{End}_K V$ as equivalent notations.)

(D) Let F be any polynomial function on V in K. Then, the set

$$G \equiv \{g \mid g \in \text{Aut } V, F(g(v)) = F(v), \text{ any } v \in V\} \quad (16)$$

is an algebraic linear group.

Justification:

(i) $$\begin{aligned}F((g_1 \cdot g_2)(v)) &= F(g_1(g_2(v))) \\ &= F(g_2(v)) \\ &= F(v), \quad \text{any } g_1, g_2 \in G.\end{aligned}$$

(ii) Simply consider the mapping

$$\text{End}_K V \to K,$$

defined by

$$g \mapsto F(g(v)) - F(v) = 0, \quad \text{any } v \in V.$$

(E) The special linear group, $SL(n, K)$, is algebraic.

(F) As an example of a nonalgebraic group, consider the 2-dimensional real vector space V with a basis $\{V_1, V_2\}$. Consider the set G of all K-endomorphisms of V describable in the form

$$\begin{pmatrix} a & 0 \\ 0 & a^t \end{pmatrix}$$

with real positive a and an irrational number t. In this case, G is not an algebraic group though G is a Lie group.

VI. EXAMPLES OF ALGEBRAIC LIE ALGEBRAS

These follow simply as the Lie algebras of algebraic groups; for example, the known algebraic Lie algebras are:

(1) $sl(n, K)$, with K being algebraically closed and of characteristic zero.

(2) $gl(n, K)$.

(3) All nilpotent Lie algebras as sub-Lie algebra of $gl(n, K)$.

(4) All semisimple Lie algebras over a field of characteristic zero (e.g., real or complex number fields).

(5) Representations of Lie algebras of diagonal matrices as Lie algebras of nilpotent matrices.

(6) Representations of semisimple Lie algebras as derivations of solvable Lie algebras.

VII. ON NOETHER RING, ORE CONDITION, AND QUOTIENTS

We summarize in this section some of the standard mathematical concepts in ring theory to be used later.

Def: Ring: A *ring* is an Abelian group and also a (noncommutative) multiplicative semigroup with or without a unit element.

Def: Noether ring: A ring R is (left) *Noetherian* if *every* chain of (left) ideals of R:

$$R_1 \subseteq R_2 \subseteq \cdots \quad (17)$$

terminates (i.e., $\exists$ an index n: $R_n = R_{n+1} = \cdots$).

Def: Ore condition: A ring R is said to satisfy the (left) *Ore condition* if for $\forall a, b \in R$ where b is a nonzero divisor (a *nonzero divisor* b is defined as $\nexists c \neq 0$, $c \in R$: $cb = 0$ or $bc = 0$), then

$$\exists a', b' \in R: b' \cdot a = a' \cdot b, \quad (18)$$

where b' is also a nonzero divisor.

Def: A ring satisfying (left) Ore condition is called a (left) *Ore ring*.

Now we can give the formal definition of quotients:

Def: Quotients associated with an Ore ring: A quotient associated with an Ore ring R is defined as an ordered pair (a, b) with $a, b \in R$ and b being a nonzero divisor, equipped with an *equivalence relation* defined by: two quotients (a, b) and (c, d) are said to be "equivalent" (it is easy to verify that all axioms of equivalence are satisfied) if

$$\exists \text{ nonzero } x, y: (xa, xb) = (yc, yd) \quad (19)$$

and we shall denote the (a, b) quotient by the notation

$$a^{-1}b. \quad (20)$$

Definition for right quotients is similar and we denote them by ab^{-1}.

Proposition 11: For any two left quotients $a^{-1}b$, $c^{-1}d$, defined in an Ore ring without zero divisor, we can find their "least common denominator" in the sense that

$$\exists x, y_1, y_2 \in R: a^{-1}b = x^{-1}y_1 \quad \text{and} \quad c^{-1}d = x^{-1}y_2. \tag{21}$$

Proof: For a, $c \in R$, the (left) Ore condition implies that

$$\exists t, s \in R: t \cdot a = s \cdot c, \tag{22}$$

i.e.,

$$c^{-1} \cdot s^{-1} = a^{-1} \cdot t^{-1}. \tag{23}$$

Next, we have

$$a^{-1} \cdot b = a^{-1} \cdot t^{-1} \cdot t \cdot b \tag{24}$$

and, similarly,

$$c^{-1} \cdot d = c^{-1} \cdot s^{-1} \cdot s \cdot d. \tag{25}$$

We set [using (23)]:

$$a^{-1} \cdot t^{-1} = c^{-1} \cdot s^{-1} \equiv x^{-1} \tag{26}$$

and

$$t \cdot b \equiv y_1, \quad s \cdot d \equiv y_2; \tag{27}$$

thus, substituting (26) and (27) into (24) and (25) gives

$$a^{-1}b = x^{-1}y_1$$

and

$$c^{-1}d = x^{-1}y_2.$$

Next, we note that if R is an Ore ring without zero divisor then an equivalence relation can be established between any given left quotient and some right quotient. This is quite obvious since R is Oreian, thus for any nonzero a, $b \in R$,

$$\exists c \in R \quad \text{and} \quad d\,(\neq 0) \in R: ca = db \tag{28}$$

which provides the equivalence relation

$$\overbrace{d^{-1}c}^{\text{(left quotient)}} = \overbrace{ba^{-1}}^{\text{(right quotient)}} \tag{29}$$

Def: The quotient field: Since any right quotient (i.e., of the form cd^{-1}) can be written as a left quotient, we shall only deal with left quotients. The set of all left quotients associated with an Ore ring without zero divisor forms a *noncommutative field* (i.e., a division ring) w.r.t. the compositions of addition, subtraction, multiplication, and division as defined below:

(1) *addition:*

$$a^{-1}b_1 + a^{-1}b_2 \equiv a^{-1}(b_1 + b_2), \tag{30}$$

(2) *subtraction:*

$$a^{-1}b_1 - a^{-1}b_2 \equiv a^{-1}(b_1 - b_2), \tag{31}$$

(3) *division:*

$$(a^{-1}b_1)^{-1} \cdot (a^{-1}b_2) \equiv b_1^{-1}b_2, \tag{32}$$

(4) *multiplication:*

$$(a_1^{-1}b_1) \cdot (a_2^{-1}b_2) \equiv (b_1^{-1}a_1)^{-1} \cdot (a_2^{-1}b_2). \tag{33}$$

VIII. ON FILTRATION AND GRADED STRUCTURE

Def: Filtration: Let R be a ring. If a chain of subgroups (w.r.t. the additive structure of the ring) of R,

$$R_0 \subset R_1 \subset \cdots, \tag{34}$$

satisfies the condition

$$\bigcup_{i=0}^{\infty} R_i = R, \tag{35}$$

then this chain is called an "increasing filtration" of R.

Def: Graded Ring: If a filtration of a ring R satisfies the condition

$$R_i \cdot R_j \subset R_{i+j}, \tag{36}$$

then the ring R is said to have a "graded ring" structure w.r.t. this filtration.

Def: gr R: For a graded ring, w.r.t. the filtration

$$R_0 \subset R_1 \subset \cdots, \tag{37}$$

we define

$$\text{gr}^{(i)}R \equiv R_i/R_{i-1}$$

with

$$R_i \equiv 0, \quad \text{for} \quad \forall i < 0, \tag{38}$$

and

$$\text{gr}\, R \equiv \sum_{i=0}^{\infty} \text{gr}^{(i)}R, \tag{39}$$

where the last expression signifies that $\forall \alpha \in \text{gr}\, R$ can be written as

$$\alpha = \sum_{i=0}^{\infty} \alpha_i, \quad \alpha_i \in \text{gr}^{(i)}R \tag{40}$$

with only a finite number of α_i being nonzero.

Now, let π_i denote the "canonical projection" mapping,

$$\pi_i: R_i \to \text{gr}^{(i)}R \tag{41}$$

defined by

$$\pi_i: a_i \mapsto a_i \bmod R_{i-1}, \tag{42}$$

then it follows from

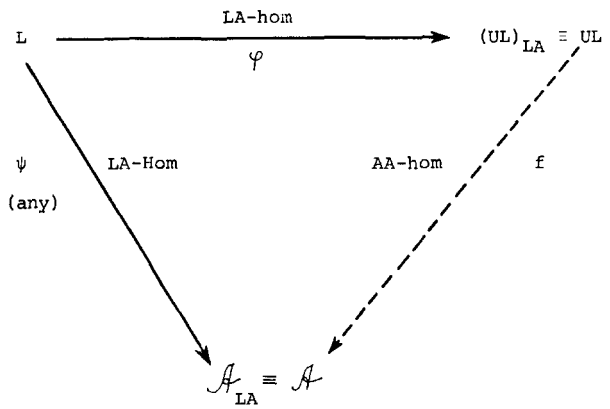
$$R_i \cdot R_j \subset R_{i+j} \tag{43}$$

that we have the mapping

$$\text{gr}^{(i)}R \times \text{gr}^{(j)}R \to \text{gr}^{(i+j)}R \tag{44}$$

which equips gr R with a "graded ring" structure.

Next, we mention briefly some properties of the *universal enveloping algebra* of a Lie algebra. First, we recall the universal enveloping algebra UL of a Lie algebra L (over a commutative field F) is defined as an associative algebra (over F) satisfying the commutative diagram

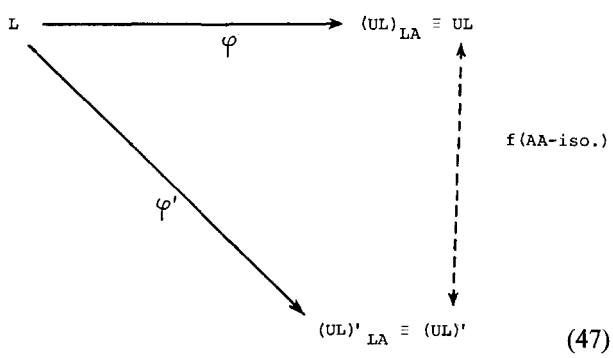


(45)

That is, for any given AA $\mathcal{A}$ and any LA-hom ψ, there exists a unique AA-hom f satisfying

$$f \circ \varphi = \psi. \tag{46}$$

One important property of UL is its *uniqueness up to isomorphism*; this is rather obvious because in the defining commutative diagram the $\mathcal{A}$ is arbitrary; therefore, we can draw



(47)

where $(UL)'$ is another universal enveloping algebra of L. Therefore, it is only necessary to look for *one* construction of UL. It can be easily shown that we can achieve this by setting

$$UL \equiv TL/IL, \tag{48}$$

where TL is the tensor algebra on L, i.e.,

$$TL = \bigoplus_{i=0}^{\infty} L_i \tag{49}$$

with (F is the ground field of L):

$$L_0 \equiv F1, \quad L_1 \equiv L, \quad L_2 \equiv L \otimes L, \quad \text{etc.}, \tag{50}$$

and IL is the *two-sided ideal* (associative) *generated* by all elements of the form

$$l \otimes l' - l' \otimes l - [l, l']. \tag{51}$$

Now we shall talk about filtration of UL.

Def: Filtration of UL: Let $(UL)_n$ be defined as a sub F-module (of UL) generated by the set of all $\varphi(l_1) \cdot \varphi(l_2) \cdots \varphi(l_j), j \leq n, \forall l_i \in L$. Thus, we have

$$\begin{aligned} (UL)_{-1} &= 0, \\ (UL)_0 &= F1, \\ (UL)_1 &= F1 \oplus \varphi(L), \\ (UL)_m &= F1 \oplus \sum_{j=1}^{m} \underbrace{\varphi(L) \cdots \varphi(L)}_{j \text{ copies}}, \end{aligned} \tag{52}$$

and it is obvious that

$$(UL)_0 \subset (UL)_1 \subset \cdots \tag{53}$$

which defines an *increasing filtration* of UL.

Define, now,

$$\text{gr}^{(n)}UL \equiv (UL)_n/(UL)_{n-1} \tag{54}$$

and define

$$\text{gr } UL \equiv \sum_{n=0}^{\infty} \text{gr}^{(n)}UL. \tag{55}$$

This newly constructed

$$\text{gr } UL \tag{56}$$

can be made into an *algebra* w.r.t. Ω_*:

$$\Omega_*: \text{gr}^{(m)}UL \times \text{gr}^{(n)}UL \to \text{gr}^{(m+n)}UL, \tag{57}$$

defined by means of passing to quotient from

$$(UL)_m \times (UL)_n \to (UL)_{m+n} \tag{58}$$

defined by

$$(u, u') \mapsto u \cdot u', \quad u \in (UL)_m, \quad u' \in (UL)_n, \tag{59}$$

i.e.,

$$\begin{aligned} \Omega_*: ((u \bmod (UL)_{m-1}), (u' \bmod (UL)_{n-1})) \\ \mapsto (u \bmod (UL)_{m-1}) * (u' \bmod (UL)_{n-1}) \\ \equiv u \cdot u' \bmod (UL)_{m+n-1}. \end{aligned} \tag{60}$$

The algebra gr UL (with Ω_*) is called the *graded algebra* associated with UL.

gr UL has the properties:
(i) gr UL is generated by φL;
(ii) gr UL is an Abelian associative algebra.

IX. THE ALGEBRA $A_n(R)$ GENERATED BY CANONICAL ELEMENTS

Let $A_n(Z)$ be the associative algebra generated over the ground field of integers Z by the set of canonical elements

$$\{p_1, \cdots, p_n, q_1, \cdots, q_n\} \equiv \{p, q\}, \tag{61}$$

satisfying

$$[p_i, q_j] \equiv p_i \cdot q_j - q_j \cdot p_i = \delta_{ij} 1_A, \tag{62}$$

$$[p_i, p_j] = 0, \quad [q_i, q_j] = 0, \tag{63}$$

where 1_A is the unit element of the algebra $A_n(Z)$.

It is obvious, the ground ring can be extended from Z to an arbitrary Ore ring R by tensor product in the usual way. That is, we construct

$$A_n(Z) \otimes_Z R \equiv A_n(R) \tag{64}$$

by considering $A_n(Z)$ as a right Z-module and R as a left Z-module.

The algebra $A_n(R)$ thus constructed has the following important property.

Proposition 12: The algebra $A_n(R)$ is a free R-module and it has a basis consisting of all the monomials of the form

$$(p_1)^{i_1} \cdots (p_n)^{i_n} (q_1)^{j_1} \cdots (q_n)^{j_n} \equiv p^{(i)} q^{(j)}, \tag{65}$$

where $(p_i)^0 = 1_A$ and $(q_i)^0 = 1_A$.

Proof: Our proof will be restricted to the case of R being a commutative field (i.e., commutative division ring) of characteristic zero, since this is the only case to be considered later on.

(A) First, we want to show that the monomials of the form

$$p^{(i)} q^{(j)} \tag{66}$$

generate the R-module $A_n(R)$. Let us introduce a filtration of $A_n(R)$,

$$[A_n(R)]_0 \subset [A_n(R)]_1 \subset [A_n(R)]_2 \subset \cdots, \tag{67}$$

defined by:

$$[A_n(R)]_0 \equiv R, \tag{68}$$

$$[A_n(R)]_i \equiv \text{the set of all elements, in } A_n(R), \text{ which can be written as polynomials on } \{p, q\} \text{ with coefficients in } R \text{ and of degree} \leq i. \tag{69}$$

We have, obviously,

$$A_n(R) = \bigcup_k [A_n(R)]_k \tag{70}$$

and we use mathematical induction on the index k to prove that the statement of (66) is applicable to all $[A_n(R)]_k$. The statement is obviously true for $k \equiv 0$. We assume now, by induction, that the statement is true for all $k < k_0$. But, by using relations (62) and (63),

$$[p_i, q_j] = \delta_{ij} 1_A,$$
$$[p_i, q_j] = 0, \quad [q_i, q_j] = 0,$$

we can always reduce any monomial of degree k_0 to a monomial of the form (65) (i.e., with all p's on the left side) modulo $[A_n(R)]_{k_0-2}$. This completes the induction.

Finally, since $A_n(R)$ is just the union of all $[A_n(R)]_k$, (66) is proved.

(B) Next, we can show that the different monomials of type (65) are R-linearly independent. This can be proved by contradiction; let

$$\sum_{(i,j)} a_{ij} p^{(i)} q^{(j)} = 0 \tag{71}$$

where $a_{ij} \in R$ and at least one of the a_{ij} is nonzero. Let us introduce a lexicographic ordering of (i, j) which represents the set

$$(i_1, \cdots, i_n, j_1, \cdots, j_n). \tag{72}$$

Let (s, t) be the greatest set (according to the lexicographic ordering) such that

$$a_{st} \neq 0. \tag{73}$$

Using the notation of "adjoint" mapping, we write

$$(\text{adj } \alpha)\beta \equiv [\alpha, \beta]. \tag{74}$$

By operating the quantity

$$\prod_{i=1} (\text{adj } q_i)^{s_i} \prod_{j=1} (\text{adj } p_j)^{t_j} \tag{75}$$

on the left of (71), we get

$$\left[\prod_{i=1} (\text{adj } q_i)^{s_i} \prod_{j=1} (\text{adj } p_j)^{t_j}\right] \sum_{(i,j)} a_{ij} p^{(i)} q^{(j)} = 0, \tag{76}$$

i.e.,

$$\prod_{i=1} (-1)^{s_i} (s_i!)(t_i!) a_{st} = 0. \tag{77}$$

But the lhs of (77) cannot be zero by nature of a_{st}, thus the contradiction. Q.E.D.

X. THE FIELDS OF QUOTIENTS ASSOCIATED WITH $A_n(R)$ AND WITH UL

Before discussing the fields of quotients we state two useful properties of the Noether ring; the proofs of them are given in the appendices.

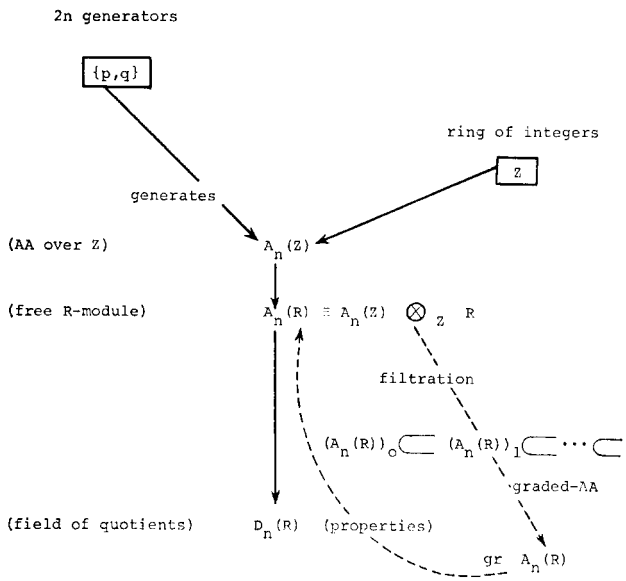


FIG. 1. Field of quotients $D_n(R)$.

Proposition 13: Let R be a ring. If gr R is a (left) Noether ring without zero-divisors then R is also a (left) Noether ring without zero-divisors (see Appendix A).

Proposition 14: A (left) Noether ring without zero-divisors satisfies the (left) Ore condition (see Appendix B).

Proposition 15: $A_n(R)$ is an Ore ring.

Proof: Consider the filtration given by Eqs. (67)–(70) and introduce its corresponding graded-ring structure

$$\text{gr}^{(i)} A_n(R) \equiv [A_n(R)]_i / [A_n(R)]_{i-1} \tag{78}$$

and

$$\text{gr } A_n(R) = \sum_{i=0}^{\infty} \text{gr}^{(i)} A_n(R). \tag{79}$$

By Proposition 12, we see

$$\text{gr } A_n(R) \leftrightarrow R[p, q], \tag{80}$$

where $R[p, q]$ denotes the polynomial ring of (p, q). But $R[p, q]$ is obviously a Noether ring without zero-divisors (because it is a polynomial ring); therefore, due to isomorphism, gr $A_n(R)$ is also a Noether ring without zero-divisors. Then, by Propositions 13 and 14, we know that $A_n(R)$ is an Ore ring without zero-divisors. Q.E.D.

The above property permits us to construct the associated *field of quotients*, to be denoted by $D_n(R)$.

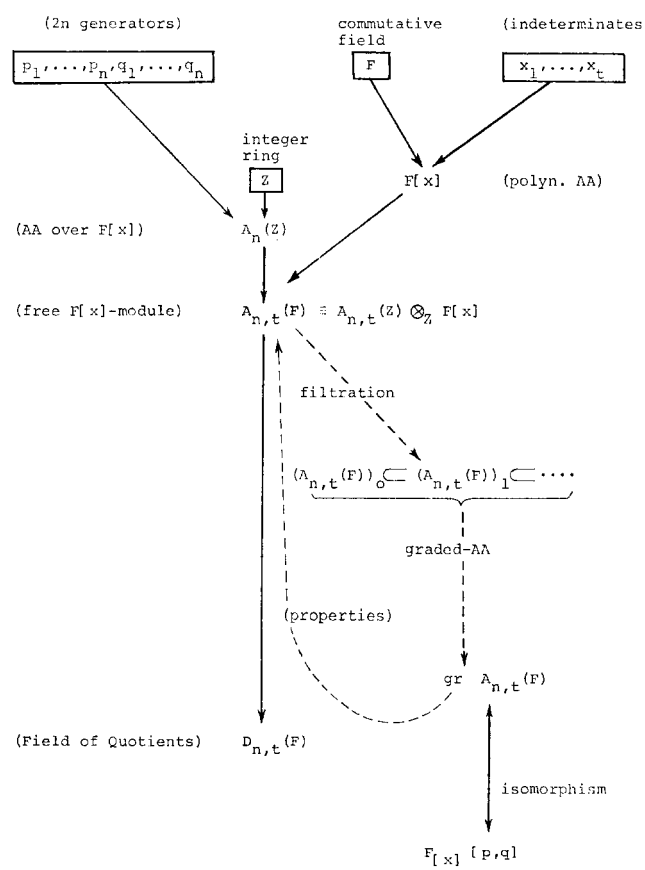


FIG. 2. Field of quotients $D_{n,t}(F)$.

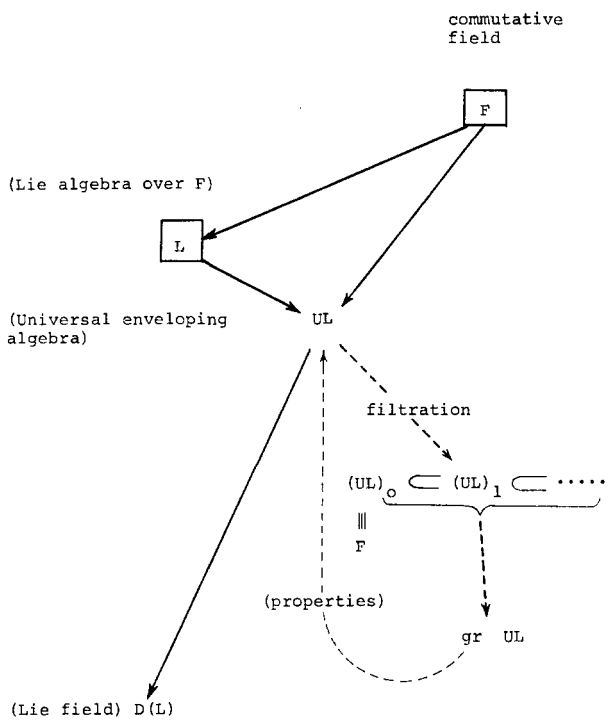


FIG. 3. Lie field $D(L)$.

In particular, if the ground ring R is taken to be the polynomial algebra $F[x]$ on the set of indeterminates

$$\{x_1, \cdots, x_r\} \equiv x$$

with coefficients in a commutative field F, then we have $A_n(F[x])$:

$A_{n,t}(F) \equiv A_n(F[x])$,

$D_{n,t}(F) \equiv$ field of quotients associated with $A_{n,t}(F)$

(see Figs. 1 and 2).

Next, let us discuss the field of quotients associated with UL, where L is a Lie algebra over a commutative field F. The increasing filtration on which we define the graded structure, gr UL, was given by (52) and (53). The Poincaré–Birkoff–Witt theorem says that gr UL can be expressed as the ring (or rather an associative algebra) of all polynomials on m variables ($m = \dim L$). Therefore, gr UL is clearly a (two-sided) Noether ring without zero-divisors. By Propositions 13 and 14, we conclude that UL is a (two-sided) Ore ring without zero divisors. This permits us to construct the *field of quotients* from UL; this field of quotients will be referred to as *Lie field* of L:

$$D(L) = \text{Lie field of } L$$

(see Fig. 3).

XI. A CONCEPT OF DIMENSION

We shall introduce some different concepts about dimensions in the sense that they are of invariant nature but differ from the *usual definition* which will be denoted by "dim" to avoid confusion. We shall define the new concepts for the case of an *associative algebra* A over a commutative field F and for the case of a *field* D (a division ring) over a commutative field F. First, let α be any finite set,

$$\alpha \equiv \{a_1, \cdots, a_s\}, \quad \forall a_i \in A, \tag{81}$$

and

$$(\alpha; N) \equiv \{p \mid p \in F[\alpha], p \in A, \dim p \leq N\}, \tag{82}$$

with $F[\alpha]$ being the ring of all the noncommutative polynomials on α. Now we define

$$\operatorname{Dim}_F A \equiv \operatorname*{Sup}_{\alpha} \overline{\lim_{N\to\infty}} \frac{\ln [\dim (\alpha; N)]}{\ln N} \tag{83}$$

where $\overline{\lim}$ refers to the *least upper bound.*

Similarly, for the case of a division ring D we form any finite set

$$\alpha \equiv \{a_1, \cdots, a_s\}, \quad \forall a_i \in D, \tag{84}$$

and form also

$$\alpha b \equiv \{a_1 \cdot b, a_2 \cdot b, \cdots, a_s b\}, \quad \forall b \in D, \quad b \neq 0; \tag{85}$$

then we define

$$\operatorname{Dim}_F D \equiv \operatorname*{Sup}_{\alpha} \operatorname*{Inf}_{b} \overline{\lim_{N\to\infty}} \frac{\ln [\dim (\alpha b; N)]}{\ln N}. \tag{86}$$

As an illustration of the definitions we shall discuss two examples which are also useful.

Proposition 16:

$$\operatorname{Dim}_F A_{n,t}(F) = 2n + t. \tag{87}$$

Proof: As in the definition of Dim, we consider a set

$$\alpha \equiv \{a_1, \cdots, s_s\}, \quad \forall a_i \in A_{n,t}(F). \tag{88}$$

In virtue of definition of $A_{n,t}(F)$, the set

$$\beta \equiv \{p_1, \cdots, p_n, q_1, \cdots, q_n, x_1, \cdots, x_t\} \tag{89}$$

is the set of generators, with F as the ground field (instead of $F[x]$) for $A_{n,t}(F)$. Thus, any element a_i of (88) can be written into the form of a polynomial on β. Let m be the highest degree among these polynomials, then obviously

$$\dim (\alpha; N) \leq \dim (\beta; mN). \tag{90}$$

Next, we know the dimension of the subspace of all polynomials of degree $\leq N$ on j indeterminates is just

$$\sum_{i=0} C^i_{i+j-1} = C^j_{N+j}. \tag{91}$$

This result in fact also applies to polynomials on β, even the polynomials are now noncommutative. The reason for this is that any two of the generators (from β) which do not commute with each other are always pairs like $\{p_i, q_i\}$; in this case we have

$$p_i q_i = q_i p_i + 1 \tag{92}$$

which does not change the dimension of the subspace since the appearance of 1 on the rhs of (92) only give rise to an extra term of *lower* degree w.r.t. the generators.

Therefore, (91) applies, that is

$$\dim(\beta; N) = C^{2n+t}_{N+(2n+t)}, \tag{93}$$

$$\begin{aligned}
\mathrm{Dim}_F A_{n,t}(F) &\equiv \operatorname*{Sup}_{\beta} \overline{\lim_{N\to\infty}} \frac{\ln(\dim(\beta; N))}{\ln N} \\
&\geq \lim_{N\to\infty} \frac{\ln C^{2n+t}_{N+(2n+t)}}{\ln N} \\
&= \lim_{N\to\infty} \left(\frac{1}{\ln N} \cdot \ln \frac{(N+2n+t)(N+2n+t-1)\cdots(N+1)}{(2n+t)(2n+t-1)\cdots 1}\right) \\
&= \lim_{N\to\infty} \left\{\frac{1}{\ln N}\left[\ln\left(1+\frac{N}{2n+t}\right) + \ln\left(1+\frac{N}{2n+t-1}\right) + \cdots + \ln(1+N)\right]\right\} \\
&= \lim_{N\to\infty} \left[\frac{1}{\ln N}\left(\ln\frac{N}{2n+t} + \ln\frac{N}{2n+t-1} + \cdots + \ln N\right)\right] \\
&= \lim_{N\to\infty} \frac{1}{\ln N}\left(\ln\frac{N^{2n+t}}{(2n+t)!}\right) \\
&= \lim_{N\to\infty} \frac{1}{\ln N}\{(2n+t)\cdot \ln N - \ln[(2n+t)!]\} \\
&= \lim_{N\to\infty} \frac{(2n+t)\ln N}{\ln N} \\
&= 2n+t.
\end{aligned} \tag{94}$$

On the other hand, from (90) and the definition of Dim,

$$\begin{aligned}
\mathrm{Dim}_F A_{n,t}(F) &\equiv \operatorname*{Sup}_{\alpha} \overline{\lim_{N\to\infty}} \frac{\ln(\dim(\alpha; N))}{\ln N} \\
&\leq \lim_{N\to\infty} \frac{\ln(\dim(\beta; mN))}{\ln N} \\
&= \lim_{N\to\infty} \frac{C^{2n+t}_{mN+(2n+t)}}{\ln N} \\
&= 2n+t.
\end{aligned} \tag{95}$$

Finally, (94) and (95) together imply

$$2n + t \leq \mathrm{Dim}_F A_{n,t}(F) \leq 2n + t, \tag{96}$$

therefore,

$$\mathrm{Dim}_F A_{n,t}(F) = 2n + t. \quad \text{Q.E.D.} \tag{97}$$

A property similar to Proposition 16 can be proved for the field of quotients $D_{n,t}(F)$ but we shall introduce the concept of "leading term" of a nonzero element $h \in A_{n,t}(F)$.

Def: "leading term": For any given nonzero $h \in A_{n,t}(F)$, it is obvious that

$$\underset{1!}{\exists} \text{ integer } i\colon \pi_i(h) \in \mathrm{gr}^{(i)} A_{n,t}(F), \tag{98}$$

with $\pi_i(h) \neq 0$ and $\pi_i(h)$ being well defined. We recall that the canonical projection π_i [defined by (41)] maps

$$\pi_i\colon h \mapsto h \bmod [A_{n,t}(F)]_{i-1}, \tag{99}$$

i.e.,

$$\pi_i(h) = h \bmod [A_{n,t}(F)]_{i-1} \tag{100}$$

and we call $\pi_i(h)$ the "leading part" of h, and it will be denoted by $\hat{h}$.

We have the following properties of leading terms:

Proposition 17:

(i) $\hat{h}$ is a homogeneous polynomial on p and q with coefficients in $F[x]$ for any $h \in A_{n,t}(F)$.

(ii) For any $h_1, h_2 \in A_{n,t}(F)$,

$$\widehat{h_1 \cdot h_2} = \hat{h}_1 \cdot \hat{h}_2. \tag{101}$$

Proof: It is obvious.

Proposition 18: For $d \in D_{n,t}(F)$, we can write d into the form (not unique because of equivalence relation):

$$h_1^{-1} \cdot h_2, \quad h_i \in A_{n,t}(F); \tag{102}$$

then, the rational function

$$\hat{h}_2/\hat{h}_1 \tag{103}$$

depends only on the element d and not upon the way d is cast into the quotients.

Proof: Let $h_3^{-1} \cdot h_4$ be an alternative quotient form for d, then it implies

$$h_1^{-1} \cdot h_2 = h_3^{-1} \cdot h_4. \tag{104}$$

Taking their leading terms,

$$(h_1^{-1} \cdot h_2)\hat{} = (h_3^{-1} \cdot h_4)\hat{},$$

i.e.,

$$\widehat{(h_1^{-1})} \cdot \hat{h}_2 = \widehat{(h_3^{-2})} \cdot \hat{h}_4,$$

i.e.,

$$\hat{h}_2/\hat{h}_1 = \hat{h}_4/\hat{h}_3, \tag{105}$$

where (101) was used. Equation (105) proved the proposition. Q.E.D.

Proposition 19:

(i) For $d_1, d_2 \in D_{n,t}(F)$,

$$\widehat{d_1 \cdot d_2} = \hat{d}_1 \cdot \hat{d}_2. \tag{106}$$

(ii) The function d is invariant under inner automorphisms of the field; i.e.,

$$(d \cdot d' \cdot d^{-1})\hat{} = \hat{d}' \tag{107}$$

Proof:

(i) We write d_1 and d_2 into the quotient forms $h_1^{-1} \cdot h$ and $h_2^{-1} \cdot h'$, respectively. Then,

$$\begin{aligned}\widehat{d_1 \cdot d_2} &= ((h_1^{-1} \cdot h) \cdot (h_2^{-1} \cdot h'))\hat{} \\ &= (h_1^{-1} \cdot h)\hat{} \cdot (h_2^{-1} \cdot h')\hat{} \\ &= \hat{d}_1 \cdot \hat{d}_2,\end{aligned}$$

where we used (101).

(ii) We have

$$(d \cdot d' \cdot d^{-1})\hat{} = \hat{d} \cdot \hat{d}' \cdot (\hat{d})^{-1} = \hat{d}'. \quad \text{Q.E.D.}$$

Lemma: For any nonzero $b \in D_{n,t}(F)$, the set $\mathfrak{T}(b)$ of all the distinct monomials

$$\begin{aligned}&(p_1 \cdot b)^{i_1} \cdots (p_n \cdot b)^{i_n} \cdot (q_1 \cdot b)^{j_1} \cdots \\ &\qquad (q_n \cdot b)^{j_n} \cdot (x_1 \cdot b)^{k_1} \cdots (x_t \cdot b)^{k_t} \\ &\equiv P(i, j, k; b) \equiv (p \cdot b)^i \cdot (q \cdot b)^j \cdot (x \cdot b)^k\end{aligned} \tag{108}$$

is linearly dependent only if

$$\text{degree } (\hat{b}) = -1. \tag{109}$$

Proof: We prove this by contradiction. First, consider

$$\text{degree } (\hat{b}) > -1 \tag{110}$$

and assume a linear dependence among $P(i, j, k; b)$'s, then

$$\sum_{(i,j,k)} C_{ijk} P(i, j, k; b) = 0, \tag{111}$$

where C_{ijk} are coefficients from F and

$$(i, j, k) \equiv (i_1, \cdots, i_n, j_1, \cdots, j_n, k_1, \cdots, k_n). \tag{112}$$

For convenience, hereafter, we denote the set of all distinct $P(i, j, k; b)$ by $\mathfrak{T}(b)$, then

$$\exists w \in A_{n,t}(F): P(i, j, k; b) \cdot w \in A_{n,t}(F) \quad \text{for} \quad \forall P(i, j, k; b) \in \mathfrak{T}(b). \tag{113}$$

From (111), we have:

$$\sum_{(i,j,k)} C_{ijk} P(i, j, k; b) \cdot w = 0. \tag{114}$$

Taking the "leading part" we get

$$\sum C_{ijk} (P(i, j, k; b) \cdot w)\hat{} = 0, \tag{115}$$

where the sum is now carried over all (i, j, k) such that

$$\text{degree } P(i, j, k; b) = \max \{\text{degree } P(i, j, k; b) \text{ with nonzero } C_{ijk}\}. \tag{116}$$

On the other hand, explicitly,

$$P(i, j, k; b) \cdot w = (p \cdot b)^i \cdot (q \cdot b)^j \cdot (x \cdot b)^k \cdot w \tag{117}$$

whose "leading part" is

$$(P(i, j, k; b) \cdot w)\hat{} = p^i q^j x^k (\hat{b})^d \hat{w} \tag{118}$$

with

$$d \equiv \sum_{h=1}^{n} i_h + j_h + \sum_{h=1}^{t} k_h. \tag{119}$$

In view of (118), it is clear that (115) is *impossible* except if the C_{ijk}'s are *all* zero. Further, we know the mapping

$$u \mapsto \hat{u}, \quad \forall u \in D_{n,t}(F)$$

preserves the "multiplication" composition of the division ring; therefore $\mathfrak{T}(b)$ is a linearly independent set if

$$\text{degree } (\hat{b}) > -1.$$

Similarly, we can prove this for the case of

$$\text{degree } (\hat{b}) < -1. \quad \text{Q.E.D.}$$

Proposition 20:

$$\mathrm{Dim}_F\, D_{n,t}(F) = 2n + t. \tag{120}$$

Proof: First, we proceed as in the proof of Proposition 16; consider an arbitrary, finite set

$$\alpha \equiv \{a_1, \cdots, a_s\}, \quad a_i \in D_{n,t}(F). \tag{121}$$

The same argument used in the proof of Proposition 16 is applied here to establish

$$\mathrm{Dim}_F\, D_{n,t}(F) \leq 2n + t. \tag{122}$$

Next, we shall prove

$$\mathrm{Dim}_F\, D_{n,t}(F) \geq 2n + t \tag{123}$$

so as to establish the equality.

By the preceding lemma, for any nonzero h, $b \in D_{n,t}(F)$, the set $\mathfrak{S}(h \cdot b)$ of all distinct monomials

$$P(i, j, k; h \cdot b)$$

is linearly dependent *only if*

$$\text{degree } \widehat{(h \cdot b)} = -1. \tag{124}$$

Thus, if

$$\text{degree } \hat{h} \neq 0, \tag{125}$$

then clearly the two conditions (124) and (109) are *not compatible.* This means either $\mathfrak{S}(b)$ or $\mathfrak{S}(h \cdot b)$ has to be a set of linearly independent monomials. Using this fact, and considering the set

$$\gamma \equiv \beta \cup (\beta \cdot h), \quad \beta \equiv \{p, q, x\}, \tag{126}$$

we obtain easily (following the same argument as in the proof of Proposition 16):

$$\dim(\gamma b; N) \geq C^{2n+t}_{N+2n+t} \tag{127}$$

which leads to

$$\mathrm{Dim}_F\, D_{n,t}(F) \geq 2n + t. \qquad \text{Q.E.D.}$$

Proposition 21: If A is an associative algebra over a commutative division ring (field) F with the property that there exists an increasing filtration of A,

$$A_0\,(\equiv F) \subset A_1 \subset \cdots \tag{128}$$

with

$$A_i \cdot A_j \subset A_{i+j}, \tag{129}$$

such that its corresponding graded algebra

$$\mathrm{gr}\, A \equiv \sum_{i=0}^{\infty} A_i / A_{i-1} \tag{130}$$

is isomorphic to the algebra of polynomials in t indeterminates; then

$$\mathrm{Dim}_F\,(D(A)) = t. \tag{131}$$

Proof: Proof follows that of Proposition 20.

Proposition 22:

$$\text{(i)}\quad A_{n,t}(F) \leftrightarrow A_{n',t'}(F) \Leftrightarrow n = n', \quad t = t', \tag{132}$$

$$\text{(ii)}\quad D_{n,t}(F) \leftrightarrow D_{n',t'}(F) \Leftrightarrow n = n', \quad t = t'. \tag{133}$$

Proof: We know, by definition of $A_{n,t}(F)$, that

$$\mathrm{Cen}\,(A_{n,t}(F)) = A_{0,t}(F)$$

and

$$\mathrm{Cen}\,(A_{n',t'}(F)) = A_{0,t'}(F),$$

where Cen denotes the *center.*

An isomorphism between $A_{n,t}$ and $A_{n',t'}$ implies trivially

$$\mathrm{Cen}\,(A_{n,t}(F)) \leftrightarrow \mathrm{Cen}\,(A_{n',t'}(F))$$

i.e.,

$$A_{0,t}(F) \leftrightarrow A_{0,t'}(F)$$

which is possible if and only if

$$\mathrm{Dim}_F\,(A_{0,t}(F)) = \mathrm{Dim}_F\,(A_{0,t'}(F)),$$

i.e.,

$$t = t'. \tag{134}$$

On the other hand,

$$A_{n,t}(F) \leftrightarrow A_{n',t'}(F)$$

requires necessarily

$$\mathrm{Dim}_F\,(A_{n,t}(F)) = \mathrm{Dim}_F\,(A_{n',t'}(F)),$$

i.e.,

$$2n + t = 2n' + t'.$$

Equation (134) and above imply

$$n = n' \quad \text{and} \quad t = t'. \qquad \text{Q.E.D.}$$

Remark: Proposition 22 gives the important result that no two $A_{n,t}(F)$ and $A_{n',t'}(F)$ are isomorphic unless $n = n'$ and $t = t'$, and similarly for $D_{n,t}(F)$.

Proposition 23: Let L be a Lie algebra over a commutative field F, then

$$\mathrm{Dim}_F\,(D(L)) = \dim L, \tag{135}$$

where "dim" is the *ordinary* dimension and "Dim_F" is the one defined previously by (86).

Proof: The following facts:

(i) UL is an associative algebra over F,

(ii) UL has the graded structure given by (52),

(iii) According to Poincaré–Birkhoff–Witt theorem, the graded algebra gr UL is the algebra of all polynomials on m indeterminates (or "variables," $m = \dim L$), allow us to use Proposition 21, therefore, according to (131),

$$\mathrm{Dim}_F\,(D(L)) = \dim L. \qquad \text{Q.E.D.} \tag{136}$$

XII. THE GEL'FAND-KIRILLOV CONJECTURE

We first note a few properties concerning the concept of *orbit* before the statement of the Gel'fand-Kirillov conjecture. The concept of orbit is well known; it is defined for a manifold M acted upon by a transformation group G, as:

$$\begin{aligned} \mathcal{O}_v &\equiv \text{orbit of a point } v \in M \\ &= \{z \mid z \in M, \exists g \in G: gv = z\}. \end{aligned} \tag{137}$$

A connection between orbits and representations was very extensively explored by Kirillov.[4] We note here an important property:

Proposition 24: For an algebraic Lie algebra L,

$$\dim \Omega(L) = \text{even}, \tag{138}$$

where $\Omega(L)$ denotes an orbit of general position in the vector space dual to the adjoint representation of the Lie group G_L whose Lie algebra is L.

Proof: By dual space here we mean the representation $\bar{\gamma}$ of G_L in the vector space of all real linear functionals on L dual to the adjoint representation γ.

Introducing the bilinear form B_f on L w.r.t. a given functional f,

$$B_f: L \times L \to F, \tag{139}$$

given by

$$\begin{aligned} B_f: \{x, y\} \mapsto B_f(x, y) &\equiv (f, [x, y]) \in F, \\ \text{for} \quad \forall x, \quad y \in L, \quad f \in \Omega. \end{aligned} \tag{140}$$

First, an orbit Ω containing the functional f can be expressed in the form of the factor group,

$$G_L/G_f \tag{141}$$

where G_f is the *stability group* of $f \in \Omega$. This is rather obvious because, by definition,

$$G_f \equiv \{g \mid \bar{\gamma}(g)f = f, g \in G_L\}; \tag{142}$$

thus,

$$G_L \bmod G_f \tag{143}$$

gives all the other transformations that send f to all those elements, of the dual space, of the orbit Ω. It is easy to see that the Lie algebra L_f of the Lie group G_f is simply the set of

$$\{x \mid x \in L, B_f(x, y) = 0, \text{ for all } y \in L\}, \tag{144}$$

because, corresponding to the action of stability group, we have its Lie algebra counterpart as "adjoint" mapping

$$(\text{adj } x)y = [x, y] = 0. \tag{145}$$

[4] A. A. Kirillov, Dokl. Akad. Nauk SSSR **128**, 886 (1959); **130**, 966 (1960); **138**, 283 (1961) [Sov. Math. **2**, 588 (1961)].

Therefore, in the language of Lie algebra, we say the subspace L_f is *orthogonal* to L w.r.t. the bilinear form B_f; thus,

$$\dim L_f = \dim L - \text{rank } B_f, \tag{146}$$

but, by (141),

$$\begin{aligned} \dim \Omega &= \dim G_L - \dim G_f, \\ \dim \Omega &= \dim L - \dim L_f. \end{aligned} \tag{147}$$

We get finally that

$$\dim \Omega = \text{rank } B_f. \tag{148}$$

But B_f is obviously antisymmetric, remembering that

$$B_f(x, y) \equiv (f, [x, y]), \tag{149}$$

thus the rank of B_f must be even, i.e.,

$$\dim \Omega = \text{even}. \quad \text{Q.E.D.} \tag{150}$$

Remark: We also note for an algebraic Lie algebra

$$\text{codim } \Omega(L) = \text{Dim}_F (\text{Cen } D(L)). \tag{151}$$

where "codim" means codimension of the sub vector space.

Now we are in a position to state the Gel'fand-Kirillov conjecture.

Gel'fand-Kirillov Conjecture[5]*:*

If L is an *algebraic* Lie algebra over a commutative field F, then we have the following unique isomorphism:

$$D(L) \leftrightarrow D_{n,t}(F), \tag{152}$$

where

$$n \equiv \tfrac{1}{2}(\dim L - \text{codim } \Omega(L)) = \tfrac{1}{2} \dim \Omega(L) \tag{153}$$

and

$$t = \text{codim } \Omega(L). \tag{154}$$

Discussions:

(i) Equation (153) is justified to be meaningful due to Proposition 24.

(ii) The conjecture was verified by Gel'fand and Kirillov for the following special cases:

(a) L is any nilpotent Lie algebra over an algebraically closed F of characteristic zero,
(b) L is the Lie algebra of $GL(n, F)$,
(c) L is the Lie algebra of $SL(n, F)$,

[5] I. M. Gel'fand and A. A. Kirillov, Dokl. Akad. Nauk SSSR **167**, 503 (1966) [Sov. Math. **7**, 403 (1966)]; Preprint, V. A. Steyklova Institute of Mathematics, Akad. Nauk SSSR, 1965 [French transl.: Publ. Math. France **31** (1966)]. In regard to the latter, I am grateful to two persons: Professor A. Böhm, for helping in the translation from the original, and Professor H. Bacry, for making a prepublication copy of the French version available to me.

(d) L is a semisimple Lie algebra of rank two (no full detail; the only indication of proof is mentioned in Ref. 5).

XIII. PROOF OF THE CONJECTURE FOR LIE ALGEBRAS $gl(m, C)$ AND $sl(m, C)$

The conjecture can be proved for these two cases and, in fact, the proofs are similar. To facilitate the discussion, we prove first the following property:

Proposition 25: Let $gl(m, C)_0$ be the Lie algebra of all $m \times m$ matrices with last rows consisting of zeros only; then

$$D(gl(m, C)_0) \leftrightarrow D_{\frac{1}{2}m(m-1),0}(C). \tag{155}$$

Proof: Let us first introduce a basis for $gl(m+1, C)_0$:

$$e_{ik}, \quad i = 1, \cdots, m, \quad k = 1, \cdots, m+1, \tag{156}$$

where e_{ik} has unity as the element of ith row and kth column and zero everywhere else. We can show that the following choice of canonical generators is the correct one:

$$q_i \equiv e_{i,m+1}, \quad i = 1, \cdots, m, \tag{157}$$

$$p_i \equiv e_{ii}q_i^{-1}, \quad i = 1, \cdots, m. \tag{158}$$

We note that q_i^{-1} is not defined in the ordinary sense as a matrix (since $\det q_i = 0$), but it is defined as a formal quantity whose rules of manipulation are conformed to the equivalence relation of quotients by Ore condition (cf. Sec. VII).

First, we have

$$\begin{aligned}[q_i, q_j] &= q_iq_j - q_jq_i \\ &= e_{i,m+1}e_{j,m+1} - e_{j,m+1}e_{i,m+1} \\ &= 0 \end{aligned} \tag{159}$$

(by straight matrix multiplications). This also implies, by definition of quotients,

$$q_jq_i^{-1} = q_i^{-1}q_j, \tag{160}$$

that is, q_i^{-1} and q_j commute.

Next, let us compute

$$\begin{aligned}[p_i, q_j] &= p_iq_j - q_jp_i \\ &= e_{ii}q_i^{-1}q_j - q_je_{ii}q_i^{-1} \\ &= e_{ii}q_jq_i^{-1} - q_ie_{ii}q_i^{-1} \\ &= e_{ii}q_jq_i^{-1} - \underbrace{e_{j,m+1}e_{ii}}_{0}q_i^{-1} \\ &= e_{ii}q_jq_i^{-1} \\ &= \delta_{ij}1, \end{aligned} \tag{161}$$

where Eq. (160) was used in obtaining the third equality.

Similarly, we can verify

$$[p_i, p_j] = 0, \tag{162}$$

$$[b_{ik}, q_j] = \delta_{kj}q_j, \tag{163}$$

and

$$[b_{ik}, p_j] = \delta_{kj}p_j, \quad i, j, k = 1, \cdots, m, \tag{164}$$

where

$$b_{ik} \equiv e_{ik}q_i^{-1}q_k. \tag{165}$$

Next, we introduce a matrix H whose entries satisfy the condition

$$\sum_i h_{ij} = 0, \quad j = 1, \cdots, m, \tag{166}$$

and denote the set of all such matrices by $\mathcal{H}$. It can be verified that the mapping

$$\alpha: \mathcal{H} \to D(gl(m+1, C)_0)$$

defined by

$$\alpha: H \mapsto \sum_{i,j} h_{ij}b_{ij} \tag{167}$$

satisfies the condition

$$\alpha([H, H']) = [\alpha(H), \alpha(H')] \quad \text{for} \quad \forall H, H' \in \mathcal{H}. \tag{168}$$

It is easy to see that

$$[\alpha(H), q_i] = 0, \quad i = 1, \cdots, m, \tag{169}$$

$$[\alpha(H), p_i] = 0, \quad i = 1, \cdots, m, \tag{170}$$

since we have, from (167),

$$\begin{aligned}[\alpha(H), q_i] &= \sum_{j,k} h_{jk}[b_{jk}, q_i] \\ &= \sum_{j,k} h_{jk}\delta_{ki}q_i \\ &= \sum_j h_{ji}q_i, \end{aligned} \tag{171}$$

where (166) and (163) were used. Similarly, we get

$$\begin{aligned}[\alpha(H), p_i] &= \sum_j h_{ji}p_i \\ &= 0, \end{aligned} \tag{172}$$

where (164) and (166) were used.

It is clear that the Lie field

$$D(gl(m+1, C)_0)$$

is generated by

$$\{\alpha(\mathcal{H}), p_1, \cdots, p_m, q_1, \cdots, q_m\}. \tag{173}$$

Now we can use mathematical induction on the number m; i.e., we assume the lemma is true for m and want to show it is also true for $m+1$. But this is not difficult to see since we can easily show

$$\mathcal{H} \leftrightarrow gl(m, C)_0, \tag{174}$$

because $\mathcal{H}$ consists of all $m \times m$ matrices such that

the m conditions

$$\sum_i h_{ij} = 0, \quad j = 1, \cdots, m \tag{175}$$

are satisfied. The isomorphism is, in fact, an obvious one; we simply choose the correspondence

$$h_{ij} = M_{ij}, \quad i = 1, \cdots, m-1, \quad j = 1, \cdots, m, \tag{176}$$

and

$$h_{mj} = -\sum_{i=1}^{m-1} M_{ij}, \quad j = 1, \cdots, m, \tag{177}$$

where M_{ij} are matrix entries of any $M \in gl(m, C)_0$.

By mathematical induction we assume now

$$D(gl(m, C)_0) \leftrightarrow D_{\frac{1}{2}m(m-1),0}(C). \tag{178}$$

With (173), (174), and the fact that

$$\alpha: U\mathfrak{K} \to D(gl(m+1, C)_0) \tag{179}$$

is injective (which can be proved by choosing a base B' for $\mathfrak{K}$ such that $\alpha(B') = \{b'_{ij}\}$ with $b'_{ij} \equiv b_{ij} - e_{im}$, $i = 1, \cdots, m$, and $j = 1, \cdots, m-1$), then we get

$$D(gl(m+1, C)_0) \leftrightarrow D_{\frac{1}{2}m(m-1)+m,0}(C) = D_{\frac{1}{2}m(m+1),0}(C),$$

which completes the induction. Q.E.D.

Proposition 26:

$$D(gl(m, C)) = D_{\frac{1}{2}m(m-1),m}(C), \tag{180}$$

$$D(sl(m, C)) = D_{\frac{1}{2}m(m-1),m-1}(C). \tag{181}$$

Proof: This follows easily from Proposition 25. First, it is well known[6] that

$$\text{Cen } [U(gl(m, C))] \text{ is generated by the generalized Casimir operators } \Delta_1, \cdots, \Delta_m, \tag{182}$$

where Δ_k is defined as follows: choose the matrices e_{ij} [the (i,j)th entry of the matrix e_{ij} is $+1$, whereas the other elements of e_{ij} are zero] as the base of $gl(m, C)$; then,

$$\Delta_1 \equiv \sum_{i=1}^{m} e_{ii}, \tag{183}$$

$$\Delta_2 \equiv \sum_{i,j} e_{ij}e_{ji}. \tag{184}$$

$$\vdots$$

We note that e_{ij} is a *matrix* and not just an entry of it.

Since $D(gl(m, C))$ is generated by

$$\Delta_1, \cdots, \Delta_n \quad \text{and} \quad D(gl(m, C)_0), \tag{185}$$

Proposition 25 implies

$$D(gl(m, C)) = D_{\frac{1}{2}m(m-1),m}(C). \tag{186}$$

[6] I. M. Gel'fand, Mat. Sb. **26**, 193 (1950).

Similarly, $D(sl(m, C))$ is generated by

$$\Delta_2, \cdots, \Delta_m \quad \text{and} \quad D(gl(m, C)_0) \tag{187}$$

and, by Proposition 25, we get

$$D(sl(m, C)) = D_{\frac{1}{2}m(m-1),m-1}(C). \quad \text{Q.E.D.}$$

XIV. PROOF OF THE CONJECTURE FOR NILPOTENT LIE ALGEBRAS

Let the L in this section be a nilpotent Lie algebra over a commutative field F of characteristic zero. Then the conjecture can be proved by means of Proposition 28. First, we mention the following theorem due to Dixmier.[7]

Proposition 27:

(i) $$\text{Cen } [D(L)] = D(\text{Cen } UL), \tag{188}$$

(ii) $$D(\text{Cen } UL) \leftrightarrow D(R_t[x]), \tag{189}$$

where $R_t[x]$ is the set of all rational functions of t variables. t is even (odd) if dim L is even (odd).

(iii) If L_0 is an ideal of codimension 1 in L then either

$$\text{Cen } UL_0 \subset \text{Cen } UL \tag{190}$$

or

$$\text{Cen } UL_0 \supset \text{Cen } UL. \tag{191}$$

Proposition 28:

$$\exists x_1, \cdots, x_n, y_1, \cdots, y_n, z_1, \cdots, z_t \in UL \tag{192}$$

(where the integers n and t depend on L) such that:

(i) $D(L)$ is generated by:

$$x_i, y_j, z_k, \quad \text{with} \quad i, j = 1, \cdots, n, \quad k = 1, \cdots, t; \tag{193}$$

(ii) x_i, y_j, and z_k satisfy

$$\begin{aligned} [x_i, x_j] &= 0, \\ [y_i, y_j] &= 0, \\ [z_i, z_j] &= 0, \\ [x_i, z_j] &= 0, \\ [y_i, z_j] &= 0, \end{aligned} \tag{194}$$

and

$$[x_i, y_j] = \delta_{ij}w,$$

where $w \in \text{Cen } UL$ and $w \neq 0$.

Proof: The proof is carried out by a mathematical induction on dim L. Let L_0 be an ideal in L with

$$\text{codim } L_0 = 1; \tag{195}$$

then we have, as according to Proposition 27 (iii), two possibilities and they are treated separately as follows.

[7] J. Dixmier, Bull. Soc. Math. France **85**, 325 (1957); Arch. Math. **10**, 321 (1959).

(A) Cen $UL_0 \subset$ Cen UL. In this case,

$$\exists y \in \text{Cen } UL: y \notin \text{Cen } UL_0. \tag{196}$$

Let $x \in L$ and $x \notin L_0$; then we can write

$$y \equiv \sum_{i=0}^{n} u_i x^{n-i}, \quad u_i \in UL_0, \tag{197}$$

with $x^0 \equiv 1$ understood.

Since y is in the center of UL, by definition we have

$$[y, v] = 0 \quad \text{for any} \quad v \in L. \tag{198}$$

Thus (197) gives

$$[u_0, v]x^n + (nu_0[x, v] + [u_1, v])x^{n-1} + \cdots = 0 \tag{199}$$

which implies

$$[u_0, v] = 0, \tag{200}$$

$$nu_0[x, v] + [u_1, v] = 0 \quad \text{(etc.)}. \tag{201}$$

From (200) and (201), we conclude, respectively,

$$u_0 \in \text{Cen } UL_0 \tag{202}$$

and

$$nu_0 x + u_1 \in \text{Cen } UL. \tag{203}$$

By mathematical induction we now assume that

$$\exists x_1, \cdots, x_n, y_1, \cdots, y_n, z_1, \cdots, z_t \in UL. \tag{204}$$

Then from (203) we introduce

$$z_{t+1} \equiv nu_0 x + u_1 \in \text{Cen } UL. \tag{205}$$

It can be immediately verified that

$$x_1, \cdots, x_n, y_1, \cdots, y_n, z_1, \cdots, z_{t+1}$$

satisfy the required conditions of (194). This completes our mathematical induction on t. Q.E.D.

(B) Cen $UL_0 \supset$ Cen UL. The proof for this is slightly more complicated and we refer to the original paper of Gel'fand and Kirillov for details.

Proof of the Conjecture: Now we are in a position to prove the conjecture, i.e.,

$$D(L) \leftrightarrow D_{n,t}(F),$$

for a nilpotent L over F. The proof is easily carried out by putting simply

$$p_i \equiv x_i w^{-1}$$

and

$$q_i \equiv y_i$$

where w is defined in (194). Q.E.D.

XV. FURTHER REMARKS

It is clear that the Gel'fand–Kirillov conjecture is actually much more ambitious than the solid proofs they gave for the case of

(i) $gl(m, C)$,
(ii) $sl(m, C)$,
(iii) nilpotent Lie algebras over an algebraically closed commutative field of characteristic zero.

Gel'fand and Kirillov have also indicated in their paper[8] that they have a proof for the case of a semisimple Lie algebra L of rank two over an algebraically closed commutative field by direct construction of a proper basis in $D(L)$. They indicated the choice of such a basis is facilitated by the fact that $D(L)$ is generated *by the subfield spanned by* the maximal solvable subalgebra L' in L *and by the subfield consisting of* the elements that commute with the elements of L'.

It is clear that algebraic Lie algebras cover many more cases than the cases proved by Gel'fand and Kirillov and, therefore, the conjecture leaves a large domain as an open question subject to further investigation.

For those who are interested in the Gel'fand–Kirillov conjecture as an indication of possible realizations of algebraic Lie algebras by p's and q's satisfying the so-called "canonical commutation relations" [i.e., relations (62) and (63)], the explicit constructions are hinted at in the proof of Proposition 25 for the cases of $gl(m, C)$ and $sl(m, C)$ which are particularly interesting to physicists working in particle physics or dynamic groups. The conjecture also serves a useful purpose to identify a "dynamic group" (i.e., if $A_{n,t}(F)$ is known as an algebra generated by $2n$ generators p's and q's and t indeterminates) with an algebraic Lie algebra, up to isomorphisms, under certain circumstances.

As to the general statement concerning classification of algebraic Lie algebras, the present knowledge seems to be,[9] if we restrict ourselves to an algebraically closed ground field of characteristic zero, that the following are the known cases of *algebraic* Lie algebras:

(i) nilpotent linear Lie algebras,
(ii) representations of diag (n, C) as derivations of nilpotent linear Lie algebras, where "diag" means all the diagonal matrices,
(iii) representations of semisimple Lie algebras as derivations of solvable Lie algebras.

[8] I. M. Gel'fand and A. A. Kirillov, Dokl. Akad. Nauk SSSR **167**, 503 (1966).
[9] G. Seligman (private communication).

The proofs of some of the facts, e.g., $sl(n, C)$ is algebraic, can be found in Ref. 10.

ACKNOWLEDGMENTS

I am very grateful to Professor E. C. G. Sudarshan for introducing me into this interesting subject and for his encouragement and kind hospitality during my leave of absence with his High-Energy Theory Group at Syracuse. I am also much indebted to the following mathematicians: to Professor Bruno Harris and Professor Daya-Nand Verma for discussions at Princeton, to Professor G. Seligman for a communication about classification of algebraic Lie algebras, and to Professor K. T. Chen of Illinois for a useful comment.

APPENDIX A: PROOF OF PROPOSITION 13

The proof of Proposition 13 is as follows.

Proof: Consider an increasing filtration

$$R_0 \subset R_1 \subset R_2 \subset \cdots, \tag{A1}$$

with

$$R_i \cdot R_j \subset R_{i+j} \tag{A2}$$

and

$$\bigcup_{i=0}^{\infty} R_i = R. \tag{A3}$$

We define, as usual,

$$\mathrm{gr}^{(i)} R \equiv R_i / R_{i-1} \tag{A4}$$

and

$$\mathrm{gr}\, R \equiv \sum_{i=0}^{\infty} \mathrm{gr}^{(i)} R. \tag{A5}$$

Let B be an ideal of R, then we can also introduce an increasing filtration inherited naturally from that of R:

$$B_0 \subset B_1 \subset B_2 \cdots, \tag{A6}$$

by requiring that

$$B_i \equiv B \cap R_i; \tag{A7}$$

then, we can proceed to define the graded structure associated with the increasing filtration we just introduced for B:

$$\mathrm{gr}\, B = \sum_{i=0}^{\infty} \mathrm{gr}^{(i)} B, \tag{A8}$$

where

$$\mathrm{gr}^{(i)} B \equiv B_i / B_{i-1}. \tag{A9}$$

We can see that

$$\mathrm{gr}\, B \Subset \mathrm{gr}\, R \tag{A10}$$

and also,

$$\mathrm{gr}\, B_0 \Subset \mathrm{gr}\, B_1 \Subset \cdots. \tag{A11}$$

Since gr R is Noetherian by assumption,

$$\exists n: \mathrm{gr}\, B_n = \mathrm{gr}\, B_{n+1} = \cdots. \tag{A12}$$

We want to show, then, that

$$B_n = B_{n+1} = \cdots. \tag{A13}$$

Consider $\forall x \in B_{n+1}$ (let $I_{(+)}$ be the set of all nonnegative integers) and let $k \in I_{(+)}$ be such that

$$x \in B_{n+1}^{(k)} \equiv B_{n+1} \cap R_k,$$

and

$$x \notin B_{n+1}^{(k-1)} \equiv B_{n+1} \cap R_{k-1}.$$

Since

$$\frac{\mathrm{gr}\, B_{n+1}}{\oplus_j \mathrm{gr}^{\,j} B_{n+1}} = \frac{\mathrm{gr}\, B_n}{\oplus_j \mathrm{gr}^{\,j} B_n},$$

$$\therefore \quad \mathrm{gr}^{\,j} B_{n+1} = \mathrm{gr}^{\,j} B_n, \quad \text{for} \quad \forall j \in I_{(+)}. \tag{A14}$$

In particular, for $j = k$,

$$\mathrm{gr}^{\,k} B_{n+1} = \mathrm{gr}^{\,k} B_n \tag{A15}$$

$$\therefore \quad \exists y_1 \in B_n : \underbrace{x \bmod B_{n+1}^{(k-1)}}_{\in \mathrm{gr}^{\,k} B_{n+1}} = \underbrace{y_1 \bmod B_{n+1}^{(k-1)}}_{\in \mathrm{gr}^{\,k} B_n = \mathrm{gr}^{\,k} B_{n+1}},$$

i.e.,

$$x - y_1 \in B_{n+1}^{(k-1)}. \tag{A16}$$

Similarly,

$$\mathrm{gr}^{\,k-1} B_{n+1} = \mathrm{gr}^{\,k-1} B_n, \tag{A17}$$

therefore,

$$\exists y_2 \in B_b : \underbrace{x - y_1 \bmod B_{n+1}^{(k-2)}}_{\in \mathrm{gr}^{\,k-1} B_{n+1}} = \underbrace{y_2 \bmod B_{n+1}^{(k-2)}}_{\in \mathrm{gr}^{\,k-1} B_n = \mathrm{gr}^{\,k-1} B_{n+1}},$$

i.e.,

$$\exists y_2 \in B_n : (x - y_1 - y_2) \in B_{n+1}^{(k-2)}. \tag{A18}$$

Proceeding in this manner, we get, finally,

$$\exists y_1, \cdots, y_k \in B_n : \left[x - \sum_{i=1}^{k} y_i \right] \in B_{n+1}^{(0)}. \tag{A19}$$

Using

$$\mathrm{gr}^{\,0} B_{n+1} = \mathrm{gr}^{\,0} B_n, \tag{A20}$$

i.e.,

$$B_{n+1}^{(0)} = B_n^{(0)}, \tag{A21}$$

we obtain, therefore,

$$\exists y_{k+1} \in B_n : x - \sum_{i=1}^{k} y_i = y_{k+1}, \tag{A22}$$

but

$$B_{n+1}^{(-1)} = 0 \quad (\because \quad R_{-1} \equiv 0)$$

$$\therefore \quad x = \sum_{i=1}^{k+1} y_i \in B_n. \tag{A23}$$

Remembering we started with $\forall x \in B_{n+1}$, it follows that

$$B_{n+1} = B_n. \tag{A24}$$

[10] C. Chevalley, *Théorie de groupes de Lie; Tome II: groupes algebriques* (Hermann, Paris, 1955); G. Seligman, "Algebraic Groups" (Yale University lecture notes, available from the Mathematics Dept.).

Next, it is not difficult to see that R of the theorem does not contain any zero divisors: Consider any two nonzero elements a_1 and a_2 in R, and let k_1 and k_2 be the smallest integers such that

$$a_1 \in R_{k_1} \quad (\text{i.e., } a_1 \notin R_{k_1-1}) \tag{A25}$$

and

$$a_2 \in R_{k_2} \quad (\text{i.e., } a_2 \notin R_{k_2-1}) \tag{A26}$$

Therefore, by the canonical projection mapping defined in (41), we have, obviously, that

$$\pi_{k_1}(a_2) \neq 0 \tag{A27}$$

and

$$\pi_{k_2}(a_2) \neq 0. \tag{A28}$$

Equations (A27) and (A28) lead to

$$\pi_{k_1+k_2}(a_1 \cdot a_2) = \pi_{k_1}(a_1) \cdot \pi_{k_2}(a_2) \neq 0; \tag{A29}$$

$$\therefore \quad a_1 \cdot a_2 \neq 0. \quad \text{Q.E.D.} \tag{A30}$$

APPENDIX B: PROOF OF PROPOSITION 14

The proof of Proposition 14 is as follows.

Proof: Let R be a Noether ring without zero divisors, then, by definition any chain of ideals of R,

$$R_1 \subseteq R_2 \subseteq \cdots \subseteq R_n \tag{B1}$$

terminates at some integer n. That is,

$$\exists n\colon R_{n+1} = R_n. \tag{B2}$$

Now consider the particular case of the (left) ideal R_n generated by the set of elements

$$\{a, a \cdot b, a \cdot b^2, \cdots, a \cdot b^n\}, \tag{B3}$$

where a and b are any two nonzero elements of R. Therefore, (B2) allows us to write

$$a \cdot b^{n+1} = \sum_{i=0}^{n} c_i \cdot a \cdot b^i, \tag{B4}$$

where c_i are elements in R and $b^0 = 1$ (unit element of the ring). Let k be the smallest integer such that, in (B4),

$$c_k \neq 0 \quad (\text{i.e., } c_{k-1} = 0). \tag{B5}$$

Equation (B4) can be written as

$$a \cdot b^{n+1} = \sum_{i=k}^{n} c_i \cdot a \cdot b^i,$$

i.e.,

$$a \cdot b^{n+1-k} = \sum_{i=k}^{n} c_i \cdot a \cdot b^{i-k},$$

i.e.,

$$a \cdot b^{n+1-k} - \sum_{i=k+1}^{n} c_i \cdot a \cdot b^{i-k} = c_k \cdot a,$$

i.e.,

$$\left[a \cdot b^{n-k} - \sum_{i=k+1}^{n} c_i \cdot a \cdot b^{i-k-1}\right] \cdot b = c_k \cdot a,$$

i.e.,

$$h \cdot b = c_k \cdot a, \tag{B6}$$

where

$$h \equiv a \cdot b^{n-k} - \sum_{i=k+1}^{n} c_i \cdot a \cdot b^{i-k-1}. \tag{B7}$$

Therefore we have, from (B6),

$$\exists c_k \neq 0, \quad c_k, h \in R\colon c_k \cdot a = h \cdot b,$$

which is the (left) Ore condition. Q.E.D.

Broken Symmetry and Generalized Bose Condensation in Restricted Geometries

M. D. GIRARDEAU
Institute of Theoretical Science and Department of Physics, University of Oregon, Eugene, Oregon

(Received 21 June 1968)

It is suggested, on the basis of soluble models, that generalized Bose condensation results from a broken symmetry associated with a nonvanishing pairing amplitude with very large low-momentum components, leading to a nonzero "pairing density" p_c as well as a nonzero generalized condensate density ρ_c. For systems of interacting bosons it is proved that (1) nonzero p_c implies nonzero ρ_c, and (2) $p_c = 0$ in one or two dimensions and, more generally, for geometries finite in one dimension and infinite in the other two (films) or finite in two dimensions and infinite in one (pores). It is pointed out that this does not exclude superfluidity in such geometries, but does show the need of a new mechanism to explain it.

There exist proofs[1,2] that the usual type of Bose condensation cannot occur at any nonzero temperature unless the system is infinite in at least three dimensions. This presents a problem for the theoretical interpretation of experimental results, which show, e.g., that superfluidity occurs at sufficiently low temperatures in films of ^{4}He as thin as a few atomic diameters.[3,4]

However, the type of Bose condensation usually assumed, namely "simple Bose condensation" (SBC) characterized by the presence of a nonzero fraction of the particles with momentum exactly zero, is only a very special case of a generalized type of Bose condensation (GBC) characterized by the presence of a nonzero fraction of the particles in an infinitesimal neighborhood of the origin of momentum space.[5,6] There exist simplified models which exhibit a thermodynamic phase transition at a temperature $T_c > 0$ below which GBC is present, but nevertheless SBC is absent.[5,7] It has been pointed out by Krueger[2] that GBC might be present in restricted geometries, since the existing proofs[1,2] do not exclude this possibility. The main purpose of the present paper is to show that the hope of explaining the experimental results on films and pores by such a generalization from SBC to GBC is illusory; we extend Hohenberg's proof[1] to show that the type of broken symmetry expected in the case of GBC also cannot occur in restricted geometries. This does not absolutely rule out GBC or, more generally, superfluidity, just as the existing proofs[1,2] do not absolutely rule out SBC. However, it does show that the superfluidity observed in restricted geometries cannot be associated with a broken symmetry of the type expected in a system with GBC, just as it cannot be associated with SBC.

A system of interacting bosons has a Hamiltonian of the form

$$H = \sum_{\mathbf{k}} \tfrac{1}{2}k^2 N_{\mathbf{k}} + V, \tag{1}$$

where $N_{\mathbf{k}} = a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$ and V is some functional of the density operator[8]

$$\rho(\mathbf{r}) = \Omega^{-1} \sum_{\mathbf{q}} \rho_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}, \quad \rho_{\mathbf{q}} = \sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}+\mathbf{q}}. \tag{2}$$

We assume periodic boundary conditions with periodicity volume (or length or area) Ω. In case a degenerate phase of the type associated, e.g., with superfluidity is present, the usual grand canonical ensemble does not lead to correct results for thermodynamic averages of observables which have an infinite susceptibility to a degeneracy-breaking perturbation; in this case the grand-canonical average should be replaced by the Bogoliubov quasiaverage[9]

$$\langle \cdots \rangle = \frac{\mathrm{Tr}\,(\cdots e^{-\beta H_\epsilon})}{\mathrm{Tr}\, e^{-\beta H_\epsilon}}, \tag{3}$$

where

$$H_\epsilon = H - \mu \sum_{\mathbf{k}} N_{\mathbf{k}} + V_\epsilon \tag{4}$$

and V_ϵ is an appropriate symmetry-breaking term which vanishes as $\epsilon \to 0$. This limit is to be taken *after* the thermodynamic limit ($\Omega \to \infty$ for fixed μ).

GBC is characterized by a nonvanishing condensate density ρ_c, where ρ_c is defined by[5,6]

$$\rho_c = \lim_{k_0 \to 0} \lim_{\epsilon \to 0} \lim \mathrm{therm}\, \Omega^{-1} \sum_{k < k_0} \langle N_{\mathbf{k}} \rangle \tag{5}$$

[1] P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).
[2] D. A. Krueger, Phys. Rev. Letters **19**, 563 (1967).
[3] E. Long and L. Meyer, Phil. Mag. Suppl. **2**, 1 (1953).
[4] D. F. Brewer and K. Mendelssohn, Proc. Roy. Soc. (London) **A260**, 1 (1961).
[5] M. Girardeau, Phys. Fluids **5**, 1468 (1962), Eq. (48) ff.
[6] M. D. Girardeau, J. Math. Phys. **6**, 1083 (1965).
[7] Reference 6, Sec. 4.

[8] In the n-particle Schrödinger representation, this means that $V = V(\mathbf{r}_1 \cdots \mathbf{r}_n)$. This includes two-particle interactions as a special case, as well as any interactions with the walls.
[9] N. N. Bogoliubov, Physica **26**, 1 (1960).

and "lim therm" denotes the thermodynamic limit. One can also define a condensate pairing order parameter or "pairing density" p_c by[10]

$$p_c = \lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}|\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle|. \tag{6}$$

The symmetry-breaking term is taken to be

$$V_\epsilon = \sum_{\mathbf{k}}\sigma(\mathbf{k})(\epsilon a_{\mathbf{k}}a_{-\mathbf{k}} + \epsilon^* a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger), \tag{7}$$

where $\sigma(\mathbf{k})$ is a real, even function of $\mathbf{k}$. The motivation for the definitions (6) and (7) is that the known models which exhibit GBC ($\rho_c > 0$ for $T < T_c$) also have $p_c > 0$ for $T < T_c$, provided that the symmetry-breaking term is of the form (7). This is rather obvious for the model of reference[5] which contains pairing interactions even before inclusion of the symmetry-breaking term. However, it is also true for the plane-wave Hartree–Fock model.[7] This model remains exactly soluble in the thermodynamic limit after addition of V_ϵ, and the solution has the property that $|\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle| \to \langle N_{\mathbf{k}}\rangle$ as $\mathbf{k}\to 0$; this is shown in the Appendix. On the other hand, $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ would vanish identically if averages were taken with an ordinary grand canonical ensemble.

In fact, any Bose system with $p_c > 0$ exhibits GBC. To prove this we note that the quasiaverage $\langle A^\dagger B\rangle$, defined according to (3), satisfies the necessary properties of an inner product, the operators $A, B, \cdots$ being interpreted here as elements of a vector space. Thus, the Schwartz inequality

$$|\langle A^\dagger B\rangle|^2 \le \langle A^\dagger A\rangle\langle B^\dagger B\rangle \tag{8}$$

is satisfied. Taking $A = a_{\mathbf{k}}^\dagger$, $B = a_{-\mathbf{k}}$ gives

$$\begin{aligned}|\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle|^2 &\le \langle a_{\mathbf{k}}a_{\mathbf{k}}^\dagger\rangle\langle a_{-\mathbf{k}}^\dagger a_{-\mathbf{k}}\rangle \\ &= (\langle N_{\mathbf{k}}\rangle + 1)\langle N_{-\mathbf{k}}\rangle = \langle N_{\mathbf{k}}\rangle^2 + \langle N_{\mathbf{k}}\rangle,\end{aligned} \tag{9}$$

the last equality following from the assumption that the geometry of the system is symmetrical enough that $\langle N_{\mathbf{k}}\rangle = \langle N_{-\mathbf{k}}\rangle$. Now

$$(\langle N_{\mathbf{k}}\rangle + \langle N_{\mathbf{k}}\rangle^{\frac12})^2 \ge \langle N_{\mathbf{k}}\rangle^2 + \langle N_{\mathbf{k}}\rangle, \tag{10}$$

so that

$$|\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle| \le \langle N_{\mathbf{k}}\rangle + \langle N_{\mathbf{k}}\rangle^{\frac12}. \tag{11}$$

Then (5) and (6) imply

$$0 \le p_c \le \rho_c + \lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}\langle N_{\mathbf{k}}\rangle^{\frac12}. \tag{12}$$

Define an inner product

$$(\alpha, \beta) \equiv \sum_{k<k_0}\alpha_{\mathbf{k}}\beta_{\mathbf{k}}. \tag{13}$$

[10] One might be tempted, instead, to define a generalized order parameter in terms of the values of $\langle a_{\mathbf{k}}\rangle$. However, the spatial homogeneity of a liquid implies that $\langle a_{\mathbf{k}}\rangle = 0$ for $\mathbf{k} \neq 0$.

Then the Schwartz inequality, with $\alpha_{\mathbf{k}} = \langle N_{\mathbf{k}}\rangle^{\frac12}$, $\beta_{\mathbf{k}} = 1$, gives

$$\begin{aligned}0 &\le \lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}\langle N_{\mathbf{k}}\rangle^{\frac12} \\ &\le \rho_c^{\frac12}\lim_{k_0\to 0} f(k_0) = 0,\end{aligned} \tag{14}$$

where[11]

$$f(k_0) = \left[\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}1\right]^{\frac12}. \tag{15}$$

Hence (12) reduces to

$$\rho_c \ge p_c \ge 0. \tag{16}$$

In particular, $p_c > 0$ implies $\rho_c > 0$.

We now prove that $p_c > 0$ is impossible in restricted geometries. The proof is a modification of Hohenberg's proof[1] of the impossibility of superconductivity in one or two dimensions. We start with the Bogoliubov inequality[12–14]

$$\tfrac12\langle\{A, A^\dagger\}\rangle \ge \frac{\kappa T\,|\langle[C, A]\rangle|^2}{\langle[[C, H_\epsilon], C^\dagger]\rangle}, \tag{17}$$

where κ is Boltzmann's constant. Defining

$$C = \rho_{-\mathbf{q}}, \quad A = \sum_{\mathbf{k}}S(\mathbf{k})a_{\mathbf{q}-\mathbf{k}}a_{\mathbf{k}}, \tag{18}$$

where $S(\mathbf{k})$ is a function which will be specified later, one has by (1), (2), (4), (7), and (17) that

$$\begin{aligned}&\tfrac12\Omega^{-1}\sum_{\mathbf{k}}S(\mathbf{k})[S^*(\mathbf{k}) + S^*(\mathbf{q}-\mathbf{k})] \\ &\quad + \tfrac12\Omega^{-1}\sum_{\mathbf{k}}|S(\mathbf{k}) + S(\mathbf{q}-\mathbf{k})|^2\langle N_{\mathbf{k}}\rangle \\ &\quad + \Omega^{-1}\sum_{\mathbf{k}\mathbf{k}'}S(\mathbf{k})S^*(\mathbf{k}')\langle a_{\mathbf{k}'}^\dagger a_{\mathbf{q}-\mathbf{k}'}^\dagger a_{\mathbf{q}-\mathbf{k}}a_{\mathbf{k}}\rangle \\ &\ge \frac{\kappa T\left|\Lambda^{-1}\sum_{\mathbf{k}}[S(\mathbf{k})+S(\mathbf{q}-\mathbf{k})]\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle\right|^2}{\rho q^2 - 2\Omega^{-1}\sum_{\mathbf{k}}[\sigma(\mathbf{k})+\sigma(\mathbf{q}-\mathbf{k})](\epsilon\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle+\epsilon^*\langle a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger\rangle)},\end{aligned} \tag{19}$$

where $\rho = \Omega^{-1}\langle N\rangle = \Omega^{-1}\sum_{\mathbf{k}}\langle N_{\mathbf{k}}\rangle$. Choose $S(\mathbf{k})$ to vanish for $k > k_0$ and to have unit modulus and phase opposite to that of $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ for $k < k_0$:

$$\begin{aligned}|S(\mathbf{k})| = 1, \quad S(\mathbf{k})\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle &= |\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle|, \quad k < k_0, \\ S(\mathbf{k}) &= 0, \quad k > k_0.\end{aligned} \tag{20}$$

If one applies the three limits involved in the definitions (5) and (6), then most of the summations in (19)

[11] For systems of one, two, or three dimensions, $f(k_0)$ is equal to $(k_0/\pi)^{\frac12}$, $(k_0^2/\pi)^{\frac12}$, or $(k_0^3/6\pi^2)^{\frac12}$, respectively.

[12] N. N. Bogoliubov, Physik. Abhandl. Sowjetunion **6**, 1, 113, 229 (1962).

[13] H. Wagner, Z. Physik **195**, 273 (1966).

[14] N. D. Mermin and H. Wagner, Phys. Rev. Letters **17**, 1133 (1966).

will vanish for fixed $q \neq 0$ provided that

$$\lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}\langle N_{\mathbf{q}-\mathbf{k}}\rangle = 0, \quad \mathbf{q}\neq 0,$$
$$\lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ \Omega^{-1}\sum_{k<k_0}|\langle a_{\mathbf{q}-\mathbf{k}}a_{\mathbf{k}-\mathbf{q}}\rangle| = 0, \quad \mathbf{q}\neq 0. \tag{21}$$

These conditions could be violated only if there were GBC into the neighborhood of $\mathbf{q} \neq 0$; however, we are assuming that GBC, if present at all, occurs in the neighborhood of the origin of $\mathbf{k}$ space. Thus (19) reduces, with (5) and (6), to

$$\lim_{k_0\to 0}\lim_{\epsilon\to 0}\lim \text{therm}\ F(\mathbf{q}) \geq \frac{\kappa T p_c^2}{\rho q^2} - \tfrac{1}{2}\rho_c, \tag{22}$$

where

$$F(\mathbf{q}) \equiv \Omega^{-1}\sum_{k<k_0}\sum_{k'<k_0} S(\mathbf{k})S^*(\mathbf{k}')\langle a_{\mathbf{k}'}^\dagger a_{\mathbf{q}-\mathbf{k}'}^\dagger a_{\mathbf{q}-\mathbf{k}}a_{\mathbf{k}}\rangle. \tag{23}$$

Multiplying (22) by $e^{-i\mathbf{q}\cdot\mathbf{r}}$ and integrating over the region of momentum space satisfying $q < q_0$, where q_0 is independent of Ω, ϵ, and k_0, one finds in the thermodynamic limit

$$\lim_{k_0\to 0}\lim_{\epsilon\to 0}\int f(\mathbf{r}-\mathbf{r}')\sigma(r', q_0)\,d^\nu r' \geq \frac{\kappa T p_c^2}{\rho}\int_{q<q_0}\frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{q^2}\,d^\nu q - \tfrac{1}{2}\rho_c\sigma(r, q_0), \tag{24}$$

where $f(\mathbf{r})$ is the Fourier transform of $F(\mathbf{q})$:

$$F(\mathbf{q}) = \int f(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}\,d^\nu r \tag{25}$$

and

$$\sigma(r, q_0) \equiv \int_{q<q_0} e^{i\mathbf{q}\cdot\mathbf{r}}\,d^\nu q; \tag{26}$$

here ν is the number of dimensions of the system (1, 2, or 3). The explicit forms of (26) are

$$\begin{aligned}\sigma(r, q_0) &= 2r^{-1}\sin(q_0 r), && \nu = 1,\\ &= 2\pi q_0 r^{-1}J_1(q_0 r), && \nu = 2,\\ &= 4\pi r^{-3}[\sin(q_0 r) - q_0 r\cos(q_0 r)], && \nu = 3,\end{aligned} \tag{27}$$

where J_1 is the Bessel function of order one.

To obtain a bound on the left side of (24), we substitute for the $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ in terms of the Bose field operators $\psi(\mathbf{r})$, $\psi^\dagger(\mathbf{r})$ of which they are the Fourier transforms, obtaining, with (23), (25), and (20),

$$\begin{aligned}F(\mathbf{q}) &= \Omega^{-1}\int f(\mathbf{r}-\mathbf{r}')e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\,d^\nu r\,d^\nu r',\\ f(\mathbf{r}_1-\mathbf{r}_2) &= \int s(\mathbf{r}'-\mathbf{r}_2)s^*(\mathbf{r}-\mathbf{r}_1)\\ &\quad\times\langle\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r})\psi(\mathbf{r}')\psi(\mathbf{r}_2)\rangle\,d^\nu r_1\,d^\nu r_2,\\ s(\mathbf{r}) &= \Omega^{-1}\sum_{k<k_0}e^{-i\vartheta_{\mathbf{k}}}e^{i\mathbf{k}\cdot\mathbf{r}},\end{aligned} \tag{28}$$

where $\vartheta_{\mathbf{k}}$ is the phase of $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$. The dependence of $\vartheta_{\mathbf{k}}$ on $\mathbf{k}$ is determined by the geometry of the system and the interparticle interactions. The simplest case (analogous to S-wave pairing in superconductors) is that in which $\vartheta_{\mathbf{k}} = \vartheta$, independent of $\mathbf{k}$. Then by (26)

$$s(\mathbf{r}) = (2\pi)^{-\nu}e^{-i\vartheta}\sigma(r, k_0) \tag{29}$$

in the thermodynamic limit, so that

$$f(\mathbf{r}_1-\mathbf{r}_2) = (2\pi)^{-2\nu}\int\sigma(|\mathbf{r}'-\mathbf{r}_2|, k_0)\sigma(|\mathbf{r}-\mathbf{r}_1|, k_0) \times\langle\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r})\psi(\mathbf{r}')\psi(\mathbf{r}_2)\rangle\,d^\nu r_1\,d^\nu r_2. \tag{30}$$

The Schwartz inequality (8) implies

$$\begin{aligned}&|\langle\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r})\psi(\mathbf{r}')\psi(\mathbf{r}_2)\rangle|\\ &\quad\leq|\langle\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r})\psi(\mathbf{r})\psi(\mathbf{r}_1)\rangle|^{\frac{1}{2}}\,|\langle\psi^\dagger(\mathbf{r}_2)\psi^\dagger(\mathbf{r}')\psi(\mathbf{r}')\psi(\mathbf{r}_2)\rangle|^{\frac{1}{2}}\\ &\quad=|\langle\rho(\mathbf{r}_1)\rho(\mathbf{r})\rangle - \langle\rho(\mathbf{r})\rangle\delta(\mathbf{r}-\mathbf{r}_1)|^{\frac{1}{2}}\\ &\qquad\times|\langle\rho(\mathbf{r}_2)\rho(\mathbf{r}')\rangle - \langle\rho(\mathbf{r}')\rangle\delta(\mathbf{r}'-\mathbf{r}_2)|^{\frac{1}{2}}\\ &\quad=|\langle\rho(\mathbf{r}_1)\rho(\mathbf{r})\rangle|^{\frac{1}{2}}\,|\langle\rho(\mathbf{r}_2)\rho(\mathbf{r}')\rangle|^{\frac{1}{2}},\end{aligned} \tag{31}$$

the last equality (when inserted into an integral) following from the facts that $\langle\rho(\mathbf{r})\rangle$ is finite and that the square root of the Dirac delta function integrates to zero. Since the density–density correlation function $\langle\rho(\mathbf{r})\rho(\mathbf{r}')\rangle$ is integrable (in fact, everywhere finite), it follows that the same is true of $\langle\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r})\psi(\mathbf{r}')\psi(\mathbf{r}_2)\rangle$.

To proceed we separately consider the small- and large-distance contributions to (30). From (27) one sees that as $k_0 \to 0$, $\sigma(r, k_0)$ vanishes like k_0^ν for all finite r. Hence the contributions to (30) from finite values of $\mathbf{r}_1$ and $\mathbf{r}_2$ vanish in the same limit. On the other hand, $\sigma(|\mathbf{r}-\mathbf{r}_1|, k_0)$ and $\sigma(|\mathbf{r}'-\mathbf{r}_2|, k_0)$ exhibit a damped oscillatory behavior as $|\mathbf{r}_1-\mathbf{r}| \to \infty$ and $|\mathbf{r}_2-\mathbf{r}'| \to \infty$ for fixed k_0. Noting that the density–density correlation function must approach the uncorrelated form at large separations,

$$\langle\rho(\mathbf{r}_1)\rho(\mathbf{r})\rangle \xrightarrow[|\mathbf{r}_1-\mathbf{r}|\to\infty]{} \rho^2, \tag{32}$$

and making use of (27), one can estimate the contribution to (30) from the region $r_1 > r_0$, $r_2 > r_0$, where r_0 is large but finite, as

$$\begin{aligned}\nu = 1:&\quad\left[\rho\int_{k_0 r_0}^\infty x^{-1}\sin x\,dx\right]^2\\ &\qquad=[\rho\ \mathrm{si}(k_0 r_0)]^2 \xrightarrow[k_0\to 0]{} (\tfrac{1}{2}\pi\rho)^2,\\ \nu = 2:&\quad\left[\rho\int_{k_0 r_0}^\infty J_1(x)\,dx\right]^2\\ &\qquad\sim\left[\rho\int_{k_0 r_0}^\infty x^{-\frac{1}{2}}\cos x\,dx\right]^2 \xrightarrow[k_0\to 0]{} \tfrac{1}{2}\pi\rho^2,\\ \nu = 3:&\quad\left[\rho\int_{k_0 r_0}^\infty x^{-1}(\sin x - x\cos x)\,dx\right]^2\\ &\qquad=\text{indeterminate}.\end{aligned} \tag{33}$$

We thus conclude that $f(\mathbf{r}_1 - \mathbf{r}_2)$ is finite and of order ρ^2 for $\nu = 1$ or 2. It is unlikely that this conclusion would be changed in cases where the pairing amplitude does not have S-wave symmetry, giving an angular dependence of $\vartheta_{\mathbf{k}}$; this would be expected to increase the cancellations in the integral (28), making $f(\mathbf{r}_1 - \mathbf{r}_2)$ still smaller.

Substitution of these estimates for $f(\mathbf{r}_1 - \mathbf{r}_2)$ into (24) and repetition of the same argument leads to the conclusion that the integral on the left side of (24) is itself finite and of order ρ^2 for the cases of one and two dimensions; thus,

$$O(\rho^2) \geq \frac{\kappa T p_c^2}{\rho} \int_{q<q_0} \frac{e^{-i\mathbf{q}\cdot\mathbf{r}}}{q^2} d^\nu q - \tfrac{1}{2}\rho_c \sigma(r, q_0), \qquad \nu = 1 \text{ or } 2. \quad (34)$$

The integral on the right side diverges at low momenta in both one and two dimensions. This is compatible with (34) only if

$$p_c = 0, \quad \nu = 1 \text{ or } 2. \quad (35)$$

This completes the proof of the impossibility of a nonzero pairing density in one or two dimensions. For $n = 3$, no such conclusion follows, and in fact counterexamples are known.[5,7] Since SBC is a special case of GBC, our proof also excludes SBC for $\nu = 1$ or 2, although the simpler proofs[1,2] are certainly adequate for that purpose. The connection between our proof and a proof of the impossibility of SBC becomes clearer if one regards the order parameter in the case of SBC as $|\langle a_0^2 \rangle|/\Omega \equiv p_0$, rather than the usual choice of $\langle a_0 \rangle$ as order parameter.

The generalization to the case of pore geometries ($d \times d \times L$, $L \to \infty$) and film geometries ($d \times L \times L$, $L \to \infty$) is straightforward. It is clear from Figs. 1 and 2 and the definition (20) of $S(\mathbf{k})$ that, as soon as $k_0 < \pi/d$ and $q < \pi/d$, the sums on the left side of (19) and in the numerator of the right side will have contributions only from the plane through the origin of $\mathbf{k}$ space ($\nu = 2$) or from the line through the origin ($\nu = 1$). Thus the proof reduces in all essentials to that

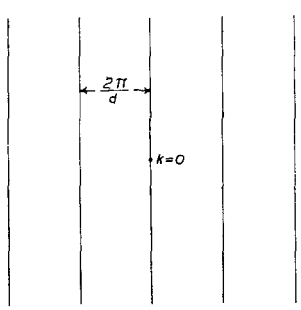


FIG. 1. Geometry of k space, seen edge on, for a film of thickness d.

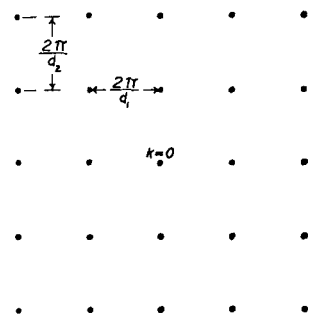


FIG. 2. Geometry of k space, seen end on, for a pore with dimensions $d_1 \times d_2$.

already given. We conclude that films and pores also cannot have $p_c > 0$. If d is large enough that

$$\kappa T \gg 2\pi^2\hbar^2/md^2$$

at the experimentally relevant temperatures, then the combined process of integration over individual planes or lines and summation over different planes or lines is well approximated by integration over a three-dimensional $\mathbf{k}$ space, so that the thermodynamic properties become experimentally indistinguishable from those of a system macroscopic in all three dimensions, except for a rounding of the thermodynamic singularities in a temperature interval of order $2\pi^2\hbar^2/\kappa md^2$ about the λ point. In such a case, one expects that true symmetry breaking, i.e., $p_c > 0$, would be replaced by a "nearly-broken symmetry" in that the pairing amplitude $\langle a_{\mathbf{k}} a_{-\mathbf{k}} \rangle$ (or, in the case of SBC, $\langle a_0 \rangle$) will retain its bulk value as ϵ is decreased, until ϵ becomes so small that the symmetry-breaking term makes an energy contribution (per particle) small compared to $2\pi^2\hbar^2/md^2$.

The experimental observations[3] do in fact show that the logarithmic specific-heat singularity disappears as the temperature falls below $\sim 2\pi^2\hbar^2/\kappa md^2$, and the specific-heat anomaly becomes more diffuse as the temperature falls. The results on the flow properties of thin films are, however, quite different. It is observed[3,4] that even for films a few monolayers thick, there is a well-defined temperature T_s below which the film flows as a superfluid at sufficiently low velocities. For such thin films T_s is considerably lower than the temperature of any observed specific heat anomaly[15] and the specific heat does not exhibit any observable anomaly at $T = T_s$.

How is this observed superfluidity to be reconciled with the proofs of the impossibility of a broken symmetry in such geometries? It would seem that the crucial point here is that the existence of Bose–Einstein condensation, either simple or generalized,

[15] D. L. Goodstein and W. D. McCormick, Phys. Rev. Letters 16, 8 (1966), plus other references cited there.

has not been shown to be necessary[16] for superfluidity. The Landau criterion[17] shows that a necessary condition for superfluidity is that the quasiparticle excitation spectrum $E(k)$ must satisfy

$$\min_k k^{-1}E(k) > 0. \tag{36}$$

In all presently known many-boson models which satisfy this criterion, Bose–Einstein condensation plays a crucial role. However, (36) could arise as a result of some mechanism other than Bose condensation, or it could arise from a relaxed definition of condensation in which not all the limiting processes involved in (5) and (6) enter.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under grant GP-7689.

APPENDIX: BROKEN SYMMETRY IN THE HARTREE–FOCK MODEL

Taking H in (4) to be the plane-wave Hartree–Fock model Hamiltonian,[7] one finds

$$\begin{aligned} H_\epsilon &= 2\pi(N-1)\rho\alpha + \sum_{\mathbf{k}}(\tfrac{1}{2}k^2 - \mu)N_{\mathbf{k}} \\ &\quad + 2\pi\Omega^{-1}\alpha\sum_{\mathbf{k}\mathbf{k}'}{}' N_{\mathbf{k}}N_{\mathbf{k}'} \\ &\quad + \sum_{\mathbf{k}}\sigma(\mathbf{k})(\epsilon a_{\mathbf{k}}a_{-\mathbf{k}} + \epsilon^* a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger). \end{aligned} \tag{A1}$$

Write

$$\begin{aligned} N_{\mathbf{k}}N_{\mathbf{k}'} &= (N_{\mathbf{k}} - n_{\mathbf{k}})(N_{\mathbf{k}'} - n_{\mathbf{k}'}) \\ &\quad + (n_{\mathbf{k}'}N_{\mathbf{k}} + n_{\mathbf{k}}N_{\mathbf{k}'}) - n_{\mathbf{k}}n_{\mathbf{k}'}, \end{aligned} \tag{A2}$$

where the $n_{\mathbf{k}}$ are c-number parameters to be determined. Then

$$\begin{aligned} H_\epsilon &= H_0 + H_1, \\ H_0 &= 2\pi(N-1)\rho\alpha - 2\pi\Omega^{-1}\alpha\sum_{\mathbf{k}\mathbf{k}'}{}' n_{\mathbf{k}}n_{\mathbf{k}'} \\ &\quad + \sum_{\mathbf{k}} w(\mathbf{k})N_{\mathbf{k}} + \sum_{\mathbf{k}}\sigma(\mathbf{k})(\epsilon a_{\mathbf{k}}a_{-\mathbf{k}} + \epsilon^* a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger), \\ H_1 &= 2\pi\Omega^{-1}\alpha\sum_{\mathbf{k}\mathbf{k}'}{}' (N_{\mathbf{k}} - n_{\mathbf{k}})(N_{\mathbf{k}'} - n_{\mathbf{k}'}), \end{aligned} \tag{A3}$$

where

$$\begin{aligned} w(\mathbf{k}) &= \tfrac{1}{2}k^2 - \mu + 4\pi\Omega^{-1}\alpha\sum_{\mathbf{k}'}{}' n_{\mathbf{k}'} \\ &= \tfrac{1}{2}k^2 - \mu + 4\pi\rho\alpha - 4\pi\Omega^{-1}\alpha n_{\mathbf{k}}. \end{aligned} \tag{A4}$$

Then H_ϵ may be replaced by H_0 in computing thermodynamic quasiaverages (3) with a fractional error which vanishes in the thermodynamic limit, provided that the $n_{\mathbf{k}}$ are taken to be[18]

$$n_{\mathbf{k}} = \frac{\mathrm{Tr}\,(N_{\mathbf{k}}e^{-\beta H_0})}{\mathrm{Tr}\,e^{-\beta H_0}}. \tag{A5}$$

H_0 can be diagonalized by a Bogoliubov transformation:

$$\begin{aligned} U^{-1}H_0U &= E_0 + \sum_{\mathbf{k}}\omega(\mathbf{k})N_{\mathbf{k}}, \\ E_0 &= 2\pi(N-1)\rho\alpha - 2\pi\Omega^{-1}\alpha\sum_{\mathbf{k}\mathbf{k}'}{}' n_{\mathbf{k}}n_{\mathbf{k}'} \\ &\quad - \tfrac{1}{2}\sum_{\mathbf{k}}[w(\mathbf{k}) - \omega(\mathbf{k})], \\ \omega(\mathbf{x}) &= [w^2(\mathbf{k}) - 4|\epsilon|^2\sigma^2(\mathbf{k})]^{\frac{1}{2}}, \end{aligned} \tag{A6}$$

where

$$\begin{aligned} U^{-1}a_{\mathbf{k}}U &= (1 - |\varphi_{\mathbf{k}}|^2)^{-\frac{1}{2}}(a_{\mathbf{k}} - \varphi_{\mathbf{k}}a_{-\mathbf{k}}^\dagger), \\ U^{-1}a_{\mathbf{k}}^\dagger U &= (U^{-1}a_{\mathbf{k}}U)^\dagger, \\ \varphi_{\mathbf{k}} &= \frac{w(\mathbf{k}) - \omega(\mathbf{k})}{2\epsilon\sigma(\mathbf{k})}. \end{aligned} \tag{A7}$$

The thermal averages of $N_{\mathbf{k}}$ and $a_{\mathbf{k}}a_{-\mathbf{k}}$ are easily computed[19] and found to be

$$\begin{aligned} \langle N_{\mathbf{k}}\rangle = n_{\mathbf{k}} &= \frac{\mathrm{Tr}\,(N_{\mathbf{k}}e^{-\beta H_0})}{\mathrm{Tr}\,e^{-\beta H_0}} \\ &= \frac{w(\mathbf{k}) - \omega(\mathbf{k})}{2\omega(\mathbf{k})} + \frac{w(\mathbf{k})}{\omega(\mathbf{k})}\frac{1}{e^{\beta\omega(\mathbf{k})} - 1} \end{aligned} \tag{A8}$$

and

$$\begin{aligned} \langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle &= \frac{\mathrm{Tr}\,(a_{\mathbf{k}}a_{-\mathbf{k}}e^{-\beta H_0})}{\mathrm{Tr}\,e^{-\beta H_0}} \\ &= -\frac{\epsilon^*\sigma(\mathbf{k})}{\omega(\mathbf{k})}\left[1 + \frac{2}{e^{\beta\omega(\mathbf{k})} - 1}\right]. \end{aligned} \tag{A9}$$

In view of (A4), Eq. (A8) is really a transcendental equation for $n_{\mathbf{k}}$, which is to be solved simultaneously with the equation

$$\Omega^{-1}\sum_{\mathbf{k}} n_{\mathbf{k}} = \rho, \tag{A10}$$

which determines μ. It can be shown that in one and two dimensions, the solution is such that no Bose condensation, either SBC or GBC, occurs at any nonzero temperature.[20] On the other hand, a straightforward extension of the previous analysis[7] shows that in three dimensions, if ϵ is small enough and $\alpha < 0$, the system will exhibit GBC at all temperatures $T < T_c$, where T_c reduces in the limit $\epsilon \to 0$ to the condensation temperature of the ideal Bose gas. Thus for $T < T_c$ in three dimensions, $n_{\mathbf{k}}$ becomes very large

[16] It is perhaps not superfluous to mention that it is not sufficient either, since an ideal Bose gas in three dimensions has $\rho_0 > 0$ (hence $\rho_c > 0$) for $T < T_c$, yet it is not superfluid at any temperature.

[17] L. Landau, J. Phys. (USSR) **5**, 71 (1941), Sec. 4.

[18] G. Wentzel, Phys. Rev. **120**, 1572 (1960).

[19] In view of (A6), these are most easily computed with the aid of the identity $\mathrm{Tr}\,(Oe^{-\beta H_0}) = \mathrm{Tr}\,(U^{-1}OUe^{-\beta U^{-1}H_0U})$.

[20] This is true in spite of the fact that the plane-wave Hartree–Fock Hamiltonian cannot be expressed in the form (1).

in a very small neighborhood of $\mathbf{k} = 0$, corresponding to the fact that $\beta\omega_{\mathbf{k}}$ becomes very small in that neighborhood.[7] Then it follows from (A6) that $w(\mathbf{k})$ differs infinitesimally from $2|\epsilon|\,\sigma(\mathbf{k})$ in this neighborhood.[21] Thus by (A8) and (A9) one finds, upon replacing $e^{\beta\omega(\mathbf{k})}$ by the first two terms of its power-series expansion,

$$n_{\mathbf{k}} \approx 2|\epsilon|\,\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}),$$
$$\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle \approx -2\epsilon^*\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}), \qquad \text{(A11)}$$

for $T < T_c$ and $\mathbf{k} \to 0$. More explicit results can be obtained by solving (A4) for $n_{\mathbf{k}}$ after the substitutions $\alpha = -|\alpha|$, $w(\mathbf{k}) \approx 2|\epsilon|\,\sigma(\mathbf{k})$ [the latter following from $\omega(\mathbf{k}) \approx 0$]. This gives

$$n_{\mathbf{k}} = \tfrac{1}{2}(k_s^2 - k^2)\Omega/4\pi|\alpha|, \quad T < T_c, \quad k < k_s, \qquad \text{(A12)}$$

where

$$k_s = \{2[\mu + 4\pi\rho|\alpha| + 2|\epsilon|\,\sigma(\mathbf{k})]\}^{\frac{1}{2}}. \qquad \text{(A13)}$$

[21] We assume that $\sigma(\mathbf{k}) \geq 0$.

Then, by (A11),

$$\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle = -\tfrac{1}{2}e^{-i\vartheta}(k_s^2 - k^2)\Omega/4\pi|\alpha|, \quad T < T_c, \quad k < k_s, \qquad \text{(A14)}$$

where $\epsilon = |\epsilon|\,e^{i\vartheta}$. Equations (5) and (6) then give

$$p_c = \rho_c = f\rho, \quad T < T_c,$$
$$p_c = \rho_c = 0, \quad T > T_c, \qquad \text{(A15)}$$

where[11]

$$(1 - f)\rho = 2.612(\kappa T/2\pi)^{\frac{3}{2}}, \quad T < T_c, \qquad \text{(A16)}$$
$$\kappa T_c = 2\pi(\rho/2.612)^{\frac{2}{3}},$$

and

$$\lim_{\epsilon\to 0} k_s = [15(2\pi)^3 f\rho|\alpha|/\Omega]^{\frac{1}{5}}, \quad T < T_c. \qquad \text{(A17)}$$

The infinite susceptibility of $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ to the symmetry-breaking perturbation V_ϵ is clear in this model; with $\epsilon = 0$ a grand canonical ensemble calculation gives $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle = 0$; on the other hand, if $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ is interpreted as the quasiaverage (3), it is not zero and in fact gets large like $\langle N_{\mathbf{k}}\rangle$ as $\mathbf{k} \to 0$, for arbitrarily small but nonzero ϵ and $T < T_c$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Nonlinear Perturbations*

TIMOTHY P. COFFEY† AND G. W. FORD
Department of Physics, The University of Michigan, Ann Arbor, Michigan

(Received 12 June 1967)

The perturbation theory of Bogoliubov and Mitropolsky for systems having a single rapid phase is generalized to systems having several rapid phases. It is shown that one can avoid the classic problem of small divisors to all orders in the perturbation theory. The method has the advantage of providing a single approach to many problems conventionally treated by a variety of specialized techniques.

1. INTRODUCTION

The techniques of perturbation theory for nonlinear systems, initiated by Poincaré three quarters of a century ago, have been extended and developed by many workers. One such technique, the method of averaging, was introduced by Krylov and Bogoliubov thirty years ago.[1] The essential feature of this method is the separation of a given motion into a secular motion plus a rapidly fluctuating motion of small amplitude; the given motion is then expressed in terms of the solution of a system of differential equations which describe the secular motion alone. A wide variety of physical problems may be handled by this method, e.g., Case in a recent publication has shown how the method can be applied to time-dependent perturbation theory in quantum mechanics.[2]

Bogoliubov and Mitropolsky have presented a form of the method of averaging, called the method of rapidly rotating phase, which is especially convenient for systems in which a single variable, called the phase, has a rapid secular motion.[3] Our purpose in this paper is to extend this method to systems with

* Supported in part by the National Science Foundation.

† Present address: E. G. and G. Inc., Arlington, Va.

[1] N. N. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Academy of Sciences of the Ukranian S.S.R., Kiev, 1937), trans. by S. Lefschetz in *Annals of Mathematics Studies*, No. 11 (Princeton University Press, Princeton, N.J., 1947).

[2] K. M. Case, Suppl. Progr. Theoret. Phys. (Kyoto) **37**, 1 (1966). See also R. Y. Y. Lee, "On a New Perturbation Method" Thesis, The University of Michigan, Ann Arbor, 1964.

[3] N. N. Bogoliubov and D. N. Zubarev, Ukrain. Mat. Zh. **7**, 5 (1955); N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations* (Hindustan Publishing Co., Delhi, India, 1961), Chaps. 5 and 6. See also N. Minorsky, *Nonlinear Oscillations* (D. van Nostrand Co., Princeton, N.J., 1962), and M. Kruskal, J. Math. Phys. **3**, 806 (1962).

in a very small neighborhood of $\mathbf{k} = 0$, corresponding to the fact that $\beta\omega_{\mathbf{k}}$ becomes very small in that neighborhood.[7] Then it follows from (A6) that $w(\mathbf{k})$ differs infinitesimally from $2|\epsilon|\,\sigma(\mathbf{k})$ in this neighborhood.[21] Thus by (A8) and (A9) one finds, upon replacing $e^{\beta\omega(\mathbf{k})}$ by the first two terms of its power-series expansion,

$$n_{\mathbf{k}} \approx 2|\epsilon|\,\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}),$$
$$\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle \approx -2\epsilon^*\sigma(\mathbf{k})\kappa T/\omega^2(\mathbf{k}), \qquad \text{(A11)}$$

for $T < T_c$ and $\mathbf{k} \to 0$. More explicit results can be obtained by solving (A4) for $n_{\mathbf{k}}$ after the substitutions $\alpha = -|\alpha|$, $w(\mathbf{k}) \approx 2|\epsilon|\,\sigma(\mathbf{k})$ [the latter following from $\omega(\mathbf{k}) \approx 0$]. This gives

$$n_{\mathbf{k}} = \tfrac{1}{2}(k_s^2 - k^2)\Omega/4\pi|\alpha|, \quad T < T_c, \quad k < k_s, \qquad \text{(A12)}$$

where

$$k_s = \{2[\mu + 4\pi\rho|\alpha| + 2|\epsilon|\,\sigma(\mathbf{k})]\}^{\frac{1}{2}}. \qquad \text{(A13)}$$

[21] We assume that $\sigma(\mathbf{k}) \geq 0$.

Then, by (A11),

$$\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle = -\tfrac{1}{2}e^{-i\vartheta}(k_s^2 - k^2)\Omega/4\pi|\alpha|, \quad T < T_c, \quad k < k_s, \qquad \text{(A14)}$$

where $\epsilon = |\epsilon|\,e^{i\vartheta}$. Equations (5) and (6) then give

$$p_c = \rho_c = f\rho, \quad T < T_c,$$
$$p_c = \rho_c = 0, \quad T > T_c, \qquad \text{(A15)}$$

where[11]

$$(1 - f)\rho = 2.612(\kappa T/2\pi)^{\frac{3}{2}}, \quad T < T_c, \qquad \text{(A16)}$$
$$\kappa T_c = 2\pi(\rho/2.612)^{\frac{2}{3}},$$

and

$$\lim_{\epsilon\to 0} k_s = [15(2\pi)^3 f\rho|\alpha|/\Omega]^{\frac{1}{5}}, \quad T < T_c. \qquad \text{(A17)}$$

The infinite susceptibility of $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ to the symmetry-breaking perturbation V_ϵ is clear in this model; with $\epsilon = 0$ a grand canonical ensemble calculation gives $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle = 0$; on the other hand, if $\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle$ is interpreted as the quasiaverage (3), it is not zero and in fact gets large like $\langle N_{\mathbf{k}}\rangle$ as $\mathbf{k} \to 0$, for arbitrarily small but nonzero ϵ and $T < T_c$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Nonlinear Perturbations*

TIMOTHY P. COFFEY† AND G. W. FORD
Department of Physics, The University of Michigan, Ann Arbor, Michigan

(Received 12 June 1967)

The perturbation theory of Bogoliubov and Mitropolsky for systems having a single rapid phase is generalized to systems having several rapid phases. It is shown that one can avoid the classic problem of small divisors to all orders in the perturbation theory. The method has the advantage of providing a single approach to many problems conventionally treated by a variety of specialized techniques.

1. INTRODUCTION

The techniques of perturbation theory for nonlinear systems, initiated by Poincaré three quarters of a century ago, have been extended and developed by many workers. One such technique, the method of averaging, was introduced by Krylov and Bogoliubov thirty years ago.[1] The essential feature of this method is the separation of a given motion into a secular motion plus a rapidly fluctuating motion of small amplitude; the given motion is then expressed in terms of the solution of a system of differential equations which describe the secular motion alone. A wide variety of physical problems may be handled by this method, e.g., Case in a recent publication has shown how the method can be applied to time-dependent perturbation theory in quantum mechanics.[2]

Bogoliubov and Mitropolsky have presented a form of the method of averaging, called the method of rapidly rotating phase, which is especially convenient for systems in which a single variable, called the phase, has a rapid secular motion.[3] Our purpose in this paper is to extend this method to systems with

* Supported in part by the National Science Foundation.

† Present address: E. G. and G. Inc., Arlington, Va.

[1] N. N. Krylov and N. N. Bogoliubov, *Introduction to Non-Linear Mechanics* (Academy of Sciences of the Ukranian S.S.R., Kiev, 1937), trans. by S. Lefschetz in *Annals of Mathematics Studies*, No. 11 (Princeton University Press, Princeton, N.J., 1947).

[2] K. M. Case, Suppl. Progr. Theoret. Phys. (Kyoto) **37**, 1 (1966). See also R. Y. Y. Lee, "On a New Perturbation Method" Thesis, The University of Michigan, Ann Arbor, 1964.

[3] N. N. Bogoliubov and D. N. Zubarev, Ukrain. Mat. Zh. **7**, 5 (1955); N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations* (Hindustan Publishing Co., Delhi, India, 1961), Chaps. 5 and 6. See also N. Minorsky, *Nonlinear Oscillations* (D. van Nostrand Co., Princeton, N.J., 1962), and M. Kruskal, J. Math. Phys. **3**, 806 (1962).

several rapid phases and to succinctly describe how the method works, first in the nondegenerate case, and then in the more interesting degenerate case. In a following paper one of us (T. P. C.) will use the method to discuss the perturbation by a weak, transverse, spatially periodic magnetic field of the motion of a charged particle gyrating in a uniform magnetic field.[4]

2. NONDEGENERATE PERTURBATION THEORY

We consider the following set of coupled differential equations[5]:

$$\dot{x}_i = \epsilon A_i(\mathbf{x},\boldsymbol{\psi}), \qquad i = 1, 2, \cdots, r, \quad (2.1a)$$

$$\dot{\psi}_j = \omega_j(x) + \epsilon B_j(\mathbf{x},\boldsymbol{\psi}), \quad j = 1, 2, \cdots, s, \quad (2.1b)$$

where ϵ is a small parameter, $\mathbf{x} = (x_1, \cdots, x_r)$, $\boldsymbol{\psi} = (\psi_1, \cdots, \psi_s)$, and the A_i's and B_j's are periodic functions of each of the ψ_k's with period 2π. The dot represents differentiation with respect to time.

When $\epsilon = 0$, the x_i's will be constants and the ψ_j's will be linear functions of time. When ϵ is small but finite, the x_i's will experience a slow secular growth on which is superimposed small-amplitude rapid fluctuations. Similarly, the ψ_j's will experience a rapid secular growth on which is superimposed small-amplitude rapid fluctuations. Our aim is to separate this secular motion from the rapid fluctuating motion. To do this we seek a solution in the form

$$x_i = y_i + \sum_{n=1}^{\infty} \epsilon^n F_i^{(n)}(\mathbf{y},\boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \quad (2.2a)$$

$$\psi_j = \phi_j + \sum_{n=1}^{\infty} \epsilon^n G_j^{(n)}(\mathbf{y},\boldsymbol{\phi}), \quad j = 1, 2, \cdots, s, \quad (2.2b)$$

where the $F_i^{(n)}$'s and $G_j^{(n)}$'s are periodic functions of each of the ϕ_k's with period 2π. We further require that the new variables y_i and ϕ_j satisfy the following differential equations:

$$\dot{y}_i = \sum_{n=1}^{\infty} \epsilon^n a_i^{(n)}(\mathbf{y}), \qquad i = 1, 2, \cdots, r, \quad (2.3a)$$

$$\dot{\phi}_j = \omega_j(\mathbf{y}) + \sum_{n=1}^{\infty} \epsilon^n b_j^{(n)}(\mathbf{y}), \quad j = 1, 2, \cdots, s, \quad (2.3b)$$

where the right-hand sides of Eq. (2.3) are required to be independent of the ϕ_k's. The idea here is that the y_i and ϕ_j exhibit only secular motion, since they are solutions of a system of differential equations which are independent of the rapidly increasing (or decreasing) phases ϕ_j. The rapid fluctuations of the x_i and ψ_j about the y_i and ϕ_j are given by the terms in the series in (2.2). We must now show that we can construct the function $F_i^{(n)}$, $G_j^{(n)}$, $a_i^{(n)}$, and $b_j^{(n)}$ so that (2.2) is indeed a solution of the set of differential equations (2.1).

If we insert Eq. (2.2) in Eq. (2.1) and then use Eq. (2.3) we find, upon equating equal powers of ϵ,

$$a_i^{(1)}(\mathbf{y}) + \sum_{k=1}^{s} \omega_k \frac{\partial F_i^{(1)}(\mathbf{y},\boldsymbol{\phi})}{\partial \phi_k} = A_i(\mathbf{y},\boldsymbol{\phi}), \quad (2.4a)$$

$$b_j^{(1)}(\mathbf{y}) + \sum_{k=1}^{s} \omega_k \frac{\partial G_j^{(1)}(\mathbf{y},\boldsymbol{\phi})}{\partial \phi_k} = B_j(\mathbf{y},\boldsymbol{\phi}) + \sum_{l=1}^{r} F_l^{(1)}(\mathbf{y},\boldsymbol{\phi}) \frac{\partial \omega_j(\mathbf{y})}{\partial y_l}, \quad (2.4b)$$

from the first power of ϵ, and

$$a_i^{(2)} + \sum_{k=1}^{s} \omega_k \frac{\partial F_i^{(2)}}{\partial \phi_k} = \sum_{k=1}^{s} G_k^{(1)} \frac{\partial A_i}{\partial \phi_k} + \sum_{l=1}^{r} F_l^{(1)} \frac{\partial A_i}{\partial y_l} - \sum_{l=1}^{r} a_l^{(1)} \frac{\partial F_i^{(1)}}{\partial y_l} - \sum_{k=1}^{s} b_k^{(1)} \frac{\partial F^{(1)}}{\partial \phi_k}, \quad (2.5a)$$

$$b_j^{(2)} + \sum_{k=1}^{s} \omega_k \frac{\partial G_j^{(2)}}{\partial \phi_k} = \frac{1}{2} \sum_{l=1}^{r} \sum_{m=1}^{r} F_l^{(1)} F_m^{(1)} \frac{\partial^2 \omega_j}{\partial y_l \partial y_m} + \sum_{l=1}^{r} F_l^{(2)} \frac{\partial \omega_j}{\partial y_l} + \sum_{k=1}^{s} G_k^{(1)} \frac{\partial B_j}{\partial \phi_k} + \sum_{l=1}^{r} F_l^{(1)} \frac{\partial B_j}{\partial y_l} - \sum_{l=1}^{r} a_l^{(1)} \frac{\partial G_j^{(1)}}{\partial y_l} - \sum_{k=1}^{s} b_k^{(1)} \frac{\partial G^{(1)}}{\partial \phi_k}, \quad (2.5b)$$

from the second power of ϵ, and so on. We thus obtain a sequence of equations for the determination of the unknown functions.

Each of these equations is of the general form

$$a(\mathbf{y}) + \sum_{k=1}^{s} \omega_k(y) \frac{\partial F(\mathbf{y},\boldsymbol{\phi})}{\partial \phi_k} = A(\mathbf{y},\boldsymbol{\phi}), \quad (2.6)$$

where $a(\mathbf{y})$ and $F(\mathbf{y},\boldsymbol{\phi})$ are to be determined and $A(\mathbf{y},\boldsymbol{\phi})$ is a periodic function of the ϕ_k which is known in terms of the solutions of the previous equations. Note that the dependence upon $\mathbf{y}$ is trivial, the y_i behaving as parameters in this equation, so we may suppress this dependence for the moment and write the equation in the form

$$\sum_{k=1}^{s} \omega_k \frac{\partial F(\boldsymbol{\phi})}{\partial \phi_k} = A(\boldsymbol{\phi}) - a. \quad (2.7)$$

This equation, viewed as an equation for determining $F(\boldsymbol{\phi})$, is a first order, linear, inhomogeneous partial differential equation with constant coefficients. Solutions of such an equation exist only if the inhomogeneous term is orthogonal to all solutions of the

[4] T. P. Coffey, J. Math. Phys. **10**, 1362 (1969).

[5] The generalization to the general case where the right-hand sides of (2.1) are power series in ϵ is straightforward.

homogeneous equation:

$$\sum_{k=1}^{s} \omega_k \frac{\partial F(\boldsymbol{\phi})}{\partial \phi_k} = 0. \tag{2.8}$$

But the solutions of this equation are all of the form

$$F(\boldsymbol{\phi}) = \exp\left\{ i \sum_{k=1}^{s} p_k \phi_k \right\}, \tag{2.9}$$

where, because $F(\boldsymbol{\phi})$ must be periodic in each of the ϕ_k, the p_k must be integers and, because (2.9) must be a solution of (2.8), these integers must satisfy the identity

$$\sum_{k=1}^{s} p_k \omega_k = 0. \tag{2.10}$$

In the nondegenerate case we assume there are no sets of integers satisfying this identity except for the trivial set in which all the p_k are zero, i.e., $F(\boldsymbol{\phi})$ is a constant. If there is a nontrivial set of integers satisfying (2.10) we say there is a degeneracy; we discuss this case in Sec. 4.

We see, therefore, that in the nondegenerate case a must be chosen so there is no constant term on the right-hand side of (2.7); $F(\boldsymbol{\phi})$ is then the solution of the resulting differential equation. To exhibit this solution more explicitly, we return to Eq. (2.6) where the $\mathbf{y}$ dependence is indicated. The given function $A(\mathbf{y}, \boldsymbol{\phi})$, since it is periodic in the ϕ_k, must be expressible in the form

$$A(\mathbf{y}, \boldsymbol{\phi}) = \sum_{\mathbf{p}} A_{\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \tag{2.11}$$

where the sum is over all sets of integers $\mathbf{p} = (p_1, p_2, \cdots, p_s)$ and

$$\mathbf{p}\cdot\boldsymbol{\phi} \equiv \sum_{k=1}^{s} p_k \phi_k. \tag{2.12}$$

The function $a(\mathbf{y})$ must be chosen to cancel the terms corresponding to $\mathbf{p} = \mathbf{0}$, in which all the p_k are zero:

$$a(\mathbf{y}) = A_0(\mathbf{y}) = \frac{1}{(2\pi)^s} \int_0^{2\pi} d\phi_1 \int_0^{2\pi} d\phi_2 \cdots \int_0^{2\pi} d\phi_s A(\mathbf{y}, \boldsymbol{\phi}). \tag{2.13}$$

The solution of (2.6) is then

$$F(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{\mathbf{p}}{}' \frac{A_{\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} e^{i\mathbf{p}\cdot\boldsymbol{\phi}} + f(\mathbf{y}), \tag{2.14}$$

where the prime indicates that the term $\mathbf{p} = \mathbf{0}$ is absent from the sum and

$$\mathbf{p}\cdot\boldsymbol{\omega} \equiv \sum_{k=1}^{s} p_k \omega_k(y). \tag{2.15}$$

The function $f(\mathbf{y})$ in (2.14) is arbitrary; the solution of an inhomogeneous, linear, partial differential equation is determined only up to an arbitrary solution of the homogeneous equation. We usually choose $f(\mathbf{y})$ to be zero.

Thus, we see how the two functions $a(\mathbf{y})$, given by (2.13), and $F(\mathbf{y}, \boldsymbol{\phi})$, given by (2.14), are determined from the single equation (2.6). Since each of the equations in the sequence for the determination of the functions $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $a_i^{(n)}(\mathbf{y})$, and $b_j^{(n)}(\mathbf{y})$ is of the form (2.6), we may, in the nondegenerate case, successively solve to determine these functions. To be more explicit, we first note that the given functions $A_i(\mathbf{x}, \boldsymbol{\psi})$ and $B_j(\mathbf{x}, \boldsymbol{\psi})$ in (2.1) are periodic in each of the ψ_k and so may be expanded in the form

$$A_i(\mathbf{x}, \boldsymbol{\psi}) = \sum_{\mathbf{p}} A_{i,\mathbf{p}}(\mathbf{x}) e^{i\mathbf{p}\cdot\boldsymbol{\psi}}, \tag{2.16a}$$

$$B_j(\mathbf{x}, \boldsymbol{\psi}) = \sum_{\mathbf{p}} B_{j,\mathbf{p}}(\mathbf{x}) e^{i\mathbf{p}\cdot\boldsymbol{\psi}}. \tag{2.16b}$$

Then from (2.4a) we find

$$a_i^{(1)}(\mathbf{y}) = A_{i,0}(\mathbf{y}) \tag{2.17}$$

and

$$F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{\mathbf{p}}{}' \frac{A_{i,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} e^{i\mathbf{p}\cdot\boldsymbol{\phi}}. \tag{2.18}$$

Using this solution in (2.4b) we then find

$$b_j^{(1)}(\mathbf{y}) = B_{j,0}(\mathbf{y}) \tag{2.19}$$

and

$$G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{\mathbf{p}}{}' \left[\frac{B_{j,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} - i \frac{\sum_{l=1}^{r} \frac{\partial \omega_j(\mathbf{y})}{\partial y_l} A_{l,\mathbf{p}}(\mathbf{y})}{(\mathbf{p}\cdot\boldsymbol{\omega})^2} \right] e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \tag{2.20}$$

and so on. The expressions become increasingly cumbersome, but we can, in principle, solve to obtain explicit expressions for the $F_i^{(n)}$, $G_j^{(n)}$, $a_i^{(n)}$, and $b_j^{(n)}$ so (2.2) is a solution of the system of equations (2.1) to any desired order in ϵ.

3. THE VAN DER POL EQUATION

As a simple example illustrating the working of the general method for the nondegenerate case, we consider the van der Pol equation

$$\ddot{z} + \epsilon(z^2 - 1)\dot{z} + z = 0. \tag{3.1}$$

We cast this equation into the standard form (2.1) by introducing variables x and ψ through the substitution:

$$z = x^{\frac{1}{2}} \cos\psi, \qquad \dot{z} = -x^{\frac{1}{2}} \sin\psi, \tag{3.2}$$

that is,

$$x = z^2 + \dot{z}^2, \qquad \psi = -\arctan \dot{z}/z. \tag{3.3}$$

Forming the time derivative of both sides of this last pair of equations and using (3.1) and (3.2) on the right-hand sides, we find

$$\dot{x} = 2\epsilon x(1 - x\cos^2\psi)\sin^2\psi$$
$$= \epsilon x(1 - \tfrac{1}{4}x - \cos 2\psi + \tfrac{1}{4}x\cos 4\psi) \quad (3.4a)$$

$$\dot{\psi} = 1 + \epsilon(1 - x\cos^2\psi)\sin\psi\cos\psi$$
$$= 1 + \epsilon[(\tfrac{1}{2} - \tfrac{1}{4}x)\sin 2\psi - \tfrac{1}{8}x\sin 4\psi]. \quad (3.4b)$$

These equations are in the standard form (2.1) for applying the method of rapidly rotating phase, with ϵ a small parameter.

According to our general method, we seek a solution in the form

$$x = y + \epsilon F^{(1)}(y, \phi) + \epsilon^2 F^{(2)}(y, \phi) + \cdots, \quad (3.5a)$$

$$\psi = \phi + \epsilon G^{(1)}(y, \phi) + \epsilon^2 G^{(2)}(y, \phi) + \cdots, \quad (3.5b)$$

where

$$\dot{y} = \epsilon a^{(1)}(y) + \epsilon^2 a^{(2)}(y) + \cdots, \quad (3.6a)$$

$$\dot{\phi} = 1 + \epsilon b^{(1)}(y) + \epsilon^2 b^{(2)}(y) + \cdots. \quad (3.6b)$$

Inserting (3.5) in (3.4), using (3.6), and equating powers of ϵ, we get the following sequence of equations.

$$a^{(1)} + \frac{\partial F^{(1)}}{\partial\phi} = y(1 - \tfrac{1}{4}y - \cos 2\phi + \tfrac{1}{4}y\cos 4\phi), \quad (3.7a)$$

$$b^{(1)} + \frac{\partial G^{(1)}}{\partial\phi} = (\tfrac{1}{2} - \tfrac{1}{4}y)\sin 2\phi - \tfrac{1}{8}y\sin 4\phi, \quad (3.7b)$$

from the first power of ϵ, and

$$a^{(2)} + \frac{\partial F^{(2)}}{\partial\phi} = (1 - \tfrac{1}{2}y - \cos 2\phi + \tfrac{1}{2}y\cos 4\phi)F^{(1)}$$
$$+ (2y\sin 2\phi - y^2\sin 4\phi)G^{(1)}$$
$$- a^{(1)}\frac{\partial F^{(1)}}{\partial y} - b^{(1)}\frac{\partial F^{(1)}}{\partial\phi}, \quad (3.8a)$$

$$b^{(2)} + \frac{\partial G^{(2)}}{\partial\phi} = -(\tfrac{1}{4}\sin 2\phi + \tfrac{1}{8}\sin 4\phi)F^{(1)}$$
$$+ [(1 - \tfrac{1}{2}y)\cos 2\phi - \tfrac{1}{2}y\cos 4\phi]G^{(1)}$$
$$- b^{(1)}\frac{\partial G^{(1)}}{\partial\phi} - a^{(1)}\frac{\partial G^{(1)}}{\partial y}, \quad (3.8b)$$

from the second power of ϵ, and so on.

We solve this sequence of equations as indicated in the previous section. From (3.7a) we find

$$a^{(1)}(y) = y(1 - \tfrac{1}{4}y), \quad (3.9a)$$

$$F^{(1)}(y, \phi) = y(-\tfrac{1}{2}\sin 2\phi + \tfrac{1}{16}y\sin 4\phi), \quad (3.9b)$$

while from (3.7b) we find

$$b^{(1)}(y) = 0, \quad (3.9c)$$

$$G^{(1)}(y) = -\tfrac{1}{4}(1 - \tfrac{1}{2}y)\cos 2\phi + \tfrac{1}{32}y\cos 4\phi. \quad (3.9d)$$

From (3.8a), using the solutions (3.9), we find

$$a^{(2)}(y) = 0, \quad (3.10a)$$

$$F^{(2)}(y, \phi) = y^2(\tfrac{1}{32}(y - 5)\cos 2\phi - \tfrac{1}{64}\cos 4\phi + \tfrac{1}{96}y\cos 6\phi), \quad (3.10b)$$

while from (3.8b) we find

$$b^{(2)}(y) = -\frac{1}{8} + \frac{3y}{16} - \frac{11y^2}{256}, \quad (3.10c)$$

$$G^{(2)}(y, \phi) = -\frac{y(1 + y)}{128}\sin 2\phi - \frac{16 - 4y + 3y^2}{512}\sin 4\phi$$
$$+ \frac{y(3 - 2y)}{384}\sin 6\phi - \frac{y^2}{2048}\sin 8\phi. \quad (3.10d)$$

These expressions, when inserted in (3.5) and (3.6), give the complete reduction of the problem through second order in ϵ.

The method of rapidly rotating phase does not in general lead to an explicit solution of the original set of differential equations. Rather, it is a method for separating the secular motion from the rapid periodic fluctuations and reducing the problem to that of solving the differential equations for the secular motion alone. The solution of these equations, i.e., in the general case the equations (2.3), may be a very difficult problem, but in the case of the van der Pol equation it is quite simple. Using (3.9a) and (3.10a), the differential equation (3.6a) becomes

$$\dot{y} = \epsilon y(1 - \tfrac{1}{4}y) \quad (3.11)$$

through second order in ϵ. The solution is

$$y(t) = \frac{4y(0)}{y(0) + [4 - y(0)]e^{-\epsilon t}}. \quad (3.12)$$

Here we see the well-known feature of the van der Pol equation: for long times the amplitude approaches a constant independent of the initial amplitude. Inserting this solution in (3.6b) we can integrate to find

$$\phi(t) = \phi(0) + \left(1 - \frac{\epsilon^2}{16}\right)t$$
$$+ \frac{\epsilon}{16}\log\frac{y(0) + [4 - y(0)]e^{-\epsilon t}}{4}$$
$$+ \frac{11\epsilon}{64}y^2(0)\frac{1 - e^{-\epsilon t}}{y(0) + [4 - y(0)]e^{-\epsilon t}}. \quad (3.13)$$

Here we see there is a shift in the frequency of the rapid phase together with a slow secular shift of the phase.

This discussion of the van der Pol equation is only intended to be illustrative of the method. We refer, for example, to a recent paper by Struble and Fletcher, who give a much more thorough discussion of the van der Pol equation using a somewhat different method.[6]

4. DEGENERATE PERTURBATION THEORY

In the degenerate case we must consider what changes must be made when there is a nontrivial set of integers satisfying (2.10). More generally, we must consider the situation when

$$|\mathbf{p}\cdot\boldsymbol{\omega}| < O(\epsilon). \tag{4.1}$$

That is, the case when the factors in the denominators of our solutions, e.g., (2.13) or (2.16) or (2.18), are small of order ϵ. When this occurs the successive terms in the series (2.2) are no longer small if ϵ is small; they no longer represent small amplitude fluctuations of the given motion about the mean motion. This so-called problem of small divisors is, of course, closely related to the degeneracy problem for which the divisors are zero.

The solution of this problem is already indicated by our discussion of Eq. (2.6) in the nondegenerate case. There we saw that the function $a(\mathbf{y})$ has to be chosen to cancel the terms in $A(\mathbf{y}, \boldsymbol{\phi})$ which correspond to solutions of the homogeneous equation (2.8). In the nondegenerate case, the only such term was the constant term, but in the degenerate or near degenerate case we must cancel all the terms corresponding to sets of integers satisfying (4.1). That is, we generalize to allow $a(\mathbf{y}, \boldsymbol{\phi})$ to depend upon those combinations of the ϕ_j which give rise to small divisors and then choose $a(\mathbf{y}, \boldsymbol{\phi})$ to cancel those terms in $A(\mathbf{y}, \boldsymbol{\phi})$.

Our procedure is formally similar to the nondegenerate case. We seek a solution of (2.1) in the form

$$x_i = y_i + \sum_{n=1}^{\infty} \epsilon^n F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \tag{4.2a}$$

$$\psi_j = \phi_j + \sum_{n=1}^{\infty} \epsilon^n G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad j = 1, 2, \cdots, s, \tag{4.2b}$$

where the $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ and $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ are periodic functions of each of the ϕ_k. We further require that

$$\dot{y}_i = \sum_{n=1}^{\infty} \epsilon^n a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad i = 1, 2, \cdots, r, \tag{4.3a}$$

$$\dot{\phi}_j = \omega_j(\mathbf{y}) + \sum_{n=1}^{\infty} \epsilon^n b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi}), \quad j = 1, 2, \cdots, s. \tag{4.3b}$$

[6] R. A. Struble and J. E. Fletcher, J. Math. Phys. **2**, 880 (1961). See also N. Minorsky, Ref. 2, pp. 219–224 and pp. 329–338.

Inserting (4.2) in (2.1) and using (4.3) we find, upon equating powers of ϵ,

$$a_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) + \sum_{k=1}^{s} \omega_k(\mathbf{y}) \frac{\partial F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k} = A_i(\mathbf{y}, \boldsymbol{\phi}), \tag{4.4}$$

$$b^{(1)}(\mathbf{y}, \boldsymbol{\phi}) + \sum_{k=1}^{s} \omega_k(\mathbf{y}) \frac{\partial G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi})}{\partial \phi_k} = B_j(\mathbf{y}, \boldsymbol{\phi}) + \sum_{l=1}^{r} \frac{\partial \omega_j(\mathbf{y})}{\partial y_l} F_l^{(1)}(\mathbf{y}, \boldsymbol{\phi}), \tag{4.5}$$

and so on. The sequence of equations we obtain differs from that in the nondegenerate case only in that the $a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ and $b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ depend upon $\boldsymbol{\phi}$ as well as $\mathbf{y}$. The formal solution of these equations is straightforward. Using again the expansions (2.16), from (4.4) we obtain:

$$a_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = \sum_{|\mathbf{p}\cdot\boldsymbol{\omega}| < O(\epsilon)} A_{i,\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \tag{4.6}$$

where the sum is over all sets of integers fulfilling (4.1), and

$$F_i^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{|\mathbf{p}\cdot\boldsymbol{\omega}| > O(\epsilon)} \frac{A_{i,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} e^{i\mathbf{p}\cdot\boldsymbol{\phi}}, \tag{4.7}$$

where the sum is over all sets of integers not contained in the sum in (4.6). Continuing, from (4.5) we obtain

$$b_j^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = \sum_{|\mathbf{p}\cdot\boldsymbol{\omega}| < O(\epsilon)} B_{j,\mathbf{p}}(\mathbf{y}) e^{i\mathbf{p}\cdot\boldsymbol{\phi}} \tag{4.8}$$

and

$$G_j^{(1)}(\mathbf{y}, \boldsymbol{\phi}) = -i \sum_{|\mathbf{p}\cdot\boldsymbol{\omega}| > O(\epsilon)} \left[\frac{B_{j,\mathbf{p}}(\mathbf{y})}{\mathbf{p}\cdot\boldsymbol{\omega}} - i \frac{\sum_{l=1}^{r} \frac{\partial \omega_j(\mathbf{y})}{\partial y_l} A_{l,\mathbf{p}}(\mathbf{y})}{(\mathbf{p}\cdot\boldsymbol{\omega})^2} \right] e^{i\mathbf{p}\cdot\boldsymbol{\phi}}. \tag{4.9}$$

It should be clear that in this manner we can successively solve the equations for the determination of the $F_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $G_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, $a_i^{(n)}(\mathbf{y}, \boldsymbol{\phi})$, and $b_j^{(n)}(\mathbf{y}, \boldsymbol{\phi})$ to obtain explicit expressions in which small divisors do not occur. Of course, the equations (4.3) for the determination of the secular motion are more complicated than the corresponding equations in the nondegenerate case; they explicitly involve certain combinations of the ϕ_k. However, these equations still describe the slowly varying secular motion, since those combinations of the ϕ_k which do appear are themselves slowly varying. Thus, if the combination $(\mathbf{p}\cdot\boldsymbol{\phi})$ appears in (4.3), then

$$(\mathbf{p}\cdot\dot{\boldsymbol{\phi}}) \cong (\mathbf{p}\cdot\boldsymbol{\omega}) < O(\epsilon), \tag{4.10}$$

i.e., this combination is slowly varying in exactly the

same sense that the y_k are slowly varying. The basic idea of the expansion (4.2), or of (2.2), is the separation of the secular motion from the rapidly fluctuating motion, and this separation is preserved in the degenerate case.

There is, however, a serious difficulty in our general formulation of the degenerate perturbation problem. This is the question of deciding which combinations of the ϕ_k are to be included in the secular motion. For any particular set of ω_k, we can always find a set of integers p_k such that $(\mathbf{p} \cdot \boldsymbol{\omega})$ is as close as we please to any real number. That is, the values of $(\mathbf{p} \cdot \boldsymbol{\omega})$ are dense in the whole range

$$-\infty < \mathbf{p} \cdot \boldsymbol{\omega} < \infty. \tag{4.11}$$

This means that we cannot, in general, make a sharp separation between the terms for which $|\mathbf{p} \cdot \boldsymbol{\omega}| < O(\epsilon)$, which we put into the secular motion, and the remaining terms, which we put into the fluctuating motion. We can do so only if the coefficients $A_{i,\mathbf{p}}(\mathbf{y})$ and $B_{j,\mathbf{p}}(\mathbf{y})$ vanish sufficiently rapidly for large values of $|\mathbf{p}| \equiv (p_1^2 + p_2^2 + \cdots + p_s^2)^{\frac{1}{2}}$. The point here is that, for the most general functions $A_i(\mathbf{x}, \boldsymbol{\phi})$ and $B_j(\mathbf{x}, \boldsymbol{\phi})$ in Eq. (2.1), it is not possible to sharply separate the secular motion from the periodic fluctuations; when these functions are such as to allow a sharp separation, the method we have outlined will work.

5. CONCLUSION

The method of rapidly rotating phase which we have presented here is applicable to a wide range of physical problems. On the one hand, it can be shown to be equivalent to classical perturbation theory of Hamiltonian systems, at least in the nondegenerate case.[7] On the other hand, Rayleigh–Schrödinger perturbation theory in quantum mechanics is also a special case. In both cases the treatment of degeneracy or near-degeneracy is generally simplest by the method of rapidly rotating phase. Thus, the advantages of the method are its generality and its simplicity.

Of course, not all perturbation problems can be cast into the form of a set of coupled differential equations in the standard form (2.1) appropriate for the method. In general, we can say that the method is suited for the discussion of small perturbations of periodic or multiple-periodic motions, but we cannot precisely characterize such problems.

As we remarked earlier, an aspect of the method which may cause difficulty in applications is that the Eqs. (2.3) or (4.2), which describe the slow secular motion, may not be appreciably easier to solve than the original equations. (Here we are speaking of the finite versions of (2.3) or (4.2) which are obtained by truncating the series on the right.) The point is that the method is designed to separate the secular motion from the fluctuating motion; it gives no help in the discussion of the equations for the secular motion. This is a characteristic feature of all averaging methods.

We close with a few remarks about convergence. It should be clear that the method of rapidly rotating phase is asymptotic in the sense that the approximate solution is intended to be valid for long times, i.e., for times of order ϵ^{-1}, the characteristic time of the secular motion. What can be proved is a typical asymptotic convergence theorem: With suitable restrictions on the perturbing functions, the approximate solutions obtained by solving the differential Eqs. (2.3) or (4.2), truncated at a finite order in ϵ, and inserting the resulting solution in (2.2) or (4.1), also truncated, differ from the exact solution by an error which is small but which grows in time like $\exp\{c\epsilon t\}$, with c a constant.[8] This is a rather weak theorem, but we have not been able to improve it in the general case. The question of the convergence of the infinite series in Eqs. (2.2), (2.3), (4.2), and (4.3) is still open.

[7] For a proof see T. P. Coffey, "Analytical Methods in the Theory of Non-Linear Oscillations," thesis, The University of Michigan, Ann Arbor, 1966.

[8] For a proof, see Ref. 7.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Phase-Integral Approximation in Momentum Space and the Bound States of an Atom. II

MARTIN C. GUTZWILLER
IBM Watson Laboratory, Columbia University, New York, New York

(Received 15 November 1968)

The phase-integral approximation of the Green's function in momentum space is investigated for a particle of negative energy (bound state) which moves in a spherically symmetric potential. If this potential has a Coulomb-like singularity at the origin, it is shown that any two momenta can be connected by an infinity of classical trajectories with a fixed energy. The summation of the usual phase and amplitude factors over these trajectories is the approximate Green's function. If there is orbital precession, there are not only poles along the negative energy axis, but also weaker singularities which are not examined in detail. The poles are found at the energies which are given by the semiclassical quantum conditions: angular momentum $= (l + \frac{1}{2})\hbar$ and action integral for the radial motion $= (n + \frac{1}{2})2\pi\hbar$, where l and n are integers ≥ 0. The residues at these poles give the approximate bound-state wavefunctions as a product of the asymptotic formula for Legendre polynomials with the asymptotic solution of the radial Schrödinger equation. It is conjectured that the occurrence of poles in the approximate Green's function is a direct consequence of the periodic character of the classical motion.

INTRODUCTION

The present work continues to pursue an idea which was first discussed in an earlier paper[1] (to be referred henceforth as I). The phase-integral approximation, commonly known as WKB method, was applied to Schrödinger's equation in order to find the approximate Green's function for single electrons in an atomic potential. Attention was focused on momentum space (rather than coordinate space) because there was good reason to believe that such a procedure would improve the approximation for bound states. This hope was completely vindicated in the case of the hydrogen atom where not only the correct bound-state energies, but even the correct bound-state wavefunctions were obtained as the residues at the poles of the approximate Green's function in momentum space.

The formalism in I can be used to discuss the phase-integral approximation for the Green's function in any spherically symmetric potential. The corresponding calculations will be carried out and discussed in the present paper. The emphasis will be again on momentum space, because classical mechanics in momentum space forms a much more convincing basis for an approximation to quantum mechanics in the case of negative energies. The reason is simply that any two momenta p' and p'' can be connected by an infinity of classical trajectories with fixed negative energy E, provided the potential V has singularities of the Coulomb type, i.e., $V \sim |q|^{-1}$ for small q. The phase-integral approximation $\tilde{F}(p''p'E)$ for the Green's function $F(p''p'E)$ does not have any of the obvious limitations which beset the phase-integral approximation $\tilde{G}(q''q'E)$ of the Green's function $G(q''q'E)$ in coordinate space.

[1] M. C. Gutzwiller, J. Math. Phys. 8, 1979 (1967).

The singularities of $\tilde{F}(p''p'E)$ will be examined in order to find the approximate eigenvalues and eigenstates of the particle in the potential $V(q)$. It was found in I that $\tilde{F}(p''p'E)$ has only poles along the negative E axis in the case of the pure Coulomb potential. It will turn out, however, that $\tilde{F}(p''p'E)$ has not only poles, in general, but other singularities of a weaker kind like branch cuts. The poles of $\tilde{F}$ will be determined in this paper. They are given exactly by the ordinary quantum conditions: angular momentum $M = (l + \frac{1}{2})\hbar$ and action integral for the radial motion $\oint p_r\, dr = (n + \frac{1}{2})2\pi\hbar$ with integers l and $n \geq 0$. The corresponding approximate eigenfunctions are the asymptotic expression for the Legendre polynomials and the asymptotic solution of the radial Schrödinger equation. These results hold equally well in momentum and in coordinate space. We find, therefore, the well-known formulas for the WKB method in a completely different manner without the need to make certain additional assumptions which are usually necessary for the separation of variables.

Nevertheless, it is only fair to make the following remarks. The spherical symmetry of the potential $V(q)$ has been used not only to solve the equations of motion in classical mechanics, but also to make use of the periodicity in the angular and radial motion. Indeed, it seems at this point that the phase-integral approximation for the Green's function has poles only if the classical motions are periodic in some sense. If the periodic character of the classical orbits is complete as in the Coulomb potential where all orbits are closed, one has only poles. In the more general case where orbital precession occurs, there are weaker singularities besides the poles. When the classical trajectories have no periodicity at all, the poles will presumably disappear from the approximate Green's

function and only weaker singularities remain along the negative E axis. The understanding and the calculation of these weak singularities presents, obviously, a challenge which is not met in this paper. It would appear that the generalization of the WKB method to nonseparable potentials is found by tackling this difficult problem.

The results of this paper do not seem to throw any light in that direction. This is particularly regrettable because the approximate wavefunctions from the poles of $\tilde{F}$ cover only a limited domain in momentum space exactly as the approximate wavefunctions in coordinate space cover only the classical domain and have strong singularities at its boundary. The original motive for going into momentum space seems thereby lost. Actually, the capability of $\tilde{F}(p''p'E)$ to be different from zero for any pair of momenta p' and p'' is hidden in the weak singularities. They have to be investigated if one tries to go beyond the ordinary WKB method even in the case where the separation of variables is feasible as for a spherically symmetric potential.

The present paper has been arranged as follows. The notations and results of I are recalled in Sec. 1. The possibility of connecting any two momenta by an infinity of trajectories at a given negative energy is discussed in Sec. 2. The principal argument is purely geometric because the trajectories in momentum can be put into a one-to-one correspondence with the geodesics on a surface in Euclidean space which is topologically equivalent to a sphere. A complete enumeration of all the trajectories from p' to p'' is accomplished in Sec. 3. The number of "conjugate times" between p' and p'' is also determined because it fixes the extra phase losses which occur every time the trajectory touches a focal line.

The sum over all classical trajectories can then be written formally in Sec. 4 with the help of the representation that was established in the preceding section. This summation is transformed with the help of Poisson's formula in Sec. 5. The result is particularly simple if we consider only the discontinuity of $\tilde{F}(p''p'E)$ across the negative E axis. One of the new variables of summation is immediately recognized as the discreet angular momentum $(l+\frac{1}{2})\hbar$. The other variable corresponds to the radial quantum number n. The poles of $\tilde{F}(p''p'E)$ are finally obtained in Sec. 6 and are found from the ordinary quantum conditions. The residues are simply the products of the corresponding approximate wavefunctions for p' and p'', including the correct angular dependence.

1. THE BASIC FORMULAS AND NOTATIONS

The semiclassical approximation $\tilde{F}(p''p'E)$ will be constructed for the quantum-mechanical propagator $F(p''p'E)$ of a particle which starts out with a momentum p' and ends up with a momentum p'' while propagating with energy E. The classical Hamiltonian $H(pq)$ is given and we can, therefore, calculate the classical trajectories in momentum space. Along a particular trajectory which goes from p' to p'' at the energy $E = H(pq)$ we can calculate the virial

$$T(p''p'E) = -\int_{p'}^{p''} q\, dp. \tag{1}$$

The approximate propagator $\tilde{F}(p''p'E)$ is given by summing over all classical trajectories from p' to p'' with energy E, namely,

$$\tilde{F}(p''p'E) = -\frac{1}{2\pi\hbar^2} \sum_{\text{classical paths}} (|D_T|)^{\frac{1}{2}} \times \exp\left[i\frac{T}{\hbar} - \text{phases}\right]. \tag{2}$$

D_T is the following 4 by 4 determinant

$$D_T = \begin{vmatrix} \dfrac{\partial^2 T}{\partial p' \partial p''} & \dfrac{\partial^2 T}{\partial p' \partial E} \\ \dfrac{\partial^2 T}{\partial E \partial p''} & \dfrac{\partial^2 T}{\partial E^2} \end{vmatrix}. \tag{3}$$

The "phases" in the formula (2) are related to the number and multiplicity of conjugate times along the particular trajectory in the following manner. Suppose that the particle starts at time $t = 0$ with momentum p' and coordinate q'. Consider the trajectories of the same energy which start simultaneously with momentum p', but whose initial coordinate lies in a element of 2-dimensional surface $d\Omega'$ around q'. At any later time $t > 0$, the momenta of these trajectories lie in a surface element $d\Omega$ of momentum space around the momentum p of the particle. In general, $d\Omega$ will again be two-dimensional; but at certain special times, called "conjugate times," the dimension of $d\Omega$ will be reduced by 1 or 2; these are the simply or doubly conjugate times. The trajectory of the particle is then tangent to a focal surface or focal line. The "phases" are given by $\pi/2$ multiplied by the number of conjugate times along the trajectory from p' to p'', each counted with the proper multiplicity.

The physical significance of these extra phases in (2) is the following. Each focal surface or focal line represents a boundary for those classical trajectories which start with momentum p' and energy E. The particle is allowed, however, to "tunnel" beyond this boundary. The net effect of this quantum-mechanical penetration into the classically inaccessible region is a loss of phase. It is as if the particle had been able to

penetrate by an eighth of a wavelength before being reflected back into the classically accessible region by an infinitely hard wall.

The Hamiltonian is assumed to be the sum of kinetic and potential energy

$$H(pq) = p^2/2m + V(q). \tag{4}$$

In this paper, $V(q)$ is assumed to be spherically symmetric. Moreover, $V(q)$ is assumed to be a monotonically increasing function of $r = |q|$, so that one can invert the relation between V and r. The inverse function will be written as $r(V)$. Each classical trajectory lies in the plane which is defined by the momenta p' and p''. Also, each classical trajectory is characterized by its angular momentum M in addition to its energy E. The coordinates in the plane of the trajectory are chosen such that $M \geq 0$ for the shortest trajectory.

The equations of motion can readily be integrated in polar coordinates. Thus, let $\rho = |p|$ and η be the polar angle of p in the orbital plane, with $\eta = 0$ for p'. The projection σ of q along p is the conjugate variable for ρ and the angular momentum M is the conjugate variable for η. The new Hamiltonian in the orbital plane is given by

$$H = \frac{\rho^2}{2m} + V\left(\left[\sigma^2 + \frac{M^2}{\rho^2}\right]^{\frac{1}{2}}\right), \tag{5}$$

which is to be considered as a function of ρ, σ, and M. The equations of motion

$$\frac{d\sigma}{dt} = +H_\rho, \quad \frac{d\rho}{dt} = -H_\sigma \tag{6}$$

can be greatly simplified if the time t is eliminated, and if ρ (or σ) is used as parameter after eliminating σ (or ρ) with the help of energy conservation $H(\rho\sigma M) = E$. The resulting integrals, however, have singularities whenever ρ (or σ) reach their extreme values compatible with the fixed energy E and angular momentum M. These singularities are harmless and can always be circumvented by going from ρ (or σ) as parameter of integration to its conjugate variable σ (or ρ).

The quantities of interest to the propagator (2) can now be written as

$$T(p''p'E) = \int_{t'}^{t''} \sigma H_\sigma \, dt + M(\eta'' - \eta'), \tag{7}$$

$$\eta'' - \eta' = \eta = \int_{t'}^{t''} H_M \, dt, \tag{8}$$

$$\frac{1}{D_T} = \frac{\rho'\rho'' \sin\eta}{M} \cdot H'_\sigma H''_\sigma \int_{t'}^{t''} \left[H_{MM} - \left(\frac{H_M^2}{H_\sigma}\right)_\sigma\right] dt. \tag{9}$$

Lower indices always mean partial differentiation, whereas primes or double primes indicate that the quantity is to be evaluated at the beginning or at the end of the trajectory, respectively. The integral in (9) can be transformed with the help of the identity

$$\int_{t'}^{t''} \left(\frac{H_M^n}{H_\sigma}\right)_\sigma dt = \int_{t'}^{t''} \left(\frac{H_M^n}{H_\rho}\right)_\rho dt - \frac{H''^n_M}{H''_\sigma H''_\rho} + \frac{H'^n_M}{H'_\sigma H'_\rho}, \tag{10}$$

which is valid for $n = 0, 1, 2$ and serves to avoid the singularities in the integrand of (9).

The conjugate times for t' are determined by either one of two conditions which are sometimes satisfied simultaneously and then yield a doubly counting conjugate time. One condition requires the vanishing of $\sin\eta$, whereas the other condition requires the vanishing of the integral in (9) or, equivalently, the vanishing of $\partial\eta/\partial M$ taken at constant ρ' and ρ''. The occurrence of conjugate times coincides with the vanishing of (9). Since the $-\frac{1}{2}$ power of (9) enters into the amplitude of $\tilde{F}(p''p'E)$, the phases in (2) can be interpreted as follows. Each time D_T changes sign as one follows the particle along its trajectory, the $-\frac{1}{2}$ power of D_T implies a factor $e^{-i\pi/2}$.

2. THE CLASSICAL TRAJECTORIES IN MOMENTUM SPACE

The classical trajectories in coordinate space from q' to q'' at the energy E can be obtained from the variational principle

$$\delta \int_{q'}^{q''} \{2m[E - V(q)]\}^{\frac{1}{2}} \cdot |dq| = 0 \tag{11}$$

due to Jacobi (cf. Ref. 2). The variation is to be taken over all rectifiable curves from q' to q''. The first factor in the integrand of (11) is the absolute value of the momentum of a particle at q with total energy E. The stationary value of the integral gives the action $\int p \, dq$ along the particular trajectory from q' to q''.

Since we are interested mostly in the trajectories in momentum space, it is natural to ask for a generalization of (11) to momentum space. For spherically symmetric potentials one finds in analogy to (11) the variational principle

$$\delta \int_{p'}^{p''} r\left(E - \frac{p^2}{2m}\right) \cdot |dp| = 0. \tag{12}$$

The stationary value of the integral equals the classical virial $T(p''p'E)$. The proofs of either (11) or (12) are

[2] D. Laugwitz, *Differential and Riemannian Geometry* (Academic Press Inc., New York, 1965), p. 172.

straightforward and can be omitted. The simplicity of (11) comes from having p parallel to dq along the trajectory, whereas (12) is guaranteed if q is parallel to dp along the trajectory, i.e., if the potential is spherically symmetric.

The variational principles have a simple geometric interpretation. In (11) we consider the inside of the sphere $V(q) \leq E$ as endowed with the Riemannian metric $2m[E - V(q)]\,dq^2$. The trajectories are then identical with the geodesics in this Riemannian space. Similarly, we endow the momentum space with the metric $[r(E - p^2/2m)]^2\,dp^2$, so that the trajectories can equally well be considered as the geodesics of that second space. If the potential $V(q)$ has an infinity at $q = 0$, the point $q = 0$ has to be removed from the sphere $V(q) \leq E$. There is no such difficulty in momentum space since $r = 0$ is only reached for $|p| = \infty$.

The trajectories in momentum space can be further investigated and understood if an additional assumption is made. $V(q)$ is from now on assumed to have an atomiclike singularity for small r, i.e.,

$$V(q) \approx -Z_0e^2/r, \quad \text{for small } r. \tag{13}$$

Since the potential is spherically symmetric, nothing is lost by considering the particular plane through the origin and the two end points in either coordinate space or momentum space. The crucial question arises whether either one of these 2-dimensional subspaces with their Riemannian metrics can be mapped isometrically onto a surface in 3-dimensional Euclidean space. If so, we would gain much better insight into the geodesics between the two given end points.

This kind of isometric mapping can be accomplished for the momentum space in the following manner. Let the vector p be given by $(\rho \cos \varphi, \rho \sin \varphi)$ with $0 \leq \rho < +\infty$ and $0 \leq \varphi < 2\pi$. There is a one-to-one relationship between ρ and r which is given by

$$\rho^2/2m + V(r) = E \tag{14}$$

for a fixed value of $E < 0$. Since our mind is used to thinking in coordinate space rather than momentum space, r is preferable as a parameter when it comes to intuitive understanding. Therefore, we shall use r rather than ρ to define the isometric surface for momentum space.

Let us consider a cylindrical coordinate system (R, φ, z) in a Euclidean space. The isometric surface is described by giving R and z as functions of r only, and identifying φ with the polar angle in momentum space. Thus the isometric surface has rotational symmetry around the z axis, and the circles of constant z correspond to the circles of constant ρ (or constant r) in momentum space. The Riemannian metric in momentum space has to agree with the natural metric of the surface $(R(r)\cos\varphi,\ R(r)\sin\varphi,\ z(r))$ in Euclidean space, i.e.,

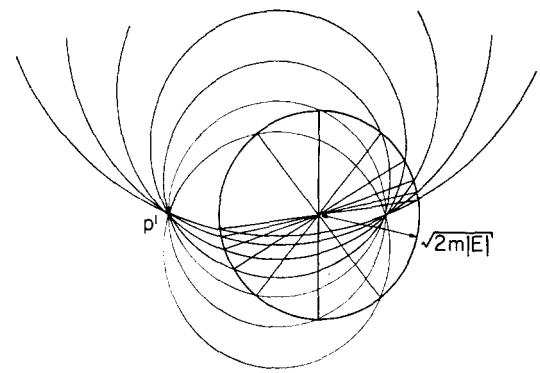


FIG. 1. The Kepler orbits in momentum space through a given initial momentum p' and with energy E. There is exactly one trajectory connecting any given pair p' and p''.

$$r^2(d\rho^2 + \rho^2\,d\varphi^2) = dR^2 + R^2\,d\varphi^2 + dz^2. \tag{15}$$

With the help of (14) we find immediately that

$$R = r\{2m[E - V(r)]\}^{\frac{1}{2}},$$
$$\left(\frac{dz}{dr}\right)^2 = 2m\left(\frac{d(rV)}{dr} - E\right). \tag{16}$$

As a particular example, let us consider the Coulomb potential $V(r) = -Z_0e^2/r$. The second equation (16) gives $z = (2m\,|E|)^{\frac{1}{2}}r$ if we chose $z(0) = 0$. The first equation (16) then yields $R = (-z^2 + 2az)^{\frac{1}{2}}$ with $a = [mZ_0^2e^4/2\,|E|]^{\frac{1}{2}}$. The isometric surface for momentum space turns out to be a sphere of radius a in the case of a Coulomb potential. The mapping from momentum space onto this sphere is a simple stereographic projection, so that circles are mapped into circles. In particular, the great circles of the isometric surface are mapped into those circles in momentum space which intersect the circle $\rho = (2m\,|E|)^{\frac{1}{2}}$ in diametrically opposite points as shown in Fig. 1. Thus we find again the result of I about the shape of the Kepler orbits in momentum space.[3]

The second equation (16) requires that the derivative of $rV(r)$ be larger than E. Such an inequality can be proved in the stronger form $d(rV)/dr \geq 0$ if we make more detailed, but physically reasonable, assumptions

[3] The circular shape of the Kepler orbits in momentum space is, of course, a classical result. It seems, however, that only very few books on classical mechanics mention this important fact. Among them is A. Sommerfeld, *Lectures on Theoretical Physics, Vol. 1: Mechanics* (Academic Press Inc., New York, 1952). A recent discussion of Kepler orbits in momentum space and many relevant references can be found in a series of papers by A. Norcliffe and I. C. Percival, J. Phys. B, Ser. 2, **1**, 774, 784 (1968).

about V. If we let $Z(r)$ be the number of elementary charges inside the sphere of radius r around the nucleus, the electrostatic potential $V(r)$ is given by

$$V(r) = -e^2\int_r^\infty Z(r)\frac{dr}{r^2} = -e^2\left[\frac{Z(r)}{r} + \int_r^\infty \frac{dZ}{dr}\frac{dr}{r}\right]. \tag{17}$$

$Z(r)$ is assumed to be a nonincreasing function of r. Its limit for $r \to 0$ is Z_0. The decrease of $Z(r)$ from Z_0 with increasing r describes the screening of the nuclear charge Z_0 by the electron cloud. We find from (17) that

$$\frac{d(rV)}{dr} = -e^2\int_r^\infty \frac{dZ}{dr}\frac{dr}{r} \geq 0, \tag{18}$$

where the integral is a nonincreasing function of r. Thus, $d(rV)/dr$ reaches its maximum at $r = 0$ where it is given by the integral $-e^2\int_0^\infty (dZ/dr)\,dr/r$. We shall assume that this integral converges or, equivalently, that the screening charge density $(4\pi r^2)^{-1}\,dZ/dr$ goes to infinity less fast than r^{-2} as r approaches zero.

Thus we arrive at the inequality

$$2m\,|E| \leq \left(\frac{dz}{dr}\right)^2 \leq 2m\,|E| - 2me^2\int_0^\infty \frac{dZ}{dr}\frac{dr}{r}, \tag{19}$$

and $z(r)$ is, therefore, a monotonically increasing function of r with a bounded slope no less than $(2m\,|E|)^{\frac{1}{2}}$. The profile of the isometric surface, i.e., the plot of R versus z, is essentially given by the plot of R versus r; in particular, the maxima and minima of R can be obtained from the latter plot. Also, the shape of the profile at its two ends, $r = 0$ and $r = r_0 = r(E)$, is well represented by the behavior of $R(r)$ which goes to zero as $(r)^{\frac{1}{2}}$ and $(r_0 - r)^{\frac{1}{2}}$. The $r^{\frac{1}{2}}$ dependence for small r is a direct consequence of (13).

An instructive example of a typical profile is given by the potential for a nucleus of charge Z_0 which is screened by an infinitely thin shell of electronic charge $Z_0 - 1$ at a distance $r = b$ from the nucleus:

$$\begin{aligned} V(r) &= -e^2[Z_0/r - (Z_0 - 1)/b], && \text{for } 0 < r < b, \\ &= -e^2/r, && \text{for } b < r. \end{aligned} \tag{20}$$

Since the potential is Coulombic in each region, we get two spheres, of radii R_{in} for the inner region and

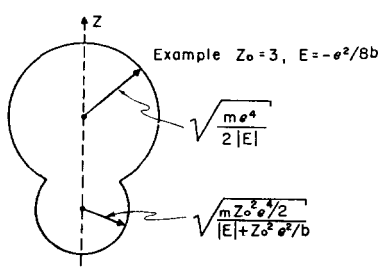


FIG. 2. Cross section through the isometric surface of a hypothetical Li atom where the nuclear charge $3e$ is screened by a thin shell of electronic charge $-2e$ at a distance b.

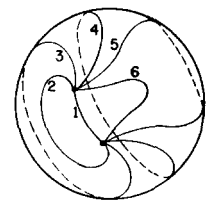


FIG. 3. A one-parameter family of curves on a sphere between given end points whose least upper bound for the lengths is realized by a geodesic of index 1, i.e., not a shortest connection.

R_{out} for the outer region. They join at a radius R which is given according to (16) and (20) with $r = b$. The values of z cover a range z_{in} for the inner region and z_{out} for the outer region. For small energies, i.e., $|E| < e^2/2b$, and a nuclear charge $Z_0 \geq 2$, one finds that $z_{\text{in}} > R_{\text{in}}$ and $z_{\text{out}} > R_{\text{out}}$. The resulting profile is shown in Fig. 2 for a hypothetical Li atom where $Z_0 = 3$ and $E = -e^2/12b$. The discontinuous derivative dR/dz in this profile at $r = b$ is obviously due to the concentrated screening charge. If the screening charge were smeared out over a shell of nonvanishing thickness, the derivative dR/dz would be continuous. But even with such a discontinuous profile the geodesics are still well defined. The presence of a screening charge is seen to introduce a region of negative curvature into the isometric surface and we have to reckon with such negative curvature in typical atomic potentials.

The variational principle (12) is, therefore, identical to finding the geodesics on a surface in 3-dimensional Euclidean space which is topologically equivalent to a sphere. We can immediately conclude that there are an infinity of classical trajectories connecting any two given momenta p' and p'' at a fixed energy E. These trajectories can be distinguished by the signature (i.e., number of negative eigenvalues) of the second variation, also called the index. The existence of a trajectory of index 0 is intuitively obvious, whereas the existence of a trajectory of higher index can be inferred from Morse's theory.[4] For example, if we consider all families of paths between two given endpoints of the type depicted in Fig. 3, the least upper bound for the lengths of the paths in all these families is realized by a geodesic of index 1. By considering appropriate families of paths depending on 2 and more parameters, we find geodesics of index 2 and more.

The topological arguments of Morse show the existence of at least one trajectory of each index ≥ 0. It is clear, however, that an isometric surface of the type given in Fig. 2 may give more than one trajectory of index 0, namely, when the two end points lie near the "waist-line," as shown in Fig. 4. Also, if we apply the Morse argument as shown in Fig. 3 to end

[4] J. Milnor, *Morse Theory* (Princeton University Press, Princeton, N.J., 1963).

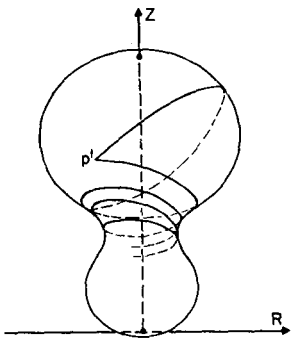


FIG. 4. Geodesics of arbitrary length, but of index 0, i.e., shorter than any neighboring curve, and of index 1, on an isometric surface which is typical for effective atomic potentials.

points near the waistline of the isometric surface, the existence of at least two trajectories of index 1 becomes apparent. The screening charge is, therefore, shown to give more trajectories of given index between given end points than the simple Coulomb potential.

The procedure which gave an isometric surface of rotational symmetry for momentum space can be applied equally well to coordinate space. The first formula (16) remains the same, but the second formula (16) has a different right-hand side. If one takes for $V(r)$ the Coulomb potential, the new expression on the right-hand side is found to become negative for values of r in the neighborhood of $r = r_0 = r(E)$. Our construction fails, therefore, to give an imbedding of the whole Riemannian space $0 < r < r_0$ into the 3-dimensional Euclidean space.

It seems likely that no other construction would succeed, even if we abandon the requirement of a rotationally symmetric imbedding. Indeed, if we had any imbedding for the whole space, it would be intuitively compelling that any two end points can be connected by a trajectory which has the index 0. Since we know the shape of all these trajectories in coordinate space, however, we can check immediately whether such a shortest, physically acceptable trajectory does exist. We have given in Fig. 5 a sketch of all Kepler orbits with the same major axis (same energy) through a given initial point. These orbits fill an ellipse with the origin $r = 0$ and the initial point as foci, and which touches the limiting circle $r = r_0$. The points outside this ellipse cannot be reached by any Kepler orbit. We expect, therefore, that the approximate Green's function in coordinate space is more complicated than in momentum space for energies $E < 0$.

3. THE TRAJECTORIES BETWEEN TWO GIVEN MOMENTA

The representation of the classical trajectories as geodesics on a surface in 3-dimensional Euclidean space gives a vivid picture of the situation and ties our problem to some important results of modern mathematics. In order to perform explicit calculations it seems more convenient, however, to revert back to a more conventional presentation. In any case we think that it can only be of help to our understanding if the same situation can be viewed in more than one way.

The new representation involves a Cartesian space whose three mutually orthogonal axes will be labeled by an angular momentum M, an angle η which varies from $-\infty$ to $+\infty$, and a radius r which varies from 0 to $r_0 = r(E)$. The angular momentum M is the one associated with a particular trajectory. The angle η is the polar angle of the momentum p (or the coordinate q) for the particular trajectory in momentum space (or coordinate space), and we shall think of it as varying monotonically (either increasing or decreasing) if we keep going along the same trajectory. The radius r gives the distance $|q|$ from the nucleus along the trajectory or, equivalently, r gives the total momentum $\rho = |p|$ with the help of Eq. (14).

The equations of motion result from the Hamiltonian (5) if we remember that $r = (\sigma^2 + M^2/\rho^2)^{\frac{1}{2}}$. In addition to $dM/dt = 0$, we get the following equations:

$$\frac{d\eta}{dt} = H_M = \frac{M}{r\rho^2}\frac{dV}{dr} = \frac{M}{2mr(E-V)}\frac{dV}{dr},$$

$$m\frac{dr}{dt} = \frac{1}{r}\left(\sigma\frac{d\sigma}{dt} - \frac{M^2}{\rho^3}\frac{d\rho}{dt}\right)$$

$$= \frac{\sigma\rho}{mr} = \pm[2m(E-V) - M^2/r^2]^{\frac{1}{2}}, \quad (21)$$

where we have used (14) to eliminate ρ. The double

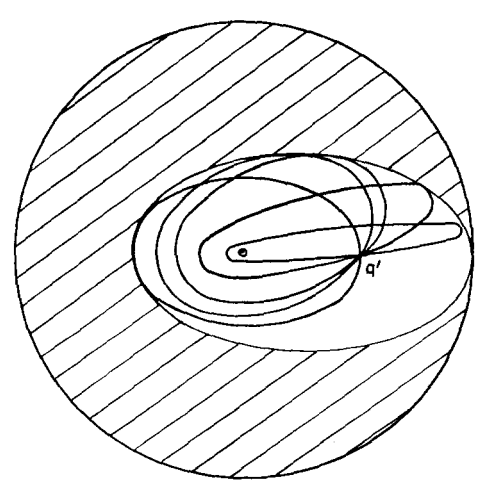


FIG. 5. The Kepler orbits in coordinate space through a given initial position q' and with energy E. Only positions q'' inside the elliptical region can be connected with q' by a trajectory.

sign in the last equation is fixed by the sign of σ. Since σ is the projection of the vector q onto the direction of p, the double sign is determined by the sign of the scalar product pq. Every time r reaches a maximum or a minimum as t varies monotonically the sign of pq changes. After having chosen a particular sign for pq at $t = 0$, there is no ambiguity left.

In order to investigate the set of trajectories which go from p' to p'', we imagine that all those solutions of (21) which go through $r = r'$ and $\eta = 0$ at $t = 0$ are appropriately plotted in the (M, η, r) space. Because of the ambiguity in the sign of pq at $t = 0$, each plane $M = \text{const}$ carries two trajectories. The trajectories in two planes $M = \text{const}$, whose angular momenta are equal but of opposite sign, are identical except that the direction of motion with increasing time is opposite. We can, therefore, restrict our attention to the trajectories with $M \geq 0$, i.e., the trajectories with counterclockwise motion.

The time can be eliminated between Eqs. (21) and the angle η can be expressed as an integral over r:

$$\eta = \int_{r'}^{r} dr \frac{(\pm)\, dV/dr}{2(E - V)} \frac{M}{[2mr^2(E - V) - M^2]^{\frac{1}{2}}}. \quad (22)$$

The double sign takes care of the changing sign of pq as r goes through the extremal values compatible with M and E. These are given by the condition $2mr^2(E - V) = M^2$, and r varies between two solutions of this equation in a range for which

$$2mr^2(E - V) > M^2$$

and which contains the initial r'.

Since $r[2m(E - V)]^{\frac{1}{2}} = R(r)$ has, in general, a local minimum $\bar{M}$ at $\bar{r}$ in the region of electrostatic screening, there are various cases to be distinguished accordingly as $R(r') \gtrless \bar{M}$, where we will assume that we have only one such local minimum. Since $R(r)$ vanishes at $r = 0$ and $r = r_0$, the equation $R(r) = \bar{M}$ has two more solutions, $\bar{r}_1$ and $\bar{r}_2$, besides $\bar{r}$, where $\bar{r}_1 < \bar{r} < \bar{r}_2$. If $|M| < \bar{M}$, the variable r in (22) can vary between two limits r_1 and r_2 with $r_1 < \bar{r}_1 < \bar{r}_2 < r_2$. If $|M| > \bar{M}$, we have either $\bar{r}_1 < r' < \bar{r}$ or $\bar{r} < r' < \bar{r}_2$. In the first case r varies between two limits r_1 and r_2 such that $\bar{r}_1 < r_1 < r' < r_2 < \bar{r}$, whereas in the second case the limits r_1 and r_2 are restricted by $\bar{r} < r_1 < r' < r_2 < \bar{r}_2$. Thus, the trajectory goes through the whole electrostatic field if $|M| < \bar{M}$. But if $|M| > \bar{M}$, the trajectory is either restricted to the inside of the screening charge (if $\bar{r}_1 < r' < \bar{r}$) or to the outside (if $\bar{r} < r' < \bar{r}_2$).

As r goes from r_1 to r_2 and back to r_1, the angle η increases by an angle 2γ, which is given by the integral

$$\gamma = \int_{r_1}^{r_2} dr \frac{dV/dr}{2(E - V)} \frac{M}{[2mr^2(E - V) - M^2]^{\frac{1}{2}}}. \quad (23)$$

The electrostatic screening has the effect of making γ larger than π, the value which is obtained for the pure Coulomb potential. This so-called precession of the orbits depends very much on the type of trajectory, whether it goes through the whole electrostatic field or only through the region inside or outside the screening charge. In any one of these cases there is a strong dependence on the angular momentum M. Two limiting cases are of particular interest and will be discussed in detail, the case of very small M and the case of $|M|$ close to $\bar{M}$.

For $|M| \ll \bar{M}$, Eqs. (21) tell us that $d\eta/dt$ is small, whereas dr/dt remains bounded away from zero as long as r does not approach $r = 0$ or $r = r_0$. The corresponding geodesic on the isometric surface of the previous section gets closer to the meridian as M is smaller. It is intuitively clear that as M goes to zero, the geodesic approaches the "south pole" for small r and the "north pole" for r close to r_0. The angle γ goes to π as M vanishes so that there is no precession in that limit. This result will be derived directly from Eqs. (21) in Appendix A. The necessity of the assumption (13) will become evident once more.

If $|M|$ is very close to $\bar{M}$, the derivative $d\eta/dr$ given by (22) becomes very large as r gets close to $\bar{r}$. The corresponding geodesic keeps winding around the "waist" of the isometric surface while gaining little height z, as shown in Fig. 5. This occurs in both cases $|M| \lesssim \bar{M}$. Obviously, the gain 2γ during one period can become arbitrarily large. The precession as a function of M in the neighborhood of $\bar{M}$ will be discussed in Appendix B. If we expand

$$R(r) = r[2m(E - V)]^{\frac{1}{2}} = \bar{M} + (\mu/2)(r - \bar{r})^2 + \cdots,$$

the leading contribution to γ for each passage near $\bar{r}$ becomes

$$2\gamma \cong \left[\frac{dV/dr}{2(E - V)}\right]_{r=\bar{r}} \left(\frac{M}{\mu}\right)^{\frac{1}{2}} \log \left|\frac{\mu \bar{r}^2}{2(M - \bar{M})}\right|. \quad (24)$$

If $|M| < \bar{M}$, the trajectory passes the radius $\bar{r}$ twice in each period so that it will have twice the contribution (24). If $|M| > \bar{M}$, the trajectory passes near the radius $\bar{r}$ only once in each period so that it will have just the contribution (24). If μ becomes very large as in a potential with a strongly localized screening charge, the contribution (24) to γ becomes small again if we keep the difference $|M - \bar{M}|$ fixed. A potential like (20) shows, therefore, no infinitely large precession, as one expects from an isometric surface of the type shown in Fig. 2. The occurrence of an infinitely large precession is easily understood with the help of the isometric surface, although an estimate like (24) still

requires integrating the equations of motion. The trajectories in coordinate space do not allow this kind of geometric insight.

The approximate Green's function $\tilde{F}(p''p'E)$ depends on the absolute values $\rho' = |p'|$ and $\rho'' = |p''|$, or r' and r'' according to (14), as well as on the spherical distance ζ between p' and p''. The angle ζ is defined to lie between 0 and π, whereas the angle $\eta = \eta'' - \eta'$ can be anywhere from $-\infty$ to $+\infty$. A solution of the Eqs. (21) is an acceptable trajectory from p' to p'' if $\eta = \zeta \bmod 2\pi$. For $\eta > 0$ we may call the trajectory direct, and for $\eta < 0$ indirect, because the two main examples of each case are $\eta = \zeta$ and $\eta = -(2\pi - \zeta)$. The direct trajectories have $M > 0$, whereas the indirect ones have $M < 0$. Finally, it should be remembered that the approximate Green's function $\tilde{F}$ carries the particle forward in time as it goes from p' and p'', and that the virial $T(p''p'E)$ in (2) is, therefore, positive for the direct as well as for the indirect trajectory. This is borne out by the formula (7) together with the above description of the relevant trajectories.

The solutions of the equation $\eta = \zeta \bmod 2\pi$ can be obtained as follows. We consider again the trajectories through $r = r'$ and $\eta = 0$ in the (M, η, r) space. Their intersections with the plane $r = r''$ form a set of curves which is shown in Figs. 6 and 7 for two typical situations. The intersections of these η-vs-M plots with the set of horizontal lines $\eta = \zeta \bmod 2\pi$ give the values of M for a trajectory from p' to p''.

Since Figs. 6 and 7 are fairly complicated, it may be useful to discuss their origin. In order to accomplish this, the trajectories through $r = r'$ and $\eta = 0$ can be projected onto the plane $M = 0$. Figure 8 shows the situation schematically when $R(r') < \bar{M}$. We have assumed that the precession γ increases with increasing $|M|$ which is, indeed, natural as long as $|M|$ remains below $\bar{M}$. Figure 6 can be obtained from Fig. 8 if we follow along a line $r = r''$ and note the angular momentum M of the trajectory which goes through a particular value of the angle η. M is a monotonically increasing function of the direction cosine for the particular trajectory at $\eta = 0$. The range of M in Fig. 6 is determined by the smaller among $R(r')$ and $R(r'')$. If we had no precession, the triangular regions in Fig. 8, which are bounded by the lines $\eta =$ multiple of π and two envelopes, would reduce to points. Also, the branches of the η-vs-M plot in Fig. 6 would stay horizontal, instead of turning away from the M axis.

Figures 9 and 10 should be superimposed on each other, but they are separated to keep the confusion of the trajectories from becoming total. In both Figs. 9 and 10 we have $R(r') > \bar{M}$, but in Fig. 9 only the

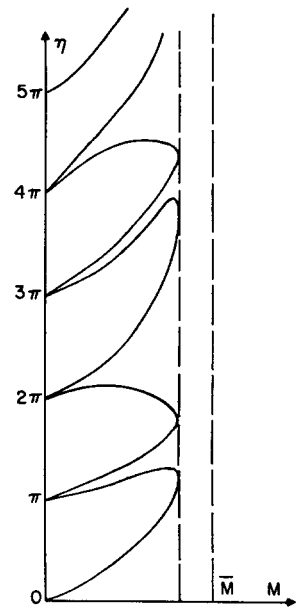


FIG. 6. The angle η between initial and final momentum, p' and p'', is plotted as a function of the angular momentum M for the trajectory from p' to p''. The energy E is given and the absolute values of p' and p'' are fixed so that M never reaches the critical value $\bar{M}$. For $M < 0$, these curves have to be inverted through the origin.

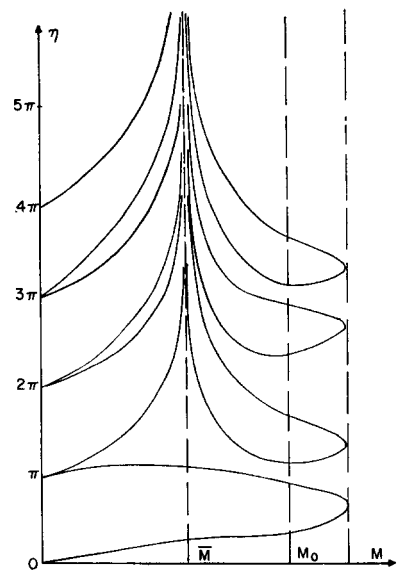


FIG. 7. Same as Fig. 6, but the absolute values of p' and p'' are now fixed so that M can exceed the critical value $\bar{M}$. Also, there is an interval for large M with no precession at all.

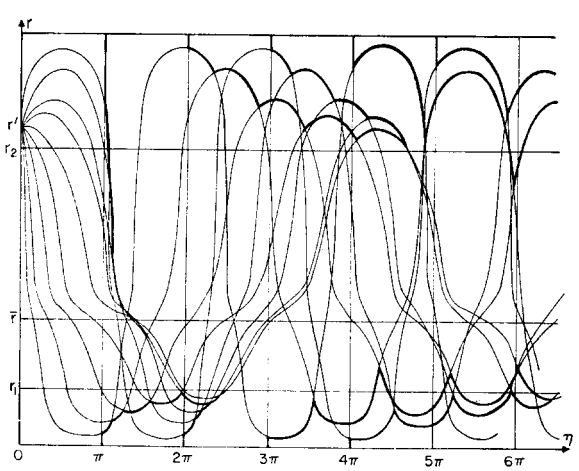


FIG. 8. Schematic plot of the trajectories through a given initial momentum $\{2m[E - V(r')]\}^{\frac{1}{2}}$, assumed to be so small that none of the trajectories can have an angular momentum as large as $\bar{M}$.

trajectories with $|M| < \bar{M}$, and in Fig. 10 only trajectories with $|M| > \bar{M}$, are plotted. The infinite precession for $|M| = \bar{M}$ produces the envelopes which touch the horizontal lines $r = \bar{r}$ and $r = \bar{r}_2$ if $\bar{r} < r' < \bar{r}_2$ as in Figs. 9 and 10. Again we can follow a line $r = r''$ and notice the angular momentum M associated with a particular angle η, in order to construct the η-vs-M plot. If $R(r'') < \bar{M}$, nothing new is obtained as compared with Fig. 6. If $\bar{r}_1 < r'' < \bar{r}$ whereas $\bar{r} < r' < \bar{r}_2$, the range of $|M|$ in the η-vs-M plot is still limited to $\bar{M}$ because the trajectories with $|M| > \bar{M}$ are either limited to $\bar{r}_1 < r < \bar{r}$ or to $\bar{r} < r < \bar{r}_2$. Therefore, only the case $\bar{r} < r'' < \bar{r}_2$ produces a η-vs-M plot like Fig. 7 which is quite different from Fig. 6. The infinite peaks at $M = \bar{M}$ are actually less dramatic than might appear from Fig. 7 since they have only the logarithmic infinity which is indicated by formula (24). We have assumed that the precession γ decreases with increasing $|M| > \bar{M}$, which is reasonable since the trajectory stays outside most of the screening charge. In Fig. 10 we even assumed no precession at all for $|M|$ larger than some critical value M_0 such that if $R(r) > M_0$, the potential $V(r)$ is again purely Coulombic. The various branches in Fig. 7 all come down again into the range of η from which they originated at $M = 0$.

If we were interested in the approximate Green's function $\tilde{G}(q''q'E)$ in coordinate space, we could construct similar pictures, with some significant differences, however. Most notable would be that the trajectories in Figs. 8, 9, and 10 would not cover the whole strip between $r = 0$ and $r = r_0$, and that the branches in Figs. 6 and 7 would not form one continuous curve, but would be disconnected.

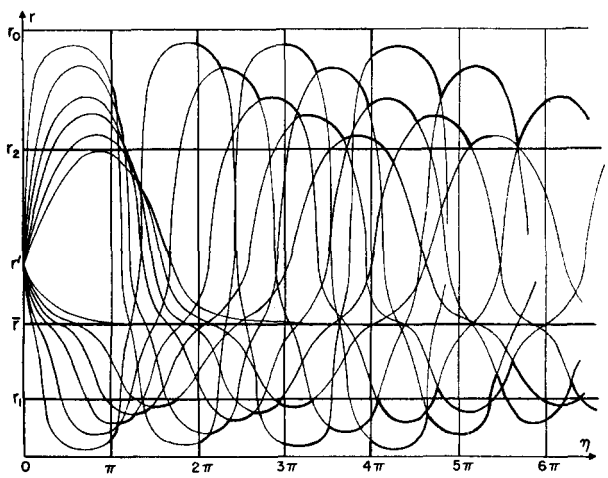


FIG. 9. Same as Fig. 8, but now $r'\{2m[E - V(r')]\}^{\frac{1}{2}} > \bar{M}$, and r' lies between $\bar{r}$ and r_2. Only the trajectories with $M < \bar{M}$ are plotted to avoid confusion.

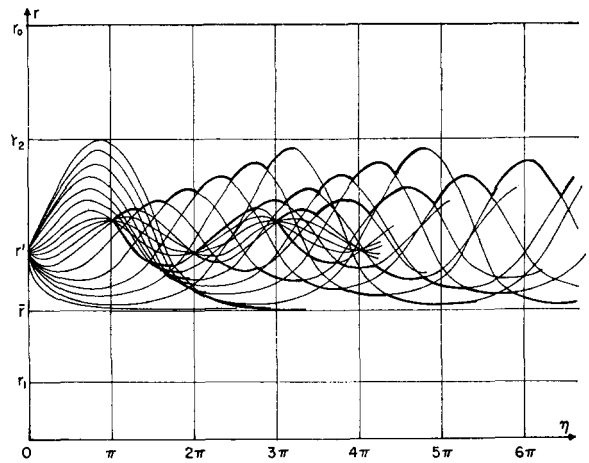


FIG. 10. Same as Fig. 9, but only the trajectories with $M > \bar{M}$ are plotted. There is a family of trajectories without orbital precession, i.e., with periodicity 2π.

4. THE SUMMATION OVER ALL THE POSSIBLE TRAJECTORIES

On the basis of the preceding discussion we can now make the complete enumeration of all the trajectories from p' to p'' which is required in formula (2) for $\tilde{F}(p''p'E)$. To be definite we shall assume that $\rho'' < \rho'$ or, equivalently, $r' < r''$. In the opposite case we can use the symmetry relation

$$\tilde{F}(p'p''E) = \tilde{F}(p''p'E), \tag{25}$$

which is a direct consequence of the formula (2). The amplitude factor (9) can be written as

$$\frac{1}{D_T} = \frac{\rho'\rho'' \sin\eta}{Mr'r''} \left\{\frac{dV}{dr}\,[2mr^2(E - V) - M^2]^{\frac{1}{2}}\right\}' \times \left\{\frac{dV}{dr}\,[2mr^2(E - V) - M^2]^{\frac{1}{2}}\right\}'' \left(\frac{\partial\eta}{\partial M}\right)_{\rho'\rho''} \tag{26}$$

with the help of (5) and (14). The derivative $(\partial\eta/\partial M)_{\rho'\rho''}$ can be obtained directly from Fig. 6 or Fig. 7. It will be shown in Appendix C that this expression remains finite throughout the trajectory, except near $\bar{M}$ where $(\partial\eta/\partial M)_{\rho'\rho''} \approx (M - \bar{M})^{-1}$ as can be recognized from (24).

In addition to γ as given by (23), we shall use the angles

$$\alpha = \int_{r_1}^{r'} dr\, \frac{dV/dr}{2(E - V)}\, \frac{M}{[2mr^2(E - V) - M^2]^{\frac{1}{2}}},$$

$$\beta = \int_{r''}^{r_2} dr\, \frac{dV/dr}{2(E - V)}\, \frac{M}{[2mr^2(E - V) - M^2]^{\frac{1}{2}}}, \tag{27}$$

which can be considered as a function of r' and M (respectively, r'' and M), always at a fixed energy E.

Similarly, we have the virials corresponding to α, β, and γ given by

$$\{\sigma,\tau,\theta\} = \left\{\int_{r_1}^{r'}, \int_{r''}^{r_2}, \int_{r_1}^{r_2}\right\} dr \frac{mr^2\, dV/dr}{[2mr^2(E-V)-M^2]^{\frac{1}{2}}}, \tag{28}$$

again to be considered as functions of r' and M (r'' and M, respectively) and in the case of θ as function of M only. In case of ambiguity all radii $r_1 < r' < r'' < r_2$ are assumed to be larger than the screening radius $\bar{r}$.

For any particular trajectory from p' to p'' we can immediately determine two integers λ and ν as follows. We write $\eta = 2\lambda\pi + \zeta$, where $0 \leq \zeta < \pi$, to obtain λ and $\eta = 2\nu\gamma + \delta$, where $0 \leq \delta < \gamma$, to obtain ν. There are four different types of trajectories accordingly as δ is made up of the angles α, β, and γ. In terms of the spherical distance ζ between p' and p'', we find that

$$\begin{aligned}
&\text{(i)} && (2\nu+1)\gamma - \alpha - \beta = 2\lambda\pi + \zeta,\\
&\text{(ii)} && (2\nu+1)\gamma - \alpha + \beta = 2\lambda\pi + \zeta,\\
&\text{(iii)} && (2\nu+1)\gamma + \alpha - \beta = 2\lambda\pi + \zeta,\\
&\text{(iv)} && (2\nu+1)\gamma + \alpha + \beta = 2\lambda\pi + \zeta.
\end{aligned} \tag{29}$$

Since α, β, γ, and λ have the same sign as M, positive for direct and negative for indirect trajectories, whereas $0 \leq \zeta \leq \pi$, the integer ν is always ≥ 0. The virial along the four types of trajectories is given by

$$T(p''p'E) = (2\nu+1)\theta \pm \sigma \pm \tau, \tag{30}$$

with the same signs in front of σ and τ as in front of α and β in (29). The four types can easily be recognized on sketches like Figs. 8, 9, and 10.

The four types of trajectories (29) are associated with the various branches of Figs. 6 and 7 in the following manner. Because of the absence of orbital precession for $M = 0$, the behavior of the η versus M is particularly simple at $M = 0$ and can serve to characterize each particular branch. Two such branches come together at each multiple of π for $M = 0$ corresponding to type (i) and (iv) at even multiples of π, and to type (ii) and (iii) for odd multiples. This fact results from Appendix A where the angles α, β, and γ are shown to go to the values $\pm\pi/2$, $\pm\pi/2$, and $\pm\pi$ in the limit $M \to 0$, with the upper sign for $M > 0$ and the lower sign for $M < 0$. Furthermore, it is evident from Figs. 8, 9, and 10 that the type (iii) meets the type (ii) from above (below) for $M > 0$ ($M < 0$), as indicated in Figs. 6 and 7, although it does not always have to be that way. Similarly, the type (i) meets the type (iv) from above (below). As we follow any branch (type) from $M = 0$ out to the largest value of M which is compatible with r' and r'', type (i) meets with the nearest type (ii) and type (iii) with the nearest type (iv). Therefore, if one starts with the lowest (highest) branch for $M > 0$ ($M < 0$) which is of type (i), the next higher (lower) branches in Figs. 6 and 7 come in the order (ii), (iii), (iv), (i), (ii), etc. Even across the peaks at $M = \bar{M}$ in Fig. 7, this identification of the branches is easily accomplished, and we shall use it from now on whenever it seems convenient.

The identification of the branches in Figs. 6 and 7 is important because it helps in finding the number of conjugate times between p' and p'' for each trajectory. The contribution from the angular motion, i.e., the number of times the angle η goes through a multiple of π, is given by 2λ for the direct trajectories and $2|\lambda| + 1$ for the indirect trajectories. Again consulting Figs. 8, 9, and 10, the rules for the contribution of the radial motion become evident. A conjugate time for the radial motion along any trajectory is located wherever the trajectory touches one of the various envelopes. For small M that number equals 2ν for type (i), $2\nu + 1$ for types (ii) and (iii), and $2\nu + 2$ for type (iv), provided $(\partial\eta/\partial M)_{\rho'\rho''} > 0$ at $M = 0$. The last condition may not hold, particularly for the types (ii) and (iv) if there is little precession or if ν is small. In that case, the numbers $2\nu + 1$ and $2\nu + 2$ have to be reduced by 1. This initial number of conjugate times due to the radial motion changes only when going from $M = 0$ to the maximum value of M along any branch in Figs. 6 and 7, if the slope $(\partial\eta/\partial M)_{\rho'\rho''}$ changes sign.

Let us assume that a particular branch, whatever its type, starts with $(\partial\eta/\partial M)_{\rho'\rho''} > 0$ at $M = 0$. If we move away from the η axis by increasing M for $M > 0$ or by decreasing M for $M < 0$, one conjugate time is lost when the slope $(\partial\eta/\partial M)_{\rho'\rho''}$ becomes negative. This conjugate time is regained as soon as $(\partial\eta/\partial M)_{\rho'\rho''}$ becomes positive upon moving further away from the η axis. It will be lost again if $(\partial\eta/\partial M)_{\rho'\rho''}$ becomes negative once more, and so on, until M reaches the extremal value which is compatible with r' and r''. There the slopes in Figs. 6 and 7 are positive for the branches of types (i) and (iii), negative for the branches of types (ii) and (iv), so that the number of conjugate times does not change as we go from a branch of type (i) to one of type (ii), or from types (iii) to (iv) at the extreme value of M. Finally, if the slope $(\partial\eta/\partial M)_{\rho'\rho''} < 0$ at $M = 0$, we may think of it as having already changed from positive to negative and, therefore, having caused the loss of a conjugate time. In this manner, the index of a particular trajectory, i.e., the total number of conjugate times between

p' and p'', can be obtained if we know the type of trajectory, the sign of the slope $(\partial\eta/\partial M)_{p'p''}$, as well as the integers λ and ν.

In order to evaluate the formula (2) for $\tilde{F}(p''p'E)$ we have first to find the solutions of the Eqs. (29); i.e., given r', r'', ζ, and E, we have to find all the triples (λ, ν, M) which satisfy one of the four equations (29). For each triple we can evaluate $T(p''p'E)$ according to (30), the phases according to the above rules, and $(\partial\eta/\partial M)_{p'p''} = (2\nu + 1)\dot{\gamma} \pm \dot{\alpha} \pm \dot{\beta}$, where the dot indicates the partial derivative with respect to M, to be inserted into (26). In this manner we obtain the formula

$$\begin{aligned}\tilde{F}(p''p'E) &= -\frac{1}{2\pi\hbar^2}\sum_{\nu=0}^{\infty}\sum_{M\gtrless 0}\sum_{\text{(i)}}^{\text{(iv)}}\sum_{\lambda} \\ &\times \left(\frac{\rho'\rho''\sin\zeta}{r'r''}\left\{\frac{dV}{dr}[2mr^2(E-V)-M^2]^{\frac{1}{2}}\right\}'\right. \\ &\times \left.\left\{\frac{dV}{dr}[2mr^2(E-V)-M^2]^{\frac{1}{2}}\right\}''\right. \\ &\times \left.\left|\frac{(2\nu+1)\dot{\gamma}\pm\dot{\alpha}\pm\dot{\beta}}{M}\right|\right)^{-\frac{1}{2}} \\ &\times \exp i\left\{\frac{(2\nu+1)\theta\mp\sigma\mp\tau}{\hbar} - (\nu+\tfrac{1}{2}\mp\tfrac{1}{4}\mp\tfrac{1}{4})\pi\right. \\ &- (\lambda+\tfrac{1}{4})\pi\,\operatorname{sgn} M \\ &\left.- \frac{\pi}{4}\operatorname{sgn}[(2\nu+1)\dot{\gamma}\mp\dot{\alpha}\mp\dot{\beta}] + \frac{\pi}{2}\right\}. \qquad (31)\end{aligned}$$

5. THE ORIGIN OF THE QUANTUM NUMBERS

The last expression (31) for the approximate Green's function in momentum space has to be investigated in order to find its singularities. This can be accomplished most directly by examining Eqs. (29) in the limit of very large ν so as to find the conditions for E and M which make the terms in (31) all add up in phase. Since we would like to find also the behavior of $\tilde{F}(p''p'E)$ "between" the singularities, a more circuitous, although formally more elegant, route seems preferable.

The first step is to rewrite (31) so as to get rid of the annoying limitation $\nu \geq 0$. If we keep the expression (31) as it stands, but add to it all the terms that arise from allowing $\nu < 0$, we find the following: Each term with $\nu < 0$ can be interpreted as (-1) times the complex conjugate of a term with $\nu \geq 0$. In this correspondence a term with $\nu < 0$ which belongs to the angular momentum M and the type (i), (ii), (iii), or (iv), is associated with the term $\nu^+ = -2(\nu + 1) \geq 0$ of angular momentum $M^+ = -M$ and of type (iv), (iii), (ii), or (i). Indeed, Eqs. (29) can be written as $(2\nu + 1)\gamma \mp \alpha \mp \beta = 2\lambda\pi + \zeta$ in abbreviated form, and become

$$(2\nu^+ + 1)\gamma^+ \pm \alpha^+ \pm \beta^+ = 2\lambda\pi + \zeta$$

after replacing $(2\nu + 1)$ with $-(2\nu^+ + 1)$ and M with $-M^+$, etc. Similarly, the phase of (31), i.e., the inside of the braces in $\exp i\{\ \}$, can be rewritten in terms of ν^+, M^+, etc. as

$$\begin{aligned}&-\frac{(2\nu^+ + 1)\theta^+ \pm \sigma^+ \pm \tau^+}{\hbar} + (\nu^+ + \tfrac{1}{2} \pm \tfrac{1}{4} \pm \tfrac{1}{4})\pi \\ &\quad + \pi(\lambda + \tfrac{1}{4})\operatorname{sgn} M^+ \\ &\quad + \frac{\pi}{4}\operatorname{sgn}[(2\nu^+ + 1)\dot{\gamma}^+ \pm \dot{\alpha}^+ + \dot{\beta}^+] - \frac{\pi}{2} + \pi.\end{aligned}$$

Therefore, if we formally let $(2\nu + 1) < 0$ in (31), we get terms with $(2\nu^+ + 1) > 0$ whose phases have the wrong sign plus an extra π. Formula (31) can then be used in the form

$$2\,\operatorname{Im}\tilde{F}(p''p'E) = \frac{1}{2\pi\hbar^2}\sum_{\gamma=-\infty}^{\infty}\sum_{M\gtrless 0}\sum_{\text{(i)}}^{\text{(iv)}}\sum_{\lambda}\cdots \qquad (32)$$

to give the discontinuity of $\tilde{F}(p''p'E)$ across the E axis.

The next step replaces the summation over λ in (32) by a summation over an integer l that will turn out to be the angular quantum number. To this end we shall first change the meaning of λ and make it a continuous variable which is a function of M for any fixed ν and given type (i), (ii), (iii), or (iv) in accordance with Eqs. (29). Then we shall replace any integration over λ by an integration over M extending through the range which is compatible with r', r'', and E. Thus we write the following sequence of equations in somewhat symbolic form:

$$\begin{aligned}\sum_{\lambda} &= \int d\lambda \sum_{\Lambda=-\infty}^{+\infty}\delta(\lambda-\Lambda) \\ &= \frac{1}{2\pi}\int dM\,|(2\nu+1)\dot{\gamma}\mp\dot{\alpha}\mp\dot{\beta}| \\ &\quad\times\sum_{l=-\infty}^{+\infty}\exp\{-il[(2\nu+1)\gamma\mp\alpha\mp\beta-\zeta]\operatorname{sgn}M\},\end{aligned} \qquad (33)$$

with $2\pi\lambda$ to be replaced by $(2\nu + 1)\gamma \mp \alpha \mp \beta - \zeta$ wherever it occurs in (31) and (33).

The sgn M has been inserted into (33) for convenience. The new exponential in (32) becomes, after

inserting (33),

$$\exp i\Big\{\frac{(2\nu+1)\theta \mp \sigma \mp \tau}{\hbar} - (\nu + \tfrac{1}{2} \mp \tfrac{1}{4} \mp \tfrac{1}{4})\pi - (l+\tfrac{1}{2})[(2\nu+1)\gamma \mp \alpha \mp \beta]\operatorname{sgn} M + \Big[(l+\tfrac{1}{2})\zeta - \frac{\pi}{4}\Big]\operatorname{sgn} M - \frac{\pi}{4}\operatorname{sgn}[(2\nu+1)\dot\gamma \mp \dot\alpha \mp \dot\beta] + \frac{\pi}{2}\Big\}.$$

The quantities θ, σ, τ, γ sgn M, α sgn M, β sgn M, $\dot\gamma$, $\dot\alpha$, $\dot\gamma$ are even functions of M, so that only the term $[(l+\frac{1}{2})\zeta - \pi/4]$ sgn M changes its sign with M. The summation over M in (32) can be carried out and gives the ζ dependence

$$2\frac{\cos[(l+\frac{1}{2})\zeta - \pi/4]}{(\sin\zeta)^{\frac{1}{2}}}, \tag{34}$$

if we include the ζ dependence of the amplitude, of (31).

For $l \geq 0$, this function of ζ is actually the first term in the asymptotic expansion of the Legendre polynomial $P_l(\cos\theta)$ for large l. According to Whittaker and Watson,[5] we have for the normalized Legendre polynomial

$$\Big(\frac{2l+1}{4\pi}\Big)^{\frac{1}{2}} P_l(\cos\zeta) \sim \frac{\Gamma(l+1)}{\Gamma(l+\frac{1}{2})\cdot(l+\frac{1}{2})^{\frac{1}{2}}}\frac{\cos[(l+\frac{1}{2})\zeta - \pi/4]}{\pi(\sin\zeta)^{\frac{1}{2}}}. \tag{35}$$

The first factor goes to 1 in the limit $l = \infty$, so that (34) corresponds to 2π times the normalized Legendre polynomial for large $l > 0$. The functions (34) for $l < 0$, unfortunately, have no such interpretation. Since the functions (34) with $l \geq 0$ are presumably a complete, although not an orthogonal, system in the interval $0 < \zeta < \pi$, the functions (34) with $l < 0$ are not independent.

The summation over ν can be carried out by writing

$$\begin{aligned}\operatorname{Im}\tilde F(p''p'E) &= \frac{-i}{(2\pi\hbar)^2}\sum_l \frac{\cos[(l+\frac{1}{2})\zeta - \pi/4]}{(\sin\zeta)^{\frac{1}{2}}}\int_{M>0} dM \sum_{\text{(i)}}^{\text{(iv)}}\sum_\nu \\ &\times\Big\{\frac{\rho'\rho''}{Mr'r''}\Big[\frac{dV}{dr}(2mr^2(E-V)-M^2)^{\frac{1}{2}}\Big]'\Big[\frac{dV}{dr}(2mr^2(E-V)-M^2)^{\frac{1}{2}}\Big]''\Big\}^{-\frac{1}{2}} \\ &\times|(2\nu+1)\dot\gamma \mp \dot\alpha \mp \dot\beta|^{\frac{1}{2}} \\ &\times\exp i\Big\{(2\nu+1)\Big[\frac{\theta}{\hbar} - \frac{\pi}{2} - (l+\tfrac{1}{2})\gamma\Big] - \frac{\pi}{4}\operatorname{sgn}[(2\nu+1)\dot\gamma \mp \dot\alpha \mp \dot\beta] \\ &\mp\Big[\frac{\sigma}{\hbar} - \frac{\pi}{4} - (l+\tfrac{1}{2})\alpha\Big] \mp \Big[\frac{\tau}{\hbar} - \frac{\pi}{4} - (l+\tfrac{1}{2})\beta\Big]\Big\}.\end{aligned} \tag{36}$$

The sum over ν is evaluated with the help of Poisson's formula[6]

$$\sum_\nu \Phi(\nu) = \sum_n \int dx\,\Phi(x)e^{-2n\pi i x}, \tag{37}$$

where both summations go from $-\infty$ to $+\infty$ and so does the integration over x. The integral is treated in the context of the theory of generalized functions.

With the abbreviation

$$\Delta = (2l+1)\gamma + (2n+1)\pi - 2\theta/\hbar, \tag{38}$$

the integral to be evaluated is given by

$$\begin{aligned}&\int_{-\infty}^{+\infty} dx\,|(2x+1)\dot\gamma + \dot\delta|^{\frac{1}{2}}\exp i\Big\{[(2x+1)\dot\gamma + \dot\delta]\frac{\Delta}{2\dot\gamma} - \frac{\pi}{4}\operatorname{sgn}[(2x+1)\dot\gamma + \dot\delta] + \frac{\dot\delta\Delta}{2\dot\gamma} + n\pi\Big\} \\ &= (2\pi|\dot\gamma|)^{\frac{1}{2}}|\Delta|^{-\frac{3}{2}}\exp i\Big(n\pi + \frac{\dot\delta\Delta}{2\dot\gamma}\Big) \times\begin{cases}-1, & \text{for } \dot\gamma\Delta > 0,\\ 0, & \text{for } \dot\gamma\Delta < 0,\end{cases}\end{aligned} \tag{39}$$

where δ stands for $\mp\alpha \mp\beta$. The sum from (i) to (iv) is trivial at this point and we get the final expression

$$\begin{aligned}\operatorname{Im}\tilde F(p''p'E) &= \frac{i}{\pi^2\hbar^2}\sum_l \frac{\cos[(l+\frac{1}{2})\zeta - \pi/4]}{(\sin\zeta)^{\frac{1}{2}}} \\ &\times\sum_n(-1)^n\int dM\,\frac{\sqrt{2\pi|\dot\gamma|M}}{|\Delta|^{\frac{3}{2}}} \\ &\times\Big\{\Big[\rho\frac{dV}{dr}\Big(2m(E-V)-\frac{M^2}{r^2}\Big)^{\frac{1}{2}}\Big]'\Big\}^{-\frac{1}{2}} \\ &\times\cos\Big[\frac{\sigma}{\hbar} - (l+\tfrac{1}{2})\alpha + \frac{\dot\alpha\Delta}{2\dot\gamma} - \frac{\pi}{4}\Big] \\ &\times\Big\{\Big[\rho\frac{dV}{dr}\Big(2m(E-V)-\frac{M^2}{r^2}\Big)^{\frac{1}{2}}\Big]''\Big\}^{-\frac{1}{2}} \\ &\times\cos\Big[\frac{\tau}{\hbar} - (l+\tfrac{1}{2})\beta + \frac{\dot\beta\Delta}{2\dot\gamma} - \frac{\pi}{4}\Big], \qquad \dot\gamma\Delta > 0.\end{aligned} \tag{40}$$

[5] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, England, 1927), 4th ed., p. 316.

[6] M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, England, 1958), p. 69.

The second line depends only on r', and the third line only on r'', as can be checked with the help of (27) and (28). The summation over n is limited to the region where $\dot{\gamma}\Delta > 0$.

If we are interested in $\tilde{F}(p''p'E)$ itself, not only its imaginary part, all the steps in the preceding derivation remain the same except that the integral (39) goes now only from $2x + 1 = 0$ to ∞, exactly as the sum over ν in (37) goes only from 0 to ∞. The integer n still takes on all values from $-\infty$ to $+\infty$. The expression (40) remains the same with the exception of the factor $(2\pi\,|\dot{\gamma}|\,M)^{\frac{1}{2}} \cdot \Delta^{-\frac{3}{2}}$, which is replaced by a much more complicated function of $\dot{\gamma}$, $\dot{\delta}$, and Δ.

Since we are interested in the discontinuity of $\tilde{F}(p''p'E)$ across the real E axis, it is convenient to rewrite (4) so as to give it the appearance of being just such a discontinuity. Accordingly, we should try to introduce complex values of the energy E. It is obvious, however, that this would be difficult to achieve because the potential V is not necessarily an analytic function of r, and most of our reasoning depends on writing expressions like $r(E - p^2/2m)$, etc. Therefore, we take another approach where the integer n appearing in (38) can take on arbitrary complex values although n is restricted to integer values in (40). The quantity Δ becomes complex, and we can write each term in the summation over l and n as the discontinuity across the real axis of an integral with a factor

$$\frac{(2\pi M)^{\frac{1}{2}}}{|\dot{\gamma}|}\left(\frac{\Delta}{\dot{\gamma}}\right)^{-\frac{3}{2}} \quad \text{instead of} \quad \frac{(2\pi\,|\dot{\gamma}|\,M)^{\frac{1}{2}}}{|\Delta|^{\frac{3}{2}}}.$$

The function $(\Delta/\dot{\gamma})^{-\frac{3}{2}}$ is defined in the whole complex $\Delta/\dot{\gamma}$ plane with the exception of the positive real axis. $(\Delta/\dot{\gamma})^{-\frac{3}{2}}$ is defined to be positive just above the positive $\Delta/\dot{\gamma}$ axis and negative below.

6. THE SINGULARITIES OF THE APPROXIMATE GREEN'S FUNCTION

The eigenstates for the bound electron can be obtained from the poles of the Green's function along the negative real E axis. We shall use this approach to determine the approximate eigenfunctions from the singularities of the approximate Green's function. If one examines the expression (40) for the discontinuity of $\tilde{F}(p''p'E)$ across the negative real E axis, he finds Dirac δ functions of the energy [corresponding to the poles of $\tilde{F}(p''p'E)$ in the complex E plane]. But besides these strong singularities there seem to be weaker ones, corresponding to branch-cuts in the complex E plane, which have no counterpart in the exact Green's function, nor are they easily calculated and interpreted. We shall, therefore, determine only the strong singularities in this work.

Before discussing the general case, we shall dispose of the special circumstance where there is no orbital precession in some closed domain of energy E and angular momentum M. This may happen if the potential is purely Coulombic in some domain of the radius r as for example in the case (20). The formula (40) breaks down in that domain of E and M, because $\dot{\gamma} = 0$. If we go back to (39), however, we find immediately that the right-hand side becomes a Dirac δ function of Δ, so that we can concentrate (40) on $\Delta = 0$. Also the terms $\dot{\alpha}\Delta/2\dot{\gamma}$ and $\dot{\beta}\Delta/2\dot{\gamma}$ disappear from the arguments of the cosines in (40). The summation over l and n can be performed subject to $\Delta = 0$ as given by (38). This leads to another Dirac δ function which relates the angle ζ with α, β, γ (assumed to equal π). Thus, the integration over M becomes trivial, and we are left with the same result as in I, restricted to the closed domain of E and M where $\dot{\gamma} = 0$.

It should be noted that $\dot{\gamma} = 0$ implies that the action θ does not depend on M, so that the quantum condition $\Delta = 0$, or $2\theta = 2\pi\hbar$ times an integer > 0, fixes only the energy. This fact follows from the general formula

$$\left(\frac{\partial T}{\partial \eta}\right)_{\rho'\rho''} = M, \tag{41}$$

where the action T along a particular trajectory is written as a function of ρ', ρ'', and η. If this formula is rewritten in terms of T and η which are now considered to be function of ρ', ρ'', and M, one finds that

$$\left(\frac{\partial T}{\partial M}\right)_{\rho'\rho''} = M\left(\frac{\partial \eta}{\partial M}\right)_{\rho'\rho''}. \tag{42}$$

This last relation can also be checked directly with the help of (30) and the integrals (23), (27), and (28). As a special case we find that $M\dot{\gamma} = d\theta/dM$, so that $\dot{\gamma} = 0$ implies $d\theta/dM = 0$.

If we exclude the case where $\dot{\gamma}$ vanishes identically in a closed domain of E and M, the integrand in (40) can still present difficulties in the neighborhood of special values of M. Whenever the integral over M seems to diverge, we shall treat it with the help of the theory of generalized functions. In this manner we can eliminate two critical values of M as sources of singularities in (40). A first critical value $\tilde{M}$ arises where $\dot{\gamma} = 0$ because $\dot{\gamma}$ occurs in the argument of the cosines in (40). If $\Delta \neq 0$ and $\ddot{\gamma} \neq 0$ at $\tilde{M}$, we can use $\mu = (M - \tilde{M})^{-1}$ as variable of integration in the neighborhood of $\tilde{M}$. The resulting integrand diverges

with a $-\frac{5}{2}$ power of μ and gives, therefore, a finite result.

A second critical value M_0 arises where $\Delta = 0$ because of the $-\frac{3}{2}$ power of $|\Delta|$ in the integrand of (40). If we assume that $d\Delta/dM \neq 0$ at M_0, we can conclude also that $\dot{\gamma} \neq 0$ and $M \neq (l + \frac{1}{2})\hbar$ because of the relation

$$\frac{d\Delta}{dM} = 2\dot{\gamma}\left(l + \frac{1}{2} - \frac{M}{\hbar}\right), \tag{43}$$

which again follows from (42). Actually the change of the arguments in the cosines of (40) comes out equally simple:

$$\frac{d}{dM}\left[\frac{\sigma}{\hbar} - (l + \tfrac{1}{2})\alpha + \frac{\dot{\alpha}\Delta}{2\dot{\gamma}}\right] = \left(\frac{\dot{\alpha}}{2\dot{\gamma}}\right)\cdot\Delta,$$
$$\frac{d}{dM}\left[\frac{\tau}{\hbar} - (l + \tfrac{1}{2})\beta + \frac{\dot{\beta}\Delta}{2\dot{\gamma}}\right] = \left(\frac{\dot{\beta}}{2\dot{\gamma}}\right)\cdot\Delta, \tag{44}$$

i.e., the phase of the cosines in (40) is actually stationary at the critical value M_0 where $\Delta = 0$. Again, the integral around M_0 diverges with a fractional power, and leads to no singularity. This remains true even when the two critical values $\tilde{M}$ and M_0 coincide, i.e., if simultaneously $\dot{\gamma} = 0$ and $\Delta = 0$, because the divergence remains fractional.

If $\Delta = 0$ coincides with $M = (l + \frac{1}{2})\hbar$, then the expansion of Δ in powers of $M - (l + \frac{1}{2})\hbar$ starts with a quadratic term. If simultaneously $\dot{\gamma} \neq 0$ for $M = (l + \frac{1}{2})\hbar$, the arguments in the cosines of (40) remain well behaved. The divergence of the integrand goes as the -3 power of $M - (l + \frac{1}{2})\hbar$ and one ends up with a singularity. This singularity is a Dirac δ function of Δ and is hard to recognize directly from (40). In discussing this singularity it seemed, therefore, more convenient to consider the expression (40) as the discontinuity across the real n axis of the complex-valued function which was defined at the end of the preceding section. If the integrals (40) are calculated in this form, one obtains a pole in the complex n plane rather than a δ function along the real n axis.

If Δ is expanded around $M = (l + \frac{1}{2})\hbar$, one finds that

$$\Delta = (2n + 1)\pi - \frac{2\omega_l}{\hbar} - \frac{\dot{\gamma}_l}{\hbar}[M - (l + \tfrac{1}{2})\hbar]^2 + \cdots. \tag{45}$$

The subscript l indicates that M has been replaced by $(l + \frac{1}{2})\hbar$. The quantity ω is defined by

$$\omega = \theta - \gamma M = \int_{r_1}^{r_2} dr \frac{dV/dr}{2(E - V)}[2mr^2(E - V) - M^2]^{\frac{1}{2}}; \tag{46}$$

if we transform the variable of integration to ρ as given by (14), ω becomes simply $\int \sigma\, d\rho$ between the limits ρ_1 and ρ_2. Since σ, as defined in (5), is the projection of q along p, the quantity ω is the action integral for the radial motion. Indeed, by a partial integration the expression (46) can be written in the more familiar form $\int dr[2m(E - V) - M^2/r^2]^{\frac{1}{2}}$ from r_1 to r_2. The condition $\Delta = 0$ for the singularity becomes, therefore,

$$2\int_{r_1}^{r_2} dr[2m(E - V) - M^2/r^2]^{\frac{1}{2}} = (n + \tfrac{1}{2})2\pi\hbar \tag{47}$$

in addition to $M = (l + \frac{1}{2})\hbar$. Equation (47) is, of course, the well-known quantum condition for the radial motion. Usually, it is obtained by treating the radial motion as a 1-dimensional problem after Schrödinger's equation has been separated into angular and radial coordinates. This leads to the replacement of M^2 by $l(l + 1)\hbar^2$ rather than by $(l + \frac{1}{2})^2\hbar^2$, as in the present derivation, and consequently to erroneous results even for the simple hydrogen atom.

The evaluation of the integral (40) in the neighborhood of $\Delta = 0$ is fairly tricky. As shown in Appendix D, the result can be written as

$$i\int dM \frac{(2\pi M)^{\frac{1}{2}}}{|\dot{\gamma}|}\left(\frac{\Delta}{\dot{\gamma}}\right)^{-\frac{3}{2}} = \operatorname{sgn}\dot{\gamma}\,\frac{[(2l + 1)\pi]^{\frac{1}{2}}\hbar^2}{\omega_l - (n + \frac{1}{2})\pi\hbar}, \tag{48}$$

which multiplies the second and third line of (40) with M everywhere replaced by $(l + \frac{1}{2})\hbar$. Formula (48) represents the leading term of the singularity which is given by (47).

If the last formula is inserted into (40), we get the following approximation:

$$\begin{aligned} F(p''p'E) \cong{}& \sum_{l=0}^{\infty} \frac{(2l + 1)^{\frac{1}{2}}}{\pi^{\frac{3}{2}}} \frac{\cos[(l + \frac{1}{2})\zeta - \pi/4]}{(\sin\zeta)^{\frac{1}{2}}} \\ &\times \sum_n \frac{2(-1)^n \operatorname{sgn}\dot{\gamma}}{\omega_l - (n + \frac{1}{2})\pi\hbar} \\ &\times \left\{\left[\rho\frac{dV}{dr}\left(2m(E - V) - \frac{M^2}{r^2}\right)^{\frac{1}{2}}\right]'\right\}^{-\frac{1}{2}} \cos\left(\frac{\omega'}{\hbar} - \frac{\pi}{4}\right) \\ &\times \left\{\left[\rho\frac{dV}{dr}\left(2m(E - V) - \frac{M^2}{r^2}\right)^{\frac{1}{2}}\right]''\right\}^{-\frac{1}{2}} \cos\left(\frac{\omega''}{\hbar} - \frac{\pi}{4}\right) \end{aligned} \tag{49}$$

with ω' and ω'' integrals like (46) taken from r_1 to r' and from r'' to r_2. The factor between the two summation signs is the asymptotic formula for

$$(2l + 1/4\pi)P_l(\cos\zeta),$$

as is evident from (35). In polar coordinates (ρ, θ, ψ)

for p we have

$$\frac{2l+1}{4\pi} P_l(\cos\zeta) = \sum_{m=-l}^{+l} Y_{lm}(\theta''\varphi'')Y^*_{lm}(\theta'\varphi'), \quad (50)$$

where $Y_{lm}(\theta\varphi)$ are the normalized spherical harmonics (cf. Ref. 7). The ζ dependence in (49) represents, therefore, the sum over the $2l+1$ eigenstates of the total angular momentum $(l+\frac{1}{2})\hbar$.

The sum over n for a given value of l has also a simple interpretation. It constitutes the phase-integral approximation for the Green's function of the 1-dimensional Hamiltonian (5) with M fixed to the value $(l+\frac{1}{2})\hbar$. The conjugate variables are ρ and σ, and the action is defined by $\int \sigma\, d\rho$. These are different from the Hamiltonian and the action corresponding to the usual radial variables p_r and r. But the two treatments are similar in that neither ρ nor p_r cover the full range from 0 to ∞ because of the centrifugal potential which prevents the particle from getting into the center $r=0$ and increasing its kinetic energy indefinitely. The assumption (13) is crucial at this point since the centrifugal potential is able to overwhelm the attractive potential $V(q)$.

The radial dependence at the main singularities of $\tilde{F}(p''p'E)$, as given by (49), corresponds exactly to the results of the ordinary WKB method. The present treatment does not go beyond the usual one, although the discussion of momentum space is new, and there are fewer *ad hoc* assumptions. The apparent advantage of momentum space for the phase-integral approximation can only be exploited if a way is found to understand not only the δ-function singularities of $\tilde{F}(p''p'E)$ but also its branch-cuts.

ACKNOWLEDGMENT

The author would like to thank the Physics Division of the Aspen Institute for their generous hospitality during the summer of 1967 where part of this work was done.

APPENDIX A

The absence of orbital precession for small angular momentum can be shown most directly if we break up the integral (23) for γ into three parts: from r_1 to r' as in α of formula (27), from r' to r'', and from r'' to r_2 as in β. The intermediate points r' and r'' are arbitrary and independent of M. Since the integral of (27) is well behaved, except near r_1 and r_2, the integral from r' to r'' vanishes as M goes to zero.

In order to discuss the integrals α and β, we introduce the variables ξ_1 near r_1 and ξ_2 near r_2 through the common formula

$$\cos\xi = M/r[2m(E-V)]^{\frac{1}{2}}, \quad (A1)$$

with the understanding that $\xi_1 = 0$ for $r = r_1$, $\xi_2 = 0$ for $r = r_2$, and $0 < \xi < \pi/2$ in both cases. Since

$$d\xi = \frac{dr}{r}\frac{M}{[2mr^2(E-V)-M^2]^{\frac{1}{2}}} - \frac{dV/dr}{2(E-V)}\frac{M\,dr}{[2mr^2(E-V)-M^2]^{\frac{1}{2}}}, \quad (A2)$$

we can write the formulas

$$\alpha = \left(\arccos\frac{M}{r[2m(E-V)]^{\frac{1}{2}}}\right)_{r'} + \int_{r_1}^{r'}\frac{M\,dr}{[2mr^2(E-V)-M^2]^{\frac{1}{2}}}\left(\frac{dV/dr}{E-V}-\frac{1}{r}\right),$$
$$\beta = \left(\arccos\frac{M}{r[2m(E-V)]^{\frac{1}{2}}}\right)_{r''} + \int_{r''}^{r_2}\frac{dr}{r}\frac{M}{[2mr^2(E-V)-M^2]^{\frac{1}{2}}}, \quad (A3)$$

whose integrands are sufficiently well behaved even in the limit $M=0$. In the case of α, the assumption (13) has been used in order to show that the integrand is regular.

As M approaches zero, while r' and r'' remain fixed, the arguments of the arc cos become small. Therefore, we can write the expansion

$$\arccos x = \pi/2 - \arcsin x = \pi/2 - x - x^3/6 + \cdots, \quad (A4)$$

with $x = M/r[2m(E-V)]^{\frac{1}{2}}$ for $r = r'$ and $r = r''$. Thus, we get

$$\alpha = \frac{\pi}{2} - M\left[\left(\frac{1}{r[2m(E-V)]^{\frac{1}{2}}}\right)_{r'} - \int_{r_1}^{r'}\frac{dr}{r[2m(E-V)]^{\frac{1}{2}}}\left(\frac{dV/dr}{E-V}-\frac{1}{r}\right)\right] + \cdots,$$
$$\beta = \frac{\pi}{2} - M\left[\left(\frac{1}{r[2m(E-V)]^{\frac{1}{2}}}\right)_{r''} - \int_{r''}^{r_2}\frac{dr}{r}\frac{1}{r[2m(E-V)]^{\frac{1}{2}}}\right] + \cdots. \quad (A5)$$

α and β approach $\pi/2$ as M vanishes, independently of r' and r'', and γ approaches π.

APPENDIX B

If M is close to $\bar{M}$, we shall compute η according to (22) under the assumption that r' is close enough to

[7] Reference 6, p. 328.

$\bar{r}$ so as to guarantee the expansion $2mr^2(E-V) = \bar{M}[\bar{M} + \mu(r-\bar{r})^2 + \cdots]$ throughout the interval from r' to r. Also, we assume that all the factors in the integrand of (22) can be replaced by their value at $\bar{r}$, except the root in the denominator. Thus, we find for $M < \bar{M}$ that

$$\eta \cong \pm\left[\frac{dV/dr}{2(E-V)}\right]_{\bar{r}}\left(\int_{r'}^{\bar{r}} + \int_{\bar{r}}^{r}\right) \times \frac{\bar{M}\,dr}{\{2\bar{M}[\bar{M} - M + \mu(r-\bar{r})^2]\}^{\frac{1}{2}}},$$
$$\cong \pm\left[\frac{dV/dr}{2(E-V)}\right]_{\bar{r}}\left(\frac{\bar{M}}{2\mu}\right)^{\frac{1}{2}}\log\frac{\mu(\bar{r}-r')(r-\bar{r})}{\bar{M}-M}, \quad \text{(B1)}$$

if we keep only the leading terms as M approaches $\bar{M}$. If $M > \bar{M}$, we find, similarly, for β in (27) that

$$\beta \cong \pm\left[\frac{dV/dr}{2(E-V)}\right]_{\bar{r}} \times \int_{r''}^{r_2}\frac{\bar{M}\,dr}{\{2\bar{M}[\bar{M} - M + \mu(r-\bar{r})^2]\}^{\frac{1}{2}}}, \quad \text{(B2)}$$

where the upper limit of integration r_2 is given by the condition $\mu(r_2 - \bar{r})^2 = M - \bar{M}$. The leading term turns out to be

$$\beta \cong \pm\left[\frac{dV/dr}{2(E-V)}\right]_{\bar{r}}\left(\frac{\bar{M}}{2\mu}\right)^{\frac{1}{2}}\log\left(\frac{\mu(\bar{r}-r'')^2}{M-\bar{M}}\right)^{\frac{1}{2}}. \quad \text{(B3)}$$

In both (B1) and (B3), the distances $\bar{r} - r'$, $r - \bar{r}$, and $\bar{r} - r''$ can be replaced by any characteristic radius such as $\bar{r}$ to give the formula (24).

APPENDIX C

The equivalence of the formulas (9) and (26) can be demonstrated as follows. First, we find from (5) that

$$H_M = \frac{M}{\rho^2 r}\frac{dV}{dr} \quad \text{with} \quad r = \left(\sigma^2 + \frac{M^2}{\rho^2}\right)^{\frac{1}{2}}. \quad \text{(C1)}$$

By differentiating with respect to M and σ we get

$$H_{MM} = \frac{1}{\rho^2 r}\frac{dV}{dr} + \frac{M^2}{\rho^4 r^2}\frac{d^2V}{dr^2} - \frac{M^2}{\rho^4 r^3}\frac{dV}{dr}, \quad \text{(C2)}$$

$$\left(\frac{H_M^2}{H_\sigma}\right)_\sigma = \frac{M^2}{\rho^4 r^2}\frac{d^2V}{dr^2} - \frac{M^2}{\rho^4 r^3}\frac{dV}{dr} - \frac{M^2}{\rho^4\sigma^2 r}\frac{dV}{dr}. \quad \text{(C3)}$$

We can combine these two expressions into

$$\left[H_{MM} - \left(\frac{H_M^2}{H_\sigma}\right)_\sigma\right]dt = \left(\frac{1}{\rho^2 r}\frac{dV}{dr} + \frac{M^2}{\rho^4\sigma^2 r}\frac{dV}{dr}\right)\frac{r\,d\rho}{\sigma\,dV/dr}$$
$$= \left(\frac{1}{\rho^2\sigma} + \frac{M^2}{\rho^4\sigma^3}\right)d\rho = \frac{\partial}{\partial M}\left(\frac{M}{\rho^2\sigma}\right)d\rho, \quad \text{(C4)}$$

where σ in the last step is considered as a function of E, M, and ρ, which follows from (5). The formula (22) can also be written as $\int M\,d\rho/\rho^2\sigma$ if we remember that $\rho\,d\rho = m(dV/dr)\,dr$ according to (14). Since this change of variables does not involve M, the formulas (9) and (26) become, indeed, equivalent after H_σ' and H_σ'' have been expressed in terms of r' and r''.

Since σ is the projection of q along p, it vanishes when ρ reaches its extremal values because p is perpendicular to the radius vector q. As is apparent from (C4), the integrals (9) and (26) do not converge when σ vanishes. The troublesome term comes from the integral over $(H_M^2/H_\sigma)_\sigma$, and can be eliminated with the help of formula (10). This requires the calculation of $(H_M^2/H_\rho)_\rho$, which proceeds exactly as the calculation (C3), but turns out to be more complicated. The important feature is, however, that σ does not occur in any denominator anymore. The integral on the right-hand side of (10) is, therefore, as well behaved as the integral over H_{MM}. The singular parts in (9) and (26) are isolated in the last two terms of (10). More specifically, since both H_M and H_ρ are well behaved near the extremal values of ρ, the singularities in (9) and (26) arise from H_σ' and H_σ'' in the denominators of (10). If ρ'' is assumed to approach an extremal value of ρ, but not ρ', formula (9) shows that $1/D_T$ approaches the finite value

$$\frac{1}{D_T} \to \frac{\rho'\rho''\sin\eta}{M}(H_M'')^2 H_\sigma'/H_\rho''. \quad \text{(C5)}$$

Also, $(\partial\eta/\partial M)_{\rho'\rho''}$ has the leading term $(H_M'')^2/H_\sigma''H_\rho''$, which depends only on the end point ρ'', but not on the preceding trajectory.

APPENDIX D

In view of (45) we can write

$$\frac{\Delta}{\dot{\gamma}} = \frac{(2n+1)\pi - 2\omega_l/\hbar}{\dot{\gamma}} - \frac{1}{\hbar}[M - (l+\tfrac{1}{2})\hbar]^2 + \cdots, \quad \text{(D1)}$$

where n is assumed to have a small imaginary part ϵ. Since we assume $2\omega_l/\hbar$ to be close to $(2n+1)\pi$, the ratio $\Delta/\dot{\gamma}$ is always negative when M is sufficiently far away from $(l+\frac{1}{2})\hbar$. With our conventions about the discontinuity of $(\Delta/\dot{\gamma})^{-\frac{3}{2}}$ along the real positive $\Delta/\dot{\gamma}$ axis, the integrand in (48) is always positive (after including the factor i) when M is not close to $(l+\frac{1}{2})\hbar$.

If $\mathrm{Re}\,[(2n+1)\pi\hbar - 2\omega_l]/\dot{\gamma} < 0$, the integral becomes

$$\int dM\left\{\frac{2\omega_l - (2n+1)\pi\hbar}{\dot{\gamma}} + [M - (l+\tfrac{1}{2})\hbar]^2\right\}^{-\frac{3}{2}} \times \frac{\hbar^{\frac{3}{2}}(2\pi M)^{\frac{1}{2}}}{|\dot{\gamma}|} \times (\text{r.f.}). \quad \text{(D2)}$$

The "regular factors" (r.f.) represent all the remaining terms which appear in (40) and have no singular behavior at $M = (l + \frac{1}{2})\hbar$. The $-\frac{3}{2}$ power has been written so as to make it obviously positive, as it has to be in this case, even when M is close to $(l + \frac{1}{2})\hbar$. Since the main variation comes from the $-\frac{3}{2}$ power in (D2), the leading term is obtained if the "regular factors" including $(2\pi M)^{\frac{1}{2}}/|\dot{\gamma}|$ are replaced by their values at $M = (l + \frac{1}{2})\hbar$. The integration over M is trivial and gives the expression on the right of (48) times the "regular factors" at $M = (l + \frac{1}{2})\hbar$.

If Re $[(2n + 1)\pi\hbar - 2\omega_l]/\dot{\gamma} > 0$, we get two integrals, one where $\Delta/\dot{\gamma}$ has a negative real part and one where the real part of $\Delta/\dot{\gamma}$ is positive. The integrand of (48) is positive in the former region, as in the preceding case, but it is purely imaginary in the latter region. In conformity with the definition of the $-\frac{3}{2}$ power, we get

$$\int dM\left([M - (l + \tfrac{1}{2})\hbar]^2 + \frac{2\omega_l - (2n + 1)\pi\hbar}{\dot{\gamma}}\right)^{-\frac{3}{2}} \times \frac{\hbar^{\frac{3}{2}}(2\pi M)^{\frac{1}{2}}}{|\dot{\gamma}|} \times (\text{r.f.}) + i \operatorname{sgn}\left(\frac{\epsilon}{\dot{\gamma}}\right) \times \int dM\left(\frac{(2n + 1)\pi\hbar - 2\omega_l}{\dot{\gamma}} + [M - (l + \tfrac{1}{2})\hbar]^2\right)^{-\frac{3}{2}} \times \frac{\hbar^{\frac{3}{2}}(2\pi M)^{\frac{1}{2}}}{|\dot{\gamma}|} \times (\text{r.f.}), \tag{D3}$$

where first integral goes over $[M - (l + \frac{1}{2})\hbar]^2 > [(2n + 1)\pi\hbar - 2\omega_l]/\dot{\gamma}$, and the second over

$$[M - (l + \tfrac{1}{2})\hbar]^2 < [(2n + 1)\pi\hbar - 2\omega_l]/\dot{\gamma},$$

in the limit of very small ϵ.

Each integral is really divergent at the limits of integration, and we have to perform the formal integration by parts which is used in the theory of generalized functions so as to arrive at a finite result.[6] Thus, we introduce the new variable

$$\mu = \frac{M - (l + \frac{1}{2})\hbar}{\{[(2n + 1)\pi\hbar - 2\omega_l]/\dot{\gamma}\}^{\frac{1}{2}}} - 1, \tag{D4}$$

and the first integral becomes

$$\frac{2\hbar^{\frac{3}{2}} \operatorname{sgn} \dot{\gamma}}{(2n + 1)\pi\hbar - 2\omega_l}\left[\int \frac{2\,d\mu}{\mu^{\frac{1}{2}}} \frac{d}{d\mu}\left(\frac{(2\pi M)^{\frac{1}{2}} \times (\text{r.f.})}{(2 + \mu)^{\frac{3}{2}}}\right) - \frac{2}{\sqrt{\mu}} \frac{(2\pi M)^{\frac{1}{2}} \times (\text{r.f.})}{(2 + \mu)^{\frac{3}{2}}}\right], \tag{D5}$$

where the factor 2 takes into account the double range of integration, for $M > (l + \frac{1}{2})\hbar$ and for $M < (l + \frac{1}{2})\hbar$. The part which is integrated out has to be evaluated at some upper limit of $\mu \gg 1$ and does not contribute to any singular behavior near $M = (l + \frac{1}{2})\hbar$. If we take the derivative with respect to μ in the first part of (D5), we get a contribution from the derivative of $(2\pi M)^{\frac{1}{2}} \times$ r.f. and one from the derivative of $(2 + \mu)^{-\frac{3}{2}}$. In the former contribution a factor $\{[(2n + 1)\pi\hbar - 2\omega_l]\dot{\gamma}\}^{\frac{1}{2}}$ appears as soon as we revert to M as variable of integration, thereby lowering the singular behavior near $M = (l + \frac{1}{2})\hbar$. Therefore, only the derivative of $(2 + \mu)^{-\frac{3}{2}}$ is of interest and we are finally led to the expression

$$\frac{2\hbar^{\frac{3}{2}} \operatorname{sgn} \dot{\gamma}}{(2n + 1)\pi\hbar - 2\omega_l} (2\pi M)^{\frac{1}{2}} \times (\text{r.f.}) \int_0^\infty \frac{2\,d\mu}{(\mu)^{\frac{1}{2}}} \frac{-\frac{3}{2}}{(2 + \mu)^{\frac{5}{2}}}. \tag{D6}$$

After the elementary integral is evaluated, we find again the formula (48) times the "regular factors" at $M = (l + \frac{1}{2})\hbar$.

The second integral in (D3) is treated exactly the same way, but the integrated part [corresponding to the second term in (D5)] cannot be thrown away because it is to be evaluated at $\mu = 1$, not $\mu \gg 1$. Therefore, it contributes to the main singularity at $M = (l + \frac{1}{2})\hbar$. After taking the derivative with respect to μ [corresponding to the first term (D5)] and keeping only the leading term, it turns out that the two parts cancel each other exactly. Only the first integral in (D3) contributes to the leading singularity.

Dynamics near Equilibrium of Systems Described in Thermal Hartree–Fock Approximation

B. U. FELDERHOF*
Baker Laboratory, Cornell University, Ithaca, New York 14850

(Received 9 August 1968)

The dynamics of systems described in Hartree–Fock approximation is studied near a stationary point of the free energy. It is shown that the second-order free-energy functional is a constant for the linearized self-consistent equation of motion. This leads to the stability criterion derived by Mermin. A simple collision model is constructed and is shown to satisfy the H theorem. It exhibits the effect of critical slowing down. The formalism is shown to be applicable to superconductors.

1. INTRODUCTION

The stability properties of many-particle systems described in Hartree–Fock approximation[1] have been studied by Thouless.[2] The relation between thermodynamic and dynamic stability in the finite temperature case has been investigated by Mermin.[3] It is the purpose of this paper to extend their work in several directions. Our first objective is to study the linearized self-consistent equation of motion from a formal point of view. It is shown that the second-order free-energy functional is a constant of the motion. This leads to the stability criterion already derived by Mermin. Also, a more general class of constants of the motion may be found.

The self-consistent equation of motion for the single-particle-density matrix derived in the thermal Hartree–Fock approximation is formally time-reversible. One may hope that corrections to the approximation lead to a collision term in the equation of motion. A simple collision model is constructed which satisfies an H theorem, i.e., it leads to a monotonic decrease in free energy. Moreover, it is shown that the model implies critical slowing down near a critical point.

Although the self-consistent field approximation often is a very good approximation away from the critical point, it usually breaks down in its immediate neighborhood. Only in the case of superconductors is the approximation valid up to very close to the critical point.[4] For this reason it seems worthwhile to show that the formalism may be extended to incorporate the pairing effects present in superconductors.

2. THERMAL HARTREE–FOCK APPROXIMATION

According to a general statistical-mechanical definition, a many-particle system with Hamiltonian $\mathcal{H}$ in contact with a heat bath at temperature T in a state described by the statistical operator $\mathcal{S}$ has a free energy[5]

$$\mathcal{F} = Sp\mathcal{S}\mathcal{H} + k_BTSp\mathcal{S}\ln\mathcal{S}. \tag{2.1}$$

The free energy takes its minimum value $\mathcal{F}_{eq} = -k_BT\ln Z$ in the equilibrium state characterized by the canonical distribution $\mathcal{S}_{eq} = \exp(-\beta\mathcal{H})/Z$, where $Z = Sp\exp(-\beta\mathcal{H})$ is the partition function and $\beta = 1/k_BT$. We shall consider a system of fermions with Hamiltonian (in obvious notation)

$$\mathcal{H} = \sum K_{ij}c_i^+c_j + \tfrac{1}{2}\sum V(ij\,|\,kl)c_i^+c_j^+c_kc_l. \tag{2.2}$$

The two-particle interaction V may be taken to have the symmetry properties

$$V(ij\,|\,kl) = -V(ji\,|\,kl) = V(ji\,|\,lk) = V(kl\,|\,ij)^*. \tag{2.3}$$

In Hartree–Fock approximation, one considers states given by statistical operators of the form

$$\mathcal{S} = \exp[-\sum\Omega_{ij}c_i^+c_j]/Sp\exp[-\sum\Omega_{ij}c_i^+c_j], \tag{2.4}$$

where Ω is an Hermitian matrix. For these states the reduced density matrix ρ defined by $\rho_{ij} = Sp\mathcal{S}c_j^+c_i$ is related to Ω by

$$\rho = [\exp\Omega + 1]^{-1}. \tag{2.5}$$

The free energy (2.1) for these states becomes a functional of ρ given by

$$\begin{aligned}\mathcal{F}(\rho) &= \mathcal{E}(\rho) - T\mathcal{S}(\rho),\\ \mathcal{E}(\rho) &= \mathrm{Tr}\,\rho K + \tfrac{1}{2}\,\mathrm{Tr}\,\rho W,\\ \mathcal{S}(\rho) &= -k_B[\mathrm{Tr}\,\rho\ln\rho + \mathrm{Tr}\,(1-\rho)\ln(1-\rho)],\end{aligned} \tag{2.6}$$

* On leave of absence from the Institute for Theoretical Physics, Rijksuniversiteit, Utrecht, The Netherlands.

[1] For a general introduction see J. G. Valatin, *Lectures in Theoretical Physics, Boulder, 1961* (Interscience Publishers, Inc., New York, 1962), Vol. IV, p. 1; D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic Press Inc., New York, 1961); D. Pines and P. Nozières, *The Theory of Quantum Liquids* (W. A. Benjamin, Inc., New York, 1966), Vol. I.

[2] D. J. Thouless, Nucl. Phys. **21**, 225 (1960); Nucl. Phys. **22**, 78 (1961).

[3] N. D. Mermin, Ann. Phys. (N.Y.) **21**, 99 (1963).

[4] L. P. Kadanoff *et al.*, Rev. Mod. Phys. **39**, 395 (1967).

[5] See, e.g., J. M. Blatt, *Theory of Superconductivity* (Academic Press Inc., New York, 1964).

where the trace involves a sum over single-particle states and where W is determined self-consistently by ρ according to

$$W_{il} = 2\sum_{jk} V(ij \mid kl)\rho_{kj}. \tag{2.7}$$

Writing $\rho = \rho_0 + \rho_1$, one may expand the free-energy functional in terms of ρ_1, so that $\mathcal{F}(\rho) = \mathcal{F}_0 + \mathcal{F}_1 + \mathcal{F}_2 + \cdots$, and ask for ρ_0 for which the free energy is stationary, i.e., for which the first variation $\mathcal{F}_1$ vanishes identically. From (2.6) it follows that the first-order energy $\mathcal{E}_1$ is given by $\mathrm{Tr}\,\rho_1 H_0$, where $H_0 = K + W_0$ is determined self-consistently from ρ_0 by (2.7). The first-order entropy $\mathcal{S}_1$ is given by

$$\mathcal{S}_1 = -k_B \,\mathrm{Tr}\,\rho_1[\ln \rho_0 - \ln(1 - \rho_0)]. \tag{2.8}$$

Hence the free energy is stationary for ρ_0, satisfying

$$\rho_0 = [\exp(\beta H_0) + 1]^{-1}. \tag{2.9}$$

This is a Fermi-distribution of quasiparticles (for convenience, the chemical potential has been included in K, so there is no restriction from the total number of particles).

The second-order free-energy functional $\mathcal{F}_2$ determines the thermodynamic stability of the state ρ_0. At this point, it is convenient to introduce a Hilbert space $\mathfrak{H}$ with kets $|a\rangle$ given by matrices a of the same dimensionality as ρ and with a scalar product defined by

$$\langle a \mid b\rangle = \mathrm{Tr}\, a^+ b. \tag{2.10}$$

In this notation, the matrix W given by (2.7) defines a linear operator U on $\mathfrak{H}$ by

$$|W_1\rangle = U\,|\rho_1\rangle. \tag{2.11}$$

U is Hermitian in the scalar product (2.10) on account of the symmetry properties (2.3). The second-order energy functional is given by

$$\mathcal{E}_2 = \tfrac{1}{2}\langle \rho_1|\, U\,|\rho_1\rangle. \tag{2.12}$$

The calculation of the second-order entropy functional from (2.6) is complicated by the fact that the matrices ρ_0 and ρ_1 in general do not commute. First variation of the identity $[\varphi(\rho), \rho] = 0$ yields

$$[\varphi_0, \rho_1] + [\varphi_1, \rho_0] = 0, \tag{2.13}$$

so that in the representation where φ_0 and ρ_0 are diagonal

$$(\varphi_1)_{ij} = \frac{\varphi_{0i} - \varphi_{0j}}{\rho_{0i} - \rho_{0j}}(\rho_1)_{ij}. \tag{2.14}$$

Moreover, expanding φ as $\varphi_0 + \varphi_1 + \varphi_2 + \cdots$, one may show that $\mathrm{Tr}\,\rho_0\varphi_2 = -\tfrac{1}{2}\,\mathrm{Tr}\,\rho_1\varphi_1$. Hence one obtains

$$\mathcal{S}_2 = \frac{-1}{2T}\langle \rho_1|\, G\,|\rho_1\rangle, \tag{2.15}$$

where the linear operator G is defined by its action in the representation where ρ_0 and H_0 are diagonal as

$$(G\,|a\rangle)_{ij} = -\frac{\epsilon_i - \epsilon_j}{f_i - f_j}a_{ij}, \tag{2.16}$$

where ϵ_i are the eigenvalues of H_0 and

$$f_i = [\exp(\beta\epsilon_i) + 1]^{-1},$$

the eigenvalues of ρ_0. A more formal definition of G not bound to a particular representation is obtained as follows. Corresponding to any Hermitian, positive-definite, and bounded matrix $p = \exp l$, define the linear operator $\Phi(p)$ by

$$\Phi(p)\,|a\rangle = \left|\int_0^1 e^{(1-s)l} a e^{sl}\, ds\right\rangle. \tag{2.17}$$

It is easily shown that $\Phi(p)$ is Hermitian and positive-definite. In terms of its inverse, the operator G may be written

$$G = k_B T[\Phi^{-1}(\rho_0) + \Phi^{-1}(1 - \rho_0)]. \tag{2.18}$$

G is also Hermitian and positive-definite. From (2.12) and (2.15), it follows that the second-order free-energy functional $\mathcal{F}_2$ is given by

$$\mathcal{F}_2 = \tfrac{1}{2}\langle \rho_1|\, F\,|\rho_1\rangle, \tag{2.19}$$

where $F = U + G$ is a Hermitian operator.

3. DYNAMICS IN HARTREE–FOCK APPROXIMATION

In Hartree–Fock approximation, ρ satisfies the nonlinear equation of motion

$$i\hbar\frac{\partial \rho}{\partial t} = [H, \rho], \tag{3.1}$$

where $H = K + W$ depends self-consistently on ρ. From (2.9) it follows that a stationary point of the free energy ρ_0 is a time-independent solution of this equation. Linearizing about ρ_0, one obtains for ρ_1 the linear equation

$$i\hbar\frac{\partial \rho_1}{\partial t} = [H_0, \rho_1] + [H_1, \rho_0]. \tag{3.2}$$

In the notation of the previous section, this may be written

$$i\hbar\frac{\partial}{\partial t}|\rho_1\rangle = S\,|\rho_1\rangle, \tag{3.3}$$

where S is a linear streaming operator. Defining the Hermitian operator C by

$$C\,|a\rangle = |[a, \rho_0]\rangle, \tag{3.4}$$

one obtains from (2.11), (2.16), and (3.2) that the

streaming operator is given by

$$S = C(G + U) = CF. \tag{3.5}$$

C and G commute, but C and U do not, so that S is a non-Hermitian operator.

From the fundamental relation (3.5), one may derive some useful properties of the linearized motion. In particular, we shall be interested in the constants of the motion. We ask for operators A with expectation values $\langle\rho_1| A |\rho_1\rangle$ which are independent of time. From (3.3) it follows that the necessary and sufficient condition for this to be valid for all ρ_1 is $S^+A = AS$, or, using the fact that C and F are both Hermitian,

$$FCA = ACF. \tag{3.6}$$

An obvious solution of this equation is $A_n = FS^n$, when n is a positive integer or, by linear combination $A = F\varphi(S)$, where $\varphi(z)$ is an arbitrary function of z, analytic at $z = 0$. The most important of this class of constants is F itself. Evidently, the second-order free-energy functional $\mathcal{F}_2$ is constant in time. As an immediate consequence, one may derive a stability criterion. Suppose one has found an eigenmode $|\rho_\omega\rangle$ for which $S|\rho_\omega\rangle = \hbar\omega|\rho_\omega\rangle$, with a corresponding solution of (3.3). From the fact that $\langle\rho_1(t)| F |\rho_1(t)\rangle$ does not depend on time, it follows that

$$\langle\rho_\omega| F |\rho_\omega\rangle = 0 \quad \text{if} \quad \omega \neq \omega^*. \tag{3.7}$$

Hence, for unstable modes, $\mathcal{F}_2$ vanishes identically. Consequently, the stationary state ρ_0 is dynamically stable if F is positive-definite. This is a sufficient but not a necessary condition for dynamic stability, as is evident from the fact that ρ_0 is also dynamically stable if F is negative-definite.

Further constants of the motion may exist. Suppose there is an Hermitian matrix q which commutes with H_0. Hence it follows that

$$S^+|q\rangle = FC|q\rangle = F|[q, \rho_0]\rangle = 0, \tag{3.8}$$

which implies

$$S^+|q\rangle\langle q| = |q\rangle\langle q| S, \tag{3.9}$$

so that $Q = |q\rangle\langle q|$ is a constant of the motion. More generally, suppose q does not commute with H_0, but corresponds to a broken symmetry, i.e., q is the generator of a transformation which leaves the equilibrium free energy invariant,

$$\rho_{0\lambda} = e^{i\lambda q}\rho_0 e^{-i\lambda q} = \rho_0 + i\lambda[q, \rho_0] + \cdots,$$
$$\mathcal{F}(\rho_{0\lambda}) = \mathcal{F}(\rho_0). \tag{3.10}$$

To second order in λ, this implies, since ρ_0 is a stationary point,

$$\langle[q, \rho_0]| F |[q, \rho_0]\rangle = 0. \tag{3.11}$$

If the equilibrium is stable but indifferent against the symmetry operation, the free-energy operator F is positive-semidefinite, and from (3.11) one may infer that (3.8) holds. Hence again $Q = |q\rangle\langle q|$ is a constant of the motion. From (3.11) it also follows that

$$S|[q, \rho_0]\rangle = 0. \tag{3.12}$$

Hence $|[q, \rho_0]\rangle$ is a (collective) eigenmode of zero frequency (Goldstone's theorem[6]).

4. RELAXATION TOWARDS EQUILIBRIUM

The nonlinear self-consistent equation of motion (3.1) does not describe the approach to equilibrium of the system. It is easily shown from the commutator nature of the right-hand side that both the energy $\mathcal{E}(\rho)$ and the entropy $\mathcal{S}(\rho)$ are constants of the motion. One may hope that corrections to the Hartree–Fock approximation lead to a collision term expressed solely in terms of ρ. Thus we formally write

$$\frac{\partial\rho}{\partial t} = \frac{1}{i\hbar}[H, \rho] - R\rho, \tag{4.1}$$

where the relaxation operator R is a nonlinear operator acting on ρ which one expects to be such that an H theorem is valid ensuring a monotonic decrease of free energy,

$$\frac{\partial\mathcal{F}}{\partial t} = -\mathrm{Tr}\frac{\delta\mathcal{F}}{\delta\rho}(R\rho) \leq 0, \tag{4.2}$$

with equality only if

$$\frac{\delta\mathcal{F}}{\delta\rho} \equiv H + k_B T[\ln\rho - \ln(1 - \rho)] = 0, \tag{4.3}$$

i.e., at a stationary point ρ_0.

In the neighborhood of ρ_0, the equation of motion (4.1) may be linearized to

$$\frac{\partial}{\partial t}|\rho_1\rangle = -(R' + iS/\hbar)|\rho_1\rangle = -M|\rho_1\rangle, \tag{4.4}$$

where R' is the Fréchet derivative of R at ρ_0. We define an Onsager operator L by

$$M = LF = (B + iC/\hbar)F. \tag{4.5}$$

If the relaxation mechanism is such that $R' = BF$ satisfies a detailed balance relation expressed by

$$B^+ = B, \tag{4.6}$$

and if furthermore B is positive-definite, one may

[6] J. Goldstone, Nuovo Cimento **19**, 154 (1961); H. Stern, Phys. Rev. **147**, 94 (1966).

derive an H theorem for the linear regime

$$\frac{\partial \mathcal{F}_2}{\partial t} = -\tfrac{1}{2}\langle\rho_1|\,(M^+F + FM)\,|\rho_1\rangle$$
$$= -\langle\rho_1|\,FBF\,|\rho_1\rangle \leq 0, \tag{4.7}$$

where the equality sign obtains only for $F\,|\rho_1\rangle = 0$. We restrict ourselves to stable equilibrium ρ_0; hence F is positive-semidefinite and

$$|\rho_1\rangle = 0 \quad \text{and} \quad |\rho_1\rangle = i\lambda\,|[q, \rho_0]\rangle. \tag{4.8}$$

The latter case corresponds to a broken symmetry; in this case the parameter λ is determined by the initial conditions.

The actual construction of the relaxation operator R requires a detailed kinetic theory. We shall here be content with a single relaxation-time model and show that this model has the desired properties. A natural choice is

$$R\rho = \frac{1}{\tau}[\rho - \rho_{\text{loc}}], \tag{4.9}$$

where τ is a relaxation time and ρ_{loc} is defined by

$$\rho_{\text{loc}} = [\exp(\beta H) + 1]^{-1}, \tag{4.10}$$

where H is determined by ρ according to (2.7) and therefore contains the local self-consistent field rather than its equilibrium value. By comparison with (2.9), it follows that (4.10) is of the form of the equilibrium density matrix, but is not self-consistent.

This collision model satisfies the H theorem (4.2). In order to prove the inequality, it suffices to show that

$$\operatorname{Tr}(\ln\rho - \ln\rho_{\text{loc}})(\rho - \rho_{\text{loc}}) \geq 0, \tag{4.11}$$

with equality only for $\rho = \rho_{\text{loc}}$. This monotonicity property is proved by writing

$$\rho_{\text{loc}} = \exp a,$$
$$\rho(\lambda) = \exp(a + \lambda b), \quad \rho = \rho(1), \tag{4.12}$$

and defining the function $f(\lambda)$ by

$$f(\lambda) = \lambda^{-1}\operatorname{Tr}(\ln\rho(\lambda) - \ln\rho_{\text{loc}})(\rho(\lambda) - \rho_{\text{loc}})$$
$$= \operatorname{Tr} b(e^{a+\lambda b} - e^a). \tag{4.13}$$

Differentiating with respect to λ, one obtains

$$df/d\lambda = \langle b|\,\Phi(\rho(\lambda))\,|b\rangle, \tag{4.14}$$

where we have used the notation (2.17). Since $df/d\lambda \geq 0$ and $f(0) = 0$, it follows that $f(1) \geq 0$, which proves (4.11).

In the linear regime, the H theorem (4.7) is satisfied. Linearizing (4.9) about a stationary point ρ_0, one obtains

$$R'\,|\rho_1\rangle = \frac{1}{\tau}\left|\left(\rho_1 + \rho_0\left\{\int_0^1 \exp[(1-s)\beta H_0]\beta H_1 \times \exp(s\beta H_0)\,ds\right\}\rho_0\right)\right\rangle. \tag{4.15}$$

Hence,

$$R' = \tau^{-1}[I + G^{-1}U], \tag{4.16}$$

where I is the unit operator and we have used (2.12) and (2.19). Writing $R' = BF$, we obtain

$$B = \frac{1}{\tau}G^{-1}. \tag{4.17}$$

Since G^{-1} is Hermitian and positive-definite, the operator B in this model has the desired properties derived previously.

It is of interest to note that this model exhibits the effect of critical slowing down. The relative free-energy decrease is bounded by

$$-\frac{\partial \ln \mathcal{F}_2}{\partial t} \leq \frac{2\beta}{\tau}\frac{\langle\rho_1|\,F^2\,|\rho_1\rangle}{\langle\rho_1|\,F\,|\rho_1\rangle}, \tag{4.18}$$

where we have employed (4.7), (4.17), and (2.19). Near the critical point, the free-energy minimum becomes very shallow and hence the right-hand side of (4.18) becomes very small for a class of density matrices. Hence the corresponding transport processes become very slow near the critical point.[7]

5. EXTENSION TO SUPERCONDUCTORS

The formalism of the previous sections may be extended to incorporate the pairing effects present in superconductors. For simplicity, we describe the electrons by the Gor'kov Hamiltonian[8] and omit the Coulomb interaction. The class of Hartree–Fock statistical operators defined in (2.4) must now be generalized to include pairs of operators $c_i c_j$ and $c_i^+ c_j^+$ in the exponent. If the one-electron Hamiltonian is independent of spin, one may use the reduced density matrix (in r-representation),

$$\rho(\mathbf{r}, \mathbf{r}') = \begin{pmatrix} \langle\psi^+(\mathbf{r}'\uparrow)\psi(\mathbf{r}\uparrow)\rangle & \langle\psi(\mathbf{r}'\downarrow)\psi(\mathbf{r}\uparrow)\rangle \\ \langle\psi^+(\mathbf{r}'\uparrow)\psi^+(\mathbf{r}\downarrow)\rangle & \langle\psi(\mathbf{r}'\downarrow)\psi^+(\mathbf{r}\downarrow)\rangle \end{pmatrix}, \tag{5.1}$$

where ψ^+, ψ are the electron-field operators and the pointed brackets indicate average over the statistical operator related to ρ by (2.5). The reduced density matrix (5.1) satisfies a self-consistent equation of

[7] M. Fixman, in *Advances in Chemical Physics*, I. Prigogine, Ed. (Interscience Publishers, Inc., New York, 1963), Vol. VI; K. Kawasaki, Phys. Rev. **150**, 291 (1966).

[8] See, e.g., P. G. de Gennes, *Superconductivity of Metals and Alloys* (W. A. Benjamin, Inc., New York, 1966).

motion of the form (3.1) with H given by

$$H(\mathbf{r}, \mathbf{r}') = \begin{pmatrix} \mathcal{H}_e(\mathbf{r}, \mathbf{r}') & 0 \\ 0 & -\mathcal{H}_e^*(\mathbf{r}, \mathbf{r}') \end{pmatrix} + \begin{pmatrix} U_\downarrow(\mathbf{r}, t) & \Delta(\mathbf{r}, t) \\ \Delta^*(\mathbf{r}, t) & -U_\uparrow(\mathbf{r}, t) \end{pmatrix} \delta(\mathbf{r} - \mathbf{r}'), \quad (5.2)$$

where $\mathcal{H}_e(\mathbf{r}, \mathbf{r}')$ is the one-electron Hamiltonian, including the chemical potential. $U_\sigma(\mathbf{r}, t)$ are the self-consistent Hartree–Fock potentials ($\sigma = \uparrow, \downarrow$) and $\Delta(\mathbf{r}, t)$ is the pair potential,

$$U_\sigma(\mathbf{r}, t) = -V\langle\psi^+(\mathbf{r}, \sigma)\psi(\mathbf{r}, \sigma)\rangle,$$
$$\Delta(\mathbf{r}, t) = -V\langle\psi(\mathbf{r}\downarrow)\psi(\mathbf{r}\uparrow)\rangle, \quad (5.3)$$

where V is the strength of the Gor'kov interaction. If one introduces the Hermitian matrices $D_j(\mathbf{x})$ defined by

$$D_0(\mathbf{x}; \mathbf{r}, \mathbf{r}') = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \delta(\mathbf{x} - \mathbf{r})\delta(\mathbf{r} - \mathbf{r}'),$$
$$\mathbf{D}(\mathbf{x}; \mathbf{r}, \mathbf{r}') = \boldsymbol{\sigma}\delta(\mathbf{x} - \mathbf{r})\delta(\mathbf{r} - \mathbf{r}'), \quad (5.4)$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli spin matrices, one may write H in the form $H = K + W$, where K is the first matrix in (5.2), and W is the second matrix, which may also be written

$$W = \tfrac{1}{2}V \sum_j \eta_j \int d\mathbf{x} D_j(\mathbf{x}) \operatorname{Tr} (\rho - O)D_j(\mathbf{x}). \quad (5.5)$$

The sign operators η_j are defined by

$$\eta_0 = +1, \quad \eta_x = \eta_y = \eta_z = -1, \quad (5.6)$$

and the matrix O by

$$O(\mathbf{r}, \mathbf{r}') = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \delta(\mathbf{r} - \mathbf{r}'). \quad (5.7)$$

In (5.5), the difference of ρ and O has to be formed before taking the trace in order to avoid divergencies. The free energy is again a functional $\mathcal{F}(\rho)$ given by (2.6), with ρ replaced by $\rho - O$ in the expression for $\mathcal{E}(\rho)$.

One may again define a Hilbert space. The energy operator U defined by (2.11) in this case is given by

$$U = \tfrac{1}{2}V \sum_j \eta_j \int d\mathbf{x} \, |D_j(\mathbf{x})\rangle\langle D_j(\mathbf{x})|, \quad (5.8)$$

which is obviously Hermitian in the scalar product (2.10). From (2.12) and (5.8), one obtains the second-order energy functional

$$\mathcal{E}_2 = -\frac{1}{V}\int d\mathbf{x}[|\Delta_1(\mathbf{x}, t)|^2 + U_{\uparrow 1}(\mathbf{x}, t)U_{\downarrow 1}(\mathbf{x}, t)]. \quad (5.9)$$

The second-order entropy functional $\mathcal{S}_2$ may be calculated explicitly from (2.15) and (2.16) in a representation where H_0 is diagonal. In the spatially homogeneous case, it is convenient to first transform to a plane-wave representation and to define

$$\rho(\mathbf{k}, \mathbf{k}') = \frac{1}{\mathcal{V}}\int \exp(-i\mathbf{k}\cdot\mathbf{r})\rho(\mathbf{r}, \mathbf{r}') \exp(i\mathbf{k}'\cdot\mathbf{r}') \, d\mathbf{r} \, d\mathbf{r}', \quad (5.10)$$

where $\mathcal{V}$ is the system volume and each element $\rho(\mathbf{k}, \mathbf{k}')$ is itself a 2×2 matrix. Below the transition temperature, the system condenses into a superconducting state ρ_0 characterized by a spatially homogeneous order parameter Δ_0, which may be chosen real and is to be determined self-consistently. In this representation, H_0 is given by

$$H_0(\mathbf{k}, \mathbf{k}') = \begin{pmatrix} \xi(\mathbf{k}) & \Delta_0 \\ \Delta_0 & -\xi(\mathbf{k}) \end{pmatrix} \delta_{\mathbf{k},\mathbf{k}'}, \quad (5.11)$$

where $\xi(\mathbf{k}) = \hbar^2k^2/2m - \mu$, with m the electron mass and μ the chemical potential. Finally, one may diagonalize $H_0(\mathbf{k}, \mathbf{k})$ by a 2×2 transformation and calculate $\mathcal{S}_2$. According to the theory of Sec. 3, the second-order free energy $\mathcal{E}_2 - T\mathcal{S}_2$ is a constant of the linearized motion.

The undetermined phase of the equilibrium order parameter implies a broken symmetry. We have chosen Δ_0 real and hence the matrix q which generates the transformation leaving the equilibrium free energy invariant is given by

$$q(\mathbf{k}, \mathbf{k}') = \sigma_y\delta_{\mathbf{k},\mathbf{k}'}. \quad (5.12)$$

From Sec. 3 it follows that $\operatorname{Im}\int \Delta_1(\mathbf{x}, t) \, d\mathbf{x}$ is the corresponding constant of the motion.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation and by the Advanced Research Projects Agency through the Materials Science Center at Cornell University.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Next-Nearest-Neighbor Ising Model

R. W. GIBBERD
Theoretical Physics Institute, Physics Department, University of Alberta, Edmonton, Alberta

(Received 10 September 1968)

The next-nearest-neighbor two-dimensional Ising model is cast into a form which resembles a one-dimensional interacting many-fermion system. An approximation, which has previously been shown to be successful in giving the critical properties of the two-dimensional ferroelectric problem, is used. From the approximate expression obtained, the critical indices are found to be $\alpha = \alpha' = 0$, $\beta = \frac{1}{8}$, which agrees with the results obtained from series expansions and plausible physical arguments. The critical temperature obtained agrees to within 6% of the series expansion results.

I. INTRODUCTION

For many years now, it has been known that the algebraic and combinatorial methods used to solve exactly the two-dimensional square-lattice Ising model break down when one tries to extend them to three-dimensional or nonplanar-type lattices. Not only do the techniques fail for these more complicated lattices, but they have not been useful in obtaining approximate or perturbation solutions. Recently, the combinatorial or Pfaffian method has been recast into a form which resembles a one-dimensional many-fermion system.[1] It is found that the expressions obtained for lattice problems which are soluble by the Pfaffian technique resemble a noninteracting fermion system and hence can be evaluated exactly, whereas the lattice problems that have not been solved resemble interacting fermion systems. Although these latter expressions cannot be treated exactly, the similarity to many-fermion problems enables one to use the techniques of quantum many-fermion theory to generate approximate solutions.

One problem that has already been considered in this formalism is the two-dimensional ferroelectric and antiferroelectric problem,[2] where it was shown that the first-order or free-fermion approximation reproduced correctly most of the critical properties of this model. In this paper, we examine an unsolved model, the two-dimensional next-nearest-neighbor Ising model and look at the critical properties as given by the first-order approximation. It is found that the results agree very well with the series-expansion results of Dalton and Wood.[3]

In Sec. II, the partition function for the next-nearest-neighbor Ising model is written as the vacuum-to-vacuum expectation value of a time-ordered product of exponentials of Fermi operators. This expression is called the S matrix and cannot be evaluated exactly because of the quartic terms of Fermi operators that appear. However, the S matrix may be written as a perturbation series about this term, and in Sec. III the critical indices and critical temperature are calculated from the first-order approximation to the S matrix. Since the critical point T_c is determined by the temperature at which the partition function has a singularity, and since the critical indices are determined by the nature of the singularity, in making this approximation we are assuming that the analytic structure of the exact S matrix is directly related to the analytic structure of the approximate S matrix. This assumption is the basis of many calculations in quantum field theory, where, for example, in the case of strongly interacting particles, the perturbation method is inapplicable because the perturbation series cannot necessarily be expected to converge. However, despite this, many people[4] consider that the singularity structure of some of the first few perturbation terms may contain useful information about the analytical properties of the complete S matrix. This indeed appears to be the case for the model considered here, since we obtain $\alpha = \alpha' = 0$, $\beta = \frac{1}{8}$ from the first approximation. These are the results obtained by series expansions, and are also expected from the conjecture that the critical indices should not be affected by the minor details of the interaction, and hence should be the same as the simple square lattice. Thus, we can assume that further approximations to the S matrix will not change the critical indices. In the conclusion in Sec. IV, we indicate what this approximation means in terms of counting closed polygons on the lattice.

II. THE PARTITION FUNCTION

In this section the partition function of the next-nearest-neighbor Ising lattice is cast into the field-theoretic formalism. In this lattice the jth spin interacts

[1] C. A. Hurst, J. Math. Phys. **7**, 305 (1966); R. W. Gibberd and C. A. Hurst, *ibid.* **8**, 1427 (1967).
[2] R. W. Gibberd, Phys. Rev. **171**, 563 (1968).
[3] N. W. Dalton and D. W. Wood, report of work prior to publication.
[4] See, for example, R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix* (Cambridge University Press, London, 1966).

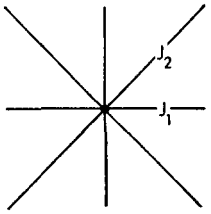


FIG. 1. Each spin interacts with eight neighbors with interaction strengths J_1 and J_2.

with its eight nearest neighbors as shown in Fig. 1. The horizontal and vertical interaction strength is denoted by J_1, and the diagonal interaction as J_2.

Following the combinatorial approach of Kac and Ward,[5] the partition function Z can be written as

$$Z = \sum_{r,s} g(r, s)x^r y^s,$$

where $x = \tanh \beta J_1$, $y = \tanh \beta J_2$, and $g(r, s)$ is the number of closed polygons that can be constructed from r horizontal and vertical bonds and s diagonal bonds on the lattice. Green and Hurst[6] have shown that the counting of closed polygons on a planar lattice is equivalent to the vacuum-to-vacuum expectation value of a product of Fermi operators. However, the above lattice is not planar because the diagonal bonds cross each other at a nonlattice point, and to overcome this it is first necessary to introduce an extra lattice point as shown in Fig. 2.

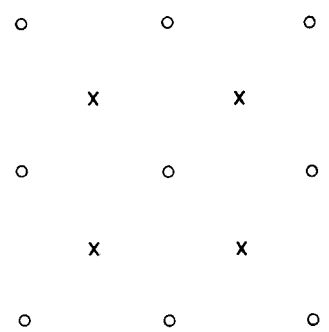


FIG. 2. The new lattice points are designated by x.

Now assigning fermion creation and annihilation operators according to the creation and annihilation of bonds on the lattice, where the particular ordering of the operators chosen is shown in Fig. 3, we can write the partition function as

$$Z = \langle 0| \prod_{j=1}^{N} (1 + \text{all possible products of operators corresponding to all possible vertices at the original } j\text{th lattice point}) \times (1 + a_j^{6*}a_{j-m+1}^{4} + a_j^{5*}a_{j-m}^{3} + a_j^{6*}a_j^{5*}a_{j-m+1}^{4}a_{j-m}^{3})\,|0\rangle, \quad (1)$$

where $|0\rangle$ is the vacuum state defined by $a_j\,|0\rangle = 0$.

The terms in the second bracket of the above expression represent the allowed vertices at the inserted lattice points, and correspond to those shown in Fig. 4. The other possible vertices at this lattice point are not included, since they would produce graphs which were not on the original lattice. Using the Fermi anticommutation rules, we can write Eq. (1) in the following form:

$$Z = \langle 0|\, T \exp\left[\sum_{j=1}^{N} H_0(j) + H_1(j)\right] |0\rangle, \quad (2)$$

where T is the usual time-ordering operator which orders the operators associated with site j from $j = 1$ to $j = N$. It can be shown that for this model

$$\begin{aligned} H_0(j) = {} & ya_j^{4*}(xa_j^{2*} + ya_j^{3*} + xa_j^{1*} \\ & + a_j^6 + a_{j-m}^2 + a_{j-1}^5 + a_{j-1}^1) \\ & + xa_j^{2*}(ya_j^{3*} + xa_j^{1*} + a_j^6 \\ & + a_{j-m}^2 + a_{j-1}^5 + a_{j-1}^1) \\ & + ya_j^{3*}(xa_j^{1*} + a_j^6 + a_{j-m}^2 + a_{j-1}^5 + a_{j-1}^1) \\ & + xa_j^{1*}(a_j^6 + a_{j-m}^2 + a_{j-1}^5 + a_{j-1}^1) \\ & + a_j^6(a_{j-m}^2 + a_{j-1}^5 + a_{j-1}^1) \\ & + a_{j-m}^2(a_{j-1}^5 + a_{j-1}^1) \\ & + a_{j-1}^5 a_{j-1}^1 + a_j^{5*}a_{j-m}^3 + a_j^{6*}a_{j-m+1}^4 \end{aligned}$$

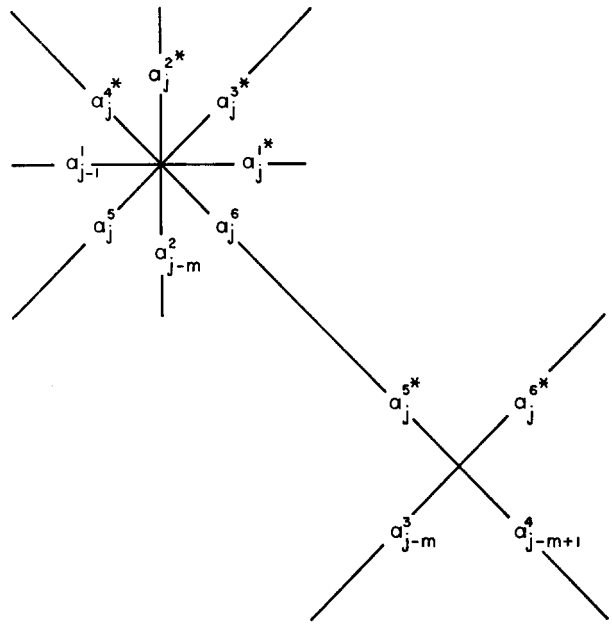


FIG. 3. The bonds and their associated Fermi operators for the original and extra lattice points.

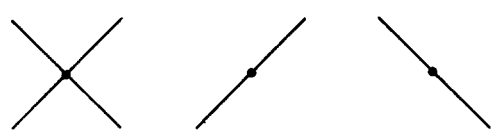


FIG. 4. The allowed vertices at the inserted lattice point.

[5] M. Kac and J. C. Ward, Phys. Rev. **88**, 1332 (1952).
[6] H. S. Green and C. A. Hurst, *Order-Disorder Phenomena* (Interscience Publishers, Inc., New York, 1964).

and

$$H_1(j) = 2a_j^{6*}a_j^{5*}a_{j-m+1}^4 a_{j-m}^3. \quad (3)$$

Equation (2) is an exact expression for the partition function, but it is also intractable because of the presence of the quartic term $H_1(j)$.

III. AN APPROXIMATE SOLUTION

Although the expression in Eq. (2) is not easily evaluated exactly, it is very amenable to the approximations used in quantum field theory. In particular, first- and higher-order Green's-function techniques have been developed[7] which allow one to make many different approximations. In this section we are going to make the simplest approximation possible, which, however, gives surprisingly good results. The approximation is to neglect completely the $H_1(j)$ term in Eq. (2) and to evaluate Z_1, which is given by

$$Z_1 = \langle 0|\, T \exp\left[\sum_{j=1}^{N} H0(j)\right] |0\rangle. \quad (4)$$

This expression can be evaluated using the techniques given in Ref. 1, where the calculation is straightforward though tedious. However, Green[8] has already done an equivalent calculation of Eq. (4) using the Pfaffian technique and so we here only present the results:

$$\frac{1}{N}\log Z_1^2 = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} d\theta\, d\phi \log D(\theta, \phi),$$

where

$$\begin{aligned} D(\theta, \phi) = {} & (1 + x^2)^2(1 + y^2)^2 + 16x^2y(1 - y^2) \\ & + 2x(1 - x^2)[4y^2 - (1 - y^2)^2] \\ & \times (\cos\theta + \cos\phi) - 2y(1 - x^2)(1 - y^2) \\ & \times [\cos(\theta + \phi) + \cos(\theta - \phi)]. \end{aligned} \quad (5)$$

The derivatives of this expression have been evaluated by Green and Hurst[6] in terms of the complete elliptic integrals of the first and third kinds. Using these results, the average energy $\bar{E}$, which is defined by

$$\bar{E} = -\frac{1}{N}\frac{\partial}{\partial\beta}\log Z,$$

can be written as

$$\bar{E}(T) = \bar{E}(T_c) + A(T - T_c)\log|T - T_c| + \cdots,$$

for T close to T_c. Thus the average energy per spin is continuous at the critical point, and the specific heat has a logarithmic divergence at $T = T_c$.

The magnetization can also be calculated for this approximation, and the equivalent calculation has been done by Green.[8] Using Green's Eq. (76), it can be seen that as $T \to T_c$, the magnetization goes to zero such that $\beta = \frac{1}{8}$.

[7] R. W. Gibberd, Can. J. Phys. **47**, 809 (1969).
[8] H. S. Green, Z. Phys. **171**, 129 (1963).

TABLE I. Values of J_1/kT_c for different values of $\alpha = J_2/J_1$: (a) results from this paper; (b) results of Dalton and Wood; (c) results of Fan and Wu.

α	(a)	(b)	(c)
0	0.4407	0.4407	0.4407
0.1	0.3879	0.3870	0.3864
0.2	0.3484	0.3451	0.3444
0.3	0.3174	0.3118	0.3109
0.4	0.2922	0.2849	0.2834
0.5	0.2711	0.2625	0.2605
0.6	0.2532	0.2436	0.2411
0.7	0.2377	0.2274	0.2244
0.8	0.2242	0.2134	0.2099
0.9	0.2123	0.2010	0.1972
1.0	0.2016	0.1902	0.1859
1.2	0.1834		0.1669
1.4	0.1684		0.1514
1.6	0.1557		0.1386
1.8	0.1449		0.1278
2.0	0.1356		0.1185

To obtain the critical temperature we use the following equations, which Hurst[9] has shown will determine the position of the singularity in $\log Z$:

$$D(\theta, \phi) = 0, \quad (6)$$

$$\frac{\partial D(\theta, \phi)}{\partial\theta} = \frac{\partial}{\partial\phi} D(\theta, \phi) = 0. \quad (7)$$

Taking the solutions of Eq. (7) as θ, ϕ equal to either 0 or π, then Eq. (6) becomes

$$1 - 2x - 2y - x^2 - y^2 - 4xy + 2xy^2 + 2x^2y + x^2y^2 = 0$$

or

$$1 + 2x - 2y - x^2 - y^2 + 4xy - 2xy^2 + 2x^2y + x^2y^2 = 0.$$

These equations were solved numerically for different values of α, where $\alpha = J_2/J_1$, and the results are shown in Table I and Fig. 5, where the critical temperature divided by the critical temperature of the square lattice is plotted against α. In a recent preprint Fan and Wu[10] have used a transformation due to Lieb, where the next-nearest-neighbor lattice is transformed onto a square lattice. Using techniques similar to those described here, they make an equivalent approximation and obtain the approximate critical temperatures shown. It is surprising to see how good both approximations are, and we conjecture that Fan and Wu's result is an upper bound and ours a lower bound to the exact result. We also note that for $\alpha = 0$ and in the limit $\alpha \to \pm\infty$, the approximation presented here gives the exact critical temperature, whereas the approximation of Fan and Wu is exact for $\alpha = 0$, but

[9] C. A. Hurst, J. Chem. Phys. **38**, 2558 (1963).
[10] C. Fan and F. Y. Wu, report of work prior to publication.

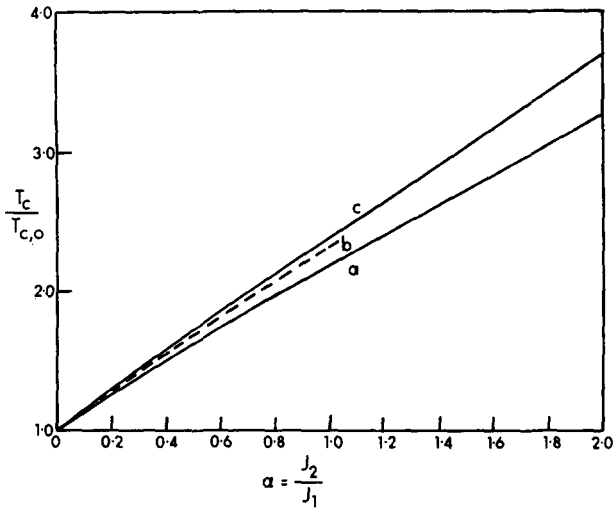


FIG. 5. Plot of values for critical temperature against different values for α. (See Table I.)

is too high in the case $\alpha = \infty$. However, our approximation does not give any exact results for the antiferromagnetic region.

IV. CONCLUSION

It can be shown that the first-order approximation which we have considered here counts all the closed polygons that can be drawn on the lattice, but those which contain an odd number of vertices of the crossed type shown in Fig. 4 are counted with a negative weight. It appears from the results obtained in this paper that the analytical behavior of the partition function is not greatly affected by the incorrect counting of some of these graphs. As already mentioned, there are, in other fields of physics, precedents for assuming that the partial summations of graphs will contain the correct analytical behavior. However, one would like to have a criterion which would indicate when the neglected terms do not contribute.

In conclusion, we mention that a similar S-matrix expression can be derived for the three-dimensional Ising models. However, the simple approximation of neglecting the perturbing Hamiltonian H_1 does not give such realistic results as the two-dimensional problems. This probably relates to the perturbing Hamiltonian containing in three dimensions a much larger number of quartic terms than in the two-dimensional cases. Thus, a more sophisticated approximation is required.

Note Added in Proof: The author has recently shown that such first-order approximations as considered here give rigorous lower (or upper) bounds to the critical temperature when based on a high- (or low-) temperature expansion. Hence the conjecture mentioned in Sec. III is correct.

ACKNOWLEDGMENTS

The author wishes to acknowledge fruitful discussions with Professor C. A. Hurst, during whose supervision much of this work was completed. He also wishes to thank Dr. F. Y. Wu, for discussions and the preprint of his work, and Professor D. D. Betts.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Geometrization of the Brans–Dicke Scalar Field*

P. C. PETERS
University of Washington, Seattle, Washington

(Received 16 October 1968)

The Brans–Dicke gravitational scalar field is geometrized in the spirit of the Rainich–Misner–Wheeler geometrization of electromagnetism. Geometric equations are derived which imply that the Brans–Dicke field is present and an explicit expression is given for this field in terms of geometrical quantities.

The general theory of relativity geometrizes the gravitational field in the sense that the properties of the gravitational interaction are described in terms of the geometry of space–time rather than as an independent field. The appeal of this approach led to attempts to geometrize other fields, notably electromagnetism. However, the fact that the charge-to-mass ratio of particles varies in nature meant that charged particles do not follow geodesics, and one can therefore distinguish electromagnetic effects from geometrical effects. The program of geometrodynamics[1]

* Work supported in part by the National Science Foundation.

[1] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962).

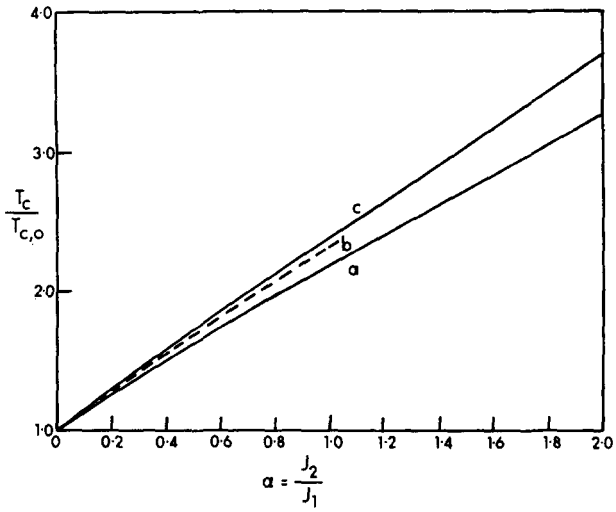


FIG. 5. Plot of values for critical temperature against different values for α. (See Table I.)

is too high in the case $\alpha = \infty$. However, our approximation does not give any exact results for the antiferromagnetic region.

IV. CONCLUSION

It can be shown that the first-order approximation which we have considered here counts all the closed polygons that can be drawn on the lattice, but those which contain an odd number of vertices of the crossed type shown in Fig. 4 are counted with a negative weight. It appears from the results obtained in this paper that the analytical behavior of the partition function is not greatly affected by the incorrect counting of some of these graphs. As already mentioned, there are, in other fields of physics, precedents for assuming that the partial summations of graphs will contain the correct analytical behavior. However, one would like to have a criterion which would indicate when the neglected terms do not contribute.

In conclusion, we mention that a similar S-matrix expression can be derived for the three-dimensional Ising models. However, the simple approximation of neglecting the perturbing Hamiltonian H_1 does not give such realistic results as the two-dimensional problems. This probably relates to the perturbing Hamiltonian containing in three dimensions a much larger number of quartic terms than in the two-dimensional cases. Thus, a more sophisticated approximation is required.

Note Added in Proof: The author has recently shown that such first-order approximations as considered here give rigorous lower (or upper) bounds to the critical temperature when based on a high- (or low-) temperature expansion. Hence the conjecture mentioned in Sec. III is correct.

ACKNOWLEDGMENTS

The author wishes to acknowledge fruitful discussions with Professor C. A. Hurst, during whose supervision much of this work was completed. He also wishes to thank Dr. F. Y. Wu, for discussions and the preprint of his work, and Professor D. D. Betts.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Geometrization of the Brans–Dicke Scalar Field*

P. C. PETERS
University of Washington, Seattle, Washington

(Received 16 October 1968)

The Brans–Dicke gravitational scalar field is geometrized in the spirit of the Rainich–Misner–Wheeler geometrization of electromagnetism. Geometric equations are derived which imply that the Brans–Dicke field is present and an explicit expression is given for this field in terms of geometrical quantities.

The general theory of relativity geometrizes the gravitational field in the sense that the properties of the gravitational interaction are described in terms of the geometry of space–time rather than as an independent field. The appeal of this approach led to attempts to geometrize other fields, notably electromagnetism. However, the fact that the charge-to-mass ratio of particles varies in nature meant that charged particles do not follow geodesics, and one can therefore distinguish electromagnetic effects from geometrical effects. The program of geometrodynamics[1]

* Work supported in part by the National Science Foundation.

[1] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962).

takes the point of view that what we call charge is only a gross feature ascribed to some geometry, in the same manner as we ascribe a mass to the Schwarzschild solution, even though there may be no real source present. Thus geometrodynamics restricts our considerations to source-free regions of space–time. The geometrization of source-free electromagnetism was first carried out by Rainich[2] and was later elaborated by Misner and Wheeler.[3] The important feature of their approach to geometrization is that the field considered, e.g., electromagnetism, leaves such a characteristic imprint on the geometry that one can determine what field is present from a knowledge of the geometry. Geometric equations are found which guarantee that only electromagnetism is present and one finds a prescription for extracting the field from the metric. The program has been extended to other fields, e.g., the massless scalar field[4] and neutrino field.[5]

The program of geometrodynamics rests on the foundations of general relativity. Recently, however, observations of the solar oblateness by Dicke and Goldenberg[6] have cast doubt as to the validity of general relativity. An alternate theory of gravitation, the Brans–Dicke[7] or scalar–tensor theory, can be made to agree with the consequences of the solar oblateness measurement by a suitable choice of coupling constant. This theory is no longer purely geometric, however, since it requires the introduction of a real scalar field φ in addition to the metric.[8] In this paper we show that this Brans–Dicke scalar field can be geometrized in the spirit of geometrodynamics.

In a source-free region, in standard units, the Brans–Dicke equations can be derived from the requirement that the variations of the action

$$A = \int [\varphi R + \omega(\varphi_{;\mu}\varphi^{;\mu}/\varphi)](-g)^{\frac{1}{2}}\, d^4x \tag{1}$$

with respect to φ and with respect to $g_{\mu\nu}$ vanish. This leads to the field equation for φ (where we have defined $\Psi = -\ln \varphi$, which is real since φ is positive):

$$\Psi^{;\lambda}_{;\lambda} - \tfrac{1}{2}\Psi_{;\lambda}\Psi^{;\lambda} + (2\omega)^{-1}R = 0 \tag{2}$$

and the gravitational field equations

$$R_{\mu\nu} = \Psi_{;\mu;\nu} - (1+\omega)\Psi_{;\mu}\Psi_{;\nu} + \tfrac{1}{2}g_{\mu\nu}(\Psi^{;\lambda}_{;\lambda} - \Psi_{;\lambda}\Psi^{;\lambda}) \tag{3}$$

whose trace is

$$R = 3\Psi^{;\lambda}_{;\lambda} - (\omega+3)\Psi_{;\lambda}\Psi^{;\lambda}. \tag{4}$$

The identity $(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R)^{;\nu} = 0$ leads, from (3), to the requirement that Ψ satisfy the equation

$$(\omega + \tfrac{3}{2})\Psi_{;\mu}[\Psi^{;\lambda}_{;\lambda} - \Psi_{;\lambda}\Psi^{;\lambda}] = 0. \tag{5}$$

Thus for $\omega \neq -\frac{3}{2}$, the quantity in the brackets in (5) must vanish, and then the field equation for Ψ, Eq. (2), can be derived from (4) and (5). For $\omega = -\frac{3}{2}$, the field equation for Ψ, Eq. (2), is identical to the trace of the gravitational field equations (4). Therefore, in any case, the field equation for Ψ, Eq. (2), can be derived from the gravitational field equations (3) and we need consider only the geometric form of Eq. (3).

Assuming that the field equations (3) are satisfied, it is possible to solve explicitly for $\Psi_{;\mu}$ in terms of geometrical quantities. To this end we compute $C^{\lambda\mu\nu\alpha}R_{\mu\nu;\alpha}$, using $R_{\mu\nu}$ from (3), where $C_{\alpha\beta\gamma\delta}$ is the Weyl conformal tensor defined by

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + g_{\alpha[\gamma}R_{\delta]\beta} + g_{\beta[\delta}R_{\gamma]\alpha} - \tfrac{1}{3}g_{\alpha[\gamma}g_{\delta]\beta}R, \tag{6}$$

where $[\alpha\beta] = \frac{1}{2}(\alpha\beta - \beta\alpha)$. Making use of the symmetry and trace properties of (6) we then have that

$$C^{\lambda\mu\nu\alpha}R_{\mu\nu;\alpha} = \tfrac{1}{2}C^{\lambda\mu\nu\alpha}R^{\sigma}{}_{\mu\nu\alpha}\Psi_{;\sigma} + (1+\omega)C^{\lambda\mu\nu\alpha}R_{\mu\nu}\Psi_{;\alpha}. \tag{7}$$

We now express the Riemann tensor in (7) in terms of the conformal tensor and Ricci tensor using (6). This allows us to make use of the identity

$$C^{\mu\alpha\beta\gamma}C_{\nu\alpha\beta\gamma} = \tfrac{1}{4}\delta^{\mu}_{\nu}C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} \tag{8}$$

to write Eq. (7) in the form

$$C^{\lambda\mu\nu\alpha}R_{\mu\nu;\alpha} = \tfrac{1}{8}\Psi^{;\lambda}C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} + (\omega+\tfrac{3}{2})C^{\lambda\mu\nu\alpha}R_{\mu\nu}\Psi_{;\alpha}. \tag{9}$$

If $\omega = -\frac{3}{2}$, the last term in (9) vanishes and, dividing by $C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta}$, we find an explicit expression for $\Psi_{;\mu}$ in terms of geometrical quantities.

The last term of (9) may be reduced by observing that, in the coefficient of $(\omega + \frac{3}{2})$, we may assume that $\omega \neq -\frac{3}{2}$. In this case Eq. (5), together with (3) and (4), gives

$$R_{\mu\nu} = \Psi_{;\mu;\nu} - (1+\omega)\Psi_{;\mu}\Psi_{;\nu}, \tag{10}$$

$$\Psi^{;\lambda}_{;\lambda} = \Psi_{;\lambda}\Psi^{;\lambda} = -R/\omega. \tag{11}$$

Computing $R_{\mu[\nu;\alpha]}$ from (10) yields

$$R^{\sigma}{}_{\mu\nu\alpha}\Psi_{;\sigma} = 2R_{\mu[\nu;\alpha]} - 2(1+\omega)R_{\mu[\nu}\Psi_{;\alpha]}, \tag{12}$$

where we have used $\Psi_{;\mu;[\alpha;\beta]} = \frac{1}{2}R^{\sigma}{}_{\mu\alpha\beta}\Psi_{;\sigma}$ and also have used (10) to eliminate terms involving second

[2] G. Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[3] C. W. Misner and J. A. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1957).
[4] D. R. Brill, Nuovo Cimento Suppl. **2**, 1 (1964).
[5] R. Penney, J. Math. Phys. **6**, 1309 (1965).
[6] R. H. Dicke and H. M. Goldenberg, Phys. Rev. Letters **18**, 313 (1967).
[7] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[8] The theory retains the postulate that particles follow geodesics, so one cannot observe φ directly from the equations of motion. However, one can perform a Cavendish experiment to measure G and thus φ.

derivatives of Ψ. Multiplying (12) by $g^{\mu\nu}$ gives, after rearrangement,

$$R^{\sigma}_{\alpha}\Psi_{;\sigma} = -\frac{1}{2\omega}R_{;\alpha} + \frac{(1+\omega)}{\omega}\Psi_{;\alpha}\,, \qquad (13)$$

which can also be found by multiplying (10) by $\Psi^{;\nu}$ and using (11) to eliminate $\Psi_{;\lambda}\Psi^{;\lambda}$. Equations (12) and (13) allow us to reduce the expression in the last term of (9) to a form where $\Psi^{;\lambda}$ only appears with a free index, which then allows us to explicitly solve for $\Psi^{;\lambda}$.

From (6), we can write

$$C^{\lambda\mu\nu\alpha}R_{\mu\nu}\Psi_{;\alpha} = R^{\lambda\mu\nu\alpha}R_{\mu\nu}\Psi_{;\alpha} + \Psi^{;\lambda}[\tfrac{1}{6}R^2 - \tfrac{1}{2}R_{\alpha\beta}R^{\alpha\beta}] + R^{\lambda}_{\mu}R^{\mu\alpha}\Psi_{;\alpha} - \tfrac{2}{3}RR^{\lambda\alpha}\Psi_{;\alpha}. \qquad (14)$$

Equation (12) can be used to eliminate the $R^{\lambda\mu\nu\alpha}\Psi_{;\alpha}$ and Eq. (13) can be used to eliminate all terms in $R^{\mu\alpha}\Psi_{;\alpha}$. The result of this is that

$$C^{\lambda\mu\nu\alpha}R_{\mu\nu}\Psi_{;\alpha} = (\omega+\tfrac{3}{2})\Psi^{;\lambda} \times\left[\frac{\omega^2+2\omega+\frac{4}{3}}{\omega^2}R^2 - R_{\mu\nu}R^{\mu\nu}\right] + R^{\mu\nu}[R_{\mu\nu}{}^{;\lambda} - R^{\lambda}_{\mu;\nu}] - \left(\frac{\omega+2}{2\omega}\right)R^{\lambda}_{\mu}R^{;\mu} - \left(\frac{\omega^2+\frac{7}{3}\omega+2}{2\omega}\right)R^{;\lambda}R. \qquad (15)$$

Substituting (15) into (9), we find that $\Psi_{;\lambda} = \alpha_{\lambda}$, where α_{λ} is defined to be the geometrical quantity

$$\alpha_{\lambda} = \left\{C_{\lambda}{}^{\mu\nu\alpha}R_{\mu\nu;\alpha} + (\omega+\tfrac{3}{2})\left[\frac{\omega^2+\frac{7}{3}\omega+2}{2\omega}R_{;\lambda}R + \left(\frac{\omega+2}{2\omega}\right)R_{\lambda\mu}R^{;\mu} + 2R^{\mu\nu}R_{\mu[\lambda;\nu]}\right]\right\}\Big/\left\{\tfrac{1}{8}C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} + (\omega+\tfrac{3}{2})\left[\frac{\omega^2+2\omega+\frac{4}{3}}{\omega^2}R^2 - R_{\mu\nu}R^{\mu\nu}\right]\right\}. \qquad (16)$$

We can now state the geometrical equations which imply that the Brans–Dicke field is present in terms of the geometrical quantity α_{μ}. The geometrical equations to be satisfied are, from Eq. (3),

$$R_{\mu\nu} = \tfrac{1}{2}(\alpha_{\mu;\nu} + \alpha_{\nu;\mu}) - (1+\omega)\alpha_{\mu}\alpha_{\nu} + \tfrac{1}{2}g_{\mu\nu}(\alpha^{;\mu}_{\mu} - \alpha_{\mu}\alpha^{\mu}), \qquad (17)$$

where α_{μ} is defined in (16) and satisfies

$$\alpha_{[\mu;\nu]} = 0. \qquad (18)$$

Then Ψ is obtained in the same manner as the complexion of the electromagnetic field[3]

$$\Psi = \int \alpha_{\mu}\, dx^{\mu}. \qquad (19)$$

Since $\varphi = -\ln\Psi$, we have that the Brans–Dicke field φ is obtained from the geometrical vector α_{μ} as

$$\varphi(x) = \varphi_0 \exp\left(-\int_0^x \alpha_{\mu}\, dx^{\mu}\right), \qquad (20)$$

where φ_0 is the value of φ at some initial point. It is clear that one could not hope to obtain the value of φ_0, since the field equations (2) and (3) are invariant under a constant change of scale of φ. Also one should note that in the case that the denominator of (16) vanishes, α_{μ} is not defined. Therefore, we restrict ourselves to regions where the denominator is nonzero, the analog of excluding null fields by Misner and Wheeler.[3]

Although the observations seem to imply a value of $\omega \approx 6$, the case $\omega = -\frac{3}{2}$ is of some interest in a different context. If we relax the Brans–Dicke assumption that particles follow geodesics, then in a source-free region the equations for $\omega = -\frac{3}{2}$ are just the geometric equations of the conformal scalar field.[9] In this case a conformal transformation of the form $g_{\mu\nu} \to g_{\mu\nu}/\varphi$, with φ given by (20), reduces the field equations in the transformed space to $R_{\mu\nu} = 0$. Thus, our geometrical equations for $\omega = -\frac{3}{2}$ are the conditions that a geometry be conformally related to a geometry which is a solution of the source-free equations of general relativity. From (16) we see that this information can be extracted if $C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} \neq 0$.

[9] P. C. Peters, Phys. Letters **20**, 641 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Expansion of the Density Matrix of an N-Fermion System in Terms of the Correlation Densities of Fermion Clusters

TAI-ICHI SHIBUYA AND OKTAY SINANOĞLU
Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

(Received 8 October 1967; Revised Manuscript Received 30 October 1968)

The N-fermion density operator $D = |\Psi\rangle\langle\Psi|$ is decomposed into the densities of clusters of correlating fermions. The m-body cluster densities $\hat{\gamma}^{(m)}$ are "orthogonal" to the one-body "Fock–Dirac"-type densities ρ_1, i.e., $\mathrm{Tr}_1\,(\rho_1\hat{\gamma}^{(m)}_{1,2,\cdots,m}) = 0$. The reduced-correlation-density matrices obtained differ from the conventional reduced-density-matrices and are particularly convenient for the treatment of fermion–fermion correlations.

INTRODUCTION

In finite many-fermion problems, it is convenient to treat an independent-particle model, orbital wavefunction (w.f.) part, and the remaining fermion–fermion correlations part separately. In the particular case of the finite many-electron problem, electron correlation in atoms and molecules, most traditional methods have worked with the total wavefunction without separating the orbital and correlation parts. In the customary configuration–interaction (C.I.) method, if one started with the Hartree–Fock self-consistent field (SCF) wavefunction, added configurations introduced some of the correlations, but here too, as in other methods, the N-fermion system was studied as a whole, with the result that difficulty of the problem rapidly increased with N. Different types of correlations involving fewer electrons were not calculated and studied separately as subsystems and with respect to their relative importance in building up the total correlation.

The density matrix (density operator) in the form $|\Psi\rangle\langle\Psi|$ does not exhibit the orbital (model) theory versus the (orthogonal) correlation subspaces separated. The reduced density matrices too, which have been studied by several authors, do not contain the model versus the correlation effects separately, nor are the different types of correlation effects separated from one another.

Over the past few years, an approach developed by Sinanoğlu[1] and co-workers for correlation in N-electron systems has been applied in various atomic and molecular structure problems. Here, after separating the orbital theory and the correlation parts of the total wavefunction (and of the N-electron Hilbert space) into mutually orthogonal parts, the correlation part is successively decomposed into subspaces involving one, two, three, $\cdots$ electron-correlations at a time. Methods for examining or evaluating these different correlation parts separately are given. It is shown, for example, that in closed-shell systems, the N-electron problem reduces to good approximation into $N(N-1)/2$ "heliumlike" two-electron problems.[1a,b,e] In excited states, nonclosed shells, additional correlation effects arise and have recently been evaluated.[2] The latter were used to obtain such atomic properties as electron affinities, negative-ion excitation energies, and oscillator strengths.

It was noted[1b,e] previously that the closed-form methods used in the above approach could be useful also for finite nuclei, but they were not applied in that context.

For some purposes, it is convenient to have formulations directly in terms of the density (matrix) operator. For example, in molecular orbital (MO) theory, one deals with electron populations in different atomic orbitals which in linear combination yield the LCAO MO's. A density (matrix) operator formulation is compact and general in such cases, making transformations from one basis set to another particularly easy.

The present paper gives a new type of reduced-density (matrices) operators suitable for N-fermion correlation problems. The exact N-body-density (matrix) operator of an N-fermion system is decomposed into successive 1, 2, 3, $\cdots$, n, $\cdots$ N-fermion-correlation densities. The forms of the exact density and energy and their relation to the correlation functions of the previous theory are obtained. The m-fermion-correlation density operators have simple properties under the unitary transformations which transform one type of atomic or molecular orbitals into another (e.g., MO's into localized, chemical-bond-like orbitals in molecules).

[1] (a) O. Sinanoğlu, Proc. Roy. Soc. (London) **A260**, 379 (1961); (b) J. Chem. Phys. **36**, 706 (1962); (c) O. Sinanoğlu and D. F. Tuan, *ibid.* **41**, 2677 (1964); (d) O. Sinanoğlu and V. McKoy, *ibid.* **41**, 2689 (1964); (e) O. Sinanoğlu, Advan. Chem. Phys. **6**, 315 (1964).

[2] O. Sinanoğlu and İ. Öksüz, Phys. Rev. Letters **21**, 507 (1968). See also H. J. Silverstone and O. Sinanoğlu, J. Chem. Phys. **44**, 1899, 3608 (1966).

FORM OF THE EXACT WAVEFUNCTION

For a single determinantal state,

$$\phi_0 = \mathcal{A}_N(123\cdots i\cdots N),$$

where $i = i(\mathbf{x}_1)$ are the N occupied spin orbitals, $\mathbf{x}_i$ space, spin (and isotopic spin) variables, and $\mathcal{A}_N$ the antisymmetrizer, it has been shown previously that the exact wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)$ of an N-fermion system can be written in the form[1e]

$$\Psi = \phi_0 + \chi,$$
$$\langle\phi_0 \,|\, \chi\rangle = 0, \quad \langle\Psi \,|\, \Psi\rangle > 1, \quad \langle\phi_0 \,|\, \phi_0\rangle = 1,$$
$$\Psi = \phi_0 + \sum_{i=1}^{N} \{\hat{f}_i\} + \sum_{j>i=1}^{N} \{\hat{U}'_{ij}\} + \sum_{k>j>i=1}^{N} \{\hat{U}'_{ijk}\} + \cdots + \{\hat{U}'_{123\cdots N}\}, \quad (1)$$

where, e.g.,

$$\{\hat{U}'_{ijk}\} = \frac{1}{(3!)^{\frac{1}{2}}} \mathcal{A}_N\left\{\frac{123\cdots N}{ijk}\, \hat{U}'_{ijk}\right\}.$$

The $\hat{U}$'s are closed-form "cluster" functions with the properties

$$\hat{U}'_{ij}(\mathbf{x}_i, \mathbf{x}_j) = -\hat{U}'_{ij}(\mathbf{x}_j, \mathbf{x}_i),$$
$$\hat{U}'_{ijk}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) = -\hat{U}'_{ijk}(\mathbf{x}_j, \mathbf{x}_i, \mathbf{x}_k), \quad (2)$$

and so on, and

$$\langle\hat{f}_i \,|\, k\rangle = 0,$$
$$\langle\hat{U}'_{ij} \,|\, k\rangle = 0, \quad (3)$$
$$\langle\hat{U}'_{ijk} \,|\, k\rangle = 0,$$

where $k = 1, 2, 3, \cdots, N$ are the occupied spin-orbitals in the "model" or "orbital theory" wavefunction (w.f.) ϕ_0. Note that, e.g.,

$$\langle\hat{U}'_{ij} \,|\, k\rangle \equiv \int \hat{U}'^{*}_{ij}(\mathbf{x}_i, \mathbf{x}_j) k(\mathbf{x}_i)\, dx_i.$$

In N-electron systems, a very convenient ϕ_0 is the Hartree–Fock (H.F.) one. In finite nuclei, ϕ_0 may be based on some "generalized H.F." method. For the formal development that will follow, however, ϕ_0 needs to be only some $\mathcal{A}_N(123\cdots N)$, with N orthonormal spin orbitals based on any orbital theory. If ϕ_0 is some type of a H.F. w.f., the χ may then be defined as the fermion–fermion correlation wavefunction. The above form, Eqs. (1–3), has turned out very convenient in the treatment of electron correlation in atoms and molecules.[3] The successive terms in Eq. (1) correspond to correlation between successively larger number of electrons in the H.F. (or ϕ_0) "sea."

The form of the Ψ above is derived most directly by successively Schmidt-orthogonalizing the Ψ to the products of $N, N-1, N-2, \cdots, 1$ H.F. spin-orbitals at a time (the "method of successive partial orthogonalizations").[4] The same derivation that gives Eq. (1) gives also the important and rigorous "orbital orthogonality" condition of the cluster functions, i.e., Eqs. (3).

We now derive the form of the exact N-body density of an N-fermion system in terms of correlation density clusters by a generalization of the method of "successive partial orthogonalizations."[4] The density operators obtained differ from the various types of reduced-density matrices that have been used previously by other authors.[5,6] We shall give the derivation for single determinantal states only.

THE DENSITY OPERATORS

Suppose that f and g are wavefunctions of n fermions, $1, 2, \cdots, n$. Then, for the nth-order density matrix $|f\rangle\langle g|$ of those w.f.'s, we define the traces of fermions, $1, 2, \cdots, m$, by

$$\begin{aligned}\mathrm{Tr}_{1,2,\cdots,m}(|f\rangle\langle g|) &\equiv \mathrm{Tr}_{\neq m+1,m+2,\cdots,n}(|f\rangle\langle g|)\\ &\equiv \int dx_1 \int dx_2 \cdots \int dx_m \\ &\quad \times \langle \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m \,|f\rangle\langle g|\, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m\rangle\\ &\equiv \langle g \,|f\rangle_{1,2,\cdots,m} \equiv \langle g \,|f\rangle_{\neq m+1,m+2,\cdots,n} \quad (4)\end{aligned}$$

and thus obtain an $(n-m)$th-order reduced-density matrix. The symbol "Tr" without subindices will be used to stand for the trace over all the fermions in the density matrix. Thus,

$$\mathrm{Tr}\,(|f\rangle\langle g|) \equiv \mathrm{Tr}_{1,2,\cdots,n}(|f\rangle\langle g|) = \langle g \,|f\rangle$$

which is just a number. We also have the relations

$$\mathrm{Tr}_{1,2,\cdots,m}\left(\sum_{f,g} |f\rangle\langle g|\right) = \sum_{f,g} \mathrm{Tr}_{1,2,\cdots,m}(|f\rangle\langle g|), \quad (5)$$

$$\begin{aligned}\mathrm{Tr}\,(\mathcal{O}\,|f\rangle\langle g|) &= \langle g|\,\mathcal{O}\,|f\rangle \equiv \langle g \,|\, (\mathcal{O}f)\rangle,\\ \mathrm{Tr}\,(|f\rangle\langle g|\,\mathcal{O}^{\dagger}) &= \langle g|\,\mathcal{O}^{\dagger}\,|f\rangle \equiv \langle (\mathcal{O}g) \,|\, f\rangle,\end{aligned} \quad (6)$$

where $\mathcal{O}$ can be any linear operator and $\mathcal{O}^{\dagger}$ its adjoint.

[3] See Refs. 1 and 2. Other references can be found through Refs. 1 and 2.

[4] O. Sinanoğlu, Rev. Mod. Phys. **35**, 517 (1963). Discussed also in Ref. 1(e).

[5] (a) K. Husimi, Proc. Phys.-Math. Soc. Japan **22**, 264 (1940); (b) P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955); Advan. Chem. Phys. **2**, 207 (1959); (c) R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960); (d) D. ter Haar, Rept. Progr. Phys. **24**, 304 (1961); (e) B. C. Carlson and J. M. Keller, Phys. Rev. **121**, 659 (1961); (f) S. Cho, Science Rept. Gunma University (Japan) **9**, No. 5 (1961); **11**, No. 3 (1962); (g) A. J. Coleman, Can. Math. Bull. **4**, 209 (1961); Rev. Mod. Phys. **35**, 668 (1963); (h) T. Ando, *ibid.* **35**, 690 (1963); (i) F. Weinhold and E. B. Wilson, Jr., J. Chem. Phys. **46**, 2752; **47**, 2298 (1967); (j) A. J. Coleman and R. M. Erdahl, Eds., *Reduced Density Matrices with Applications to Physical and Chemical Systems* (Queen's University, Kingston, Ontario, Canada, 1968).

[6] A. Primas, in *Modern Quantum Chemistry—Istanbul Lectures* O. Sinanoğlu, Ed. (Academic Press Inc., New York, 1965).

Let us introduce the density matrices D and D_0, respectively corresponding to the w.f.'s Ψ and ϕ_0, as follows:

$$D \equiv |\Psi\rangle\langle\Psi|, \quad D_0 \equiv |\phi_0\rangle\langle\phi_0|. \tag{7}$$

By the constraints in Eq. (1),

$$\mathrm{Tr}\,(D_0) = \mathrm{Tr}\,(D_0 D_0) = \mathrm{Tr}\,(D_0 D) = 1, \tag{8}$$

but $\mathrm{Tr}\,(D) \neq 1$ when $N \geq 2$. We then define the mth-order density matrices $\rho^{(m)}$ by

$$\rho^{(m)}_{1,2,\cdots,m} \equiv m!\binom{N}{m} \mathrm{Tr}_{\neq 1,2,\cdots,m}\,(D_0). \tag{9}$$

Substitution of

$$|\phi_0\rangle = \mathcal{A}_N\,|123\cdots k\cdots N\rangle \tag{10}$$

with the orthonormality conditions

$$\langle k \mid l\rangle = \delta_{k,l}, \text{ for any } k, l \text{ in the set } \{1, 2, \cdots, N\}, \tag{11}$$

into D_0 of Eq. (9) gives

$$\begin{aligned}
&\rho^{(m)}_{1,2,\cdots,m} \\
&\quad = m! \sum_{k_m\cdots k_2 k_1=1}^{N} \mathcal{A}_m\,|k_1 k_2\cdots k_m\rangle\langle k_1 k_2\cdots k_m|\,\mathcal{A}_m^\dagger) \\
&\quad = \sum_{k_1,k_2,\cdots,k_m=1}^{N} \mathcal{A}_m\,|k_1 k_2\cdots k_m\rangle\langle k_1 k_2\cdots k_m|\,\mathcal{A}_m^\dagger \\
&\quad = \mathcal{A}_m(\rho_1\rho_2\cdots\rho_m)\mathcal{A}_m^\dagger,
\end{aligned} \tag{12}$$

where

$$\rho \equiv \rho^{(1)} = \sum_{k=1}^{N} |k\rangle\langle k|. \tag{13}$$

In this derivation, it should be noted that

$$\mathcal{A}_m \equiv \sum_{P\in S_m} \delta_P P, \tag{14}$$

where $\delta_P = +1$ or -1, respectively, for the even or odd permutations P. If ϕ_0 is H.F., then $\rho^{(m)}$ is the mth-order Fock–Dirac density matrix, otherwise an analog of it for other "model" orbitals. The term "Fock–Dirac density" is used below in this more general sense. We have

$$\mathrm{Tr}\,(\rho^{(m)}) = m!\binom{N}{m}. \tag{15}$$

It can be easily shown that

$$D_0 = \frac{1}{N!}\,\rho^{(N)} = \frac{1}{N!}\,\mathcal{A}_N(\rho_1\rho_2\cdots\rho_N)\mathcal{A}_N^\dagger. \tag{16}$$

ANALYSIS OF THE DENSITY (MATRIX) OPERATOR BY THE METHOD OF "SUCCESSIVE PARTIAL ORTHOGONALIZATIONS"

Suppose that D_χ, a portion of D, is "orthogonal" to D_0 in a sense that

$$\mathrm{Tr}\,(D_0 D_\chi) = 0. \tag{17}$$

The way of finding such a portion D_χ is similar to the Schmidt orthogonalization method. This orthogonality condition is satisfied if

$$\begin{aligned}
D_\chi &= D - D_0\,\mathrm{Tr}\,(D_0 D) = D - D_0 \\
&= |\phi_0\rangle\langle\chi| + |\chi\rangle\langle\phi_0| + |\chi\rangle\langle\chi|.
\end{aligned} \tag{18}$$

Thus, D_χ is the whole portion left over in D after D_0 is taken out and, therefore, may be called the "total correlation density (matrix)" of the N-fermion system.

We next find $D_{\chi'}$, a portion of D_χ, "orthogonal" to products of $(N-1)$ ρ's in the sense

$$\mathrm{Tr}_{\neq N}\,(\rho_1\rho_2\cdots\rho_{N-1}D_{\chi'}) = 0. \tag{19}$$

Suppose that this condition is satisfied if

$$D_{\chi'} = D_\chi - D_\gamma(1), \tag{20}$$

where

$$D_\gamma(1) \equiv \left[\binom{N}{1}\Big/1!\,N!\right]\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-1}\hat{\gamma}_N(1))\mathcal{A}_N^\dagger, \tag{21}$$

$$\hat{\gamma}_N(1) = C'\,\mathrm{Tr}_{\neq N}\,(\rho_1\rho_2\cdots\rho_{N-1}D_\chi). \tag{22}$$

Requiring that $\hat{\gamma}^{(1)}$ has the same permutation symmetry as $\rho^{(1)}$, we can easily show that

$$\mathrm{Tr}_{\neq N}\,[\rho_1\rho_2\cdots\rho_{N-1}D_\gamma(1)] = \left[1/1!\binom{N}{1}\right]\hat{\gamma}_N^{(1)}. \tag{23}$$

Thus, substituting Eq. (20) into Eq. (19) and combining Eqs. (22) and (23), we see that the condition (19) is satisfied if

$$C' = 1!\binom{N}{1}. \tag{24}$$

Then, Eq. (20) can be written as

$$\begin{aligned}
D_{\chi'} = D_\chi &- \left[\binom{N}{1}^2\Big/N!\right] \\
&\times \mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-1}[\mathrm{Tr}_{\neq N}\,(\rho_1\rho_2\cdots\rho_{N-1}D_\chi)]\mathcal{A}_N^\dagger.
\end{aligned} \tag{25}$$

We continue with $D_{\chi''}$, a portion of $D_{\chi'}$, "orthogonal" to products of $(N-2)$ ρ's in the sense

$$\mathrm{Tr}_{\neq N-1,N}\,(\rho_1\rho_2\cdots\rho_{N-2}D_{\chi''}) = 0. \tag{26}$$

Suppose that this condition is satisfied if

$$D_{\chi''} = D_{\chi'} - D_\gamma(2), \tag{27}$$

where

$$D_\gamma(2) = \left[\binom{N}{2}\Big/2!\,N!\right]\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-2}\hat{\gamma}^{(2)}_{N-1,N})\mathcal{A}_N^\dagger, \tag{28}$$

$$\hat{\gamma}^{(2)}_{N-1,N} = C''\,\mathrm{Tr}_{\neq N-1,N}\,(\rho_1\rho_2\cdots\rho_{N-2}D_{\chi'}). \tag{29}$$

Requiring that $\hat{\gamma}^{(2)}$ has the same permutation symmetry as $\rho^{(2)}$, we can show that

$$\mathrm{Tr}_{\neq N-1,N}(\rho_1\rho_2\cdots\rho_{N-2}D_\gamma(2)) = \left[1/2!\binom{N}{2}\right]\hat{\gamma}^{(2)}_{N-1,N}. \tag{30}$$

Thus, substituting Eq. (27) into Eq. (26) and combining Eqs. (29) and (30), we see that the condition (26) is satisfied if

$$C'' = 2!\binom{N}{2}. \tag{31}$$

Then, Eq. (27) can be written as

$$D_{\chi''} = D_{\chi'} - \left[\binom{N}{2}^2\Big/N!\right]\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-2} \times [\mathrm{Tr}_{\neq N-1,N}(\rho_1\rho_2\cdots\rho_{N-2}D_{\chi'})])\mathcal{A}_N^\dagger. \tag{32}$$

Note that $D_{\chi'}$ was already found in Eq. (25).

Similarly, in general, we can show that $D_\chi^{(m)}$, [where (m) as a superscript to this term indicates m primes] a portion of $D_\chi^{(m-1)}$,

$$D_\chi^{(m)} = D_\chi^{(m-1)} - D_\gamma(m) \tag{33}$$

is "orthogonal" to products of $(N-m)$ ρ's in the sense

$$\mathrm{Tr}_{\neq N-m+1,\cdots,N}(\rho_1\rho_2\cdots\rho_{N-m}D_\chi^{(m)}) = 0, \tag{34}$$

if

$$D_\gamma(m) = \left[\binom{N}{m}\Big/m!\,N!\right] \times \mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-m}\hat{\gamma}^{(m)}_{N-m+1,\cdots,N})\mathcal{A}_N^\dagger. \tag{35}$$

$D_\gamma(m)$ is to be compared with

$$D_0 = \frac{1}{m!\,N!}\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-m}\rho^{(m)}_{N-m+1,\cdots,N})\mathcal{A}_N^\dagger;$$

and

$$\hat{\gamma}^{(m)}_{N-m+1,\cdots,N} = m!\binom{N}{m}\mathrm{Tr}_{\neq N-m+1,\cdots,N}(\rho_1\rho_2\cdots\rho_{N-m}D_\chi^{(m-1)}) \tag{36}$$

to be compared with

$$\rho^{(m)}_{N-m+1,\cdots,N} = m!\binom{N}{m}\mathrm{Tr}_{\neq N-m+1,\cdots,N}(\rho_1\rho_2\cdots\rho_{N-m}D_0).$$

We then have

$$D_\chi^{(m)} = D_\chi^{(m-1)} - \left[\binom{N}{m}^2\Big/N!\right]\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-m} \times [\mathrm{Tr}_{\neq N-m+1,\cdots,N}(\rho_1\rho_2\cdots\rho_{N-m}D_\chi^{(m-1)})])\mathcal{A}_N^\dagger. \tag{37}$$

When the above "successive orthogonalizations" terminate, we obtain the finite cluster expansion of the density operators. Thus,

$$D - D_0 = D_\chi = \sum_{n=1}^{N} D_\gamma(n), \tag{38}$$

with $D_\gamma(n)$ given by Eq. (35).

Also

$$D_\chi^{(m)} = \sum_{n\geq m+1}^{N} D_\gamma(n), \quad 1 \leq m \leq N-1, \quad D_\chi^{(N)} = 0. \tag{39}$$

SOME PROPERTIES OF THE CLUSTER DENSITIES

From Eqs. (35) and (39) we see that $D_\chi^{(m)}$ contains all $\hat{\gamma}^{(n)}$'s of $n \geq m+1$, but does not contain any $\hat{\gamma}^{(n)}$'s of $n \leq m$. Therefore, the "orthogonality" condition (34) for $m = N-1$, i.e.,

$$\mathrm{Tr}_1(\rho_1 D_\chi^{(N-1)}) = 0,$$

implies that $\mathrm{Tr}_1[\rho_1 D_\gamma(n)] = 0$ or

$$\mathrm{Tr}_1(\rho_1\hat{\gamma}^{(N)}_{1,2,\cdots,N}) = 0. \tag{40}$$

The "orthogonality" condition for $m = N-2$, i.e.,

$$\mathrm{Tr}_{1,2}(\rho_1\rho_2 D_\chi^{(N-2)}) = 0,$$

implies that

$$\mathrm{Tr}_{1,2}[\rho_1\rho_2 D_\gamma(n)] + \mathrm{Tr}_{1,2}[\rho_1\rho_2 D_\gamma(N-1)] = 0,$$

which, by the above result (40), reduces to

$$\mathrm{Tr}_{1,2}[\rho_1\rho_2 D_\gamma(N-1)] = 0,$$

which further reduces to

$$\mathrm{Tr}_2(\rho_2\hat{\gamma}^{(N-1)}_{2,3,\cdots,N}) = 0$$

or

$$\mathrm{Tr}_1(\rho_1\hat{\gamma}^{(N-1)}_{1,2,\cdots,N-1}) = 0. \tag{41}$$

By similar procedures, we get, in general,

$$\mathrm{Tr}_1(\rho_1\hat{\gamma}^{(m)}_{1,2,\cdots,m}) = 0. \tag{42}$$

This "orthogonality" is similar to the orbital orthogonality, previously discussed by one of us.[1a,e] It represents the "exclusion effect,"[1] and here it is again designated by the caret ($\hat{}$) on the top of γ. It is natural to call $\hat{\gamma}^{(m)}$ the "m-fermion correlation density" in D.

To find the expressions of $\hat{\gamma}^{(m)}$'s in terms of the correlation functions $\hat{U}'$'s introduced previously,[1e,4] we first review the following relations:

$$|\hat{U}_{k_1k_2\cdots k_m}\rangle \equiv \left(\frac{N!}{m!}\right)^{\frac{1}{2}}\left\langle\frac{12\cdots N}{k_1k_2\cdots k_m}\middle|\chi^{(m-1)}\right\rangle, \quad \hat{f}_k \equiv \hat{U}'_k, \tag{43}$$

$$|\{\hat{U}'_{k_1k_2\cdots k_m}\}\rangle \equiv \frac{1}{(m!)^{\frac{1}{2}}}\mathcal{A}_N\left|\frac{12\cdots N}{k_1k_2\cdots k_m}\,\hat{U}'_{k_1k_2\cdots k_m}\right\rangle, \tag{44}$$

and

$$\chi^{(m)} = \chi^{(m-1)} - \sum_{k_m>\cdots>k_2>k_1=1}^{N} \{\hat{U}'_{k_1k_2\cdots k_m}\}, \quad (45)$$

where each $\hat{U}'_{k_1k_2\cdots k_m}$ has the same permutation symmetry as $\mathcal{A}_m(k_1k_2\cdots k_m)$. Each $\hat{U}'_{k_1k_2\cdots k_m}$ replaces only $\mathcal{A}_m(k_1k_2\cdots k_m)$ in the w.f. $\mathcal{A}_N(12\cdots N)$, and does not replace, for instance, $\mathcal{A}_m(k'_1k_2\cdots k_m)$ if $k'_1 = k_1$. The $\mathcal{A}_m$ is the m-fermion antisymmetrizer.

Noting these relations and Eq. (18), we can easily derive the following expression for $\hat{\gamma}^{(1)}$ from Eq. (36):

$$\hat{\gamma}^{(1)} = \sum_{k=1}^{N}(|k\rangle\langle\hat{f}_k| + |\hat{f}_k\rangle\langle k| + |\hat{f}_k\rangle\langle\hat{f}_k|). \quad (46)$$

Substitution of this result into Eq. (21) or (35) gives

$$D_{\gamma(1)} = \sum_{k=1}^{N}(|\phi_0\rangle\langle\{\hat{f}_k\}| + |\{\hat{f}_k\}\rangle\langle\phi_0| + |\{\hat{f}_k\}\rangle\langle\{\hat{f}_k\}|), \quad (47)$$

and then, by Eq. (20) or (33),

$$\begin{aligned} D_{\chi'} = {} & |\phi_0\rangle\langle\chi'| + |\chi'\rangle\langle\phi_0| + |\chi'\rangle\langle\chi'| \\ & + \sum_{k=1}^{N}(|\{\hat{f}_k\}\rangle\langle\chi'| + |\chi'\rangle\langle\{\hat{f}_k\}|) \\ & + \sum_{l>k=1}(|\{\hat{f}_k\}\rangle\langle\{\hat{f}_l\}| + |\{\hat{f}_l\}\rangle\langle\{\hat{f}_k\}|). \end{aligned} \quad (48)$$

Using this result in Eq. (36) for the case $m = 2$, we obtain

$$\begin{aligned} \hat{\gamma}(2) = 2! \sum_{l>k=1}^{N} & [(\mathcal{A}_2|kl\rangle\langle\hat{U}'_{kl}| + |\hat{U}'_{kl}\rangle \\ & \times \langle kl|\mathcal{A}_2^{\dagger} + |\hat{U}'_{kl}\rangle\langle\hat{U}'_{kl}|) + (1!)^{-\frac{1}{2}}(\mathcal{A}_2|k\hat{f}_l\rangle \\ & \times \langle\hat{U}'_{kl}| + |\hat{U}'_{kl}\rangle\langle\hat{f}_l k|\mathcal{A}_2^{\dagger} + \mathcal{A}_2|\hat{f}_k l\rangle \\ & \times \langle\hat{U}'_{kl}| + |\hat{U}'_{kl}\rangle\langle\hat{f}_k l|\mathcal{A}_2^{\dagger}) + (1!)^{-1}(\mathcal{A}_2|k\hat{f}_l\rangle \\ & \times \langle\hat{f}_k l|\mathcal{A}_2^{\dagger} + \mathcal{A}_2|\hat{f}_k l\rangle\langle\hat{f}_l k|\mathcal{A}_2^{\dagger})]. \end{aligned} \quad (49)$$

In the same manner, similar expressions for the other $\hat{\gamma}^{(m)}$'s of higher orders can be found. These expressions clearly show the structures of $\hat{\gamma}^{(m)}$ in terms of the functions $\hat{U}'_{k_1\cdots k_n}$ $(n \le m)$. The $D_\gamma(m)$ represents a density matrix corresponding to the

$$\sum_{k_m>\cdots>k_1=1}^{N} \{\hat{U}'_{k_1\cdots k_m}\}.$$

However, it should be noted that $D_{\gamma(m)}$ in general contains some extra terms whose component wavefunctions do not appear in

$$\sum_{k_m>\cdots>k_1=1}^{N} \{\hat{U}'_{k_1\cdots k_m}\}$$

but do appear in

$$\sum_{k_n>\cdots>k_1=1}^{N} \{\hat{U}'_{k_1\cdots k_n}\} \quad (n < m).$$

This can be seen, for example, in Eq. (49) of $\hat{\gamma}^{(2)}$ [i.e., $\mathcal{A}_2(kl)$, $\mathcal{A}_2(k\hat{f}_l)$, etc.]

THE EXACT ENERGY

Consider an N-fermion Hamiltonian of the form

$$H = H_0 + H_1, \quad (50)$$

where

$$H_0 = \sum_{i=1}^{N} h'_i, \quad H_1 = \sum_{j>i=1}^{N} g'_{ij}. \quad (51)$$

Suppose that ϕ_0 is the eigenfunction of H_0, in other words the spin-orbitals k are the eigenfunctions of h'. Then, we have the Schrödinger equations, $H|\Psi\rangle = E|\Psi\rangle$ and $H_0|\phi_0\rangle = E_0|\phi_0\rangle$, or

$$HD = ED; \quad H_0D_0 = E_0D_0. \quad (52)$$

Introducing the energy E_{ϕ_0} such as

$$E_{\phi_0} \equiv \mathrm{Tr}\,(HD_0) = E_0 + E_1, \quad (53)$$

where

$$E_0 = \mathrm{Tr}\,(H_0D_0); \quad E_1 = \mathrm{Tr}\,(H_1D_0), \quad (54)$$

we can write the exact energy E as

$$E = \frac{\mathrm{Tr}\,(HD)}{\mathrm{Tr}\,(D)} = E_{\phi_0} + \frac{\mathrm{Tr}\,[(H - E_{\phi_0})D]}{\mathrm{Tr}\,(D)}, \quad (55)$$

or the exact "correlation" energy E_{corr} as

$$\begin{aligned} E_{\mathrm{corr}} &= E - E_{\phi_0} \\ &= \frac{\mathrm{Tr}\,[(H - E_{\phi_0})D]}{\mathrm{Tr}\,(D)}. \end{aligned} \quad (56)$$

For any operator G such that

$$G = \sum_{i_m>\cdots>i_2>i_1=1}^{N} g^{(m)}_{i_1i_2\cdots i_m}, \quad (57)$$

where $g^{(m)}$ is an m-particle operator, we have

$$\mathrm{Tr}\,(GD_0) = (1/m!)\,\mathrm{Tr}\,(g^{(m)}\rho^{(m)}). \quad (58)$$

Hence, E_{ϕ_0} can be written as

$$E_{\phi_0} = (1/1!)\,\mathrm{Tr}\,(h'\rho^{(1)}) + (1/2!)\,\mathrm{Tr}\,(g'_{12}\rho^{(2)}). \quad (59)$$

Now, to find a simpler expression of E_{corr} than Eq. (56), we go back to the Schrödinger equations in Eq. (52) and derive the following equation:

$$E_{\mathrm{corr}}D_0 = D_0(H_1 - [E_1 + E_{\mathrm{corr}}])D_\chi. \quad (60)$$

Taking the trace of both sides, we obtain

$$E_{\mathrm{corr}} = \mathrm{Tr}\,(D_0H_1D_\chi). \quad (61)$$

Substitution of Eq. (38) into the right-hand side of this equation gives

$$E_{\mathrm{corr}} = \sum_{m=1}^{N} \mathrm{Tr}\,[D_0H_1D_\gamma(m)].$$

Each term in this expression is reduced to

$$\mathrm{Tr}\,(D_0 H_1 D_\gamma^{(1)}) = \frac{1}{1!}\,\mathrm{Tr}\,[\rho_{12}^{(2)} g'_{12} \mathcal{A}_2(\rho_1^{(1)}\hat{\gamma}_2^{(1)})\mathcal{A}_2^\dagger],$$

$$\mathrm{Tr}\,(D_0 H_1 D_\gamma^{(m)}) = \frac{1}{m!}\,\mathrm{Tr}\left[\rho_{12\cdots m}^{(m)}\left(\sum_{j>i=1}^{m} g'_{ij}\right)\hat{\gamma}_{12\cdots m}^{(m)}\right] \quad (m \geq 2). \quad (62)$$

However, because of the "sea"-orthogonality (42), the terms of $m \geq 3$ vanish. Hence,

$$E_{\mathrm{corr}} = \frac{1}{1!}\,\mathrm{Tr}\,[\rho_{12}^{(2)} g'_{12} \mathcal{A}_2(\rho_1^{(1)}\hat{\gamma}_2^{(1)})\mathcal{A}_2^\dagger] + \frac{1}{2!}\,\mathrm{Tr}\,[\rho_{12}^{(2)} g'_{12}\hat{\gamma}_{12}^{(2)}]. \quad (63)$$

If the Schrödinger equation $H_0\,|\phi_0\rangle = E_0\,|\phi_0\rangle$ corresponds to the ordinary Hartree–Fock scheme, the first term of this expression vanishes because of Brillouin's theorem, and also, the "residual interaction" or "fluctuation potential" g'_{12} in the second term can be replaced by the bare interaction g_{12}.

SUMMARY OF THE MAIN EQUATIONS ON CORRELATION DENSITY CLUSTERS

The main equations which decompose an N-fermion density D into correlation cluster densities are summarized below:

$$D = D_0 + D_\chi, \quad \mathrm{Tr}\,(D_0 D_\chi) = 0,$$

$$D_\chi = D_\gamma(1) + D_\gamma(2) + \cdots + D_\gamma(N),$$

$$D_\gamma(m) = \frac{\binom{N}{m}}{m!\,N!}\,\mathcal{A}_N(\rho_1\rho_2\cdots\rho_{N-m}\hat{\gamma}_{N-m+1,\cdots,N}^{(m)})\mathcal{A}_N^\dagger,$$

where

$$\rho_i = \sum_{k\geq 1}^{N} |k\rangle\langle k| \quad \text{for the } i\text{th fermion.}$$

There is the "sea"-orthogonality of m-clusters:

$$\mathrm{Tr}_1\,(\rho_1\hat{\gamma}_{1,2,\cdots,m}^{(m)}) = 0.$$

The correlation energy in the H.F. scheme of an N-electron case is given by:

$$E_{\mathrm{corr}} = \frac{1}{2!}\,\mathrm{Tr}\left(\rho_{12}^{(2)}\frac{1}{r_{12}}\hat{\gamma}_{12}^{(2)}\right); \quad \rho_{12}^{(2)} = \mathcal{A}_2(\rho_1\rho_2)\mathcal{A}_2^\dagger.$$

CONCLUSION

We decomposed the total N-fermion density (matrix) operator D into a "model" (like Hartree–Fock) part D_0 and a correlation part D_χ. The D_χ itself is expressed in terms of newly defined n-fermion correlation densities $\hat{\gamma}^{(n)}$ where $n = 1, 2, 3, \cdots, N$. The D_χ also contains products of "Fock–Dirac" type densities ρ. As it has been shown previously by one of us[1b,e] that in ground state atoms and molecules correlation is made up mainly of $N(N-1)/2$ decoupled pair correlations and their products, the corresponding D_χ would contain mainly the two-body correlation density $\hat{\gamma}^{(2)}$ and its products. The $\hat{\gamma}^{(2)}$ differs from the conventional second-order reduced density matrix.[7] The two-body correlation density $\hat{\gamma}^{(2)}$ contains only the pair correlation functions $\hat{U}_{ij}(\mathbf{x}_i, \mathbf{x}_j)$, whereas the *reduced density matrix* has in it all the correlation effects, as well as the "model" wavefunction.

The transformation properties of correlation functions under basis set or unitary transformations of the "model" wavefunction orbitals are particularly transparent in the correlation density operators formulation.

ACKNOWLEDGMENTS

The support of this research by the U.S. National Science Foundation and the Alfred P. Sloan Foundation is gratefully acknowledged.

[7] See, especially, Refs. 5 (b), (c), (g), (i), and (j).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Algebraic Matrices for the Configurations $(d + s)^n p$

C. ROTH
Department of Mathematics, McGill University, Montreal, Canada

(Received 24 June 1968)

The construction of the algebraic matrices for the configurations $(d + s)^n p$ in L–S coupling is discussed. Particular emphasis is given to configurations with a half-filled d shell. From a total of 42 parameters specifying the various interactions for the configurations $(d + s)^n p$ only the matrices of 8 parameters need to be calculated explicitly for the complementary configurations $(d + s)^n p$, $n > 6$. The matrices of the other 34 parameters can be obtained from the corresponding algebraic matrices of $(d + s)^n p$, $n \leq 6$, either directly or by simple changes in sign.

1. INTRODUCTION

The algebraic matrices of $(d + s)^n p$ comprise the electrostatic and spin–orbit interaction matrices of the configurations $d^n p$, $d^{n-1}sp$, and $d^{n-2}s^2 p$, the matrices of the correction parameters representing two- and three-body interactions of the core d electrons, as well as the matrices of the interactions between configurations. The energy matrix (for a particular n) is then a linear combination of these matrices, the coefficients of which are parameters usually obtained empirically by fitting the experimental levels to the eigenvalues of the energy matrix. The following lists these parameters and gives their significance (unprimed quantities denote the configuration $d^n p$, primes denote $d^{n-1}sp$ and double primes denote $d^{n-2}s^2 p$):

A, A', A''—the heights of the configurations

$$S' = A' - A, \quad S'' = A'' - A.$$

B, B', B''—linear combinations of the Slater parameters $F^2(dd)$ and $F^4(dd)$:

$$B = \tfrac{1}{441}[9F^2(dd) - 5F^4(dd)] = F_2(dd) - 5F_4(dd).^1$$

C, C', C''—multiples of the Slater parameter $F^4(dd)$:

$$C = \tfrac{5}{63}F^4(dd) = 35F_4(dd).^1$$

G'_{ds}—the parameter of the d–s interaction in the configuration $d^{n-1}sp$:

$$G'_{ds} = \tfrac{1}{5}G^2(ds).^2$$

F_2, F'_2, F''_2—parameters of the direct part of the d–p interaction

$$F_2 = \tfrac{1}{35}F^2(dp).^1$$

G_1, G'_1, G''_1—parameters of the exchange part of the d–p interaction

$$G_1 = \tfrac{1}{15}G^1(dp).^1$$

G_3, G'_3, G''_3—parameters of the exchange part of the d–p interaction

$$G_3 = \tfrac{3}{245}G^3(dp).^1$$

G'_{ps}—the parameter of the p–s interaction in the configuration $d^{n-1}sp$:

$$G'_{ps} = \tfrac{1}{3}G^1(ps).^3$$

α, α', α''—correction parameters multiplying $L_1(L_1 + 1)$, where L_1 is the angular momentum of the core of $(d + s)^n p$, i.e., $(d + s)^n$.[4–6]

β, β', β''—correction parameters multiplying the seniority operator of Racah.[2,7]

T, T', T''—parameters of Trees multiplying the squared matrix of the interaction $3s^2 3d^n$–$3s3d^{n+1}$.[8,9]

H—parameter of the d^n–$d^{n-1}s$ interaction

$$H = R^2(dd, ds)/35.^2$$

H'—parameter of the $d^{n-1}s$–$d^{n-2}s^2$ interaction, defined the same as H.

J—parameter of the direct part of the $d^n p$–$d^{n-1}sp$ interaction

$$J = R^2(dp, sp)/5.^{10,11}$$

J'—parameter of the direct part of the $d^{n-1}sp$–$d^{n-2}s^2 p$ interaction, defined the same as J.

K—parameter of the exchange part of the $d^n p$–$d^{n-1}sp$ interaction

$$K = R^1(dp, ps)/3.^{10,11}$$

K'—parameter of the exchange part of the $d^{n-1}sp$–$d^{n-2}s^2 p$ interaction, defined the same as K.

[1] G. Racah, Phys. Rev. **62**, 438 (1942), referred to as R II.
[2] G. Racah, Phys. Rev. **63**, 367 (1943), referred to as R III.
[3] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England, 1935), referred to as TAS.
[4] R. E. Trees, Phys. Rev. **83**, 756 (1951); **84**, 1089 (1951).
[5] G. Racah, Phys. Rev. **85**, 381 (1952).
[6] G. Racah, Lunds Univ. Arsskr. Ard. (2) **50**, 31 (1955).
[7] G. Racah and Y. Shadmi, Phys. Rev. **119**, 156 (1960).
[8] R. E. Trees and C. K. Jorgensen, Phys. Rev. **123**, 1278 (1961).
[9] R. E. Trees, Phys. Rev. **129**, 1220 (1963).
[10] N. Rosenzweig, Phys. Rev. **88**, 580 (1952).
[11] G. Racah, unpublished material, 1952.

G—parameter of the $d^n p$–$d^{n-2}s^2 p$ interaction

$$G = R^2(dd, ss)/5 = R^2(ds, sd)/5 = G^2(ds)/5 = G'_{ds}.^2$$

ζ_d, ζ'_d, ζ''_d—parameters of the spin–orbit interaction of the d electrons.[2]

ζ_p, ζ'_p, ζ''_p—parameters of the spin–orbit interaction of the p electron.[2]

The algebraic matrices of $(d+s)^n p$ were constructed and then checked by the author with the purpose of using them to explain and predict the spectra of the configurations $(3d+4s)^n 4p$ in neutral and singly-ionized atoms of the iron group. The checked algebraic matrices of $(d+s)^n p$ for all permissible n are available and can be obtained by request.

2. THE CONSTRUCTION OF THE ALGEBRAIC MATRICES IN L–S COUPLING

The matrix elements of the electrostatic and spin–orbit interactions for the configurations $d^n p$ were obtained by Racah,[2] Rohrlich and others.[12–14] General formulas for the matrix elements of the electrostatic and spin–orbit interactions for the configurations $d^{n-1}sp$ were obtained by the author.[15,16] Characterizing the states of $d^n p$ by $d^n(v_1 S_1 L_1)pSLJM$, and d^{n-1} by $(v_2 S_2 L_2)$, and using the same methods as for $d^{n-1}sp$ we obtain the general formula for the d–p interaction of the configuration $d^n p$:

$$\begin{aligned}
&nF_0(dp) + n\sum_{v_2S_2L_2}\Big[\langle d^n v_1 S_1 L_1\{|\, d^{n-1}(v_2S_2L_2)\, dS_1L_1\rangle\langle d^{n-1}(v_2S_2L_2)\, dS_1L'_1\,|\}\, d^n v'_1 S_1 L'_1\rangle\delta(S_1, S'_1)\\
&\quad\times[\exp\pi i(L_1+L'_1+L_2+L+1)][2100(2L_1+1)(2L'_1+1)]^{\frac12}\,\overline{W}\begin{pmatrix}L_1 & L'_1 & 2\\ 2 & 2 & L_2\end{pmatrix}\overline{W}\begin{pmatrix}L_1 & L'_1 & 2\\ 1 & 1 & L\end{pmatrix}F_2(dp)\Big]\\
&\quad+n\sum_{v_2S_2L_2}\Big[\langle d^n v_1 S_1 L_1\{|\, d^{n-1}(v_2S_2L_2)\, dS_1L_1\rangle\langle d^{n-1}(v_2S_2L_2)\, dS'_1L'_1\,|\}\, d^n v'_1 S'_1 L'_1\rangle[\exp\pi i(S_1+S'_1)][(2S_1+1)(2S'_1+1)\\
&\quad\times(2L_1+1)(2L'_1+1)]^{\frac12}\overline{W}\begin{pmatrix}S_2 & \frac12 & S_1\\ S & \frac12 & S'_1\end{pmatrix}\left\{30X\begin{pmatrix}1 & L & L_1\\ 2 & L'_1 & L_2\\ 1 & 1 & 2\end{pmatrix}G_1(dp)+105X\begin{pmatrix}1 & L & L_1\\ 2 & L'_1 & L_2\\ 3 & 1 & 2\end{pmatrix}G_3(dp)\right\}\Big].
\end{aligned}\tag{1}$$

Similarly, the general formula for the spin–orbit interaction of the electrons d of $d^n p$ is given by

$$\begin{aligned}
&n\zeta_d\sum_{v_2S_2L_2}\{\langle d^n v_1 S_1 L_1\{|\, d^{n-1}(v_2S_2L_2)\, dS_1L_1\rangle\\
&\quad\times\langle d^{n-1}(v_2S_2L_2)\, dS'_1L'_1\,|\}\, d^n v'_1 S'_1 L'_1\rangle\\
&\quad\times[\exp\pi i(2S_1+2S'+S_2+L_2+L'+L+J)]\\
&\quad\times\overline{W}\begin{pmatrix}S & S' & 1\\ L' & L & J\end{pmatrix}\overline{W}\begin{pmatrix}S & S' & 1\\ S'_1 & S_1 & \frac12\end{pmatrix}\\
&\quad\times\overline{W}\begin{pmatrix}S_1 & S'_1 & 1\\ \frac12 & \frac12 & S_2\end{pmatrix}\overline{W}\begin{pmatrix}L_1 & L'_1 & 1\\ 2 & 2 & L_2\end{pmatrix}\overline{W}\begin{pmatrix}L & L' & 1\\ L'_1 & L_1 & 1\end{pmatrix}\\
&\quad\times[45(2L_1+1)(2L'_1+1)(2L+1)(2L'+1)\\
&\quad\times(2S_1+1)(2S'_1+1)(2S+1)(2S'+1)]^{\frac12}\}.
\end{aligned}\tag{2}$$

The spin–orbit interaction of the electron p of $d^n p$ is

$$\begin{aligned}
&3\zeta_p[\exp\pi i(S_1+S+S'+L_1+J+\tfrac32)]\\
&\times[(2L+1)(2L'+1)(2S+1)(2S'+1)]^{\frac12}\\
&\times\overline{W}\begin{pmatrix}S & S' & 1\\ L' & L & J\end{pmatrix}\overline{W}\begin{pmatrix}S & S' & 1\\ \frac12 & \frac12 & S_1\end{pmatrix}\overline{W}\begin{pmatrix}L & L' & 1\\ 1 & 1 & L_1\end{pmatrix}.
\end{aligned}\tag{3}$$

As expected, the above result is identical to Eq. (17) in II giving the matrix elements of the spin–orbit interaction of the electron p for $d^{n-1}sp$.

The d–d interaction of $d^n p$ is the same as that of d^n and hence is given by Racah.[1]

For the configurations $d^{n-2}s^2 p$ the electrostatic energy matrix elements are

$$\begin{aligned}
&\langle d^{n-2}s^2 p|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-2}s^2 p\rangle\\
&\quad=[n(n+1)/2]\,\langle d^{n-2}s^2 p|\,(e^2/r_{n,n+1})\,|d^{n-2}s^2 p\rangle.
\end{aligned}\tag{4}$$

After expanding each side of the matrix element in terms of antisymmetric eigenfunctions, we obtain contributions which characterize the d–d, d–p, s–p, and d–s interactions. For the above matrix element, the d–d and d–p interactions are given by those contributions for which the nth and $(n+1)$th electrons are either both d or one d and the other p.

[12] F. Rohrlich, Phys. Rev. **74**, 1372 (1948).
[13] T. Ishidzu and S. Obi, J. Phys. Soc. Japan **5**, 124 (1950).
[14] E. Shimoni, S. Hollander, and B. Z. Abraham, M.Sc. Theses, The Hebrew University of Jerusalem (1960).
[15] C. Roth, J. Math. Phys. **9**, 686 (1968), referred to as I.
[16] C. Roth, J. Math. Phys. **9**, 1832 (1968), referred to as II.

Then using (7.10) and (11.10) ITS,[17] we obtain

$$\begin{aligned}\langle d^{n-2}(v_1S_1L_1)s^2pSL|\,(e^2/r_{n,n+1})\,|d^{n-2}(v_1'S_1'L_1')s^2pSL\rangle &= [\exp \pi i(L_1+1+L)][(2L_1+1)(2L+1)]^{\frac{1}{2}} \\ &\times \overline{W}\begin{pmatrix} L_1 & 0 & L_1 \\ L & 1 & L \end{pmatrix}[\exp \pi i(L_1'+1+L)] \\ &\times [(2L'+1)(2L+1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix} L_1' & 0 & L_1' \\ L & 1 & L \end{pmatrix} \\ &\times [\exp \pi i(S_1+\tfrac{1}{2}+S)][(2S_1+1)(2S+1)]^{\frac{1}{2}} \\ &\times \overline{W}\begin{pmatrix} S_1 & 0 & S_1 \\ S & \frac{1}{2} & S \end{pmatrix}[\exp \pi i(S_1'+S+\tfrac{1}{2})] \\ &\times [(2S_1'+1)(2S+1)]^{\frac{1}{2}}W\begin{pmatrix} S_1' & 0 & S_1' \\ S & \frac{1}{2} & S \end{pmatrix} \\ &\times \langle s^2\,{}^1S\, d^{n-2}(v_1S_1L_1)pSL|\,(e^2/r_{n,n+1}) \\ &\times |s^2\,{}^1S\, d^{n-2}(v_1'S_1'L_1')pSL\rangle. \qquad (5)\end{aligned}$$

Remembering that either the d–d or d–p interactions are considered and using (11.12) ITS, we obtain

$$\begin{aligned}&\langle d^{n-2}(v_1S_1L_1)s^2pSL|\,(e^2/r_{n,n+1})\,|d^{n-2}(v_1'S_1'L_1')s^2pSL\rangle \\ &\quad = \langle d^{n-2}(v_1S_1L_1)pSL|\,(e^2/r_{n,n+1})\,|d^{n-2}(v_1'S_1'L_1')pSL\rangle. \qquad (6)\end{aligned}$$

Thus, the d–d and d–p interactions of $d^{n-2}s^2p$ are the same as the corresponding interactions for $d^{n-2}p$.

The interactions of the p electron and s electrons with the closed d shell can be obtained from Eqs. (11) and (9) and Eq. (10).[18] The results are $2F_0(ps) - G_{ps}$ and $(n-2)[2F_0(ds) - G_{ds}]$, respectively. Both contributions can be incorporated into the height of the configuration A''.

The spin–orbit interactions for the configurations $d^{n-2}s^2p$ are the same as those for $d^{n-2}p$.

The matrix elements for the electrostatic interactions between configurations were derived by Rosenzweig[10] and Racah.[11] As the latter paper was not published and the parameters J and K first introduced there we quote the results:

$$\begin{aligned}\langle d^n(v_1S_1L_1)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-1}(v_2S_2L_2)s(S'L_2)pSL\rangle &= \langle d^nv_1S_1L_1|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}} \\ &\times |d^{n-1}(v_2S_2L_2)sS_1L_1\rangle\,\delta(S_1S')\delta(L_1L_2) \\ &+ \langle d^nv_1S_1L_1\{|\,d^{n-1}(v_2S_2L_2)\,dS_1L_1\rangle \\ &\times [6n(2L_1+1)]^{\frac{1}{2}}[\exp \pi i(L_2+L)] \\ &\times \overline{W}\begin{pmatrix} L_2 & 2 & L_1 \\ 1 & L & 1 \end{pmatrix}[(J-K)\delta(S_1S') \\ &+ [\exp \pi i(S_1-S')][(2S_1+1)(2S'+1)]^{\frac{1}{2}} \\ &\times [1/(2S+1)]K\delta(S_2S)] \qquad (7)\end{aligned}$$

and

$$\begin{aligned}\langle d^{n-1}(v_1S_1L_1)s(S'L_1)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-2}(v_2S_2L_2)s^2pSL\rangle &= \langle d^{n-1}(v_1S_1L_1)sS_2L_2|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}} \\ &\times |d^{n-2}(v_2S_2L_2)s^2S_2L_2\rangle\,\delta(S'S_2)\delta(L_1L_2) \\ &+ \langle d^{n-1}v_1S_1L_1\{|\,d^{n-2}(v_2S_2L_2)\,dS_1L_1\rangle \\ &\times [6(n-1)(2L_1+1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix} L_2 & 2 & L_1 \\ 1 & L & 1 \end{pmatrix} \\ &\times [\exp \pi i(L_2+L+S_1-S'-\tfrac{1}{2})] \\ &\times \{[(2S_1+1)/(2S'+1)]^{\frac{1}{2}}J'\delta(S'S_2) \\ &- [(2S'+1)/(2S_1+1)]^{\frac{1}{2}}K'\delta(S_1S)\}. \qquad (8)\end{aligned}$$

The matrix elements of the interactions between the cores, i.e.,

$$\langle d^nv_1S_1L_1|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-1}(v_2S_2L_2)sS_1L_1\rangle$$

and

$$\langle d^{n-1}(v_1S_1L_1)sS_2L_2|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-2}(v_2S_2L_2)s^2S_2L_2\rangle$$

have been calculated by Racah.[2]

The matrix elements for the interaction between the configurations d^np and $d^{n-2}s^2p$ are the same as those for d^n and $d^{n-2}s^2$. The latter have also been calculated by Racah.[2] Explicitly,

$$\begin{aligned}\langle d^n(v_1S_1L_1)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-2}(v_1'S_1'L_1')s^2pSL\rangle &= \langle d^nv_1S_1L_1|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-2}s^2v_1'S_1'L_1'\rangle\,\delta(S_1S_1')\delta(L_1L_1') \\ &= [(n-v)(12-n-v)/4]^{\frac{1}{2}}\delta(vv')\delta(S_1S_1')\delta(L_1L_1')G. \qquad (9)\end{aligned}$$

Racah and Trees[4–6] have shown that second-order effects caused by perturbations on the configuration l^n by configurations differing from l^n by two electrons can be described by a model interaction of the form

$$\sum_{i<j}[2\alpha(\bar{l}_i\circ\bar{l}_j)+\beta q_{ij}],$$

where q_{ij} is the seniority operator.[2] For the configuration d^n this becomes

$$\alpha[L(L+1)-6n]+\beta Q,$$

where Q is the total seniority operator.[2] If the constant $-6n\alpha$ is incorporated into the height of the configuration, the above correction reduces to

$$\alpha L(L+1)+\beta Q.$$

[17] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959), referred to as ITS.

[18] Reference 3, pp. 182 and 176, respectively.

Trees and Jorgensen[8] have shown that the main perturbing configuration on $3s^2 3p^6 3d^n$ is the configuration $3s^2 3p^4 3d^{n+2}$. Trees[9] also remarked that the configuration $3s3p^6 3d^{n+1}$ should give a perturbation of the same magnitude as $3s^2 3p^4 3d^{n+2}$. This perturbation is not included in

$$\sum_{i<j} [2\alpha(\bar{l}_i \circ \bar{l}_j) + \beta q_{ij}],$$

since now the configurations differ by only one electron. By second-order perturbation theory, this effect depends upon the ratio $H^2/\Delta E$, where H is the interaction parameter that appears in the nondiagonal term and ΔE is the energy difference between the two configurations. The parameter $H^2/\Delta E$ is denoted by T. When calculating the model interaction, one uses second-order perturbation theory of degenerate configurations which permits the introduction of these interactions before diagonalizing the energy matrices of the separate configurations. Hence, the matrices of T are not diagonal.

As a particular example, consider the perturbation $3s3p^6 3d^4$ on the terms $3s^2 3p^6 3d^3{}_1{}^2D$ and $3s^2 3p^6 3d^3{}_3{}^2D$. The terms of d^4 which have to be considered are ${}^1_2D$, ${}^1_4D$, and ${}^3_4D$.

From $d^4{}_2{}^1D$ we get, using (81) (R III) and Table XXI (R III),

$$\langle d^4{}_2{}^1D s^2 D | \sum_{i<j} \frac{e^2}{r_{ij}} | d^3{}_1{}^2D s^2\,{}^2D \rangle$$

$$= -[\tfrac{1}{2}]^{\frac{1}{2}} \langle d^4{}_2{}^1D | \sum_{i<j} \frac{e^2}{r_{ij}} | d^3{}_1{}^2D s\,{}^1D \rangle = [{}^{1\,0\,5}_{\frac{1}{2}}]^{\frac{1}{2}} H$$

and

$$\langle d^4{}_2{}^1D s\,{}^2D | \sum_{i<j} \frac{e^2}{r_{ij}} | d^3{}_3{}^2D s^2\,{}^2D \rangle = [{}^{4\,5}_{\frac{5}{2}}]^{\frac{1}{2}} H.$$

Thus by perturbation theory, the diagonal elements are $[{}^{1\,0\,5}_{\frac{1}{2}}]H^2/\Delta E$ and $[{}^{4\,5}_{\frac{5}{2}}]H^2/\Delta E$, whereas the nondiagonal element from the contribution of $d^4{}_2{}^1D$ equals $[{}^{4\,7\,2\,5}_{\frac{4}{4}}]^{\frac{1}{2}} H^2/\Delta E$.

Similarly, the terms $d^4{}_4{}^1D$ and $d^4{}_4{}^3D$ contribute diagonal elements only for $d^3{}_3{}^2D$ of $180H^2/\Delta E$ and $120H^2/\Delta E$, respectively.

Hence, the matrix of T for $d^3{}_1{}^2D$ and $d^3{}_3{}^2D$ can be written as

$$\begin{matrix} {}^2_1D \\ {}^2_3D \end{matrix} \begin{bmatrix} {}^{1\,0\,5}_{\frac{2}{2}} & ({}^{4\,7\,2\,5}_{\frac{4}{4}})^{\frac{1}{2}} \\ ({}^{4\,7\,2\,5}_{\frac{4}{4}})^{\frac{1}{2}} & {}^{6\,4\,5}_{\frac{2}{2}} \end{bmatrix}.$$

The matrix of A is a unit matrix for the three configurations $d^n p$, $d^{n-1}sp$, and $d^{n-2}s^2 p$. The matrices of S' and S'' are unit matrices for the configurations $d^{n-1}sp$ and $d^{n-2}s^2 p$, respectively.

3. COMPLEMENTARY CONFIGURATIONS

In TAS,[19] it is shown that the electrostatic interaction matrices of the configurations d^n and d^{10-n} are the same. From this result and Eq. (6), it is evident that the matrices of the parameters B and C for the configurations $d^n p$, $d^{n-1}sp$, and $d^{n-2}s^2 p$ are equal to the corresponding matrices of $d^{10-n}s^2 p$, $d^{11-n}sp$, and $d^{12-n}p$, respectively. From the above article it is also evident that the electrostatic energy matrices of the configurations $d^n s$ and $d^{10-n}s$ are the same. Thus, the matrix of G'_{ds} for the configuration $d^{n-1}sp$ is the same as that of $d^{11-n}sp$.

From R II,[20] it is seen that the matrices of the parameters $F_2(dp)$ and ζ_d for the configurations $d^n p$, $d^{n-1}sp$, and $d^{n-2}s^2 p$ are equal in magnitude but of opposite sign to the corresponding matrices of $d^{10-n}s^2 p$, $d^{11-n}sp$, and $d^{12-n}p$, respectively.

For the matrices of $G_1(dp)$ and $G_3(dp)$, Eqs. (1) and (21) of I are used in a slightly modified form. For the case of an almost closed shell and an electron outside the shell, it is customary to subtract from the energy matrix the interaction of the outer electron with the closed shell. Then, the resulting matrices of $l^{4l+1}l'$ have a particularly simple form. Here we subtract the exchange interaction contribution of the electron p with the closed shell of d electrons when using Eqs. (1) and (21) of I to obtain the matrices of $G_1(dp)$ and $G_3(dp)$ for an almost complete shell. From TAS,[21] this interaction is given by $-10G_1 - 35G_3$. The coefficients of fractional parentage for an almost complete shell which are required to calculate the matrices of $G_1(dp)$ and $G_3(dp)$ can be obtained from (19) (R III) and Tables II, III, and IV of (R III). Rosenzweig[10] calculated the coefficients of fractional parentage by this method for d^n, $n > 5$. Checking his results, one mistake was found. From (19) (R III),

$$\langle d^6{}_4{}^3D \{| d^5{}_5{}^2S\, d^3D \rangle = -\tfrac{1}{3}\langle d^5{}_5{}^2S \{| d^4{}_4{}^3D\, d^2S \rangle.$$

By Table IV (R III), the above coefficient equals $-[\frac{1}{1\,5}]^{\frac{1}{2}}$, whereas in Table VI of Rosenzweig's paper it is given as $[\frac{1}{1\,5}]^{\frac{1}{2}}$.

Since each term of $d^{11-n}sp$ has a corresponding term formed from the same ancestors as $d^{n-1}sp$, it is evident from Eq. (34) of I that the matrix of G'_{ps} for $d^{n-1}sp$ is the same as the matrix of G'_{ps} for $d^{11-n}sp$.

From (3), giving the spin–orbit interaction for the electron p of $(d+s)^n p$, it is evident that this formula is independent of n. Hence the matrices of ζ_p for $d^n p$, $d^{n-1}sp$, and $d^{n-2}s^2 p$ are equal to the corresponding matrices of $d^{10-n}s^2 p$, $d^{11-n}sp$, and, $d^{12-n}p$, respectively.

[19] Reference 3, Chap. XIII, Sec. 1.
[20] Reference 1, Sec. 6.
[21] Reference 3, Eqs. (9) and (10), p. 182, and Eq. (10), p. 176.

Also, from the definitions of the correction parameters α and β, it is evident that the matrices of these parameters for d^np, $d^{n-1}sp$, and $d^{n-2}s^2p$ are equal to the corresponding matrices of $d^{10-n}s^2p$, $d^{11-n}sp$, and $d^{12-n}p$, respectively. On the other hand, the matrices for the correction parameter T must be calculated separately for each n using the method outlined in the previous section.

We now derive the transformation relationships from the original configurations $(d+s)^np$ to the complementary configurations $(d+s)^{12-n}p$ for the d^n–$d^{n-1}s$ and $d^{n-1}s$–$d^{n-2}s^2$ interactions.

Using (74) (R II) to transform the reduced matrix elements of the tensor $U^{(2)}$ and (19) (R III) to transform the coefficients of fractional parentage we obtain from (79) (R III):

$$\begin{aligned}
&\langle d^n vSL|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-1}(v'S'L)sSL\rangle\\
&\quad=[n/14]^{\frac12}\Big\{\sum_{v''L''}[\exp\pi i(L-L')][\exp\pi i(L-L'+1)]\\
&\qquad\times\langle d^{10-n}vSL\|\,U^{(2)}\,\|d^{10-n}v''SL''\rangle\\
&\qquad\times[\exp\pi i(S+S'+L+L''-\tfrac12)][(11-n)\\
&\qquad\times(2S'+1)(2L+1)/n(2S+1)(2L''+1)]^{\frac12}\\
&\qquad\times\langle d^{11-n}v'S'L\,\{|\,d^{10-n}(v''SL'')\,dS'L\rangle\\
&\qquad\times(2L''+1)^{\frac12}/(2L+1)+\sum_{v''L''}[\exp\pi i(L-L'')]\\
&\qquad\times[\exp\pi i(L-L'+1)]\\
&\qquad\times\langle d^{11-n}v''S'L''\|\,U^{(2)}\,\|d^{11-n}v'S'L\rangle\\
&\qquad\times[\exp\pi i(S+S'+L+L''-\tfrac12)][(11-n)\\
&\qquad\times(2S'+1)(2L''+1)/n(2S+1)(2L+1)]^{\frac12}\\
&\qquad\times\langle d^{11-n}v''S'L''\,\{|\,d^{10-n}(vSL)\,dS'L''\rangle/(2L+1)^{\frac12}\Big\}\\
&\qquad\times R^2(dd,ds).
\end{aligned}\tag{10}$$

The above expression can be written as

$$\begin{aligned}
&[\exp 2\pi iS'][\exp\pi i(S+\tfrac12-S')]\\
&\quad\times[(2S'+1)/(2S+1)]^{\frac12}[(11-n)/14]^{\frac12}\\
&\quad\times\Big\{\sum_{v''L''}[\exp\pi i(L-L')]\langle d^{10-n}vSL\|\,U^{(2)}\,\|d^{10-n}v''SL''\rangle\\
&\quad\times\langle d^{11-n}v'S'L\,\{|\,d^{10-n}(v''SL'')\,dS'L\rangle\\
&\quad\times\Big[\frac{1}{2L+1}\Big]^{\frac12}+\sum_{v''L''}[\exp\pi i(L-L'')]\\
&\quad\times\langle d^{11-n}v''S'L''\|\,U^{(2)}\,\|d^{11-n}v'S'L\rangle\\
&\quad\times\langle d^{11-n}v''S'L''\,\{|\,d^{10-n}(vSL)\,dS'L''\rangle\\
&\quad\times(2L''+1)^{\frac12}/(2L+1)\Big\}R^2(dd,ds).
\end{aligned}$$

By (79) (R III), the above equals

$$\begin{aligned}
&[\exp 2\pi iS'][\exp\pi i(S+\tfrac12-S')][(2S'+1)/(2S+1)]^{\frac12}\\
&\qquad\times\langle d^{11-n}v'S'L|\sum_{i<j=1}^{11-n}\frac{e^2}{r_{ij}}|d^{10-n}(vSL)sS'L\rangle.
\end{aligned}$$

Then by using (81) (R III) we finally obtain

$$\begin{aligned}
&\langle d^nvSL|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-1}(v'S'L)sSL\rangle\\
&\quad=[\exp 2\pi iS']\,\langle d^{11-n}(v'S'L)sSL|\sum_{i<j=1}^{12-n}\frac{e^2}{r_{ij}}|d^{10-n}s^2vSL\rangle.
\end{aligned}\tag{11}$$

On replacing n by $(12-n)$ in (11) we obtain

$$\begin{aligned}
&\langle d^{n-1}(v'S'L)sSL|\sum_{i<j=1}^{n}\frac{e^2}{r_{ij}}|d^{n-2}s^2vSL\rangle\\
&\quad=[\exp 2\pi iS']\,\langle d^{12-n}vSL|\sum_{i<j=1}^{12-n}\frac{e^2}{r_{ij}}|d^{11-n}(v'S'L)sSL\rangle.
\end{aligned}\tag{12}$$

From (11) [(12)], it is evident that for $n>6$ the matrix of $H[H']$ for $(d+s)^np$ equals that of $H'[H]$ for $(d+s)^{12-n}p$ if n is odd, and equals the negative of the matrix of $H'[H]$ for $(d+s)^{12-n}p$ if n is even.

We now derive similar relationships as (11) and (12) for the matrices of J and J'.

From (7) we obtain by using (19) (R III) for the transformation of the fractional parentage coefficients that the coefficient of J in the matrix element

$$\langle d^n(v_1S_1L_1)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-1}(v_2S_2L_2)s(S'L_2)pSL\rangle$$

is

$$\begin{aligned}
&[\exp\pi i(S_1+L_1+S_2+L_2-\tfrac12)]\\
&\qquad\times[(11-n)(2S_2+1)(2L_2+1)/n(2S_1+1)\\
&\qquad\times(2L_1+1)]^{\frac12}[6n(2L_1+1)]^{\frac12}\\
&\qquad\times\langle d^{11-n}v_2S_2L_2\,|\}\,d^{10-n}(v_1S_1L_1)\,dS_2L_2\rangle\\
&\qquad\times[\exp\pi i(L_2+L)]\overline{W}\begin{pmatrix}L_2&2&L_1\\1&L&1\end{pmatrix}\delta(S_1S')\\
&\quad=\langle d^{11-n}v_2S_2L_2\,|\}\,d^{10-n}(v_1S_1L_1)\,dS_2L_2\rangle\\
&\qquad\times[6(11-n)(2L_2+1)]^{\frac12}[\exp 2\pi iS_1]\\
&\qquad\times[\exp\pi i(L_1+L+S_2-S_1-\tfrac12)]\\
&\qquad\times\overline{W}\begin{pmatrix}L_1&2&L_2\\1&L&1\end{pmatrix}[(2S_2+1)/(2S_1+1)]^{\frac12}\delta(S_1S')\\
&\quad=[\exp 2\pi iS_1]\,\langle d^{11-n}v_2S_2L_2\,|\}\,d^{10-n}\\
&\qquad\times(v_1S_1L_1)\,dS_2L_2\rangle\,[6(11-n)(2L_2+1)]^{\frac12}\\
&\qquad\times[\exp\pi i(L_1+L+S_2-S'-\tfrac12)]\\
&\qquad\times\overline{W}\begin{pmatrix}L_1&2&L_2\\1&L&1\end{pmatrix}[(2S_2+1)/(2S'+1)]^{\frac12}\delta(S'S_2).
\end{aligned}$$

By (8), the above expression is just $\exp(2\pi i S_1)$ multiplied by the coefficient of J' in

$$\langle d^{11-n}(v_2S_2L_2)s(S'L_2)pSL|\sum_{i<j=1}^{13-n}\frac{e^2}{r_{ij}}|d^{10-n}(v_1S_1L_1)s^2pSL\rangle$$

Thus,

$$\langle d^n(v_1S_1L_1)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-1}(v_2S_2L_2)s(S'L_2)pSL\rangle_J$$
$$=\exp(2\pi iS_1)\langle d^{11-n}(v_2S_2L_2)s(S'L_2)pSL|$$
$$\times\sum_{i<j=1}^{13-n}\frac{e^2}{r_{ij}}|d^{10-n}(v_1S_1L_1)s^2pSL\rangle_J. \quad (13)$$

On replacing n by $(12-n)$ in (13) we get

$$\langle d^{n-1}(v_2S_2L_2)s(S'L_2)pSL|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-2}(v_1S_1L_1)s^2pSL\rangle_J$$
$$=\exp(2\pi iS_1)\langle d^{12-n}(v_1S_1L_1)pSL|$$
$$\times\sum_{i<j=1}^{13-n}\frac{e^2}{r_{ij}}|d^{11-n}(v_2S_2L_2)s(S'L_2)pSL\rangle_J. \quad (14)$$

From (13) [(14)], it is evident that for $n>6$, the matrix of $J[J']$ for $(d+s)^n p$ equals that of $J'[J]$ for $(d+s)^{12-n}p$ if n is even, and equals the negative of the matrix of $J'[J]$ for $(d+s)^{12-n}p$ if n is odd.

The matrices of K and K' must be calculated for all n using (7) and (8). The coefficients of fractional parentage are transformed using (19) (R III).

From (9) directly,

$$\langle d^n p|\sum_{i<j=1}^{n+1}\frac{e^2}{r_{ij}}|d^{n-2}s^2p\rangle=\langle d^{10-n}s^2p|\sum_{i<j=1}^{13-n}\frac{e^2}{r_{ij}}|d^{12-n}p\rangle.$$

However, as pointed out after (75) (R III), a minus sign must be introduced in the above expression for n equal to 6 and terms of seniority number 2 as well as for n equal to 7 and seniority number 3.

Thus, with the exception of G_1, G_1', G_1'', G_3, G_3', G_3'', T, T', T'', K, and K', the configuration $d^n p$ corresponds to $d^{10-n}s^2p$, $d^{n-1}sp$ to $d^{11-n}sp$ and $d^{n-2}s^2p$ to $d^{12-n}p$. However, as the matrices of G_1'', G_3'', and T'' of $d^{n-2}s^2p$ correspond to the matrices of G_1, G_3, and T of $d^{n-2}p$, the matrices of only 8 parameters need to be calculated anew for complementary configurations. Of these, the matrices of T and T' are particularly easy to evaluate.

It should be noted that for $n>6$ the height of the configuration $d^n p$ is $A+S''$, the height of $d^{n-1}sp$ is $A+S'$, and the height of $d^{n-2}s^2p$ is given by A.

4. THE CONFIGURATIONS d^5p, d^5sp, AND d^5s^2p

The half-filled d shell can be considered as either the configuration comprising five electrons or as the configuration with five holes in the d shell. From (65) (R III), it is evident that the only difference in the energy matrices in these two cases is a change in sign of all nondiagonal matrix elements connecting terms of seniority number 1 or 5 with terms of seniority number 3. Thus, in the configuration d^5 or, equivalently d^5s^kp [$k=0$, 1, or 2], to transform from the scheme of d^5 considered as five electrons to the scheme of d^5 as five holes, it is only necessary to change the sign of those rows and corresponding columns based on d^5 with seniority number 3. This transformation leaves the eigenvalues invariant. Thus, when considering the configurations d^5s^kp alone, it is irrelevant whether the core d^5 is considered as comprising five electrons or as five holes in the d shell. Also in $(d+s)^5p$, $(d+s)^6p$, and $(d+s)^7p$ we can define d^5 in each case as either consisting of five electrons or as five holes in the d shell. However, after choosing the particular scheme for d^5 in d^5s^kp, it is necessary to be consistent when calculating the electrostatic interaction of d^5s^kp, the spin–orbit interaction of d^5s^kp and the interactions between configurations.

For $(d+s)^5p$, i.e., $d^5p+d^4sp+d^3s^2p$, in the interaction d^5p–d^4sp as given by (7) and in the interaction d^5p–d^3s^2p as given by (9), the configuration d^5p is considered as having five d electrons. Thus, using (7) and (9) for the interactions between configurations, it is necessary to be consistent and calculate d^5p also in the scheme of five electrons d.

For $(d+s)^6p$, it is equally convenient to choose d^5 of d^5sp as comprising five electrons or as five holes in the d shell. Since for the physical applications of $(d+s)^6p$ the main configurations are d^6p and d^5sp, it seems logical to consider d^5sp analogously to d^6p, i.e., as consisting of five holes in the d shell.

The matrix elements for the interaction between the configurations d^6p and d^5sp are obtained from (7) with the simplifications (11) and (13) for the matrices of H and J, respectively.

In all these relations, (19) (R III) is used to transform the coefficients of fractional parentage $\langle d^6\{|\,d^5\,d\rangle$ to $\langle d^4\,d\,|\}\,d^5\rangle$. In the first coefficient d^5 is defined in the same manner as d^6, i.e., consisting of five holes in the d shell.

The interaction between the configurations d^5sp and d^4s^2p is given by (8). The fractional parentage coefficients $\langle d^5\{|\,d^4\,d\rangle$ as tabulated in R III are calculated with d^5 being considered as five electrons in the d shell. Thus in the matrices of $H'(d^5sp$–$d^4s^2p)$, $J'(d^5sp$–$d^4s^2p)$, and $K'(d^5sp$–$d^4s^2p)$ as obtained from (8) and using (81) (R III) and Tables IV and XXII of R III, it is necessary to change the signs of all those matrix elements connecting terms of d^4s^2p

with those of d^5sp based on d^5 with seniority number 3.

For the configuration d^5sp, the only parameters which have nondiagonal matrix elements connecting states of seniority number 1 or 5 of d^5 with those based on d^5 with seniority number 3 are F_2', G_1', G_3', T', and ζ_d'. As the matrices of the parameters F_2' and ζ_d' for d^5sp are equal in magnitude and opposite in sign of the matrices of F_2' and ζ_d' for the complementary configuration $d^{-5}s^{-1}p$, it is evident that the *only* matrix elements of F_2' and ζ_d' are those connecting states based on d^5 with seniority numbers 1 or 5 and states based on d^5 with seniority number 3. Now, the fractional parentage coefficients $\langle d^5 \| \{ d^4 d \rangle$ appearing in (17) of I and (13) of II are tabulated in R III with d^5 considered as comprising five electrons. Thus, the matrix elements of F_2' and ζ_d' calculated with the aid of the fractional parentage coefficients given in Table IV (R III) must have their signs changed in order to comply with the choice of d^5 as consisting of five holes in the d shell. Similarly if the fractional parentage coefficients appearing in (21) of I are taken from Table IV (R III), the matrices of G_1' and G_3' are those of the configuration d^5sp, where the core d^5 is defined as consisting of five electrons. Then it is necessary to change the signs of all the matrix elements of G_1' and G_3' which connect terms based on d^5 with seniority numbers either 1 or 5 and terms based on d^5 with seniority number 3. In addition the interaction of the p electron with the closed d shell should be subtracted from (21) of I. This expression is given by $-10G_1' - 35G_3'$. Furthermore, also the matrix elements of T connecting terms of seniority number 1 or 5 with terms of seniority number 3 must have their signs changed. Finally, a minus sign must be introduced in the matrices of $G(d^6p\text{–}d^4s^2p)$ for terms of seniority number 2.

Since, in $(d + s)^7p$, the configurations d^7p and d^6sp are complementary configurations of d^3s^2p and d^4sp, respectively, it is logical to consider the configuration d^5s^2p as complementary to d^5p. Thus the core d^5 of d^5s^2p is defined as comprising five holes. By the same reasoning as for the interaction $d^6p\text{–}d^5sp$ of $(d + s)^6p$, the interaction $d^6sp\text{–}d^5s^2p$ as obtained from (8) with the simplifications (12) and (14) for $H'(d^6s\text{–}d^5s^2)$ and $J'(d^6sp\text{–}d^5s^2p)$, respectively, refers to five holes in the core d^5 of d^5s^2p. In the interaction $d^7p\text{–}d^5s^2p$ as given by (9), a minus sign must be introduced for terms of seniority number 3. If the matrices of d^5p are to be used for the configuration d^5s^2p then it is necessary to change the signs of the nondiagonal matrix elements connecting states of d^5 with seniority number 1 or 5 and those states with seniority number 3. As for d^5sp these matrices are of F_2, G_1, G_3, T, and ζ_d. The matrices of F_2 and ζ_d simply reverse their sign. For the matrices of G_1 and G_3 it is also necessary to subtract from the height $-10G_1 - 35G_3$ which represents the interaction of the p electron with the closed d shell.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

General Interaction Picture from Action Principle for Mechanics

L. M. GARRIDO
Instituto de Física Teórica, Barcelona, Spain

(Received 28 March 1968)

In this paper we consider a general action principle for mechanics written by means of the elements of a Lie algebra. We study the physical reasons why we have to choose precisely a Lie algebra to write the action principle. By means of such an action principle we work out the equations of motion and a technique to evaluate perturbations in a general mechanics that is equivalent to a general interaction picture. Classical or quantum mechanics come out as particular cases when we make realizations of the Lie algebra by derivations into the algebra of products of functions or operators, respectively. Later on we develop in particular the applications of the action principle to classical and quantum mechanics, seeing that in this last case it agrees with Schwinger's action principle. The main contribution of this paper is to introduce a perturbation theory and an interaction picture of classical mechanics on the same footing as in quantum mechanics.

1. INTRODUCTION

We present in this paper a general action principle for mechanics, valid for classical or quantum problems. From such a principle the equations of motion may be derived, but its main application is the possibility of deducing an interaction picture, valid quite generally, from which perturbation expansions can be obtained. In particular, of course, we get a perturbation method for the two kinds of mechanics mentioned above.

We look for the "intersection" of the various dynamical structures in a common formalism. This common abstract mathematical structure is that of the realizations of a Lie algebra $\mathfrak{L}$, by derivations in an associative linear algebra D. All dynamical theories can be unified in the above-mentioned manner, since they have enough features in common. We start from an initially very general presentation of the dynamical principle to obtain, later on, as realization of our principle, action principles for each one of the mentioned mechanics. But the main aim of this paper is the application of this technique to the evaluation of perturbations.[1] The elements of the Lie algebra are abstract mathematical entities isomorphically associated with the physical dynamical variables.

Let us examine the case for quantum mechanics. If we have only one irreducible representation of the "algebra of observables," all relevant information of the theory is contained in the algebraic structure alone. Hilbert space representations are not needed since they add nothing to our knowledge of the physical world: this is certainly the case when the number of degrees of freedom is finite. We may say, therefore, that for ordinary quantum mechanics, the purely algebraic approach should prevail. However, in quantum field theories we have infinitely many degrees of freedom, and it is well known that there exist, indeed, many inequivalent irreducible representations of the same algebra. Nevertheless, the differences between inequivalent representations of dynamics in quantum field theory are too fine and they do not have any physical importance. Any faithful representation of the algebra of observables will give the same physical results, and therefore, none of them is needed. Whether the number of degrees of freedom of quantum mechanics is finite or infinite, our discussion shows that the answer that we find is in favor of the purely algebraic approach. We conclude that all faithful representations are "physically" equivalent, even though they may be mathematically strong inequivalent, and conclude that none of them is needed.[2]

The vector space of the Lie algebra of the general dynamical structure of mechanics has a dual space whose elements are called states. The states determine the mapping of the Lie algebra $\mathfrak{L}$ onto the field of real numbers, which are the elements that can be compared with the physical reality. They correspond to the expectation values of the observables for a state—a vector in Hilbert space—that are commonly used in quantum mechanics. The selection of a particular (faithful) representation is a matter of convenience without physical implications. It may provide a more or less handy analytical apparatus.

We can find many mappings of a Lie algebra into the field of real numbers. It is, therefore, possible to define states in many different ways, and so we can have many kinds of mechanics from the same dynamical Lie algebra structure. To obtain classical or quantum mechanics we have to specify clearly what kind of mapping has to be used for each case. However, a Lie algebra may have additional mappings,

[1] E. C. Sudarshan, *Lectures in Theoretical Physics, 1961 Brandeis Summer Institute* (W. A. Benjamin, Inc., New York, 1962), p. 144.

[2] R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964).

unexplored by physics as yet, into the field of real numbers, that eventually may generate another kind of mechanics. Of course, we can compare, and we here do so, the action principle presented in this paper only with action principles and perturbation methods for the two kinds of mechanics mentioned that are the ones used in physical problems. But we hope that the action principle presented is valid more generally, even though we are not able at the present time to check these further applications.

We do not study in this paper classical or quantum statistical mechanics, because we are essentially concerned with dynamics and they offer nothing new to the action principle that we present. Statistical mechanics differs from other kinds of mechanics not in the action principle but in the mapping of the elements of the Lie algebra into the field of real numbers; that is done by means of density operators or distribution functions, kinematical aspects to which we do not pay special attention here.

Dynamical variables and states are duals to each other. In a most general sense, states are the mappings of the Lie algebra onto the field of real numbers. Besides the action principle, which is purely dynamical, there is another aspect in all mechanics—namely, the choice of admissible states belonging to the dual space of the dynamical Lie algebra—a kinematical aspect that limits the mappings onto the field of real numbers which have physical meanings. Generally, there are additional requirements, most frequently imposed to preserve the meaning of probability, so that not every element of the vector dual space is an admissible physical state. The admissible states form a manifold that usually has to be convex, in order not to have negative probabilities. This manifold of states is in general not a subspace because the convexity conditions limit the number of admissible linear combinations that one may make. The natural determination of the admissible manifold of states imposes additional conditions to the Lie algebra $\mathfrak{L}$, or to its realizations into another linear associative algebra D, by means of derivations.

To determine the convex manifold of states, which is physically admissible, further additional information not included in the Lie algebra specifications is needed. The convex manifold of states must be so chosen that, in a Schrödinger-like picture of dynamics, the changes compatible with the action principle will not throw them out of the admissible manifold.

We do not study in this paper a Schrödinger-like picture of dynamics but rather a Heisenberg-like picture of dynamics deduced from the action principle that we here introduce.

There are dynamical theories which have to be Lorentz-covariant. Physically we have to require that for every element of the Poincaré group an automorphism of the algebra has to be introduced. The requirement that the Lorentz transformations be represented by unitary operators in Hilbert space for quantum mechanics is a very powerful restriction that may not be completely justified on physical grounds,[3] and in the same way, intimately connected with the action principle are questions about symmetry properties of the physical system. This means that a Lie algebra may have additional, unexplored structural features, the existence of which is inherent in the special form of its action element.

In Sec. 2 we present as a postulate the general action principle for a quite general mechanics without specifying whether it is classical or quantum mechanics. The action principle is written by means of the elements of an abstract algebra that is a Lie algebra. We examine immediately which is the physical meaning of all the properties of the Lie bracket multiplication. We apply the action principle to obtain the equations of motion and to arrive at an interaction picture in a general scheme of mechanics. Later on we examine the consistency requirements between both applications—for deduction of the equations of motion and for the evaluation of perturbations—of the action principle.

In Sec. 3 we make concrete the realization of the action principle into the algebra that is proper for classical mechanics. A perturbation theory valid for classical mechanics is presented as deduced from our action principle. In Sec. 4 we do the same for quantum mechanics; in particular we observe how Schwinger's action principle can be deduced from the action principle postulated here.

We conclude this paper in Sec. 5 with a discussion of the possibility of extending the application of the present action principle to other mechanics that may eventually be derived.

The main contribution of this paper is to introduce an interaction picture, and from it a perturbation theory of classical mechanics on the same footing as in quantum mechanics.

2. ACTION PRINCIPLE

We plan to introduce the action principle as a postulate by means of the elements of a Lie algebra, which we designate by $\mathfrak{L}$. For any three elements A, B, C, such that

$$A, B, C \in \mathfrak{L}$$

[3] R. Haag, *Lectures in Theoretical Physics. 1964* (University of Colorado Press, Boulder, Colorado), p. 107.

of the Lie algebra, the distributive and nonassociative product of any two of the elements of $\mathfrak{L}$, which we write down as $[A, B]$, has to satisfy the following properties to generate a Lie algebra:

$$[A, B] = -[B, A], \tag{2.1}$$

that is, the antisymmetry condition, and

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0, \tag{2.2}$$

called the Jacobi identity.

Later on we examine which is the physical meaning of these two conditions on the elements of the algebra. Such a study gives us the reasons why we choose precisely a Lie algebra as the mathematical structure most fit to postulate the action principle. (The symbol [,] is called a Lie bracket.)

For any element A of the algebra, the action principle that we postulate is written as

$$\delta A = dA - \partial A = [\delta W, A], \tag{2.3}$$

where δA is the total infinitesimal variation of the element A of the algebra in relation to a certain parameter λ of a certain class of parameters that we study later. We should write more carefully as follows:

$$\delta A \equiv \delta_\lambda A, \tag{2.4}$$

notation that we use as it is needed. From the total variation dA, we have to subtract ∂A, which is the change in A associated with the explicit appearance in A of the parameter λ, since the latter cannot be produced by any action principle, but can be deduced immediately once we are given the explicit dependence of A on the parameter λ: it corresponds to the partial derivative of A with respect to the parameter λ. Without loss of generality and in order not to complicate the equations, we always suppose that the elements A of the Lie algebra do not depend explicitly on the parameters λ, so that

$$\partial A = 0. \tag{2.5}$$

The difference $\delta A = dA - \partial A$ is always the dynamical variation of the element A of the algebra, equal to the total variation less the explicit variation.

W is also an element of the algebra which plays a very special role and which we call Action. We study the general properties of W for any mechanics. The concrete specification of W depends on the kind of mechanics that we are considering and, more specifically, on the problem that we study. We call δW the variations of W in relation to a parameter λ of a class of parameters, some examples of which are presented later on. The elements δW are such that

$$\delta W \in \mathfrak{L}. \tag{2.6}$$

As we saw, we designate by [,] the combination or multiplication law for any ordered pair of elements of the algebra. The actual nature of the bracket [,] has to be specified for each kind of mechanics, as we see later on.

The requirement that a Lie bracket, the multiplication of the elements of the Lie algebra, be always expressible as a linear combination of the elements of the Lie algebra by means of the structure constants ensures, according to the action principle presented above, that the variations of any dynamical symbol are a linear combination of these same elements. Therefore, structure constants govern the dynamics.

We should present now the reasons for choosing Lie algebras to express the general action principle. This Lie algebra contains two elements W and $\mathcal{H}$ (so that $W, \mathcal{H} \in \mathfrak{L}$), called respectively Action and Hamiltonian of the system. The time variation of the Action yields the Hamiltonian, whose Lie bracket with any element of the algebra provides us with the dynamical time derivative, since the explicit time derivative of an element of the algebra cannot be generated by a Lie algebra bracket. Generally any element of the algebra A generates a certain dynamical variation of all the other elements of the algebra in relation to a certain parameter. We impose the physical condition that no element can produce a dynamical variation of itself, a condition that implies that the Lie bracket of an element A with itself is always zero:

$$[A, A] = 0. \tag{2.7}$$

If A and B are elements of $\mathfrak{L}$, then since an algebra is a vector space, $A + B$ will also be an element of $\mathfrak{L}$. From the above result and the fact that the combination relation of any algebra is distributive, we have

$$\begin{aligned} 0 &= [A + B, A + B] \\ &= [A, A] + [A, B] + [B, A] + [B, B] = 0, \end{aligned}$$

which gives

$$[A, B] = -[B, A]. \tag{2.8}$$

Therefore, the antisymmetry requirement of the Lie-bracket multiplication is equivalent to the physical condition that no element can produce the dynamical variations of itself.

Next, let us see where the Jacobi identity comes from. We should indeed require that the bracket-composition law be consistent with the dynamical variations of the elements of our algebra in relation to any parameter. This requirement is equivalent to the statement that any functional relationship, such as $C = [A, B]$, existing between any three elements

of our algebra A, B, C for a certain value of the parameters that determine the dynamical variations, should be preserved for any other value of these parameters. For the sake of concreteness, we consider the dynamical time evolution of the system produced by $\mathcal{H}$, the Hamiltonian of the system, which is an element of the algebra we are considering. Then the elements A, B, C, $\mathcal{H}$ are considered at an instant of time t, and they satisfy

$$C = [A, B] \tag{2.9}$$

at time instant t. We would like to find out the requirements which our algebra has to satisfy in order that this relation be also valid at another time instant $t + dt$, infinitesimally different from t. The element A becomes $A + dt[A, \mathcal{H}]$ as deduced from an action principle as we see later. We would have similar expressions for the changes of B and C. In particular, C becomes $C + dt[C, \mathcal{H}]$. But from the relationship between C and the bracket $[A, B]$ that we want to preserve for the time instant $t + dt$, we should have

$$\begin{aligned} C + dt[C, \mathcal{H}] &= [A, B] + dt[[A, B], \mathcal{H}] \\ &= [A + dt[A, \mathcal{H}]], [B + dt[B, \mathcal{H}]]. \end{aligned} \tag{2.10}$$

If we keep only terms linear in dt and use the antisymmetry of the brackets already assumed in our algebra, then the Jacobi identity between A, B, and $\mathcal{H}$ immediately follows. If we had considered other kinds of dynamical variations, we would have obtained in a similar manner the Jacobi identity among any three elements of our algebra. This result is completely general, since the Lie-bracket multiplication is the only combination law to obtain from any pair of elements A, B, and a third element C.

We consider, therefore, that the Jacobi identity expresses the consistency between the algebra whose elements describe the physical system and the action principle that we have presented; i.e., the Jacobi identity guarantees that the variations of the elements of the algebra compatible with the dynamical action principle do not throw these elements out of the algebra.

We have, therefore, to write the action principle between elements of a Lie algebra in order that the dynamical evolution of the system produce new elements within the same algebra. We remember that we have used a Heisenberg-like picture of the dynamics of a system to arrive at these conclusions.

Besides the general form of the action principle as a Lie bracket, the practical basis for the applications of this dynamical principle is the fact that there exists a class of parameters λ such that the variations $\delta_\lambda W$ are obtained by appropriate variation of a single element W of the Lie algebra. The action principle must be complemented by the explicit specification of such a class.

Of the whole class of variation parameters that can be considered, we study here only the instances when λ is the time t of the system and, secondly, the case in which λ is the coupling parameter g between two systems. Variations with respect to the time yield the equations of motion, while when we change the coupling parameter infinitesimally we get a perturbation expansion that, as we said, is the main aim of this paper.

Let us consider the temporal evolution of the system. We designate by $A(t)$ any element of the algebra at instant t. There is an automorphism between the set of elements $A(t)$ and those of $A(t^1)$ considered at another instant t^1 of time. The action principle, in the form that we have presented it, implies that the dynamical time evolution of any element of the algebra is obtained by multiplying such an element by another δW of the same algebra, i.e., by an element δW evaluated at the same instant of time t. An element δW evaluated at another instant of time t^1 cannot generate according to (2.8) the time evolution at instant t. This deduction from the action principle (2.3) is equivalent to the principle of stationary action. It states that W, whose meaning is $\delta(W) = \delta W$, must be stationary with respect to variations at another time instant t^1, $t^1 \neq t$, since δW can only contain elements of the algebra associated with instant t. Therefore, we write

$$\delta_t W = \mathcal{H}(t)\delta t, \tag{2.11}$$

where $\mathcal{H}(t)$ is called the Hamiltonian of the system. The fact that the dynamical temporal variations of the elements of the algebra at an instant t can only be generated by an element of the algebra evaluated at the same time instant t, implies the existence of equations of the motion. The general equation of motion is

$$\delta A/\delta t = [A, \mathcal{H}]. \tag{2.12}$$

Since $[\mathcal{H}, A]$ is a linear combination of elements of the Lie algebra $\mathfrak{L}$, determined by the structure constants, the same equation of motion can be applied to $\delta A/\delta t$. We get

$$\delta^2 A/\delta t^2 = [[A, \mathcal{H}], \mathcal{H}], \tag{2.13}$$

and, in general,

$$\delta^n A/\delta t^n = [\cdots, [A, \mathcal{H}], \mathcal{H}, \cdots, \mathcal{H}] \tag{2.14}$$

with n multiplication brackets.

Applying Taylor's theorem, we can write

$$A(t) = \{e^{-t\mathcal{H}}A(0)e^{t\mathcal{H}}\}, \tag{2.15}$$

where the braces indicate that the expression is only symbolic in the sense that its only meaning is

$$A(t) = A(0) + \frac{t}{1!}[A(0), \mathcal{H}] + \frac{t^2}{2!}[[A(0), \mathcal{H}], \mathcal{H}] + \cdots. \tag{2.16}$$

Indeed, the exponentials of $\mathcal{H}$ in (2.15) are not defined, since the only multiplication that we have introduced in the algebra is the Lie-bracket multiplication according to which the powers of any element of the algebra are identically zero, given the antisymmetry of the brackets required by physical conditions. Indeed, for instance, so far

$$\mathcal{H}^2 \equiv [\mathcal{H}, \mathcal{H}] = 0. \tag{2.17}$$

The fact that the expression (2.15) is only symbolic is a serious inconvenience for the practical applications of the action principle, since we do not have an analytical apparatus to use in our calculations. This is the reason why we have to introduce realizations of the Lie algebra defining a new algebra and a new product (,) that, since it does not enter into the action principle, does not have to be antisymmetric as it is required on physical grounds for the Lie-bracket product [,]. Then, powers of an element are defined by means of this new kind of product. This is the reason why the dynamical Lie algebras are realized by means of derivations though, evidently, these realizations are not required by the physical content of the theory; they simply are convenient ways of performing the calculations that appear in the action principle and of mapping the Lie algebra into the field of real numbers.

Let us consider the introduction of a general interaction picture to study perturbations. We consider that the action element W can be divided into two parts coupled by the parameter g, so that

$$W = W_0 + gW_1, \tag{2.18}$$

and we want to obtain the change of any element of the algebra when the coupling parameter changes from g to $g + \delta g$. Action principle (2.3) yields

$$\delta_g A = [\delta_g W, A] = [W_1, A]\delta g, \tag{2.19}$$

where the Lie bracket $[A, W_1]$ has to be evaluated for the value $\lambda = g$ of the coupling parameter. W_0 is the unperturbed action; it corresponds to $g = 0$. The action corresponding to $g = 1$, $W = W_0 + W_1$, is the fully perturbed action, since we consider W_1 to be the perturbation.

We need to study the boundary conditions for the application of the perturbation. Undoubtedly the action W should contain two labels to indicate when the interaction begins and when it ends. So

$$W \equiv W(t, t_0), \tag{2.20}$$

where t_0 is the instant when the perturbation starts and t the final moment of action of the perturbation. The physical consistency requirement implies that

$$W(t_0, t_0) = 0 \tag{2.21}$$

and that

$$W(t, t_0) + W(t_0, t_1) = W(t, t_1), \tag{2.22}$$

from which we deduce

$$W(t, t_0) = -W(t_0, t). \tag{2.23}$$

The action element $W(t, t_0)$ evidently possesses the form

$$W(t, t_0) = \int_{t_0}^{t_1} dt_1 L(t_1), \tag{2.24}$$

where $L(t_1)$ is the Lagrangian. As we see later, the action $W(t, t_0)$ has to be varied in relation to the upper limit t in order to obtain the Hamiltonian at the instant, i.e.,

$$\delta_t W(t, t_0) = -\mathcal{H}(t)\delta t. \tag{2.25}$$

Taking the action principle (2.3) to evaluate perturbations, we deduce that

$$\delta^n A/\delta g^n = 0, \quad \text{if} \quad t = t_0 \quad \text{for any } n, \tag{2.26}$$

since $W(t_0, t_0) = 0$. From here we deduce that, as assumed before, t_0 is the instant when the perturbation starts to act and that, therefore, the perturbation acts during the interval $t - t_0$.

If we write the time labels explicitly, Eq. (2.19) has the following form:

$$\frac{\delta A(t)}{\delta g} = [W_1(t, t_0), A(t)] = \int_{t_0}^{t} dt_1 [L(t_1), A(t)]. \tag{2.27}$$

From here we also have, as before,

$$\frac{\delta^2 A(t)}{\delta g^2} = \int_{t_0}^{t} dt_1 \left[\frac{\delta L(t_1)}{\delta g}, A(t)\right] + \int_{t_0}^{t} dt_1 \left[L(t_1), \frac{\delta A(t)}{\delta g}\right], \tag{2.28}$$

which is a procedure that can be continued so as to evaluate $[\delta^n A(t)]/(\delta g^n)$ for any value of n.

The explicit expression for the element A for $g = 1$, i.e., fully perturbed, is obtained from the same element A for $g = 0$, i.e., from the unperturbed element, by means of a Taylor's expansion in powers of $\Delta g = 1$,

and so

$$A(t)|_{g=1} = A(t)\,|_{g=0} + \frac{1}{1!}\frac{\delta A(t)}{\delta g}\Big|_{g=0} + \cdots + \frac{1}{n!}\frac{\delta^n A(t)}{\delta g^n}\Big|_{g=0} + \cdots, \quad (2.29)$$

which is an expansion that can also be written in a symbolic way by means of exponentials of W_1. This is the general interaction picture, since here all the successive Lie brackets have to be evaluated for $g = 0$, i.e., for the elements of the Lie algebra calculated for the unperturbed motion generated by W_0.

To apply the above formula to a perturbation expansion we have to suppose that the motion of the system generated by the unperturbed action W_0 has been solved exactly. Then we can calculate exactly the different successive Lie brackets that appear in (2.29). The term in this formula that contains n times the perturbing action W_1 is the nth perturbation. Indeed, to apply expression (2.29) to a concrete perturbation problem, we have to define for each kind of mechanics the Lie bracket. But, undoubtedly, we have written a general perturbation expansion.

The combination or multiplication law of any two elements A and B of the Lie algebra $\mathfrak{L}$, by means of which the action principle is introduced, is written as $[A, B]$. We see in the applications that, as a matter of fact, such a product becomes the Poisson bracket or the commutator between any two elements, respectively, in each one of the mechanics in which the action principle is applied.

In the vector space of the elements of the abstract Lie algebra $\mathfrak{L}$ we define a second combination law of the two elements, that we design by $(\ ,\ )$, which maps pairs of elements of the dynamical abstract Lie algebra, A and B, into another such element (A, B), under which the vector space becomes an associative algebra D. This implies that the new product (A, B) is also distributive. We also further require that the two product operations satisfy

$$[(A, B), C] = (A, [B, C]) + ([A, C], B) \quad (2.30)$$

for any three elements of the Lie algebra.[4] This property is referred to by saying that the Lie bracket is a derivation in a linear associative algebra with the product (A, B).

In a linear associative algebra, powers of an element are uniquely defined. The associative and distributive product $(\ ,\ )$ is often referred to as the ordinary product. As a matter of fact, the product $(\ ,\ )$ is either the ordinary product of analytic functions in classical mechanics or the ordinary product of operators in quantum mechanics.

By virtue of the derivation property of the Lie bracket, it follows that algebraic relations among the elements of the Lie algebra, involving either the ordinary product $(\ ,\ )$ or the Lie bracket $[\ ,\ ]$, are preserved by infinitesimal transformations.

A Lie algebra provides only an abstract framework for the dynamical properties of a physical system, and even if this framework is supplemented by the dual space of physical states, it is not enough for the complete and practical specification of the physical situation. We have to introduce also the additional structure of an associative algebra D, and an explicit realization of the Lie algebra $\mathfrak{L}$, by derivations in this associative algebra D. And so, as we see, for classical mechanics we use analytic functions where $(\ ,\ )$ is the ordinary product of the same and $[\ ,\ ]$ is the Poisson bracket; but for quantum mechanics we introduce operators in Hilbert space where $(\ ,\ )$ is now the ordinary product of the same operators, while $[\ ,\ ]$ is proportional to the commutator. For instance, powers of the dynamical variables will, in general, have a meaning in the explicit realization of the algebra not being defined in the algebra $\mathfrak{L}$ itself.

Our dynamical scheme is as follows. We have an abstract Lie algebra $\mathfrak{L}$, whose elements constitute the dynamical variables, and a concrete linear associative algebra D, which furnishes a realization of $\mathfrak{L}$ by derivations.

We note in passing that classical and quantum mechanics, in order to be discussed, fall within this characterization. As a matter of fact, in Sec. 3 of this paper we examine the case for classical mechanics while in Sec. 4 we study, from the viewpoint of this paper, quantum mechanics.

The most important point that we want to make clear in this paper is that considering all different mechanics as different realizations of one and the same algebra $\mathfrak{L}$, we obtain a unified apparatus to formulate the *dynamical* properties of all mechanical systems, to introduce a general *interaction picture* for dynamics, and to deduce a general method for evaluating *perturbations* in all kinds of mechanical systems.

The entities that form the associative algebra D, in which we obtain realizations of the Lie algebra $\mathfrak{L}$ by derivations, have composition laws of their own, only part of which will reflect, in a homomorphic manner, the composition table of $\mathfrak{L}$. In general, we are able to define functions of the representatives of the elements, additional relations that it may not be possible to define in the original algebra $\mathfrak{L}$, and that give rise to elements that do not belong to the algebra

[4] J. P. Serre, *Lie Algebras and Lie Groups* (W. A. Benjamin, Inc., New York, 1965).

$\mathfrak{L}$. This additional content of the realization has always a definite physical meaning, allowing the introduction of physical degrees of freedom, and so it is not only a matter of mathematical freedom.

A canonical transformation is a mapping which leaves Lie-bracket relations invariant; they are essentially automorphisms of the algebra. The dynamical evolutions of the elements of the algebra are canonical transformations. The set of all automorphisms of a given Lie algebra constitutes the corresponding Lie group, which, accordingly, consists of sets of canonical transformations.

All these Lie algebras contain an identity I, which is an element whose Lie bracket with any other element of the algebra is zero. Normalization of the states is achieved, requiring that the identity be mapped into the real number 1. In this paper we do not study the mappings of the Lie-algebra elements into the field of real numbers, since we are mostly concerned with dynamical questions and not with the states.

We want to add, however, that the states can be characterized in classical and quantum mechanics in a very similar manner. From the physical viewpoint, the possibility to obtain this lies in the form that Ehrenfest gave to the principle of correspondence: the expectation values in quantum mechanics of dynamical operators obey the same equations of motion as the corresponding classical dynamical variables.

3. CLASSICAL MECHANICS

The action principle that we have established as a postulate is

$$\delta A = [\delta W, A], \tag{3.1}$$

where by [,] we indicate the Lie-bracket multiplication.

In classical mechanics we have to introduce the sets of canonical conjugate variables q_k, p_k, where $k = 1, 2, 3, \cdots, n$, by means of which we define the Poisson bracket between two analytical functions U and V of the sets q_k, p_k that we designate by $[U, V]_c$

$$[U, V]_c = \sum_k \left(\frac{\partial U}{\partial q_k} \frac{\partial V}{\partial p_k} - \frac{\partial U}{\partial p_k} \frac{\partial V}{\partial q_k} \right), \tag{3.2}$$

where the subscript c comes from classical mechanics.

Our action principle is translated into classical mechanics when the Lie-bracket multiplication is the Poisson bracket between analytic functions of the canonical set of conjugate variables, as follows:

$$\delta A = [\delta W, A] = [\delta W, A]_c \tag{3.3}$$

that, when A represents the variables q_k and p_k, yield

$$\delta q_k(t) = \frac{-\partial \delta W(t, t_0)}{\partial p_k(t)}, \quad \delta p_k(t) = \frac{\partial \delta W(t, t_0)}{\partial q_k(t)}, \tag{3.4}$$

which are equations of motion already obtained before.[5]

To simplify notation, when we deal with classical mechanics, the notation indicating Lie bracket will denote the Poisson bracket, i.e., we do not write from now on a special sign to specify that in classical mechanics the Lie bracket is interpreted always as the Poisson bracket. We see that the elements of the Lie algebra for classical mechanics are represented by analytic functions where the multiplication (,) is simply ordinary multiplication of functions and the Lie-bracket multiplication [,] is the Poisson bracket between the elements that are multiplied, with the notation $(u, v) = uv$. We see quite easily that relation (2.30) is satisfied between these two kinds of products.

As has been shown,[5] the temporal evolution of the physical system is obtained from action principle (3.1) or its equivalent (3.4). Indeed, to see this fact we study complete variations of the action integral that corresponds to an intrinsic variation $\Delta q(t)$ of the dynamical variables and to a change of the upper limit of the action integral

$$\delta W(t, t_0) = \int_{t_0}^{t+\delta t} L(\bar{q}(t_1), \dot{\bar{q}}(t_1))\, dt_1 - \int_{t_0}^{t} L(q(t_1), \dot{q}(t_1))\, dt_1 . \tag{3.5}$$

The intrinsic variation

$$\Delta q(t_1) = \bar{q}(t_1) - q(t_1) \tag{3.6}$$

is supposed to be zero at $t_1 = t_0$. To evaluate the complete variation of the dynamical variable at time t, we have to add to $\Delta q(t)$ the variation due to the shift $\bar{t} = t + \delta t$ of the upper limit

$$q(\bar{t}) = q(t) + \delta t \dot{q}(t), \tag{3.7}$$

so that its complete variation is

$$\delta q(t) = \Delta q(t) + \delta t \dot{q}(t). \tag{3.8}$$

Now the evaluation of $\delta W(t, t_0)$ is straightforward. We get

$$\delta W(t, t_0) = -\delta t(\mathcal{H} q(t), p(t)) + p(t)\delta q(t), \tag{3.9}$$

where the Hamiltonian $\mathcal{H}(q, p)$ is defined, as is usually done, by

$$-\mathcal{H} = L - p\dot{q}. \tag{3.10}$$

[5] L. M. Garrido, J. Math. Anal. Appl. 3, 295 (1961).

The time evaluation corresponds to

$$\delta_t W(t, t_0) = -\delta t \mathcal{H}(q(t), p(t)), \tag{3.11}$$

the case in which the action principle yields

$$\dot{q} = \frac{\delta_t q(t)}{\delta t} = \frac{\partial \mathcal{H}}{\partial p}, \quad -\dot{p} = -\frac{\delta_t p(t)}{\delta t} = \frac{\partial \mathcal{H}}{\partial q}, \tag{3.12}$$

which are Hamilton's equations of motion.

The variation

$$\delta_q W(t, t_0) = p(t)\delta q(t) \tag{3.13}$$

gives the kinematic independence of $q(t)$ and $p(t)$, since then

$$\delta_q q(t) = \frac{\partial[p(t)\delta q(t)]}{\partial p(t)} = \delta q(t),$$

$$-\delta_q p(t) = \frac{\partial p(t)}{\partial q(t)} \delta q(t) = 0. \tag{3.14}$$

So far we have done nothing new. The preceding formulas are well known and so the postulated action principle appears as a different way of writing the equations of motion. Such a postulate is only meaningful if we can also obtain from it other results beyond the equations of motion. This is the case, since our action principle yields also perturbation theory and provides a means of writing an interaction picture for classical mechanics. And this is what we do next.

We would like to study the system whose action suffers the effect of a perturbing Lagrangian so that the new action becomes

$$W(t, t_0) = W^{(g)}(t, t_0) = W_0(t, t_0) + gW(t, t_0)$$

$$= \int_{t_0}^{t} L^{(g)}(q_g(t'), p_g(t'))\, dt', \tag{3.15}$$

where

$$L^{(g)} = L_0 + gL_1. \tag{3.16}$$

Here t_0 is the time instant when the perturbation starts out. For the study of perturbations we fix t_0 and t, but change the coupling parameter g from g to $g + \delta g$, so that

$$\delta W^{(g)}(t, t_0) = \delta g W^{(g)}(t, t_0), \tag{3.17}$$

where

$$W_1^{(g)}(t, t_0) = \int_{t_0}^{t} L_1(q_g(t'), p_g(t'))\, dt_1, \tag{3.18}$$

since in the evaluation of $W_1^{(g)}(t, t_0)$ we have to use the canonical conjugate variables evaluated at the value g of the coupling parameter q_g and p_g in order to calculate the Poisson bracket that appears in classical mechanics. Usually, however, the Lagrangian is not given as presented above but in terms of a set of generalized coordinates q and the time derivatives of the same $\dot{q}$. Using the definition of generalized momenta canonically conjugate to a given generalized coordinate, we can eliminate the time derivative of the coordinate $\dot{q}$ and write the Lagrangian as a function of sets of canonical conjugate variables q and p. If this variation is applied to the action principle, we obtain

$$\delta_g A_g = \delta g[W_1^{(g)}(t, t_0), A_g] \tag{3.19}$$

where now $A \equiv A_g$, i.e., the element of the Lie algebra depends on the coupling parameter. To have a clear idea of how to calculate the Poisson bracket that appears here we work out two examples later on.

Since the perturbation is analytic in the coupling parameter g, we make use of Taylor's expansion

$$A_1 \equiv A_g|_{g=1} = A_0 + \frac{1}{1!}\frac{\delta_g A_g}{\delta g}\bigg|_{g=0} + \frac{1}{2!}\frac{\delta_g^2 A_g}{\delta g^2}\bigg|_{g=0}$$

$$+ \cdots + \frac{1}{n!}\frac{\delta_g^n A_g}{\delta g^n}\bigg|_{g=0} + \cdots, \tag{3.20}$$

where the unperturbed system is obtained for $g = 0$ and the fully perturbed motion corresponds to $g = 1$. We have to evaluate $(\delta_g^n A_g)/\delta g^n$ from our action principle starting from

$$\frac{\delta A}{\delta g} = [W_1, A]. \tag{3.21}$$

In this way we obtain the general interaction picture for classical mechanics and a procedure to evaluate perturbations in classical mechanics to any order in the perturbing action.

The perturbation method that results from the action principle (3.3) in classical mechanics does not follow the same steps as the technique deduced before for Hamiltonian's equations of motion.[6,7]

We now clarify most of our ideas, working out some examples that indicate how the perturbation techniques deduced from the action principle postulated above can be applied to classical mechanics. The action integral is defined as follows for the unperturbed motion:

$$W_0(t, t_0) = \int_{t_0}^{t} L_0(q_0, \dot{q}_0, t_1)\, dt_1. \tag{3.22}$$

The Lagrangian may depend on time explicitly. For simplicity, we limit ourselves to the case when the motion of the system is properly described by a single generalized coordinate and its time derivative, but the principle is valid for motions with any fixed number of generalized coordinates.

[6] L. M. Garrido, Proc. Phys. Soc. (London) **76**, 33 (1960).

[7] L. M. Garrido and F. Gascon, Proc. Phys. Soc. (London) **81**, 1115 (1963).

If we call $p_0 \equiv \partial L_0(q_0, \dot{q}_0)/\partial \dot{q}_0$ the canonical conjugate momentum, we can eliminate the time derivative $\dot{q}_0$ of the coordinate and write the Lagrangian as a function of q_0 and p_0. Using the same symbol for the new function, we get

$$L_0(q_0, p_0) = L_0(q_0, \dot{q}_0). \tag{3.23}$$

The action integral becomes

$$W_0(t, t_0) = \int_{t_0}^{t} L_0(q_0, p_0, t_1)\, dt_1, \tag{3.24}$$

which cannot be evaluated until the equations of motion are solved. Let us suppose that we have solved such equations exactly, equations corresponding to the unperturbed motion, and have written their solutions in terms of the boundary values at the time origin

$$\begin{aligned} q_0 &\equiv q_0(t) = q_0(q_0(0), p_0(0), t), \\ p_0 &\equiv p_0(t) = p_0(q_0(0), p_0(0), t). \end{aligned} \tag{3.25}$$

By means of these expressions we can evaluate q_0 and p_0 at any time $t = t'$, and consider $q_0(t')$ and $p_0(t')$ as boundary values, so that we are able to write the solutions of the unperturbed equations of motion as

$$\begin{aligned} q_0 &\equiv q_0(t) = q_0(q_0(t'), p_0(t'), t - t'), \\ p_0 &\equiv p_0(t) = p_0(q_0(t'), p_0(t'), t - t'). \end{aligned} \tag{3.26}$$

Now we define derivatives with respect to the boundary values at any instant, and evaluate expressions like $[\partial p_0(t)]/[\partial p_0(t')]$, which is the derivative of a function with respect to the same function at any other time.

Evidently

$$\begin{aligned} \frac{\partial p_0(t)}{\partial p_0(t)} &= 1, \quad \frac{\partial p_0(t)}{\partial q_0(t)} = 0, \\ \frac{\partial q_0(t)}{\partial q_0(t)} &= 1, \quad \frac{\partial q_0(t)}{\partial p_0(t)} = 0. \end{aligned} \tag{3.27}$$

To be more concrete, we evaluate derivatives with respect to boundary values when the unperturbed system is the harmonic oscillator, whose Lagrangian is

$$L_0(q_0, p_0) = \tfrac{1}{2}p_0^2 - \omega^2 \tfrac{1}{2} q_0^2. \tag{3.28}$$

The equation of motion yield the following solutions:

$$\begin{aligned} q_0(t) &= q_0(0) \cos \omega t + \frac{p_0(0)}{\omega} \sin \omega t, \\ p_0(t) &= -q_0(0)\omega \sin \omega t + p_0(0) \cos \omega t, \end{aligned} \tag{3.29}$$

and therefore

$$\begin{aligned} \frac{\partial q_0(t)}{\partial q_0(t')} &= \cos \omega(t - t'), \quad \frac{\partial p_0(t)}{\partial q_0(t)} = -\omega \sin \omega(t - t'), \\ \frac{\partial q_0(t)}{\partial p_0(t')} &= \frac{1}{\omega} \sin \omega(t - t'), \quad \frac{\partial p_0(t)}{\partial p_0(t')} = \cos \omega(t - t'). \end{aligned} \tag{3.30}$$

Making use of the above calculations, let us evaluate the first term of the perturbation series of the harmonic oscillator perturbed by the Lagrangian

$$L_1(q, p) = -\mu^2 \tfrac{1}{2} q^2, \tag{3.31}$$

when $t_0 = 0$.

With the help of expansions obtained before, we get

$$\begin{aligned} q(t) &= q_0(t) + \frac{\mu^2}{2} \int_{t_0}^{t} \frac{\partial q_0^2(t_1)}{\partial p_0(t)}\, dt_1 + \cdots \\ &= q_0(t) + \frac{\mu^2}{\omega} \int_0^t q_0(t_1) \sin \omega(t_1 - t)\, dt_1 + \cdots \\ &= q_0(t) - \frac{\mu^2}{\omega} t q_0(0) \sin \omega t + \frac{\mu^2}{2\omega^2} t p_0(0) \cos \omega t \\ &\quad - \frac{\mu^2}{2\omega^3} p_0(0) \sin \omega t + \cdots, \\ p(t) &= p_0(t) - \frac{\mu^2}{2} \int_0^t \frac{\partial q_0^2(t_1)}{\partial q_0(t)}\, dt_1 + \cdots \\ &= p_0(t) - \mu^2 \int_0^t q_0(t_1) \cos \omega(t_1 - t)\, dt_1 + \cdots \\ &= p_0(t) - \frac{\mu^2}{2} t q_0(0) \cos \omega t - \frac{\mu^2}{2\omega} t p_0(0) \sin \omega t \\ &\quad - \frac{\mu^2}{2\omega} q_0(0) \sin \omega t + \cdots. \end{aligned} \tag{3.32}$$

As a final application, we deduce from our action principle the definition of the Poisson bracket introduced by Peierls[8] for the nonrelativistic case. To define the Poisson bracket between $A(q, p)$ and $B(q, p)$ at time $t = T$, Peierls introduces a perturbing Lagrangian

$$L_1 = A(q(t), p(t))\delta(t - T), \tag{3.33}$$

where $\delta(t - T)$ is Dirac's delta, and considers an infinitesimal variation of the coupling parameter around $g = 0$. He then evaluates variations corresponding to two-boundary conditions, $t_0 = \infty$ and $t_0 = -\infty$, called advanced and retarded perturbations, respectively. Correspondingly, we have to calculate the changes induced in $B(q, p)$ that we call respectively $\delta^- B$ and $\delta^+ B$. Peierls' definition of the

[8] R. Peierls, Proc. Roy. Soc. (London) **A214**, 143 (1952).

Poisson bracket $[A, B]$ is

$$[A, B] = \lim_{\delta g \to 0} \frac{1}{\delta g} (\delta^+ - \delta^-)B. \tag{3.34}$$

To show the validity of Peierls' definition is quite easy if we utilize the action principle that we have postulated, since then

$$\lim_{\delta g \to 0} \frac{1}{\delta g} (\delta^+ - \delta^-)q(T)$$
$$= -\left[\partial \int_{-\infty}^{\infty} A(q(t_1), p(t_1))\delta(t_1 - T)\, dt_1\right] \Big/ [\partial p(T)]$$
$$= -\frac{\partial A(q(T), p(T))}{\partial p(T)},$$
$$\lim_{\delta g \to 0} \frac{1}{\delta g} (\delta^+ - \delta^-)p(T) = \frac{\partial A(q(T), p(T))}{\partial q(T)}, \tag{3.35}$$

and therefore

$$\lim_{\delta g \to 0} \frac{1}{\delta g} (\delta^+ - \delta^-)B(q(T), p(T))$$
$$= \frac{\partial B}{\partial p}\frac{\partial A}{\partial q} - \frac{\partial B}{\partial q}\frac{\partial A}{\partial p}\bigg|_{t=T}. \tag{3.36}$$

That justifies Peierls' statement.

4. QUANTUM MECHANICS

The equations of motion of quantum mechanics can be formulated in a form which is isomorphic to the Poisson-bracket formulation of classical mechanics, with quantities proportional to commutators taking the place of Poisson brackets. The quantum scheme introduces the associated algebra by means of all "analytic functions" of operators in Hilbert space. The expression "analytic function" is here understood with the meaning of convergent symmetrized power series.

We have stated our action principle as follows:

$$\delta A = [\delta W, A], \tag{4.1}$$

where $[\ ,\ ]$ is the Lie-bracket multiplication. In quantum mechanics the Lie-algebra multiplication is realized by means of commutation of operators in Hilbert space as follows:

$$[\delta W, A] \equiv [\delta W, A]_q = (1/i\hbar)(\delta W A - A\delta W). \tag{4.2}$$

Here the associative algebra D has the product equal to the ordinary product of operators in Hilbert space. The Poisson-bracket multiplication is equal to the factor $(1/i\hbar)$ multiplied by the commutator of operators as specified in (4.2). The ordinary product $(\ ,\ )$ is in quantum mechanics

$$(\delta W, A) = \delta W A, \tag{4.3}$$

where δW and A are Hilbert space operators. With these two definitions of $[\ ,\ ]$ and of $(\ ,\ )$, we can see quite easily that the relation (2.30) is satisfied.

As a particular case, we can consider the time evolution by means of the relation $\delta_t W = -\mathcal{H}\delta t$ that yields

$$i\hbar \frac{\delta A}{\partial t} = A\mathcal{H} - \mathcal{H}A, \tag{4.4}$$

which is the well-known Heisenberg equation of motion in Heisenberg picture. Usually we would have to specify the time limits in the variation of the action as done in (2.28), the case when the action principle becomes

$$\delta A(t) = [\delta W(t, t_0), A(t)], \tag{4.5}$$

where t_0 is the instant when the perturbation starts to act.

We have stated the action principle by means of variations of the elements of the Lie algebra, that in quantum mechanics are operators of Hilbert space. In this way such a principle is applicable both to classical and to quantum mechanics. But in quantum mechanics only, we want to transform this action principle to another, written by means of the variation of the transformation function as it was done by Schwinger.[9]

The quantities that in quantum mechanics are related to the physical reality are the matrix elements. To transfer from a Heisenberg-like picture in which we stated the action principle, to a Schrödinger-like picture as Schwinger stated it for quantum mechanics, we have to remark that the infinitesimal unitary transformation of the observables given by

$$\delta A(t) = \bar{A}(t) - A(t), \tag{4.6}$$

$$\bar{A}(t) = \left(1 + \frac{1}{i\hbar}\delta W(t, t_0)\right) A(t) \left(1 - \frac{1}{i\hbar}\delta W(t, t_0)\right) \tag{4.7}$$

$$= A(t) + [\delta W(t, t_0), A(t)] \tag{4.8}$$

induces in the eigenstates a transformation from $|a\rangle$ to $|\bar{a}\rangle$ given by

$$\langle \bar{a}|\, \bar{A}\, |\bar{a}\rangle = \langle a|\, A\, |a\rangle, \tag{4.9}$$

$$\delta\, |a\rangle = |\bar{a}\rangle - |a\rangle \equiv |\delta a\rangle, \tag{4.10}$$

$$\delta\, |a\rangle = \frac{\delta W(t_a, t_0)}{i\hbar}\, |a\rangle, \tag{4.11}$$

where t_a is the time instant at which the state vector $|a\rangle$ is evaluated.

[9] J. Schwinger, Phys. Rev. **91**, 713 (1953).

The variation of the transformation function is given by the following expression if the two eigenvectors of the transformation function are varied independently:

$$\begin{aligned}\delta\langle a \mid b\rangle &= \langle \delta a \mid b\rangle + \langle a \mid \delta b\rangle \\ &= i\frac{1}{\hbar}\langle a|\, \delta W(t_a, t_0) - \delta W(t_b, t_0)\,|b\rangle \\ &= i\frac{1}{\hbar}\langle a|\, \delta W(t_a, t_b)\,|b\rangle, \end{aligned} \tag{4.12}$$

which is the action principle for quantum mechanics as stated by Schwinger.[9]

Schwinger considers (4.12) as the definition of the infinitesimal operator $\delta W(t_a, t_b)$, from which he deduces that the requirement that any infinitesimal variation maintains the multiplicative composition law of transformation functions implies the additive composition law for the infinitesimal operators (2.25). The fundamental dynamical principle is contained in the postulate that there exists a class of transformation-function alterations for which the characterizing operators δW are obtained by appropriate variation of a single operator W given by (2.27).

The fact that the action element has to be the integral of the Lagrangian can be deduced in quantum mechanics from the requirement that any infinitesimal alteration of the transformation function maintains the multiplicative composition law of the same transformation functions. This conclusion, however, is a consequence of the fact that usually, in quantum mechanics, the dynamical Lie-algebra principle is realized by means of an algebra of operators in Hilbert space. We deduce immediately that the action principle (4.12) is also valid for the calculation of perturbations in quantum mechanics.

Indeed, if we use action principle (4.12) and apply Taylor's theorem

$$\langle a \mid b\rangle = \langle a \mid b\rangle|_{g=0} + \frac{\partial}{\partial g}\langle a \mid b\rangle|_{g=0} + \frac{\partial^2}{\partial g^2}\langle a \mid b\rangle|_{g=0} + \cdots, \tag{4.13}$$

where g is the coupling parameter between the two parts of the Action

$$W(t_a, t_b) = \int_{t_b}^{t_a} dt\{L_0(t) + gL_1(t)\}. \tag{4.14}$$

Action principle (4.12) yields immediately

$$\frac{\partial}{\partial g}\langle a \mid b\rangle = \frac{i}{\hbar}\langle a|\int_{t_b}^{t_a} dt L_1(t)\,|b\rangle, \tag{4.15}$$

which is a relation valid for any value of the coupling constant g and, in particular, for $g = 0$. We have also

$$\frac{\partial^2}{\partial g^2}\langle a \mid b\rangle = \frac{i}{\hbar}\frac{\partial}{\partial g}\sum_{c,d}\int_{t_b}^{t_a} dt\langle a \mid c\rangle\langle c|\, L_1(t)\,|d\rangle\langle d \mid b\rangle \tag{4.16}$$

with the restriction $t_c = t_d = t$, that implies

$$\frac{\partial}{\partial g}\langle c|\, L_1(t)\,|d\rangle = 0. \tag{4.17}$$

Therefore, after applying again the result (4.15), we obtain

$$\begin{aligned}&\frac{\partial^2}{\partial g^2}\langle a \mid b\rangle \\ &= \langle a|\left(\frac{i}{\hbar}\right)^2\int_{t_b}^{t_a} dt\left\{\int_{t_b}^{t_a} dt' L_1(t')L_1(t) + \int_{t_b}^{t_a} dt' L_1(t)L_1(t')\right\}|b\rangle \\ &= \langle a|\left(\frac{i}{\hbar}\right)^2\int_{t_b}^{t_a} dt\int_{t_b}^{t_a} dt'[L_1(t)L_1(t')]_+\,|b\rangle, \end{aligned} \tag{4.18}$$

where we have introduced the time-ordering operation $[\ ]_+$, which has the property that in operating on a product of time-labeled operators, it rearranges them in the same order as the time sequence of their labels, the latest one in time occurring first in the product.

In general, we have

$$\begin{aligned}\frac{\partial^n}{\partial g^n}\langle a \mid b\rangle = \langle a|\left(\frac{i}{\hbar}\right)^n\int_{t_b}^{t_a} dt^{(1)}\int_{t_b}^{t_a} dt^{(2)}\cdots\int_{t_b}^{t_a} dt^{(n)} \\ \times [L_1(t^{(1)})L_1(t^{(2)})\cdots L_1(t^{(n)})]_+\,|b\rangle \end{aligned} \tag{4.19}$$

that has to be evaluated for $g = 0$ to obtain the expression for $\langle a \mid b\rangle$ in (4.13). So we have

$$\langle a \mid b\rangle = \langle a|\left(\exp\left\{\frac{i}{\hbar}\int_{t_b}^{t_a} L_1(t)\,dt\right\}\right)_+|b\rangle|_{g=0}, \tag{4.20}$$

which is an expression from which we deduce the well-known formula for the evolution operator in the interaction picture with which we can calculate perturbations in quantum mechanics.

Indeed, if U is the evolution operator in the interaction picture we have

$$\langle a \mid b\rangle = \langle a|\, U(t_a, t_b)\,|b\rangle|_{g=0}, \tag{4.21}$$

and, therefore,

$$U(t_a, t_b) = \left(\exp\left\{\frac{i}{\hbar}\int_{t_b}^{t_a} L_1(t)\,dt\right\}\right)_+, \tag{4.22}$$

which is a very well-known expression.

5. CONCLUSION

The general structural features of dynamical theories that we have exhibited have a profound physical meaning. Classical and quantum dynamics

require particular realizations (functions or operators) as the natural realizations of the Lie algebra of dynamics common to both of them. The realization that is important for kinematics and for the physical interpretation of the theory is not important for the dynamical structure analysis. That is why we can obtain, as we have done above, a general interaction picture valid for any kind of mechanics.

The relative importance of the selected realization of the kinematic part of any mechanics is illustrated, considering the possibility of a transcription of classical and quantum mechanics each into the natural realization of the other.[10] Doing so we see that many of the features of the formalisms of all kinds of mechanics become identical. From this point of view the main difference between the two mechanics is in the choice of the Lie bracket. The difference between classical and quantum mechanics resides mainly in the choice of realization for the dynamical group. But we see also that each mechanics is very awkward in the natural representation of the other.

As has been shown before,[10] there is a general form of a Lie bracket which includes the brackets of classical and quantum mechanics as special cases. This fact suggests the existence of more general mechanical formalisms.

After we have examined the validity of the action principle for classical and quantum mechanics, a question that arises quite naturally is whether there exists a superscheme beyond, and inclusive of, the two kinds of mechanics that we have specifically studied. The problem of the existence of a universal superscheme has to be answered affirmatively, so far as we know now; however, it has to be left open in this paper.

Such a superscheme will be obtained when the action principle is extended to yield also the variations of the Lie-algebra elements induced when we change the realization by derivations of the algebra. It can be seen that action principle (2.3) gives also these variations when δW is interpreted as the change in the action that is induced by the change of realization.

The different realizations of the algebra will be mapped isomorphically into a set of parameters that have continuous or discrete values. We obtain classical and quantum mechanics when we give to these sets of parameters a concrete, fixed set of values. Since the action principle is valid also for the variations of the elements of the algebra corresponding to the variations of these sets of parameters which determine the realization, we can obtain, in a form compatible with the action principle, continuously or discretely different kinds of mechanics. This process allows us to obtain, in a quite natural way, a semiclassical approximation to quantum mechanics, for instance.

The principal aim of this paper has been to write an action principle (2.3) from which an interaction picture valid for classical and quantum mechanics could be deduced, and from it to write down a general procedure to evaluate perturbations in both kinds of mechanics mentioned.

[10] T. F. Jordan and E. C. Sudarshan, Rev. Mod. Phys. 33, 515 (1961).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Quadratic Action Principle of Relativity

C. LANCZOS
Dublin Institute for Advanced Studies, Dublin, Ireland,
and
Department of Physics, North Carolina State University, Raleigh, North Carolina

(Received 20 November 1968)

By an improved mathematical technique, the field equations derivable from a Lagrangian which is quadratic in the curvature components can now be studied in greater detail. The intercalation of a high-frequency metrical plateau between the flat Minkowskian metric and the macroscopic physical world has the consequence that the resulting perturbation equations are no longer of fourth but only of second order, thus making a comparison with Einstein's equations possible. The principal difference is that the free vector of Einstein's theory is now restricted by a divergence condition, with the result that the equations of the electromagnetic field, expressed in terms of the vector potential, become solutions of the macroscopic field equations. The cases of fourfold symmetry with imaginary time and 3 + 1 symmetry with real time are discussed. The gravitational phenomena and the second-order interaction terms, needed for the construction of material particles, remain outside the limits of the present investigation.

1. INTRODUCTION

The great discoveries of Einstein opened a new perspective in our speculative outlook on the phenomena of nature. Although it was the phenomenon of universal gravitation which led Einstein to his deep-rooted analysis of the problem of geometry and the discovery of the Riemannian nature of world geometry, one could hardly stop here and relegate the realm of geometry to gravitational phenomena alone. The field equations

$$R_{ik} = 0, \tag{1.1}$$

so fundamental for the description of the gravitational field, can hardly be considered as more than the first step toward a much more comprehensive structure. These equations put the matter tensor equal to zero, whereas we can hardly doubt that metrical tensor and matter tensor are of equal significance and that it is their interaction which has to be considered as the true battleground of physical phenomena.

Einstein arrived at the equations (1.1) with such convincing necessity that he doubted the possibility of making Riemannian geometry responsible for more than purely gravitational events. But if we believe in the fundamental significance of action principles, then the field equations (1.1) follow from a special choice of the basic action integral, namely the scalar curvature R, multiplied by the volume element, and integrated over the entire manifold.

It was Weyl[1] who first called attention to the fact that a rational action integral should have invariance not only with respect to arbitrary coordinate transformations, but also with respect to the arbitrary units in which we measure lengths (gauge invariance). Einstein's action integral is dimensioned (the square of a length), whereas we should demand that the basic action integral become a *pure number*. This means that the basic Lagrangian must depend *quadratically* on the curvature components. The general possibilities can be reduced to the choice[2]

$$L = \tfrac{1}{2}(R_{ik}R^{ik} - \mu R^2)g^{\frac{1}{2}}, \tag{1.2}$$

where μ is an *a priori* undetermined (positive) constant.

Here we have an action principle which promises the possibility of erecting world geometry on a basis of maximum rationality. We would hardly be justified to abandon the tenets of Riemannian geometry before we have convinced ourselves that this action principle cannot lead to sufficiently general results.

From a purely gravitational standpoint, one would scarcely see the necessity of modifying Einstein's linear invariant. Nevertheless, occasionally such efforts were made. Buchdahl[3] investigated the gravitational equations, which follow if Einstein's R is replaced by R^2. Gregory,[4] and later Pechlaner and Sexl,[5] employed a linear combination of R and R^2. Penney[6] refers to the quadratic action principle in his investigation of the classical electron.

The principal aim of Weyl in his discussion of gauge invariance was to find a natural place for the electromagnetic vector potential in the geometry of nature. He endeavored to generalize the Riemannian basis of geometry by demanding that the field equations should determine the ratios of the g_{ik} only, leaving a free factor at every point. One can show, however,

[1] H. Weyl, Math. Z. **2**, 384 (1918); Ann. Physik **59**, 101 (1919); Physik. Z. **22**, 473 (1921).

[2] C. Lanczos, Ann. Math. **39**, 842 (1938).
[3] H. A. Buchdahl, Nuovo Cimento **23**, 141 (1962).
[4] R. C. Gregory, Phys. Rev. **72**, 72 (1947).
[5] E. Pechlaner and R. Sexl, Commun. Math. Phys. **2**, 165 (1966).
[6] R. Penney, Phys. Rev. **137**, B1385 (1965).

that the vector potential thus introduced exists already within a purely Riemannian framework.[7] The calculations of Pauli[8] concerning the theory of Weyl did not lead to promising results.

The author's early papers[9] demonstrated the existence of a field vector which satisfied the potential equation, but the identification of this vector with the vector potential foundered on the difficulty that it leads to a vanishing of the free electric charge. In the later phase of his speculations he came to the realization that no progress can be made as long as we adhere to the purely historically motivated idea that cosmic geometry must be erected on a flat Minkowskian background, considering the Riemannian line element as nearly flat, whenever we are away from the central core of material particles. By this assumption we throw away the valuable quadratic terms of the curvature tensor and reduce our problem to an oversimplified structure. A much richer structure is obtained if we assume that there exists a highly agitated metrical plateau of such high frequency that, for all macroscopic purposes (including even nuclear events), only the average values of the g_{ik} are at our disposal which thus become constants, although they hide the existence of very high curvatures.[10]

The present paper develops a mathematical treatment of the quadratic action principle, which makes it possible to draw definite conclusions concerning the interaction between the basic plateau and the macroscopic superposition effects. The detailed structure of the basic lattice eludes our present mathematical possibilities. We can argue, however, on the basis of *symmetry considerations*. They permit us to exhibit certain general consequences of the basic hypotheses which are interesting in themselves, although they have to be corroborated by future research.

2. THE BASIC LAGRANGIAN

The following viewpoint will guide us in the subsequent discussions of this section. A Lagrangian, which is not more than *quadratic* in the action variables, leads to *linear* field equations. This excludes any interaction and is thus unable to cope with the physical facts. On the other hand, a Lagrangian of high algebraic order would make the study of interactions exceedingly difficult. Can we perhaps succeed with a Lagrangian which is not higher than *cubic* in the action variables? Such a Lagrangian would guarantee the existence of weak and also strong interactions and, at the same time, facilitate the mathematical study of the interaction terms. We will see that such a program can actually be carried out up to a last term, which is a determinant, and thus of fourth rather than of third order.

We define our first action variables h_i^k by putting

$$h_i^k = R_i^k g^{\frac{1}{4}}. \tag{2.1}$$

Our quadratic Lagrangian (1.2) now becomes

$$L_1 = \tfrac{1}{2}(h_i^k h_k^i - \mu h^2). \tag{2.2}$$

The metrical tensor g_{ik} appears in the form $g_{ik}g^{-\frac{1}{4}}$ consistently throughout this paper. We should choose a special notation for this quantity, but it will be more convenient to remember that we denote this modified tensor by g_{ik} while the usual g_{ik} becomes $g_{ik}g^{\frac{1}{4}}$. In fact, we put

$$g^{\frac{1}{2}} = e^{\varphi} \tag{2.3}$$

and consider as our action variables the new g_{ik}, which now satisfy the auxiliary condition

$$\|g_{ik}\| = 1, \tag{2.4}$$

and the scalar φ. The actual metrical tensor, to be denoted by g_{ik}^*, now becomes

$$g_{ik}^* = e^{\varphi/2} g_{ik}. \tag{2.5}$$

Similarly,

$$h_i^k = R_i^k e^{\varphi/2} \tag{2.6}$$

and

$$R_{ik} = h_i^\alpha e^{-\varphi/2} g_{h\alpha}^* = h_i^\alpha g_{\alpha k}. \tag{2.7}$$

We see that our new action variables have no tensor significance, although they are closely related to tensor quantities.

The Lagrangian (2.2) holds under the auxiliary condition

$$R_{ik} - (\varphi_{,ik} - \varphi_{,\alpha}\Gamma_{ik}^\alpha - \Gamma_{ik,\alpha}^\alpha + \Gamma_{i\alpha}^\beta \Gamma_{k\beta}^\alpha) = 0. \tag{2.8}$$

This gives rise to a Lagrangian factor p^{ik} and the added Lagrangian

$$L_2 = -p^{ik}(h^{i\alpha}g_{\alpha k} - \varphi_{,ik} + \varphi_{,\alpha}\Gamma_{ik}^\alpha + \Gamma_{ik,\alpha}^\alpha - \Gamma_{i\alpha}^\beta \Gamma_{k\beta}^\alpha). \tag{2.9}$$

By joining the Γ_{ik}^m to the previous action variables, our Lagrangian is still of not higher than third degree. But the freedom of varying the Γ_{ik}^m entails the auxiliary condition

$$\Gamma_{ik}^\alpha g_{\alpha m} - e^{-\varphi/2}\begin{bmatrix} ik \\ m \end{bmatrix} = 0, \tag{2.10}$$

where

$$\begin{bmatrix} ik \\ m \end{bmatrix} = \tfrac{1}{2}[(e^{\varphi/2}g_{im})_{,k} + (e^{\varphi/2}g_{km})_{,i} - (e^{\varphi/2}g_{ik})_{,m}]. \tag{2.11}$$

[7] C. Lanczos, Rev. Mod. Phys. **29**, 337 (1957).
[8] W. Pauli, Physik. Z. **20**, 457 (1919).
[9] C. Lanczos, Phys. Rev. **39**, 716 (1932); **61**, 713 (1942).
[10] C. Lanczos, J. Math. Phys. **4**, 951 (1963); **7**, 316 (1966); **8**, 829 (1967).

Hence we have to add the further Lagrangian

$$L_3 = -u^{ikm}[\Gamma^\alpha_{ik}g_{\alpha m} - \tfrac{1}{2}(g_{im,k} + g_{km,i} - g_{ik,m}) - \tfrac{1}{4}(g_{im}\varphi_{,k} + g_{km}\varphi_{,i} - g_{ik}\varphi_{,m})]. \quad (2.12)$$

Here again the added Lagrangian does not surpass the third degree.

In addition, however, the auxiliary condition (2.4) gives rise to a last Lagrangian of the form

$$L_4 = C(\|g_{ik}\| - 1), \quad (2.13)$$

which is of the fourth degree, in view of the four-dimensional character of the space–time world.

3. THE FIELD EQUATIONS

Our final action variables become

$$(g_{ik}, \varphi), \quad (h^i_k, p^{ik}), \quad (\Gamma^m_{ik}, u^{ikm}). \quad (3.1)$$

We can now carry out the variations and obtain the basic field equations. These equations fall into two categories. The variation with respect to the variables in the second and third brackets do not restrict the type of geometry we are interested in. We may call these equations the "morphological equations," because they establish certain fundamental quantities which will be helpful in the study of that particular geometrical structure which is established by the quadratic action principle without specifying yet that geometry beyond its Riemannian character. The decisive statements concerning the specific structure of world geometry will be obtained by the variations of φ and the g_{ik}.

We tabulate the morphological equations in the sequence of varying with respect to h^i_k, p^{ik}, u^{ikm}, Γ^m_{ik}, obtaining

$$p^{k\alpha}g_{\alpha i} = h^k_i - \mu h\delta^k_i, \quad (3.2)$$

$$h^\alpha_i g_{\alpha k} = \varphi_{,ik} - \varphi_{,\alpha}\Gamma^\alpha_{ik} - \Gamma^\alpha_{ik,\alpha} + \Gamma^\beta_{i\alpha}\Gamma^\alpha_{k\beta}, \quad (3.3)$$

$$\Gamma^\alpha_{ik}g_{\alpha m} = \tfrac{1}{2}(g_{im,k} + g_{km,i} - g_{ik,m}) + \tfrac{1}{4}(g_{im}\varphi_{,k} + g_{km}\varphi_{,i} - g_{ik}\varphi_{,m}), \quad (3.4)$$

$$u^{ik\alpha}g_{\alpha m} = p^{ik}_{,m} - p^{ik}\varphi_{,m} + p^{i\alpha}\Gamma^k_{\alpha m} + p^{k\alpha}\Gamma^i_{\alpha m}. \quad (3.5)$$

In absolute terms these equations have the following significance, if we use the symbol | for covariant differentiation:

$$p^{ik} = (R^{ik} - \mu R g^{*ik})e^\varphi, \quad (3.6)$$

$$\Gamma^m_{ik} = \hat{\Gamma}^m_{ik} + \tfrac{1}{4}(\delta^m_i\varphi_{,k} + \delta^m_k\varphi_{,i} - g_{ik}\varphi^{,m}), \quad (3.7)$$

$$u^{ik\alpha}g_{\alpha m} = u^{ik}_m = e^\varphi(R^{ik} - \mu R g^{*ik})_{|m}. \quad (3.8)$$

The notation $\hat{\Gamma}^m_{ik}$ is employed for the usual definition of the Γ^m_{ik}, but now formed with respect to our present g_{ik}. They satisfy the condition

$$\hat{\Gamma}^\alpha_{i\alpha} = 0, \quad (3.9)$$

and thus

$$\Gamma^\alpha_{i\alpha} = \varphi_{,i}. \quad (3.10)$$

We now come to the study of the specific field equations which follow by varying the basic action variables φ and g_{ik}. The variation of φ yields the following scalar relation:

$$p^{ik}_{,ik} + (p^{ik}\Gamma^\alpha_{ik})_{,\alpha} - \tfrac{1}{2}(u^{i\alpha k}g_{ik})_{,\alpha} + \tfrac{1}{4}(u^{ik\alpha}g_{ik})_{,\alpha} = 0. \quad (3.11)$$

The sum of the first two terms has already invariant significance, if we take into account Eq. (3.6):

$$p^{ik}_{,ik} + (p^{ik}\Gamma^\alpha_{ik})_{,\alpha} = e^\varphi(R^{ik} - \mu R g^{*ik})_{|ik}. \quad (3.12)$$

In view of the divergence identity

$$(R^{ik} - \tfrac{1}{2}Rg^{*ik})_{|k} = 0, \quad (3.13)$$

the right-hand side becomes

$$(\tfrac{1}{2} - \mu)e^\varphi\Delta R. \quad (3.14)$$

The operator Δ is the usual invariant Laplace operator (extended to four dimensions):

$$\Delta = {}_{|ik}g^{ik}. \quad (3.15)$$

Similarly, by (3.8), we obtain

$$(u^{i\alpha k}g_{ik})_{,\alpha} = (u^{ik}_i)_{,k} = e^\varphi(\tfrac{1}{2} - \mu)\Delta R \quad (3.16)$$

and

$$(u^{ik\alpha}g_{ik})_{,\alpha} = e^\varphi(1 - 4\mu)\Delta R. \quad (3.17)$$

Hence, the field equation caused by the variation of φ becomes

$$(\tfrac{1}{4} - \tfrac{1}{2}\mu + \tfrac{1}{4} - \mu)\Delta R = \tfrac{1}{2}(1 - 3\mu)\Delta R = 0, \quad (3.18)$$

which yields the result (well known from earlier treatments) that the scalar curvature R must satisfy the scalar wave equation

$$\Delta R = 0. \quad (3.19)$$

This equation is lost only for the singular value $\mu = \frac{1}{3}$, which would lead to an unacceptable underdetermination, although exactly this singular value plays a characteristic role in Weyl's theory.[11] Hence

$$R = \text{const} = R_0 \quad (3.20)$$

is an exact first integral of the field equations. In a Riemannian geometry of positive-definite signature, this is the *only* possible solution of the potential equation which is free of singularities and behaves

[11] Confer Ref. 7, p. 343.

properly in infinity. Even in the indefinite Minkowskian case, it seems highly improbable that the sourceless eigensolutions of the scalar wave equation can have physical significance. We thus assume that (3.20) is not only a *possible*, but a *necessary* consequence of the field equations and a characteristic property of the quadratic action principle.

This result has immediately profound consequences. It shows that the scalars φ and h are coupled by the relation

$$e^{-\varphi/2}h = R_0. \tag{3.21}$$

Furthermore, from (3.8) by contraction we obtain

$$\begin{aligned} u_m^{mk} &= e^{\varphi}(R^{mk} - \mu R g^{*mk})_{|m} \\ &= e^{\varphi}(\tfrac{1}{2} - \mu)R_{,m}g^{*mk} = 0 \end{aligned} \tag{3.22}$$

and likewise, multiplying by g^*_{ik}, we get

$$g^*_{ik}u_m^{ik} = e^{\varphi}(1 - 4\mu)R_{,m} = 0. \tag{3.23}$$

Hence the two vectors, which could be obtained by contraction from the third-rank tensor μ_m^{ik} (symmetric in i, k), both *vanish.*

We also notice from (3.8) that on the right-hand side the term proportional to μ *cancels out.* Hence we can replace the definition (3.5) of μ_m^{ik} by putting

$$u_m^{ik} = h_{,m}^{ik} - h^{ik}\varphi_{,m} + h^{i\alpha}\Gamma^k_{\alpha m} + h^{k\alpha}\Gamma^i_{\alpha m}. \tag{3.24}$$

Finally, we vary the basic Lagrangian with respect to g_{ik}, obtaining the following set of fundamental field equations:

$$\begin{aligned} &-\tfrac{1}{2}(p^{\alpha i}h^k_\alpha + p^{\alpha k}h^i_\alpha) + Cg^{ik} \\ &\quad = \tfrac{1}{2}(u^{\alpha ik} + u^{\alpha ki} - u^{ik\alpha})_{,\alpha} + \tfrac{1}{2}(u^{\alpha\beta i}\Gamma^k_{\alpha\beta} + u^{\alpha\beta k}\Gamma^i_{\alpha\beta}) \\ &\qquad - \tfrac{1}{4}(u^{\alpha ik} + u^{\alpha ki} - u^{ik\alpha})\varphi_{,\alpha}. \end{aligned} \tag{3.25}$$

If this equation is multiplied on both sides by g_{ik}, we find that the right-hand side cancels out identically, in consequence of the previous relations (3.22) and (3.23). This determines the Lagrangian factor C in the sense of

$$\begin{aligned} C &= \tfrac{1}{4}p^\alpha_\beta \cdot h^\beta_\alpha \\ &= \tfrac{1}{4}(h^\beta_\alpha h^\alpha_\beta - \mu h^2). \end{aligned} \tag{3.26}$$

4. THE PERTURBATION EQUATIONS

We cannot hope to obtain the general solution of a highly nonlinear set of ten partial differential equations. We can follow, however, Einstein's procedure in the investigation of his gravitational equations. We start with a fundamental field (which satisfies the field equations) which we take for granted and investigate a small *perturbation* of this field. This must lead to a linear set of equations for the perturbation.

Einstein identified the basic field with the flat Minkowskian manifold. This is not our program. If we assume the existence of a metrical plateau of very high frequency, then the quadratic terms of the Riemann tensor can cause by resonance very high average curvatures, although the amplitudes of these vibrations remain very small. Our field equations are such that they allow the so-called "cosmological equations"

$$R_{ik} = \sigma g^*_{ik} = \sigma e^{\varphi/2}g_{ik} \tag{4.1}$$

as exact solutions. Ordinarily we would assume that the constant σ must be very *small.* In our highly agitated field, however, this constant—whose dimension is the reciprocal square of a length—might become excessively *large*, if measured in ordinary units, because the order of magnitude of this constant becomes proportional to

$$\alpha^2\omega^2, \tag{4.2}$$

where ω denotes the lattice frequency and α the amplitude of the vibrations.

In our present variables the relation (4.1) appears in the form

$$h_{ik} = \sigma e^{\varphi/2}g_{ik}, \tag{4.3}$$

and we can easily convince ourselves that the field equations are *trivially* fulfilled, because the tensor u_m^{ik} vanishes *identically*. For this purpose we write (3.24) in the form

$$\begin{aligned} u_m^{ik} = h_{,m}^{ik} &- \tfrac{1}{2}h^{ik}\varphi_{,m} + h^{i\alpha}\hat{\Gamma}^k_{\alpha m} + h^{k\alpha}\hat{\Gamma}^i_{\alpha m} \\ &+ \tfrac{1}{4}h^i_\alpha(\delta^k_m\varphi^\alpha_{,} - \delta^\alpha_m\varphi^k_{,}) + \tfrac{1}{4}h^k_\alpha(\delta^i_m\varphi^\alpha_{,} - \delta^\alpha_m\varphi^i_{,}). \end{aligned} \tag{4.4}$$

The substitution of (4.3) yields the vanishing of u_m^{ik}, and the field equations (3.25) are satisfied, because both the left and the right side vanish separately.

We now take the *perturbation* of the field equation (3.25), i.e., a small modification of the basic solution (4.3). We generally denote a small perturbation of a quantity by an overbar, e.g.,

$$\hat{g}_{ik} = g_{ik} + \bar{g}_{ik}, \tag{4.5}$$

and so on, with the understanding that we neglect quantities which are of second order in $\bar{g}_{ik}$. We make, however, the further assumption that the perturbation is small not only in *amplitude*, but also in *frequency*, in comparison to the frequency of the original quantity. This has the consequence that the perturbation equations need not be taken locally, but can be integrated over a lattice cell. At this point we abandon the principle of general covariance, because averaging over a domain is not a covariant operation.

If we consider Eq. (4.1) in the average sense, then the g_{ik} become *constants.* The same can be said of the

scalar φ, in view of (3.21). This constant can be normalized to zero which brings the basic field into the form

$$h_{ik} = \sigma g_{ik}. \tag{4.6}$$

(We neglect here cosmological effects which would cause a slow, secular change of the g_{ik}.)

We get three types of terms: those proportional to σ^2, those proportional to σ, and absolute terms which are independent of σ. Dimensional considerations show us that terms of the first type are not differentiated, terms of the second type are twice differentiated, terms of the third type are four times differentiated. In view of the symmetry of the basic solution with respect to the four axes, the σ^2 terms are reduced to a single term, proportional to $\bar{g}_{ik}$. This term must cancel out. If this were not the case, then our lattice would not be able to transmit signals of low frequency. We have the numerical constant μ at our disposal and must assume that this constant is adjusted in such a way that the undifferentiated term proportional to $\bar{g}_{ik}$ *drops out*. This leaves us with the twice differentiated terms proportional to σ, and the four-times-differentiated absolute terms. The latter ones are practically negligible in comparison to the former, in view of the largeness of σ. The resulting field equations will thus be of *second order* only. If we proceed in the customary fashion by erecting the perturbation equations on a flat space, then σ vanishes and the only remaining terms are those which are four times differentiated. The basic reason which made Einstein lukewarm toward the quadratic action principle was that the resulting differential equations became of fourth order, which did not seem to jibe with our physical experiences. This difficulty is overcome by the presence of the agitated metrical plateau.

The terms which are decisive for the perturbation of the field (3.25) become on the left-hand side (after canceling out the terms with σ^2):

$$\begin{aligned} &-\tfrac{1}{2}[(p^{\alpha i}g^{\beta k} + p^{\alpha k}g^{\beta i})\bar{h}_{\alpha\beta} + (h^{\alpha i}g^{\beta k} + h^{\alpha k}g^{\beta i})\bar{p}_{\alpha\beta}] \\ &= -\sigma[(1 - 4\mu)g^{\alpha i}g^{\beta k}\bar{h}_{\alpha\beta} + g^{\alpha i}g^{\beta k}(\bar{h}_{\alpha\beta} - \mu\bar{h}_{\rho\sigma}g^{\rho\sigma}g_{\alpha\beta})] \\ &= -\sigma[(2 - 4\mu)g^{\alpha i}g^{\beta k}\bar{h}_{\alpha\beta} - \mu\bar{h}_{\alpha\beta}g^{\alpha\beta}g^{ik}], \end{aligned} \tag{4.7}$$

while on the right-hand side we obtain (considering that the original u^{ik}_m vanish and also the average values of Γ^m_{ik} are zero):

$$\tfrac{1}{2}(\bar{u}^{\alpha ik} + \bar{u}^{\alpha ki} - \bar{u}^{ik\alpha})_{,\alpha}. \tag{4.8}$$

The g^{ik}, although locally highly agitated, are for our purposes constants. If we multiply on both sides by g_{ik} and take into account the fact that the perturbation of Eq. (3.22) and (3.23) gives

$$\bar{u}^{\alpha i}_{\alpha} = 0 \tag{4.9}$$

and

$$\bar{u}^{ik\alpha}g_{ik} = 0, \tag{4.10}$$

we obtain on the right-hand side zero, and this yields for the left-hand side

$$(2 - 8\mu)\bar{h}_{\alpha\beta}g^{\alpha\beta} = 0. \tag{4.11}$$

Hence we can omit the second term on the right-hand side of (4.7) and write down our perturbation equation as follows:

$$-\sigma(2 - 4\mu)g^{\alpha i}g^{\beta k}\bar{h}_{\alpha\beta} = \tfrac{1}{2}(\bar{u}^{\alpha ik} + \bar{u}^{\alpha ki} - \bar{u}^{ik\alpha})_{,\alpha}. \tag{4.12}$$

The perturbation of (4.4) yields

$$\begin{aligned} \bar{u}^{ik}_m &= \bar{h}^{ik}_{,m} - \tfrac{1}{2}h^{ik}\bar{\varphi}_{,m} + h^{i\alpha}\hat{\Gamma}^k_{\alpha m} + h^{k\alpha}\hat{\Gamma}^i_{\alpha m} \\ &= \bar{h}^{ik}_{,m} - \tfrac{1}{2}\sigma g^{ik}\bar{\varphi}_{,m} + \sigma\bar{g}_{\alpha\beta,m}g^{\alpha i}g^{\beta k}. \end{aligned} \tag{4.13}$$

Under these circumstances, we can put

$$\bar{u}^{ik}_m = v^{ik}_{,m}, \tag{4.14}$$

where

$$v^{ik} = \bar{h}^{ik} - \tfrac{1}{2}\sigma g^{ik}\bar{\varphi} + \sigma\bar{g}_{\alpha\beta}g^{\alpha i}g^{\beta k}. \tag{4.15}$$

The condition (4.10) imposes the condition

$$v^{ik}g_{ik} = \bar{h}^{ik}g_{ik} + \sigma\bar{g}_{\alpha\beta}g^{\alpha\beta} - 2\sigma\bar{\varphi} = 0. \tag{4.16}$$

This is, in fact, in harmony with the demands of Eq. (3.21):

$$h^{ik}g_{ik}e^{-\varphi/2} = R_0 \tag{4.17}$$

whose perturbation gives exactly the condition (4.16). To this has to be added the condition

$$\bar{g}_{ik}g^{ik} = 0, \tag{4.18}$$

which is an immediate consequence of (2.4). Hence (4.16) is reduced to

$$2\sigma\bar{\varphi} = \bar{h}^{ik}g_{ik}. \tag{4.19}$$

We have to add the remark that the $\bar{h}_{ik}$ appearing on the left-hand side and the right-hand side of (4.12) are *not* the same quantities. On the left $\bar{h}_{ik}$ is multiplied by σ, but not on the right. On the left the *linear* terms of the operator (2.8) are activated, while on the right the perturbation of the highly agitated *quadratic* terms of the curvature tensor become operative, which are proportional to the large constant σ.

5. THE EINSTEIN OPERATOR

In his investigation of weak gravitational fields, Einstein employed the field equations (1.1) in the form of a perturbation of the flat Minkowskian metric. Let us, for the sake of mathematical simplicity, introduce the imaginary coordinate

$$x_4 = ict, \tag{5.1}$$

thus reducing the basic line element to the Euclidean normal values

$$g_{ik} = \delta_{ik}. \tag{5.2}$$

Then we can dispose of the distinction between covariant and contravariant components and write in our present notation:

$$\begin{aligned}\bar{R}_{ik} &= \bar{\varphi}_{,ik} - \Gamma^{\alpha}_{ik,\alpha} \\ &= \tfrac{1}{2}\bar{\varphi}_{,ik} + \tfrac{1}{4}\bar{\varphi}_{,\alpha\alpha}\delta_{ik} - \tfrac{1}{2}(\bar{g}_{i\alpha,k} + \bar{g}_{k\alpha,i} - \bar{g}_{ik,\alpha})_{,\alpha}\end{aligned} \tag{5.3}$$

with the auxiliary condition

$$\bar{g}_{\alpha\alpha} = 0. \tag{5.4}$$

The Einstein operator $\bar{R}_{ik}$ has two remarkable properties. The first is the divergence condition

$$\bar{R}_{i\alpha,\alpha} - \tfrac{1}{2}\bar{R}_{\alpha\alpha,i} = 0. \tag{5.5}$$

The second is that $\bar{R}_{ik}$ is automatically zero, if we put

$$\bar{g}_{ik} = \varphi_{i,k} + \varphi_{k,i} - \tfrac{1}{2}\varphi_{\alpha,\alpha}\delta_{ik}, \tag{5.6}$$

$$\bar{\varphi} = \varphi_{\alpha,\alpha}. \tag{5.7}$$

The fact that the field equations (1.1) leave a free vector undetermined was originally a great puzzle to Einstein,[12] since it appeared to him as an impermissible underdetermination, until he realized that it is a natural consequence of general covariance, expressing the freedom of infinitesimal coordinate transformations. Hence the vector φ_i is void of physical significance. The freedom of choosing φ_i can be utilized for a natural *normalization* of our reference system. After the normalization, the field equations of Einstein for infinitesimal gravitational fields can be written as follows:

$$\Delta\gamma_{ik} = \gamma_{ik,\alpha\alpha} = 0 \tag{5.8}$$

with the coordinate condition

$$\gamma_{i\alpha,\alpha} = 0, \tag{5.9}$$

where

$$\gamma_{ik} = \bar{g}^{*}_{ik} - \tfrac{1}{2}\bar{g}^{*}_{\alpha\alpha}\delta_{ik} \tag{5.10}$$

or, expressed in our g_{ik},

$$\gamma_{ik} = \bar{g}_{ik} - \tfrac{1}{2}\bar{\varphi}\delta_{ik}. \tag{5.11}$$

Let us investigate what our field equations give under corresponding circumstances.

Our basic perturbation equation is (4.12) and we will start with the left-hand side which, in view of (5.2), becomes proportional to $\bar{h}_{ik}$. In this $\bar{h}_{ik}$, we have to utilize those terms which are independent of σ. The definition (2.7) of h_{ik} shows that these terms are identical with Einstein's $\bar{R}_{ik}$ expressed in the form (5.3). Had we no right-hand side, our field equations would coincide with Einstein's equations for infinitely weak gravitational fields.

Now we come to the right-hand side of (4.12) which becomes

$$\tfrac{1}{2}(v_{\alpha i,k} + v_{\alpha k,i} - v_{ik,\alpha})_{,\alpha}, \tag{5.12}$$

with

$$v_{ik} = \bar{h}^{ik} - \tfrac{1}{2}\sigma\bar{\varphi}\delta_{ik} + \sigma\bar{g}_{ik}, \tag{5.13}$$

where

$$\bar{h}^{ik} = \overline{h_{\alpha\beta}g^{\alpha i}g^{\beta k}} = \bar{h}_{ik} - 2\sigma\bar{g}_{ik}, \tag{5.14}$$

and thus

$$v_{ik} = \bar{h}_{ik} - \tfrac{1}{2}\sigma\bar{\varphi}\delta_{ik} - \sigma\bar{g}_{ik}. \tag{5.15}$$

The $\bar{h}_{ik}$ which appear here originate from the perturbation of the high-frequency lattice field. They are caused by the quadratic terms of the tensor R_{ik} and have the following general structure:

$$\bar{h}_{ik} = \sigma(\bar{g}_{\alpha\beta}a_{\alpha\beta ik} + \bar{\varphi}b_{ik}), \tag{5.16}$$

where the coefficients $a_{\alpha\beta ik}$ and b_{ik} are the components of two numerical tensors, obtained by averaging certain quadratic resonance terms over the lattice. We have no right to dispose freely of these coefficients, because they are determined by the structure of the basic lattice. Since, however, that structure is beyond our present knowledge, we can rely only on symmetry considerations. If we assume that our lattice is macroscopically homogeneous in all the four axes, then we can put

$$\bar{h}_{ik} = \sigma_1\bar{g}_{ik} + \sigma_2\bar{\varphi}\delta_{ik}. \tag{5.17}$$

Then (4.19) gives

$$(2\sigma - 4\sigma_2)\bar{\varphi} = \sigma_1\bar{g}_{ik}\delta_{ik} = 0, \tag{5.18}$$

and hence by (5.15) we obtain

$$v_{ik} = (\sigma_1 - \sigma)\bar{g}_{ik}, \tag{5.19}$$

to which we have to add the condition (4.9):

$$v_{i\alpha,\alpha} = (\sigma_1 - \sigma)\bar{g}_{i\alpha,\alpha} = 0. \tag{5.20}$$

The resulting field equations now become

$$\bar{\varphi} = 0, \tag{5.21}$$

$$\bar{g}_{i\alpha,\alpha} = 0, \tag{5.22}$$

$$\Delta\bar{g}_{ik} = 0. \tag{5.23}$$

We compare these equations with Einstein's equations (5.8) and (5.9) for infinitesimal fields. These equations remain once more valid if the definition (5.10) of γ_{ik} is modified to

$$\gamma_{ik} = \bar{g}^{*}_{ik} - \tfrac{1}{4}\bar{g}^{*}_{\alpha\alpha}\delta_{ik}. \tag{5.24}$$

[12] A. Einstein and M. Grossmann, Z. Math. Physik **62**, 225 (1913).

The modification of Einstein's equations is twofold. First, the $\frac{1}{2}$ in Einstein's normalization equation (5.10) is changed to $\frac{1}{4}$. Second, the condition (5.9) is *no longer a matter of normalization but a consequence of the field equations.* For this reason the Einsteinian equations have now a second solution. Previously, the solution

$$\bar{g}_{ik} = \varphi_{i,k} + \varphi_{k,i} \tag{5.25}$$

had no physical significance, since the free vector φ_i was purely caused by coordinate transformations. In our noncovariant theory (covariant in principle, but made noncovariant by the existence of a strong basic matter field), the vector φ_i is no longer free but subject to the equation

$$\Delta\varphi_i = 0 \tag{5.26}$$

together with the condition

$$\varphi_{\alpha,\alpha} = 0. \tag{5.27}$$

In (5.26) we recognize the equation of the *electromagnetic vector potential*, while (5.27) is the *Lorentz condition.* This condition is here again demanded by the field equations and is not a matter of normalization.

The change of $\frac{1}{2}$ to $\frac{1}{4}$ in (5.24) has the consequence that it annihilates the mass in Einstein's static and spherically symmetric solution of the gravitational equations. This is not necessarily an absurd result. In view of the excessive smallness of the gravitational effects compared with the electromagnetic effects, it is entirely plausible that the free mass should be treated as a second-order phenomenon, which does not come into evidence in the first approximation, particularly if this approximation assumes complete symmetry of the basic field with respect to the four axes.

6. THE CASE OF 3 + 1 SYMMETRY

If we pursue a purely pragmatic philosophy, then *any* mathematical structure is acceptable if it fits the sum total of observed phenomena as they exist at a certain period of historical evolution. With such a notion, rationalistic or aesthetic arguments can have no place in theoretical considerations. If, on the other hand, in experiencing the ever-widening unification of the fundamental principles, and particularly under the influence of Einstein's great discoveries, we arrive at the viewpoint that the basic structure of nature is rationally comprehensible, then we cannot start out with the hypothesis of an indefinite metric which from the beginning destroys the basic tenets of a rational geometry. It is true that the mathematical artifice of operating with an imaginary time makes the four coordinates apparently homogeneous. Yet the very fact of the $+++-$ signature of the Minkowskian metric indicates that the symmetry pattern of the physical universe is not of the $4 + 0$, but of the $3 + 1$ variety.

Let us assume that we do not deviate from the rationalism of a genuinely Riemannian geometry and operate with a line element which is positive-definite in all the four axes. We also assume our high-frequency metrical plateau which should satisfy the field equations

$$R_{ik} = \sigma g^*_{ik}, \tag{6.1}$$

at least in the macroscopic sense. This lattice need not be locally symmetric with respect to the four axes; and thus it is conceivable that in the macroscopic superposition effects, the $\bar{h}_{ik}$ of the general form (5.16) will not necessarily follow a fourfold, but a $3 + 1$ symmetry pattern. In the present section we want to investigate what consequences could be drawn from such a hypothesis. Can we come to a model which could demonstrate that a Minkowskian metric for the perturbation is nevertheless reconcilable with a strictly Riemannian structure of the basic plateau?

Once more we want to assume that the macroscopically constant g_{ik} are normalized in the sense of (5.2). But now all our four coordinates are *real.* The only change, in comparison with our previous treatment, is that we abandon the too-special assumption (5.17) and leave $\bar{h}_{ik}$ for the time being unspecified. However, the condition (4.14) is still valid, and hence we can once more reduce the tensor u^{ik}_m to the second-rank tensor v^{ik}. But then there exists a remarkable isomorphism between the Einstein operator on the left-hand side of (4.12) and the right-hand side of the same equation. We want to exploit this similarity.

Let us denote

$$-2\sigma(1 - 2\mu) = \rho. \tag{6.2}$$

The entire field equation (4.12) can now be written in the following form:

$$\rho(\tfrac{1}{2}\bar{\varphi}_{,ik} + \tfrac{1}{4}\varphi_{\alpha,\alpha}\delta_{ik}) - \begin{bmatrix} ik \\ \alpha \end{bmatrix}_{,\alpha} = 0, \tag{6.3}$$

if the Christoffel symbol [ ] is applied to the following quantity:

$$g'_{ik} = \rho\bar{g}_{ik} + v_{ik} \tag{6.4}$$

with the auxiliary condition

$$g'_{\alpha\alpha} = 0. \tag{6.5}$$

As we have seen in (5.6) and (5.7), these equations are identically satisfied by putting

$$\rho\bar{g}_{ik} + v_{ik} = \rho(\varphi_{i,k} + \varphi_{k,i} - \tfrac{1}{2}\varphi_{\alpha,\alpha}\delta_{ik}), \tag{6.6}$$

$$\bar{\varphi} = \varphi_{\alpha,\alpha}, \tag{6.7}$$

to which the condition

$$v_{i\alpha,\alpha} = 0 \tag{6.8}$$

has to be added.

The difficulty in making further progress lies in the fact that the v_{ik}, defined by (5.15), depend on h_{ik}, which again is determined by the structure of the high-frequency metrical lattice. At present we will be satisfied with a tentative solution of Eq. (6.6), which is interesting from the physical viewpoint, although its real motivation would require a much deeper investigation.

We agree that in the next set of formulas the indices i, k, α only assume the values 1, 2, 3, while the index 4 shall be considered separately. We now put

$$\begin{aligned} v_{ik} &= \varphi_{i,k} + \varphi_{k,i} - 2(\varphi_{\alpha,\alpha} + \eta\varphi_{4,4})\delta_{ik}, \\ v_{i4} &= \eta(\varphi_{i,4} + \varphi_{4,i}), \\ v_{44} &= 2\eta^2\varphi_{4,4} = 2\eta(\varphi_{\alpha,\alpha} + \eta\varphi_{4,4}) = -2\eta\varphi_{\alpha,\alpha}. \end{aligned} \tag{6.9}$$

The constant η is a numerical constant—possibly very large—which we assume to be *negative.*

Now we have no difficulty in showing that Eqs. (6.8) become identical with the Maxwell vacuum equations for the electromagnetic field strength

$$F_{ik} = \varphi_{i,k} - \varphi_{k,i} \tag{6.10}$$

(the subscripts go once more from 1 to 4), usually written in the form

$$F^{i\alpha}_{\ ,\alpha} = 0. \tag{6.11}$$

The constant η, taken with a negative sign, has the significance of $1/c^2$.

The metrical dissonance. In the last section we have exhibited a possible solution of the perturbation equations which can be identified with Maxwell's vacuum equations for the electromagnetic field, in spite of the fact that our metric is a genuine Riemannian metric (of the signature $+++-$). This establishes a dissonance between the metric of the basic plateau and the macroscopic superposition. Can we allow such a dissonance in view of the empirically established fact that in all differential equations of mathematical physics the time coordinate has always a hyperbolic and not an elliptic character? Do we have any evidence of a four-dimensional Laplace operator which is elliptic in all the four variables?

In order to answer this question we remark the following. In our ordinary physical theories, we consider the light velocity c as a dimensioned quantity, which can be normalized to any value we like by our choosing the proper unit for the time t. The "second" is a very large unit; its $1/3 \cdot 10^{10}$ part makes the light velocity equal to 1. In our present considerations η is a *pure number*, since the fourth coordinate x_4 is already normalized by the macroscopic equation $g_{ik} = \delta_{ik}$. Hence η is an *absolute quantity*, determined by the structure of the metrical lattice. It is possible that the natural unit of time is still *many orders of magnitude smaller* than the usual relativistic normalization. This is the case if η happens to be a very *large* constant which makes the light velocity c very *small.* If this is the case, then even the fastest and most rapid physical events are still very *slow* in absolute units. This means that in absolute units the universe is in a quasistationary state. Under such circumstances the derivative $\partial/\partial x_4$ becomes practically negligible in comparison to the other $\partial/\partial x_i$, which means that the four-dimensional Laplace operator becomes practically reduced to the usual three-dimensional potential operator. In macroscopic relations the operation $\partial^2/\partial x_4^2$ becomes noticeable only if magnified by the very large factor η which, being negative, explains the hyperbolic character of time in all observed phenomena of physics.

7. CONCLUSIONS

The present investigation was motivated by the following thought. Can we, encouraged by the great speculative victories of Einstein, apply mathematical principles of maximum rationality to the exploration of the physical world? In particular, can we establish a Riemannian geometry of maximum rationality as a unifying basis of all physical phenomena? A quadratic action principle seems to fit the criterion of such rationality, particularly if we do not abandon the positive-definiteness of a genuine Riemannian line element.

In carrying out this program, this paper deviates from Einstein not only by replacing Einstein's linear action integral by a quadratic one—which is a pure number and thus independent of the units in which lengths are measured—but also by replacing the flat field, on which the fields of material particles are erected, by a highly agitated periodic field, macroscopically characterized by the cosmological equations with a very large cosmological constant. In this case the resulting perturbation equations are no longer of fourth but only of second order, making a comparison with Einstein's infinitesimal gravitational equations possible. The result is that these equations now possess a second solution, which can be correlated to the electromagnetic field. This opens new possibilities toward the understanding of the relation between gravitational and electromagnetic phenomena. Moreover, replacing the fourfold symmetry of the

Minkowskian metric by a 3 + 1 symmetry of a real metric, a model was obtained which simulates the Minkowskian metric for the superposition field, although the basic metric is strictly Riemannian, i.e., positive-definite.

The present investigation does not go beyond the linear approximation. The second-order interaction terms, which must be made responsible for the construction of material particles as excited eigenstates of the field equations, have not been taken into account in this paper.

ACKNOWLEDGMENT

I remember with pleasure the many stimulating discussions with Professor W. R. Davis and his group of the Physics Department, North Carolina State University at Raleigh, North Carolina, during the Spring term, 1968.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Representations of the Orthogonal Group. II. Polynomial Bases for the Irreducible Representations of the Orthogonal Group

M. K. F. WONG*
Research Institute for Natural Sciences, Woodstock College, Woodstock, Maryland

(Received 25 February 1968)

Polynomial bases for the irreducible representations of the orthogonal group, which are characterized by the Gel'fand pattern, have been obtained. The method used is very similar to Moshinsky's and is a generalization from the unitary group to the orthogonal group. The Wigner coefficients of $O(3)$, commonly called the Clebsch–Gordan coefficients of $R(3)$, are rederived by means of the polynomial bases obtained in this paper.

I. INTRODUCTION

In a previous paper,[1] henceforth referred to as I, we obtained normalized lowering and raising operators of the orthogonal group in the group chain $O(n) \supset O(n-1) \supset \cdots \supset O(2)$. As a result, the Gel'fand–Zetlin matrix elements for the generators of the orthogonal group have been obtained. In this note we show how to construct polynomial bases for the irreducible representations of the orthogonal group as represented by the Gel'fand pattern.

Alcarás and Ferreira[2] have constructed bases for a restricted class of the irreducible representations of the orthogonal group. They consider only the case where the first Casimir invariant $I_1^{(n)}$ is not equal to zero. The rest are put to zero, and therefore their representation has only a single row. In this note we obtain representations for any Gel'fand pattern where m_{ij} need not be zero.

The procedure is as follows. Since we already possess from I the normalized lowering operators, it is only necessary to work out a polynomial for the highest weight, in order to obtain all the other polynomials in the irreducible representations of the group. For once the highest-weight polynomial is known, the rest can be simply obtained by applying the normalized lowering operators successively to the highest-weight polynomial.

It is found that the highest-weight polynomial can be obtained by a method very similar to Moshinsky's[3] in the case of the unitary group. This is presented in Sec. 2. In Sec. 3 we show how to obtain the Wigner coefficients of $O(3)$, commonly called the Clebsch–Gordan coefficients of $R(3)$, by means of the polynomials thus obtained. Computation of some Wigner coefficients of $O(5)$ is now in progress.

II. CONSTRUCTION OF THE HIGHEST-WEIGHT POLYNOMIAL P

Since the method we use to find the highest-weight polynomial P is parallel to Moshinsky's method in finding the highest-weight polynomial of the unitary group, we shall first review briefly Moshinsky's argument in the case of the unitary group[4] and then carry over his arguments step by step into the case of the orthogonal group.

Moshinsky defined two kinds of operators $C_{\mu}^{\mu'}$ and

* Present address: Department of Physics, St. Louis University, St. Louis, Mo.

[1] M. K. F. Wong, J. Math. Phys. **8**, 1899 (1967); see also, S. C. Pang and K. T. Hecht, J. Math. Phys. **8**, 1233 (1967).

[2] J. A. C. Alcarás and P. L. Ferreira, J. Math. Phys. **6**, 578 (1965).

[3] M. Moshinsky, Nucl. Phys. **31**, 384 (1962).

[4] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

Minkowskian metric by a 3 + 1 symmetry of a real metric, a model was obtained which simulates the Minkowskian metric for the superposition field, although the basic metric is strictly Riemannian, i.e., positive-definite.

The present investigation does not go beyond the linear approximation. The second-order interaction terms, which must be made responsible for the construction of material particles as excited eigenstates of the field equations, have not been taken into account in this paper.

ACKNOWLEDGMENT

I remember with pleasure the many stimulating discussions with Professor W. R. Davis and his group of the Physics Department, North Carolina State University at Raleigh, North Carolina, during the Spring term, 1968.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Representations of the Orthogonal Group. II. Polynomial Bases for the Irreducible Representations of the Orthogonal Group

M. K. F. WONG*
Research Institute for Natural Sciences, Woodstock College, Woodstock, Maryland

(Received 25 February 1968)

Polynomial bases for the irreducible representations of the orthogonal group, which are characterized by the Gel'fand pattern, have been obtained. The method used is very similar to Moshinsky's and is a generalization from the unitary group to the orthogonal group. The Wigner coefficients of $O(3)$, commonly called the Clebsch–Gordan coefficients of $R(3)$, are rederived by means of the polynomial bases obtained in this paper.

I. INTRODUCTION

In a previous paper,[1] henceforth referred to as I, we obtained normalized lowering and raising operators of the orthogonal group in the group chain $O(n) \supset O(n-1) \supset \cdots \supset O(2)$. As a result, the Gel'fand–Zetlin matrix elements for the generators of the orthogonal group have been obtained. In this note we show how to construct polynomial bases for the irreducible representations of the orthogonal group as represented by the Gel'fand pattern.

Alcarás and Ferreira[2] have constructed bases for a restricted class of the irreducible representations of the orthogonal group. They consider only the case where the first Casimir invariant $I_1^{(n)}$ is not equal to zero. The rest are put to zero, and therefore their representation has only a single row. In this note we obtain representations for any Gel'fand pattern where m_{ij} need not be zero.

The procedure is as follows. Since we already possess from I the normalized lowering operators, it is only necessary to work out a polynomial for the highest weight, in order to obtain all the other polynomials in the irreducible representations of the group. For once the highest-weight polynomial is known, the rest can be simply obtained by applying the normalized lowering operators successively to the highest-weight polynomial.

It is found that the highest-weight polynomial can be obtained by a method very similar to Moshinsky's[3] in the case of the unitary group. This is presented in Sec. 2. In Sec. 3 we show how to obtain the Wigner coefficients of $O(3)$, commonly called the Clebsch–Gordan coefficients of $R(3)$, by means of the polynomials thus obtained. Computation of some Wigner coefficients of $O(5)$ is now in progress.

II. CONSTRUCTION OF THE HIGHEST-WEIGHT POLYNOMIAL P

Since the method we use to find the highest-weight polynomial P is parallel to Moshinsky's method in finding the highest-weight polynomial of the unitary group, we shall first review briefly Moshinsky's argument in the case of the unitary group[4] and then carry over his arguments step by step into the case of the orthogonal group.

Moshinsky defined two kinds of operators $C_{\mu}^{\mu'}$ and

* Present address: Department of Physics, St. Louis University, St. Louis, Mo.

[1] M. K. F. Wong, J. Math. Phys. **8**, 1899 (1967); see also, S. C. Pang and K. T. Hecht, J. Math. Phys. **8**, 1233 (1967).

[2] J. A. C. Alcarás and P. L. Ferreira, J. Math. Phys. **6**, 578 (1965).

[3] M. Moshinsky, Nucl. Phys. **31**, 384 (1962).

[4] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

$C_{ss'}$ by

$$\mathcal{C}_{\mu}^{\mu'} = \sum_{s=1}^{n} a_{\mu s}^{\dagger} a_{s}^{\mu'}, \tag{1}$$

$$C_{ss'} = \sum_{\mu=1}^{n} a_{\mu s}^{\dagger} a_{s'}^{\mu}. \tag{2}$$

He then points out that the highest-weight polynomial should satisfy Cartan's theorem, i.e., it should be unique. This means that the highest-weight polynomial P should satisfy both Eqs. (3) and (4):

$$C_{ss}P = h_s P, \quad C_{ss'}P = 0, \quad s < s', \tag{3}$$

$$\mathcal{C}_{\mu}^{\mu}P = h_s P, \quad \mathcal{C}_{\mu}^{\mu'}P = 0, \quad \mu < \mu'. \tag{4}$$

It is to be noted that $C_{ss'}$ $(s < s')$ and $\mathcal{C}_{\mu}^{\mu'}$ $(\mu < \mu')$ are the raising generators of the unitary group corresponding to the positive roots $e_s - e_{s'}$ or $e_{\mu} - e_{\mu'}$. These can all be obtained by combining the n primitive roots of U_{n+1}, which are C_{ss+1}. Thus, the second parts of (3) and (4) can be further simplified to read

$$C_{ss+1}P = 0, \quad s = 1, 2, \cdots, n-1, \tag{3'}$$

$$\mathcal{C}_{\mu}^{\mu+1}P = 0, \quad \mu = 1, 2, \cdots, n-1, \quad \text{for } U_n. \tag{4'}$$

The equivalence of Eqs. (3) and (4) to the well-known case of the three-dimensional rotation group has also been pointed out by Moshinsky. In that case, the highest-weight polynomial must satisfy Eqs. (5) and (6), corresponding to Eqs. (3) and (4):

$$I_0 P = -\tfrac{1}{2}(j + \tfrac{3}{2})P, \quad I_+ P = 0, \tag{5}$$

where

$$I_0 = (\tfrac{1}{4}i)(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r}) = -\tfrac{1}{2}(\mathbf{r}\cdot\nabla + \tfrac{3}{2}) \quad I_+ = \tfrac{1}{2}p^2,$$

$$L_z = jP, \quad L_+ P = 0, \tag{6}$$

where

$$\mathbf{L} = (\mathbf{r}\times\mathbf{p}), \quad \mathbf{p} = i^{-1}\nabla, \quad L_+ = L_x + iL_y.$$

We now carry over Moshinsky's method from the unitary group to the orthogonal group. Using the notation of I and making use of the concept of primitive roots, we find that the highest-weight polynomial P is defined by the following equations:

For $O(2k+1)$,

$$J_{2\alpha}^{2\alpha-1}P = m_{2k+1,\alpha}P, \quad \alpha = 1, 2, \cdots, k, \tag{7}$$

$$D_{p+1}^{p}P = 0, \quad p = 1, 2, \cdots, k-1, \tag{8}$$

$$E_{2k+1}^{k}P = 0. \tag{9}$$

For $O(2k)$,

$$J_{2\alpha}^{2\alpha-1}P = m_{2k,\alpha}P, \quad \alpha = 1, 2, \cdots, k, \tag{10}$$

$$D_{p+1}^{p}P = 0, \quad p = 1, 2, \cdots, k-1, \tag{11}$$

$$A_{k}^{k-1}P = 0, \tag{12}$$

where the generators J_p^q can be expressed in two different ways:

$$J_p^q(1) = -i\sum_{\mu=1}^{n}\left(x_{\mu p}\frac{\partial}{\partial x_{\mu q}} - x_{\mu q}\frac{\partial}{\partial x_{\mu p}}\right), \tag{13}$$

$$J_p^q(2) = -i\sum_{\mu=1}^{n}\left(x_{p\mu}\frac{\partial}{\partial x_{q\mu}} - x_{q\mu}\frac{\partial}{\partial x_{p\mu}}\right). \tag{14}$$

The definitions of A, D, and E are the same as in I. Cartan's theorem requires that P should satisfy Eqs. (7), (8), (9), or (10), (11), (12), whether J_p^q is expressed by (13) or by (14). If such a polynomial can be found, then it is unique and is therefore the correct highest-weight polynomial.

We now show how P can be obtained. First, in the case of $O(2k)$, define

$$a_{v,t} \equiv x_{2v-1,2t-1} - ix_{2v-1,2t} - ix_{2v,2t-1} - x_{2v,2t}. \tag{15}$$

Using $J_p^q(1)$ as expressed by (13) and rewriting Eqs. (10), (11), and (12) in terms of $a_{v,t}$, we obtain

$$J_{2\alpha}^{2\alpha-1}(1)P = \sum_{v=1}^{k} a_{v\alpha}\frac{\partial}{\partial a_{v\alpha}}P = m_{2k,\alpha}P, \quad \alpha = 1, 2, \cdots, k, \tag{10'}$$

$$D_{p+1}^{p}(1)P = \sum_{v=1}^{k} a_{vp}\frac{\partial}{\partial a_{vp+1}}P = 0, \quad p = 1, 2, \cdots, k-1, \tag{11'}$$

$$A_{k}^{k-1}(1)P = 0 \Rightarrow 0 = 0. \tag{12'}$$

Equation (12') means that any function of $a_{v,t}$ satisfies Eq. (12).

Using $J_p^q(2)$ as expressed by (14) and rewriting Eqs. (10), (11), and (12) in terms of $a_{v,t}$ again, we obtain

$$J_{2\alpha}^{2\alpha-1}(2)P = \sum_{v=1}^{k} a_{\alpha v}\frac{\partial}{\partial a_{\alpha v}}P = m_{2k,\alpha}P, \quad \alpha = 1, 2, \cdots, k, \tag{10''}$$

$$D_{p+1}^{p}(2)P = \sum_{v=1}^{k} a_{pv}\frac{\partial}{\partial a_{p+1,v}}P = 0, \quad p = 1, 2, \cdots, k-1, \tag{11''}$$

$$A_{k}^{k-1}(2)P = 0 \Rightarrow 0 = 0. \tag{12''}$$

Comparing, now, (10'), (11'), and (10''), (11'') with the equivalent Eqs. (3), (3'), and (4), (4') in the case of the unitary group, we find that they are exactly the same. That is, our $J_{2\alpha}^{2\alpha-1}(1)$ is equivalent to Moshinsky's C_{ss} and D_{p+1}^{p} is equivalent to his C_{ss+1}. $J_{2\alpha}^{2\alpha-1}(2)$ is equivalent to his $\mathcal{C}_{\mu}^{\mu}$ and $D_{p+1}^{p}(2)$ is equivalent to his $\mathcal{C}_{\mu}^{\mu+1}$. Thus the arguments used by Moshinsky[3] to obtain the highest-weight polynomial of the unitary group can be completely carried over to the case of the orthogonal group. Moreover, the uniqueness of the

highest-weight polynomial as characterizing an irreducible representation of the group and expressed by the determinants Δ_1^1, Δ_{12}^{12}, etc., is also established following Moshinsky's proof.

Thus, the highest-weight polynomial P, for $O(2k)$, is

$$P = (\Delta_1^1)^{m_{2k,1}-m_{2k,2}}(\Delta_{12}^{12})^{m_{2k,2}-m_{2k,3}} \cdots (\Delta_{12\cdots k}^{12\cdots k})^{m_{2k,2}}, \tag{16}$$

where

$$\Delta_{\mu_1\mu_2\cdots\mu_r}^{s_1s_2\cdots s_r} = \sum_p [(-1)^p p a_{\mu_1 s_1} a_{\mu_2 s_2} \cdots a_{\mu_r s_r}]. \tag{17}$$

(p here means permutation over $s_1 s_2 \cdots s_r$.)

For $O(2k+1)$, Eqs. (7) and (8) are unchanged, i.e., they still go to Eqs. (10′), (11′) and (10″), (11″), with $m_{2k,\alpha}$ replaced by $m_{2k+1,\alpha}$, while Eq. (9) becomes

$$E_{2k+1}^k(1)P = -\sum_{\nu=1}^{k+1} a_{\nu k} \frac{\partial}{\partial a_{\nu,k+1}} P = 0, \tag{9′}$$

$$E_{2k+1}^k(2)P = -\sum_{\nu=1}^{k+1} a_{k\nu} \frac{\partial}{\partial a_{k+1,\nu}} P = 0. \tag{9″}$$

Equations (9′) and (9″) can be satisfied as long as P does not contain $a_{\nu,k+1}$ or $a_{k+1,\nu}$. In this case, $a_{\nu,k+1}$ is defined as follows:

$$a_{\nu,k+1} = x_{2\nu-1,2k+1} - i x_{2\nu,2k+1},$$

since $x_{2\nu-1,2k+2}$ and $x_{2\nu,2k+2}$ do not exist in $O(2k+1)$. Similarly,

$$a_{k+1,\nu} = x_{2k+1,2\nu-1} - i x_{2k+1,2\nu} \quad \text{for} \quad O(2k+1).$$

Therefore, in the case of $O(2k+1)$, P is also expressed by

$$P = (\Delta_1^1)^{m_{2k+1,1}-m_{2k+1,2}} \times (\Delta_{12}^{12})^{m_{2k+1,2}-m_{2k+1,3}} \cdots (\Delta_{12\cdots k}^{12\cdots k})^{m_{2k+1,k}}. \tag{18}$$

It remains only for us to normalize P. This can be done easily. We expand the function P into a polynomial of x_{ij}, and use the condition that

$$(x_{ij}^h, x_{i'j'}^{h'}) = \delta_{ii'}\delta_{jj'}\delta_{hh'}h!$$

When h and h' are half-integers, we write

$$h! = \Gamma(1+h)$$

and using the relations $\Gamma(1+z) = z\Gamma(z)$ and $\Gamma(\frac{1}{2}) = (\pi)^{\frac{1}{2}}$, we obtain the normalization factors for both integer and half-integer eigenvalues of m_{ij}. The procedure of writing $h! = \Gamma(1+h)$ agrees with the prescription given by Bargmann[5] for half-integers h and h', since, according to Bargmann,

$$(z^h, z^{h'}) = \frac{1}{\pi}\int_0^{2\pi} \exp[i(h'-h)\phi]\, d\phi \times \int_0^\infty r^{h+h'+1} e^{-r^2} dr = \delta_{hh'}\Gamma(1+h).$$

In the case of $O(3)$, for example, we obtain for the normalized highest-weight polynomial $\mathfrak{F}$ ($m_{31} = J$):

$$\mathfrak{F} = \frac{1}{2^J\sqrt{J!}} a_{11}^J \tag{19}$$

such that

$$(\mathfrak{F}, \mathfrak{F}) = 1. \tag{20}$$

III. WIGNER COEFFICIENTS OF $O(3)$

In this section we calculate the Wigner coefficients of $O(3)$ by means of the polynomials just obtained. By applying the normalized lowering operators obtained in I, i.e., $N^{-1}L_3^1$, $(J-m)$ times to the normalized highest-weight polynomial $\mathfrak{F}$ in (19), we obtain the normalized polynomial for the irreducible representation $\mathfrak{F}(J, m)$:

$$\mathfrak{F}(J,m) = \frac{1}{2^J\sqrt{J!}}\left[2^{J-m}\frac{J!\,J!}{m!\,m!}\frac{(J-m)!\,2J!}{(J+m)!}\right]^{-\frac{1}{2}} \times [\sqrt{2}\, i(J_3^2 - iJ_3^1)J_2^1]^{J-m} a_{11}^J. \tag{21}$$

When two angular momenta $J^{(1)}$ and $J^{(2)}$ are coupled together, we obtain, for the coupled angular momentum J,

$$J_p^q = J_p^{q(1)} + J_p^{q(2)}; \quad p, q = 1, 2, 3; \quad p > q. \tag{22}$$

Note that $J_p^{q(1)}$ and $J_p^{q(2)}$ are not to be confused with $J_p^q(1)$ and $J_p^q(2)$ in Eqs. (13) and (14). Then the highest weight $P(J, J)$ of the coupled angular-momentum system satisfies the following two equations:

$$J_2^1 P(J,J) = JP(J,J), \tag{23}$$

$$E_3^1 P(J,J) = 0. \tag{24}$$

We see that (23) is satisfied if we put

$$P(J,J) = \sum_{m_1} A_{m_1}\mathfrak{F}^{(1)}(J_1, m_1)\mathfrak{F}^{(2)}(J_2, J-m_1), \tag{25}$$

where $\mathfrak{F}^{(1)}$ and $\mathfrak{F}^{(2)}$ are restricted to the systems (1) and (2), respectively. Applying now Eq. (24) to Eq. (25) and remembering that

$$E_3^1 = E_3^{1(1)} + E_3^{1(2)}, \tag{26}$$

we obtain a recurrence relation for A_{m_1}, i.e.,

$$\frac{A_{m_1+1}}{A_{m_1}} = -\left[\frac{(J_1-m_1)(J_1+m_1+1)}{(J_2-J+m_1+1)(J_2+J-m_1)}\right]^{\frac{1}{2}}. \tag{27}$$

[5] V. Bargmann, Comm. Pure Appl. Math. **14**, 187 (1961); Rev. Mod. Phys. **34**, 829 (1962).

This equation is the same as Eq. (2.17) of Moshinsky,[4] and is the same as Eq. (3.6.1) of Edmonds,[6] if we write (27) as

$$\frac{A_{m_1}}{A_{m_1-1}} = -\left[\frac{(J_1 - m_1 + 1)(J_1 + m_1)}{(J_2 - J + m_1)(J_2 + J - m_1 + 1)}\right]^{\frac{1}{2}}. \tag{27'}$$

In fact A_{m_1} is already the Wigner coefficient

$$(J_1 J_2 m_1 J - m_1 \,|\, JJ)$$

for $m = J$.

If we put $m_1 = J_1$ in (27') first, and then apply the recurrence relation (27') $(J_1 - m_1)$ times to A_{J_1}/A_{J_1-1}, we obtain

$$\frac{A_{J_1}}{A_{m_1}} = (-1)^{J_1-m_1} \times \left[\frac{(J_2 + J - J_1)!\,(2J_1)!\,(J_2 - J + m_1)!\,(J_1 - m_1)!}{(J_2 + J - m_1)!\,(J_1 + m_1)!\,(J_2 - J + J_1)!}\right]^{\frac{1}{2}}. \tag{28}$$

By normalizing $P(J, J)$ in Eq. (24), we obtain

$$\sum_{m_1=-J_1}^{J} A_{m_1} A_{m_1}^{*} = 1. \tag{29}$$

Combining (28) and (29), we obtain

$$A_{J_1} = \left[\frac{(2J_1)!\,(2J+1)!}{(J_1 + J_2 + J + 1)!\,(J_1 - J_2 + J)!}\right]^{\frac{1}{2}}. \tag{30}$$

Combining (28) and (30) we obtain A_{m_1}.

To obtain the general Wigner coefficient for any m, we can apply the lowering operator $F_3^1 = F_3^{1(1)} + F_3^{1(2)}$ to Eq. (25) and obtain a difference equation between the Wigner coefficients, and applying the finite-difference technique used by Edmonds,[6] we obtain from A_{m_1} the general Wigner coefficients $(J_1 J_2 m_1 m_2 \,|\, Jm)$. The formula so obtained agrees with that first obtained by Racah.[7]

CONCLUSION

In conclusion, we see that our polynomial basis is a function of x_{ij} $(i, j = 1, 2, \cdots, n)$ for $O(n)$, which has dimension n^2, whereas before polynomials for $O(n)$ have only been constructed in an n-dimensional space, e.g., the spherical harmonics, from which only a restricted class of the irreducible representations of $O(n)$ can be obtained. By enlarging the dimensionality of the space from n to n^2, we are able to obtain all the irreducible representations of $O(n)$. Our construction of the highest-weight polynomial P is greatly helped by the work of Moshinsky on the unitary group. In this formalism one has only to solve some first-order partial differential equations. It is remarkable that our highest-weight polynomial for $O(2k)$ and $O(2k + 1)$ has the same form as the highest-weight polynomial for the unitary group $U(k)$. In order to gain some confidence in the polynomial basis we have so obtained, we have shown how to rederive the Wigner coefficients of $O(3)$ with the help of these normalized polynomials.

ACKNOWLEDGMENT

This work was supported by the National Aeronautics and Space Administration, under Contract No. NsG-670.

[6] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1957).

[7] G. Racah, Phys. Rev. **62**, 438 (1942).

Codiagonal Perturbations

H. S. GREEN
Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia
AND
T. TRIFFET*
Michigan State University, East Lansing, Michigan

(Received 14 November 1968)

Matrix methods are developed for calculating the eigenvalues and eigenvectors of a large class of quantum-mechanical operators which may be regarded as perturbed forms of special-function operators. Specific representations are obtained for the latter, including all of the important cases treated by Infeld and Hull. To these are added representations for terms sufficient to generate forms corresponding to the Mathieu equation, the Lamé equation, and others. A rapidly convergent computational scheme applicable to asymmetric matrices, which retains its stability even when the perturbing terms become large, is described; and its use is illustrated by application to the operator $p(1-q^2)p-\alpha^2q^2$, corresponding to the Legendre-like form $(d/dx)(1-x^2)(d/dx)+\alpha^2x^2$. Though group-theoretic considerations are stressed, appropriate correlations with differential and integral equations are presented throughout.

1. INTRODUCTION

Because Schrödinger's formulation of quantum mechanics leads to irreducible divergent integrals for the infinite-degree-of-freedom systems of quantum electrodynamics, Dirac has expressed the view that it should be regarded as a limiting case of the matrix mechanics originally proposed by Heisenberg.[1] We believe this to be true, and also feel that the algebraic character of the latter gives it a considerable computational (and philosophical) advantage, in spite of the fact that the former has been developed much more extensively. Accordingly, we have undertaken to extend the matrix methods initiated by Born and Jordan[2] and to develop a compatible perturbation technique. The results are presented here, along with certain other features which appeared while pursuing these objectives—most notably, a systematic procedure for calculating specialized step operators and a simple method of handling large perturbations.

It is a curious historical circumstance that one of the basic ideas on which our developments depend was suggested by Schrödinger himself.[3] This is the factorization concept, exploited from the differential equation point of view by Infeld and Hull.[4] In its algebraic form, it has previously been stated in a preliminary way and used to solve a number of fundamental quantum-mechanical problems by Green.[5] Also, it seems to have evolved from differential equation considerations and led a separate existence in the development of computational theory.[6] Another essential concept, that of a codiagonal matrix, represents an adaptation for our purposes of a general definition originally framed by Lanczos,[7] but since used in a more specialized sense by others.[8] Both of these terms will be fully explained as they appear in subsequent sections.

Since our emphasis will be on an operator algebra in which differential forms do not explicitly appear, it may be helpful to display in advance the type of differential equations which may be solved by the codiagonal perturbation method. One example is the Legendre-like eigenequation

$$\left[\frac{d}{dx}(1-x^2)\frac{d}{dx}+\alpha^2x^2+a_l\right]u_l(x)=0, \quad (1.1)$$

which in operational form, featuring the quantum-mechanical conjugate variables q and p satisfying $qp-pq=i$ (with $h=1$), becomes

$$[A-a^{(l)}]\psi^{(l)}=[p(1-q^2)p-\alpha^2q^2-a^{(l)}]\psi^{(l)}=0, \quad (1.2)$$

with α representing a constant of any magnitude. The solution of this equation is used to illustrate each part of the present work. A second example, which should suffice to give some idea of the range of problems which can be solved, is the special Lamé

* Robertson Research Fellow in Mathematical Physics at the University of Adelaide for 1966–67, during which time most of this work was completed; currently associated with the Department of Metallurgy, Mechanics and Materials Science and the Center for Applied Mathematics at Michigan State University.

[1] P. A. M. Dirac, Nature **203**, 115 (1964).

[2] M. Born and P. Jordan, *Elementare Quantenmechanik* (Springer-Verlag, Berlin, 1930).

[3] E. Schrödinger, Proc. Roy. Irish Acad. **A46**, 9 (1940).

[4] L. Infeld and T. E. Hull, Rev. Mod. Phys. **23**, 21 (1951).

[5] H. S. Green, *Matrix Mechanics* (P. Noordhoff Ltd., Groningen, The Netherlands, 1965).

[6] J. G. F. Francis, Computer J. **4**, 265, 332 (1961).

[7] C. Lanczos, *Applied Analysis* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1956).

[8] C. Strachey and J. G. F. Francis, Computer J. **4**, 168 (1961).

eigenequation

$$\left[(1-\kappa x^2)\frac{d}{dx}(1-x^2)\frac{d}{dx} - (3-\kappa-2\kappa x^2)x\frac{d}{dx} - \lambda x^2 + a_l\right]u_l(x) = 0 \quad (1.3)$$

or

$$[A - a^{(l)}]\psi^{(l)} = [(1-\kappa q^2)\{p(1-q^2)p + 3iqp\} - i\kappa(1-q^2)qp + \lambda q^2 - a^{(l)}]\psi^{(l)} = 0, \quad (1.4)$$

where $\kappa < 1$ and λ is an unrestricted constant. A general form of this equation is one of the cases treated in Sec. 4.

2. THE OPERATOR $A = \Gamma B + \Delta$

A. General Properties

Examination of the operators given in (1.2) and (1.4) above will establish that both are of the form

$$A = \Gamma(p,q)B(p,q) + \Delta(p,q). \quad (2.1)$$

In the first example,

$$B = p(1-q^2)p, \quad \Gamma = 1, \quad \Delta = -\alpha^2 q^2, \quad (2.2)$$

while in the second

$$B = p(1-q^2)p + 3iqp, \quad \Gamma = (1-\kappa q^2), \quad \Delta = -i\kappa(1-q^2)qp + \lambda q^2, \quad (2.3)$$

with $\kappa < 1$. We shall now formulate the general requirements to be satisfied by the operators B, Γ, and Δ.

We may identify B as an unperturbed linear operator, and Γ and Δ as perturbations. The operator B is defined on an associated Hilbert space H. It is not necessarily Hermitian, but is self-adjoint in the sense that

$$\bar{B} \equiv \eta B^* \eta^{-1} = B, \quad (2.4)$$

where B^* denotes the Hermitian conjugate, and η is a positive-definite linear operator ($\eta = 1$ in both examples cited above). The eigenvalues $b^{(j)}$ of B will then be real, and we shall suppose that they are distinct and bounded below, but not above, so that, suitably ordered, they form an indefinitely increasing sequence starting with a finite value $b^{(1)}$. The operators Γ and Δ are nonsingular (or satisfy certain weaker conditions which will be formulated later) on the Hilbert space H; we shall suppose further that Γ is positive-definite.

Although Γ and Δ are not required to be self-adjoint, we suppose that A is self-adjoint in the sense that

$$A \equiv \eta' A^* \eta'^{-1} = A, \quad (2.5)$$

where η' is possibly different from η. Under these circumstances, the eigenvalues $a^{(j)}$ of A, like those of B, are real and form an indefinitely increasing sequence, starting with a finite value $a^{(1)}$. If A' is an operator which does not satisfy these conditions, it will normally be possible to find a transformation

$$A = F(A'), \quad (2.6)$$

such that the transformed operator A is of the required type.

It should be pointed out that most of the developments to follow will apply to any operator A of the general form described. But, because of the wealth of important applications, we have in mind especially the possibility that A depends, as in the above examples, on a coordinate q and the conjugate momentum p or, more generally, on a number of such coordinates and momenta.

In many of the examples, the operator B will be of the form

$$B = f_1(q)p^2 + 2if_2(q)p + f_3(q). \quad (2.7)$$

The "singularities" of operators of this type are zeros of $f_1(q)$ or poles of $f_3(q)$. If $q^{(1)}$ and $q^{(2)}$ are such singularities (numbers), and there is no other singularity in the interval $(q^{(1)}, q^{(2)})$, the Hilbert space will be defined in such a way that

$$q^{(1)} < (\phi, q\phi) < q^{(2)} \quad (2.8)$$

for an arbitrary normalized vector ϕ, where $(\ ,\ )$ represents the scalar product. Because of the restrictions placed on Γ and Δ, the singularities of A will be identical with those of B.

B. Eigenvalues and Eigenvectors

The equation to be solved is

$$(\Gamma B + \Delta)\psi^{(l)} = a^{(l)}\psi^{(l)}, \quad (2.9)$$

wherein it is assumed that the eigenvalues $b^{(j)}$ and the corresponding eigenvectors $\phi^{(j)}$ of B can be found without approximation. Let us expand the perturbed eigenvectors in terms of $\phi^{(j)}$ and a set of undetermined numerical coefficients $\psi_j^{(l)}$:

$$\psi^{(l)} = \sum_j \psi_j^{(l)}\phi^{(j)}. \quad (2.10)$$

Now, if we premultiply the eigenequation of B by Γ,

$$\Gamma B\phi^{(j)} = \Gamma b^{(j)}\phi^{(j)}, \quad (2.11)$$

and make the appropriate substitutions in (2.9), the

result is

$$\sum_j \psi_j^{(l)} \Delta \phi^{(j)} = \sum_j \psi_j^{(l)} (a^{(l)} - b^{(j)} \Gamma) \phi^{(j)}. \quad (2.12)$$

Applying the adjoint $\bar{\phi}^{(i)} = \phi^{(i)*} \eta^{-1}$, and supposing that $\phi^{(i)}$ is normalized in the usual way ($\bar{\phi}^{(i)} \phi^{(j)} = \delta_{ij}$), this becomes

$$\sum_j \Delta_{ij} \psi_j^{(l)} = a^{(l)} \psi_i^{(l)} - \sum_j b^{(j)} \Gamma_{ij} \psi_j^{(l)}, \quad (2.13)$$

where

$$\Delta_{ij} = \bar{\phi}^{(i)} \Delta \phi^{(j)}, \quad (2.14)$$

$$\Gamma_{ij} = \bar{\phi}^{(i)} \Gamma \phi^{(j)}. \quad (2.15)$$

It remains to solve the system of equations

$$\sum_j (\Delta_{ij} + b^{(j)} \Gamma_{ij}) \psi_j^{(l)} = a^{(l)} \psi_i^{(l)}; \quad (2.16)$$

but now it is apparent that we may proceed by separately diagonalizing the matrix B, then determining the elements of Γ and Δ in terms of the resulting matrix to facilitate diagonalization of the total expression bracketed above. The perturbed eigenvalues $a^{(l)}$ will be obtained as a direct result, while the perturbed eigenvectors are determined by the $\psi_j^{(l)}$.

In what follows, we develop methods of identifying these three matrices, of calculating their elements, and of diagonalizing them in such a way that rapid convergence will occur in most cases irrespective of the magnitude of the elements. Matrix representations of the coordinate and momentum variables q and p in terms of the diagonal form of B are also presented, so that the elements of $\Gamma(q, p)$ and $\Delta(q, p)$ can be obtained without recourse to the relations (2.14) and (2.15).

3. THE OPERATOR B

A. Eigenvalues and Eigenvectors

We consider first the problem of finding the eigenvalues and eigenvectors of the unperturbed operator B. Here, our object will be to obtain and tabulate a large number of exact results, using only matrix techniques, to simplify the identification of this operator.

The eigenvalues of any matrix which, like those of B, are bounded below, can be found by constructing a sequence of operators, $\theta_1, \theta_2, \cdots$ and corresponding numerical multiples of the unit operator, $b^{(1)}$, $b^{(2)}, \cdots$, such that

$$B = B_1 = \bar{\theta}_1 \theta_1 + b^{(1)},$$
$$B_2 = \theta_1 \bar{\theta}_1 + b^{(1)},$$
$$\cdots\cdots\cdots$$
$$B_j = \bar{\theta}_j \theta_j + b^{(j)}, \quad (3.1)$$
$$B_{j+1} = \theta_j \bar{\theta}_j + b^{(j)}, \quad (3.2)$$
$$\cdots\cdots\cdots$$

where $\bar{\theta}_j = \eta \theta_j^* \eta^{-1}$ is the adjoint of θ_j (defined so that B is self-adjoint), and $b^{(j)}$ has its maximum value at each step. When this is done, the jth eigenvalue of B, in numerically ascending order, is $b^{(j)}$. The proof is as follows.

By using the relations (3.1) and (3.2) we may easily establish that

$$\bar{\theta}_1 \bar{\theta}_2 \cdots \bar{\theta}_m \theta_m \cdots \theta_2 \theta_1 = \prod_{j=1}^{m} (B - b^{(j)}). \quad (3.3)$$

The left side of (3.3) can be written as $\eta \theta^* \eta^{-1} \theta$, where $\theta = \theta_m \cdots \theta_2 \theta_1$ or, since η is positive-definite, as

$$\eta^{\frac{1}{2}} (\eta^{-\frac{1}{2}} \theta \eta^{\frac{1}{2}})^* (\eta^{-\frac{1}{2}} \theta \eta^{\frac{1}{2}}) \eta^{-\frac{1}{2}}; \quad (3.4)$$

its eigenvalues must therefore be nonnegative. Thus, if b is any eigenvalue of B,

$$\prod_{j=1}^{m} (b - b^{(j)}) \geq 0 \quad (3.5)$$

for all m. It follows from this set of inequalities that the only eigenvalues less than the upper bound of the increasing sequence $b^{(1)}, b^{(2)}, \cdots$, are the $b^{(j)}$ themselves. To determine the eigenvector corresponding to $b^{(j)}$, we note that if ϕ_j is the vector satisfying

$$\theta_j \phi_j = 0, \quad (3.6)$$

and

$$\phi^{(j)} = \bar{\theta}_1 \bar{\theta}_2 \cdots \bar{\theta}_{j-1} \phi_j, \quad (3.7)$$

then it follows from (3.1) and (3.2) that

$$B\phi^{(j)} = \bar{\theta}_1 \bar{\theta}_2 \cdots \bar{\theta}_{j-1} (\bar{\theta}_j \theta_j + b^{(j)}) \phi_j$$
$$= b^{(j)} \phi^{(j)}. \quad (3.8)$$

This argument is a simple generalization of one given for Hermitian operators by one of the authors elsewhere.[5] It is the algebraic formulation of the factorization concept referred to earlier. Here we wish to draw attention to the fact that the θ_j, and the B_j, determined by this method are not unique. If U_j is unitary adjoint ($\bar{U}_j U_j = U_j \bar{U}_j = 1$), we can replace θ_j by $U_j \theta_j$ at any stage of the factorization, so that B_{j+1} is replaced by $U_j B_{j+1} \bar{U}_j$. This, of course, does not affect the eigenvalues obtained. One particular choice of the θ_j is especially useful for setting up matrix representations of the various operators, however: that which makes the B_j all commute with one another.

To illustrate this matter, let us return to the operator listed under (2.2),

$$B = p(1 - q^2)p, \quad (3.9)$$

whose eigenvectors in the coordinate (Schrödinger's) representation are the Legendre polynomials. The

singularities of this operator are $q^{(1)} = -1$ and $q^{(2)} = 1$. The θ_j obtained most naturally are

$$\theta_1 = (1 - q^2)^{\frac{1}{2}}p, \tag{3.10}$$

$$\theta_j = (1 - q^2)^{\frac{1}{2}}p - i(j - 1)q(1 - q^2)^{-\frac{1}{2}}, \tag{3.11}$$

and these are sufficient to obtain the eigenvalues

$$b^{(j)} = j(j - 1) \tag{3.12}$$

of B. However, the

$$B_j = p(1 - q^2)p + (j - 1)^2/(1 - q^2) \tag{3.13}$$

obtained with this choice of the θ_j do not commute with one another. On the other hand, as will be demonstrated below, if we write

$$B = M(M + 1), \tag{3.14}$$

we can use alternatively

$$\theta_j = [(2M + 1)(M + 1)(M + 2j)/(2M + 3)]^{\frac{1}{2}} \times [q + i(1 - q^2)p/M] \tag{3.15}$$

and this will yield

$$B_j = (M + j - 1)(M + j), \tag{3.16}$$

which commutes with B.

In the present section, we shall generally restrict ourselves to use of the particular θ_j associated with B_j which do commute with one another. Then we shall be able to express θ_j in the form

$$\theta_j = \epsilon(B_j - b^{(j)})^{\frac{1}{2}}, \tag{3.17}$$

where

$$\epsilon\bar{\epsilon} = 1, \quad \bar{\epsilon}\epsilon = 1 - \rho; \tag{3.18}$$

$$(B_j - b^{(j)})\rho = 0; \tag{3.19}$$

$$\epsilon B_j = B_{j+1}\epsilon. \tag{3.20}$$

The isolation of the "step" operator ϵ allows us to establish a representation in which the elements of B_j and θ_j are

$$(B_j)_{kl} = b^{(j+k+1)}\delta_{kl}, \tag{3.21}$$

$$(\theta_j)_{kl} = (b^{(j+k)} - b^{(j)})^{\frac{1}{2}}\delta_{k+1\,l}, \tag{3.22}$$

$$(\bar{\theta}_j)_{kl} = (b^{(j+l)} - b^{(j)})^{\frac{1}{2}}\delta_{k\,l+1}. \tag{3.23}$$

B. Matrix Representations

Let us now consider the problem of finding a matrix representation for a linear operator B, self-adjoint in the sense already described, and a self-adjoint operator σ which does not commute with it. In the applications, B will be a function of one or more coordinates and their conjugate momenta, and then σ may be identified with some elementary function of the coordinates. From B and σ, we construct a sequence of linearly independent operators $\sigma_1 = \sigma$, $\sigma_2, \sigma_3, \cdots$, by forming successive commutations with B:

$$\begin{aligned} B\sigma_1 - \sigma_1 B &= \sigma_1 c_{11} + \sigma_2 c_{21}, \\ B\sigma_2 - \sigma_2 B &= \sigma_1 c_{12} + \sigma_2 c_{22} + \sigma_3 c_{32}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ B\sigma_k - \sigma_k B &= \sum_j \sigma_j c_{jk}, \end{aligned} \tag{3.24}$$

where the c_{jk} are either numerical constants or operators which commute with B; they will, for instance, often depend on B itself. The choice of $\sigma_2, \sigma_3, \cdots$ is not unique; therefore, neither are the coefficients c_{jk}. But the c_{jk} always form an almost triangular matrix in which c_{jk} vanishes if $j > k + 1$. They are analogous to the structure constants of a Lie algebra; however, the σ_j and B may not always constitute elements of such an algebra. The commutators of the σ_j among themselves need not be expressible in terms of the σ_j and need not form a finite sequence, even when the above sequence of equations terminates. In the applications which we shall consider in this section, the latter sequence will, in fact, always terminate; and since most of the above considerations apply equally well to the general operator A, it is this termination which in practice distinguishes B from A.

For the Legendre operator (3.9),

$$B = p(1 - q^2)p,$$

taking $\sigma_1 = q$, we find

$$\begin{aligned} [B, \sigma_1] &= Bq - qB \\ &= -2i(1 - q^2)p + 2q \\ &= 2(\sigma_1 + \sigma_2), \end{aligned} \tag{3.25}$$

where $\sigma_2 = -i(1 - q^2)p$, and

$$\begin{aligned} [B, \sigma_2] &= [p, -i(1 - q^2)p](1 - q^2)p \\ &= 2\sigma_1 B. \end{aligned} \tag{3.26}$$

Suppose, in general, that the eigenvalues of the matrix c_{jk} are $\lambda^{(l)}$, and that the corresponding right- and left-eigenvectors are $\xi^{(l)}$ and $\zeta^{(l)}$:

$$\sum_k c_{jk}\xi_k^{(l)} = \lambda^{(l)}\xi_j^{(l)}, \tag{3.27}$$

$$\sum_j \zeta_j^{(l)}c_{jk} = \lambda^{(l)}\zeta_j^{(l)} \tag{3.28}$$

where we may normalize the $\zeta^{(l)}$ so that

$$\sum_j \zeta_j^{(k)}\xi_j^{(l)} = \delta_{kl}. \tag{3.29}$$

Furthermore, let

$$\sigma^{(l)} = \sum_k \sigma_l \xi_k^{(l)}, \tag{3.30}$$

so that

$$\sigma_k = \sum_l \sigma^{(l)}\zeta_k^{(l)}. \tag{3.31}$$

Then

$$[B, \sigma^{(l)}] = \sigma^{(l)}\lambda^{(l)}, \tag{3.32}$$

where in general $\lambda^{(l)} = \lambda^{(l)}(B)$. It follows that if $\psi^{(j)}$ is an eigenvector of B, corresponding to the eigenvalue $b^{(j)}$,

$$\begin{aligned} B(\sigma^{(l)}\psi^{(j)}) &= \sigma^{(l)}(B + \lambda^{(l)})\psi^{(j)} \\ &= [b^{(j)} + \lambda^{(l)}(b^{(j)})]\sigma^{(l)}\psi^{(j)}. \end{aligned} \tag{3.33}$$

Thus, if $\sigma^{(l)}\psi^{(j)}$ does not vanish, it is also an eigenvector of B, and $b^{(j)} + \lambda^{(l)}(b^{(j)})$ is the corresponding eigenvalue.

There are several possibilities to be considered, depending on the number of $\lambda^{(l)}$ and the relations between them. In most of the applications to be considered, there will be only two nonvanishing $\lambda^{(l)}$. Then we choose the one which is positive in the range of eigenvalues of interest; calling this $\lambda^{(l)}$, the relation

$$b^{(j+1)} = b^{(j)} + \lambda^{(1)}(b^{(j)}) \tag{3.34}$$

is a difference equation connecting the eigenvalues in that range. It follows from the fact that $\sigma_1 = \sum_l \sigma^{(l)}\zeta_1^{(l)}$ is self-adjoint, that the other nonvanishing $\lambda^{(l)}$, which we call $\lambda^{(2)}$, is negative and has the property

$$b^{(j)} = b^{(j+1)} + \lambda^{(2)}(b^{(j+1)}). \tag{3.35}$$

There may be a third $\lambda^{(l)}$, $\lambda^{(3)}$, which vanishes. Successive normalized eigenvectors are related by

$$\sigma^{(1)}\psi^{(j)} = \alpha_j\psi^{(j+1)}, \tag{3.36}$$

$$\sigma^{(2)}\psi^{(j+1)} = \beta_j\psi^{(j)}, \tag{3.37}$$

where α_j and β_j are constants which determine the matrix elements of $\sigma^{(1)}$ and $\sigma^{(2)}$ in a representation in which B is diagonal. The first eigenvector $\psi^{(1)}$ satisfies $\sigma^{(2)}\psi^{(1)} = 0$.

This possibility is illustrated by the example $B = p(1 - q^2)p$. There we obtain $\lambda^{(1)} = 1 + (1 + 4B)^{\frac{1}{2}}$ and $\lambda^{(2)} = 1 - (1 + 4B)^{\frac{1}{2}}$; but, as mentioned in connection with (3.14), to remove the square root we set $B = M(M + 1)$, so that $\lambda^{(1)} = 2(M + 1)$ and $\lambda^{(2)} = -2M$. Then, since $\xi^{(1)} = (M + 1, 1)$ and $\xi^{(2)} = (M, -1)$,

$$\sigma^{(1)} = q(M + 1) - i(1 - q^2)p, \tag{3.38}$$

$$\sigma^{(2)} = qM + i(1 - q^2)p. \tag{3.39}$$

These are the recurrence relations of the Legendre polynomials in operational form. Also,

$$M\sigma^{(1)} = \sigma^{(1)}(M + 1), \tag{3.40}$$

$$M\sigma^{(2)} = \sigma^{(2)}(M - 1). \tag{3.41}$$

The introduction of an operator like M, whose eigenvalues differ by unity, is a device which in many instances simplifies the algebra.

If there are more than two nonvanishing $\lambda^{(l)}$, including two which are positive, suppose these to be $\lambda^{(1)}$ and $\lambda^{(3)}$. It may happen that the corresponding $\sigma^{(1)}$ and $\sigma^{(3)}$ are connected by a relation of the type

$$\sigma^{(1)n} = \sigma^{(3)m}s(B). \tag{3.42}$$

In this case, it is convenient to introduce an operator σ such that

$$\sigma^{(1)} = \sigma^m s^{(1)}(B), \tag{3.43}$$

$$\sigma^{(3)} = \sigma^n s^{(3)}(B), \tag{3.44}$$

$$B\sigma = \sigma[B + \lambda(B)]. \tag{3.45}$$

The operator σ then plays the part of $\sigma^{(1)}$ in the previous discussion, and λ plays the part of $\lambda^{(1)}$. However, it may happen that there is no relation of the type mentioned between $\sigma^{(1)}$ and $\sigma^{(3)}$; a simple example arises if

$$B = p_1^2 + p_2^2 + \alpha q_1^2 + 2\gamma q_1 q_2 + \beta q_2^2, \tag{3.46}$$

where $[q_1, p_1] = [q_2, p_2] = i$ and other commutators of the q's and p's vanish. The most direct way of dealing with this possibility is to introduce a product representation, in which the eigenvectors of B are denoted by $\psi^{(j,k)}$, with

$$\sigma^{(1)}\psi^{(j,k)} = \alpha_j^{(1)}\psi^{(j+1,k)}, \tag{3.47}$$

$$\sigma^{(3)}\psi^{(j,k)} = \alpha_j^{(3)}\psi^{(j,k+1)}. \tag{3.48}$$

Since

$$\sigma^{(1)}\sigma^{(2)}\psi^{(j+1)} = \alpha_j\beta_j\psi^{(j+1)}, \tag{3.49}$$

$$\sigma^{(1)}\sigma^{(2)}\psi^{(1)} = 0, \tag{3.50}$$

$\sigma^{(1)}\sigma^{(2)}$ is an operator which both commutes with B and can be expressed as a function of B,

$$\sigma^{(1)}\sigma^{(2)} = f(B). \tag{3.51}$$

The first eigenvalue of B is determined from the condition $\sigma^{(1)}\sigma^{(2)}\psi^{(1)} = 0$, which yields

$$f(b^{(1)}) = 0; \tag{3.52}$$

subsequent eigenvalues can then be obtained by using the relation (3.34),

$$b^{(j+1)} = b^{(j)} + \lambda^{(1)}(b^{(j)}).$$

To illustrate the method of determining the function $f(B)$, we consider again the example

$$B = p(1 - q^2)p = M(M + 1),$$

noting that

$$\begin{aligned}\sigma^{(1)}\sigma^{(2)} &= q(M+1)\sigma^{(2)} - i(1-q^2)p\sigma^{(2)} \\ &= q\sigma^{(2)}M - i(1-q^2)p\sigma^{(2)} \\ &= M^2 - (1-q^2)M(M+1) \\ &\quad + (1-q^2)p(1-q^2)p \\ &= M^2. \end{aligned} \tag{3.53}$$

According to the prescription $f(b^{(1)}) = 0$, the least eigenvalue of M must be zero, which means that the least eigenvalue of B must be zero also.

To determine the matrix elements of $\sigma^{(1)}$ and $\sigma^{(2)}$, we also need a relation between $\sigma^{(2)}$ and the adjoint $\bar{\sigma}^{(1)}$ of $\sigma^{(1)}$. It is clear that such a relation must exist, for it is easy to show that B commutes with $\sigma^{(1)}\bar{\sigma}^{(1)}$, which must therefore be a function of B, in view of our hypothesis that B is nondegenerate. Thus, we write

$$\bar{\sigma}^{(1)} = \sigma^{(2)}g(B), \tag{3.54}$$

where $g(B)$ is a function which, like $f(B)$, is easily calculated. Now,

$$\begin{aligned}\sigma^{(1)}\bar{\sigma}^{(1)} &= f(B)g(B) \\ &= [h^{(1)}(M-1)]^2, \end{aligned} \tag{3.55}$$

say, and

$$\begin{aligned}\bar{\sigma}^{(2)}\sigma^{(2)} &= f(B)/g(B) \\ &= [h^{(2)}(M-1)]^2. \end{aligned} \tag{3.56}$$

Clearly, if ϵ is the step operator introduced in Eqs. (3.17)–(3.20),

$$\sigma^{(1)} = h^{(1)}(M-1)\bar{\epsilon} = \bar{\epsilon}h^{(1)}(M), \tag{3.57}$$

$$\sigma^{(2)} = \epsilon h^{(2)}(M-1) = h^{(2)}(M)\epsilon. \tag{3.58}$$

The matrix elements of $\sigma^{(1)}$ and $\sigma^{(2)}$, in the representation in which B is diagonal, are given by

$$(\sigma^{(1)})_{jk} = \delta_{j\,k+1}h^{(1)}(m^{(k)}), \tag{3.59}$$

$$(\sigma^{(2)})_{jk} = \delta_{j+1\,k}h^{(2)}(m^{(j)}), \tag{3.60}$$

where $m^{(j)} = m^{(1)} + j - 1$ is the jth eigenvalue of M. Matrix elements of the other $\sigma^{(l)}$, if any, can be inferred in a similar way. Since the $\sigma_k = \sum_l \sigma^{(l)}\zeta_k^{(l)}$ are expressed in terms of the $\sigma^{(l)}$, their matrix elements are also readily available.

We again illustrate the procedure with the help of the example $B = p(1-q^2)p$. In this particular instance the adjoint is simply the Hermitian conjugate and we have

$$\begin{aligned}\bar{\sigma}^{(1)} &= (M+1)q + ip(1-q^2) \\ &= (M-1)q + i(1-q^2)p. \end{aligned} \tag{3.61}$$

To shift the multiplier $(M-1)$ to a position on the right of the term in which it occurs, we use

$$\begin{aligned} q = \sigma_1 &= (\sigma^{(1)} + \sigma^{(2)})/(2M+1), \\ (M-1)q &= [\sigma^{(1)}M + \sigma^{(2)}(M-2)]/(2M+1), \\ -i(1-q^2)p &= \sigma_2 \\ &= [\sigma^{(1)}M - \sigma^{(2)}(M+1)]/(2M+1), \end{aligned} \tag{3.62}$$

whence

$$\bar{\sigma}^{(1)} = \sigma^{(2)}(2M-1)/(2M+1). \tag{3.63}$$

Thus, if $B = M(M+1)$,

$$h^{(1)}(M) = M[(2M+1)/(2M+3)]^{\frac{1}{2}}, \tag{3.64}$$

$$h^{(2)}(M) = M[(2M+3)/(2M+1)]^{\frac{1}{2}}. \tag{3.65}$$

C. Exactly Soluble Problems

Most operators of the type $B(q,p)$, characterized by the fact that they can be diagonalized exactly, are connected with the hypergeometric differential equation or one of its limiting forms. Listed below are results obtained by the preceding method for a variety of operators of this type, which are bounded below but unbounded above. Many additional results can be obtained from them by one or the other of two transformations, which we shall first describe.

The first type includes canonical transformations of the coordinate and momentum. If

$$Q = Q(q), \tag{3.66}$$

$$P = \tfrac{1}{2}[Q'(q)]^{-1}p + \tfrac{1}{2}[Q'(q)]^{-1}, \tag{3.67}$$

where Q is any function with a derivative Q', then $[Q,P] = i$ follows from $[q,p] = i$. Some of the more frequently used transformations are:

$$Q = \beta q + \gamma, \tag{3.68}$$

$$Q = q^{\beta}, \tag{3.69}$$

$$Q = (q-\beta)/(q-\gamma). \tag{3.70}$$

If the operator B whose eigenvalues are required involves a transcendental function $Q(q)$, such as

$$Q = \cos(\beta q + \gamma), \tag{3.71}$$

$$Q = \exp(-\beta q), \tag{3.72}$$

then an obvious first step will be to eliminate this function by the appropriate canonical transformation. One or two examples are given below, but the list is not intended to be exhaustive.

The second type of transformation which can be usefully employed includes unitary and similarity transformations on B itself:

$$B \to B' = SBS^{-1}. \tag{3.73}$$

There is a corresponding transformation,

$$\psi^{(j)} \to \psi^{(j)\prime} = S\psi^{(j)}, \tag{3.74}$$

of the eigenvectors of B. An arbitrary nonsingular S can be resolved into two factors, one of which is unitary-adjoint and the other self-adjoint. If S is unitary-adjoint ($S\bar{S} = 1$), the form of η is not affected by the transformation, i.e., $B' = \eta B'^{*}\eta^{-1}$; but if S is self-adjoint,

$$\eta = \eta' = S^2\eta. \tag{3.75}$$

In the following list of exactly soluble examples, we tabulate (a) the operator B, its singularities $q^{(1)}$ and $q^{(2)}$, and η; (b) the relation between B and M, whose eigenvalues differ by an integer; (c) appropriate forms of σ_1 and σ_2 ($\sigma_3 = 1$, wherever it is needed), and their relation to the step operators $\sigma^{(1)}$ and $\sigma^{(2)}$; (d) the functions $f(B)$ and $g(B)$, which determine the matrix elements of $\sigma^{(1)}$ and $\sigma^{(2)}$; and (e) the jth eigenvalue $b^{(j)}$ of B.

1. The Harmonic Oscillator

$$B = p^2 + q^2, \quad \eta = 1,$$
$$q^{(1)} = -\infty, \quad q^{(2)} = \infty, \tag{3.76a}$$
$$B = 2M, \tag{3.76b}$$
$$\sigma_1 = q = \tfrac{1}{2}(\sigma^{(1)} + \sigma^{(2)}),$$
$$\sigma_2 = -ip = \tfrac{1}{2}(\sigma^{(1)} - \sigma^{(2)}), \tag{3.76c}$$
$$f(B) = 2M - 1, \quad g(B) = 1, \tag{3.76d}$$
$$b^{(j)} = 2j - 1. \tag{3.76e}$$

2. The Three-Dimensional Oscillator

$$B = p^2 + q^2 + \beta q^{-2}, \quad \eta = 1,$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty, \tag{3.77a}$$
$$B = 2M, \tag{3.77b}$$
$$\sigma_1 = q^2 = \tfrac{1}{2}(\sigma^{(1)} + \sigma^{(2)}) + M,$$
$$\sigma_2 = -2iqp - 1 = \sigma^{(1)} - \sigma^{(2)}, \tag{3.77c}$$
$$f(B) = 4\beta + 1 - 4(M - 1)^2, \quad g(B) = 1, \tag{3.77d}$$
$$b^{(j)} = 2j + (4\beta + 1)^{\frac{1}{2}}. \tag{3.77e}$$

The different range $(q^{(1)}, q^{(2)})$ explains why the eigenvalues differ from those obtained for the simple oscillator, even when $\beta = 0$.

3. Generalized Associated Legendre Harmonics

$$B = (1 - q^2)p^2 + (\beta + 2\gamma q)/(1 - q^2),$$
$$\eta = (1 - q^2),$$
$$q^{(1)} = -1, \quad q^{(2)} = 1, \tag{3.78a}$$
$$B = M(M + 1), \tag{3.78b}$$
$$\sigma_1 = q = (\sigma^{(1)} + \sigma^{(2)})(2M + 1)^{-1} - \gamma[M(M + 1)]^{-1},$$
$$\sigma_2 = -i(1 - q^2)p = [\sigma^{(1)}(M + 1) - \sigma^{(2)}M] \times (2M + 1)^{-1}, \tag{3.78c}$$
$$f(B) = M^2 + \gamma^2/M^2 - (\beta + 1),$$
$$g(B) = (2M - 1)/(2M + 1), \tag{3.78d}$$
$$b^{(j)} = (m^{(1)} + j - 1)(m^{(1)} + j), \tag{3.78e}$$

where $m^{(1)2} + \gamma^2/m^{(1)2} = \beta + 1$.

4. Hypergeometric Harmonics

$$B = p(1 - q^2)p + 2i(\mu + \nu q)p,$$
$$\eta = (1 - q)^{-(\mu+\nu)}(1 + q)^{-(\nu-\mu)},$$
$$q^{(1)} = -1, \quad q^{(2)} = 1, \tag{3.79a}$$
$$B = (M - \nu)(M + \nu + 1), \tag{3.79b}$$
$$\sigma_1 = q = (\sigma^{(1)} + \sigma^{(2)})(2M + 1)^{-1} - \mu\nu[M(M + 1)]^{-1},$$
$$\sigma_2 = -i(1 - q^2)p = [\sigma^{(1)}(M - \nu) - \sigma^{(2)}(M + \nu + 1)](2M + 1)^{-1} + \mu(M - \nu)(M + \nu + 1)[M(M + 1)]^{-1}, \tag{3.79c}$$
$$f(B) = M^2 + \mu^2\nu^2/M^2 - \mu^2 - \nu^2,$$
$$g(B) = (2M - 1)/(2M + 1), \tag{3.79d}$$
$$b^{(j)} = (m^{(1)} + j + \nu)(m^{(1)} + j - \nu - 1), \tag{3.79e}$$

where $m^{(1)2} + \mu^2\nu^2/m^{(1)2} = \mu^2 + \nu^2$.

This operator is related to that listed under Example 3 above by a similarity transformation, but is listed independently because of its wide applications. The relations between the constants β, γ, μ, and ν are obviously

$$\beta = \mu^2 + \nu^2, \quad \gamma = \mu\nu.$$

5. Coulomb Harmonics

$$B = qp^2 + \beta^2 q + \gamma q^{-1}, \quad \eta = q,$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty, \tag{3.80a}$$
$$B = 2\beta M, \tag{3.80b}$$
$$\sigma_1 = q = \tfrac{1}{2}(\sigma^{(1)} + \sigma^{(2)})/\beta + M/\beta,$$
$$\sigma_2 = -iqp = \tfrac{1}{2}(\sigma^{(1)} - \sigma^{(2)}), \tag{3.80c}$$
$$f(B) = M(M - 1) - \gamma, \quad g(B) = 1, \tag{3.80d}$$
$$b^{(j)} = \beta[2j - 1 + (4\gamma + 1)^{\frac{1}{2}}]. \tag{3.80e}$$

If we set $\gamma = l(l + 1)$, then

$$b^{(j)} = 2\beta(j + l).$$

6. The Hydrogen Atom

$$B' = p^2 - 2\alpha q^{-1} + \gamma q^{-2}, \quad \eta = 1,$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty. \tag{3.81a}$$

This example will serve to illustrate the type of procedure required when the discrete eigenvalues are bounded above. Then the discrete eigenvectors of B' do not form a complete set and are unsuitable as a basis for a matrix representation. We first notice that there is no difficulty in finding the eigenvalues and eigenvectors of B' by the factorization method. If we set $\gamma = l(l+1)$, so that l is the angular momentum, we see the equations

$$B' = B_1 = \theta_1^*\theta_1 + b^{(1)},$$
$$B_{j+1} = \theta_j\theta_j^* + b^{(j)} = \theta_{j+1}^*\theta_{j+1} + b^{(j+1)}$$

to be satisfied by taking

$$\theta_j = p + i(l+j)q^{-1} - i\alpha/(l+j),$$
$$B_j = B' + (2l+j)(j-1)q^{-2},$$
$$b^{(j)} = -\alpha^2/(l+j)^2,$$

But, as the B_j do not commute with one another, this decomposition is not readily adapted to calculating matrix elements. We therefore consider the transformation

$$Q = \mu q, \quad P = \mu^{-1}p,$$

which transforms the eigenvalue equation

$$B'\psi = -\mu^2\psi$$

to

$$(QP^2 + \gamma Q^{-1} + Q)\psi = (2\alpha/\mu)\psi.$$

The operator $B = QP^2 + \gamma Q^{-1} + Q$, whose eigenvalues are related to those of B' and whose eigenvectors (with the coordinate Q) are the same as those of B', is of the form considered in Example 5 above, with $\beta = 1$. We infer that

$$B' = -(\alpha/M)^2, \tag{3.81b}$$

since the operators on the two sides of this equation, applied to the typical eigenvector ψ, yield the same eigenvalue $-\mu^2$. Because of the coordinate transformation involved, the operator M is here different from that used in Example 5. We can use the same matrix representation, however, on the understanding that B' is not diagonal, except for a single element corresponding to the eigenvalue $-(\alpha/\mu)^2$. Finally,

$$f(B') = M(M-1) - l(l+1), \quad g(B') = 1, \tag{3.81d}$$
$$b^{(j)} = -\alpha^2/(l+j)^2. \tag{3.81e}$$

7. Confluent Hypergeometric Harmonics

$$B = qp^2 + 2i(\beta q + \nu)p, \quad \eta = q^{1+2\nu}e^{2\beta q},$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty, \tag{3.82a}$$
$$B = 2\beta M, \tag{3.82b}$$
$$\sigma_1 = q = \tfrac{1}{2}(\sigma^{(1)} + \sigma^{(2)})/\beta + (M-\nu)/\beta,$$
$$\sigma_2 = -iqp = -\sigma^{(2)} - M, \tag{3.82c}$$
$$f(B) = M(M - 2\nu - 1), \quad g(B) = 1, \tag{3.82d}$$
$$b^{(j)} = 2\beta(j + 2\nu). \tag{3.82e}$$

8. Associated Legendre Harmonics

$$B = (1+q^2)^2p^2 + 4\mu(1+q^2), \quad \eta = (1+q^2)^2,$$
$$q^{(1)} = -\infty, \quad q^{(2)} = \infty, \tag{3.83a}$$
$$B = 4M(M+1), \tag{3.83b}$$
$$\sigma_1 = (1+q^2)^{-1}$$
$$= \tfrac{1}{2}(\sigma^{(1)} + \sigma^{(2)})(2M+1)^{-1} + \tfrac{1}{2} + \tfrac{1}{2}\mu[M(M+1)]^{-1}, \tag{3.83c}$$
$$\sigma_2 = iqp + \tfrac{1}{2}$$
$$= [\sigma^{(1)}M + \sigma^{(2)}(M+1)](2M+1)^{-1} + 1 - \mu[M(M+1)]^{-1},$$
$$f(B) = (M+1)^2 + \mu^2/(M+1)^2 - 2\mu - \tfrac{1}{4},$$
$$g(B) = (2M+3)/(2M+1), \tag{3.83d}$$
$$b^{(j)} = 4(j + m^{(1)} - 1)(j + m^{(1)}), \tag{3.83e}$$

where $(m^{(1)}+1)^2 + \mu^2/(m^{(1)}+1)^2 = 2\mu + \tfrac{1}{4}$.

9. Particle in a Box

$$B = p^2 + 4\mu\tan^2 q, \quad \eta = 1,$$
$$q^{(1)} = -\pi/2, \quad q^{(2)} = \pi/2. \tag{3.84a}$$

The transformation

$$Q = \tan q, \quad P = \tfrac{1}{2}(\cos^2 q\, p + p\cos^2 q)$$

yields

$$B = (1+Q^2)P^2(1+Q^2) + 1 + 4\mu Q^2,$$

and, with a similarity transformation, reduces this problem to that considered in Example 8 above. Thus,

$$B = (2M+1)^2, \tag{3.84b}$$
$$\sigma_1 = \cos^2 q,$$
$$\sigma_2 = i\sin q\cos q\, p - \sin^2 q + \tfrac{1}{2}, \tag{3.84c}$$
$$f(B) = (M+1)^2 + \mu^2/(M+1)^2 - 2\mu - \tfrac{1}{4},$$
$$g(B) = (2M+3)/(2M+1), \tag{3.84d}$$
$$b^{(j)} = 4(j + m^{(1)} - 1)(j + m^{(1)}), \tag{3.84e}$$

where $(m^{(1)}+1)^2 + \mu^2/(m^{(1)}+1)^2 = 2\mu + \tfrac{1}{4}$.

10. Kummer Harmonics

$$B = q^2p^2 + \beta^2q^2 - 2\alpha q, \quad \eta = q^2,$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty. \tag{3.85a}$$

This case resembles Example 6 in that it is not suited as it stands for defining a complete matrix representation,

B being a decreasing function of M:

$$B = M(1 - M), \tag{3.85b}$$

$$\sigma_1 = q^{-1} = (\sigma^{(1)} + \sigma^{(2)})/(1 - 2M) - \alpha[M(1 - M)]^{-1},$$

$$\sigma_2 = ip = [\sigma^{(1)}(1 - M) - \sigma^{(2)}M]/(1 - 2M) - \alpha[M(1 - M)]^{-1}, \tag{3.85c}$$

$$f(B) = \alpha^2/(1 - M)^2 - \beta^2,$$

$$g(B) = (3 - 2M)/(1 - 2M), \tag{3.85d}$$

$$b^{(j)} = (m^{(1)} + 1 - j)(j - m^{(1)}), \tag{3.85e}$$

where $(1 - m^{(1)})^2 = \beta^2/\alpha^2$.

11. Coupled Harmonic Oscillators

$$B = p_1^2 + p_2^2 + (1 + \alpha)q_1^2 + 2\beta q_1 q_2 + (1 - \alpha)q_2^2,$$
$$\eta = 1, \quad q_1^{(1)} = q_2^{(1)} = -\infty, \quad q_1^{(2)} = q_2^{(2)} = \infty. \tag{3.86a}$$

This example, involving two degrees of freedom, is easily generalized to any number of degrees of freedom; thus it is especially useful in dealing with quantal systems containing several particles.[9]

$$B = 2(1 + \gamma)^{\frac{1}{2}}M_1 + 2(1 - \gamma)^{\frac{1}{2}}M_2, \tag{3.86b}$$

where $\gamma = (\alpha^2 + \beta^2)^{\frac{1}{2}}$,

$$\sigma_1 = q_1 = \beta(\sigma^{(1)} - \sigma^{(2)}) + \beta(\sigma^{(3)} - \sigma^{(4)}),$$

$$\sigma_2 = -ip_1 = \beta(1 + \gamma)^{\frac{1}{2}}(\sigma^{(1)} + \sigma^{(2)}) + \beta(1 - \gamma)^{\frac{1}{2}}(\sigma^{(3)} + \sigma^{(4)}),$$

$$\sigma_3 = q_2 = (\alpha - \gamma)(\sigma^{(1)} - \sigma^{(2)}) + (\alpha + \gamma)(\sigma^{(3)} - \sigma^{(4)}),$$

$$\sigma_4 = -ip_2 = (\alpha - \gamma)(1 + \gamma)^{\frac{1}{2}}(\sigma^{(1)} + \sigma^{(2)}) + (\alpha + \gamma)(1 - \gamma)^{\frac{1}{2}}(\sigma^{(3)} + \sigma^{(4)}), \tag{3.86c}$$

$$f(B_1) = (M_1 - \tfrac{1}{2})/[4\beta^2\gamma(1 + \gamma)^{\frac{1}{2}}], \quad g(B_1) = 1, \tag{3.86d}$$

$$f(B_2) = (M_2 - \tfrac{1}{2})/[4\beta^2\gamma(1 - \gamma)^{\frac{1}{2}}], \quad g(B_2) = 1,$$
$$b^{(j,k)} = (2j - 1)(1 + \gamma)^{\frac{1}{2}} + (2k - 1)(1 - \gamma)^{\frac{1}{2}}. \tag{3.86e}$$

12. Cylindrical Harmonics

$$B = -qpqp + q^2, \quad \eta = q,$$
$$q^{(1)} = 0, \quad q^{(2)} = \infty, \tag{3.87a}$$

$$B = M^2, \tag{3.87b}$$

$$\sigma_1 = q^{-1} = \tfrac{1}{2}\sigma^{(1)}/[L(L + 1)] + \tfrac{1}{2}\sigma^{(2)}/[(L - 1)L],$$

$$\sigma_2 = -ip = \tfrac{1}{2}\sigma^{(1)}/(L + 1) - \tfrac{1}{2}\sigma^{(2)}/(L - 1), \tag{3.87c}$$

$$f(B) = L(L - 1), \quad g(B) = 1, \tag{3.87d}$$

$$b^{(j)} = j^2. \tag{3.87e}$$

[9] H. S. Green, Nucl. Phys. **54**, 505; **57**, 483 (1964).

This example is unusual because both

$$\sigma^{(1)} = (q^{-1}L - ip)(L + 1)$$

and

$$\sigma^{(2)} = (q^{-1}L + ip)(L - 1)$$

contain a factor, depending on B only, without which the equation $f(b^{(1)}) = 0$ would have no solution. The possibility of introducing such a factor exists only because $\bar{\sigma}^{(1)} = \sigma^{(2)}$, and there is a vector $\psi^{(1)}$ for which $\sigma^{(2)}\psi^{(1)} = 0$, while $\sigma^{(1)}\psi^{(1)} \neq 0$.

4. THE OPERATORS Γ AND Δ

A. Matrix Representation

Our discussion of the operator B has yielded a number of independent operators, the $\sigma^{(l)}$, and B itself, for which matrix elements can be written down at once. In most of the examples, there were just two $\sigma^{(l)}$, or three—including a $\sigma^{(3)}$ which was simply a constant multiple of the unit operator. Usually we were able to express B in terms of a "diagonal" operator M, and $\sigma^{(1)}$ and $\sigma^{(2)}$ in terms of M and the step operators ϵ and $\bar{\epsilon}$:

$$B = B(M), \tag{4.1}$$

$$\sigma^{(1)} = \bar{\epsilon}h^{(1)}(M), \tag{4.2}$$

$$\sigma^{(2)} = h^{(2)}(M)\epsilon. \tag{4.3}$$

The operator $\bar{\epsilon}$ is represented by the matrix $\bar{\epsilon}_{jk} = \delta_{j\,k+1}$, ϵ by $\epsilon_{jk} = \delta_{j+1\,k}$, and M by $M_{jk} = m^{(j)}\delta_{jk}$, where $m^{(j)} = m^{(1)} + j - 1$.

We wish to consider next the determination of the matrix elements of Γ and Δ as a preliminary step towards the diagonalization of $A = \Gamma B + \Delta$. Since the same considerations will apply to both perturbation matrices (as well as to many other matrices of similar type), let us agree to feature Δ in the discussion.

We denote by $\Delta^{(0)}$ the diagonal matrix whose diagonal elements are identical with those of Δ. Similarly we denote by $\Delta^{(l)}\epsilon^l$, where l is any positive integer, the matrix with elements identical with those of Δ in the lth diagonal above the center, and zero elsewhere. Finally, $\bar{\epsilon}^l\Delta^{(-l)}$ denotes the matrix with elements identical with those of Δ in the lth diagonal below the center, and zero elsewhere. Then Δ is expressed in the form

$$\Delta = \Delta^{(0)} + \sum_{l=1}^{\infty}(\bar{\epsilon}^l\Delta^{(-l)} + \Delta^{(l)}\epsilon^l). \tag{4.4}$$

Now we shall assume that

$$\Delta^{(l)} = \Delta^{(l)}(M). \tag{4.5}$$

There is a sense in which every matrix Δ can be represented in this way; all that is required is to define the function $\Delta^{(l)}$ so that $\Delta^{(l)}(m^{(1)} + j - 1)$ is the jth

element in the lth diagonal from the center. However, as our notation suggests, we shall require that $\Delta^{(l)}(m)$ be an analytic function of its argument, with a well-defined asymptotic behavior as $m \to \infty$. When this requirement is met, we will say that Δ has codiagonal form. A codiagonal perturbation can also be represented entirely in terms of M, $\sigma^{(1)}$, and $\sigma^{(2)}$:

$$\Delta = E^{(0)} + \sum_{l=1}^{\infty} (\sigma^{(1)l} E^{(-l)} + E^{(l)} \sigma^{(2)l}), \tag{4.6}$$

with

$$\Delta^{(-l)} = E^{(-l)}(M) h^{(1)}(M) \times h^{(1)}(M+1) \cdots h^{(1)}(M+l-1), \tag{4.7}$$

$$\Delta^{(l)} = E^{(l)}(M) h^{(2)}(M) \times h^{(2)}(M+1) \cdots h^{(2)}(M+l-1). \tag{4.8}$$

$$\Delta^{(0)} = E^{(0)}(M). \tag{4.9}$$

For convenience we note here that

$$h^{(1)}(M-1) = [f(B)g(B)]^{\frac{1}{2}}, \tag{4.10}$$

$$h^{(2)}(M-1) = [f(B)/g(B)]^{\frac{1}{2}}, \tag{4.11}$$

where the functions $f(B)$ and $g(B)$ are those listed in the previous section.

We may therefore define our first task as that of finding the explicit form of the right side of (4.6). When this has been done, the $\Delta^{(l)}$ are given by (4.7)–(4.9) and we can write down the matrix elements of Δ in the form

$$\Delta_{jk} = \Delta_j^{(0)} \delta_{jk} + \sum_{l=1}^{\infty} (\Delta_k^{(-l)} \delta_{j\,k+l} + \Delta_j^{(l)} \delta_{j+l\,k}), \tag{4.12}$$

$$\Delta_j^{(l)} = \Delta^{(l)}(m^{(1)} + j - 1). \tag{4.13}$$

In the examples we shall consider, only a small number of the $\Delta^{(l)}(M)$ will be different from zero. Usually the nonvanishing diagonals are $\Delta^{(-1)}$, $\Delta^{(0)}$, and $\Delta^{(1)}$ or $\Delta^{(-2)}$, $\Delta^{(0)}$, and $\Delta^{(2)}$; the matrix Δ_{jk} will then have tridiagonal form. But the methods we will develop are by no means limited to such simple instances.

We illustrate the procedure with the initial example given in Eqs. (1.2) and (2.2):

$$A = p(1-q^2)p - \alpha^2 q^2.$$

To obtain the matrix elements of this operator, which commonly arises through the use of spheroidal coordinates in quantum mechanics, we begin by noting that $B = p(1-q^2)p$ contains the singularities of the operator in the range $q^{(1)} = -1$ to $q^{(2)} = 1$, which is usually of physical interest, and that this is one of the exactly diagonalizable operators considered in the preceding section. From the discussion of this operator in the text, or by specialization from Example 3 listed there, we find

$$q = (\sigma^{(1)} + \sigma^{(2)})/(2M+1), \tag{4.14}$$

$$q^2 = \sigma^{(1)2}[(2M+1)(2M+3)]^{-1} + [(2M+3) \times (2M+5)]^{-1}\sigma^{(2)2} + \sigma^{(1)}\sigma^{(2)}[(2M+1) \times (2M-1)]^{-1} + \sigma^{(2)}\sigma^{(1)}[(2M+1) \times (2M+3)]^{-1}, \tag{4.15}$$

and

$$h^{(1)}(M) = (M+1)[(2M+1)/(2M+3)]^{\frac{1}{2}}, \tag{4.16}$$

$$h^{(2)}(M) = (M+1)[(2M+3)/(2M+1)]^{\frac{1}{2}}, \tag{4.17}$$

so that

$$\sigma^{(2)}\sigma^{(1)} = (M+1)^2, \tag{4.18}$$

$$\sigma^{(1)}\sigma^{(2)} = \epsilon\bar{\epsilon}M^2 = M^2, \tag{4.19}$$

since the jth eigenvalue of M is

$$m^{(j)} = j - 1. \tag{4.20}$$

Thus we have

$$\Delta^{(0)} = -\alpha^2 M^2[(2M+1)(2M-1)]^{-1} + (M+1)^2[(2M+1)(2M+3)]^{-1} = -\alpha^2(2M^2+2M-1)[(2M-1)(2M+3)]^{-1}, \tag{4.21}$$

$$E^{(-2)} = -\alpha^2[(2M+1)(2M+3)]^{-1} \tag{4.22}$$

$$E^{(2)} = -\alpha^2[(2M+3)(2M+5)]^{-1}, \tag{4.23}$$

while all other $E^{(l)}$ vanish. Also,

$$\Delta^{(-2)} = -\alpha^2(M+1)(M+2) \times [(2M+1)(2M+3)^2(2M+5)]^{-\frac{1}{2}}, \tag{4.24}$$

$$\Delta^{(2)} = -\alpha^2(M+1)(M+2) \times [(2M+1)(2M+3)^2(2M+5)]^{-\frac{1}{2}}. \tag{4.25}$$

It will be seen that Δ is represented by a symmetric matrix, a consequence of the fact that it is self-adjoint (Hermitian, with $\eta = 1$):

$$\Delta_{jk} = \Delta_j^{(0)} \delta_{jk} + \Delta_k^{(-2)} \delta_{j\,k+2} + \Delta_j^{(2)} \delta_{j+2\,k}, \tag{4.26}$$

$$\Delta_j^{(0)} = -\alpha^2(2j^2-2j-1)[(2j-3)(2j+1)]^{-1}, \tag{4.27}$$

$$\Delta_j^{(-2)} = \Delta_j^{(2)} = -\alpha^2 j(j+1)[(2j-1)(2j+1)^2 \times (2j+3)]^{-\frac{1}{2}}. \tag{4.28}$$

B. Further Examples

In listing a few additional examples we will include those which are particularly simple, yet have a wide range of applications in quantum mechanics and other

branches of mathematical physics. We shall not display detailed calculations, which are similar to those of the example just considered.

(1) $$A = (1 - \kappa q^2)p(1 - q^2)p + 2i(\alpha + \beta q^2)qp + \lambda q^2. \quad (4.29)$$

Let us define ν and ρ by

$$\nu(1 - \kappa) = \alpha + \beta, \quad (4.30)$$

$$\rho(1 - \kappa) = \alpha\kappa + \beta, \quad (4.31)$$

so that the operator can be written in the form

$$A = (1 - \kappa q^2)B - 2i\rho(1 - q^2)qp + \lambda q^2, \quad (4.32)$$

$$B = p(1 - q^2)p + 2i\nu qp. \quad (4.33)$$

Thus

$$\Gamma = 1 - \kappa q^2, \quad (4.34)$$

$$\Delta = -2i\rho(1 - q^2)qp + \lambda q^2. \quad (4.35)$$

In order that Γ should be nonsingular and positive-definite on the Hilbert space of B, we must have $\kappa < 1$. Given an operator of similar form but with $\kappa > 1$, we should apply the transformation $Q = \kappa^{\frac{1}{2}}q$, $P = \kappa^{\frac{1}{2}}p$. Here, the operator Δ is not bounded; but as the asymptotic increase with M is of lower order than that of ΓB, this is not objectionable.

The operator B is of the type listed under Example 4 in the previous section (with $\mu = 0$). We find

$$q^2 = (q^2)^{(0)} + \bar{\epsilon}^2(q^2)^{(-2)} + (q^2)^{(2)}\epsilon^2, \quad (4.36)$$

$$-i(1 - q^2)qp = [-i(1 - q^2)qp]^{(0)} + \bar{\epsilon}^2[-i(1 - q^2)qp]^{(-2)} + [-i(1 - q^2)qp]^{(2)}\epsilon^2, \quad (4.37)$$

where

$$(q^2)^{(0)} = \tfrac{1}{2} - (2\nu^2 + \tfrac{1}{2})[(2M - 1)(2M + 3)]^{-1}, \quad (4.38)$$

$$(q^2)^{(2)} = (q^2)^{(-2)} = R(M + 1), \quad (4.39)$$

$$[-i(1 - q^2)qp]^{(0)} = -\tfrac{1}{2}(\nu + \tfrac{1}{2}) - (2\nu + 3)(\nu^2 - \tfrac{1}{4}) \times [(2M - 1)(2M + 3)]^{-1}, \quad (4.40)$$

$$[-i(1 - q^2)qp]^{(-2)} = (M - \nu)R(M + 1), \quad (4.41)$$

$$[-i(1 - q^2)qp]^{(2)} = (M + \nu + 3)R(M + 1), \quad (4.42)$$

and

$$R(M) = \left\{\frac{(M^2 - \nu^2)[(M + 1)^2 - \nu^2]}{(2M - 1)(2M + 1)^2(2M + 3)}\right\}^{\frac{1}{2}}. \quad (4.43)$$

The jth eigenvalue of M is $m^{(j)} = \nu + j - 1$, and the matrix elements of Γ and Δ are therefore

$$\Gamma_{jk} = \Gamma_j^{(0)}\delta_{jk} + \Gamma_k^{(-2)}\delta_{j\,k+2} + \Gamma_j^{(2)}\delta_{j+2\,k}, \quad (4.44)$$

$$\Delta_{jk} = \Delta_j^{(0)}\delta_{jk} + \Delta_k^{(-2)}\delta_{j\,k+2} + \Delta_j^{(2)}\delta_{j+2\,k}, \quad (4.45)$$

where

$$\Gamma_j^{(0)} = 1 - \tfrac{1}{2} + \tfrac{1}{2}\kappa(\nu^2 - \tfrac{1}{4})[(\nu + j - \tfrac{3}{2})(\nu + j + \tfrac{1}{2})]^{-1}, \quad (4.46)$$

$$\Gamma_j^{(-2)} = \Gamma_j^{(2)} = R(\nu + j), \quad (4.47)$$

$$\Delta_j^{(0)} = [\tfrac{1}{2}\lambda - \tfrac{1}{2}\lambda(\nu^2 - \tfrac{1}{4}) - \rho(\nu + \tfrac{1}{2}) - \rho(\nu + \tfrac{3}{2})(\nu^2 - \tfrac{1}{4})][(\nu + j - \tfrac{3}{2})(\nu + j + \tfrac{1}{2})]^{-1}, \quad (4.48)$$

$$\Delta_j^{(-2)} = [2\rho(j - 1) + \lambda]R(\nu + j), \quad (4.49)$$

$$\Delta_j^{(2)} = [2\rho(2\nu + j + 3) + \lambda]R(\nu + j). \quad (4.50)$$

For $\kappa = 0$ and $\rho = 0$, the operator A is connected with the associated spheroidal and Mathieu harmonics. The applications of these harmonics are described in the tables of spheroidal and associated spheroidal (including Mathieu) wavefunctions published by Flammer[10] and by Stratton, Morse, Chu, Little, and Corbato.[11] The method of computation used for these tables is equivalent to the diagonalization of the symmetric matrix to which A_{jk} reduces when $\kappa = \rho = 0$, and the calculations could have been much simplified by the use of standard techniques available for this process.

When $\kappa \neq 0$ and $\rho = \frac{1}{2}\kappa$, $\nu = \frac{3}{2}$, the operator reduces to the one used as an example in Eqs. (1.4) and (2.3); it is connected with ellipsoidal or Lamé harmonics, which also have many physical applications. Particular forms of the eigenvalue problem arise in connection with Wick's equation and the Bethe–Salpeter equation in relativistic quantum mechanics.

(2) $$A = (1 - \kappa q)p(1 - q^2)p + 2i(\alpha + \beta q + \gamma q^2) + \lambda q. \quad (4.51)$$

If we define μ, ν, and ρ by

$$\mu(1 - \kappa^2) = \alpha + \kappa\beta + \gamma, \quad (4.52)$$

$$\nu(1 - \kappa^2) = \kappa\alpha + \beta + \kappa\gamma, \quad (4.53)$$

$$\rho = \mu - \alpha, \quad (4.54)$$

this operator assumes the form

$$A = (1 - \kappa q)B - 2i\rho(1 - q^2)p + \lambda q, \quad (4.55)$$

where

$$B = p(1 - q^2)p + 2i(\mu + \nu q)p. \quad (4.56)$$

Here

$$\Gamma = 1 - \kappa q, \quad (4.57)$$

$$\Delta = -2i\rho(1 - q^2)p + \lambda q, \quad (4.58)$$

[10] C. Flammer, *Spheroidal Wave Functions* (Stanford University Press, Palo Alto, Calif., 1957).

[11] J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave Functions* (John Wiley & Sons, Inc., New York, 1956).

and we must have $-1 < \kappa < 1$ if Γ is to be nonsingular on the Hilbert space defined by B.

The operator B is now precisely that discussed under Example 4 in the previous section, and we have immediately

$$q = q^{(0)} + \bar{\epsilon}q^{(-1)} + q^{(1)}\epsilon, \tag{4.59}$$

$$-i(1-q^2)p = [-i(1-q^2)p]^{(0)} + \bar{\epsilon}[-i(1-q^2)p]^{(-1)} + [-i(1-q^2)p]^{(1)} \tag{4.60}$$

where

$$q^{(0)} = \mu\nu[M(M+1)]^{-1}, \tag{4.61}$$

$$q^{(1)} = q^{(-1)} = S(M+1), \tag{4.62}$$

$$[-i(1-q^2)p]^{(0)} = \mu(M-\nu)(M+\nu+1)[M(M+1)]^{-1}, \tag{4.63}$$

$$[-i(1-q^2)p]^{(-1)} = (M-\nu)S(M+1), \tag{4.64}$$

$$[-i(1-q^2)p]^{(1)} = -(M+\nu+2)S(M+1), \tag{4.65}$$

and

$$S(M) = \left[\frac{M^2 + \mu^2\nu^2M^{-2} - \mu^2 - \nu^2}{(2M-1)(2M+1)}\right]^{\frac{1}{2}}. \tag{4.66}$$

The jth eigenvalue of M is $m^{(j)} = \nu + j - 1$ (if $\nu > \mu$; otherwise $\mu + j + 1$), and so we obtain the matrix elements

$$\Gamma_{jk} = \Gamma_j^{(0)}\delta_{jk} + \Gamma_k^{(-1)}\delta_{jk+1} + \Gamma_j^{(1)}\delta_{j+1k}, \tag{4.67}$$

$$\Delta_{jk} = \Delta_j^{(0)}\delta_{jk} + \Delta_k^{(-1)}\delta_{jk+1} + \Delta_j^{(1)}\delta_{j+1k}, \tag{4.68}$$

where

$$\Gamma_j^{(0)} = 1 - \kappa\mu\nu[(\nu+j-1)(\nu+j)]^{-1}, \tag{4.69}$$

$$\Gamma_j^{(-1)} = \Gamma_j^{(1)} = -\kappa S(\nu+j), \tag{4.70}$$

$$\Delta_j^{(0)} = \mu[2\rho(j-1)(2\nu+j) + \lambda\nu] \times [(\nu+j-1)(\nu+j)]^{-1}, \tag{4.71}$$

$$\Delta_j^{(-1)} = [2\rho(j-1) + \lambda]S(\nu+j), \tag{4.72}$$

$$\Delta_j^{(1)} = [2\rho(2\nu+j) + \lambda]S(\nu+j). \tag{4.73}$$

(3) $$A = (q+1)qp^2 + 2i(\beta q^2 + \gamma q + \nu)p + \lambda q. \tag{4.74}$$

Let us set

$$\mu = \gamma - (\beta + \nu) \tag{4.75}$$

so that

$$A = (q+1)B + \lambda q + 2i\mu qp, \tag{4.76}$$

$$B = qp^2 + 2i(\beta q + \nu)p; \tag{4.77}$$

and

$$\Gamma = q + 1, \tag{4.78}$$

$$\Delta = \lambda q + 2i\mu qp. \tag{4.79}$$

Again, though Δ is unbounded, its increase with M is of lower order than that of ΓB and gives rise to no difficulty. The operator B is listed under Example 7 in the previous section, which shows that

$$q = q^{(0)} + \bar{\epsilon}q^{(-1)} + q^{(1)}\epsilon, \tag{4.80}$$

$$iqp = (iqp)^{(0)} + (iqp)^{(1)}\epsilon, \tag{4.81}$$

where

$$q^{(0)} = (M-\nu)/\beta, \tag{4.82}$$

$$q^{(1)} = q^{(-1)} = \tfrac{1}{2}[(M+1)(M-2\nu)]^{\frac{1}{2}}/\beta, \tag{4.83}$$

$$(iqp)^{(0)} = M, \tag{4.84}$$

$$(iqp)^{(1)} = [(M+1)(M-2\nu)]^{\frac{1}{2}}. \tag{4.85}$$

The jth eigenvalue of M is $2\nu + j$ (assuming $\nu \geq -\frac{1}{2}$) and the matrix elements of Γ and Δ are, therefore,

$$\Gamma_{jk} = \Gamma_j^{(0)}\delta_{jk} + \Gamma_k^{(-1)}\delta_{jk+1} + \Gamma_j^{(1)}\delta_{j+1k}, \tag{4.86}$$

$$\Delta_{jk} = \Delta_j^{(0)}\delta_{jk} + \Delta_k^{(-1)}\delta_{jk+1} + \Delta_j^{(1)}\delta_{j+1k}, \tag{4.87}$$

where

$$\Gamma_j^{(0)} = 1 + (\nu+j)/\beta, \tag{4.88}$$

$$\Gamma_j^{(-1)} = \Gamma_j^{(1)} = \tfrac{1}{2}[j(2\nu+j+1)]^{\frac{1}{2}}/\beta, \tag{4.89}$$

$$\Delta_j^{(0)} = \lambda(\nu+j)/\beta + 2\mu(2\nu+j), \tag{4.90}$$

$$\Delta_j^{(1)} = \Delta_j^{(-1)} = (\tfrac{1}{2}\lambda/\beta + 2\mu)[j(2\nu+j+1)]^{\frac{1}{2}}. \tag{4.91}$$

C. Matrix Reduction

The calculation of the matrix elements of Γ and Δ allows us to write down the eigenvalue equation

$$\sum_k A_{jk}\psi_k^{(l)} = \sum_k (b^{(k)}\Gamma_{jk} + \Delta_{jk})\psi_k^{(l)} = a^{(l)}\psi_j^{(l)},$$

with explicitly numerical coefficients, and we wish next to discuss the solution of this equation. All numerical methods depend on the reduction of this infinite set to a finite set of equations, and our first concern will be to examine the effect of truncation on the eigenvalue $a^{(l)}$ and its associated eigenvector $\psi_j^{(l)}$.

As Γ is a nonsingular operator, the above equation can be written as

$$b^{(j)}\psi_j^{(l)} = \sum_k (a^{(l)}\Gamma_{jk}^{-1} - E_{jk})\psi_k^{(l)}, \tag{4.92}$$

$$E_{jk} = \sum_i \Gamma_{ji}^{-1}\Delta_{ik}. \tag{4.93}$$

If $b^{(0)}$ is a number less than the lower bound $b^{(1)}$ of B, we can also write

$$j\psi_j^{(l)} = j(b^{(j)} - b^{(0)})^{-1} \times \sum_k (a^{(l)}\Gamma_{jk}^{-1} - E_{jk} - b^{(0)}\delta_{jk})\psi_k^{(l)}. \tag{4.94}$$

The general requirements on B, Γ, and Δ which were stated in Sec. 2 can now be formulated more explicitly by demanding that the operators $M(B - b^{(0)})^{-1}\Gamma^{-1}$

and $M(B - b^{(0)})^{-1}\Gamma^{-1}\Delta$ should be bounded. When these conditions are met, we can show that

$$j\psi_j^{(l)} < |a^{(l)}|\,\delta + \epsilon, \tag{4.95}$$

where δ and ϵ are positive numbers. Then, if we define the vector $\delta\psi^{(l)}$ by

$$\delta\psi_j^{(l)} = \psi_j^{(l)}, \quad n < j \leq n + m, \tag{4.96}$$

$$\delta\psi_j^{(l)} = 0, \quad j \leq n \quad \text{or} \quad j > n + m, \tag{4.97}$$

it follows that

$$\|\delta\psi^{(l)}\| = (\delta\psi^{(l)}, \delta\psi^{(l)})^{\frac{1}{2}} < O(n^{-\frac{1}{2}}), \tag{4.98}$$

independently of the value of m. Thus, truncation of A_{jk} at the nth row and column will result in errors in $a^{(l)}$ and $\psi^{(l)}$ less than $O(n^{-\frac{1}{2}})$. It should be stressed that this theoretical estimate of the error can be reduced in order of magnitude and made precise in magnitude in most practical applications. Another point worth noticing is that the value of n required to ensure satisfactory accuracy will normally depend on l.

Our problem is now reduced to the diagonalization of the n-dimensional matrix A_{jk}. If A_{jk} is symmetric, we might well regard this as a trivial matter, since there is a variety of fast and accurate computer programs in common use which is well suited to this purpose. It is true that, since $\Gamma^{-\frac{1}{2}}A\Gamma^{\frac{1}{2}}$ is self-adjoint, our matrix can always be reduced to symmetric form; however, this is not always the most convenient form of the matrix, even when $\Gamma = 1$—as we have already seen, for instance, when the discrete eigenvectors do not form a complete set. The problem of diagonalizing an asymmetric matrix, even when the eigenvalues are real, is regarded as one of the most difficult in numerical analysis. We therefore pursue the matter in this section, showing that an extension of the factorization method is well adapted to the problem. The speed and accuracy of the method is so good that it is even competitive with existing methods of diagonalizing symmetric matrices.

The principles of the method have been described by Francis, and employ what he calls the "QR transformation," a generalization of Rutishauser's "LR transformation."[6] Both of these are manifestly applications of the factorization procedure, and depend on the following algebraic lemma. If

$$A = A_1 = X_1\theta_1 + a_1 \tag{4.99}$$

and

$$A_{j+1} = \theta_j X_j + a_j = X_{j+1}\theta_{j+1} + a_{j+1}, \tag{4.100}$$

then

$$X_1X_2 \cdots X_m\theta_m \cdots \theta_2\theta_1 = \prod_{j=1}^{m}(A - a_j). \tag{4.101}$$

If X_j is the adjoint of θ_j, this operator is positive-definite, and it follows that the only eigenvalues of A_j between the lower and upper bounds of the a_j are the a_j themselves. If the X_j are not adjoints of the θ_j, the a_j are not in general eigenvalues. However, it can be shown that, if X_j is unitary and θ_j is upper-triangular at each stage of the factorization, and the a_j are suitably chosen, $X_j\theta_j$ will approach upper-triangular form.

For suppose

$$A = UTU^*, \tag{4.102}$$

where T is triangular, and

$$U_m = X_1X_2 \cdots X_m, \tag{4.103}$$

then

$$\theta_m \cdots \theta_2\theta_1 = U_m^*U\prod_{j=1}^{m}(T - a_j)U^* \tag{4.104}$$

is upper-triangular; and if any one of the a_j's is near an eigenvalue, the last diagonal element will be zero. A series of values of a_j is chosen with the object of reducing the last diagonal element to zero. When this has been done, the last row and column of the matrix is left alone and the next to last diagonal element is reduced to zero, and so on. Ultimately $X_j\theta_j$, which is a unitary transform of A, is triangular and the eigenvalues can be read off from the diagonal. An easily determined additional transformation reduces the triangular matrix to diagonal form. The final diagonal matrix has been obtained from A by a similarity transformation, whose elements determine the eigenvectors of A. This procedure is effective even when A has complex eigenvalues; but, unless complex arithmetic is used, the complex eigenvalues appear as irreducible two-dimensional matrices along the diagonal of the matrix in its final form.

D. Computations

Our first numerical example is derived from the operator

$$A = p(1 - q^2)p - \alpha^2q^2, \tag{4.105}$$

which has been analyzed in Eqs. (4.4)–(4.28). It follows from the results listed that A can be represented by the matrix

$$\begin{aligned} A_{ij} &= B_{ij} + \Delta_{ij}, \\ B_{ij} &= j(j-1)\delta_{ij}, \\ \Delta_{jj} &= -\tfrac{1}{2}(j^2 - j - \tfrac{1}{2})\alpha^2/[(j+\tfrac{1}{2})(j-\tfrac{3}{2})], \\ \Delta_{j\,j+2} &= -\tfrac{1}{4}j(j+1)\alpha^2/[(j+\tfrac{1}{2})(j-\tfrac{1}{2})^{\frac{1}{2}}(j+\tfrac{3}{2})^{\frac{1}{2}}] \\ &= \Delta_{j+2\,j}, \end{aligned} \tag{4.106}$$

where the elements of Δ_{ij} not otherwise specified vanish. As this matrix is symmetric, it can be diagonalized by Jacobi's method; but the eigenvalues

obtained in this way should agree with those of the operator

$$A' = (2L + 1)^{\frac{1}{2}} A (2L + 1)^{-\frac{1}{2}}, \qquad (4.107)$$

where $B = L(L + 1)$. This operator is represented by the matrix

$$\begin{aligned} A'_{ij} &= B_{ij} + \Delta'_{ij}, \\ \Delta_{j\,j+2} &= -\tfrac{1}{4} j(j+1)\alpha^2/[(j+\tfrac{1}{2})(j+\tfrac{3}{2})], \\ \Delta_{j+2\,j} &= -\tfrac{1}{4} j(j+1)\alpha^2/[(j-\tfrac{1}{2})(j+\tfrac{1}{2})]. \end{aligned} \qquad (4.108)$$

A FORTRAN program based on the factorization method and described in the appendix was used to find the eigenvalues and eigenvectors of this matrix, truncated to 50 rows and columns. The eigenvalues obtained are listed in Table I corresponding to $\alpha^2 = 25$.

The determination of these eigenvalues, together with eigenvectors, took eight seconds on a CDC 6400 computer, about one half the time required by a similar program applying Jacobi's method to the corresponding symmetric matrix. Comparison of the results of the two programs and the evaluation of the asymptotic formula for the eigenvalues showed that all except the last two eigenvalues are quite accurate and that, up to the 44th eigenvalue, accuracy extends to the last significant figure. Somewhat less than two factorizations were required, on the average, to obtain each eigenvalue. Similar results have been obtained with other asymmetric matrices, including some with complex eigenvalues.

TABLE I. Eigenvalues for matrix given in Eq. (4.108), corresponding to $\alpha^2 = 25$.

k	$a^{(k)}$	k	$a^{(k)}$
1	−16.0790427	26	637.525339
2	−16.0504127	27	689.523455
3	−2.44859890	28	743.521774
4	0.0609298922	29	799.520268
5	8.63039594	30	857.518913
6	18.0845680	31	917.517689
7	29.9168823	32	979.516581
8	43.8068811	33	1043.51557
9	59.7361805	34	1109.51466
10	77.6875676	35	1177.51382
11	97.6526589	36	1247.51305
12	119.626718	37	1319.51234
13	143.606902	38	1393.51169
14	169.591415	39	1469.51109
15	197.579078	40	1547.51053
16	227.569088	41	1627.51002
17	259.560884	42	1709.50954
18	293.554062	43	1793.50910
19	329.548329	44	1879.50868
20	367.543464	45	1967.50830
21	407.539299	46	2057.50794
22	449.535707	47	2149.50771
23	493.532587	48	2243.50738
24	539.529858	49	2339.70420
25	587.527460	50	2437.70002

Our second example is used to illustrate a technique useful, and often essential, in handling slowly convergent matrices. It arises in the determination of the energies of bound states and the Regge trajectories associated with Wick's equation. The relevant operator is

$$\begin{aligned} A = [\kappa^2 - (1 + q^2)^{-1}] \\ \times [\tfrac{1}{4} p(1 + q^2) p(1 + q^2) - 4\nu^2]. \end{aligned} \qquad (4.109)$$

For bound states, ν is an integer and κ determines the binding energy; the eigenvalues of A represent possible values of the coupling constant. In the corresponding scattering problem, κ is imaginary and may be assumed to be known; ν is also imaginary and determines the Regge poles. The operator

$$B = \tfrac{1}{4} p(1 + q^2) p(1 + q^2) = 4L^2 \qquad (4.110)$$

readily yields to the technique of Sec. 3, with

$$\sigma_1 = (1 + q^2)^{-1} - \tfrac{1}{2}; \qquad (4.111)$$

the eigenvalues of L are $l^{(j)} = j - 1$ and the matrix elements of σ_1 are

$$(\sigma_1)_{ij} = \tfrac{1}{4}(\delta_{i\,j+2} + \delta_{i+2\,j}), \qquad (4.112)$$

except that $(\sigma_1)_{31} = \frac{1}{2}$. Because of the poor convergence of the solutions of

$$A\psi^{(j)} = a^{(j)}\psi^{(j)},$$

it is necessary to modify the last diagonal element of the truncated matrix A_{ij}; this is done in such a way as to ensure that

$$a^{(j)}[\kappa^2 - (1 + q^2)^{-1}]^{-1}\psi^{(j)} = (B - 4\nu^2)\psi^{(j)}$$

is a finite vector.

In examples of this kind, the components $\psi_k^{(j)}$ of $\psi^{(j)}$ satisfy a recurrence relation of the type

$$a\psi_{k+2}^{(j)} + 2b\psi_{k+1}^{(j)} + c\psi_k^{(j)} = 0 \qquad (4.113)$$

for large values of i. This difference equation has two independent solutions, and the one normally required is

$$\psi_k = \alpha\lambda^k, \qquad (4.114)$$

where λ is the smaller root of

$$a\lambda^2 + 2b\lambda + c = 0. \qquad (4.115)$$

Consequently, the last diagonal element of the truncated matrix should be modified by a factor $1 + \frac{1}{2}a\lambda/b$. Such a procedure does not apply, however, when A is given by (4.109), since equation (4.115) proves to have two complex roots of equal modulus. In this case the alternative procedure already mentioned is used.

5. DIFFERENTIAL AND INTEGRAL EQUATIONS

It will be evident from the results obtained in the preceding sections that it is in no sense essential, and perhaps even detrimental, to express quantum-mechanical eigenproblems in the form of differential or integral equations. On the other hand, because of the widespread use of wave-mechanical methods, problems may often appear in this form; the present section is included both to clarify the relationship between the two formalisms and to aid with the identification of corresponding differential and operational equations.

A. Differential Equations

The functions $u_l(x)$ of the differential equations (1.1) and (1.3) are related to the vectors $\psi^{(l)}$ of the algebraic equations (1.2) and (1.4) in the following way:

$$u_l(x) = \bar{\psi}^{(0)} e^{ipx} \psi^{(l)}. \tag{5.1}$$

From this and the fact that

$$xe^{ipx} = e^{ipx} q - qe^{ipx}, \tag{5.2}$$

where q and p have their usual quantum-mechanical meanings in distinction to the unrestricted coordinate variable x, and assuming $q\psi^{(0)} = 0$ to hold, it follows that

$$xu_l(x) = \bar{\psi}^{(0)} e^{ipx} q \psi^{(l)} \tag{5.3}$$

and

$$-i \frac{d}{dx} u_l(x) = \bar{\psi}^{(0)} e^{ipx} p \psi^{(l)}. \tag{5.4}$$

Hence, we are justified in rewriting equations of the form

$$\left[A\left(-i\frac{d}{dx}, x\right) - a_l\right] u_l(x) = 0 \tag{5.5}$$

in the form

$$[A(p, q) - a^{(l)}]\psi^{(l)} = 0. \tag{5.6}$$

Conversely, the operators $B(p, q)$, $\Gamma(p, q)$, and $\Delta(p, q)$ may be expressed as differential operators wherever they appear by means of these same substitutions. For convenience, the differential equations corresponding to the examples solved in Sec. 3 are listed below:

(1) $$\left[\frac{d^2}{dx^2} - x^2 + b_j\right] u_j(x) = 0,$$

(2) $$\left[\frac{d^2}{dx^2} - x^2 - \frac{\beta}{x^2} + b_j\right] u_j(x) = 0,$$

(3) $$\left[(1 - x^2)\frac{d^2}{dx^2} - \frac{(\beta + 2\gamma x)}{(1 - x^2)} + b_j\right] u_j(x) = 0,$$

(4) $$\left[\frac{d}{dx}(1 - x^2)\frac{d}{dx} - 2(\mu + \nu x)\frac{d}{dx} + b_j\right] u_j(x) = 0,$$

(5) $$\left[x\frac{d^2}{dx^2} - \beta^2 x - \frac{\gamma}{x} + b_j\right] u_j(x) = 0,$$

(6) $$\left[\frac{d^2}{dx^2} + \frac{2\alpha}{x} - \frac{\gamma}{x^2} + b_j\right] u_j(x) = 0,$$

(7) $$\left[x\frac{d^2}{dx^2} - 2(\beta x + \nu)\frac{d}{dx} + b_j\right] u_j(x) = 0,$$

(8) $$\left[(1 + x^2)^2 \frac{d^2}{dx^2} - 4\mu(1 + x^2) + b_j\right] u_j(x) = 0,$$

(9) $$\left[\frac{d^2}{dx^2} - 4\mu \tan^2 x + b_j\right] u_j(x) = 0,$$

(10) $$\left[x^2\frac{d^2}{dx^2} - \beta^2 x^2 + 2\alpha x + b_j\right] u_j(x) = 0,$$

(11) $$\left[\frac{d^2}{dx_1^2} + \frac{d^2}{dx_2^2} - (1 + \alpha)x_1^2 - (1 - \alpha)x_2^2 - 2\beta x_1 x_2 + b_{j,k}\right] u_{j,k}(x_1, x_2) = 0,$$

(12) $$\left[x^2\frac{d^2}{dx^2} + x\frac{d}{dx} + x^2 - b_j\right] u_j(x) = 0.$$

As indicated by the names assigned earlier, these are all forms of well-known equations.

Similarly, in addition to (1.1)

$$\left[\frac{d}{dx}(1 - x^2)\frac{d}{dx} + \alpha^2 x^2 + a_l\right] u_l(x) = 0,$$

the examples discussed in Sec. 4 may be considered to solve:

(1) $$\left[(1 - \kappa x^2)\frac{d}{dx}(1 - x^2)\frac{d}{dx} - 2\{\nu(1 - \kappa x^2) - \rho(1 - x^2)\}x\frac{d}{dx} - \lambda x^2 + a_l\right] u_l(x) = 0,$$

(2) $$\left[(1 - \kappa x)\frac{d}{dx}(1 - x^2)\frac{d}{dx} - 2\{(1 - \kappa x)(\mu + \nu x) - \rho(1 - x^2)\}\frac{d}{dx} - \lambda x + a_l\right] u_l(x) = 0,$$

(3) $$\left[x(x + 1)\frac{d^2}{dx^2} - 2\{(x + 1)(\beta x + \nu) + \mu x\}\frac{d}{dx} - \lambda x + a_l\right] u_l(x) = 0.$$

These are equations which appear in many perturbation problems of current concern.

It is also interesting to note that each of the above linear second-order differential equations, the characteristic form for the Schrödinger representation, has a first-order nonlinear counterpart, a Riccati form. Defining the functions $w_j = w_j(q)$ such that

$$w_j p - p w_j = i \frac{d}{dq} w_j \tag{5.7}$$

(since $qp - pq = i$), and taking the θ_j of equations (3.1) and (3.2) to be

$$\theta_j = p + i w_j, \tag{5.8}$$

$$\theta_j^* = p - i w_j, \tag{5.9}$$

then if

$$B_j = H_j = p^2 + v(q), \tag{5.10}$$

where H_j represents the Hamiltonian (with $m = \frac{1}{2}$) incorporating the potential function $v(q)$, it follows that

$$B_j - b^{(j)} = \theta_j^* \theta_j$$

will become

$$p^2 + v(q) - b^{(j)} = (p - i w_j)(p + i w_j) \tag{5.11}$$

or

$$\frac{d}{dq} w_j + w_j^2 - v(q) + b^{(j)} = 0. \tag{5.12}$$

This Riccati equation may be converted to the Schrödinger form featured above by applying the standard transformation[12]

$$w_j = u_j^{-1} \frac{d}{dx} u_j, \tag{5.13}$$

relating $w_j(q)$ and $u_j(x)$; the result is

$$\left[\frac{d^2}{dx^2} - v(x) + b_j\right] u_j = 0. \tag{5.14}$$

Of course the w_j, u_j, and $b^{(j)}$ may be regarded as matrix elements and these equations as matrix differential equations. Clearly, either may be solved to obtain the eigenfunctions and eigenvalues relating to a particular potential $v(q)$ or $v(x)$. In the past, the latter has most often been used, and many specialized solution techniques have been developed[13]; but perhaps the most comprehensive one is the factorization method of Infeld and Hull, referred to earlier.[4] This depends on the fact that many single equations of the type (5.14) can be "factorized" into two first-order linear differential equations:

$$\left[k(x, m+1) - \frac{d}{dx}\right] u_j^m = [b_j - L(m+1)]^{\frac{1}{2}} u_j^{m+1}, \tag{5.15}$$

$$\left[k(x, m) + \frac{d}{dx}\right] u_j^m = [b_j - L(m)]^{\frac{1}{2}} u_j^{m-1}, \tag{5.16}$$

where m represents some nonnegative integer. Cross-applying the initial operators and comparing each resulting equation with (5.14) then leads to the Riccati equations

$$\frac{d}{dx} k(x, m+1) + k^2(x, m+1) + L(m+1) = v(x, m) \tag{5.17}$$

and

$$-\frac{d}{dx} k(x, m) + k^2(x, m) + L(m) = v(x, m). \tag{5.18}$$

Subtracting these yields the necessary and sufficient condition which k and L must satisfy:

$$\frac{d}{dx}[k(x, m+1) + k(x, m)] + k^2(x, m+1) - k^2(x, m) + L(m+1) - L(m) = 0. \tag{5.19}$$

Any $k - L$ pair which satisfies this equation implies a particular $v(x, m)$ through the above relations and, thus, one factorization type whose eigenfunctions and eigenvalues follow from (5.15) and (5.16). Basically, there are only two such types, but they give rise to six others which, though nonindependent, are all of practical importance. Most common potentials, including those leading to Legendrian and hypergeometric forms, prove to be special cases associated with one or another of these six types; they have been tabulated, along with the related $k - L$ pairs, eigenfunctions, and eigenvalues, by the same authors.

The step operators bracketed in (5.15) and (5.16) correspond to the θ_j of Eqs. (3.1) and (3.2) or, more generally, to the $\sigma^{(l)}$ of Eqs. (3.36) and (3.37); hence, it is to be expected that the same kind of differential equation would have to be solved to obtain specific forms of either. Actually, the general solution of the Riccati equation

$$\frac{dT}{dt} + T N_1 T + T N_2 + N_3 T - N_4 = 0, \tag{5.20}$$

where T is an $n_1 \times n_2$ matrix and N_1, N_2, N_3, N_4 are respectively $n_2 \times n_1$, $n_2 \times n_2$, $n_1 \times n_1$, $n_1 \times n_2$ matrices, can be formulated in terms of a characteristic matrix of an associated set of linear equations and a

[12] E. L. Ince, *Ordinary Differential Equations* (Longmans, Green, and Company, London, 1926; reprinted Dover Publications, New York, 1944), pp. 23–25.

[13] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Co., New York, 1953), Vol. II, pp. 1639–1758.

matrix of initial values.[14,15] In a sense, both of the methods discussed here amount to specialized schemes for obtaining the solutions of this set of associated linear equations. However, the factorization method does not yield the step operators in a convenient form for field theory computations, where they enter as creation and destruction operators in the algebraic interpretation of the required functional differential calculus,[16,17] and does not lend itself to the calculation of matrix elements for dynamical variables.

The latter limitation means that the method cannot very easily be extended from the simpler equations, such as Examples 1–12 above, to perturbed forms of these equations; and it is just this difficulty which the codiagonal approach remedies. It is possible, of course, to obtain such matrix elements from recurrence relations whenever these are available. As mentioned in connection with Eqs. (3.38) and (3.39), the operators $\sigma^{(l)}$ necessarily define relations of this kind; with

$$\sigma^{(1)}P_n = (n+1)P_{n+1}, \tag{5.21}$$

$$\sigma^{(2)}P_n = nP_{n-1}, \tag{5.22}$$

the operational form implied by Example 3 of Sec. 3,

$$x = \frac{\sigma^{(1)} + \sigma^{(2)}}{2M+1}, \tag{5.23}$$

immediately yields the well-known formula connecting Legendre polynomials $P_n(x)$ of serial order:

$$xP_n = \frac{(n+1)P_{n+1} + nP_{n-1}}{2n+1}. \tag{5.24}$$

Hence, matrix elements for the perturbation terms of Legendre-like equations can be determined from the latter. Using Eq. (1.1) as an example,

$$\alpha^2x^2P_n = \alpha^2\left\{\left[\frac{(n+1)(n+2)}{(2n+1)(2n+3)}\right]P_{n+2} + \left[\frac{(n+1)^2}{(2n+1)(2n+3)} + \frac{n^2}{(2n+1)(2n-1)}\right]P_n + \left[\frac{n(n-1)}{(2n+1)(2n-1)}\right]P_{n-2}\right\}, \tag{5.25}$$

which represents an alternate, if somewhat inconvenient, form of the relation defined by (4.26)–(4.28).

It is useful to observe that Inui has devised a technique of writing recurrence relations for the solutions of any equation of the form (5.14) which can be factorized.[18] This is the class of Fuchsian equations whose solutions are certain "special functions." It is also the class of equations for which the operator B, as defined herein, may be diagonalized exactly; and since the perturbation operators Γ and Δ need not remain small, the present results may also be interpreted as a method of solving an extended class of special function Fuchsian equations. These are equations of the form

$$\left[\frac{d^2}{dx^2} + P(x)\frac{d}{dx} + Q(x) + a_l\right]u_l(x) = 0, \tag{5.26}$$

where $P(x)$ and $Q(x)$ lead to no more than three regular singularities, or one regular and one irregular singularity, and are otherwise analytic everywhere in the domain of interest. Our restrictions on Γ and Δ are such as to leave the essential character of these functions unchanged, while yet permitting solutions for a much broader range of types than has previously been considered.

B. Integral Equations

The solution of (5.6) by matrix techniques also provides the solution of a variety of equations of the type

$$K(A)\psi^{(l)} = k^{(l)}\psi^{(l)}, \tag{5.27}$$

where $K(A)$ is some function of A, and the eigenvalue $k^{(l)}$ of K, corresponding to the eigenvector $\psi^{(l)}$, is clearly

$$k^{(l)} = K(a^{(l)}). \tag{5.28}$$

When, as in (5.5), A is represented as a differential operator and the eigenvector $\psi^{(l)}$ as a differentiable function $u^{(l)}$, K ordinarily must be represented as an integral operator—the only exceptions arising when $K(A)$ is a polynomial. Thus, in the representation adopted in this section, (5.27) assumes the form

$$\int_\alpha^\beta K(x,y)u^{(l)}(y)\,dy = k^{(l)}u^{(l)}(x), \tag{5.29}$$

where $\alpha = q^{(1)}$ and $\beta = q^{(2)}$ are the endpoints of the range of eigenvalues of q. This is a homogeneous integral equation of a type which frequently arises in mathematical physics.

If the kernel $K(x,y)$ is nonsingular, the Hilbert–Schmidt method provides a well-known matrix technique for the solution of (5.29). Here we are more interested in the possibility of identifying the operator A corresponding to a given kernel K and, hence, effectively reducing the integral equation to a differential equation or an operational equation of a known

[14] J. J. Levin, Proc. Am. Math. Soc. **10**, 519 (1959).
[15] W. T. Reid, J. Math. Mech. **2**, 221 (1959).
[16] F. Röhrlich and M. Wilner, J. Math. Phys. **37**, 482 (1966).
[17] A. I. Akheizer and V. B. Berestetskii, *Quantum Electrodynamics* (Interscience Publishers, Inc., New York, 1965), pp. 198–205.
[18] T. Inui, Progr. Theoret. Phys. (Kyoto) **3**(2), 168 (1948) and **3**(3), 244 (1948).

type. The inverse problem is also of considerable interest, since the integral equation already incorporates boundary conditions which may be difficult to apply to the corresponding differential equation. It is reasonable to require that $K(x, y)$ should be an analytic function of both variables, since otherwise the solution of the latter problem will not be unique.

The requirement that the integral operator in (5.29) should commute with the differential operator A can be met only if $K(x, y)$ satisfies the partial differential equation

$$A\left(-i\frac{\partial}{\partial x}, x\right)K(x, y) = A^*\left(-i\frac{\partial}{\partial y}, y\right)K(x, y), \tag{5.30}$$

where $A^*(p, q)$, as previously, represents the Hermitian conjugate of $A(p, q)$. Since A has to be self-conjugate in the sense of (2.5),

$$A^*(p, q) = H(p, q)A(p, q)H^{-1}(p, q), \tag{5.31}$$

where $H(p, q)$ is Hermitian. If A is a second-order differential operator of the form

$$A(p, q) = pf(q)p + ig(q)p + h(q), \tag{5.32}$$

then H is a function $H(q)$ of q only, given by

$$f(q)H'(q) = g(q)H(q). \tag{5.33}$$

If we make use of (5.31) and define $L(x, y)$ by

$$K(x, y) = H\left(-i\frac{\partial}{\partial y}, y\right)L(x, y), \tag{5.34}$$

we find that (5.30) is satisfied, provided

$$A\left(-i\frac{\partial}{\partial x}, x\right)L(x, y) = A\left(-i\frac{\partial}{\partial y}, y\right)L(x, y). \tag{5.35}$$

We may therefore assume that $L(x, y)$ is a symmetric function of x and y.

It is not possible, of course, to give a general solution of (5.35) and we shall therefore proceed to discuss some examples arising from operators $A(p, q)$ of the general type considered elsewhere in this paper. First let us suppose that A is given by (5.32) with

$$\begin{aligned} f(q) &= 1 - q^2, \\ g(q) &= 2(aq^2 + bq + c), \\ h(q) &= \lambda q^2 + \mu q, \end{aligned} \tag{5.36}$$

so that

$$H(q) = e^{2aq}(1 - q)^{a+b+c}(1 + q)^{-a+b-c}. \tag{5.37}$$

The limits of integration α and β in (5.29) will take values $-\infty$, -1, 1, or ∞, chosen with regard to the values a, b, c, and the validity of (5.30). In this example, $L(x, y)$ can be assumed to be a function of the form

$$L(x, y) = L(x + y), \tag{5.38}$$

and (5.35) is satisfied provided

$$zL''(z) + 2(az + b + 1)L'(z) + (\lambda z + \mu)L(z) = 0. \tag{5.39}$$

The solution of this equation which is finite for $z = 0$, even when $b > -1$, is readily obtained in terms of the confluent hypergeometric function $M_{k,m}$:

$$\begin{aligned} L(z) &= e^{-az}z^{-(b+1)}M_{k,m}(\nu z), \\ \nu^2 &= 4(a^2 - \lambda), \\ \nu k &= \mu - 2a(b + 1), \\ m &= b + \tfrac{1}{2}. \end{aligned} \tag{5.40}$$

The kernel of the integral equation is thus a more elementary function than its solutions, which can be expressed in terms of ellipsoidal wavefunctions.

For our second example we suppose that

$$\begin{aligned} f(q) &= q(1 - q)(\lambda - q), \\ g(q) &= f'(q) + aq^2 + bq + \lambda c, \\ h(q) &= \mu q, \end{aligned} \tag{5.41}$$

wherein it can be assumed that $\lambda > 1$. The form of $H(q)$ is readily obtained from (5.33) and the limits of integration α and β in (5.29) now take values of $-\infty$, 0, 1, λ, or ∞. Here we can assume

$$L(x, y) = L(xy) \tag{5.42}$$

[alternatively, $L(xy - x - y)$ or $L(xy - \lambda x - \lambda y)$] and then L must satisfy the equation

$$z(1 - z)L''(z) + (c - az)L'(z) - \mu L(z) = 0, \tag{5.43}$$

whose solutions are hypergeometric functions. The solutions of the integral equation are generalized Lamé functions. Another type of differential operator which has a kernel containing a factor of the type $L(xy)$ is obtained by replacing x and y with x^2 and y^2, respectively.

As a final example, suppose that A is of the type

$$A(p, q) = qf(p^2) + ipg(p^2). \tag{5.44}$$

Then clearly (5.35) will be satisfied provided $L(x, y)$ is a suitable function of the form

$$L(x, y) = L(x^2 - 2xy + y^2). \tag{5.45}$$

It is easy to construct more complicated examples and thus build a collection of corresponding differential and integral operators.

6. DISCUSSION

A. Scattering Problems

First we shall discuss one application of the methods described to operators which do not have a complete set of discrete eigenvectors. Examples 6 and 10 of Sec. 3 are of this kind, a type which occurs quite commonly in quantum mechanics. The energy spectrum of a system of particles is continuous when two or more components of the system can separate after interaction. It may even happen that the spectrum is entirely continuous, as when the potential energy has either no negative minima, or only shallow negative minima.

One way of producing a discrete spectrum for the energy in such situations is to introduce an additional potential, equivalent to an impenetrable wall containing the system. This device is not convenient for practical purposes, however, as it substitutes a very large number of closely-spaced energy levels for even a finite part of the continuous spectrum. Besides, the object of the calculation is usually to determine the scattering matrix rather than the energy eigenvalues, which are to be regarded as given in scattering problems. A more pertinent problem is, therefore, to determine the eigenvalues of (2.1),

$$A = \Gamma B + \Delta,$$

where B represents the energy and Γ and Δ are chosen so that the eigenvalues of A are discrete.

For two particles in an eigenstate of the angular momentum, the energy is

$$B = [p^2 + l(l+1)/q^2]/(2m) + V(q). \tag{6.1}$$

Supposing that, for large numerical values of q, $V(q) \sim -a/\Gamma(q)$, we take

$$\Gamma = \Gamma(q), \quad \Delta = a - b\Gamma(q), \tag{6.2}$$

where b is the known eigenvalue of B. The eigenvalue of A is then a. What is really required from this calculation is the relation

$$l = l^{(j)}(a, b), \tag{6.3}$$

obtained by assigning the value a to the jth eigenvalue of A. The bound states, which always exist for suitable values of a, correspond to negative values of b. When b is given the positive values appropriate to scattering, $l^{(j)}$ becomes complex, and $l = l^{(j)}$ is a Regge trajectory, which must appear as a pole of the S matrix. Since the S matrix is unitary, the element corresponding to the lth partial wave is

$$S_l = \prod_{j=1}^{\infty} (l - l^{(j)*})/(l - l^{(j)}) = \det [(l - L_c^*)/(l - L_c)], \tag{6.4}$$

where L_c is the analytic continuation of the angular-momentum operator whose eigenvalues are $l^{(j)}$. The phase shifts are given by

$$2i\eta_l = \log S_l = \operatorname{trace} \log [(l - L_c^*)/(l - L_c)]. \tag{6.5}$$

As a simple example of the application of this technique, consider the scattering of two charged particles due to their Coulomb interaction. Then $V(q) = -a/q$, and the operator $2mA$ coincides with that of Example 5 listed under the exactly soluble cases of Sec. 3. From the analysis given there,

$$l^{(j)} = -j + ia/(2mb)^{\frac{1}{2}}, \tag{6.6}$$

if we interpret $(-2mb)^{\frac{1}{2}}$ as $-i(2mb)^{\frac{1}{2}}$, i.e., make an analytic continuation consistent with giving the mass a negative imaginary part in the transition from negative to positive values of b. With a numerical factor of unit modulus inserted to assure convergence, the corresponding S-matrix element is

$$S_l = \prod_{j=1}^{\infty} (1 - 1/j)^{-2i\alpha}(l + j + i\alpha)/(l + j - i\alpha) = \Gamma(l + 1 - i\alpha)/\Gamma(l + 1 + i\alpha), \tag{6.7}$$

in agreement with the known result.

B. General Remarks

Because the Infeld–Hull factorization method depends on one particular coordinate representation, various studies of its algebraic substructures have been made; and because the method utilizes differential equations, these have featured Lie algebras, which inherently involve groups associated with differential equations. Kaufman has exploited this last fact in a very direct way, by using the differential operators from recurrence relations for several special functions to form the corresponding finite operator in a Lie algebra.[19] She shows that none of the examples actually treated by Infeld and Hull imply a Lie group with more than three parameters.

The six factorization types are, of course, more general; but Miller has established that these may be derived from the representations of four Lie algebras, whose groups are all special cases of a four-parameter Lie group: O_3 for the three-dimensional rotation group, T_3 for the two-dimensional Euclidean group, T_6 for the three-dimensional Euclidean group, and H_4 representing a special four-dimensional Lie algebra.[20] Making the proper identifications, the recurrence relations of the form (5.15), (5.16) for each

[19] B. Kaufman, J. Math. Phys. **7**, 447 (1966).
[20] W. Miller, Jr., Mem. Am. Math. Soc. **50** (1964).

factorization type follow immediately from the commutation relations of its algebra. As might be expected, irreducible representations of the required algebras are obtained by solving Riccati equations in the unknown functions.

Inui, in the work referred to earlier,[18] has proved that any Fuchsian equation of the hypergeometric or confluent hypergeometric type can be factorized. Accordingly, the algebra underlying any such equation must be one of the four Lie algebras given above. All examples of the operator B treated in the present paper are of this kind. However, the method we have developed is more general, allowing equations to be solved whose algebras may be of a different type. For instance, Example 11 of Sec. 3 readily generalized to n coupled oscillators, so that the group O_{2n+1} is involved.

The system of Eqs. (3.24) will terminate whenever the operator B is exactly diagonalizable and σ is properly chosen, regardless of whether or not it corresponds to a factorizable differential equation with its associated Lie algebra. Actually, the method adopted here appears to be more general than the factorization method, where the commutators of the σ_j must in addition be expressible in terms of the σ_j, and must form a finite sequence. Our method implies either a covering algebra of a Lie algebra, or even an algebra of a more general kind, when the sequence of equations does not terminate at all. Indeed, it seems reasonable to believe that a Lie-algebra structure is no more essential than a differential-equation formulation for quantum-mechanical problems.

On the other hand, a Hilbert space appears to be desirable, and generalizing from the Hermitian operator case mentioned in connection with equations (3.1)–(3.8), we have taken the operators B, σ_k, and, ultimately, A to be self-adjoint. Our results depend strongly on this fact. Also, independently pursuing a suggestion contained in work by Coish,[21] Joseph and Coulson have recently developed a method featuring self-adjoint step operators by means of which many results similar to ours for the unperturbed operator B may be obtained.[22,23]

Perhaps in conclusion it is worth remarking that, while the new solutions which the present method makes possible are not exact in the traditional sense associated with differential equations, they may nevertheless provide a superior description of the material world. The methods of conventional analysis do not necessarily correspond to the processes of nature,[24] and there is some reason to think that they are not entirely compatible with the physical requirements of quantum mechanics.

APPENDIX

We shall summarize here a computational procedure for the determination of eigenvalues and eigenvectors, based on the factorization method described in the text. It is easily translated into a computer program, consisting of three parts.

(1) The first part of the program reduces the matrix A to almost triangular form by Householder's method,[25] which uses only orthogonal transformations. This has to be done only once, and improves the speed and accuracy of the whole program. If the matrix A is n-dimensional, a succession of $n-2$ orthogonal transformations is applied, reducing the elements A_{ij} with $i > j+1$ to zero for $j = 1, 2, \cdots, n-2$.

Define

$$\begin{aligned} A^{(1)} &= A, \\ A^{(k+1)} &= P^{(k)}A^{(k)}P^{(k)}, \quad (k = 1, 2, \cdots, n-2) \\ A' &= A^{(n-1)}, \end{aligned} \tag{A1}$$

with $P^{(k)}$ defined by

$$\begin{aligned} P_{ij}^{(k)} &= \delta_{ij} - w_i^{(k)}w_j^{(k)}, \\ w_i^{(k)} &= 0, \quad i \leq k, \\ w_{k+1}^{(k)} &= (1 + \lambda^{(k)}A_{k+1\,k}^{(k)})^{\frac{1}{2}}, \\ w_i^{(k)} &= \lambda^{(k)}A_{ik}^{(k)}/w_{k+1}^{(k)}, \\ \lambda^{(k)} &= \pm\Big[\sum_{i>k}(A_{ik}^{(k)})^2\Big]^{-\frac{1}{2}}, \end{aligned} \tag{A2}$$

where the sign of $\lambda^{(k)}$ is the same as that of $A_{k+1\,k}^{(k)}$. If eigenvectors are required,

$$P' = P^{(n-2)} \cdots P^{(2)}P^{(1)} \tag{A3}$$

should be computed as well as A'.

(2) The second and most essential part of the program uses a factorization procedure, closely related to Francis's "QR transformation," to reduce the almost triangular A' to triangular form, again by orthogonal transformations only. If some of the eigenvalues of A are complex, the resulting matrix A'' will not be strictly triangular, but will contain two-dimensional submatrices on the diagonal whose eigenvalues are complex eigenvalues of A.

Define

$$U^{(k)}, T^{(k)}, \text{ and } A^{(k)}$$

[21] H. R. Coish, Can. J. Phys. **34**, 343 (1956).
[22] A. Joseph, Rev. Mod. Phys. **39**, 829 (1967).
[23] C. A. Coulson and A. Joseph, Rev. Mod. Phys. **39**, 838 (1967).
[24] P. A. M. Dirac, Phys. Rev. **139**, 684 (1965).
[25] A. S. Householder and F. L. Bauer, Numerische Math. **1**, 29 (1959).

for

$$k = n, n+1, n+2, \cdots$$

by

$$A^{(k-1)} = U^{(k)}T^{(k)} + \mu^{(k)},$$
$$A^{(k)} = T^{(k)}U^{(k)} + \mu^{(k)}, \qquad \text{(A4)}$$

where $U^{(k)}$ is unitary, $T^{(k)}$ is upper-triangular, and $\mu^{(k)}$ is a suitably chosen multiple of the unit matrix. The unitary matrix $U^{(k)}$ is a product of $n-1$ orthogonal matrices:

$$U^{(k)} = P_1^{(k)}P_2^{(k)} \cdots P_{n-1}^{(k)} \quad (k \geq n), \qquad \text{(A5)}$$

where $P_j^{(k)}$ is determined from the elements B_{jj}, $B_{j\,j+1}$, $B_{j+1\,j}$, and $B_{j+1\,j+1}$ of

$$B = \bar{P}_{j-1}^{(k)} \cdots \bar{P}_1^{(k)}A^{(k-1)} \quad (j > 1),$$

or $B = A^{(k-1)}$ for $j = 1$. In fact, $P_j^{(k)}$ is obtained by replacing the elements δ_{jj}, $\delta_{j\,j+1}$, $\delta_{j+1\,j}$, and $\delta_{j+1\,j+1}$ of the unit matrix with

$$(B_{jj} - \mu^{(k)})/N, \quad -B_{j+1\,j}/N, \quad B_{j+1\,j}/N$$

and

$$(B_{j+1\,j+1} - \mu^{(k)})/N,$$

where

$$N^2 = (B_{jj} - \lambda^{(k)})^2 + B_{j+1\,j}^2.$$

The only part of this procedure which is unspecified now is the determination of $\mu^{(k)}$, and this has an important bearing on the convergence of the sequence $A^{(n)}, A^{(n+1)}, A^{(n+2)}, \cdots$ to triangular form. The nearer $\mu^{(k)}$ is to an eigenvalue of A (or the real part of a complex eigenvalue), the more quickly this eigenvalue will be isolated. Usually it is sufficient to estimate the eigenvalue (or its real part) from the lowest unresolved two-dimensional matrix on the diagonal. However, it is important to isolate the eigenvalues in descending numerical order, and therefore better to start with an overestimate of $\mu^{(k)}$, approaching the estimated value from above after several transformations.

When the elements below the diagonal are sufficiently small, the iterations (A4) are terminated and the results

$$A'' = A^{(k)},$$
$$P'' = \bar{U}^{(k)} \cdots \bar{U}^{(n)}P' \qquad \text{(A6)}$$

are kept for the last part of the program.

(3) The third part of the program reduces the triangular matrix to diagonal form by means of similarity transformations; this is necessary, of course, only if eigenvectors are required. A sequence of transformations of the type

$$A^{(r+1)} = T^{(r)}A^{(r)}T^{(r)-1},$$
$$T_{ij}^{(r)} = \delta_{ij} - c_{kl}\delta_{ik}\delta_{jl}, \qquad \text{(A7)}$$
$$T_{ij}^{(r)-1} = \delta_{ij} + c_{kl}\delta_{ik}\delta_{il},$$

with $i < j$ may be used, starting with $l = n$ and $k = n-1$, and progressively reducing the value of $l + k$. The only complication arises when, because of the existence of complex eigenvalues, there are some nonvanishing elements just below the diagonal. We shall consider the possibility $A_{i\,i-1}^{(r)} \neq 0$ and $A_{j+1\,j}^{(r)} \neq 0$. If A_i, A_j, Q, and C are two-dimensional matrices obtained from, first, the $(i-1)$th and ith rows and columns of $A^{(r)}$, second, the jth and $(j+1)$th rows and columns of $A^{(r)}$, third, the $(i-1)$th and ith rows and the jth and $(j+1)$th columns of $A^{(r)}$ and, fourth, the $(i-1)$th and ith rows and jth and $(j+1)$th columns of $T^{(r)-1}$, we have to determine C so that

$$Q = CA_j - A_iC.$$

This matrix equation is a set of four simultaneous equations from which c_{ij} and $c_{i,j+1}$ can be determined; the solution simplifies when, as a result of the second part of the program,

$$A_{ii}^{(r)} = A_{i-1\,i-1}^{(r)} \quad \text{and} \quad A_{jj}^{(r)} = A_{j+1\,j+1}^{(r)}.$$

Lattice Dynamics of Simple Cubic Lattices with Long-Range Interactions*

JOHN A. DAVIES AND PETER D. YEDINAK†‡
Clark University, Worcester, Massachusetts

(Received 8 August 1968)

An analytic study is made of the dispersion relations and frequency spectra of simple cubic lattices in which there exist long-range potentials of the form $1/r^p$. The rigid-ion approximation is used. An expansion of the dispersion relations about the maximum propagation vector in the first Brillouin zone is obtained for $1 \leq p \leq 3$ and the contribution of the region about this point to the vibrational frequency spectrum of the lattice is studied.

I. INTRODUCTION

When long-range interactions are present in a crystal lattice, one or more branches of the dispersion relations for its normal-mode vibrations may be nonanalytic at certain critical points in the reciprocal cell. This nonanalytic behavior is caused by nonanalytic terms in the elements of the secular determinant giving the normal-mode frequencies and is not necessarily a result of a degeneracy between branches of the dispersion relations at the critical point. At the frequency corresponding to the critical point, the behavior of the frequency spectrum is determined by the expansion properties of the dispersion relations about the critical point.[1-3] We have thus begun a study of the dispersion relations and frequency spectra of lattices in which long-range potentials of the form $1/r^p$ are present between particles.

The treatment in this paper is concerned mainly with three-dimensional simple cubic lattices when the rigid-ion approximation is used. In Sec. II, we discuss the general properties of the elements of the secular determinant for an N-dimensional simple cubic lattice when a $1/r^p$ long-range interaction is present. The elements are found to be analytic except at the maximum propagation vector in the first Brillouin zone. In Sec. III, we study the dispersion relations near this nonanalytic point for a three-dimensional simple cubic lattice when p is restricted by $1 < p \leq 3$. In Secs. IV and V, the frequency spectrum and vibrational modes of this lattice are studied. We also consider the limiting behavior of our results as p approaches one. A specific model for which $p = 1$ is introduced in Sec. VI. Finally, in Sec. VII, we briefly note that many of our results depend strongly on the cubic symmetry.

II. THE SECULAR DETERMINANT

We consider an infinite, N-dimensional, simple cubic lattice in which a Cartesian coordinate system is imbedded in such a way that the ith component of the position of any lattice point is given by al_i where l_i is an integer. The displacement of the particle at equilibrium position $a\mathbf{l}$ in the i direction is denoted by $u_i(l)$. We assume that the particles interact through a long-range and a short-range pair potential in such a way that the potential energy of the particle at the origin is given by

$$V = V_L + V_S, \tag{1}$$

where

$$V_L = G \sum_l{}' \frac{(-1)^{\Sigma_k l_k}}{[\Sigma_i (a l_i + u_i(l) - u_i(0))^2]^{\frac{1}{2}p}} \tag{2}$$

and

$$V_S = H \sum_l{}' \{\Sigma_i [a l_i + u_i(l) - u_i(0)]^2\}^{-\frac{1}{2}q}. \tag{3}$$

The primes on the summation signs indicate that $\mathbf{l} = 0$ is not included in the sums. The sum in Eq. (3) is assumed to be finite. G and H are constants and we do not assume that p in Eq. (2) is necessarily an integer.

In the harmonic approximation, the equation of motion of the particle at the origin is

$$m\ddot{u}_i(0) = -\sum_{i,j} \frac{\partial^2 V}{\partial u_i(0) \partial u_j(l)}\bigg|_0 u_j(l). \tag{4}$$

After substituting solutions of the form

$$u_i(l) = U_i \exp[i\Sigma_k (l_k \phi_k - \omega t)] \tag{5}$$

into Eq. (4), we obtain the eigenvector equation

$$\lambda U_i = \Sigma_j A_{ij} U_j, \tag{6}$$

where

$$\lambda = a^{p+2} m\omega^2 / G. \tag{7}$$

The A_{ij} are defined in the following way. First we write A_{ij} in terms of its long- and short-range contributions,

$$A_{ij} = A_{ij}^L + A_{ij}^S. \tag{8}$$

A_{ij}^L and A_{ij}^S can be expressed in the following form:

$$A_{ij}^L(\boldsymbol{\phi}) = T_{ij}^L(0) - T_{ij}^L(\boldsymbol{\phi}) \tag{9}$$

and

$$A_{ij}^S(\boldsymbol{\phi}) = [T_{ij}^S(0) - T_{ij}^S(\boldsymbol{\phi})] a^{p-q} H/G, \tag{10}$$

* Based on a thesis submitted by P. D. Y. in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Physics Department of Clark University.

† P. D. Y. was supported by a NASA Graduate Student Fellowship.

‡ Present address: Western Maryland College, Westminster, Maryland.

[1] L. Van Hove, Phys. Rev. **89**, 1189 (1953).

[2] J. C. Phillips, Phys. Rev. **104**, 1263 (1956).

[3] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press Inc., New York, 1963).

where T_{ij}^L and T_{ij}^S are defined in Eqs. (A1) and (A2) of Appendix A. The secular equation, giving the dispersion relations $\lambda(\boldsymbol{\phi})$ for the lattice, is

$$|A_{ij}(\boldsymbol{\phi}) - \delta_{ij}\lambda(\boldsymbol{\phi})| = 0. \tag{11}$$

The first Brillouin zone is defined by the condition $-\pi < \phi_i \leq \pi$.

We now wish to determine the conditions under which the expansion of A_{ij} about a critical point will be dominated by a nonanalytic term as we approach the critical point. (Here we will ignore the unlikely possibility that A_{ij} contains no second-order terms.) Referring to Eqs. (2) and (10), one sees that A_{ij}^S is analytic everywhere since it is a finite sum of analytic functions. Thus, we need study only the behavior of A_{ij}^L or equivalently T_{ij}^L.

Define

$$S_n(\boldsymbol{\phi}) = \sum_l{}' \Pi_j \cos(\phi_j - \pi) l_j/(\Sigma_k l_k^2)^n. \tag{12}$$

As long as $p > N - 2$, the series in Eqs. (A1) and (A2) are uniformly convergent and we may write

$$T_{ij}^L = -p(p+2)\frac{\partial^2 S_{p/2+2}}{\partial\phi_i\partial\phi_j} - \delta_{ij}pS_{p/2+1}. \tag{13}$$

We use the Ewald transformation[4,5] to convert $S_n(\boldsymbol{\phi})$ into the following form:

$$S_n = \frac{\pi^n}{\Gamma(n)}\left\{\Sigma_l \Phi_{-n-1+\frac{1}{2}N}\left[\pi\Sigma_j\left(l_j + \frac{\xi_j}{2\pi}\right)^2\right] + \sum_l{}' [\Phi_{n-1}(\pi\Sigma_j l_j^2)\Pi_k \cos \xi_k l_k] - \frac{1}{n}\right\}. \tag{14}$$

Here, $\xi_j = \phi_j - \pi$ and $\Phi_m(x)$ is the incomplete gamma function defined by

$$\Phi_m(x) = \int_1^\infty t^m e^{-xt}\, dt. \tag{15}$$

It is well known that $\Phi_m(x)$ is analytic for any real $x > 0$. However, its expansion from $x = 0$ into regions of positive x is not a simple power series. The form of this expansion depends upon whether or not m is a negative integer. It is given in Eqs. (B1)–(B3) of Appendix B.

Using Eqs. (13) and (14) and the fact that $\Phi_m'(x) = -\Phi_{m+1}(x)$, the $T_{ij}^L(\boldsymbol{\phi})$ are easily expressed as summations over incomplete gamma functions. It is well known that these summations are rapidly and absolutely convergent. A glance at Eq. (13) will show one that an incomplete gamma function of zero argument cannot appear in such a sum unless $\boldsymbol{\xi} = 0$ or, equivalently, $\boldsymbol{\phi} = (\pi, \pi, \cdots, \pi)$. It follows that $T_{ij}^L(\boldsymbol{\phi})$ is analytic everywhere in the first Brillouin zone except at $\boldsymbol{\xi} = 0$.

[4] P. P. Ewald, Ann. Physik (Leipzig) **64**, 253 (1921).
[5] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, London, 1954).

Expansions of the T_{ij}^L about $\boldsymbol{\xi} = 0$ are easily obtained. These are listed in Appendix A, Eqs. (A3)–(A10). From these expansions and Eqs. (9) and (10) the general form for the expansion of A_{ij} about $\boldsymbol{\xi} = 0$ can be found. It is given by the following set of relations:

(a) If $-\frac{1}{2}p + \frac{1}{2}N - 1$ is not a negative integer, then

$$A_{ij} = \delta_{ij} \times \text{const} + \frac{2p\pi^{p+2-\frac{1}{2}N}\Gamma(\frac{1}{2}N - \frac{1}{2}p)}{\Gamma(1+\frac{1}{2}p)[\Sigma_k(\xi_k/2\pi)^2]^{\frac{1}{2}N-\frac{1}{2}p}}\left(\frac{\xi_i}{2\pi}\right)\left(\frac{\xi_j}{2\pi}\right) + \text{power series}. \tag{16}$$

(b) If $-\frac{1}{2}p + \frac{1}{2}N - 1$ is a negative integer, then

$$A_{ij} = \delta_{ij} \times \text{const} + \frac{2p(-1)^{\frac{1}{2}p-\frac{1}{2}N+1}\pi^{p-\frac{1}{2}N+2}}{\Gamma(1+\frac{1}{2}p)\Gamma(\frac{1}{2}p - \frac{1}{2}N + 1)}\left(\frac{\xi_i}{2\pi}\right)\left(\frac{\xi_j}{2\pi}\right) \times \left[\Sigma_k\left(\frac{\xi_k}{2\pi}\right)^2\right]^{\frac{1}{2}p-\frac{1}{2}N} \ln \Sigma_r\left(\frac{\xi_r}{2\pi}\right)^2 + \text{power series}. \tag{17}$$

In the above equations, the power series begins with second-order terms and contains terms of even order only. Clearly, the A_{ij} are nonanalytic at $\boldsymbol{\xi} = 0$ [that is, at $\boldsymbol{\phi} = (\pi, \pi, \cdots, \pi)$]. From the cubic symmetry, we see that this point is a critical point.

From Eqs. (15) and (16) we see that the following rules govern the behavior of the A_{ij} near the critical point $\boldsymbol{\xi} = 0$:

(a) If $p > N$, the second-order terms in the power series dominate the nonanalytic term.

(b) If $N - 2 < p \leq N$, the nonanalytic term dominates the second-order terms in the power series.

(c) In the limit as p approaches $N - 2$, the nonanalytic terms approach terms which are not well defined at $\boldsymbol{\xi} = 0$. That is, the values approached by these terms as $\boldsymbol{\xi}$ approaches zero depend upon the direction of approach.

In what follows we shall be primarily interested in cases (b) and (c). In case (a), the behavior of the frequency spectrum at the frequency of the critical point will be qualitatively the same as when no long-range interaction is present.

III. DISPERSION RELATIONS FOR A THREE-DIMENSIONAL MODEL

In the remainder of this paper we shall be primarily interested in three-dimensional simple cubic lattices since much previous work has been done concerning one- and two-dimensional lattices.[3,6,7] From the discussion in Sec. II, it is clear that, if $N = 3$ and $p > 3$, the qualitative behavior of the frequency

[6] H. B. Rosenstock, Phys. Rev. **111**, 755 (1958).
[7] M. Smollett, Proc. Phys. Soc. (London) **A65**, 109 (1952).

spectra will be the same as when only short-range interactions are present. Thus, our discussion will be limited to the case of $1 \leq p \leq 3$. It is also clear that in the limit of $p = 1$, the A_{ij} are not well defined at $\boldsymbol{\xi} = 0$. As a result, we will initially assume that $1 < p \leq 3$, and then study the limiting behavior of our results as p approaches unity.

From Eqs. (16) and (17) and Appendix A, we see that the general forms of the A_{ij} near $\boldsymbol{\xi} = 0$ are given by

$$A_{11} = \alpha + \beta\xi_1^2 + \gamma(\xi_2^2 + \xi_3^2) - 2\pi\xi_1^2 h(\xi^2) + O(\xi^4) \tag{18}$$

and

$$A_{12} = \delta\xi_1\xi_2 - 2\pi\xi_1\xi_2 h(\xi^2) + O(\xi^4), \tag{19}$$

with the remaining A_{ij} following from the cubic symmetry. Here α, β, γ, and δ are constants. The function $h(\xi^2)$ is given by the following set of equations:

(a) If $1 < p < 3$, then

$$h(\xi^2) = -2^{1-p}\pi^{\frac{1}{2}p}\frac{\Gamma(\frac{3}{2} - \frac{1}{2}p)}{\Gamma(1 + \frac{1}{2}p)}\frac{1}{(\xi^2)^{\frac{3}{2}-\frac{1}{2}p}}. \tag{20}$$

(b) If $p = 3$,

$$h(\xi^2) = \ln \xi^2. \tag{21}$$

We next determine the dispersion relations for the lattice in the region near $\boldsymbol{\xi} = 0$. These are most easily obtained by applying perturbation theory to the solution of Eq. (6). Let $\mathbf{A}$ be the matrix whose elements are the A_{ij}. We write $\mathbf{A}$ as the sum of two matrices

$$\mathbf{A} = \mathbf{A}^0 + \mathbf{A}^p, \tag{22}$$

where

$$\mathbf{A}^0 = \begin{bmatrix} \alpha - 2\pi\xi_1^2 h(\xi^2) & -2\pi\xi_1\xi_2 h(\xi^2) & -2\pi\xi_1\xi_3 h(\xi^2) \\ -2\pi\xi_1\xi_2 h(\xi^2) & \alpha - 2\pi\xi_2^2 h(\xi^2) & -2\pi\xi_2\xi_3 h(\xi^2) \\ -2\pi\xi_1\xi_3 h(\xi^2) & -2\pi\xi_2\xi_3 h(\xi^2) & \alpha - 2\pi\xi_3^2 h(\xi^2) \end{bmatrix} \tag{23}$$

and

$$\mathbf{A}^p = \begin{bmatrix} \beta\xi_1^2 + \gamma(\xi_2^2 + \xi_3^2) & \delta\xi_1\xi_2 & \delta\xi_1\xi_3 \\ \delta\xi_1\xi_2 & \beta\xi_2^2 + \gamma(\xi_1^2 + \xi_3^2) & \delta\xi_2\xi_3 \\ \delta\xi_1\xi_3 & \delta\xi_2\xi_3 & \beta\xi_3^2 + \gamma(\xi_1^2 + \xi_2^2) \end{bmatrix} + O(\xi^4). \tag{24}$$

We consider $\mathbf{A}^0$ to be the unperturbed matrix and $\mathbf{A}^p$ to be the perturbation. Perturbation theory to first order is easily carried out because the secular equation giving the eigenvalues of $\mathbf{A}^0$ can be factored. Our results for each of the three branches of the dispersion relations near $\boldsymbol{\xi} = 0$ are as follows:

$$\lambda_1 = \alpha - 2\pi\xi^2 h(\xi^2) + \xi^{-2}[\beta\langle 4\rangle + 2(\gamma + \delta)\langle 2, 2\rangle] + O(\xi^2/h(\xi^2)) \tag{25}$$

and

$$\left.\begin{matrix}\lambda_2 \\ \lambda_3\end{matrix}\right\} = \alpha + \xi^{-2}[\gamma\langle 4\rangle + (\beta - \delta + \gamma)\langle 2, 2\rangle \pm |(\beta - \delta - \gamma)|\,(\langle 4, 4\rangle - \langle 4, 2, 2\rangle)^{\frac{1}{2}}] + O(\xi^2/h(\xi^2)). \tag{26}$$

The bracketed terms are certain polynomials in the ξ_i which are defined in Appendix C. The remainder in the above expressions appears because of the fourth-order remainder in the expression for $\mathbf{A}^p$ and because we neglected contributions given by second- and higher-order perturbation theory.

The term $h(\xi^2)$ does not appear explicitly in the expressions for λ_2 and λ_3. On the other hand $h(\xi^2)$ contributes to the leading behavior of λ_1 as $\boldsymbol{\xi}$ approaches zero. From Eqs. (20), (21), and (25), we see that the leading terms for the first branch are as follows:

(a) If $1 < p < 3$, then

$$\lambda_1 = \alpha + 2^{2-p}\pi^{\frac{3}{2}}p\frac{\Gamma(\frac{3}{2} - \frac{1}{2}p)}{\Gamma(1 + \frac{1}{2}p)}\xi^{p-1} + O(\xi^2). \tag{27}$$

(b) If $p = 3$, then

$$\lambda_1 = \alpha - 2\pi\xi^2 \ln \xi^2 + O(\xi^2). \tag{28}$$

Clearly, if p obeys $1 < p \leq 3$, the point $\boldsymbol{\xi} = 0$ is a minimum in λ_1.

From Eqs. (26), (27), and (28), it is clear that λ approaches α as $\boldsymbol{\xi}$ approaches zero in all three branches provided $p > 1$. However, if we take the limit of Eq. (27) as p approaches one, we see that the value of λ_1 at $\boldsymbol{\xi} = 0$ is ambiguous. For example, if we first let $\boldsymbol{\xi}$ go to zero and then let p go to one, λ_1 goes to α. However, if we first let p go to one and then let $\boldsymbol{\xi}$ go to zero, λ_1 goes to $\alpha + 4\pi$. But one thing is clear. Consider a neighborhood of $\boldsymbol{\xi} = 0$ sufficiently small that second-order terms in Eq. (27) can be neglected. As p approaches one, λ_1 approaches $\alpha + 4\pi$ unambiguously at every point in this neighborhood except $\boldsymbol{\xi} = 0$ itself.

No ambiguities occur in the second and third branches as p approaches one. Thus, we assume that Eq. (26) holds true for $p = 1$. The ambiguity in the first branch occurs only at a single point $\boldsymbol{\xi} = 0$. Thus, we assume that, if we exclude this single point, the

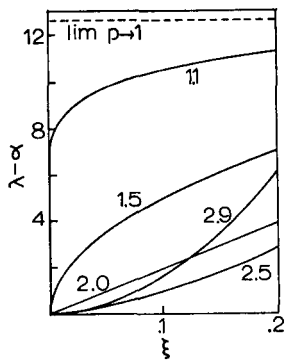


FIG. 1. The behavior of the leading term in λ_1 near $\xi = 0$ for several values of p.

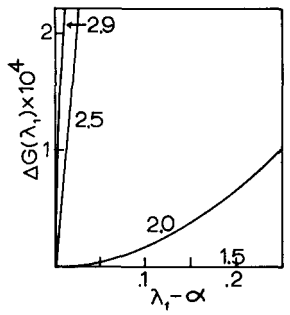


FIG. 2. The contribution $\Delta G(\lambda_1)$ of the region about $\xi = 0$ to the frequency spectrum of the first branch for several values of p. The contribution when $p = 1.5$ is too small to rise perceptively above the horizontal axis.

behavior of λ_1 when $p = 1$ is obtained by taking the limit of Eq. (25) as p approaches one. We obtain the following expansion for λ_1 when $p = 1$:

(a) In Cartesian coordinates

$$\lambda_1 = \alpha + 4\pi + \xi^{-2}[\beta\langle 4\rangle + 2(\gamma + \delta)\langle 2, 2\rangle] + O(\xi^4). \tag{29}$$

(b) In spherical coordinates

$$\lambda_1 = \alpha + 4\pi + \beta\xi^2 - 2(\beta - \gamma - \delta)\xi^2 H(\theta, \phi) + O(\xi^4), \tag{30}$$

where

$$H(\theta, \phi) = -\tfrac{1}{8}[7 + \cos(4\phi)]\sin^4\theta + \sin^2\theta. \tag{31}$$

In Fig. 1 we show the behavior of the leading terms in λ_1 as p approaches unity.

IV. THE FREQUENCY SPECTRUM FOR THE THREE-DIMENSIONAL MODEL

The contributions of the region about $\xi = 0$ to the frequency spectra of the second and third branches will depend strongly upon the values of the constants in Eq. (26). However, one can see that $\nabla_\phi\lambda_2 = \nabla_\phi\lambda_3 = 0$ at $\xi = 0$, so that these are ordinary non-analytic critical points. Their contribution to the frequency spectrum should produce the usual inverse square-root infinite slopes associated with such points.[2]

Our detailed discussion will be confined to the contribution of the first branch to the frequency spectrum. The contribution of the region about $\xi = 0$ is easily calculated provided $1 < p \leq 3$, for then the surfaces of constant frequency near the critical point are spheres. The volume enclosed by such a sphere is $V = \frac{4}{3}\pi\xi^3$ and the (unnormalized) contribution to the frequency spectrum $\Delta G(\lambda_1)$ is given by

$$\Delta G(\lambda_1) = dV/d\lambda_1. \tag{32}$$

Using Eqs. (27), (28), and (32) we obtain the following contributions for $\lambda_1 > \alpha$:

(a) If $1 < p < 3$, then

$$\Delta G(\lambda_1) = \frac{4\pi}{p-1}\left[2^{2-p}p\pi^{\frac{3}{2}}\frac{\Gamma(\frac{5}{2} - \frac{1}{2}p)}{\Gamma(1 + \frac{1}{2}p)}\right]^{3/(1-p)} \times (\lambda_1 - \alpha)^{(4-p)/(p-1)}, \tag{33}$$

and its derivative is

$$\frac{d\Delta G(\lambda_1)}{d\lambda_1} = \frac{4\pi(4-p)}{(p-1)^2}\left[2^{2-p}\pi^{\frac{3}{2}}p\frac{\Gamma(\frac{3}{2} - \frac{1}{2}p)}{\Gamma(1 + \frac{1}{2}p)}\right]^{3/(1-p)} \times (\lambda_1 - \alpha)^{(5-2p)/(p-1)}. \tag{34}$$

(b) If $p = 3$, then

$$\Delta G = [(\lambda_1 - \alpha)/2\pi]^{\frac{1}{2}}\{-\ln[(\lambda_1 - \alpha)/2\pi]\}^{-\frac{3}{2}}, \tag{35}$$

and the leading term in its derivative is

$$\frac{d\Delta G(\lambda_1)}{d\lambda_1} = \frac{1}{4\pi[(\lambda_1 - \alpha)/4\pi]^{\frac{1}{2}}[-\ln((\lambda_1 - \alpha)/2\pi)]^{\frac{3}{2}}}. \tag{36}$$

Equation (35) was obtained with the help of the method of Gillis and Weiss.[8]

If only short-range interactions were present, we would usually expect an inverse square-root infinity in $d\Delta G(\lambda_1)/d\lambda_1$ at $\lambda_1 = \alpha$. From Eqs. (34) and (36), we see that the following is true for our model:

(a) If $1 < p < 2.5$, then $d\Delta G(\lambda_1)/d\lambda_1$ approaches zero as λ_1 approaches α from above.

(b) If $p = 2.5$, then $d\Delta G(\lambda_1)/d\lambda_1$ approaches a finite value as λ_1 approaches α from above.

(c) If $2.5 < p < 3$, then $d\Delta G(\lambda_1)/d\lambda_1$ goes to infinity like $(\lambda_1 - \alpha)^{-n}$, where $0 < n < \frac{1}{2}$, as λ_1 approaches α from above.

(d) If $p = 3$, then $d\Delta G(\lambda_1)/d\lambda_1$ goes to infinity like $(\lambda_1 - \alpha)^{-\frac{1}{2}}[-\ln(\lambda_1 - \alpha)]^{-\frac{3}{2}}$ as λ_1 approaches α from above.

In Fig. 2, the leading behavior of $\Delta G(\lambda_1)$ near $\lambda_1 = \alpha$ is shown for several values of p.

In the limit of $p = 1$, the region about $\xi = 0$ makes no contribution to the frequency spectrum at $\lambda_1 = \alpha$, for at all points in a small neighborhood of $\xi = 0$, except $\xi = 0$ itself, λ_1 is approximately $\alpha + 4\pi$. Thus, this region will contribute to $\Delta G(\lambda_1)$ near the latter value of λ_1. In fact, one can easily show that the right-hand side of Eq. (33) approaches zero as p approaches one. In a later more specific example, the behavior of $\Delta G(\lambda_1)$ will be determined for the case of $p = 1$.

V. VIBRATIONAL MODES

It is well known that if ξ lies in the [100], [110], or [111] directions, our vibrational modes will be

[8] J. Gillis and G. H. Weiss, Phys. Rev. **115**, 1520 (1959).

pure transverse or pure longitudinal. We wish to determine which of our branches is the longitudinal branch and which are the transverse branches.

First consider the case where ξ lies in the [110] direction. The eigenvector equation (6) is easily solved for this case since now $A_{11} = A_{22}$ and $A_{13} = A_{23} = 0$. We find that for the longitudinal mode $\lambda = A_{11} + A_{12}$, and, referring to Eq. (25), we easily identify this frequency with the first branch. The first transverse mode is characterized by $\lambda = A_{33}$, $U_1 + U_2 = 0$, and $U_3 = 0$. Using Eq. (26), we identify this mode with the second branch. Finally, the second transverse mode is characterized by $\lambda = A_{11} - A_{12}$ and $U_1 = U_2 = 0$. It is identified with the third branch.

Similar analyses with ξ pointing in the [100] and [111] directions show that the first branch is to be identified with the longitudinal mode and the second and third branches with the transverse modes. In these cases, the transverse modes are degenerate.

So far, we have made no careful analysis of the vibrational modes for ξ pointing in an arbitrary direction.

VI. A MODEL FOR THE COULOMB INTERACTION

In Sec. IV we discussed the contribution which the region about $\xi = 0$ makes to the frequency spectrum for the first branch if p obeys the condition $1 < p \leq 3$. It was pointed out that the discussion does not apply to the case of $p = 1$. To study this case, we choose a particular model in which the long-range interaction is the Coulomb interaction. For the short-range interaction we choose $q = 9$ in Eq. (3) and assume that the sum extends over nearest neighbors only. The crystal is assumed to be under no outside pressure, so that the energy per cell is a minimum. From the latter condition we find that in Eq. (10) the factor $a^{p-q}H/G$ is equal to $\mu/54$ where μ is the Madelung constant. We then obtain the following values for the constants appearing in Eqs. (26) and (29): $\alpha = 5.11$, $\beta = -2.54$, $\gamma = 0.109$, and $\delta = -1.18$.

An analysis of Eq. (29) now shows that λ_1 approaches a maximum of $\alpha + 4\pi$ as ξ approaches zero. The surfaces of constant frequency are fluted spheres as shown by the contours in Fig. 3. This behavior of the dispersion relations is qualitatively different from that near $\xi = 0$ when $1 < p \leq 3$.

The form of the contribution $\Delta G(\lambda_1)$ of the region about $\xi = 0$ to the frequency spectrum is easily determined using Eqs. (30) and (32). We find that, if $\lambda_1 < \alpha + 4\pi$,

$$\Delta G(\lambda_1) = \tfrac{1}{2}L(\beta, \beta - \gamma - \delta)(\alpha + 4\pi - \lambda_1)^{\frac{1}{2}}, \quad (37)$$

where $L(\beta, \beta - \gamma - \delta)$ is a constant given by

$$L(\beta, \beta - \gamma - \delta) = \int_0^{2\pi}\int_0^{\pi} \frac{\sin\theta\, d\theta\, d\phi}{[2(\beta - \gamma - \delta)H(\theta, \phi) - \beta]^{\frac{3}{2}}}, \quad (38)$$

and $H(\theta, \phi)$ is defined in Eq. (31). Thus, $d\Delta G(\lambda_1)/d\lambda_1$ approaches infinity like $(\alpha + 4\pi - \lambda_1)^{-\frac{1}{2}}$ as λ_1 approaches $\alpha + 4\pi$ from above. This is qualitatively similar to the contribution of a critical point when no long-range interactions are present. It is quite different from the behavior of $d\Delta G(\lambda_1)/d\lambda_1$ when $1 < p \leq 3$.

In Figs. 3–5 we show contours of constant frequency in planes of constant ϕ_3 for all three branches obtained from Eqs. (26) and (29). We also give contours in the $\phi_1 = \phi_2$ plane for the second and third branches. The fluted maximum occurring at $\xi = 0$ in the first branch has been discussed above. From Figs. 4 and 5 we see that fluted saddle points occur at $\xi = 0$ in the second and third branches. The sector numbers[2] (P, N) of the fluted saddle point in the second branch are (1, 8). Those of the fluted saddle point in the third branch are (6, 1). We have compared these contours with contours obtained from a direct numerical calculation of the dispersion relations[9] and find good agreement near $\xi = 0$.

VII. DEPENDENCE OF RESULTS ON CUBIC SYMMETRY

Mention must be made that the above results depend strongly upon the cubic symmetry. Our discussion will be brief because our study of more general lattices is incomplete. As a simple example of a lattice in which cubic symmetry is broken, consider the above simple cubic lattice but with its particles constrained to vibrate in the x_1 direction only. Equations (6) and (18) then show that the dispersion relation is of the form

$$\lambda = \alpha - 2\pi\xi_1^2 h(\xi^2) + O(\xi^2). \quad (39)$$

We see from Eq. (20) that if we first take the limit of $p = 1$, the value approached by λ as ξ approaches one depends upon the direction of approach. Consider any neighborhood of $\xi = 0$, no matter how small. Values of λ ranging at least from α to $\alpha + 4\pi$ will be found in this neighborhood.

We have also studied an orthorhombic Bravais lattice using the rigid-ion approximation. The derivation in Sec. II can easily be carried out for such a lattice and the conclusions reached at the end of the section remain valid. If $1 < p \leq 3$, the dispersion relations are easily calculated using nondegenerate perturbation theory and the leading nonanalytic term due to the long-range interactions is found to occur in every branch, not just the first. We have not made any

[9] C. Mainville, M.A. thesis, Clark University, 1968.

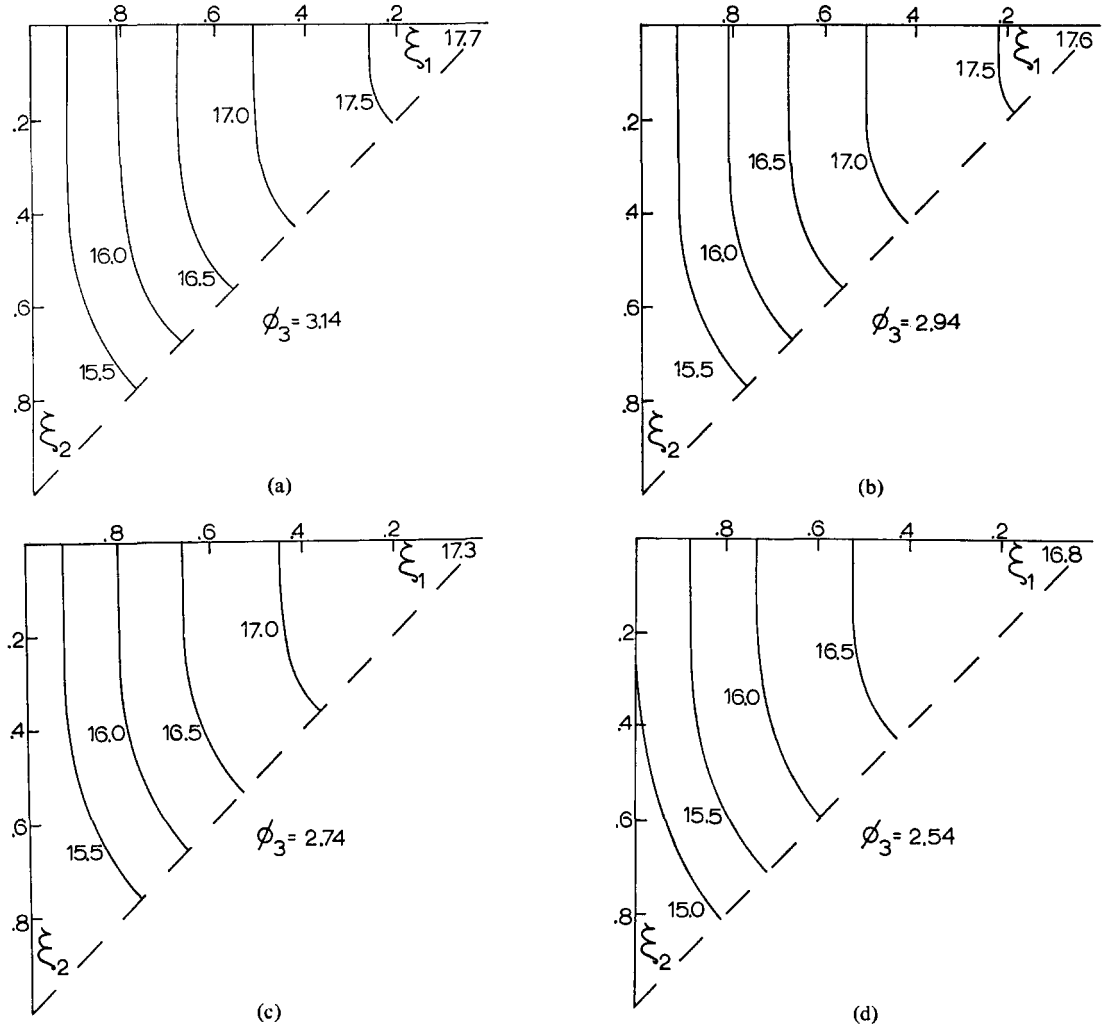


FIG. 3. Contours of constant frequency for the first branch near $\xi = 0$ when $p = 1$. In each plane of constant ϕ_3, only one-eighth of the region about $\xi = 0$ is shown.

careful analysis of the case of $p = 1$ for such lattices.

We conclude with the remark that one can easily extend the treatment in this paper to more complex lattices.

APPENDIX A

The $T^L_{ij}(\boldsymbol{\phi})$ and $T^S_{ij}(\boldsymbol{\phi})$ appearing in Eqs. (9) and (10) are defined as follows:

$$T^L_{ij}(\boldsymbol{\phi}) = -p(p+2) \times \sum_l{}' \left\{ \frac{(-1)^{\Sigma_k l_k} l_i l_j \sin(\phi_i l_i) \sin(\phi_j l_j)}{(\Sigma_k l_k^2)^{\frac{1}{2}p+2}} \times \prod_{k \neq i,j} \cos(\phi_k l_k) \right\}, \quad i \neq j, \tag{A1}$$

and

$$T^L_{ii}(\boldsymbol{\phi}) = -\sum_l{}' (-1)^{\Sigma_k l_k} \left[\frac{p}{(\Sigma_k l_k^2)^{\frac{1}{2}p+1}} - \frac{p(p+2) l_i^2}{(\Sigma_k l_k^2)^{\frac{1}{2}p+2}} \right] \times \prod_k \cos(\phi_k l_k). \tag{A2}$$

The expressions for $T^S_{ij}(\boldsymbol{\phi})$ are obtained from Eqs. (A1) and (A2) by replacing p with q, omitting the factor of minus one to the power $\Sigma_k l_k$, and regarding the sums to be finite.

The expansions of the $T^L_{ij}(\boldsymbol{\phi})$ about $\boldsymbol{\phi} = (\pi, \pi, \cdots, \pi)$ or $\boldsymbol{\xi} = 0$ are given by

$$T^L_{ij}(\boldsymbol{\phi}) = -\frac{p\pi^{\frac{1}{2}p+1}}{\Gamma(1+\frac{1}{2}p)} [g_{ij}(\xi) + h_{ij}(\xi)] + O(\xi^4), \tag{A3}$$

where $g_{ij}(\boldsymbol{\xi})$ and $h_{ij}(\boldsymbol{\xi})$ are defined below. First we write

$$g_{ij}(\boldsymbol{\xi}) = C^{(0)}_{ij} + \sum_{r,s} C^{(2)}_{ijrs} \left(\frac{\xi_r}{2\pi}\right)\left(\frac{\xi_s}{2\pi}\right). \tag{A4}$$

$C^{(0)}_{ij}$ and $C^{(2)}_{ijrs}$ are constants defined as follows:

$$C^{(0)}_{ij} = \delta_{ij}\{-2/(p+2) + \sum_l{}' [2\pi\Phi_{-\frac{1}{2}p+\frac{1}{2}N-1}(L) l_j^2 + \Phi_{\frac{1}{2}p}(L) - 2\pi\Phi_{\frac{1}{2}p+1}(L) l_j^2]\}, \tag{A5}$$

where $L = \pi\Sigma_k l_k^2$, and

$$C^{(2)}_{iirs} = (2\pi)^2 \delta_{rs} \sum_l{}' [-\tfrac{1}{2}\Phi_{\frac{1}{2}p}(L) l_r^2 + \pi\Phi_{\frac{1}{2}p+1}(L) l_i^2 l_r^2 + \pi\Phi_{-\frac{1}{2}p+\frac{1}{2}N+1}(L) l_i^2 l_r^2 - \tfrac{1}{2}\Phi_{-\frac{1}{2}p+\frac{1}{2}N}(L) l_i^2 - 2\delta_{ir}\Phi_{-\frac{1}{2}p+\frac{1}{2}N}(L) l_r^2 + (2\pi^{-1}\delta_{ir}\Phi_{-\frac{1}{2}p+\frac{1}{2}N-1}(L)]. \tag{A6}$$

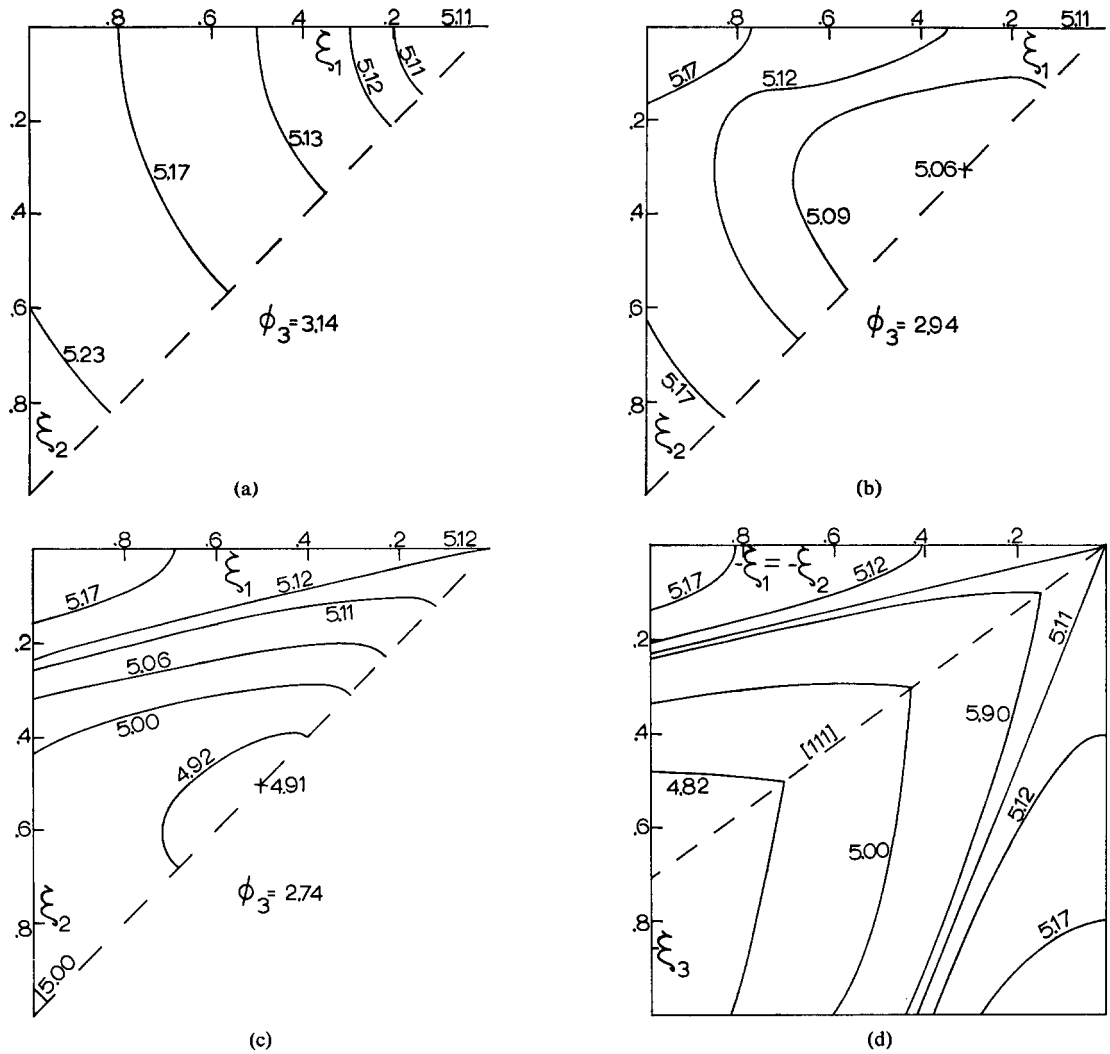


FIG. 4. Contours of constant frequency for the second branch near $\xi = 0$ when $p = 1$. Both planes of constant ϕ_3 and the $\phi_1 = \phi_2$ plane are shown.

Finally, if $i \neq j$, then

$$C^{(2)}_{ijrs} = (2\pi)^2\delta_{ir}\delta_{js}\sum_l{}' [2\pi\Phi_{\frac{1}{2}p+1}(L)l_i^2l_j^2 + 2\pi\Phi_{-\frac{1}{2}p+\frac{1}{2}N+1}(L)l_i^2l_j^2 - \Phi_{-\frac{1}{2}p+\frac{1}{2}N}(L)(l_i^2 + l_j^2) + (2\pi)^{-1}\Phi_{-\frac{1}{2}p+\frac{1}{2}N-1}(L)]. \quad \text{(A7)}$$

The second term in Eq. (A3) is defined as follows:

(a) If $-\frac{1}{2}p - 1 + \frac{1}{2}N$ is not a negative integer, then

$$h_{ij}(\xi) = \left[-\frac{4\pi}{N-p} + \frac{2\pi^{\frac{1}{2}p-\frac{1}{2}N+1}\Gamma(\frac{1}{2}N - \frac{1}{2}p)}{\Xi^{\frac{1}{2}N-\frac{1}{2}p}}\right] \times \left(\frac{\xi_i}{2\pi}\right)\left(\frac{\xi_j}{2\pi}\right). \quad \text{(A8)}$$

(b) If $p = N$, then

$$h_{ij}(\xi) = [2\pi(\gamma - \ln\pi) - 2\pi\ln(\Xi)]\left(\frac{\xi_i}{2\pi}\right)\left(\frac{\xi_j}{2\pi}\right), \quad \text{(A9)}$$

where $\gamma = 0.5772\cdots$ and $\Xi = \Sigma_k\xi_k^2/(2\pi)^2$.

(c) If $-\frac{1}{2}p - 1 + \frac{1}{2}N$ is a negative integer less than -1, then

$$h_{ij}(\xi) = \left[-\frac{4\pi}{N-p} + \frac{2(-\pi)^{\frac{1}{2}p-\frac{1}{2}N+1}}{\Gamma(\frac{1}{2}p - \frac{1}{2}N + 1)}\Xi^{\frac{1}{2}p-\frac{1}{2}N}\ln(\Xi)\right] \times \left(\frac{\xi_i}{2\pi}\right)\left(\frac{\xi_j}{2\pi}\right). \quad \text{(A10)}$$

APPENDIX B

The expansion of $\Phi_m(x)$ from zero into regions of positive x is given by the following set of formulas:

(a) If m is not a negative integer, then

$$\Phi_m(x) = \frac{\Gamma(m+1)}{x^{m+1}} - \sum_{n=0}^{\infty}\frac{(-x)^n}{(m+n+1)(n!)}. \quad \text{(B1)}$$

(b) If $m = -1$, then

$$\Phi_{-1}(x) = \gamma - \ln x + \sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{(n)(n!)}. \quad \text{(B2)}$$

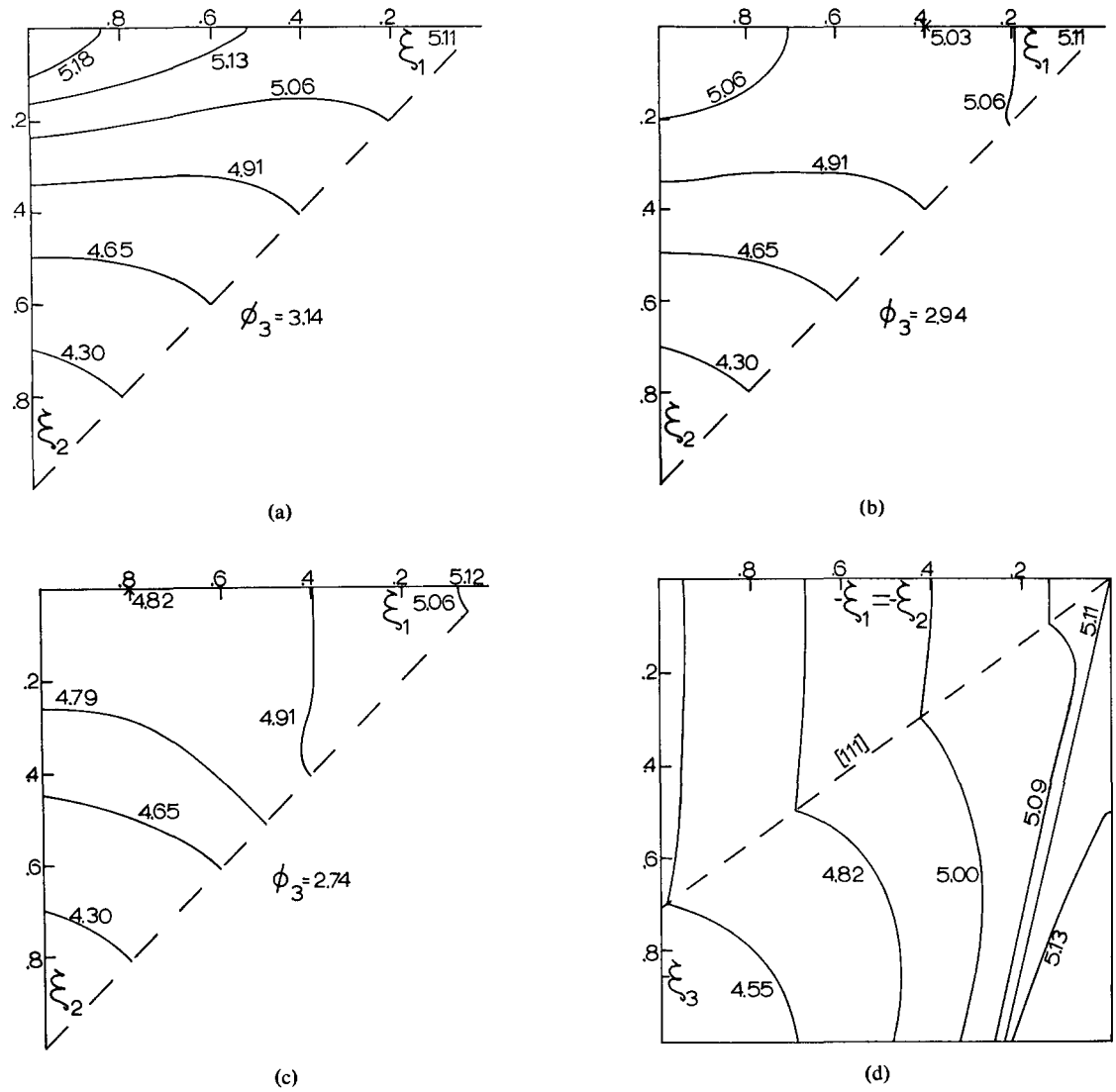


FIG. 5. Contours of constant frequency for the third branch near $\xi = 0$ when $p = 1$. Both planes of constant ϕ_3 and the $\phi_1 = \phi_2$ plane are shown.

(c) If m is a negative integer less than -1, then

$$\Phi_m(x) = \frac{1}{(-m-1)} - \frac{x}{(-m-2)} + \frac{x^2}{2!\,(-m-3)} - \cdots + \frac{(-x)^{-m-2}}{(-m-2)!} + (-1)^{-m}\frac{x^{-m-1}}{(-m-1)!}\ln x + \frac{(-x)^{-m-1}}{(-m-1)!} \times \left[\gamma + 1 + \tfrac{1}{2} + \cdots + \frac{1}{(-m-1)}\right] + (-1)^{-m+1}\left[\frac{x^{-m}}{(-m)!} - \frac{x^{-m+1}}{2(-m+1)!} + \frac{x^{-m+2}}{3(-m+2)!} - \cdots\right]. \quad \text{(B3)}$$

APPENDIX C

The bracketed terms appearing in Eqs. (25) and (26) are defined as follows. $\langle m, n, r\rangle$ is the sum of all distinct terms of the form $\xi_i^m \xi_j^n \xi_k^r$ where i, j, and k are distinct. For example,

$$\langle 4, 2, 2\rangle = \xi_1^4\xi_2^2\xi_3^2 + \xi_2^4\xi_1^2\xi_3^2 + \xi_3^4\xi_1^2\xi_2^2.$$

Finally we define $\langle m, n\rangle \equiv \langle m, n, 0\rangle$ and $\langle m\rangle \equiv \langle m, 0, 0\rangle$. For example $\langle 2\rangle = \xi^2$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Sufficient Conditions for Stability of the Faddeev Equations

W. W. ZACHARY
Naval Research Laboratory, Washington, D.C.

(Received 21 November 1968)

Functional analysis techniques are used to obtain stability theorems for the Faddeev equations in the quantum-mechanical three-particle nonrelativistic scattering theory. Sufficient conditions are obtained in order that the solutions of these equations not be sensitive to small variations of the off-shell two-particle amplitudes. These conditions provide criteria for the validity of some of the previous formal investigations of the Faddeev equations.

I. INTRODUCTION

In recent years there have been numerous applications of Faddeev's formulation of the quantum-mechanical three-particle scattering problem with pairwise interactions[1] to various aspects of nuclear and particle physics,[2] liquid helium,[3] and quantum statistical mechanics.[4] One of the advantages of this formulation is that the three-particle scattering problem is expressed in terms of the exact solutions of the scattering problems for the two-particle subsystems.

One part of the theory which is still lacking, however, is a "stability theorem" stating that the solutions of the Faddeev equations are not sensitive to small variations of the two-particle off-shell amplitudes. The need for such a result was first noted by Lovelace[5] and more recently by others.[6] As Lovelace emphasized, the stability question is particularly important in practice because the off-shell two-particle amplitudes are, in general, not exactly known. Some results pertaining to the problem of stability have been obtained by varying input parameters of computer solutions of approximate versions of the Faddeev equations.[7] The trouble with this approach is that it is a difficult matter to disentangle the properties of the calculational schemes employed to solve the Faddeev equations from the properties of the equations themselves. From a different point of view, several authors have considered modifications of the Faddeev equations resulting from the expression of the two-particle amplitudes as the sum of two parts.[8]

The purpose of the present paper is to investigate the stability of the Faddeev equations by using the rigorous formulation of perturbation theory for linear operators in Banach spaces.[9] This work may be considered the analog for the three-particle scattering problem of that of Grossmann[10] on the two-particle problem in which the stability of the Lippmann–Schwinger equation was investigated in a rigorous manner.

In Sec. II we pose the stability problem for the Faddeev equations and discuss the relationship of our approach to previous work. Then in Sec. III we lay the foundation for our work, including a short review of some known results in a form suitable for our purposes. In Sec. IV, sufficient conditions are obtained for stability of the inhomogeneous Faddeev equations.

[1] L. D. Faddeev, *Mathematical Aspects of the Three-Body Problem in the Quantum Scattering Theory* (Israel Program for Scientific Translations, Jerusalem, 1965). References to earlier papers by this author can be found here.

[2] We refer to two reviews of these developments: K. M. Watson, J. Nuttall, and J. S. R. Chisholm, *Topics in Several Particle Dynamics* (Holden-Day, Inc., San Francisco, Calif., 1967); I. Duck, in *Advances in Nuclear Physics*, M. Baranger and E. Vogt, Eds. (Plenum Press, New York, 1968), Vol. I.

[3] T. W. Burkhardt, Ann. Phys. (N.Y.) **47**, 516 (1968); E. Østgaard, Phys. Rev. **171**, 248 (1968).

[4] W. G. Gibson, Phys. Letters **21**, 619 (1966); A. S. Reiner, Phys. Rev. **151**, 170 (1966); B. J. Baumgarth, Z. Physik **198**, 148 (1967).

[5] C. Lovelace, in *Strong Interactions and High Energy Physics*, edited by R. G. Moorhouse (Plenum Press, Inc., New York, 1964); Phys. Rev. **135**, B1225 (1964).

[6] M. Fontannaz, Nuovo Cimento **53B**, 53 (1968); *International Nuclear Physics Conference*, C. D. Goodman, P. H. Stelson, and A. Zucker, Eds. (Academic Press Inc., New York, 1967), see the discussion following the talk by R. D. Amado; R. Blankenbecler, in *Theory of Three-Particle Scattering in Quantum Mechanics*, J. Gillespie and J. Nuttall, Eds. (W. A. Benjamin, Inc., New York, 1968).

[7] R. Aaron, R. D. Amado, and Y. Y. Yam, Phys. Rev. **136**, B650 (1964); A. G. Sitenko, V. F. Kharchenko, and N. M. Petrov, Phys. Letters **21**, 54 (1966); H. Hebach, P. Henneberg, and H. Kümmel, *ibid.* **24B**, 134 (1967); G. L. Shrenk and A. N. Mitra, Phys. Rev. Letters **19**, 530 (1967); W. Bierter and K. Dietrich, Nuovo Cimento **52A**, 1209 (1967); M. Fontannaz, Ref. 6; N. M. Petrov, S. A. Storozhenko, and V. F. Kharchenko, Yad. Fiz. **6**, 466 (1967) [Sov. J. Nucl. Phys. **6**, 340 (1968)].

[8] E. O. Alt, P. Grassberger, and W. Sandhas, Nucl. Phys. **B2**, 167 (1967), and references contained therein; I. H. Sloan, Phys. Rev. **165**, 1587 (1968); M. G. Fuda, *ibid.* **166**, 1064 (1968); R. J. Yaes, *ibid.* **170**, 1236 (1968). We thank Dr. Yaes for preprints of his work prior to publication.

[9] See, for example, T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Inc., New York, 1966). We shall quote some standard results from this reference, referred to as K.

[10] A. Grossmann, J. Math. Phys. **2**, 714 (1961). *Note added in proof:* Similar work was also done by A. Ya. Povzner and T. Ikebe. See T. Ikebe, Arch. Ratl. Mech. Anal. **5**, 1 (1960), where reference to Povzner's work is given.

Section V is then devoted to the discussion of the corresponding sufficient conditions for stability of the homogeneous equations. Finally, Sec. VI consists of the concluding remarks.

II. POSING THE STABILITY PROBLEM

The Faddeev equations are a set of three coupled linear integral equations which can be written in the inhomogeneous and homogeneous forms, respectively, as

$$f(z) = g(z) + A(z)f(z), \tag{1}$$

$$f(z) = A(z)f(z). \tag{2}$$

Here, z is an energy parameter which may assume complex values and the integral operator A depends upon the two-particle amplitudes in a linear manner.

These equations are defined in a scale of Banach spaces $B(\theta, \mu)$ consisting of bounded, Hölder-continuous, complex-valued functions on R^6 which are characterized by their Hölder-exponent μ and their behavior at infinity (θ). The definition of $B(\theta, \mu)$ is somewhat involved and we refer to Ref. 1 for these details. We consider instead, solely for the sake of simplicity, the scale of Banach spaces $\bar{B}(\theta, \mu)$ consisting of all bounded, continuous, complex-value functions on R^1 satisfying

$$\begin{aligned} |f(x+y) - f(x)| &\leq \text{const} \times |y|^{\mu}, \\ &\quad |y| \leq 1, \quad 0 < \mu < 1, \\ |f(x)| &\leq \text{const} \times (1 + |x|)^{-\theta}, \quad \theta > 0, \end{aligned} \tag{3}$$

normed by

$$\|f\|_{\theta,\mu} = \sup_{x,|y|\leq 1} (1 + |x|)^{\theta} \times \left[|f(x)| + \frac{|f(x+y) - f(x)|}{|y|^{\mu}} \right]. \tag{4}$$

Anything that we prove for $\bar{B}(\theta, \mu)$ can also be proved for $B(\theta, \mu)$. The procedure of using $\bar{B}$ instead of B is not necessary and does not alter the results in any way. The sole reason for introducing $\bar{B}$ is to simplify the proof of Theorem 4.

Consider an operator of the form

$$H(\kappa, z) \equiv A(z) + \kappa C(z), \quad 0 \leq \kappa \leq 1, \tag{5}$$

and write the inhomogeneous and homogeneous forms, respectively, of the corresponding equation

$$F(\kappa, z) = G(\kappa, z) + H(\kappa, z)F(\kappa, z), \tag{6}$$

$$F(\kappa, z) = H(\kappa, z)F(\kappa, z). \tag{7}$$

The stability problem can now be stated in the following form. If $G(\kappa, z)$ and $H(\kappa, z)$ are continuous in κ as $\kappa \to 0$ (in a sense to be specified), then under what conditions are the solutions $F(\kappa, z)$ of (6) and the solutions $F(\kappa, z)$ and eigenvalues $z = z(\kappa)$ of (7) continuous in κ as $\kappa \to 0$?

These questions are mathematically nontrivial because the operator A is unbounded and, moreover, is not densely defined.[11] If the function space involved here were finite-dimensional, the stability problem could readily be solved.[12] It will be shown in Sec. V that the question of the stability of the eigenvalues $z(\kappa)$ and the solutions $F(\kappa, z)$ obtained from (7) can be reduced to considerations in such a space. Unfortunately, no such reduction is possible for the solutions $F(\kappa, z)$ of (6).

The physical situation that we have in mind in formulating the stability problem in the above manner is as follows. We first consider a three-particle scattering problem described by (1) and (2). We then modify the two-particle amplitudes resulting in the corresponding Eqs. (6) and (7). The objective is then to discuss the relationship between the solutions of (6) and (7) and of (1) and (2) resulting from the modification of the operator A, as given by (5). This formulation of the problem allows us to investigate the effect of adding more terms in the separable approximation[7] and to give sufficient conditions for the validity of the previous perturbation approaches to the Faddeev equations.[8] With regard to the former application it is essential that we do not restrict C to be a single operator, but allow it to consist of a sum of operators. Our conditions will always involve the "total" operator C.

III. REVIEW OF PREVIOUS RESULTS

We first discuss some known results concerning (1) and (2).

Theorem 1 (Fredholm alternative[11,1]*):* Suppose that the two-particle potentials $V_{ij} \in \bar{B}(\theta_0, \mu_0)$, $\theta_0 > \frac{1}{2}$ [$\theta_0 > \frac{3}{2}$ in the case of $B(\theta_0, \mu_0)$], $\mu_0 > \frac{1}{2}$, $1 \leq i < j \leq 3$, are real-valued. Then for $\theta \leq \theta_0$, $\mu \leq \mu_0$, there exists a positive integer n such that A^m has a compact extension K from its domain $D(A^m)$ to the entire space $X_0 \equiv \bar{B}(\theta, \mu)$ which maps X_0 into $D(A^m)$ for all positive integers $m \geq n$. Moreover, for a given value of z, either (2) has a nontrivial solution, $0 \neq f \in D(A^m)$, or there exists a unique solution $f \in D(A)$ of (1).

Definition 1[9]*:* A linear operator T is *closed* if $x_n \in D(T)$, $x_n \to x$, and $Tx_n \to y$ imply that $x \in D(T)$ and $Tx = y$.

[11] S. Albeverio, W. Hunziker, W. Schneider, and R. Schrader, Helv. Phys. Acta **40**, 745 (1967).

[12] The relevant theorems are contained in Ref. 9, Chap. II.

Lemma 1: Given the conditions on the two-particle potentials of Theorem 1, then A is a closed linear operator.

This result is contained in Faddeev's work,[1] although it is not expressed in this terminology.

Lemma 2: If A is as in Lemma 1 and P is a polynomial, then $P(A)$ is a closed linear operator.

Proof: By virtue of Theorem 1, Lemma 1, and a well-known result,[13] we have only to show that a value of z exists such that (2) has only the trivial solution. However, Faddeev showed that this is true for any z with $\operatorname{Im} z \neq 0$.

Following Albeverio *et al.*,[11] we define a norm on $D(A^k)$ by

$$\|f\|_k = \sum_{i=0}^{k} \|A^i f\|_0 \tag{8}$$

in terms of the norm on X_0. It now follows from Lemma 2 that these domains are Banach spaces (K.IV.1.4), which will be denoted by X_k.

Following the notation of Kato,[9] we denote by $B(X, Y)$ the set of all bounded linear operators on X to Y, where X and Y are Banach spaces. If $X = Y$ we write $B(X, X) = B(X)$. $B(X, Y)$ is a Banach space with the norm of $T \in B(X, Y)$ defined in the uniform operator topology by

$$\|T\|_{XY} \equiv \sup_{0 \neq x \in X} \frac{\|Tx\|_Y}{\|x\|_X}. \tag{9}$$

Also, we denote by $C(X, Y)$ the set of all closed linear operators in X to Y. If $T \in C(X_0, Y)$, then, for simplicity, we will not use a special notation for the restriction of T to X_k. Now, we have

Theorem 2[11]*:* If Theorem 1 holds and if (1) has a unique solution $f \in X_1$, then $(I - A)^{-1} \in B(X_m, X_1)$; i.e., there exists a positive constant a such that (I is the identity operator):

$$\|(I - A)^{-1} g\|_1 \leq a \|g\|_m \quad \text{for all} \quad g \in X_m. \tag{10}$$

Indeed, by making use of the identity

$$(I - A)^{-1} = (I - A^m)^{-1} \sum_{j=0}^{m-1} A^j,$$

it is easily seen that a permissible choice for a is

$$a = 2 \|(I - K)^{-1}\|_0.$$

[13] See, for example, N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Vol. I, p. 602.

IV. STABILITY OF THE INHOMOGENEOUS FADDEEV EQUATIONS

Our first result concerning the stability of the Faddeev equation is

Theorem 3: If z is such that $[I - A(z)]^{-1}$ exists in the sense of Theorem 2, $D(C) \supset D(A)$, and

$$\|C(z)u\|_m \leq d \|u\|_1 \quad \text{for all} \quad u \in X_1, \tag{11}$$

with

$$d < 1/a, \tag{12}$$

i.e., $C(z) \in B(X_1, X_m)$, then $[I - H(\kappa, z)]^{-1}$ exists as an element of $B(X_m, X_1)$ and is continuous in the uniform operator topology for sufficiently small $\kappa \geq 0$.

Proof: Consider the formal identity

$$\begin{aligned}[I - H(\kappa, z)]^{-1} &= [I - A(z)]^{-1}[I - \kappa S(z)]^{-1} \\ &= [I - A(z)]^{-1} \sum_{j=0}^{\infty} [\kappa S(z)]^j \end{aligned} \tag{13}$$

in which $S(z) \equiv C(z)[I - A(z)]^{-1}$.

Combining (10) and (11) we see that $S(z) \in B(X_m)$. Using (10) again it follows from (13) that

$$[I - H(\kappa, z)]^{-1} \in B(X_m, X_1)$$

if it exists. Since $\kappa \in [0, 1]$, a sufficient condition for the absolute and uniform (in κ) convergence of the series (13) is

$$\|S(z)\|_m < 1. \tag{14}$$

Combining this condition with (10) and (11) gives the bound (12) on d. With this condition $[I - H(\kappa, z)]^{-1}$ exists.

Using (10) and (13) we obtain

$$\begin{aligned}\|\{[I - H(\kappa, z)]^{-1} - [I - A(z)]^{-1}\} g\|_1 \\ \leq a |\kappa| \|S(z)\|_m [I - |\kappa| \|S(z)\|_m]^{-1} \|g\|_m \end{aligned}$$

for all $g \in X_m$. Hence, $[I - H(\kappa, z)]^{-1}$ is continuous in κ for sufficiently small $\kappa \geq 0$ in the uniform topology, and the theorem is proved.

The bound on d obtained above is not very illuminating. Using an elementary identity we find

$$d < \tfrac{1}{2} \|I - K(z)\|_0 \leq \tfrac{1}{2}[1 + \|K(z)\|_0]. \tag{15}$$

A sufficient condition for the existence of $[I - K(z)]^{-1}$ is $\|K(z)\|_0 < 1$, which is satisfied for sufficiently large values of $|z|$[1]. In this case we have $d < 1$. In general however, it is only necessary that $\|K^p(z)\|_0 < 1$ for some positive integer p. In this case it is possible that $d \geq 1$.

The results of Theorem 3 derive from the sufficient condition (14). It is, of course, only necessary that

$\|S^p(z)\|_m < 1$ for some positive integer p. We shall limit consideration, in the present paper, to the investigation of the consequences of the sufficient condition (11).

We note that it follows directly from Theorem 3 that the solutions of (6) are continuous in κ for sufficiently small $\kappa \geq 0$ in the uniform topology. Thus, we have essentially[14] completed the stability problem for the scattering states and the rest of the paper will be primarily concerned with the bound-state problem.

V. STABILITY OF THE HOMOGENEOUS FADDEEV EQUATIONS

We can obtain from (11) two weaker conditions

$$\|C(z)u\|_1 \leq d\,\|u\|_1 \tag{16}$$

and

$$\|C(z)u\|_0 \leq d\,\|u\|_1, \quad \text{with} \quad u \in X_1, \quad D(C) \supset D(A), \quad \text{and} \quad d < 1/a. \tag{17}$$

Condition (17) is usually described by saying that C is A-bounded,[9] whereas (16) tells us that $C \in B(X_1)$.

Lemma 1 states that $A(z)$ is a closed linear operator. If (17) holds with $d < 1$ it follows (K.IV.1.1) that $H(\kappa, z)$ is also a closed linear operator. However, as noted above, the inequality $d < 1$ need not be satisfied. Thus, we need to impose a more stringent condition on the operator C. We will show, however, that this proposed condition actually follows from (16).

Definition 2[9]*:* Let T and S be two linear operators in a Banach space X such that $D(S) \supset D(T)$. S is *T-compact* if, for any sequence $u_n \in D(T)$ with both $\{u_n\}$ and $\{Tu_n\}$ bounded, $\{Su_n\}$ contains a convergent subsequence. Equivalently, $\{Su_n\}$ is relatively compact, i.e., it has compact closure.

We now have

Theorem 4. If (16) holds then C is A-compact.

This condition is stronger than A-boundedness, but the assumption that $d < 1$ is not required. In order to prove this theorem we use the following standard result.

Lemma 3 (Ascoli–Arzelà theorem): Let $C(R^1, \theta)$ be the Banach space of complex-valued continuous bounded functions normed by

$$\|f\|_C = \sup_{x \in R^1} (1 + |x|)^\theta\, |f(x)|, \quad \theta > \tfrac{1}{2}. \tag{18}$$

Consider an infinite sequence $\{f_n(x)\} \in C(R^1, \theta)$ such that: $f_n(x)$ is equicontinuous on each closed subinterval of R^1,

$$\lim_{\delta \to 0^+} \sup_{|x-x'| \leq \delta} |f_n(x) - f_n(x')| = 0 \quad \text{for all } n, \tag{19}$$

and there exists a uniform bound M and a positive number r such that

$$|f_n(x)| < M/|x|^\theta, \quad \theta > \tfrac{1}{2}, \tag{20}$$

for all x such that $|x| > r$ and all n. Then the sequence $\{f_n(x)\}$ is relatively compact in $C(R^1, \theta)$, i.e., from every such sequence a subsequence can be extracted which converges uniformly on every compact subset of R^1.

In the usual statement of this result the domain of the functions is taken to be a compact set. In that case the theorem is true if these functions are uniformly bounded and equicontinuous. In the case of interest here, the treatment of Epstein[15] shows that it is sufficient to require that the functions be equicontinuous on each compact subset of the domain and equibounded (in n) at (at least) one point, the uniform-boundedness property following from these conditions. In a paper on the two-particle scattering problem (Lippmann–Schwinger equation) Belinfante[16] stated the above theorem for functions on R^3 with a condition similar to (20), but with $\theta = 1$. In the three-particle problem, however, (20) is more natural, as will be seen below.

Proof of Theorem 4: Consider a bounded infinite sequence in X_1, $\{u_\nu\} \in X_1$, $\|u_\nu\|_1 \leq q = \text{const}$. We must show that the sequence $\{Cu_\nu\}$ has a subsequence which converges in the norm topology of X_1.

In terms of the norm on X_0 we have

$$\|A^j u_\nu\|_0 \leq q, \quad j = 0, 1.$$

It follows from (16) that

$$\|Cu_\nu\|_1 \leq dq$$

or

$$\|A^j C u_\nu\|_0 \leq dq, \quad j = 0, 1. \tag{21}$$

Since any Hölder-continuous function is *a fortiori* continuous it follows that the norms of $C(R^1, \theta)$ and X_0 are equivalent in the sense that two positive u-independent constants α and β exist such that

$$\alpha\,\|u\|_0 \leq \|u\|_C \leq \beta\,\|u\|_0. \tag{22}$$

[14] The uniqueness of these solutions will be established in Theorem 5.

[15] B. Epstein, *Partial Differential Equations* (McGraw-Hill Book Co., Inc., New York, 1962), p. 4.

[16] J. G. Belinfante, J. Math. Phys. 5, 1070 (1964).

This equivalence coupled with (21) implies that

$$\|A^j Cu_\nu\|_C \leq dqq', \quad q' = \text{const}, \quad j = 0, 1,$$

or

$$|A^j Cu_\nu(x)| \leq dqq'(1 + |x|)^{-\theta}, \quad j = 0, 1.$$

Hence, for sufficiently large $r > 0$,

$$|A^j Cu_\nu(x)| \leq \frac{dqq'}{|x|^\theta}, \quad |x| > r, \quad j = 0, 1,$$

and the sequences $\{A^j Cu_\nu(x)\}$, $j = 0, 1$, satisfy (20).

Now, introduce the norm

$$\|f\|_{C'} \equiv \sup_{x, |y| \leq 1} (1 + |x|)^\theta \frac{|f(x + y) - f(x)|}{|y|^\mu}, \quad \theta > 0, \quad 0 < \mu < 1. \quad (23)$$

Clearly this norm is equivalent in the sense of (22) to the norms on $C(R^1, \theta)$ and X_0. Combining this fact with (21) we obtain $\|A^j Cu_\nu\|_{C'} \leq dqq'$, $j = 0, 1$, $q'' = \text{const}$. Now using (23) we find

$$|A^j Cu_\nu(x + y) - A^j Cu_\nu(x)| \leq dqq'' |y|^\mu (1 + |x|)^{-\theta}, \quad j = 0, 1.$$

Hence, the sequences $\{A^j Cu_\nu\}$, $j = 0, 1$, are equicontinuous on each closed (and hence compact) subinterval of R^1.

It now follows from Lemma 3 that each of the sequences $\{A^j Cu_\nu\}$, $j = 0, 1$, have convergent subsequences in $C(R^1, \theta)$. By the equivalence of the norms these subsequences are also convergent in X_0 and, therefore, the sequence $\{Cu_\nu\}$ has a convergent subsequence in the norm topology of X_1, which completes the proof of the theorem.

Definition 3[9]: For $T \in C(X, Y)$ we define

nul T(*nullity* or *kernel index* of T)
≡ dimension of the null space of T,
def T (*deficiency* or *deficiency index* of T)
≡ dimension of the quotient space $Y/R(T)$, in which $R(T)$ denotes the range of T.

These quantities are either nonnegative integers or infinite. If they are both finite and $R(T)$ is closed, then T is said to be *Fredholm*.[17] In this case, the *index* of T is defined as

$$\text{ind } T \equiv \text{nul } T - \text{def } T.$$

If T satisfies the Fredholm alternative, then ind $T = 0$.

[17] This condition can be relaxed by only requiring that one of the pair nul T, def T be finite. In this case ind T is still well defined and T is said to be *semi-Fredholm*. We shall not require this more general concept.

It follows from Lemma 1 and an elementary result that $A - I$ is a closed linear operator. From the proof of Theorem 1[11] we have

$$\text{nul } (A - I) < \infty, \quad (24)$$

$$\text{def } (A - I) < \infty,$$

and

$$\text{ind } (A - I) = 0. \quad (25)$$

From these facts and the closedness of $R(A)$,[18] and consequently the closedness of $R(A - I)$, we see that $A - I$ is Fredholm.

We can now return to the discussion preceding Definition 2 and prove that $H(\kappa, z)$ is a closed linear operator without requiring the condition $d < 1$. Actually, we obtain additional results without stronger hypotheses.

Theorem 5: If (11) holds then the operator $H - I$ is linear, closed, Fredholm, and satisfies the Fredholm alternative.

Proof: The linearity is obvious and the remaining properties follow from the facts that $A - I$ is closed and Fredholm, Eq. (25), and Theorems 4 and K.IV.5.26.

This theorem states, among other things, that

$$\text{ind } (H - I) = \text{ind } (A - I),$$

but says nothing regarding the relative magnitudes of the respective kernel and deficiency indices. We now consider this question. A preliminary result is

Lemma 4: If (11) obtains then the set of z-values for which $[I - H(\kappa, z)]^{-1}$ does not exist (in the sense of Theorem 3) is at most countable, closed, and is contained within a finite interval on the real axis.

Proof: Faddeev[1] showed that these statements hold for $[I - A(z)]^{-1}$. It has been shown in Theorem 3 that, if $[I - A(z)]^{-1}$ exists and (11) obtains, then $[I - H(\kappa, z)]^{-1}$ also exists. The proof of the lemma follows immediately.

It follows from Theorem 5 that the set of z-values discussed in Lemma 4 is precisely that for which nontrivial solutions of (7) exist.

[18] This can easily be shown by using some elementary considerations and Eq. (7.8) of Ref. 1. This simple result was previously noted, by the use of a slightly different argument, by K. Mochizuki, J. Math. Soc. Japan **19**, 123 (1967).

We now have

Theorem 6: If A and H are as in Theorem 5, then

$$\text{nul}\,(H - I) \leq \text{nul}\,(A - I) < \infty.^{19}$$

Proof: Consider the equation

$$A(z)f(z) = \lambda(z)f(z), \tag{26}$$

where $\lambda(z)$ is the eigenvalue of A as a function of energy.[20] Choose z_0 to correspond to a nontrivial solution of (2), i.e.,

$$\lambda(z_0) = 1. \tag{27}$$

Since z_0 is an element of a set which is at most countable, we can draw a circle around z_0:

$$Z_\epsilon = \{z: |z - z_0| = \epsilon\}, \quad \epsilon > 0,$$

such that if $z \in Z_\epsilon$ and ϵ is sufficiently small then $\lambda(z) \neq 1$. In other words, $[\lambda(z) - A(z)]^{-1}$ exists in the sense of Theorem 2 for $z \in Z_\epsilon$.

As z ranges over Z_ϵ, $\lambda(z)$ traverses some path in the complex plane which can be taken as a circle Γ enclosing the point $\lambda = 1$. We now define the operator

$$P \equiv \frac{1}{2\pi i}\int_\Gamma [\lambda(z) - A(z)]^{-1}\, d\lambda \tag{28}$$

and note that Γ separates the spectrum, i.e., the z_0 in (27) is unique. P is the eigenprojection onto the closed linear manifold, M, consisting of all f satisfying (26) with $\lambda = 1$.[9] Furthermore, by (24),

$$\text{nul}\,(A - I) = \dim M < \infty. \tag{29}$$

Now consider the operator

$$P(\kappa) \equiv \frac{1}{2\pi i}\int_\Gamma [\lambda(\kappa, z) - H(\kappa, z)]^{-1}\, d\lambda, \tag{30}$$

where z and Γ are the same as above and $\lambda(\kappa, z)$ is the eigenvalue of $H(\kappa, z)$ as a function of κ and z.

From the estimate

$$\|P(\kappa)\|_{m1} \leq \rho\, \|(\lambda(\kappa, z) - H(\kappa, z))^{-1}\|_{m1},$$
$$\rho = \text{radius of } \Gamma,$$

and Theorem 3 it follows that $P(\kappa)$ exists as an element of $B(X_m, X_1)$ and is continuous in κ for sufficiently small $\kappa \geq 0$ in the uniform topology. Furthermore,[9] $P(\kappa)$ is a projection and is the sum of the eigenprojections of all eigenvalues of $H(\kappa, z)$ lying inside Γ. This shows that Γ separates the spectrum of $H(\kappa, z)$ and completes the proof of the theorem.

A direct consequence of this theorem is that, in physical terms, the eigenvalues $\lambda(\kappa, z(\kappa))$ of $H(\kappa, z(\kappa))$ are derived from the corresponding eigenvalue $\lambda(z)$ of $A(z)$ by splitting, i.e., by breaking the degeneracy.

We can now establish the stability of the bound state energies.

Theorem 7: If H is as in the above theorem, then the eigenvalues $z(\kappa)$, obtained from (7), are continuous functions of κ for sufficiently small $\kappa \geq 0$.

Proof: Let $P(\kappa)$ be as in the preceding theorem, and denote its range by $M(\kappa)$. It then follows from K.I.4.10 that $M(\kappa)$ and M are isomorphic. In particular,

$$\dim M(\kappa) = \dim M. \tag{31}$$

It now follows from K.VII.1.8 that the eigenvalues, $\lambda(\kappa, z(\kappa))$, of $H(\kappa, z)$ are continuous functions of κ, i.e., branches of one or several analytic functions which have at most algebraic singularities near $\kappa = 0$.[21] Moreover, Kato's discussion emphasizes that, since we are dealing with a finite system of eigenvalues (the part of the spectrum inside Γ), the problem is reduced to a problem in a finite-dimensional space. We need only restrict ourselves to the subspace $M(\kappa)$.

In view of the finite-dimensional nature of the problem we can now use the result[22] that $\lambda(\kappa, z)$ is continuous separately in the variables κ and z if $H(\kappa, z)$ has this property. The continuity of $z(\kappa)$ in κ now follows from the continuity of $\lambda(\kappa, z)$ in the two variables separately, the mapping (27), and the discussion following that equation. This completes the proof of the theorem.

Theorem 8: If H is as in Theorem 5, then the nontrivial solutions of (7), $F(\kappa, z)$, are continuous functions of κ for sufficiently small $\kappa \geq 0$.

Proof: This result follows as a corollary to the proof of the preceding theorem.

[19] This theorem is true in a finite-dimensional space (see Ref. 9, Chap. II) and is what one expects physically. In the present case with infinite dimensionality it can be proved if C is A-bounded with $d < 1$ (K.IV.5.22). This restriction on d can be lifted if C is a single operator and, furthermore, the equality holds in this case (K.IV.5.31). We pointed out in Sec. II, however, that this condition is too abusive to the physical problem. By making use of our previous results we give a different proof of this theorem, unhampered by the restrictions noted above.

[20] There should be no confusion concerning our two usages of the word "eigenvalue." In (26), λ is an eigenvalue of A, whereas z is a discrete eigenvalue of the Hamiltonian.

[21] Kato states this theorem for the case in which $H(\kappa, z)$ is holomorphic in κ near $\kappa = 0$, whereas we have been concerned with the case in which $H(\kappa, z)$ is only continuous in $\kappa \in [0, 1]$ for sufficiently small κ. Following a discussion by Kato, however, (Ref. 9, p. 365) we can consider $H(\kappa, z)$ to be real holomorphic if it admits a Taylor expansion at each $\kappa \in [0, 1]$. In this case, $H(\kappa, z)$ can be extended by the Taylor series to complex values of κ in some neighborhood D of $[0, 1]$. This extended $H(\kappa, z)$ is then holomorphic for $\kappa \in D$. We note that in the present case the Taylor expansion is entailed by (5).

[22] See Ref. 9, p. 116.

We have now completed our discussion of the stability of the Faddeev equations. The final theorem gives more information concerning the eigenvalues $z(\kappa)$.

Theorem 9: If H is as in Theorem 5, then the eigenvalues $z(\kappa)$, obtained from (7):

(a) are at most countable, form a closed set, are contained within a finite interval on the real axis,

(b) have finite multiplicity, and

(c) have the same accumulation points as the eigenvalues z obtained from (2).

Proof: It is known that the eigenvalues z, obtained from (2), have the properties (a) and (b).[1,23] It follows from Theorem 5 and Lemma 4 that (a) holds for $z(\kappa)$. The validity of (b) for $z(\kappa)$ follows from (29) and (31). Finally, (c) follows from Theorems 4 and K.IV.5.35, and the theorem is proved.

We see from this theorem that H has the spectrum of a compact operator, as does A. It does not necessarily follow, however, that some power of H has a compact extension to the entire space, X_0, although A has this property.[24] However, if we take the point of view, implied by the discussion at the end of Sec. II, that (6) and (7) evaluated at some $\kappa_0 \in (0, 1]$ are legitimate Faddeev equations, i.e., that Faddeev's estimates apply to them, then Theorem 1 will obtain and consequently the H^m, $m \geq n$, have compact extensions to all of X_0. We note, however, that all our results are independent of this additional assumption.

[23] These results were proved under different conditions on the two-particle potentials than those imposed by Faddeev by W. Hunziker, Helv. Phys. Acta **39**, 451 (1966).

[24] In this case, the Hilbert space provides a less varied situation than other Banach spaces. It is known in the case of either a bounded normal operator in an arbitrary Hilbert space or a self-adjoint operator (not necessarily bounded) in $L^2(-\infty, \infty)$ that operators with spectra of compact operators are necessarily compact. Discussion of these results and references to the original papers can be found in Ref. 13, p. 611. *Note added in proof:* However, this property does not necessarily hold if the condition of normality is removed. See C. R. Putnam, *Commutation Properties of Hilbert Space Operators* (Springer-Verlag, Inc., New York, 1967), p. 48.

VI. CONCLUDING REMARKS

It has been shown that the hypotheses of Theorem 3 are sufficient to prove the stability properties of the Faddeev equations. These conditions provide criteria for the validity of some of the previous formal investigations.[7,8]

Our results could be extended in at least two ways. First, it would be interesting to see whether or not the two-particle potentials used in practical problems[7,8] lead to operators which satisfy our sufficient conditions. Secondly, necessary conditions for stability could be obtained by considering the consequences of the requirement that the spectral radius of the operator S, introduced in Theorem 3, be less than unity, or by considering some equivalent condition. Alternatively, an attempt could be made to prove the existence of $[I - H(\kappa, z)]^{-1}$ without using the series (13).

After this paper was completed, we received a preprint of Hepp's interesting work on the N-particle problem[25] which contains, among other things, some results concerning stability. In obtaining these results stronger conditions are imposed on the two-particle potentials than in the present work. These conditions require that the homogeneous form of the generalized Faddeev equations have no nontrivial solutions. On the other hand, his results hold for the N-particle problem, whereas we have been concerned only with the case $N = 3$. It appears that, with regard to the problem of stability of the Faddeev equations, the methods and results of the two approaches are complementary.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Dr. I. Manning and Dr. A. W. Sáenz for discussions concerning this work and to Dr. Sáenz also for a critical reading of the manuscript.

[25] K. Hepp, "On the Quantum Mechanical N-Body Problem," preprint, Eidgenössische Technische Hochschule, Zürich, Switzerland. We are grateful to Dr. Hepp for making this work avialable prior to publication.

Evaluation of "Kondo" Integrals

M. L. GLASSER
Battelle Memorial Institute, Columbus, Ohio

(Received 5 September 1968)

Exact and approximate expressions are given for the integrals

$$v_{nm} = \int_{-D}^{D} x^n(x^2 + \Delta^2)^{-m}(e^{\beta x} + 1)^{-1}\, dx.$$

The integrals

$$v_{nm} = \int_{-D}^{D} \frac{\omega^n}{(\omega^2 + \Delta^2)^m} f(\omega)\, d\omega, \tag{1}$$

where $f(\omega) = (e^{\beta\omega} + 1)^{-1}$ is the Fermi function ($\beta = 1/kT$), occur throughout the theory of the Kondo effect. Most recently, the case $m = 2$ has been discussed by Klein[1] who approximated them by taking the derivative of $f(\omega)$ as a pulse function. The purpose of this note is to point out that these integrals can be evaluated "exactly."

First of all, only the cases $m = 1$, $n = 0, 1$ need be treated in detail since the others may be obtained by differentiation with respect to Δ. For $D = \infty$, these integrals have been evaluated by Adawi and the author[2] in terms of the digamma function.

Next we note that

$$v_{nm} = \int_{0}^{D} \frac{\omega^n}{(\omega^2 + \Delta^2)^m} [f(\omega) + (-1)^n f(-\omega)]\, d\omega. \tag{2}$$

Since $f(\omega) + f(-\omega) = 1$, if n is even, v_{nm} is elementary. Thus, only the case v_{11} need be considered, and we have

$$v_{11} = -\int_{0}^{D} \frac{\omega}{\omega^2 + \Delta^2} \tanh\left(\frac{\beta\omega}{2}\right) d\omega. \tag{3}$$

For the case of low temperatures, which we define as $kT < \Delta/\pi$, Fermi integrals are easily evaluated by a variety of methods. For example,[3] to order $e^{-\beta D}$,

$$v_{11} = -\tfrac{1}{2} \ln\left[1 + \left(\frac{D}{\Delta}\right)^2\right] + \sum_{k=1}^{\infty} k^{-1} B_{2k}(2^{2k-1} - 1)\left(\frac{\pi}{\beta\Delta}\right)^{2k}, \tag{4}$$

where B_{2k} is a signed Bernoulli number. For $kT \ll \Delta/\pi$,

$$v_{11} \cong -\tfrac{1}{2} \ln\left[1 + \left(\frac{D}{\Delta}\right)^2\right] + \frac{\pi^2}{6\Delta^2}(kT)^2, \tag{5}$$

which agrees [to lowest order in (Δ/D)] with Nagaoka's estimate.[4]

For the high-temperature case $kT \geq \Delta/\pi$, exact expressions can be derived as follows. We first evaluate the integral

$$J = \int_{0}^{\infty} \frac{x}{x^2 + \Delta^2} \tanh \alpha x \cos xy\, dx. \tag{6}$$

Since the integrand is even,

$$J = \tfrac{1}{2}\, \mathrm{Re} \int_{-\infty}^{\infty} \frac{x \tanh \alpha x}{x^2 + \Delta^2} e^{ixy}\, dx \tag{7}$$

and we evaluate this by residues by closing the contour in the upper half-plane. $x(x^2 + \Delta^2)^{-1}$ has a simple pole at $x = i\Delta$ with residues $\frac{1}{2}$, $\tanh \alpha x$ has simple poles at $x = k\pi i/2\alpha$ with residue α^{-1}, $k = 1, 3, 5, \cdots$ (we assume $2\alpha\Delta \neq k\pi$ so there is no double pole), so we obtain

$$J = 2\pi^2 \sum_{k} \frac{k e^{-(k\pi/2\alpha)y}}{(k\pi)^2 - (2\alpha\Delta)^2} - \frac{\pi}{2} \tan(\alpha\Delta) e^{-\Delta y}. \tag{8}$$

Thus, by the Fourier inversion theorem,

$$\frac{x}{x^2 + \Delta^2} \tanh \alpha x = \int_{0}^{\infty} \left[4\pi \sum_{k} \frac{k e^{-(k\pi/2\alpha)y}}{(k\pi)^2 - (2\alpha\Delta)^2} - \tan(\alpha\Delta) e^{-\Delta y}\right] \cos xy\, dy. \tag{9}$$

We now integrate both sides of (9) from 0 to D (taking $\alpha = \frac{1}{2}\beta$) and then integrate with respect to y (or vice versa), all integrations being trivial. Thus we obtain

$$v_{11} = \tan(\tfrac{1}{2}\beta\Delta) \tan^{-1}(D/\Delta) - 4\pi \sum_{k\ \mathrm{odd}} \frac{k \tan^{-1}(\beta D/k\pi)}{(k\pi)^2 - (\beta\Delta)^2}. \tag{10}$$

This is valid for $kT \neq \Delta/\pi, \Delta/3\pi, \cdots$. For $kT \gg \Delta/\pi$, we find easily that

$$v_{11} \sim \tfrac{1}{2}\beta\Delta \tan^{-1}(D/\Delta) - \tfrac{1}{2}\beta D. \tag{11}$$

[1] A. P. Klein, Phys. Rev. **172**, 520 (1968).
[2] I. Adawi and M. L. Glasser, J. Appl. Phys. **37**, 364 (1966).
[3] M. L. Glasser, J. Math. Phys. **5**, 1150 (1964).
[4] Y. Nagaoka, Phys. Rev. **138**, A1112 (1965).

For the case $kT = \Delta/\pi$, a double pole occurs in the integrand in (7), but in this case the series which arises can be summed. We give only the result:

$$v_{11}(\beta\Delta = \pi) = \frac{2}{\pi\Delta^2}\left\{\frac{D\Delta}{D^2+\Delta^2} - \frac{1}{\Delta}\frac{\partial}{\partial D}\left[\frac{\Delta^2(D^2-\Delta^2)}{(D^2+\Delta^2)^2} - \frac{\beta D}{2}\frac{\Delta^2}{D^2+\Delta^2}\tanh(\tfrac{1}{2}\beta D)\right]\right\}. \quad (12)$$

This cannot, of course, be differentiated with respect to Δ to obtain other v_{nm}, but the same procedure can be applied to obtain all v_{nm} ($\Delta\beta = p\pi$) in similar closed form.

In conclusion, we discuss the integrals $v_n \equiv v_{n2}$ evaluated by Klein. We have, exactly,

$$v_0 = \frac{D}{2\Delta^2(D^2+\Delta^2)} + \frac{1}{2\Delta^3}\tan^{-1}(D/\Delta), \quad (13)$$

$$v_2 = \frac{1}{2\Delta}\tan^{-1}(D/\Delta) - \frac{D}{2(D^2+\Delta^2)}, \quad (14)$$

which agree to zero order in (Δ/D) with Klein's estimates. For v_1 we obtain

$$v_1 \cong -\frac{1}{2\Delta^2}\left[\frac{D}{D^2+\Delta^2} + \frac{\pi^2}{3\Delta^2}(kT)^2\right], \quad \text{for low temperature,}$$

$$\cong \frac{1}{4\Delta kT}\left[\frac{D\Delta}{D^2+\Delta^2} - \tan^{-1}(D/\Delta)\right], \quad \text{for high temperature.} \quad (15)$$

The leading term (to zero order in D/Δ) in each case agrees with Klein's estimate but, for example, the lowest temperature-dependent correction for small T is given incorrectly by a factor of 10. For v_{11}, Klein obtains (appendix, Ref. 1), to lowest order in (Δ/D),

$$v_{11} \sim \ln(\Delta/D) + (1/2\Delta^2)(kT)^2, \quad \text{for low temperature,}$$

$$\sim (\pi/2)(\beta\Delta) - \ln(\beta D) - 1, \quad \text{for high temperature.} \quad (16)$$

Our results show

$$v_{11} \sim \ln(\Delta/D) + (\pi^2/6\Delta^2)(kT)^2$$

$$\sim (\pi/4)(\beta\Delta) - \tfrac{1}{2}\beta D. \quad (17)$$

As a final example we consider the double integral which arises in evaluating the averaged value of a localized spin[5]

$$S = \int_{-D}^{D}\int_{-D}^{D}\left\{\frac{e^{\beta(\epsilon-\epsilon')}-1}{(\epsilon-\epsilon')^2} - \frac{\beta}{\epsilon-\epsilon'}\right\} \times f(\epsilon)[1-f(\epsilon')]\,d\epsilon\,d\epsilon'. \quad (18)$$

By means of the identity

$$u^{-2}(e^u - 1) - u^{-1} = \int_0^1 (1-t)e^{ut}\,dt$$

and the invariance of the resulting integrand under the substitution $t \to 1 - t$, (13) can be written as

$$S = \tfrac{1}{2}\int_0^1 dt\{F(\beta D)\}^2, \quad (19)$$

where

$$F(\beta D) = \int_0^{\beta D}\frac{e^{tx}+e^{(1-t)x}}{1+e^x}\,dx = \int_1^{\lambda}\frac{u^{t-1}+u^{-t}}{1+u}\,du, \quad \lambda = e^{\beta D}. \quad (20)$$

The integral in (15) can easily be approximated for large as well as small λ. For example, writing

$$\int_1^{\lambda} = \int_0^{\infty} - \int_0^1 - \int_{\lambda}^{\infty},$$

from a table of Mellin transforms we have

$$F(\beta D) = \pi\csc\pi t + (t-1)^{-1}\lambda^{t-1}{}_2F_1(1, 1-t; 2-t; -\lambda^{-1}) - t^{-1}\lambda^{-t}{}_2F_1(1, t; 1+t; -\lambda^{-1}). \quad (21)$$

Thus, for $kT \ll D$,

$$S \cong (2D\ln 2)kT + 2\ln(D/kT) + \text{const}$$

and, for $kT \gg D$,

$$S \cong \tfrac{1}{2}(D/kT)^2 + O[(D/kT)^4].$$

Approximating the Fermi function by a trapezoid again gives the leading term correctly, but in each case the first correction is given incorrectly in sign and magnitude.

[5] M. L. Glasser and J. I. Kaplan, Nuovo Cimento (to be published).

Magnetization of Ising Model in Nonzero Magnetic Field*

DANIEL C. MATTIS AND MICHAEL PLISCHKE
Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

(Received 10 February 1969)

Knowing only the zero-field magnetization (e.g., Yang's result) of the Ising model in any number of dimensions, one can construct a lower bound on $m(h)$, the magnetization in finite field. Knowledge of u, the internal energy per bond, enables a more efficient lower bound to be constructed. Both are applications of the Griffiths inequality, as recently generalized by Kelly and Sherman, and should prove useful in the lattice gas problem where it is essential to know $m(h)$.

We present nontrivial *lower* bounds on the magnetization in finite magnetic field $|m(h)|$ of the Ising model. An "upper upper" bound is $|m(h)| = 1$. Subsequently we hope to derive an improved upper bound which, together with the present result, should help constrain the true $m(h)$ fairly well.

Consider an isotropic $M \times N$ lattice, with a spin at every site and periodic boundary conditions, and a Hamiltonian

$$H(h) = -J\sum_{(ij)}\sigma_i\sigma_j - h\sum_i \sigma_i. \tag{1}$$

The partition function $Z(M, N, h, \beta)$ is:

$$\begin{aligned} Z(M, N, h, \beta) &\equiv \mathrm{Tr}\,\{e^{-\beta H}\} \\ &= Z(M, N, h', \beta)\langle e^{\beta(h-h')\Sigma_i\sigma_i}\rangle_{h'}, \end{aligned} \tag{2}$$

where $\langle\ \rangle_{h'}$ indicates "thermodynamic average w.r.t. $H(h')$." Expanding:

$$\langle e^{\beta(h-h')\Sigma_i\sigma_i}\rangle_{h'} = \cosh^{MN}\beta(h-h')\left\langle\prod_1^{MN}(1+\sigma_i t)\right\rangle_{h'}, \tag{3}$$

in which $t \equiv \tanh\beta(h-h')$.

We factor the product into pairs and apply the generalized Griffiths inequality[1] due to Kelly and Sherman.[2]

$$\left\langle\prod_1^{MN}(1+\sigma_i t)\right\rangle_{h'} \geq \prod_1^{\frac{1}{2}MN}\langle[1 + t(\sigma_i + \sigma_{i+1}) + t^2\sigma_i\sigma_{i+1}]\rangle_{h'}. \tag{4}$$

By translational invariance, all factors are equal, and the rhs of (4) is

$$(1 + 2tm(h') + t^2|u(h')|)^{\frac{1}{2}MN}. \tag{5}$$

In the limit $MN \to \infty$ followed by $h' \to 0$, Griffiths has shown that m is positive and obeys

$$\lim_{h'\to 0}\lim_{MN\to\infty} m(h') \geq m_Y, \tag{6}$$

where m_Y is the (positive) magnetization calculated by Yang.[3] Similarly, the limit:

$$\lim_{h'\to 0}\lim_{MN\to\infty}|u(h')| = |u| \tag{7}$$

is the zero-field short-range correlation function—i.e., the absolute value of the internal energy per bond. Thus,

$$\begin{aligned} Z(M, N, h, \beta) \geq Z(M, N, 0, \beta)&(\cosh^{MN}\beta h) \\ &\times (1 + 2tm_Y + t^2|u|)^{\frac{1}{2}MN}. \end{aligned} \tag{8}$$

On the lhs, we have

$$Z(M, N, h, \beta) \equiv \exp\left[MN\beta\int_0^h dh''m(h'')\right]Z(M, N, 0, \beta). \tag{9}$$

Because $m(h'')$ is a nondecreasing function of its argument,

$$hm(h) \geq \int_0^h dh''m(h''). \tag{10}$$

Combine (10) and (8) to obtain

$$\begin{aligned} m(h) \geq (h\beta)^{-1}&\{\log\cosh\beta h \\ &+ \tfrac{1}{2}\log(1 + 2tm_Y + t^2|u|)\}. \end{aligned} \tag{11}$$

We illustrate this result in Fig. 1, plotting the rhs of (11) for one temperature above T_c (curve A), one at T_c (B), and two below T_c (C and D).

A lower bound, which is somewhat less efficient above T_c but almost as good as (11) below it, can be obtained with far less numerical work; according to Refs. (1) and (2), $|u| \geq m_Y^2$, therefore using this on the rhs of (11) we find

$$m(h) \geq (\beta h)^{-1}\log(\cosh\beta h + m_Y\sinh\beta h). \tag{12}$$

Above or at T_c, $m_Y = 0$, and the resultant lower bound is shown as the dotted curve in Fig. 1. Below

* This research supported by the United States Air Force, AFOSR grant No. 69-1642.

[1] R. B. Griffiths, Phys. Rev. **152**, 240 (1966).

[2] D. G. Kelly and S. Sherman, J. Math. Phys. **9**, 466 (1968).

[3] C. N. Yang, Phys. Rev. **85**, 808 (1952).

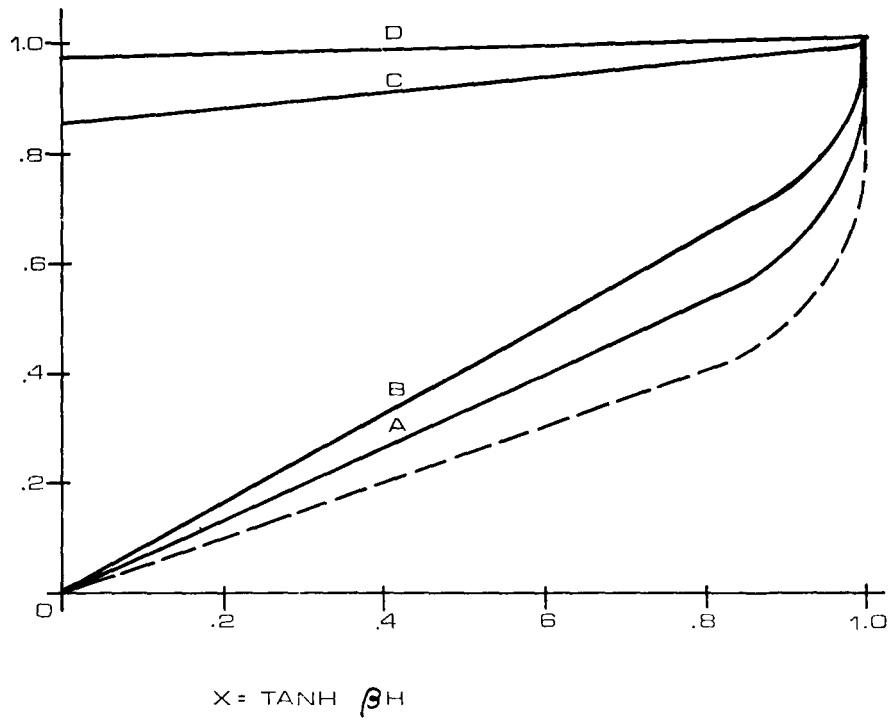


FIG. 1. Lower bounds to the magnetization at finite field $m(h)$ plotted vs $x = \tan \beta h$ at various temperatures. For A, $T/T_c = 1.83$; for B, $T = T_c$; for C, $T/T_c = 0.927$; and for D, $T/T_c = 0.61$; all using inequality (11). Dotted curve is inequality (12) at all $T \geq T_c$.

T_c, the lower bound (12) rapidly approaches (11) and would be indistinguishable from curves C and D at the temperatures we have chosen, on the scale of our graph.

It is hoped that the present results might be useful in lattice gas theory as well as in magnetism. It should be noted that they are not at all restricted to two dimensions; once a variational estimate of m in zero field is known for three dimensions, it can be used forthwith in Eq. (12).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 6 JUNE 1969

Group Representation in a Continuous Basis: An Example*

C. ITZYKSON†
Stanford Linear Accelerator Center, Stanford University, Stanford, California

(Received 2 February 1968)

Given an irreducible unitary representation of a noncompact group, what happens if one tries to diagonalize one of the noncompact generators? We study some aspects of this question on an example, chosen to be a representation of the discrete series with $j = -\frac{1}{2}$ of the special real linear group in two dimensions.

I. INTRODUCTION

Classical Fourier analysis is the standard example of diagonalization of a noncompact generator (in this case, the generator of translations along the real line). Some interesting properties arise when such an Abelian noncompact group is imbedded in a larger structure. This occurs, for instance, when one studies the group $G \equiv SL(2R)$, of two-by-two real unimodular matrices. Let us first recall elementary properties of this group that will be used. An arbitrary element $g \in G$ is of the form $z = \binom{ab}{cd}$, a, b, c, d real, $ad - bc = 1$. The Lie algebra of this simple group is realized as traceless real two-by-two matrices, a basis of which is

$$t_1 = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \quad t_2 = \frac{1}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad r = \frac{1}{2}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \tag{1}$$

satisfying the commutation rules

$$[r, t_1] = t_2, \quad [r, t_2] = -t_1, \quad [t_1, t_2] = -r. \tag{2}$$

These generators are such that, given a unitary representation of G, their representatives are skew-adjoint operators. Let us assume that we are given such a representation, and let us denote by T_1, T_2, and R these representatives. Since R generates a compact subgroup, its spectrum, though unbounded, is discrete. From the commutation rule

$$[R, (T_1 \pm iT_2)] = \pm i(T_1 \pm iT_2),$$

one sees that $T_1 \pm iT_2$ play the role of raising and lowering operators. On the other hand, suppose we diagonalize T_1. Its spectrum will be continuous of the form $i\lambda$ (λ real). The commutation rule

$$[T_1, (R \mp T_2)] = \pm(R \mp T_2)$$

seems to indicate that acting with $R \mp T_2$ on some "improper states" $|\lambda\rangle$, corresponding to the spectral value λ of $-iT_1$, will lead to the "improper state" $|\lambda \mp i\rangle$. We intend to discuss more precisely this question.

To do this, we shall specifically study one irreducible representation of G which we choose to be one of the discrete series.[1] In Sec. II, we describe this representation following Ref. 2. Section III is devoted to the diagonalization of T_1 through a Mellin transform. It turns out that we are naturally led to study some properties of a set of orthogonal polynomials, of a type introduced by Pollaczek.[3] Finally, in Sec. IV, we consider the representation of G in this new basis.[4]

It will be understood in the following that when a real positive number x is taken to complex power y, $\arg x = 0$. The complex conjugate of z will be denoted z^*.

II. A REPRESENTATION OF THE DISCRETE SERIES

As in Ref. 2, let us consider the vector space $\mathfrak{D}$ of analytic functions such that if $f \in \mathfrak{D}$:

(1) $f(z)$ is analytic for $\operatorname{Im} z > 0$, and continuous with all its derivatives in $\operatorname{Im} z \geq 0$;

(2) $\hat{f}(z) = (1/z)f(-(1/z))$ is also continuous with all its derivatives in $\operatorname{Im} z \geq 0$.

As a result of (1) and (2), one can define a norm on $\mathfrak{D}$ through

$$\|f\|^2 = \frac{1}{\pi}\int_{-\infty}^{+\infty} dx\, |f(x)|^2. \tag{3}$$

Equipped with this norm, $\mathfrak{D}$ is not complete. Its completion is a Hilbert space $\mathfrak{K}$ of analytic functions in the upper half-plane. Indeed, if $f \in \mathfrak{D}$, its value at a

* Work supported by the U.S. Atomic Energy Commission.

† Permanent address: Service de Physique Theorique, CEN Saclay, BP No. 2, 91, Gif sur Yvette, France.

[1] V. Bargmann, Ann. Math. **48**, 569 (1947).

[2] I. M. Gel'fand, M. I. Graev, and N. Y. Vilenkin, *Generalized Functions, Vol. 5* (Academic Press Inc., New York, 1966), Chap. VII, Sec. 5.

[3] G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, Providence, R.I., 1959).

[4] While completing this paper, we received a preprint from A. O. Barut and E. C. Phillips ("Matrix elements of representations of noncompact groups in a continuous basis," University of Colorado, 1967), which deals with a similar subject.

given point $z(\mathrm{Im}\, z > 0)$ is such that

$$|f(z)| = \frac{1}{2\pi}\left|\int_{-\infty}^{+\infty} f(x)\frac{1}{x-z}\,dx\right| \leq \frac{1}{2\sqrt{\pi}}\|f\|\left[\int_{-\infty}^{+\infty}\frac{dx}{x^2+\mathrm{Im}\, z^2}\right]^{\frac{1}{2}}$$

or

$$\mathrm{Im}\, z > 0: |f(z)| \leq \|f\|/2(\mathrm{Im}\, z)^{\frac{1}{2}}. \tag{4}$$

Equation (4) shows that a Cauchy sequence in $\mathfrak{D}$, with respect to the norm (3), will converge uniformly, in the ordinary sense, on any compact set of the upper half z-plane, to an analytic function. $\mathfrak{D}$ is dense in $\mathfrak{K}$, but it is clear that it is not all of $\mathfrak{K}$ as shown by the example of $[1/(z+i)]\log(1-iz)/2$,which belongs to $\mathfrak{K}$ but not to $\mathfrak{D}$; it is, however, the limit of

$$\sum_1^N \frac{1}{p}(z-i)^p(z+i)^{-p-1},$$

which belongs to $\mathfrak{D}$.

For an alternative description of $\mathfrak{K}$, we introduce the functions $f_n(z) \in \mathfrak{D}$:

$$f_n(z) = (z-i)^n(z+i)^{-n-1}, \quad n = 0, 1, \cdots,$$
$$(f_n\,|\,f_m) = \delta_{nm}. \tag{5}$$

Let us show that the system $\{f_n\}$ is complete. It is sufficient to prove that if $f \in \mathfrak{D}$ and $(f_n\,|\,f) = 0$ for all n, then $f = 0$. Indeed, an explicit computation leads to $(d/dz)^n(z+i)^n f(z)|_{z=i} = 0$, for $n = 0, 1, \cdots$. By recurrence, all derivatives of f vanish at $z = i$ and since f is analytic, $f = 0$. As a consequence, the elements of $\mathfrak{K}$ are characterized by sequences of complex numbers $\{a_n\}$, n nonnegative integer such that $\sum |a_n|^2 < \infty$; the analytic function itself is obtained as $\sum_0^\infty a_n f_n(z)$. This series converges uniformly in any compact domain of the upper half-plane, since such a domain can be enclosed in a circle $|(z-i)/(z+1)| \leq \rho < 1$, where one has

$$\left|\sum_0^N a_n f_n(z)\right| \leq \left[\sum_0^\infty |a_n|^2/(1-\rho^2)\right]^{\frac{1}{2}}.$$

Let $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv g \in G$ and $f \in \mathfrak{D}$; the set of transformations

$$f \to U(g)f, \quad U(g)f(z) = \frac{1}{bz+d} f\left(\frac{az+c}{bz+d}\right) \tag{6}$$

leave $\mathfrak{D}$ invariant and can be extended to a unitary representation of G in $\mathfrak{K}$. This representation belongs to the discrete series[1]; it is irreducible and will be studied in the following. From the global form (6), we can derive the representatives of the generators T_1, T_2, and R defined in the introduction. They are the differential operators

$$T_1 = \frac{1}{2}\left(1 + 2z\frac{d}{dz}\right),$$
$$T_2 = -\frac{1}{2}\left[z + (z^2-1)\frac{d}{dz}\right], \tag{7}$$
$$R = \frac{1}{2}\left[z + (z^2+1)\frac{d}{dz}\right].$$

The complete set $\{f_n\}$ satisfies

$$Rf_n = i(n+\tfrac{1}{2})f_n,$$
$$(T_1 + iT_2)f_n = nf_{n-1}, \tag{8}$$
$$(T_1 - iT_2)f_n = -(n+1)f_{n+1}.$$

In other words, in this basis R is diagonal with eigenvalues of the form $i(n+\frac{1}{2})$, where n is a nonnegative integer. The Casimir operator $T_1^2 + T_2^2 - R^2$ takes in this representation the value $-\frac{1}{4}$. If this is written as $j(j+1)$, it corresponds to a value of j equal to $-\frac{1}{2}$.

III. DIAGONALIZATION OF A NONCOMPACT GENERATOR

Our aim is now to diagonalize a noncompact generator, T_1, say. An eigenfunction of the corresponding differential operator (7) is a homogeneous function z^a. For no value of the exponent does such a function belong to $\mathfrak{K}$. This is to be expected: T_1 has no eigenvalue (in the sense that they would correspond to normalizable eigenstates) but we expect its spectrum to be purely imaginary, or $(izI - T_1)^{-1}$ to exist as a bounded operator for $\mathrm{Im}\, z \neq 0$.

We shall obtain this diagonal form by studying the following Mellin transform. Let $f \in \mathfrak{K}$; we introduce the function of the real variable λ, $F(\lambda)$ by

$$f \to F, \quad F(\lambda) = i\frac{\cosh \pi\lambda}{\pi}\int_0^\infty d\rho f(i\rho)\rho^{-\frac{1}{2}-i\lambda}. \tag{9}$$

It is clear that the integral converges in the ordinary sense for $f \in \mathfrak{D}$. We shall extend it with the help of the transforms $\{F_n\}$ of the basic functions $\{f_n\}$ introduced in Sec. II:

$$f_n \to F_n,$$
$$F_n(\lambda) = \frac{\cosh \pi\lambda}{\pi}\int_0^\infty d\rho(\rho-1)^n(\rho+1)^{-n-1}\rho^{-\frac{1}{2}-i\lambda}. \tag{10}$$

A convenient way of performing the integral (10) is to observe that the series

$$\sum_0^\infty t^n(\rho-1)^n(\rho+1)^{-n-1} = \frac{1}{(\rho+1)-t(\rho-1)}, \quad |t| < 1, \quad 0 \leq \rho,$$

converges absolutely and uniformly in

$$[|t| \leq 1 - \epsilon] \times [0 \leq \rho < \infty].$$

Hence, we obtain for the set $\{F_n\}$ the generating function:

$$|t| < 1, \sum_0^\infty t^n F_n(\lambda) = \frac{\cosh \pi\lambda}{\pi} \int_0^\infty d\rho \frac{\rho^{-\frac{1}{2}-i\lambda}}{\rho(1-t)+(1+t)} = (1+t)^{-\frac{1}{2}-i\lambda}(1-t)^{-\frac{1}{2}+i\lambda}, \quad (11)$$

where we have used the fact that

$$F_0(\lambda) = \frac{\cosh \pi\lambda}{\pi} \int_0^\infty d\rho \frac{\rho^{-\frac{1}{2}-i\lambda}}{\rho+1} = \lim_{\epsilon \to +0} \frac{e^{-\pi\lambda}}{2\pi} \left[\int_{-\infty-i\epsilon}^{-i\epsilon} + \int_{i\epsilon}^{\infty} \right] \frac{d\rho}{\rho+1} \rho^{-\frac{1}{2}-i\lambda} = 1$$

by Cauchy's theorem.

In formula (11), the phases of $(1+t)$ and $(1-t)$ are zero for $-1 < t < +1$.

We summarize elementary properties of the functions $F_n(\lambda)$ in the following:

Proposition 1:

(a) $F_n(\lambda)$ is polynomial in λ of precise degree n and $F_n(i/2) = 1$;

(b) $$(-)^n F_n^*(\lambda^*) = (-1)^n F_n(-\lambda) = F_n(\lambda) = F(-n, \tfrac{1}{2} + i\lambda; 1; 2); \quad (12)$$

(c) for $|t| < 1$,

$$\sum_0^\infty t^n F_n(\lambda) = (1+t)^{-\frac{1}{2}-i\lambda}(1-t)^{-\frac{1}{2}+i\lambda};$$

(d) $$\int_{-\infty}^{+\infty} F_m(\lambda)^* F_n(\lambda) \frac{d\lambda}{\cosh \pi\lambda} = \delta_{nm}; \quad (13)$$

(e) $$2i\lambda F_n(\lambda) = nF_{n-1}(\lambda) - (n+1)F_{n+1}(\lambda).$$

Proof: Proposition 1 asserts that the $F_n(\lambda)$ forms an orthonormal set of polynomials in the Hilbert space of functions $F(\lambda)$ such that

$$\|F\|^2 = \int_{-\infty}^{+\infty} \frac{d\lambda}{\cosh \pi\lambda} |F(\lambda)|^2 < \infty.$$

We have already proved (c), from which (a) and (b) easily follow. Indeed, $F_n(\lambda)$ appears equal to the polynomial of degree n:

$$F_n(\lambda) = \sum_0^n \frac{(-\frac{1}{2} - i\lambda)_{n-p}(-\frac{1}{2} + i\lambda)_p(-)^p}{(n-p)!\, p!}, \quad (14)$$

where $(x)_p = \Gamma(x+p)/\Gamma(x)$. Hence the coefficient of λ^n in $F_n(\lambda)$ is

$$(-i)^n \sum_0^n \frac{1}{p!\, n-p!} = \frac{(-2i)^n}{n!} \neq 0.$$

From the integral representation (10), we obtain the expression (12) of F_n in terms of the hypergeometric function from which the value $F_n(i/2) = 1$ follows.

To establish the orthogonality relation, we make use again of the generating function. From the equality ($|\mathrm{Im}\, x| < \pi/2$):

$$\int_{-\infty}^{+\infty} e^{2i\lambda x} \frac{d\lambda}{\cosh \pi\lambda} = \frac{1}{\cosh x},$$

for $-1 < u, t < 1$ we derive

$$\sum_{m,n=0}^\infty u^m t^n \int_{-\infty}^{+\infty} F_m(\lambda)^* F_n(\lambda) \frac{d\lambda}{\cosh \pi\lambda} = \frac{1}{[(1-u^2)(1-t^2)]^{\frac{1}{2}}} \int_{-\infty}^{+\infty} \left[\frac{(1-t)(1+u)}{(1+t)(1-u)}\right]^{i\lambda} \frac{d\lambda}{\cosh \pi\lambda} = \frac{1}{1-tu} = \sum_{m,n=\infty}^\infty \delta_{mn} u^m t^n.$$

Since we are dealing with an analytic function of t and u in $|t| < 1$, $|u| < 1$, we can identify the coefficients of its Taylor expansion and thus arrive at the desired orthogonality property.

Finally, the relation (*e*) is an immediate consequence of the representation of the Lie algebra of G. Indeed, from Eqs. (6) and (8), we have

$$\tfrac{1}{2}[nF_{n-1}(\lambda) - (n+1)F_{n+1}(\lambda)] = \frac{\cosh \pi\lambda}{\pi} \int_{-\infty}^{+\infty} d\rho \rho^{-\frac{1}{2}-i\lambda} \frac{d}{d\alpha} [e^{\alpha T} f_n(i\rho)]_{\alpha=0}$$
$$= \frac{\cosh \pi\lambda}{\pi} \int_{-\infty}^{+\infty} d\rho \rho^{-\frac{1}{2}-i\lambda} \frac{d}{d\alpha} [e^{\alpha/2} f_n(i\rho e^\alpha)]_{\alpha=0}.$$

In the last integral, interchange of the order of integration and differentiation is allowed. As a result:

$$\tfrac{1}{2}[nF_{n-1}(\lambda) - (n+1)F_{n+1}(\lambda)] = \frac{d}{d\alpha} [e^{i\alpha\lambda} F_n(\lambda)]_{\alpha=0} = i\lambda F_n(\lambda).$$

The polynomials F_n belong to a class which has been studied by Pollaczek.[3] We denote by H the Hilbert space of square-integrable functions on the real line with measure $d\lambda/\cosh \pi\lambda$. As usual, two functions which differ on a set of measure zero are identified.

Proposition 2: The polynomials F_n form a complete orthonormal basis in H.

Proof: In view of Proposition 1, it is sufficient to prove that the functions λ^n, n a nonnegative integer, form a complete set in H. Let $F \in H$ be orthogonal to

all λ^n. Consider the function

$$g(s) = \int_{-\infty}^{+\infty} \frac{d\lambda}{\cosh \pi\lambda} e^{is\lambda} F(\lambda).$$

It is analytic in the strip $|\mathrm{Im}\,\lambda| < \pi/2$ and all its derivatives vanish at the origin. As a result, $g(s) = 0$ and $F(\lambda)$ vanishes almost everywhere. This proves that the system $\{F_n\}$ is complete.

We can recover f_n from F_n through an inverse Mellin transform:

$$f_n(z) = \frac{1}{2i} \int_{-\infty}^{+\infty} F_n(\lambda)(-iz)^{-\frac{1}{2}+i\lambda} \times \frac{d\lambda}{\cosh \pi\lambda}, \quad -\frac{\pi}{2} < \arg(-iz) < \frac{\pi}{2}. \quad (15)$$

With these results, we can return to the integral transformation (9). First let $f \in \mathfrak{D}$; then the integral in (9) is absolutely convergent. Moreover, one readily shows that for real λ, $|F(\lambda)|\, e^{-\pi|\lambda|/2} \to 0$ as $|\lambda| \to \infty$ faster than any power of $|\lambda|$ and hence $F \in H$. Using (15), one finds that, if $f = \sum_0^\infty a_n f_n$, then $F = \sum_0^\infty a_n F_n$. In other words, the Mellin transform M is an isometric mapping from $\mathfrak{D} \subset \mathcal{H}$ in a dense subset of H which will be denoted D. By continuity, it is then uniquely extended to a one-to-one isometric mapping M from $\mathcal{H}$ to H.

We close this section by mentioning some properties of D. Let $f^{(n)}$ stand for the nth derivative of $f \in \mathfrak{D}$, and $\hat{f}(z) = z^{-1} f(-z^{-1})$; then:

Proposition 3: Any $F = Mf \in D$ can be extended as an entire function in the complex λ plane. Moreover,

$$F[-i(\tfrac{1}{2} + n)] = i^{(1-n)} \frac{f^{(n)}}{n!}(0),$$

$$F[i(\tfrac{1}{2} + n)] = i^{(2-n)} \frac{\hat{f}^{(n)}}{n!}(0), \quad n = 0, 1, \cdots, \quad (16)$$

and

$$\hat{F}(\lambda) = M\hat{f}(\lambda) = -iF(-\lambda).$$

Proof: Note that one can write

$$F(\lambda) = \phi_1(\lambda) + i\phi_2(\lambda),$$

where

$$\phi_1(\lambda) = i \frac{\cosh \pi\lambda}{\pi} \int_0^1 f(i\rho) \rho^{-\frac{1}{2}-i\lambda},$$

$$\phi_2(\lambda) = i \frac{\cosh \pi\lambda}{\pi} \int_0^1 \hat{f}(i\rho) \rho^{-\frac{1}{2}+i\lambda},$$

ϕ_2 is deduced from ϕ_1 by changing $\lambda \to -\lambda$ and $f \to \hat{f}$. Thus it is sufficient to consider ϕ_1. At first it is analytic for $\mathrm{Im}\,\lambda > -\frac{1}{2}$, vanishing at the points $i[\frac{1}{2} + n]$, for nonnegative integer n. Furthermore, integration by parts gives, for an arbitrary positive integer p and $\mathrm{Im}\,\lambda > -\frac{1}{2}$,

$$\phi_1(\lambda) = \frac{i \cosh \pi\lambda}{(\frac{1}{2} - i\lambda)_p \pi} \times \left\{ \sum_{r=1}^{p} (-)^{r-1} \left(\frac{d}{d\rho}\right)^{r-1} f(i\rho) \frac{d^{p-r}}{d\rho^{p-r}} \rho^{-\frac{1}{2}-i\lambda+p} \Big|_{\rho=1} + (-)^p \int_0^1 d\rho \left[\frac{d^p}{d\rho^p} f(i\rho)\right] \rho^{-\frac{1}{2}-\lambda+p} \right\}.$$

The zeros of the $(\frac{1}{2} - i\lambda)_p$ are just cancelled by those of $\cosh \pi\lambda$. We can then analytically continue this formula to $\mathrm{Im}\,\lambda > -\frac{1}{2} - p$. Since p is arbitrary, $\phi_1(\lambda)$ is an entire function of λ. If we set $p = n + 1$ and $\lambda = -i(\frac{1}{2} + n)$, n a nonnegative integer, in the above expression, we get

$$\phi_1[-i(\tfrac{1}{2} + n)] = \frac{i(-1)^n}{n!} \frac{d^n}{d\rho} f(i\rho)\Big|_{\rho=0} = \frac{i^{(1-n)}}{n!} f^{(n)}(0).$$

Combining these results with similar ones for ϕ_2, we arrive at formula (16).

IV. REPRESENTATION OF G

The isometric operator M of the preceding section enables one to carry the representation U of G, defined in $\mathcal{H}$, to an equivalent representation V, defined in H through $V(g) = MU(g)M^{-1}$. The inverse transformation V^{-1} was already indicated in (15). Hence, for $V(g)$, we obtain the following expression:

$$F \to V(g)F,$$

$$V(g)F(\lambda) = \frac{\cosh \pi\lambda}{2\pi} \int_0^\infty d\rho \rho^{-\frac{1}{2}-i\lambda} \int_{-\infty}^{+\infty} (ib\rho + d)^{-\frac{1}{2}-i\mu} \times (a\rho - ic)^{-\frac{1}{2}+i\mu} F(\mu) \frac{d\mu}{\cosh \pi\mu},$$

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with

$$\rho > 0, \quad -\frac{\pi}{2} < \arg\left(\frac{a\rho - ic}{ib\rho + d}\right) < \frac{\pi}{2}. \quad (17)$$

At first, this formula is defined when $F \in D$. It is then extended by the unitarity property to all H. Assume $F \in D$ and none of the real numbers a, b, c, d to vanish. The interchange of the order of integration is allowed in (17). Let us, therefore, compute the kernel

$$K_g(\lambda, \mu) = \frac{\cosh \pi\lambda}{2\pi} \int_0^\infty d\rho \rho^{-\frac{1}{2}-i\lambda} (ib\rho + d)^{-\frac{1}{2}-i\mu} \times (a\rho - ic)^{-\frac{1}{2}+i\mu}. \quad (18)$$

Let G_1 denote the subgroup of elements of the form

$$g(\alpha) = e^{\alpha t_1} = \begin{pmatrix} e^{\alpha/2} & 0 \\ 0 & e^{-\alpha/2} \end{pmatrix}.$$

The manifold S of elements g in G, such that $abcd \neq 0$, is invariant under right and left translation by G_1. Moreover,

$$K_{g(\alpha)gg(\beta)}(\lambda, \mu) = e^{i\lambda\alpha} K_g(\lambda\mu) e^{i\mu\beta}, \tag{19}$$

which enables one to compute K only for representatives of each type of double coset $G_1 \backslash G / G_1$.

These fall into four classes. We select representatives of the form

$$\begin{aligned} (A) &= e^{2At_2} = \begin{pmatrix} \cosh A & \sinh A \\ \sinh A & \cosh A \end{pmatrix}, \\ (B) &= e^{2Bt_1} = \begin{pmatrix} \cos B & -\sin B \\ \sin B & \cos B \end{pmatrix}, \\ (C) &= \begin{pmatrix} \sinh C & \cosh C \\ -\cosh C & -\sinh C \end{pmatrix}, \\ (D) &= \begin{pmatrix} \sinh D & -\cosh D \\ \cosh D & -\sinh D \end{pmatrix}, \end{aligned} \tag{20}$$

where the parameters A, B, C, and D are all different from zero. The last two classes are taken into account by remarking that they can be obtained from the first one by left multiplication by $g_0 = \left(\begin{smallmatrix} 0 & +1 \\ -1 & 0 \end{smallmatrix}\right)$, or right multiplication by g_0^{-1}, and that one has

$$\begin{aligned} K_{g_0 g}(\lambda, \mu) &= -iK_g(-\lambda, \mu), \\ K_{gg_0^{-1}}(\lambda, \mu) &= iK_g(\lambda, -\mu), \end{aligned} \tag{21}$$

a fact which is readily related to the properties of the mapping $F \to \hat{F}$. Let us, therefore, compute

$$K_{(A)} = K_{e^{2At_2}} \quad \text{and} \quad K_{(B)} = K_{e^{2Bt_1}}.$$

We present in some detail the calculation for case (A) with A positive. We have:

$$\begin{aligned} K_{(A)}(\lambda, \mu) = \frac{\cosh \pi\lambda}{2\pi} \int_0^\infty d\rho \rho^{-\frac{1}{2}-i\lambda} \\ \times (i\rho \sinh A + \cosh A)^{-\frac{1}{2}-i\mu} \\ \times (\rho \cosh A - i \sinh A)^{-\frac{1}{2}+i\mu}. \end{aligned}$$

We define a single-valued integrand by performing cuts in the ρ plane from 0 to $+\infty$ and from $i \tanh A$ to $i \coth A$ along the imaginary axis. The branch of the function is characterized by Eq. (17), which takes the following form for ρ slightly above the real positive axis:

$$\begin{aligned} \arg \rho = 0, \quad -\pi/2 < \arg (\rho - i \tanh A) \\ - \arg (i\rho \tanh A + 1) < \pi/2. \end{aligned}$$

With these conventions we can replace the original integration in ρ from zero to infinity by a line integral encircling the point zero clockwise, starting at $\rho = +\infty - i\epsilon$ and ending at $\rho = +\infty + i\epsilon$, provided we replace the factor $\cosh \pi\lambda$ by $\frac{1}{2}e^{-\pi\lambda}$. Observing that we can close the contour at infinity, we are left with an integral along a contour C which encloses counterclockwise the cut from $i \tanh A$ to $i \coth A$. This is equal to the integral from $i \tanh A$ to $i \coth A$ of the discontinuity of the integrand since the end points do not give any contribution. Collecting all the factors,

$$\begin{aligned} K_{(A)}(\lambda, \mu) = \cosh A^{-\frac{1}{2}+i\mu} \sinh A^{-\frac{1}{2}-i\mu} \frac{\cosh \pi\mu}{2\pi} e^{-\pi(\lambda-\mu)} \\ \times \int_{\tanh A}^{\coth A} dx x^{-\frac{1}{2}-i\lambda} \\ \times (\coth A - x)^{-\frac{1}{2}-i\mu} (x - \tanh A)^{-\frac{1}{2}+i\mu}, \end{aligned}$$

which after the change of variable

$$x = \tanh A + \frac{t}{\sinh A \cosh A}$$

yields, for $A > 0$,

$$\begin{aligned} K_{(A)}(\lambda, \mu) = \tfrac{1}{2} e^{-\pi(\lambda-\mu)} \cosh A^{i(\lambda+\mu)} \sinh A^{-1-i(\lambda+\mu)} \\ \times F(\tfrac{1}{2} + i\lambda, \tfrac{1}{2} + i\mu, 1; -1/\sinh^2 A). \end{aligned}$$

In obtaining this last expression we have made use of the classical representation of the hypergeometric function:

$$\begin{aligned} F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)} \\ \times \int_0^1 dt (1 - zt)^{-a} t^{b-1} (1 - t)^{c-b-1}. \end{aligned}$$

Proceeding along the same lines we obtain, for an arbitrary $A \neq 0$,

$$\begin{aligned} K_{(A)}(\lambda, \mu) = \tfrac{1}{2} e^{-\epsilon\pi(\lambda-\mu)/2} [\cosh A]^{i(\lambda+\mu)} [|\sinh A|]^{-1-i(\lambda+\mu)} \\ \times F(\tfrac{1}{2} + i\lambda, \tfrac{1}{2} + i\mu; 1; -1/\sinh^2 A), \\ \epsilon = A/|A|. \end{aligned} \tag{22}$$

For Class (B), we make again use of relation (21), which allows one to restrict B to $0 < B < \pi/2$, and obtain

$$\begin{aligned} K_{(B)}(\lambda, \mu) \\ = \tfrac{1}{2} e^{\pi(\mu-\lambda)/2} \frac{\Gamma(\frac{1}{2} + i\lambda)}{\Gamma(\frac{1}{2} - i\mu)\Gamma(1 + i(\lambda + \mu))} \\ \times (\sin B)^{i(\mu-\lambda)} (\cos B)^{i(\lambda+\mu)} \\ \times F(\tfrac{1}{2} + i\mu, \tfrac{1}{2} + i\mu; 1 + i(\lambda + \mu); \cos^2 B) \\ + i\tfrac{1}{2} e^{\pi(\lambda+\mu)/2} \frac{\Gamma(\frac{1}{2} + i\lambda)}{\Gamma(\frac{1}{2} + i\mu)\Gamma(1 + i(\lambda - \mu))} \\ \times (\sin B)^{i(\lambda-\mu)} (\cos B)^{-i(\lambda+\mu)} \\ \times F(\tfrac{1}{2} - i\mu, \tfrac{1}{2} - i\mu, 1 + i(\lambda + \mu), \sin^2 B). \end{aligned} \tag{23}$$

Expression (22) can be brought to a form similar to (23) using transformation properties of the hypergeometric function.

On the manifold G–S (which contains the subgroup G_1), the kernel is singular. Of particular interest is the representation of the subgroup G_1. It follows from (17) that

$$V(e^{\alpha t_1})F(\lambda) = e^{i\alpha\lambda}F(\lambda). \tag{24}$$

In other words, in this basis the representation of this subgroup is diagonal. Our calculation of the kernel K_g is not very well suited to obtain the other generators, but they can be readily recovered using, for instance, the Pollaczek polynomials of the preceding section. Indeed, we have

$$T_1F_n(\lambda) = \tfrac{1}{2}[nF_{n-1}(\lambda) - (n+1)F_{n+1}(\lambda)] = i\lambda F_n(\lambda), \tag{25a}$$

$$\begin{aligned} RF_n(\lambda) &= i(\tfrac{1}{2}+n)F_n(\lambda) \\ &= \tfrac{1}{2}i[(\tfrac{1}{2}+i\lambda)F_n(\lambda-i) + (\tfrac{1}{2}-i\lambda)F_n(\lambda+i)], \end{aligned} \tag{25b}$$

$$\begin{aligned} T_2F_n(\lambda) &= -\tfrac{1}{2}i[nF_{n-1}(\lambda) - (n+1)F_{n+1}(\lambda)] \\ &= \tfrac{1}{2}i[(\tfrac{1}{2}+i\lambda)F_n(\lambda-i) - (\tfrac{1}{2}-i\lambda)F_n(\lambda+i)]. \end{aligned} \tag{25c}$$

The first equation is the recurrence relation already proved in Sec. III, and only reflects the fact that T_1 is diagonal. The two others are derived, using the generating function (11). For instance,

$$\begin{aligned} &\sum t^n RF_n(\lambda) \\ &= i\sum_0^\infty (n+\tfrac{1}{2})t^nF_n(\lambda) \\ &= i\left(\frac{1}{2} + t\frac{d}{dt}\right)(1+t)^{-\frac{1}{2}-i\lambda}(1-t)^{-\frac{1}{2}+i\lambda} \\ &= \tfrac{1}{2}i[(\tfrac{1}{2}+i\lambda)(1+t)^{-\frac{3}{2}-i\lambda}(1-t)^{\frac{1}{2}+i\lambda} \\ &\quad + (\tfrac{1}{2}-i\lambda)(1+t)^{\frac{1}{2}-i\lambda}(1-t)^{-\frac{3}{2}+i\lambda}] \\ &= \frac{i}{2}\sum_0^\infty t^n[(\tfrac{1}{2}+i\lambda)F_n(\lambda-i) + (\tfrac{1}{2}-i\lambda)F_n(\lambda+i)], \end{aligned}$$

and similarly for T_2. As a result, wherever they are defined (and at least on D), the generators are expressed in the Hilbert space H as difference operators by the formulas

$$\begin{aligned} T_1F(\lambda) &= i\lambda F(\lambda), \\ T_2F(\lambda) &= \tfrac{1}{2}i[(\tfrac{1}{2}+i\lambda)F(\lambda-i) - (\tfrac{1}{2}-i\lambda)F(\lambda+i)], \\ RF(\lambda) &= \tfrac{1}{2}i[(\tfrac{1}{2}+i\lambda)F(\lambda-i) + (\tfrac{1}{2}-i\lambda)F(\lambda+i)]. \end{aligned} \tag{26}$$

It is easily verified that T_1, T_2, and R satisfy the correct commutation rules and are antisymmetric on D. For instance, one can directly show that, for any two F and G in D,

$$(G \mid [T_2 + R]F) + ([T_2 + R]G \mid F) = 0.$$

Indeed, the left-hand side can be written as a contour integral:

$$i\int_C \frac{d\lambda}{\cosh \pi\lambda}\, \tilde{G}(\lambda)(\tfrac{1}{2}+i\lambda)F(\lambda-i),$$

where $\tilde{G}(\lambda) = G^*(\lambda^*)$, and the contour C consists of the lines $\operatorname{Im}\lambda = 0$, $\operatorname{Im}\lambda = i$, and two infinitely remote segments joining these two lines on $\operatorname{Re}\lambda = \pm A$, $A \to \infty$. The integrand is nowhere singular inside this contour since the zero of $\cosh \pi\lambda$ for $\lambda = i/2$ is cancelled by the factor $(\frac{1}{2} + i\lambda)$ and, as a result, the integral vanishes as expected.

The relations (26) give a precise meaning to the remarks made in the introduction concerning the representatives of the other generators in the basis where T_1 is diagonal. When the generators are realized as differential operators in a Hilbert space of functions, we require the existence of an adequate supply of infinitely differentiable functions, though the whole Hilbert space need not contain only differentiable functions. In very much the same way, we are led in the present case to the existence of a sufficient number of entire functions to be able to exponentiate the generators.

Similar considerations can be extended to other representations of G or, more generally, to those of semisimple noncompact groups.[5]

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. W. K. H. Panofsky for his hospitality at SLAC. Discussions with Dr. R. Hermann and the members of the theoretical group at SLAC are gratefully acknowledged.

[5] For a general theory see T. Sherman, Can. J. Math. **18**, 159 (1969). See also N. Mukunda, J. Math. Phys. **8**, 2210 (1967); **9**, 50 (1968); J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, *ibid.* **9**, 2100 (1968); Commun. Math. Phys. **8**, 204 (1968); G. J. Iverson, "Unitary Adjoint Representations of the Lorentz Groups," University of Adelaide preprint.